NHS

National Institute for Health and Clinical Excellence

Single Technology Appraisal

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Evaluation Report

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Contents:

Any information supplied to NICE which has been marked as confidential has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Premeeting briefing

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

This briefing presents the key issues arising from the manufacturer's submission, Evidence Review Group (ERG) report and statements made by consultees and their nominated clinical specialists and patient experts. Please note that this briefing is a summary of the information available and should be read with the full supporting documents.

The manufacturer was asked to:

- confirm that the target population defined in the manufacturer's decision problem is a subgroup of the expected licensed population
- further explain why a comparison against rituximab could not be undertaken quantitatively, and describe any approaches and provide any analyses undertaken to attempt to compare the treatments quantitatively
- provide further trial data on the target population, including: demographics; baseline disease characteristics; systemic lupus erythematosus (SLE) manifestations at baseline; SLE improvements by organ system; results of the efficacy endpoints and mean of the change in Safety of Estrogen in Lupus National Assessment - Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI) score
- clarify how standard care in the trials relates to that in the UK for the high disease activity population and how it differs between trial centres
- provide a table listing all the model assumptions
- explain why the incremental cost-effectiveness ratio (ICER) reduces with the age of the patient
- clarify the reasons for non-responder status and what has been assumed for belimumab non-responders in terms of changes in their SELENA-SLEDAI score and steroid dose over time
- provide patient numbers continuing treatment, and the patient numbers continuing with treatment by responder status.

Licensed indication

Belimumab (Benlysta, GlaxoSmithKline) has a marketing authorisation as add-on therapy in adult patients with active autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g positive anti dsDNA and low complement) despite standard therapy.

Disease overview

SLE is an autoimmune rheumatic disease. It affects the skin and joints, but more serious manifestations can involve the lungs, heart, central nervous system and kidneys. The key aspects of SLE that affect patients' health-related quality of life include: disease flares, with symptoms such as joint and muscle pain, skin rash, and fever; chronic fatigue or malaise; and in more severe disease, the morbidity associated with organ damage and the side effects associated with corticosteroid therapy.

Key issues for consideration

Clinical effectiveness

- The manufacturer's submission focused on a subgroup of the licensed indication (defined as the target population). The target population includes patients with positive anti-dsDNA, low complement and a SELENA-SLEDAI score of equal to or greater than 10. Is the target population appropriate and identifiable in clinical practice?
- The Belimumab International SLE Study (BLISS) clinical trials included patients with a broader set of characteristics than both the marketing authorisation and the target population. Can inferences be made about the effectiveness of belimumab in the target population based on the clinical trial data?
- The two BLISS clinical trials were completed in different geographical regions and enrolled patients with differing characteristics.
  - Is it appropriate to pool the phase III data from the BLISS studies?

  - Are the trial results from the pooled data generalisable to a UK setting?
- The patients in the BLISS trials had a relatively narrow range of SLE manifestations (mainly restricted to musculoskeletal and mucocutaneous problems). Can the effect of belimumab be applied to all SLE manifestations?
- The manufacturer provided a quantitative analysis comparing belimumab with standard care from the BLISS trials. No such analysis was considered possible for the comparison of belimumab with rituximab; instead a narrative comparison was presented.
  - Is rituximab an appropriate comparator, or should it be standard care?
  - For the comparison with standard care, how should standard care be defined, are the BLISS trials representative of standard care in the UK?
  - For the comparison against rituximab, are the manufacturer's reasons for not providing an indirect or mixed treatment comparison considered acceptable?
  - What inferences can be drawn about the relative effectiveness of belimumab in comparison with rituximab?

Cost effectiveness

- Does the manufacturer's model adequately represent the natural history of SLE and the likely effect of treatment with belimumab on the condition?
- The ERG raised a number of uncertainties about the estimation of the effect of belimumab and calculation of costs. Does the Committee consider that the calculation of costs and effects is appropriate?
- The model includes assumptions about the maintenance of treatment effect, and treatment continuation and discontinuation.
  - Can it be assumed that treatment effect is maintained over time?
  - Is an annual rate of 8% natural discontinuation considered appropriate?
  - Is the use of a continuation rule of a change equal to or greater than 4 points in SELENA-SLEDAI score at 24 weeks appropriate?

- The manufacturer's model draws on data from a range of sources including the literature and an observational cohort from Johns Hopkins University.
  - Is the use within the model of the Johns Hopkins cohort for predicting the natural history of the disease appropriate?
  - Are the other estimates from the literature (such as those used for the standard mortality ratios) considered appropriate?
- The manufacturer has agreed a patient access scheme with the Department of Health. How does the patient access scheme affect the cost effectiveness of belimumab?
- A comparison of costs of belimumab and rituximab is provided. If belimumab is compared with rituximab, what relative cost effectiveness between the two treatments is expected?

1 Decision problem

1.1 Decision problem approach in the manufacturer's submission

Population	Evidence was provided on two populations: • Phase III trial population: adults with active autoantibody-positive SLE. • High disease activity subgroup: adults with active autoantibody-positive SLE with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥10.
Intervention	Belimumab 10 mg/kg administered as an intravenous infusion over a 1-hour period on days 0, 14 and 28, and at 4-week intervals thereafter in addition to standard therapy.
Comparators	There were two comparators: • Standard care, which comprises (alone or in combination): antimalarials, non-steroidal anti-inflammatories [NSAIDs], corticosteroids, or other immunosuppressants (azathioprine, methotrexate, and mycophenolate mofetil). • Rituximab plus standard care, for the more severe SLE subpopulation.
Outcomes	Disease activity, incidence and severity of flares, mortality, health-related quality of life, disease progression in terms of long-term organ damage, fatigue and adverse events of treatment.
Economic evaluation	The cost-effectiveness of belimumab was expressed as a cost per quality-adjusted life year (QALY). A lifetime time horizon was used. Costs were considered from an NHS and PSS perspective.

1.2 Evidence Review Group comments

1.2.1 Population

The ERG explained that the decision problem in the manufacturer's submission specified two populations: the phase III trial population (adults with active autoantibody-positive SLE), and a high disease activity subgroup.

There was also a third population, the population proposed in the marketing authorisation, but this was not covered in detail in the submission.

The high disease activity subgroup was termed the 'target population' and was the focus of the manufacturer's submission. The identification of the target population, and the evidence for clinical effectiveness of belimumab in the target population, came from post hoc analyses of the two BLISS trials. The target population represented a subpopulation (~64.5%) of the population covered by the expected marketing authorisation. The target population was defined as 'Adults with active autoantibody-positive SLE with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥ 10' (manufacturer's submission, page 53).

1.2.2 Intervention

The intervention described in the manufacturer's submission matches that in the final scope issued by NICE.

1.2.3 Comparators

The manufacturer's submission did not quantitatively consider rituximab or cyclophosphamide as comparators; only standard care was formally assessed. The manufacturer's submission did not quantitatively compare rituximab and belimumab because there has been no head-to-head trial of rituximab versus belimumab; outcome measures used in rituximab and belimumab trials differ to the extent that there is little possibility of undertaking meaningful indirect comparison. A comparison was made of the costs of belimumab (with the application of a patient access scheme) and rituximab, assuming equal efficacy. The manufacturer noted that this may be a conservative assumption given that the trial of rituximab did not achieve statistical significance compared with placebo.

The manufacturer did not consider that clophosphamide was a suitable comparator because, while it is used in the more severe SLE patient

population, it is largely reserved for the treatment of lupus nephritis (which is not the proposed target population for belimumab).

1.2.4 Outcomes

The ERG considered that the manufacturer's decision problem matches the scope issued by NICE, because it includes the outcome measures: disease activity, incidence and severity of flares, mortality, health-related quality of life including fatigue, and adverse effects of treatment. A novel composite outcome measure called the SLE Responder Index (SRI) was developed and was the primary outcome measure used in the phase III clinical trials.

1.2.5 Economic evaluation

The ERG found that the manufacturer's economic analysis was in line with that stipulated in the scope. The manufacturer's submission presented its economic assessment in terms of incremental cost per QALY and modelled outcomes using a lifetime horizon. Costs were considered from an NHS and PSS perspective.

1.2.6 Other relevant factors

The ERG noted that special considerations and issues raised in the manufacturer's decision problem include: the innovative nature of belimumab for SLE; the insensitivity of the utility measure used for capturing health-related quality of life of SLE patients; and the impact of SLE on particular ethnic groups and on women of childbearing age. The ERG noted that these issues were not included in the final scope as issued by NICE, and that the draft Summary of Product Characteristics specifies that belimumab should not be administered to pregnant women or those planning pregnancy.

2 Clinical effectiveness evidence

2.1 Clinical effectiveness in the manufacturer's submission

The manufacturer identified two phase III clinical trials for inclusion in its submission to NICE. The BLISS-52 and BLISS-76 trials were randomised, double-blind, placebo-controlled, parallel group studies with follow-up of 52 weeks and 76 weeks respectively. In the trials, belimumab plus standard care was compared with placebo and standard care. Standard care included: antimalarials, NSAIDs, corticosteroids or other immunosuppressants (azathioprine, methotrexate, and mycophenolate mofetil) either alone or in combination (see table A12.1 in the clarification response). In the BLISS-52 trial, 290 people received belimumab 10 mg/kg, 288 received belimumab 1 mg/kg and 287 received placebo. In the BLISS-76 trial, these were 273, 271 and 275 people respectively. Although each of the BLISS trials were three arm trials, only results for the 10 mg/kg belimumab dose are presented in the manufacturer's submission because this is the dose submitted for marketing authorisation. Belimumab was administered by intravenous infusion on days 0, 14 and 28, and every 28 days thereafter, for 48 weeks in BLISS-52 and for 72 weeks in BLISS-76.

Adult patients (aged 18 years or older) who met the American College of Rheumatology criteria for SLE and had active disease (score 6 or more at screening on SELENA-SLEDAI) were eligible for enrolment in the BLISS trials. In addition, patients had to have unequivocally positive antinuclear antibody (titre 1:80 or more) or anti-dsDNA antibody (30 IU/ml or more), and to have been on a stable treatment regimen for at least 30 days before the first study dose. Patients with severe active lupus nephritis or central nervous system lupus were excluded. Of the trial populations, 52% met the criteria for the marketing authorisation (that is, had positive anti-dsDNA and low complement) and 34% (n = 396) met the criteria for entering the high disease

activity subgroup (that is, had positive anti-dsDNA and low complement and SELENA-SLEDAI of 10 or more). The patient characteristics and results for the manufacturer's target population, that is the high disease activity subgroup, are described in this premeeting briefing.

Baseline demographics of the patient population can be found in table A3.1 of the manufacturer's clarification response. In summary, the BLISS-52 trial recruited people from the Americas' excluding the USA and Canada, from Asia and from Eastern Europe, whereas the BLISS-76 trial recruited people from the USA, Canada, Europe (Western and Eastern) and Israel. In the BLISS-52 trial, the majority of people were Asian or of Hispanic origin, whereas in the BLISS-76 trial the majority of people were white. Over 90% of the people included in the trials were female and the majority (>80%) were aged 45 years or younger.

In both trials, over 90% of the participants had at least 1A or 1B British Isles Lupus Assessment Group (BILAG) involvement and over 60% had at least 1A or 2B involvement. Mean SELENA-SLEDAI score was approximately 13 in both trials. Approximately 85% of people in the trials had a Physician's Global Assessment (PGA) score of between 1 and 2.5. Average prednisolone dose was between 10 mg/day and 14 mg/day.

2.1.1 Results

The primary outcome of both studies was the response rate at week 52, assessed with SRI. With the SRI criteria, a responder was defined as having: a reduction of at least 4 points in the SELENA-SLEDAI score (defined as clinically meaningful); no new BILAG A organ domain score; no more than 1 new BILAG B organ domain score; and no worsening in PGA score (increase of less than 0.3) at week 52 compared with baseline (see appendix B). Figure 1 shows the differences in SRI response between the different subgroups.

The major secondary outcomes were: percent of patients with a 4-point reduction or more in SELENA-SLEDAI at week 52; mean change in PGA at week 24; percent of patients with prednisone (equivalent) reduction 25% or more from baseline to 7.5 mg/day or less during weeks 40–52 (in patients whose prednisone equivalent dose was more than 7.5 mg/day at baseline); and mean change in SF-36 Physical Component Summary at week 24. In BLISS-76, SRI at week 76 was also assessed.

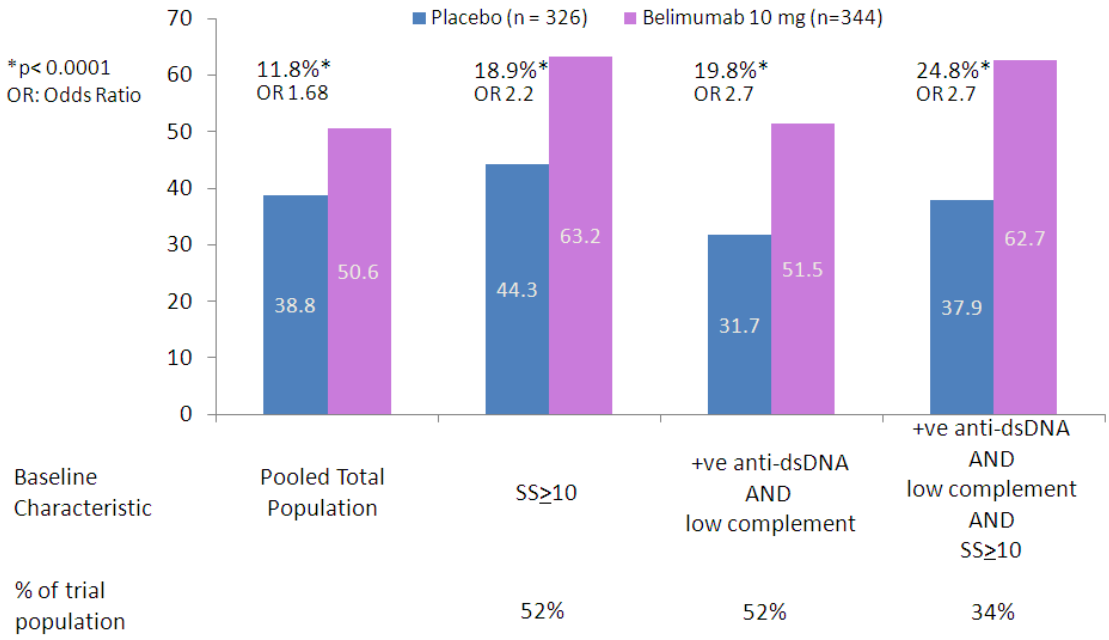


Figure 1 SRI response at week 52 by subset of trial participants. Adapted from figure 5.3 of the manufacturer's submission (page 96)

The results for the high disease activity subgroup from the two trials and the combined trial data are shown in table 1. For the primary outcome of SRI at 52 weeks, statistically significant differences were observed between the belimumab arm and standard care arm in both trials and within the combined analysis. In the combined analysis, 63% of the participants in the belimumab arm, compared with 38% of the standard care arm, were responders

according to the SRI criteria, with an odds ratio (OR) of 2.7 (95% confidence interval [CI]: 1.8–4.1). The BLISS-76 trial demonstrated that a statistically significant difference in response rate between the arms of the trial remained after 76 weeks, with an OR of 2.1 (95% CI: 1.1–3.9).

Table 1 Efficacy endpoints. Adapted from table A6.1 of the manufacturer's clarification responses, n (%). Bold indicates statistically significant result (p < 0.05)

	BLISS-52			**BLISS-76**			**Combined BLISS**		
	SC (n = 107)	**Bel (n = 112)**	**OR (95% CI)**	**SC (n = 96)**	**Bel (n = 81)**	**OR (95% CI)**	**SC (n = 203)**	**Bel (n = 193)**	**OR (95% CI)**
SRI at week 52	44 (41.1%)	75 (67.0%)	**3.0 (1.7, 5.2)**	33 (34.4%)	46 (56.8%)	**2.5 (1.3, 4.6)**	77 (37.9%)	121 (62.7%)	**2.7 (1.8, 4.1)**
SRI at week 76	–	–		30 (31.3%)	40 (49.4%)	**2.1 (1.1, 3.9)**	–	–	
SLEDAI (reduction 4 or more)	47 (43.9%)	76 (67.9%)	**2.8 (1.6, 4.8)**	37 (38.5%)	49 (60.5%)	**2.4 (1.3, 4.4)**	84 (41.4%)	125 (64.8%)	**2.6 (1.7, 3.9)**
No new BILAG 1A/2B	68 (63.6%)	88 (78.6%)	**2.3 (1.2, 4.2)**	57 (59.4%)	57 (70.4%)	1.6 (0.9, 3.1)	125 (61.6%)	145 (75.1%)	**1.9 (1.2, 3.0)**
No worsening in PGA	64 (59.8%)	86 (76.8%)	**2.3 (1.3, 4.2)**	55 (57.3%)	56 (69.1%)	1.6 (0.9, 3.0)	119 (58.6%)	142 (73.6%)	**2.0 (1.3, 3.1)**
Prednisone usage[a] (n at risk)	4 (5.3%) (n = 76)	15 (18.5%) (n = 81)	**4.11 (1.29, 13.2)**	5 (10.0%) (n = 50)	5 (11.1%) (n = 45)	0.88 (0.21, 3.60)	9 (7.1%) (n = 126)	20 (15.9%) (n = 126)	**2.43 (1.05, 5.65)**

[a] Prednisone reduction by 25% or more from baseline to 7.5 mg/day or less during weeks 40–52.

Bel, belimumab; BILAG, British Isles Lupus Assessment Group; OR, odds ratio; PGA = Physician's Global Assessment; SC, standard care; SLEDAI, Safety of Estrogens in Lupus Erythematosus National Assessment trial – Systemic Lupus Erythematosus Disease Activity Index; SRI, SLE Responder Index.

For the components of the SRI, a greater number of patients had a reduction of at least 4 points in the SELENA-SLEDAI score in the belimumab arm compared with the standard care arm, in both trials. In the combined BLISS data set, 65% of patients had a reduction of at least 4 points in the SELENA-SLEDAI score compared with 41% in the standard care arm, with an OR of 2.6 (95% CI: 1.7–3.9), which is statistically significant.

For the outcomes no new BILAG 1A/2B, no worsening in PGA, and prednisone usage, trial results from BLISS-52 showed a statistically significant difference between treatment groups, whereas trial results from BLISS-76 did not. The combined data from both trials showed that overall there is a statistically significant difference between the treatment groups.

For the group of people who were black (defined as African American or African heritage) the percentage meeting the primary endpoint was higher in the placebo group (44%) than in the belimumab arm (36%). This compares with the overall rate of response, which was 39% in the placebo group and 51% in the belimumab group (see table 5.21 in the manufacturer's submission, page 140). For all other races, the belimumab group responded in the expected direction.

Quality-of-life measures, the SF-36 (p117 of the manufacturer's submission) and EQ-5D (p132 of the manufacturer's submission), were also collected during the two phase III trials as secondary outcomes. At week 24, a significant mean change from baseline EQ-5D index was reached in the belimumab arm compared with the placebo arm, but this was not maintained at week 52. The pooled trial data for the high disease activity subgroup showed no significant difference in mean SF-36 physical component summary score between the arms of the trial at weeks 24 or 52.

2.1.2 Adverse effects

Adverse event data were taken from the entire dataset (that is, not just the high disease activity subgroup) from the two phase III clinical trials and from a phase II trial. Over 90% of patients in each arm experienced at least one adverse event. Serious adverse events were experienced by 17% in the 10 mg/kg belimumab group, compared with 16% in the placebo group. Across the double-blind treatment periods, there were 14 deaths, including 3 (0.4%) in the placebo groups, and 6 (0.9%) in the 10 mg/kg belimumab groups. Infections were the most frequent event leading to death in all treatment

groups. The most frequent (occurring in more than 10% of patients) events were headache, upper respiratory tract infection (URTI), arthralgia, nausea, urinary tract infection (UTI), diarrhoea, and fatigue. Of these events, only diarrhoea and nausea occurred slightly more frequently in the belimumab groups. There were four infection-related deaths, one with placebo, one with 1 mg/kg belimumab and two with 10 mg/kg belimumab, and infection may have contributed to the deaths of two further patients (one each of 1 mg/kg and 10 mg/kg). There were two suicides, both in patients receiving belimumab (one each of 1 mg/kg and 10 mg/kg), and one cancer-related death in a patient receiving 1 mg/kg belimumab (likely pre-existing condition). In the long-term open-label extension of the phase II trial (LBSL99), the incidence of adverse events and severe adverse events remained stable or declined over time through 5 years of exposure.

2.1.3 Comparison against rituximab

The manufacturer explained that many patients with more severe, highly active SLE routinely receive rituximab. However, no studies were identified that directly compare belimumab with rituximab. In a study identified by the manufacturer, which compared rituximab with placebo (the EXPLORER trial), no difference was noted in major clinical responses or partial clinical responses between the rituximab group (12.4% had a major clinical response, and 17.2% had a partial clinical response) and the placebo group (15.9% had a major clinical response, and 12.5% had a partial clinical response) relative to the overall response rate (29.6% versus 28.4%). In addition, the rituximab trial demonstrated no difference in secondary endpoints between the rituximab group and the placebo group over 52 weeks of treatment, in patients with moderate-to-severe SLE. The manufacturer stated that differences in the endpoints considered and the patient populations precluded the conduct of any meaningful indirect and mixed treatment comparisons between the belimumab and rituximab studies (see section 5.7, page 143 of the manufacturer's submission, and clarification response A2).

The manufacturer also provided data assessing the efficacy and safety of rituximab as part of an analysis of a prospective observational data registry from France. Overall response, defined as a reduction in SELENA-SLEDAI score of 3 or more measured over a period of 6 ± 3 months, was observed in 80 of 113 patients (71%). Efficacy was not found to differ significantly depending on whether patients received rituximab monotherapy or rituximab combination therapy.

2.2 Evidence Review Group comments

2.2.1 Clinical effectiveness

The ERG highlighted a number of concerns with the methods used to conduct the systematic review. However, it concluded that studies relevant to the decision problem had been identified and the studies representing belimumab appeared complete.

The proposed licensed population and the high disease activity target population that formed the focus of the clinical effectiveness evidence were subgroups identified from post hoc analyses aimed at identifying patients with the greatest response to belimumab in the pooled phase III trial populations. The ERG considered that the results from these trials should be viewed with some caution, as the observed results from the BLISS trials may not be the same as those that would have been seen from a randomised controlled trial in which only the target population was studied.

The ERG explained that the SRI had been developed in consultation with the US Food and Drug Administration and designed to avoid the possibility that improvement in some particular SLE manifestation or manifestations might mask deterioration in overall disease activity or involvement of new organ damage. The ERG considered that this was a positive aspect of the manufacturer's approach. However, the ERG highlighted that according to

expert clinical opinion the SELENA-SLEDAI (a component of the SRI) is not commonly used to define high disease activity in clinical practice.

The ERG considered that while both trials included adults with active autoantibody-positive SLE, the population in BLISS-76 is more likely to be similar to that in England and Wales than that in BLISS-52, and so the results from BLISS-76 may be more generalisable to the UK. The ERG found that for the target population the results from the BLISS-52 trial were more favourable to belimumab than those from BLISS-76 and additionally BLISS-52 provided more patients to the pooled target population than BLISS-76 (55% versus 45%). Therefore results favourable to belimumab for the pooled target population were more strongly driven by the contribution from the BLISS-52 target population. The ERG therefore had concerns about the relevance of the pooled results for patient populations in England and Wales.

The ERG identified that the primary endpoint (52-week SRI) was statistically significant ($p < 0.05$) in both the two phase III randomised controlled trials. In BLISS-76, of the five pre-specified secondary endpoints, a statistically significant reduction in SELENA-SLEDAI score of 4 points or more at week 52 was shown; however, none of the other major secondary outcomes were found to be statistically significant. In addition, none of the other submitted secondary outcomes were found to be statistically significant, including: time to first flare, time to first severe flare, change in Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) organ damage score at week 52, fatigue status (FACIT change from baseline), and quality of life (EQ-5D change).

A literature search undertaken by the ERG revealed published information on SLEDAI and SF-36 changes in the EXPLORER trial which could have been used for comparison with the BLISS trials. In addition, randomised controlled trials for both rituximab and belimumab recorded BILAG changes, thus offering the potential for undertaking an indirect comparison.

2.3 Statements from professional/patient groups and nominated experts

SLE was described as a rare disease with a small patient population, which is ideally managed in specialist clinics. It was explained that there is a real need for new agents in the treatment of lupus – no drug has been licensed for over 50 years, the only licensed drugs are prednisolone and hydroxychloroquine and there are no licensed products for people with severe disease.

European guidelines exist on management principles of SLE, but there is likely to be variation in the treatment of SLE. SLE is treated according to the severity of the disease, both globally and depending on which organ systems are involved. Treatment ranges from symptomatic to immunomodulatory to immunosuppressive. With more overt arthritis and pleuritic pain, for example, moderate doses of corticosteroids (10–20 mg per day), together with a drug like azathioprine and/or methotrexate are widely used. The more serious manifestations, particularly renal disease, are often treated with high doses of steroids (20 mg or more) and mycophenolate or intravenous cyclophosphamide. Rituximab is used in refractory disease.

Belimumab is thought to be appropriate for people with moderate to severely active seropositive disease, particularly people with refractory disease or for those intolerant to existing treatments, and also those who have been on high doses of corticosteroids for many years. Possible subgroups of patients include: certain ethnic groups, particularly those with aggressive disease (such as with renal involvement) at diagnosis and patients who present late with existing disease damage.

It was considered that administration of belimumab would require: day-case stay on a monthly basis for an intravenous infusion; monitoring during infusions; and training for staff in the use of belimumab.

Treatment with belimumab may reduce the steroid dose that is needed and/or may reduce the need for increases in steroid doses. Steroids are used in all treatment regimens for lupus of any severity but are thought to be the cause of much of the long-term damage accrued in patients with lupus and may account for a significant part of the increased risk of premature cardiovascular disease in patients with lupus. It was thought that any agent with proven steroid sparing/reduction capability in SLE is likely to have short- and long-term benefits.

Alternative treatments to belimumab include rituximab and cyclophosphamide. According to the professional submissions, there is a sense that rituximab is an effective agent, particularly for refractory disease, but it failed to show effect in two randomised controlled trials. Nevertheless, the submissions point out that in clinical practice, rituximab has shown promising results, especially in those with renal lupus. Rituximab is payment by results (PbR) excluded, so the decision to fund it is subject to local funding decisions and is based on whether there is exceptional case for a specific patient. Cyclophosphamide has major disadvantages, especially for the treatment of young women, as the side effects include infection, bone marrow toxicity, and infertility.

The two BLISS studies that provide the evidence base for belimumab are thought to have been well conducted, undertaken in diverse geographical settings and in patients being treated with local standard care, and so are thought to reflect UK current practice. The trials were conducted on a large scale and belimumab was shown to meet its primary endpoints. However, belimumab has principally been used to treat patients with mucocutaneous, musculoskeletal and respiratory problems: it has not yet been established how effective it will be in treating patients with renal or cerebral disease.

Two issues were identified in the design of the clinical trials. The first was that concomitant medications were limited within the trial (ACE inhibitors or angiotensin receptor blockers and statins to reduce cholesterol). The second

issue was in the use of the SRI to evaluate outcomes, which is a composite responder index involving clinical scoring that is not routinely used. However, there is recognition that such scoring is required and the introduction of a more formal, systematic assessment of disease activity and damage is likely to have a secondary impact of improving the standard of care for these patients (by focusing clinicians on outcome), regardless of drug utilisation.

3 Cost effectiveness

3.1 Cost effectiveness in the manufacturer's submission

The manufacturer identified no relevant studies on the cost effectiveness of belimumab in its review of the literature. Therefore the manufacturer developed a de novo decision-analytic model. The model is a micro-simulation model that incorporates the interaction between: patient characteristics, disease activity, medication (corticosteroid use), risk of organ damage development (a patient with SLE could potentially develop damage in 12 different organs) and mortality, as shown in figure 2.

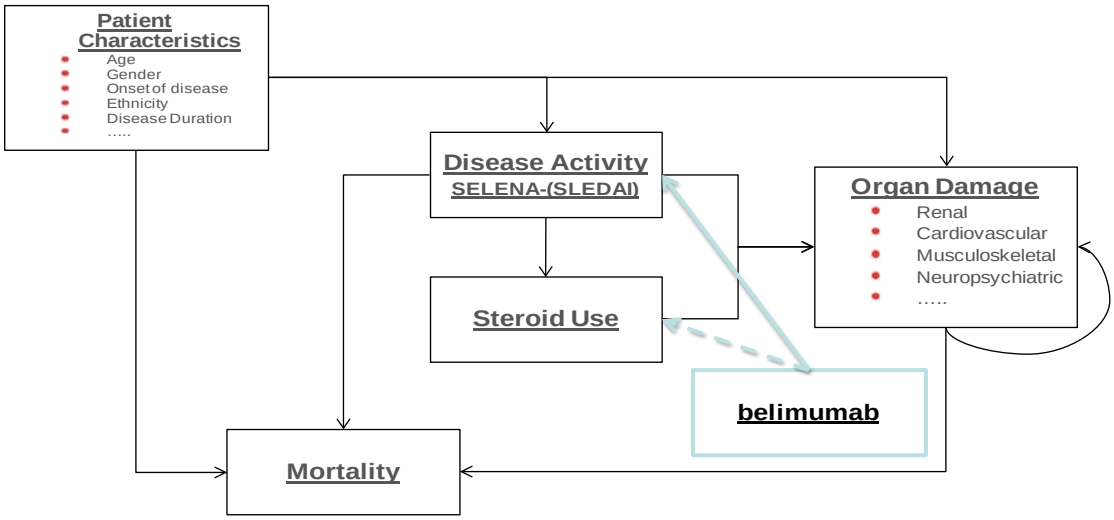


Figure 2 Schematic overview of interdependencies between baseline characteristics, treatment and outcomes in the microsimulation model. Adapted from figure 6.2 of the manufacturer's submission

3.1.1 Model states

The model states (see figure 3) are informed by data from the BLISS trials, observational cohort data, and other data from the literature.

A patient's baseline characteristics are simulated based on the population characteristics in the BLISS trials including: age, gender, ethnicity, SLE disease duration, SLICC damage index (SDI) score and SELENA-SLEDAI score (see section 6.3, page 186 of the manufacturer's submission for further details). They then enter the model in which their lifetime SLE history is simulated. A patient is 'cloned' and enters both the belimumab 10 mg/kg and standard care treatment arms and then works through the model cycle as shown in figure 3. As well as the baseline population characteristics, the BLISS clinical trials inform the likelihood of response at week 24 (based on a patient demonstrating a SELENA-SLEDAI score decrease of 4), the change in SELENA-SLEDAI score up to week 52, likelihood of discontinuation and the effect of SELENA-SLEDAI score on utility and treatment costs (see summary of data sources on page 93 of the ERG report).

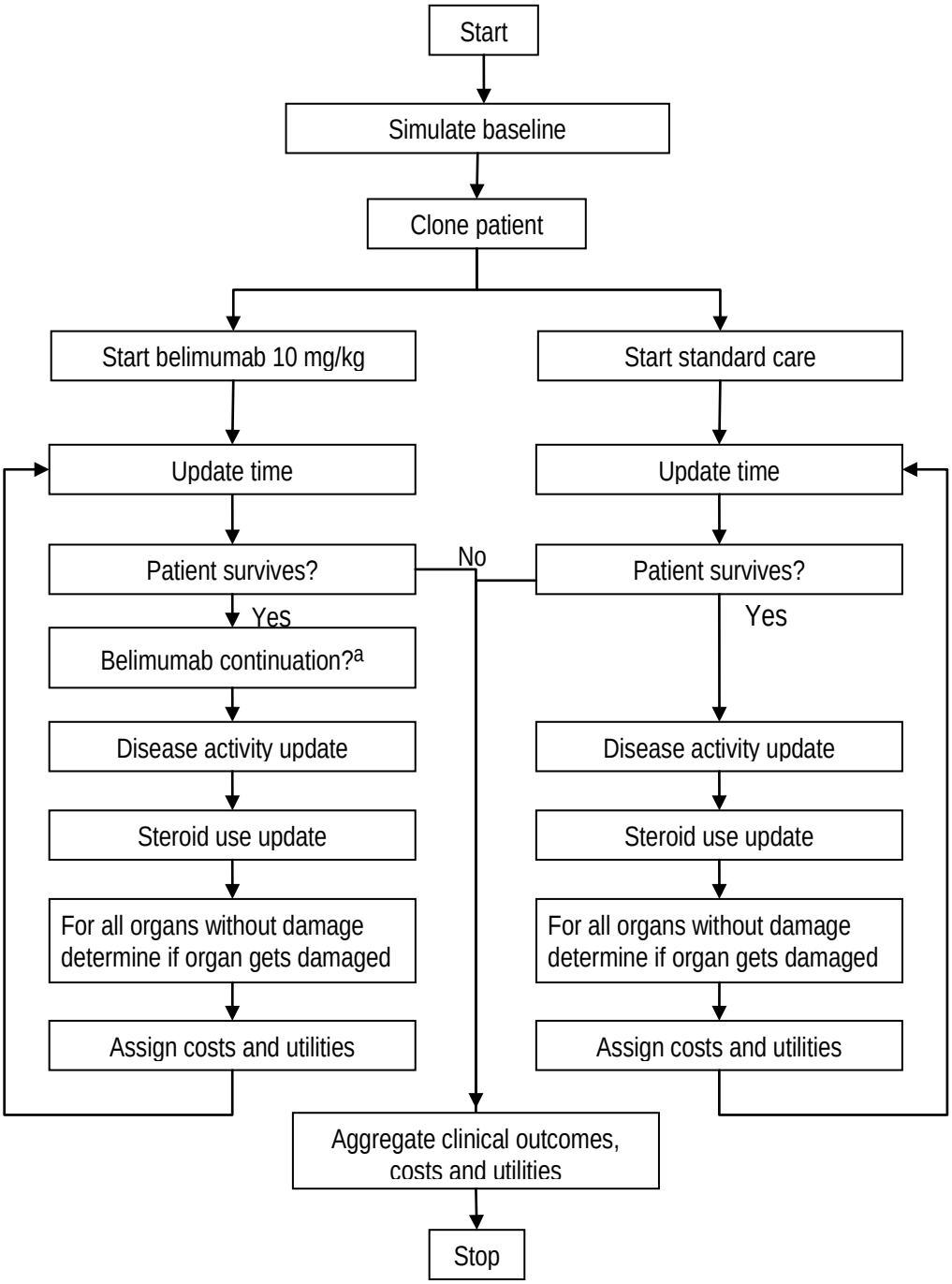


Figure 3 Patient flow through the micro-simulation model. Adapted from figure 6.3 of the manufacturer's submission

[a] if inadequate response to belimumab, the patient switches to standard care and continues through the model's yearly cycles on standard care until death

To inform other model states, prediction models are used based on data from the Johns Hopkins cohort. These are used to predict: change in mean SLEDAI score (which is used as a proxy for SELENA-SLEDAI score), average steroid dose per year, risk of organ damage and risk of death. The Johns Hopkins lupus cohort reports data on a large population of SLE patients from Baltimore, Maryland. Patients in the Johns Hopkins cohort visit the clinic every 3 months from cohort entry. 765 participants were excluded from the analysis, leaving a final sample size of 1282 patients. Of these, 93% were female, and 52% were white, and 38% of black ethnicity. Mean (standard deviation (SD)) age at diagnosis was 33 (13) years and mean age of entry into the study was 38 (13) years. SLEDAI score at first visit was 3.32 (3.7). See table 6.7 of the manufacturer's submission (page 198) for further details.

Further data from the literature were used to inform the standardised mortality rate for a given SELENA-SLEDAI score, and quality of life and cost impacts of each organ involvement.

For a patient entering the model and assigned to either belimumab or standard care, it is first determined whether the patient survives for that year. This is based on data from the Johns Hopkins cohort adjusted by standardised mortality ratios from the literature. For a surviving patient on belimumab, it is then established whether the patient continues belimumab medication. In the model belimumab treatment can be stopped due to natural discontinuation or insufficient response after the first 6 months. Figure 4 shows the estimated percentage of patients on belimumab through time.

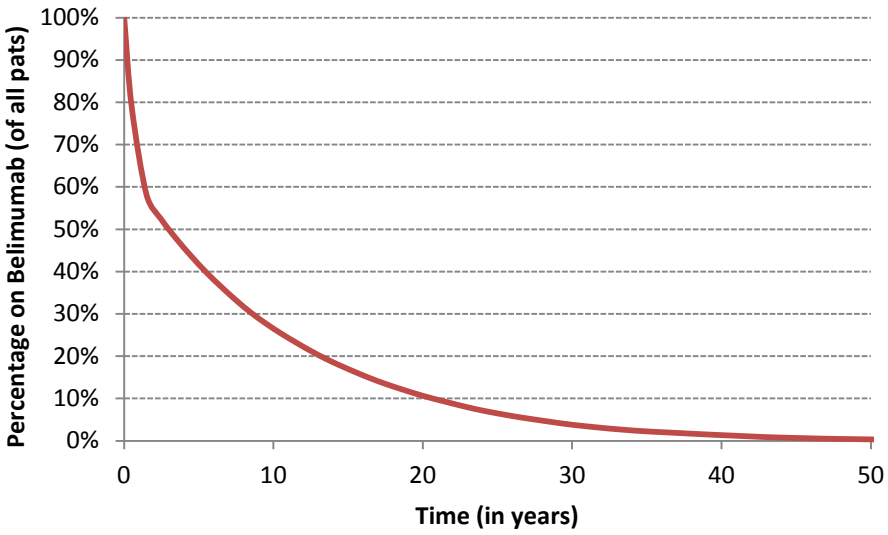


Figure 4 Discontinuation from belimumab (including death). Adapted from figure 6.35 of the manufacturer's submission

Having determined continuation of belimumab, disease activity is updated in the model. In the first year of the simulation, the effects on disease activity measured by SELENA-SLEDAI score as observed in the BLISS trials are applied. A linear regression model based on data from the BLISS trials was used to predict the change in SELENA-SEDAI score at 52 weeks. For subsequent cycles, disease activity is predicted using regression equations based on the natural history data from the Johns Hopkins cohort.

For each organ system contained within the SLICC Damage Index (SDI) (see appendix B), the probability of damage during that year is calculated based on the patient's characteristics and disease activity at that time based on Johns Hopkins data (see table 6.14 in the manufacturer's submission, page 212).

3.1.2 Model inputs

Average costs and utilities calculated from regression analyses are assigned to a patient's health state for that particular year. Costs and utilities are then recorded together with clinical outcomes for that patient. Time is then

increased by 1 year and the process is repeated for the lifetime of the patient. These yearly cycles continue until a patient dies. Utilities and costs are discounted at 3.5%. A NHS and PSS perspective was adopted. Adverse events were not included in the model because the trials did not find any important differences between the incidence of adverse events in treatment groups in the BLISS trials.

Mean EQ-5D from the BLISS trials was 0.70 (SD = 0.26) and this was used to inform the baseline utilities in the economic model. The baseline quality of life assumed in the cost-effectiveness analysis was determined by the following regression equation, which was derived from the BLISS trials:

U = 1.275 – 0.140*log e (AGE) – 0.036*BLACK – 0.009*SS

where AGE is the current age of patient, BLACK is 1 if a patient is of black African ethnicity, or 0 otherwise, and SS is the SELENA-SLEDAI score during the particular model yearly cycle. The above equation was used to determine a patient's utility (U) without organ damage.

Disutility multiplier values for each type of organ damage were identified from a search of the literature (see table 2). These disutility multipliers were applied to U if a patient had developed organ damage in the model cycle. The lowest disutility was used if multiple organs were involved. For example, for a black African SLE patient aged 40 years at entry with a SS score of 10, the baseline utility using the equation above is 0.63. If this person has ocular organ damage, this would give a disutility multiplier of 0.97 (as can be seen from table 2). So the utility for this person would be calculated as 0.63*0.97 = 0.61.

Table 2 Summary of disutility multipliers for the cost-effectiveness analysis. Adapted from table 6.19 of the manufacturer's submission

Organ damage type	Disutility multipliers year		SD
	1	2[a]	
Cardiovascular	0.72	0.76	10%
Diabetes	0.91	0.91	10%
Gastrointestinal	0.79	0.91	10%
Malignancy	0.92	0.92	10%
Musculoskeletal	0.67	0.74	10%
Neuropsychiatric	0.68	0.71	10%
Ocular	0.97	0.99	10%
Peripheral vascular	0.86	0.92	10%
Premature gonadal failure	1	1	
Pulmonary	0.69	0.69	10%
Renal	0.97	0.96	10%
Skin	0.94	0.94	10%

[a] Disutility mutipliers in year 3 and beyond were generally the same as those in year 2. See manufacturer's submission for further details.
SD, standard deviation.

Costs in the analysis were limited to direct medical costs and costs associated with disease activity and long-term organ damage. Costs related to disease activity were drawn from an analysis conducted in 2009 on the resource utilisation recorded in the 1-year belimumab phase II trial (LBSL02) in which 2005/06 NHS reference costs were used. Costs were inflated to 2010 costs. Total resource use was varied according to disease severity and calculated using a linear regression analysis (see pages 241 to 243 of the manufacturer's submission).

A literature search was conducted to identify cost of organ damage. Costs were inflated to 2010 costs. Cost for the first and second years after initial damage development are shown in table 3.

Table 3 Costs for organ damage in the first and second year after initial damage development. Adapted from table 6.26 in the manufacturer's submission

Organ damage type	Year 1	Year 2
Cardiovascular	£3440	£505
Diabetes	£2338	£2338
Gastrointestinal	£2708	£0
Malignancy	£6123	£0
Musculoskeletal	£5431	£1903
Neuropsychiatric	£3660	£1144
Ocular	£1535	£17
Peripheral vascular	£2988	£598
Premature gonadal failure	£0	£0
Pulmonary	£9679	£9603
Renal	£1765	£2453
Skin	£0	£0

The base case only considers the additional acquisition costs for belimumab. Because belimumab is given in addition to standard care, it is assumed that the costs for standard care treatments cancel one another out and so were not included (page 247 of the manufacturer's submission). The administration cost of £126 for belimumab was calculated based on 2 hours of senior hospital staff nurse time (£63/hour) from PSSRU Unit Costs of Health and Social Care 2010 (1 hour for the actual infusion and another hour for patient preparation and monitoring post-infusion). It was assumed that the first year annual cost of treatment and administration of belimumab was £10,918 and in subsequent years was £10,138. This cost assumed a price of belimumab of £114.30 for a 120mg vial and £381 for a 400mg vial. The inclusion of a cost for standard care and different costs of administration were explored in scenario analyses.

3.1.3 Results

The model shows a slower disease activity for belimumab patients than standard care patients, which leads to a decreased steroid dose and a

decreased risk of organ damage and contributes to a difference in mortality risk. The model predicts that in the belimumab arm patients live longer than patients in the standard care arm (as shown in figure 4).

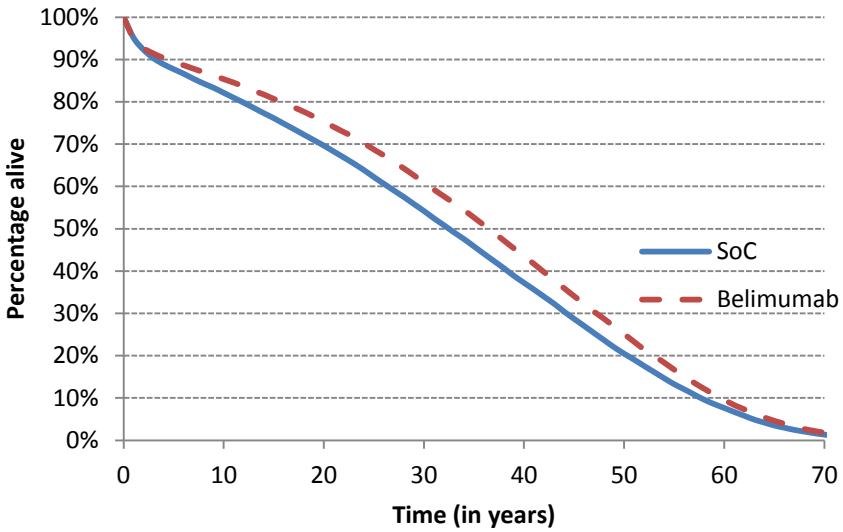


Figure 4 Survival of patients over time. Adapted from figure 6.36 of the manufacturer's submission

Because belimumab patients have an estimated longer life expectancy, the exposure to the risk of organ damage is increased for belimumab patients. For six of the organ damage types (diabetes, gastrointestinal, malignancy, musculoskeletal, neuropsychiatric and ocular) the percentage of occurrence is similar or higher in the belimumab arm than for standard care. However, for cardiovascular, peripheral vascular, premature gonadal failure, pulmonary and renal systems, fewer patients on belimumab develop damage compared with those on standard care. Although a decreased duration of damage is shown for organs on which belimumab has a large effect (cardiovascular, pulmonary and renal), the duration of damage for the other organ systems is increased because of the prolonged life-expectancy.

Table 4 Summary of health economic outcomes. Adapted from table 6.45 of the manufacturer's submission

	SC	Belimumab	Difference
Age at death	66.2	69.1	2.9
SLICC at death	4.1	4.0	–0.1
AMS	5.5	4.55	–0.9
Average monthly steroid cumulative dose	228.1	207.9	–20.2
Life years (undiscounted)	31.93	34.87	2.9
Life years (discounted)	17.05	18.11	1.1
QALYs (undiscounted)	17.31	19.17	1.9
QALYs (discounted)	9.81	10.61	0.8
AMS adjusted mean SLEDAI; QALY, quality-adjusted life year; SC, standard care; SLICC Systemic Lupus International Collaborating Clinics.			

As shown in table 4, the model predicts that belimumab-treated patients, in the subgroup with high disease activity, live on average 2.9 years longer, have a reduction in average mean SLEDAI score, and similar total SLICC organ damage score compared with standard care patients. Treatment with belimumab in the high disease activity subgroup provides an estimated additional 1.1 life years and 0.8 QALYs (discounted).

For both treatment groups, the organ damage costs are the highest expense (see table 5). In total, the organ damage costs are lower for belimumab-treated patients. The costs related to disease activity are similar in the two treatment arms. Overall, the main difference in costs is caused by belimumab acquisition and administration, amounting to £56,067 (89.6%) of the total absolute cost difference of £62,610.

Table 5 Summary of discounted costs over lifetime. Adapted from table 6.46 of the manufacturer's submission

Discounted	SC	Belimumab
Disease activity related costs	£27,882	£28,130
Belimumab drug acquisition	£0	£47,008
Belimumab administration	£0	£9059
Sum of organ damage costs	£77,483	£73,093
Total direct costs	£105,366	£157,291
SC, standard care.		

Belimumab-treated patients are estimated to live longer. However, because of their increased life expectancy and belimumab treatment, costs are higher than for standard care patients. Total costs are £157,291 for belimumab and £105,366 for standard care. Total QALYs are 10.61 for belimumab compared with 9.81 for standard care. The incremental costs are therefore £51,925, and the incremental QALYs 0.806; 2.9 life years are gained (1.05 life years (discounted)). This results in an incremental cost-effectiveness ratio (ICER) of £64,410 per life year gained for the target population (see table 6).

In comparison, the ICER for the population of the marketing authorisation is £66,170 per QALY gained. The ICER for the marketing authorisation population comprises a total cost of standard care of £103,591 compared with £143,895 for belimumab (an incremental difference of £35,584), and 9.55 and 9.98 QALYs respectively (an incremental difference of 0.43 QALYs).

The ICER for the total trial population (which includes a wider population than that specified in the marketing authorisation) is £82,909 per QALY gained. The ICER for the total trial population comprises total cost for standard care £97,583 compared with £133,167 for belimumab (an incremental difference of £35,584), and 9.55 and 9.98 QALYs respectively (an incremental difference of 0.43 QALYs). (See the ERG report page 104).

Table 6 Base-case results. Adapted from table 6.47 of the manufacturer's submission

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus baseline (QALYs)	ICER (£) incremental (QALYs)
SC	£105,366	17.05	9.81	-				
Belimumab	£157,291	18.11	10.61	£51,925	1.05	0.806	£64,410	£64,410
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; SC, standard care.								

The most influential factors on cost effectiveness were found to be: the treatment effect regression to estimate the effect of belimumab after 52 weeks, the natural discontinuation probability and the effect of the adjusted mean SLEDAI (AMS) on mortality. The scatter plot and acceptability curve for the target population are presented in figures 5 and 6 respectively.

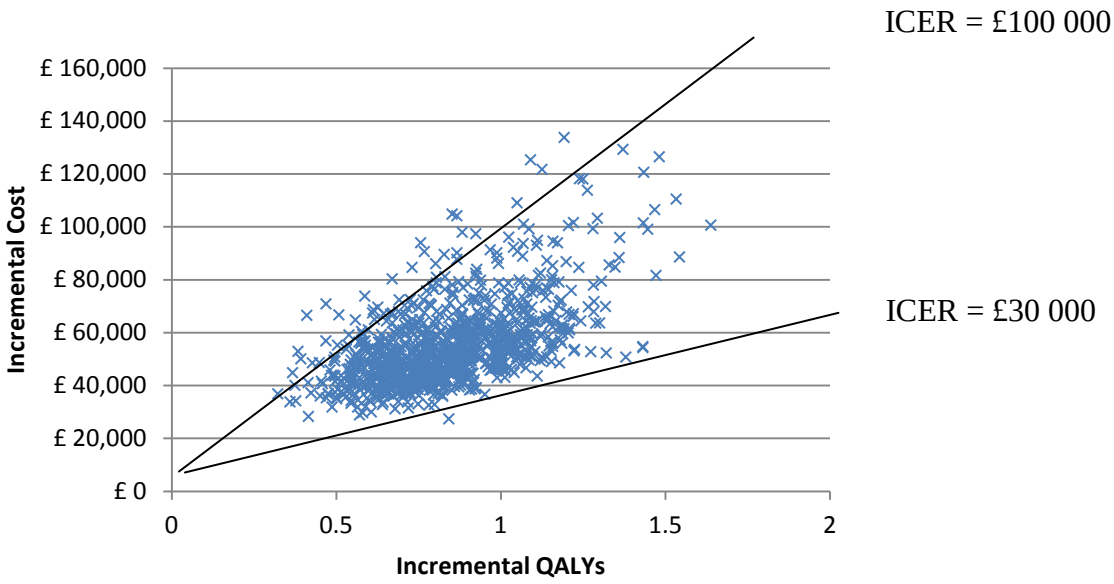


Figure 5 Scatter plot of the PSA. Adapted from figure 6.40 of the manufacturer's submission

The PSA results show that at a willingness to pay (WTP) of £30,000 per QALY gained, there is a 0% probability that belimumab is cost effective compared with standard care. With a willingness to pay of £60,000 per QALY

gained, there is a 35% probability that belimumab is cost effective compared with standard care.

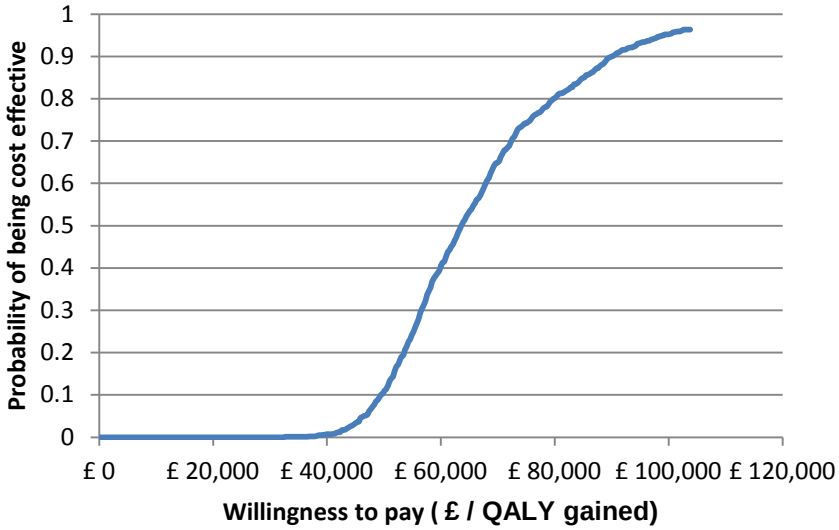


Figure 6 Acceptability curve of PSA. Adapted from figure 6.41 of the manufacturer's submission

A number of scenario analyses were conducted, with resultant ICERs ranging from £50 114 to £77 707 per QALY gained. One of the scenarios explored was an increased vial price to £127.80 for the 120mg vial and £426 for the 400mg vial. This represented the maximum expected vial price. The resulting ICER was £71 297 per QALY gained (see page 301 of the manufacturer's submission).

3.1.4 Patient access scheme

A patient access scheme, which has been accepted by the Department of Health, and provides the 120mg and 400mg vials at a cost of £[redacted] and £[redacted] respectively would make the total cost for a person taking belimumab £[redacted], rather than £157,291 without the patient access scheme. The cost of standard care is £105,366. In the standard care arm, a total of 9.81 QALYs are gained over the course of the model, compared with 10.61 QALYs for the belimumab

arm. This leads to an ICER of £[redacted] per QALY gained compared with £64,410 without the patient access scheme.

3.1.5 Comparison against rituximab

The manufacturer stated that the patient access scheme would make belimumab available at a price that is [redacted] to the cost of providing rituximab (drug acquisition cost). The annual drug cost of belimumab would be £[redacted], compared with an annual cost of rituximab of £6985. The manufacturer explained that while it is not possible to directly compare belimumab with rituximab, using the assumption that belimumab is as effective as rituximab, it is expected that belimumab would [redacted] in this patient population (see page 28 of the manufacturer's submission).

3.2 Evidence Review Group comments

The ERG considered that the manufacturer's model was well constructed and conforms to the NICE reference case and that the longer term effects of SLE had been modelled well, using the Johns Hopkins SLE cohort. An ERG cross-check of the probabilistic modelling for the target population resulted in a central estimate of £65,530 per QALY gained.

The ERG identified a number of issues that may affect the results. They also did a number of exploratory analyses around key parameters in the model as shown in table 7.

Table 7 Summary of the sensitivity analyses conducted by the ERG

Issue	Analysis	ICER	ICER (with PAS)
Base case (target pop)	-	£64,410	[redacted]
Source and calculation of steroid data not subject to exploration	Assumed constant steroid dose of10mg	£68,766	[redacted]
Uncertainty in extrapolation of 8% annual discontinuation*	Annual discontinuation: 14% 2%	 £54,518 £85,893	Not reported
Adjustment of JHU survival model by SMRs from literature. SMRs may be too high for UK cohort	Alternative UK estimates	£70,860	[redacted]
Constant in SS change regression adjusted from 2.0577 in Johns Hopkins cohort to 3.0 to improve fit*	Constant: 2.0577 2.5 3.5	 £93,654 £85,394 £80,988	Not reported
Uncertainty over administration cost of £126 included in model**	£154 as RA £432 day case	£66,907 £91,699	[redacted] [redacted]
Use of pooled trial data to estimate linear regression of SS score at 52 weeks from SS score at baseline	BLISS 52 BLISS 76	£64,960 £66,318	[redacted] [redacted]

*analysis completed by manufacturer

** £154 represents cost used in previous appraisals of rheumatoid arthritis, £432 represents the full day case cost

The ERG highlighted that the steroid use data within the trials was not used in the modelling, and that the function used was not subject to sensitivity analysis. The ERG completed a sensitivity analysis that arbitrarily applied a constant steroid dose (10mg) for all patients in both groups regardless of their SELENA-SLEDAI score. This increased the ICER by approximately £4000 to £68,766 per QALY gained.

The ERG considered that there is some lack of clarity around the reasons for patients' discontinuation and the derivation of the 8% annual discontinuation rate among belimumab week-24 responders, and of the reasonableness of extrapolating using this value. Sensitivity analyses by the manufacturer show that a low discontinuation rate, such as 2% worsens the cost effectiveness of belimumab to an ICER of £85,893, while a higher discontinuation rate, such as 14% improves it, to give an ICER of £54,518 per QALY gained.

The ERG stated that the model assumes that belimumab week-24 non-responders will experience the average SELENA-SLEDAI score within the standard care arm. The ERG considered that this assumption seems likely to have overestimated the average impact on SELENA-SLEDAI scores within the belimumab arm, which would lead to an underestimation of the ICER.

The ERG noted that it is the adjusted mean SLEDAI (AMS) score that contributes to the likelihood of a patient dying and a patient developing particular organ involvement. The economic modelling does not take into account a patient's history before entry into the trial and this may also exaggerate the impact that changes in SELENA-SLEDAI score have on the AMS for belimumab compared with standard care, with the likely result that the base case ICER is an underestimate.

The ERG stated that the requirement to adjust the Johns Hopkin's cohort survival model by standardised mortality ratios (SMRs) from the literature is unclear and may have tended to exaggerate the impact of the individual covariates within the Johns Hopkins cohort survival model. Unpublished data from a UK study obtained by the ERG also suggested that the SMR rates used by the manufacturer may not accurately represent a UK cohort. A sensitivity analysis using the lower SMRs derived from the UK study increased the ICER by approximately £6000 to £70 860 per QALY gained.

The ERG highlighted that the constant in the SS change regression equation from the Johns Hopkin's data was originally 2.0577 but was adjusted by the manufacturer to 3.0 to improve model fit. Sensitivity analyses by the manufacturer show that using the original value of the constant term increased the ICER by £29 000, to £93 654 per QALY gained.

The ERG noted that analysing the observational cost data on a 6-monthly basis in order to relate it to the maximum SELENA-SLEDAI score during that period, and then doubling it to arrive at the annual relationship, appears peculiar given that the observational cost data were collected over a year. The ERG considered that this may have also led to bias, specifically an underestimation of the ICER because of the likely exaggeration of the association between the SELENA-SLEDAI score average over the year and annual treatment costs.

The ERG was concerned that because there were separate estimations of cost per organ involved, this may have double-counted costs estimated within the SELENA-SLEDAI score cost function, the ERG considered that if there was double counting this may have also underestimated the ICER.

The ERG considered the impact of using different administration costs than were used in the model (£126). The ERG found that if costs were in line with those from previous appraisals of rheumatoid arthritis, which had an administration cost of £154, then the ICER would increase by £2500 to £66 907 per QALY gained. If the full day case cost were used (£432) then the ICER would be higher by £27 000, at £91 699 per QALY gained.

The ERG completed a sensitivity analysis that used the estimates from the single trials in the regression equation rather than the estimate from the pooled trials. This analysis demonstrated that the economic model was not particularly sensitivity to the use of the single estimates. Using the BLISS-76

as the source of the regression increased the ICER by approximately £2000 to £66 318 per QALY gained.

4 Equalities issues

In the scoping workshop consultees considered that children should have access to belimumab. The manufacturer confirmed at the scoping workshop that a marketing authorisation will not be sought for children and therefore an appraisal should focus on adult patients with SLE only. Compliance with the therapy was also raised as a potential equality issue, because patients would need hospital admission to receive this drug.

Consultees noted that certain ethnic minority groups might benefit more from therapy with belimumab than others, because the prevalence of SLE is greater in certain populations. The manufacturer highlighted that SLE is more common in women than in men. It is also more prevalent in African-Caribbean, South Asian and Chinese than in European white populations.

5 Authors

Dr Helen Starkie and Zoe Garrett, with input from the Lead Team (Peter Jones, Niru Goenka and Cliff Snelling).

Appendix A: Sources of evidence considered in the preparation of the premeeting briefing

A The Evidence Review Group (ERG) report for this appraisal was prepared by Warwick Evidence:

> Connock M, Cummins E, Sutcliffe P et al. Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (June 2011)

B Submissions or statements were received from the following organisations:

I Manufacturer/sponsor:

- GlaxoSmithKline

II Professional/specialist, patient/carer and other groups:

- British Association of Dermatologists
- British Society for Rheumatology
- Renal Association
- British Health Professionals in Rheumatology
- Primary Care Rheumatology Society
- Royal College of Nursing
- NHS Bolton

Appendix B: Description of primary and secondary endpoint measures

The SLE Responder Index (SRI) includes: a measure of the reduction in global disease activity (reduction in SELENA-SLEDAI score of 4 or more) and two measures to ensure that the improvement in disease activity score is not offset by worsening of the subject's condition overall (that is, no worsening in the PGA) or worsening in any specific organ system (that is, no new BILAG A or two new B flares).

Safety of Estrogens in Lupus Erythematosus National Assessment trial – Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI) aims to capture the subject's condition over the 10 days before the assessment. Disease activity can range from 0 to 105 (0 = no activity, ≥ 20 very high activity). A reduction of 4 points equates to elimination of a disease manifestation and a demonstration of clinical benefit.

The British Isles Lupus Assessment Group (BILAG) measures changes in disease activity over the past 28 days. A BILAG score ranges from A (very active disease) to D (no current disease activity) through to E (the organ system has never been involved). An A or 2B flare represents either an increase in disease activity sufficient to require alteration of therapy (A) or mild reversible problems in two organ systems (2B).

The Physician's Global Assessment (PGA) is a semi-quantitative test of the patient's condition. It uses a 10-cm visual analogue scale from 0 to 3 on which the physician marks his assessment. A score of 1 = mild lupus disease activity, a score of 2–2.5 = moderate disease activity, and a score of 3 = severe disease activity. A change of 1 unit on the PGA is associated with worsening of disease activity. An increase of 1 unit or more from the last assessment resulting in a PGA score of 2.5 or less is considered a mild-

moderate flare. If the increase in PGA is to more than 2.5, it is considered a severe flare.

The SLICC/SDI contains 41 damage items in 12 systems that are specific comorbidities associated with SLE or damage because of toxicity of SLE treatment (cardiovascular disease, diabetes, gastrointestinal, malignancy, musculoskeletal, neuropsychiatric, ocular, peripheral vascular, gonadal failure, pulmonary, renal and skin). Damage items are recorded irrespective of their attribution to SLE. Damage items have to persist for a minimum of 6 months, or be associated with an immediate pathological scar indicative of damage. The total score is the sum of the marked scores and ranges from 0 to 47. Since damage is irreversible, items that are marked will stay marked for the lifetime of the patient.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Health Technology Appraisal

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Final scope

Remit/appraisal objective

To appraise the clinical and cost effectiveness of belimumab within its licensed indication for the treatment of active autoantibody-positive systemic lupus erythematosus.

Background

Systemic lupus erythematosus (SLE) is a chronic autoimmune condition that causes inflammation in the body's tissues. SLE affects the whole body including the skin, joints, internal organs and serous membranes and results in chronic debilitating ill health. The cause of SLE is unknown though a combination of genetic, environmental and hormonal factors is thought to play a role in disease development and progression. Disease activity varies over time and, at the onset, symptoms are very general and may include unexplained fever, extreme fatigue, muscle and joint pain and skin rash. Active SLE involves frequent flares and more severe symptoms compared with inactive disease which is when the disease is in remission. SLE can lead to arthritis, kidney failure, heart and lung inflammation, central nervous abnormalities and blood disorders. Over 90% of people with SLE develop problems with their joints and muscles such as athralgia (joint pain) and myalgia (muscle pain). Renal disease also occurs in 40-75% of people with SLE and significantly contributes to morbidity and mortality. Long-term damage accrues as a result of persistent disease activity and also due to cumulative effects of steroids.

There are currently around 15,000 people in England and Wales with SLE and approximately 2000 people are diagnosed with SLE each year. The prevalence of SLE is significantly higher in African-Caribbean, South Asian and Chinese populations compared with European white populations. Although the severity of the disease is greater in the male population, SLE is significantly more common in women (90% of SLE) than men (10% of SLE) and mainly affects people aged 15-60 years old. After the age of 50 the percentage of women with lupus falls to 75% and the percentage of men with the disease rises to 25%.

The aim of current treatments for SLE is to control and ease symptoms. Standard therapy currently includes the use of non-steroidal anti-inflammatory drugs (NSAIDs), corticosteroids, disease-modifying drugs such as hydroxychloroquine and immunosuppressive agents such as cyclophosphamide, azathioprine, methotrexate and mycophenolate mofetil.

Rituximab and cyclophosphamide are also considered as treatment options, particularly in the case of more severe disease. Prednisolone and hydroxychloroquine are the only drugs specifically licensed for the treatment of SLE. There is currently no NICE guidance on the treatment of SLE.

The technology

Belimumab (Benlysta, GlaxoSmithKline) is a human monoclonal antibody that inhibits the biological activity of B-lymphocyte stimulator (BLyS). BLyS promotes survival and development of B-lymphocyte cells into antibody-producing mature plasma B cells. In SLE, elevated BLyS levels contribute to the production of autoantibodies and have been associated with increased SLE disease activity. Belimumab is administered intravenously.

Belimumab does not currently have UK marketing authorisation for the treatment of active autoantibody-positive systemic lupus erythematosus. It has been studied in clinical trials at different doses compared with placebo plus standard care, as an add on to standard therapy (NSAIDs, corticosteroids, disease-modifying drugs such as hydrocychloroquine and immunosuppressive agents) in people with active SLE on a stable SLE treatment regimen.

Intervention(s)	Belimumab as an add on to standard therapy
Population(s)	Adults with active autoantibody-positive systemic lupus erythematosus
Comparators	• Standard therapy alone; For people in whom it is considered appropriate: • Rituximab plus standard therapy • Cyclophosphamide plus standard therapy
Outcomes	The outcome measures to be considered include: • disease activity • incidence and severity of flares • mortality • health-related quality of life, including fatigue • adverse effects of treatment
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or

	outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.
Other considerations	Guidance will only be issued in accordance with the marketing authorisation. If evidence allows, reduction in supportive treatments, for example steroid use, will be captured in the evidence base. Standard therapy includes, but is not limited to: prednisolone, hydroxychloroquine, cyclophosphamide, azathioprine, methotrexate and mycophenolate mofetil.
Related NICE recommendations	Related Technology Appraisals: Technology Appraisal (suspended appraisal), June 2004, 'Prasterone for the treatment of systemic lupus erythematosus.'

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody positive systemic lupus erythematosus

Matrix of consultees and commentators

Consultees	**Commentators (no right to submit or appeal)**
Manufacturers/sponsors • GlaxoSmithKline (belimumab) Patient/carer groups • Action on Pain • Afiya Trust • Arthritic Association • Arthritis & Musculoskeletal Alliance • Arthritis Care • Black Health Agency • British Kidney Patient Association • British Sjögren's Syndrome Association • Changing Faces • Chinese National Healthy Living Centre • Counsel and Care • Equalities National Council • Genetic Alliance UK • Kidney Alliance • Leonard Cheshire Disability • Let's Face It • Lupus UK • Muslim Council of Britain • Muslim Health Network • National Kidney Federation • National Rheumatoid Arthritis Society • Pain Concern • Pain Relief Foundation • Raynaud's & Scleroderma Association • Skin Care Campaign • South Asian Health Foundation • Specialised Healthcare Alliance	General • Board of Community Health Councils in Wales • British National Formulary • Care Quality Commission • Commissioning Support Appraisals Service • Department of Health, Social Services and Public Safety for Northern Ireland • Medicines and Healthcare products Regulatory Agency • National Association of Primary Care • NHS Alliance • NHS Commercial Medicines Unit • NHS Confederation • NHS Quality Improvement Scotland • Public Health Wales NHS Trust • Scottish Medicines Consortium • Welsh Kidney Patients Association Comparator manufacturer(s) • Baxter Healthcare (cyclophosphamide) • Pfizer (cyclophosphamide) • Roche Pharmaceuticals (rituxumab) Relevant research groups • Arthritis Research UK • British Epidermo-Epidemiology Society • British Society for Immunology • Chronic Pain Policy Coalition • Cochrane Skin Group, Centre of Evidence-based Dermatology, University of Nottingham • Kidney Research UK • MRC Clinical Trials Unit

Consultees	**Commentators (no right to submit or appeal)**
Professional groups • Association of Renal Industries • Association of Renal Technologists • British Association of Dermatologists • British Association for Services to the Elderly • British Dermatological Nursing Group • British Geriatrics Society • British Health Professionals in Rheumatology • British Institute of Musculoskeletal Medicine • British Orthopaedic Association • British Pain Society • British Renal Society • British Skin Foundation • British Society for Allergy and Clinical Immunology • British Society for Human Genetics • British Society for Rheumatology • British Society of Rehabilitation Medicine • British Urological Foundation • National Pharmacy Association • Primary Care Rheumatology Society • Renal Association • Royal College of General Practitioners • Royal College of Nursing • Royal College of Pathologists • Royal College of Physicians • Royal Pharmaceutical Society • Royal Society of Medicine • Society for DGH Nephrologists • UK National Screening Committee • UK Renal Pharmacy Group • United Kingdom Clinical Pharmacy Association Others • Department of Health • Hywel Dda Local Health Board • NHS Herefordshire • Welsh Assembly Government	• National Institute for Health Research • Policy Research Institute on Ageing and Ethnicity • Research Institute for the Care of Older People • Skin Research Centre • START – Skin Treatment and Research Trust Evidence Review Group • Aberdeen HTA Group • National Institute for Health Research Health Technology Assessment Programme Associated Guideline Groups • National Clinical Guideline Centre Associated Public Health Groups • None

NICE is committed to promoting equality and eliminating unlawful discrimination. Please let us know if we have missed any important organisations from the lists contained within the matrix and which organisations we should include who have a particular focus on relevant equality issues.

PTO FOR DEFINITIONS OF CONSULTEES AND COMMENTATORS

Definitions:

Consultees

Organisations that accept an invitation to participate in the appraisal; the manufacturer(s) or sponsor(s) of the technology; national professional organisations; national patient organisations; the Department of Health and the Welsh Assembly Government and relevant NHS organisations in England.

The manufacturer/sponsor of the technology is invited to make an evidence submission, respond to consultations and has the right to appeal against the Final Appraisal Determination (FAD).

All non-manufacturer/sponsor consultees are invited to submit a statement[1], respond to consultations, nominate clinical specialists or patient experts and have the right to appeal against the Final Appraisal Determination (FAD).

Commentators

Organisations that engage in the appraisal process but that are not asked to prepare an evidence submission or statement, are able to respond to consultations and they receive the FAD for information only, without right of appeal. These organisations are: manufacturers of comparator technologies; NHS Quality Improvement Scotland; the relevant National Collaborating Centre (a group commissioned by the Institute to develop clinical guidelines); other related research groups where appropriate (for example, the Medical Research Council [MRC], National Cancer Research Institute); other groups (for example, the NHS Confederation, NHS Information Authority and NHS Purchasing and Supplies Agency, and the *British National Formulary*.

All non-manufacturers/sponsors commentators are invited to nominate clinical specialists or patient experts.

Evidence Review Group (ERG)

An independent academic group commissioned by the National Institute for Health Research (NIHR) Health Technology Assessment Programme (HTA Programme) to assist the Appraisal Committee in reviewing the manufacturer/sponsor evidence submission to the Institute.

[1] Non manufacturer consultees are invited to submit statements relevant to the group they are representing.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

GlaxoSmithKline

Specification for manufacturer/sponsor submission of evidence

13 April 2011

Contents

List of tables and figures

Please include a list of all tables and figures here with relevant page references.

Abbreviations

AAP	Anticardiolipid Antibodies
ACR	American College of Rheumatology
AE	Adverse Event
AIA	African/Indigenous-American
AIR	AutoImmunity and Rituximab (registry)
A-M	Antimalarial
AMS	Adjusted Mean SLEDAI
ANA	Antinuclear Antibody
ANCOVA	Analysis of Covariance
Anti-dsDNA	Anti-Double-Stranded Deoxyribonucleic Acid
Anti-La	Anti-La (antibody)
Anti-RNP	Anti-Ribonucleoprotein
Anti-Ro	Anti-Ro (antibody)
Anti-Sm	Anti-Smith (antibody)
ATC	Anatomical Therapeutic Chemical
AUCMB	Area Under the Curve Minus Baseline
AWMSG	All Wales Medicines Strategy Group
BILAG	British Isles Lupus Assessment Group
BLISS	Belimumab International SLE Study
BLyS	B-Lymphocyte Stimulator
C	Complement
C3	Complement Component 3
C4	Complement Component 4
CAPD	Cumulative Average Prednisone Dose
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence Interval
CNS	Central Nervous System
CRD	Controlled Repeat Dose
CV	Cardiovascular
DA	Disease Activity
DNA	Deoxyribonucleic Acid
EMA	European Medicines Agency
EPAR	European Public Assessment Report
EQ-5D	EuroQoL 5 dimensions
EQ VAS	EuroQoL Visual Analogue Scale
ESRD	End Stage Renal Disease
EULAR	European League Against Rheumatism
FACIT	Functional Assessment of Chronic Illness Therapy
FDA	Food and Drug Administration
GF	Gonadal Failure
GI	Gastrointestinal
GSK	GlaxoSmithKline
HCP	Health Care Professional
HE	Health Economic
HR	Hazard Ratio
HRQL	Health Related Quality of Life
ICER	Incremental Cost-Effectiveness Ratio
ICUR	Incremental Cost utility Ratio
Ig	Immunoglobulin
IgG	Immunoglobulin G
IgM	Immunoglobulin M

IS	Immunosuppressant
IV	Intravenous
JH	Johns Hopkins
LAP	Lupus Anticoagulant Positive
LOCF	Last Observation Carried Forward
LOQ	Level of Quantitation
LSM	Least Square Means
LYG	Life-Years Gained
MAA	Marketing Authorisation Application
MedDRA	Medical Dictionary for Regulatory Activities
MSK	Muskuloskeletal
NHM	Natural History Models
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NP	Neuropsychiatric
NSAID	Non-Steroidal Anti-Inflammatory Drug
OECD	Organisation for Economic Co-Operation and Development
OMERACT	Outcome Measures in Rheumatology
OR	Odds Ratio
PAS	Patient Access Scheme
PCS	Physical Component Summary
PGA	Physicians Global Assessment
PICO	People Intervention Comparison Outcome
PRO	Patient Reported Outcome
PSA	Probabilistic Sensitivity Analysis
PSS	Personal Social Services
PV	Peripheral Vascular
QALY	Quality Adjusted Life Year
QOL	Quality of Life
RCT	Randomised Controlled Trial
SAE	Serious Adverse Event
SD	Standard Deviation
SDI	Systemic Lupus International Collaborating Clinics Damage Index
SELENA	Safety of Estrogen in Lupus National Assessment
SF-8	Short Form 8-Item Health Survey
SF-36	Short Form 36-Item Health Survey
SLAM	Systemic Lupus Activity Measure
SLE	Systemic Lupus Erythematosus
SLEDAI	Systemic Lupus Erythematosus Disease Activity Index
SLICC	Systemic Lupus International Collaborating Clinics
SMC	Scottish Medicines Consortium
SMR	Standardised Mortality Ratio
SoC	Standard of Care
SPA	Special Protocol Assessment
SPC	Summary of Product Characteristics
SRI	SLE Responder Index
SS	SELENA-SLEDAI
STA	Single Technology Appraisal
TAR	Technology Assessment Report
TTE	Time To Event
UK	United Kingdom
ULN	Upper Limit of Normal

UTI	Urinary Tract Infection
URTI	Upper Respiratory Tract Infection
VAT	Value Added Tax
WTP	Willingness to Pay

Instructions for manufacturers and sponsors

This is the specification for submission of evidence to the National Institute for Health and Clinical Excellence (NICE) as part of the single technology appraisal (STA) process. It shows manufacturers and sponsors what information NICE requires and the format in which it should be presented. NICE acknowledges that for medical devices manufacturers particular sections might not be as relevant as they are for pharmaceuticals manufacturers. When possible the specification will refer to requirements for medical devices, but if it hasn't done so, manufacturers or sponsors of medical devices should respond to the best of their ability in the context of the question being addressed.

Use of the specification and completion of appendices 1 to 13 (sections 9.1 to 9.13) are mandatory (when applicable), and the format should be followed whenever possible. Reasons for not following this format must be clearly stated. Sections that are not considered relevant should be marked 'N/A' and a reason given for this response. The specification should be completed with reference to the NICE document 'Guide to the methods of technology appraisal' (www.nice.org.uk), particularly with regard to the 'reference case'. Users should see NICE's 'Guide to the single technology appraisal (STA) process' (www.nice.org.uk) for further details on some of the procedural topics referred to only briefly here.

If a submission is based on preliminary regulatory recommendations, the manufacturer or sponsor must advise NICE immediately of any variation between the preliminary and final approval.

A submission should be as brief and informative as possible. It is expected that the main body of the submission will not usually exceed **100 pages excluding the pages covered by the template**. The submission should be sent to NICE electronically in Word or a compatible format, and not as a PDF file.

The submission must be a stand-alone document. Additional appendices may only be used for supplementary explanatory information that exceeds the level of detail requested, but that is considered to be relevant to the submission. Appendices are not normally presented to the Appraisal Committee. Any additional appendices should be clearly referenced in the body of the submission and should not be used for core information that has been requested in the specification. For example, it is not acceptable to attach a key study as an appendix and to complete the clinical-effectiveness section with ‘see appendix X’. Clinical trial reports and protocols should not be submitted, but must be made available on request.

Trials should be identified by the first author or trial ID, rather than by relying on numerical referencing alone (for example, ‘Trial 123/Jones et al.[126]’ rather than ‘One trial[126]’).

For information on submitting cost-effectiveness analysis models, disclosure of information and equality and diversity, users should see ‘Related procedures for evidence submission’, appendix 10.

If a patient access scheme is to be included in the submission, please refer to the patient access scheme submission template available on request. Please submit both documents and ensure consistency between them.

Executive summary

Please provide an executive summary that summarises the key sections of the submission. All statements should be directly relevant to the decision problem, be evidence-based when possible and clearly reference the relevant section of the submission. The summary should cover the following items.

- The UK approved name, brand name, marketing status and principal mechanism of action of the proposed technology.
- The formulation(s), strength(s), pack size(s), maximum quantity(ies), anticipated frequency of any repeat courses of treatment and acquisition cost.
- The indication(s) and any restriction(s).
- The recommended course of treatment.
- The main comparator(s).
- Whether the key clinical evidence in the submission comes from head-to-head randomised controlled trials (RCTs), from an indirect and/or mixed treatment comparison, or from non-randomised studies.
- The main results of the RCTs and any relevant non-RCT evidence.
- In relation to the economic evaluation, details of:
  - the type of economic evaluation and justification for the approach used
  - the pivotal assumptions underlying the model/analysis
  - the mean costs, outcomes and incremental cost-effectiveness ratios (ICERs) from the evaluation.
- Tabulation of the base-case results as follows:
- When appropriate, please present the results for the intervention and comparator(s) incrementally to indicate when options are dominated or when there is extended dominance. For example:

- Subgroup analyses considered and clinical- and cost-effectiveness results.

Summary

SLE is a relapsing and remitting disease with disease activity fluctuating between periods of exacerbation (flares) and relative quiescence, affecting multiple organ systems in an unpredictable fashion.

There are certain patients who experience significantly high levels of disease activity, despite being managed on high dose standard of care. These patients in particular, suffer both short-term morbidity and are at increased risk of long-term organ damage.

In two pivotal Phase 3 randomised, placebo controlled trials, belimumab, when used with standard of care, demonstrated a favourable benefit/risk profile for the treatment of SLE. Belimumab plus standard of care showed a significant reduction in disease activity compared with standard of care alone, as measured by the SRI composite primary endpoint.

In the base case economic model, belimumab plus standard of care compared to standard of care alone was estimated to have an ICER of £82,909 per QALY, and for the high disease activity subgroup £64,410 per QALY.

Mindful of NHS resources, GSK is proposing a patient access scheme (PAS), designed to reflect both the value GSK believes to be inherent in this technology and the data that supports it; and the opportunity cost to the NHS of introducing a new biologic for the treatment of SLE.

When comparing belimumab to the current standard of care (NSAIDs, corticosteroids, immunosuppressants and antimalarials) the resultant ICER is £[redacted] per QALY when the PAS is included.

[redacted] highlight the additional arguments around the degree of innovation, the additional aspects of value not currently fully reflected in the Cost per QALY and the high unmet need in these patients with high disease activity.

In addition, belimumab would provide a proven, licensed alternative to the use of other licensed and unlicensed treatments in this severe patient group.

Background

Systemic Lupus Erythematosus (SLE) is a chronic autoimmune, multisystem disorder with varying manifestations characterised by an unpredictable clinical course, autoantibody production, abnormal B lymphocyte function and chronic inflammation (Manson et al. 2006).

SLE is a relapsing and remitting disease with disease activity fluctuating between periods of exacerbation (flares) and relative quiescence, affecting multiple organ systems in an unpredictable fashion (ACR Ad Hoc Committee on Systemic Lupus Erythematosus Guidelines 1999).

SLE is approximately 10 times more common in women than men (Manson et al. 2006; Manzi 2009) and more prevalent in African-Americans, South Asians and Chinese than Caucasians (Danchenko et al. 2006; Manzi 2009). The disease onset is generally between the ages of 15 and 44 years (Danchenko et al. 2006). This suggests that SLE affects predominantly women during their childbearing years. In the UK, prevalence has been estimated at 41 per 100,000 persons (Nightingale et al. 2007).

Clinical manifestations vary widely between patients with signs and symptoms evolving over time. Many patients with SLE experience general symptoms including fatigue, malaise, fever, anorexia, weight loss, skin rash and muscle and joint pain (Manson et al. 2006).

In combination with the more immediate impact of SLE on patients' HRQL (health related quality of life), ongoing disease activity correlates significantly with long-term organ damage (Swaak et al. 1999). More than half of patients develop permanent organ damage and this damage progresses over time (Danchenko et al. 2006). Renal disease is one of the commonest and most serious manifestations of SLE (Chambers et al. 2009; Cooper et al. 2007).

SLE also has a significant impact on mortality, with a 2.4-fold greater risk of mortality than the general population, with a higher risk of death due to cardiovascular disease (standardised mortality ratio [SMR] 1.7), non-Hodgkin's lymphoma (SMR 2.8), lung cancer (SMR 2.3), infections (SMR 9.0) and renal disease (SMR 7.9) (Bernatsky et al. 2006). Although the median survival rate is 90% at 5 years, 80% at 15 years, and 70% at 20 years, surviving SLE patients suffer a significant burden of disease with associated morbidity and reduction in HRQL (Abu-Shakra et al. 1995; Campbell, Jr. et al. 2008; Rahman et al. 2008).

There are certain patients who have highly active disease and experience a greater impact on their quality of life, while also being more likely to develop long-term organ damage. These patients are likely to consume significantly more health care resources, requiring more frequent health care professional (HCP) visits and hospitalisations due to exacerbations in their condition. Based on Adelphi research of clinicians treating SLE patients in the UK, 22% of patients had been hospitalised in last 12 months, increasing to over 70% in patients with severe disease (GlaxoSmithKline data on file 2010).

Although data is limited on the long-term resource burden of SLE patients, it is well established that patients with highly active disease are more likely to develop organ damage, e.g. renal dysfunction or renal failure. The costs associated with longer term outcomes such as renal disease are likely to be significant.

Current standard of care

There is no accepted SLE treatment algorithm and no relevant NICE guideline exists. Agreeing on best practice poses a significant challenge owing to the heterogeneous nature of SLE.

There is no cure for SLE and the aims of treatment are (Kalunian et al. 2009):

- Matching treatment to an accurate diagnosis of the extent of organ involvement

- Maintaining an appropriate level of therapy to control or halt the inflammatory disease activity while minimising side-effects and risk of infection
- Preventing further organ damage
- Maintaining a patient's daily function and quality of life

Patients with SLE are currently managed by a range of treatments (NSAIDs, corticosteroids, immunosuppressants and antimalarials); variously used either alone or in different combinations, constituting standard of care (SoC).

This current standard of care (SoC) may be associated with undesirable effects, either from chronic use of steroids (osteoporosis, diabetes and cardiovascular disease) or side effects associated with immunosuppressants (toxicity, infection and infertility). Many of these treatments are not licensed for use in SLE and a significant number of patients with advanced SLE do not respond to current treatments even at high doses.

Patients with more severe, highly active SLE are usually managed in tertiary centres and a significant proportion routinely receive rituximab (MabThera®), a biologic which, although appearing to have some benefit in clinical practice, failed to demonstrate statistically significant efficacy in Phase 2/3 trials and is not licensed in this indication.

According to the NICE Methods Guide (2008), '*Relevant comparators are identified, with consideration given specifically to routine and best practice in the NHS (including existing NICE guidance) and to the natural history of the condition without suitable treatment. Relevant comparator technologies may also include those that do not have a marketing authorisation (or CE mark for medical devices) for the indication defined in the scope but that are used routinely for the indication in the NHS.'*

We have tried to quantify rituximab use for SLE; research indicates that approximately 3% of rituximab use is linked to a diagnosis of SLE (GlaxoSmithKline data on file 2010). This would equate to approximately 600-700 patients in the United Kingdom. However this is likely to be a significant underestimate as rituximab is licensed for a number of indications (oncology and rheumatoid arthritis). Rheumatoid arthritis shares many of the same symptoms as SLE and may be managed by similar clinicians. Rituximab also has positive NICE guidance for the treatment of rheumatoid arthritis (TAR 195). Based on clinician feedback, of the current patient population that receives rituximab for SLE, a significant majority would be considered for treatment with belimumab. Given this use constitutes a significant proportion of more severe SLE patients, rituximab plus standard therapy should be considered within the scope of this decision problem.

Cyclophosphamide, whilst used in the more severe patient population, is largely reserved for the treatment of lupus nephritis. This is not the proposed target population for belimumab; therefore, cyclophosphamide plus standard therapy is not a relevant comparator. In addition, adverse effects associated with long-term exposure to cyclophosphamide including bladder cancer, bone marrow suppression, haematologic malignancies, infections, myelodysplasia, and infertility (Kalunian et al. 2009), limit the appropriateness of cyclophosphamide given that a high proportion of patients are women of childbearing age.

The Technology

Belimumab, a human IgG1λ monoclonal antibody that binds to soluble human B-lymphocyte stimulator (BLyS) and inhibits its biological activity, has been specifically developed for the treatment of SLE and demonstrated efficacy in two Phase 3 clinical trials.

Pharmaceutical formulation	Powder for concentrate for solution for infusion
Method of administration	Intravenous infusion over a one hour period
Doses	Vial sizes – 120mg, 400mg
Dosing frequency	The recommended dosage regimen is 10 mg/kg belimumab on Days 0, 14 and 28, and at 4-week intervals thereafter.
Acquisition cost	£8,000/annum*

* Average annual cost based on weight distribution from Phase 3 trials

BLyS inhibits B cell apoptosis and stimulates the differentiation of B cells into immunoglobulin-producing plasma cells. Over expression of BLyS by transgenic mice results in autoimmune-like disease (Cancro et al. 2009). Furthermore, BLyS is over expressed in patients with systemic lupus erythematosus (SLE) and other autoimmune diseases (Cheema et al. 2001; Zhang et al. 2001).

In patients with SLE followed for 2 years, BLyS levels correlated with changes in lupus disease activity as well as with elevated anti-dsDNA antibody titres, worsening disease activity was predicted by rises in serum BLyS concentrations (Petri et al. 2008). Belimumab has a novel mode of action that specifically inhibits the biological activity of BLyS, promoting apoptosis in autoreactive B cells. Belimumab has been specifically developed for the management of SLE.

The proposed indication for belimumab is for reducing disease activity in adult patients with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g. positive anti-dsDNA, low complement) despite standard therapy.

However, mindful of NHS resources and in order to identify patients who are most likely to benefit from belimumab, GSK proposes this appraisal focuses on the high disease activity subgroup, which is for adults with active autoantibody-positive systemic lupus erythematosus with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥10). This comprises 34% of the overall belimumab Phase

3 trial population. This subgroup combines routinely used subjective laboratory measures with a clinical measure of disease activity; allowing clinicians to identify patients with significant disease activity. SLE of this severity is managed by a small group of around 30 experts in only 20 – 25 centres, and as the marketing authorisation will limit prescribing to those experienced in the management of SLE, prescribing can be easily targeted to those patients who will benefit the most and can be well monitored.

Serological markers of low complement and positive anti-dsDNA were included as they are objective measures used routinely in SLE and accessible to physicians in clinical practice. In addition, they are widely considered important measures of disease activity and identify patients at higher risk for flares and lupus nephritis (Petri et al. 2009; Tseng et al. 2006). SELENA-SLEDAI whilst not used routinely in clinical practice, is well recognised by clinicians and would be relatively straightforward to implement. It is a direct measure of disease activity and was the most significant predictor of Week 52 response. A score of ≥10 is likely to indicate a patient with highly active disease.

Clinical evidence

Clinical data related to the decision problem was retrieved through a systematic review and supplemented by additional unpublished data.

Relevant clinical trials which compared belimumab plus standard of care to placebo plus standard of care included; Phase 2 efficacy and safety study (Wallace et al. 2009) and the two pivotal Phase 3 studies, BLISS-52 (Navarra et al. 2011) and BLISS-76 (GlaxoSmithKline data on file 2011; van Vollenhoven et al. 2010). For the purpose of this submission pooled efficacy from the two pivotal Phase 3 studies is considered most relevant to the decision problem.

In the two pivotal Phase 3 randomised, double-blind controlled trials, belimumab plus standard of care was compared to placebo and standard of care. The standard of care patients received in the clinical trials was broadly

representative of the current treatment of SLE in the UK, where patients are managed with a range of treatments (NSAIDs, corticosteroids, immunosuppressants and antimalarials); variously used either alone or in different combinations. Use of rituximab was not permitted.

The primary endpoint used in the Phase 3 belimumab studies was the SLE Responder Index (SRI). The SRI was developed to measure improvement in disease activity, while at the same time accounting for potential effects on other aspects of the disease and on patient well-being. With this in mind, the Phase 3 studies had a composite primary efficacy endpoint (SRI). A composite endpoint was used based on recommendations from the FDA, OMERACT and EULAR that an ideal responder index should detect early as well as overall changes in disease activity and should also be able to simultaneously identify improvement and worsening in the same and/or different organ systems.

The SRI comprised patient response rate at week 52:

- a reduction from baseline in the SELENA-SLEDAI score of at least 4 points (which indicates a clinically important reduction in SLE disease activity); to illustrate the clinical benefits of a 4 point improvement in SELENA-SLEDAI disease activity score experienced by patients in practice, this would, for example, equate to complete resolution of pleurisy and pericarditis (each scoring 2 points) or complete resolution of myositis (scoring 4 points) or complete resolution of arthritis (scoring 4 points);

- no new BILAG A organ domain score and no more than 1 new BILAG B organ domain score compared with baseline; to ensure that the improvement in SELENA-SLEDAI is not offset by worsening in any specific organ system;

- no worsening in Physician’s Global Assessment (with worsening defined as an increase in PGA of more than 0.30 points from baseline);

to ensure that the improvement in SELENA-SLEDAI is not offset by worsening of the subject's condition overall.

In the pooled total population, belimumab 10 mg/kg plus standard of care was shown to be superior to placebo plus standard of care as assessed by the SRI at Week 52, 50.6% versus 38.8% ($p<0.0001$), respectively (Table 1). An improved response rate was also seen in the high disease activity subgroup with belimumab 10 mg/kg plus standard of care versus placebo plus standard of care, 62.7% versus 37.9% ($p<0.0001$).

Table 1. Primary efficacy endpoint (SRI) at Week 52

	Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 562	10 mg/kg N = 563	Placebo N = 203	10 mg/kg N = 193
No. (%) Response	218 (38.8%)	285 (50.6%)	77 (37.9%)	121 (62.7%)
Observed difference vs placebo (%)	-	11.8	-	24.8
OR (95% CI) vs placebo	-	1.68 (1.3, 2.2)	-	2.7 (1.8, 4.1)
P-value	-	< 0.0001	-	< 0.0001

The Phase 3 study results for the components of the primary efficacy endpoint (SRI) are shown below (Table 2). For the pooled total population and the high disease activity subgroup, belimumab 10 mg/kg was superior to placebo for each of the 3 components.

Table 2. Components of the SRI at Week 52

	Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 562	10 mg/kg N = 563	Placebo N = 203	10 mg/kg N = 193
4-point reduction in SELENA-SLEDAI	230 (40.9%)	297 (52.8%)	84 (41.4%)	125 (64.8%)
Observed difference vs placebo (%)	-	11.9	-	23.4
OR (95% CI) vs placebo	-	1.68 (1.3, 2.2)	-	2.6 (1.7, 3.9)
P-value	-	< 0.0001	-	< 0.0001
No New 1A/2B BILAG domain scores	389 (69.2%)	425 (75.5%)	125 (61.6%)	145 (75.1%)
Observed difference vs placebo (%)	-	6.3	-	13.6
OR (95% CI) vs placebo	-	1.4 (1.1, 1.8)	-	1.9 (1.2, 3.0)
P-value	-	0.0190	-	0.0034
No worsening in PGA	372 (66.2%)	420 (74.6%)	119 (58.6%)	142 (73.6%)
Observed difference vs placebo (%)	-	8.4	-	15.0
OR (95% CI) vs placebo	-	1.5 (1.2, 2.0)	-	2.0 (1.3, 3.1)
P-value	-	0.0017	-	0.0015

The Phase 3 trials collected quality of life data. Both SF-36 and EQ-5D generic quality of life data were collected during the two Phase 3 trials. The latter instrument is consistent with the NICE reference case. The mean change in EQ-5D from baseline did not reach statistical significance for either BLISS-52 or -76 trials. However, EQ-5D may not be the most sensitive measure to assess the true impact of the disease on HRQL experienced by SLE patients. This instrument doesn't directly include certain relevant dimensions of health, such as fatigue or where the disease course is characterised by flares of unpredictable symptom severity (Wailoo et al. 2010). In SLE patients may experience disease flares at any time and not necessarily at the time the EQ-5D was completed for the pre-defined time points in the clinical trials.

Fatigue has been identified by patients as contributing significantly to the decrease in their quality of life. As a symptom of SLE, fatigue is not simply experiencing tiredness, but may be so severe as to restrict normal daily activities of patients. Within the Phase 3 trials, impact on fatigue was measured using a composite fatigue score (created from the FACIT-Fatigue questionnaire). The pooled data from both studies showed that belimumab 10

mg/kg was associated with significantly improved fatigue scores compared with placebo at Weeks 8, 12, and 52 ($p < 0.05$). In the high disease activity subgroup, belimumab 10 mg/kg was associated with significantly improved fatigue scores compared with placebo at Weeks 8 and 12 ($p < 0.05$). Fatigue is an important HRQL measure that has meaningful impact upon SLE patients. The impact of belimumab on fatigue appears to reflect the improvements observed in other clinical and biomarker measures of SLE disease activity, however, the effect on fatigue is not captured in the cost effectiveness analysis.

Consequently, the impact on HRQL is very likely to have been underestimated in the two Phase 3 studies and the corresponding cost-effectiveness analysis.

The safety profile of belimumab plus standard of care as seen in the clinical trial program is favourable. Infusion reactions were slightly higher in the belimumab group than the placebo group, but these were generally mild to moderate and managed with routine treatment. There was no increase in risk of serious infections and malignancy rate is consistent with the background rate for patients with SLE. Within the clinical trial program, the death rate per 100 patient-years was similar for placebo and belimumab treated patients.

Comparative Clinical Effectiveness

As discussed patients with more severe, highly active SLE are usually managed in tertiary centres and many routinely receive rituximab.

The inclusion criteria of the published Phase 2/3 RCT for rituximab required patients to have active disease at screening, defined as ≥ 1 organ system with a BILAG A score or ≥ 2 organ systems with a BILAG B score; and the stable use of 1 immunosupressant at study entry (Merrill et al 2010). This is likely to correspond to a slightly more severe patient population compared to both the overall BLISS trial population and the proposed high disease activity subgroup. The primary endpoint was the effect of rituximab versus placebo in achieving and maintaining a major clinical response, a partial clinical

response, or no clinical response at week 52 assessed using BILAG scores. At week 52, no difference was noted in major clinical responses or partial clinical responses between the rituximab group (12.4% had a major clinical response, and 17.2% had a partial clinical response) and the placebo group (15.9% had a major clinical response, and 12.5% had a partial clinical response) relative to the overall response rate (29.6% versus 28.4%). In addition, the rituximab trial demonstrated no difference in secondary endpoints between the rituximab group and the placebo group and over 52 weeks of treatment, in patients with moderate-to-severe SLE.

The study did not collect data on changes in SELENA-SLEDAI, which is an important short-term outcome to be able to link to longer term impact on organ damage, an important driver of cost effectiveness. So both the trial populations and the outcomes reported are different for the rituximab and BLISS trials, making any indirect comparisons of these technologies using these RCTs problematic.

The efficacy and safety of rituximab was also investigated as part of an analysis of prospective data from the French AutoImmunity and Rituximab (AIR) registry (Terrier et al 2010). One hundred and thirty-six patients received treatment for SLE. Overall response, defined as SELENA-SLEDAI reduction of ≥ 3 measured over a 6 ± 3 month period, was observed in 80 of 113 patients (71%). Efficacy did not differ significantly between patients receiving rituximab monotherapy and those receiving concomitant immunosuppressants (who had higher baseline disease activity). So although this study appears to indicate some benefit for rituximab in a more real-world setting, due to the study design it is limited in terms of the ability to make a formal comparison with belimumab. It may suggest that the full clinical benefit of the use of biologics, like rituximab may not be fully reflected in a randomised clinical trial setting.

Cost-effectiveness

In order to reflect the heterogeneity and complexity of SLE, a micro-simulation model was built. The micro-simulation model simulates individual patients over a lifelong period. The patient population entering the model reflects 1) the pooled total population of the two RCTs: BLISS-52 and BLISS-76, and 2) a subgroup of this pooled population to identify those with the highest disease activity consistent with our target population for this decision problem. As both BLISS trials were of either 52 or 76 weeks duration, the effect of treatment on long-term disease outcomes could not be determined. Long-term outcomes, however, have a major effect on health-economic assessment, and as such, these outcomes were considered important to be included in a model that estimates the cost-effectiveness of belimumab treatment.

The composite primary end point of the BLISS trials included SELENA-SLEDAI (SS) score, a measure of disease activity, as the measure of efficacy. High disease activity over time will accrue organ damage (Swaak et al. 1999). Therefore in order to address this in the model, the relationship between this short-term outcome (BLISS 52/76 week trials) and long-term outcomes was estimated based on a large Lupus registry, the Johns Hopkins cohort. Based on this data, time to event (TTE) models were estimated that describe the relation between disease activity and other covariates on the risk of dying and on the risk of developing irreversible organ damage.

The TTE models are implemented in the health-economic model to simulate a patient's future disease course based on the early outcomes observed in the BLISS trials. Health-economic consequences (quality of life impairment and health-care costs) are assigned to each long-term outcome to translate clinical outcomes to health-economic outcomes. Together with the short-term health economic consequences this allows the cost-effectiveness of belimumab and SoC relative to SoC alone to be assessed over a life-long period.

Results are presented for the base case analysis and high disease activity subgroup with and without the proposed patient access scheme (Tables 3, 4 and 5).

Table 3. Base-case results – Total Pooled Population

Technologies	Total costs (£)	Total LYG	Total QALYs	Total Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
SoC	£97,583	16.74	9.55	-	-	-	
Belimumab	£133,167	17.33	9.98	£35,584	0.59	0.43	£82,909
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years							

Table 4. Base-case results – High Disease Activity Subgroup

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
SoC	£105,366	17.05	9.81	-			
Belimumab	£157,291	18.11	10.61	£51,925	1.05	0.806	£64,410
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years							

In the high disease activity subgroup, belimumab-treated patients are estimated to live longer, however, due to their increased life expectancy and due to belimumab treatment; costs are higher than for SoC patients. The incremental costs are £51,925, resulting in 1.05 added life years or 0.806 added QALYs (discounted). This results in an incremental cost effectiveness ratio (ICER) of £64,410 per QALY gained.

Table 5. High Disease Activity Subgroup – including PAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
SoC	£105,366	17.05	9.81	-			
Belimumab	£[redacted]	18.11	10.61	£[redacted]	1.05	0.806	£[redacted]
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years							

When the patient access scheme discount is considered for the high activity subgroup, the total costs for the belimumab-treated patients are estimated to be £[redacted] and the incremental costs are £[redacted], while the incremental LYG

and incremental QALYs remain the same as those presented previously, at 1.05 and 0.806, added life years and added QALYs (discounted), respectively. This results in an incremental cost effectiveness ratio (ICER) of £[redacted] per QALY gained.

The proposed PAS would make belimumab available at [redacted] to the current opportunity cost to the NHS of providing rituximab (drug acquisition cost and administration) (Table 6).

Table 6. Drug acquisition costs

	Belimumab	**Rituximab**
Dose	10 mg/kg infusion (given over 1 hour) on days 0, 14 and 28, and at 4-week intervals thereafter	1,000mg as an infusion (given over 4-5 hours) on days, 1, 15, 168, 182 (Merrill et al. 2010a)
Price	120mg vial - £[redacted] 400mg vial - £[redacted]	10mg/ml soln in vial, 2 x 10ml=£349.25; 50ml=£873.15 (Monthly Index of Medical Specialities (MIMS) 2011)
Drug cost	£[redacted]/annum*	£6,985.20/annum

* Average annual cost based on weight distribution from BLISS trials

Although it is not possible to directly compare belimumab and rituximab, taking a conservative assumption that belimumab is at least as effective as rituximab albeit with different administration costs, one would expect belimumab to provide a [redacted] for a similar high disease activity patient population (see Table 7).

Table 7. Estimated annual budget impact for the NHS in England and Wales

	Year 1	Year 2	Year 3	Year 4	Year 5
	2011	2012	2013	2014	2015
Total eligible population	4151	4842	5114	5388	5663
Likely usage (50%)	2075	2421	2557	2694	2832
Total cost of belimumab	£21,302,157	£24,686,898	£26,066,236	£27,454,704	£28,852,348
Total cost of belimumab (PAS)	£ [redacted]	£ [redacted]	£ [redacted]	£ [redacted]	£ [redacted]
Total cost of rituximab	£16,888,144	£19,701,522	£20,807,842	£21,921,491	£23,042,503
Difference between belimumab (PAS) and rituximab	£ [redacted]	£ [redacted]	£ [redacted]	£ [redacted]	£ [redacted]

Additional considerations

Belimumab is an innovative technology as it has been developed as a targeted therapy for a specific aspect of SLE pathology through genomic science. It is designed to act on the specific pathway (BLyS) associated with immune response in SLE. There has been little therapeutic innovation in treatments for SLE, with no evidence leading to the development of new licensed treatments for several decades.

Belimumab addresses an area of significant unmet need, i.e. the management of SLE patients who have severe highly active disease despite being managed on current standard of care. This is important in addressing the short-term impact on the morbidity of highly active disease (e.g. fatigue), the impact on patients developing organ damage in the long-term (e.g. renal or cardiovascular damage) and the significant burden of side effects experienced with currently used therapies; all of which are not fully captured in the current economic modelling. As SLE is a relapsing remitting disease with long-term consequences, the full clinical benefit of belimumab may not be identified in the studies available at the time of product launch.

There are limitations of the QALY calculation that may result in certain significant and substantial health-related benefits associated with the use of belimumab in SLE not being captured.

In the Phase 3 clinical trials, belimumab plus standard of care was compared to placebo plus standard of care. The current standard of care for the management of SLE consists of relatively ‘old’ non-specific treatments (antimalarials corticosteroids and immunosuppressants), with very low acquisition costs making the demonstration of cost-effectiveness particularly challenging within the current NICE cost per QALY methodology.

The standard of care includes corticosteroids, which as mentioned above, when used chronically at high doses are associated with long-term adverse effects (osteoporosis, diabetes and cardiovascular disease). Belimumab has demonstrated corticosteroid sparing effects (reduction in corticosteroid dose) during the Phase 3 trials. This occurred against a background of a blinded clinical trial in which clinicians may have been cautious to reduce/stop corticosteroids. Due to the limited data measured in the trials and the difficulty in extrapolating to the long-term adverse effects associated with corticosteroids, there is the potential to significantly underestimate the potential benefit of avoiding chronic corticosteroid use. In addition, adverse effects associated with long-term exposure to immunosuppresants like cyclophosphamide including bladder cancer, bone marrow suppression, haematologic malignancies, infections, myelodysplasia, and infertility (Kalunian et al. 2009), therefore limiting the appropriateness of these treatments in women of childbearing age.

It is also worth considering that there is no clear association between disease activity and quality of life (e.g. fatigue). Fatigue has been identified by patients as contributing to the decrease in their quality of life. There may be significant clinical benefit in addressing fatigue associated with SLE, but this is not captured in the EQ-5D.

SLE affects patients from an age of onset of 15 to 44 years (Danchenko et al. 2006), and has a substantial impact on employment, with over half of patients

no longer working 15 years after diagnosis (Yelin et al. 2007). As these patients would otherwise have a significant portion of their working life left and the current methodology doesn't incorporate productivity loss, this is likely to result in significant benefits from the appropriate management of SLE not being accounted for.

Section A – Decision problem

Manufacturers and sponsors will be requested to submit section A in advance of the full submission (for details on timelines, see the NICE document 'Guide to the single technology appraisal (STA) process' – www.nice.org.uk). A (draft) summary of product characteristics (SPC) for pharmaceuticals or information for use (IFU) for devices, a (draft) assessment report produced by the regulatory authorities (for example, the European Public Assessment Report (EPAR)), and a (draft) technical manual for devices should be provided (see section 9.1, appendix 1).

1 Description of technology under assessment

1.1 Give the brand name, approved name and, when appropriate, therapeutic class. For devices, provide details of any different versions of the same device.

Brand name: Benlysta®

Approved name: belimumab

Therapeutic class: The ATC code for belimumab is L04AA26 (selective immunosuppressants).

1.2 What is the principal mechanism of action of the technology?

Belimumab is a human IgG1λ monoclonal antibody that binds to soluble human B-lymphocyte stimulator (BLyS; also known as B cell activating factor) and inhibits its biological activity (Baker et al. 2003).

BLyS inhibits B cell apoptosis and stimulates the differentiation of B cells into immunoglobulin-producing plasma cells. Over expression of BLyS by transgenic mice results in autoimmune-like disease (Cancro et al. 2009). Furthermore, BLyS is over expressed in patients with systemic lupus erythematosus (SLE) and other autoimmune diseases (Cheema et al. 2001; Zhang et al. 2001).

In patients with SLE followed for 2 years, BLyS levels correlated with changes in lupus disease activity as well as with elevated anti-dsDNA antibody titres, and worsening disease activity was predicted by rises in serum BLyS concentrations (Petri et al. 2008).

Inhibition of BLyS by belimumab promotes apoptosis in autoreactive B cells (Cancro et al. 2009).

1.3 Does the technology have a UK marketing authorisation/CE marking for the indications detailed in this submission? If so, give the date on which authorisation was received. If not, state current UK regulatory status, with relevant dates (for example, date of application and/or expected approval dates).

No. A Marketing Authorisation Application (MAA) was filed with the European Medicines Agency (EMA) on 4th June 2010 and is now under review via the Centralised procedure. CHMP opinion is expected in May 2011 followed by a Commission decision on European marketing authorisation in July 2011.

1.4 Describe the main issues discussed by the regulatory organisation (preferably by referring to the [draft] assessment report [for example, the EPAR]). If appropriate, state any special conditions attached to the marketing authorisation (for example, exceptional circumstances/conditions to the licence).

The Benlysta licence application is currently under assessment with the EMA. A copy of the EPAR will be submitted as soon as it is available.

1.5 What are the (anticipated) indication(s) in the UK? For devices, provide the (anticipated) CE marking, including the indication for use.

The proposed indication is as follows:

Reducing disease activity in adult patients with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g. positive anti-dsDNA, low complement) despite standard therapy.

1.6 Please provide details of all completed and ongoing studies from which additional evidence is likely to be available in the next 12 months for the indication being appraised.

There are a number of ongoing long-term continuation studies of intravenous (IV) belimumab in SLE (see Table 1.1).

Table 1.1. Ongoing long-term continuation studies

Study Number	Phase	Status	Patient Population	Number of subjects treated with belimumab	Duration of Treatment (Doses)
LBSL99	2	Ongoing	Subjects who completed the Phase II trial and achieved a satisfactory response	296	Continuation study (10mg/kg IV every 28 days)
HGS1006 -C1066	3	Ongoing	Subjects who completed the BLISS-76 trial in the United States through the Week 72 visit	268	Continuation study (1 or 10mg/kg IV every 28 days)
HGS1006 -C1074	3	Ongoing	Subjects who completed the BLISS-76 or BLISS-52 trial in Canada, the European Union, Asia Pacific and Latin America regions through the Week 72 or Week 48 visits, respectively	733	Continuation study (1 or 10mg/kg IV every 28 days)

1.7 If the technology has not been launched, please supply the anticipated date of availability in the UK.

The technology is currently undergoing review by the EMA. We anticipate that belimumab will be available in the UK from August 2011.

1.8 Does the technology have regulatory approval outside the UK? If so, please provide details.

Belimumab received FDA approval in the United States under the brand name Benlysta for the treatment of adult patients with active, autoantibody-positive, systemic lupus erythematosus who are receiving standard therapy on 9th March 2011.

1.9 Is the technology subject to any other form of health technology assessment in the UK? If so, what is the timescale for completion?

At this time, belimumab is not subject to any other form of health technology assessment in the UK. GSK will submit a New Product Assessment Form to the Scottish Medicines Consortium (SMC) within three months of marketing authorisation. GSK will also submit a Form A to the All Wales Medicines Strategy Group (AWMSG) on CHMP opinion from the EMA. AWMSG's response to the Form A will set out the requirement for a full submission.

1.10 For pharmaceuticals, please complete the table below. If the unit cost of the pharmaceutical is not yet known, provide details of the anticipated unit cost, including the range of possible unit costs.

Table 1.2. Unit costs of technology being appraised

Pharmaceutical formulation	Powder for concentrate for solution for infusion
Acquisition cost (excluding VAT)	£8,500/annum
Method of administration	Intravenous infusion over a one hour period
Doses	Vial sizes - 120mg, 400mg
Dosing frequency	The recommended dosage regimen is 10 mg/kg belimumab on Days 0, 14 and 28, and at 4-week intervals thereafter.
Average length of a course of treatment	Based on clinical judgement of response
Average cost of a course of treatment	To be determined
Anticipated average interval between courses of treatments	Not applicable
Anticipated number of repeat courses of treatments	Not applicable
Dose adjustments	None recommended

1.11 For devices, please provide the list price and average selling price. If the unit cost of the device is not yet known, provide details of the anticipated unit cost, including the range of possible unit costs.

Not applicable, as technology is a pharmaceutical.

1.12 Are there additional tests or investigations needed for selection, or particular administration requirements for this technology?

No additional tests or investigations are needed for selection of patients for belimumab treatment other than those currently used routinely in clinical practice.

Whilst tools such as the ACR criteria and SELENA-SLEDAI were designed largely for use in clinical trials, clinicians will be familiar with these measures and will be able to use these to guide their selection of suitable patients in clinical practice if required. Biomarkers such as ANA, anti-dsDNA and complement are routinely measured in clinical practice.

Belimumab is administered at a dose of 10mg/kg as an intravenous infusion over a one hour period on days 0, 14, and 28, and at 4 week intervals thereafter.

1.13 Is there a need for monitoring of patients over and above usual clinical practice for this technology?

No additional monitoring of patients over and above usual clinical practice is specified for treatment with belimumab.

1.14 What other therapies, if any, are likely to be administered at the same time as the intervention as part of a course of treatment?

The proposed indication for belimumab is for reducing disease activity in adult patients with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g. positive anti-dsDNA, low complement) despite standard therapy. It will therefore be administered alongside standard therapy for SLE, which in clinical trials has included

antimalarials, non-steroidal anti-inflammatory drugs (NSAIDs), corticosteroids and other immunosuppressants; and is reflective of clinical practice in the UK.

2 Context

In this background section the manufacturer or sponsor should contextualise the evidence relating to the decision problem.

2.1 Please provide a brief overview of the disease or condition for which the technology is being used. Include details of the underlying course of the disease.

Epidemiology

SLE is a chronic autoimmune, multisystem disorder with varying manifestations characterised by an unpredictable clinical course, autoantibody production, abnormal B lymphocyte function and chronic inflammation (Manson et al. 2006). The aetiology of SLE is unknown, although genetics, hormones and environmental conditions are thought to play a role (Kotzin 1996; Manson et al. 2006). SLE is approximately 10 times more common in women than men (Manson et al. 2006; Manzi 2009) and more prevalent in African-Americans, South Asians and Chinese than Caucasians (Danchenko et al. 2006; Manzi 2009). The disease onset is generally between the ages of 15 and 44 years (Danchenko et al. 2006). This suggests that SLE affects predominantly women during their childbearing years. In the UK, prevalence has been estimated at 41 per 100,000 persons (Nightingale et al. 2007).

Diagnosis

Diagnosis of SLE can be difficult. There are no definitive tests for diagnosing SLE and this is further complicated by the fact that clinical manifestations can occur in any organ system and therefore mimic other diseases with signs and symptoms which evolve over time. Therefore, patients can be referred to any specialty within secondary care for a number of specific symptoms (e.g. joint pain, skin rash) prior to a diagnosis of SLE being made. Patients spend an average of four years and see three physicians before the disease is correctly diagnosed (Manzi 2009).

The diagnosis of SLE is widely based on a set of clinical and laboratory criteria developed by the American College of Rheumatology (ACR). In order for a diagnosis SLE to be established, four of 11 clinical and laboratory criteria must be met (Gill et al. 2003). See Table 2.1 below.

Table 2.1. ACR classification criteria for SLE (ACR Ad Hoc Committee on Systemic Lupus Erythematosus Guidelines 1999)

Item	Definition
Malar rash	Fixed erythema, flat or raised, over the malar eminences, sparing the nasolabial folds
Discoid rash	Erythematous raised patches with adherent keratotic scaling and follicular plugging: atrophic scarring may occur in older lesions
Photosensitivity	Skin rash as a result of unusual reaction to sunlight, by patient history or physician observation
Oral ulcers	Oral or nasopharyngeal ulceration, usually painless, observed by a physician
Nonerosive arthritis	Involving 2 or more peripheral joints, characterised by tenderness, swelling, or effusion
Pleuritis or pericarditis	a. Pleuritis - convincing history of pleuritic pain or rub heard by a physician or evidence of pleural effusion *OR* b. Pericarditis - documented by electrocardiogram or rub or evidence of pericardial effusion
Renal disorder	a. Persistent proteinuria >0.5 gm per day or >3+ if quantitation not performed *OR* b. Cellular casts - may be red cell, haemoglobin, granular, tubular, or mixed
Neurologic disorder	a. Seizures - in the absence of offending drugs or known metabolic derangement, e.g., uraemia, ketoacidosis, or electrolyte imbalance *OR* b. Psychosis - in the absence of offending drugs or known metabolic derangement, e.g., uraemia, ketoacidosis, or electrolyte imbalance
Hematologic disorder	a. Haemolytic anaemia with reticulocytosis *OR* b. Leukopenia - <4,000/mm^3 on ≥2 occasions *OR* c. Lymphopenia - <1,500/mm^3 on ≥2 occasions *OR* d. Thrombocytopenia - <100,000/mm^3 in the absence of offending drugs
Immunologic	a. Anti-DNA: antibody to native DNA in abnormal titre

disorder	*OR* b. Anti-Sm: presence of antibody to Sm nuclear antigen *OR* c. Positive finding of antiphospholipid antibodies based on: 1) an abnormal serum level of IgG or IgM anticardiolipin antibodies, 2) a positive test result for lupus anticoagulant using a standard method, or 3) a false-positive test result for at least 6 months and confirmed by *Treponema pallidum* immobilisation or fluorescent treponemal antibody absorption test
Positive antinuclear antibody	An abnormal titre of antinuclear antibody by immunofluorescence or an equivalent assay at any point in time in the absence of drug

Clinical Manifestations

SLE is a relapsing and remitting disease. Disease activity fluctuates between periods of exacerbation (flares) and relative quiescence, affecting multiple organ systems in an unpredictable fashion (ACR Ad Hoc Committee on Systemic Lupus Erythematosus Guidelines 1999). Clinical manifestations vary widely between patients with signs and symptoms evolving over time, making it difficult to define a treatment pathway.

Many patients with SLE experience general symptoms including fatigue, malaise, fever, anorexia, weight loss, skin rash and muscle and joint pain. SLE can lead to arthritis, kidney failure, heart and lung inflammation, neuropsychiatric disease, vasculitis, severe skin rash and blood dyscrasias such as anaemia, leucopoenia and thrombocytopenia (Manson et al. 2006).

Organ-specific damage in SLE patients steadily advances over time (ACR Ad Hoc Committee on Systemic Lupus Erythematosus Guidelines 1999). Renal manifestations, neuropsychiatric disease and musculoskeletal disease are responsible for much of the morbidity directly related to SLE disease activity observed in the first 10 years (Chambers et al. 2009; Cooper et al. 2007).

Disease activity scores correlate significantly with organ damage in SLE patients (Swaak et al. 1999). Therefore, as long-term damage accrues due to persistent disease activity, it is important to be able link the short-term effect

of interventions on disease activity to the long-term impact on organ damage and mortality.

In addition to the autoimmune-mediated disease consequences of lupus, patients with SLE appear to be at high risk for other disease and therapy related morbidity, including infections, especially of the respiratory and urinary systems (Cervera et al. 2003; Goldblatt et al. 2009), atherosclerosis, vascular disease and coronary artery disease (Campbell, Jr. et al. 2008; Roman et al. 2003; Urowitz et al. 2008); and haematological and solid tumours (Bernatsky et al. 2005; Bernatsky et al. 2007; Parikh-Patel et al. 2008), as well as increased risk for mortality (Alarcon et al. 2001; Bernatsky et al. 2006). SLE is also associated with significant maternal and foetal morbidity, including spontaneous abortion, pre-eclampsia, intrauterine growth restriction, foetal death and pre-term delivery (Molad et al. 2005).

Burden of SLE

Patients with SLE have a 2.4-fold greater risk of mortality than the general population, with a higher risk of death due to cardiovascular disease (standardised mortality ratio [SMR] 1.7), non-Hodgkin's lymphoma (SMR 2.8), lung cancer (SMR 2.3), infections (SMR 9.0) and renal disease (SMR 7.9) (Bernatsky et al. 2006). The 5-year mortality risk has been estimated to be almost 10-fold higher in SLE patients compared with a control population, 9.7% in SLE patients versus <1% for controls (Campbell, Jr. et al. 2008). A 20 year old diagnosed with lupus has a 1 in 6 chance of dying by 35 years of age, most commonly from lupus disease complications or infection (Rahman et al. 2008). Although the median survival rate is 90% at 5 years, 80% at 15 years, and 70% at 20 years, surviving SLE patients suffer a significant burden of disease with associated morbidity and reduction in quality of life (QOL), as measured by the SF-8 Health Related Quality of Life Instrument (Abu-Shakra et al. 1995; Campbell, Jr. et al. 2008; Rahman et al. 2008).

There are certain patients who have highly active disease and experience a greater impact on their quality of life, while also being more likely to develop long-term organ damage. These patients are likely to consume significantly

more health care resources, requiring more frequent health care professional (HCP) visits and hospitalisations due to exacerbations in their condition. Based on Adelphi research of clinicians treating SLE patients in the UK, 22% of patients had been hospitalised in last 12 months, increasing to over 70% in patients with severe disease (GlaxoSmithKline data on file 2010).

More than half of patients develop permanent organ damage and this damage progresses over time (Danchenko et al. 2006). Renal disease is one of the commonest and most serious manifestations of SLE. Despite the overall improvement in the care of SLE in the past two decades, the prognosis of lupus nephritis remains unsatisfactory. Up to 25% of patients still develop end stage renal failure 10 years after onset of renal disease (Mok 2010).

2.2 How many patients are assumed to be eligible? How is this figure derived?

Based on the proposed licensed population for belimumab, it is estimated that 6,348 patients across England and Wales will be eligible for belimumab (see Table 2.2). However, we propose that belimumab would be used in a subgroup of SLE patients. In addition to having evidence for serological disease activity (low complement and positive anti-dsDNA), these patients also have high disease activity as indicated by a SELENA-SLEDAI disease activity score ≥10. Patients in this subgroup experienced an additional treatment effect to belimumab over and above the licensed population (see Section 5.3.7). This equates to 4,151 patients across England and Wales (see Table 2.2).

Table 2.2. Eligible patient population and proposed subgroup

Population	Numbers	Source
England and Wales population	*Females* - 28,189,291 *Males* - 27,412,029 *Total* - 55,601,320	(Office for National Statistics 2009)
Number of patients with SLE (71 patients per 100,000 females; 10 patients per 100,000 males)	*Females* - 20,014 *Males* - 2,741 *Total* - 22,756	(Nightingale et al. 2007)
Number of patients with active disease (58%)	*Total* - 13,198	(Caseload Data 2010)
Eligible Phase 3 trial population Patients over 18 years of age (92.5%)	*Total* -12,208	(Caseload Data 2010)
Proposed licensed population Patients with a high degree of disease activity (e.g. positive anti-dsDNA, low complement) (52% of Phase 3 trial population)	*Total* – 6,348	(GlaxoSmithKline data on file 2011)
Subgroup Patients with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥10 (34% Phase 3 trial population)	*Total* - 4,151	(GlaxoSmithKline data on file 2011)

2.3 Please give details of any relevant NICE guidance or protocols for the condition for which the technology is being used. Specify whether any specific subgroups were addressed.

No relevant NICE guidance exists.

2.4 Please present the clinical pathway of care that depicts the context of the proposed use of the technology. Explain how the new technology may change the existing pathway. If a relevant NICE clinical guideline has been published, the response to this question should be consistent with the guideline and any differences should be explained.

Given the diversity of clinical manifestations, the clinical pathway of care for SLE varies according to the individual and disease severity. To date, there is no accepted SLE treatment algorithm and no relevant NICE guideline exists.

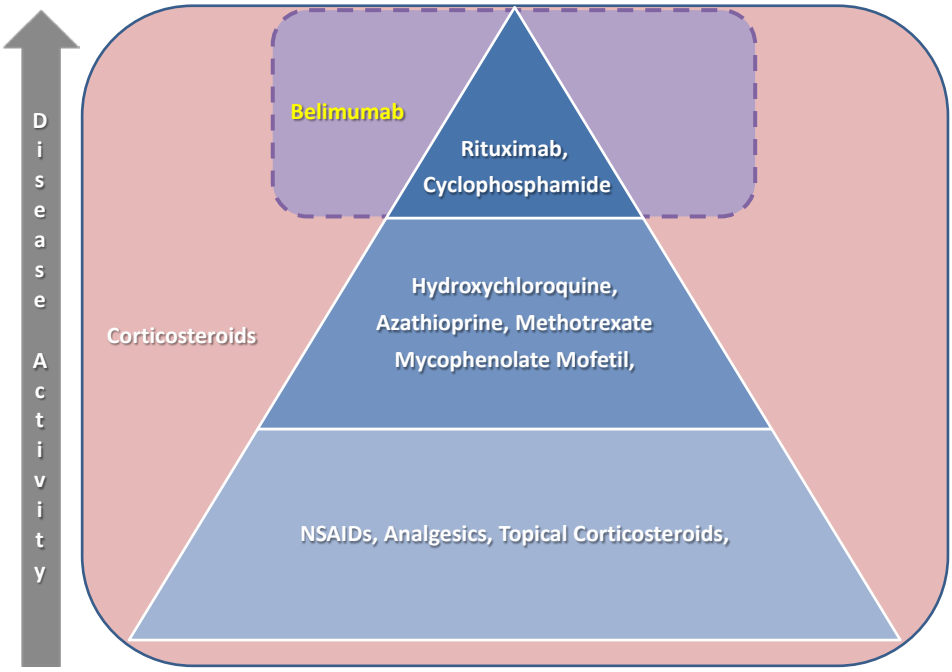
There is no cure for SLE and the aims of treatment are (Kalunian et al. 2009):

- Matching treatment to an accurate diagnosis of the extent of organ involvement
- Maintaining an appropriate level of therapy to control or halt the inflammatory disease activity while minimising side-effects and risk for infection
- Preventing further organ damage
- Maintaining a patient's daily function and quality of life

Standard therapy currently includes the use of antimalarials (hydroxychloroquine), NSAIDs, corticosteroids and immunosuppressants such as azathioprine, methotrexate and mycophenolate mofetil. Many of the treatments used for SLE are unlicensed, with only hydroxychloroquine, corticosteroids and azathioprine licensed for use in SLE. Rituximab and cyclophosphamide, although unlicensed, are used in the more severe patient population.

We have attempted to outline the clinical pathway of care for SLE in Figure 1.1. We have indicated where the belimumab trial population would be reflected within the context of this clinical pathway (dotted box). However, the proposed subgroup discussed within this submission relates to the use of belimumab in the more severe end of the clinical pathway where rituximab is currently used. Patients in the proposed subgroup are likely to be managed in tertiary centres under the care of a qualified physician experienced in the diagnosis and treatment of SLE (e.g. rheumatologist).

Figure 1.1. Clinical pathway of care



2.5 Please describe any issues relating to current clinical practice, including any variations or uncertainty about best practice.

As outlined above, there is no accepted SLE treatment algorithm and no relevant NICE guideline exists. Agreeing on best practice poses a significant challenge owing to the heterogeneous nature of SLE. There is no cure for SLE and many of the treatments used for SLE are unlicensed, with only hydroxychloroquine, corticosteroids and azathioprine licensed for use in SLE. Rituximab and cyclophosphamide, although unlicensed, are used in the more severe patient population. Some patients with advanced SLE do not respond to current treatments even at high dose.

2.6 Please identify the main comparator(s) and justify their selection.

In the two pivotal Phase 3 studies (BLISS-52 and BLISS-76), belimumab plus standard of care was compared to placebo plus standard of care. Standard of care consisted of the following (alone or in combination): antimalarials, NSAIDs, corticosteroids or other immunosuppressants (azathioprine,

methotrexate, and mycophenolate mofetil). Therefore, standard therapy alone (as defined in the BLISS-52 and BLISS-76 studies) is a relevant comparator.

Despite failing to meet primary or secondary outcomes in a Phase 2/3 SLE trial, rituximab, is used in the more severe patient population in addition to standard therapy. Therefore, rituximab plus standard therapy is a relevant comparator. The patient population and outcomes measured in the rituximab trial are not comparable to those in the belimumab trials, therefore, conducting indirect comparisons of efficacy are problematic and have not been incorporated into the cost-effectiveness model. However, the benefits of belimumab compared with rituximab will be discussed in the written submission.

Cyclophosphamide, whilst used in the more severe patient population, is largely reserved for the treatment of lupus nephritis. This is not the proposed target population for belimumab, therefore, cyclophosphamide plus standard therapy is not a relevant comparator. In addition, adverse effects associated with long-term exposure to cyclophosphamide including bladder cancer, bone marrow suppression, haematologic malignancies, infections, myelodysplasia, and infertility (Kalunian et al. 2009), limit the appropriateness of cyclophosphamide given that a high proportion of patients are women of childbearing age.

2.7 Please list therapies that may be prescribed to manage adverse reactions associated with the technology being appraised.

The most common adverse reactions observed in clinical trials (occurring in ≥1 in 100 and <1 in 10) were hypersensitivity reactions and infusion-related pyrexia.

When indicated, prescribed therapies used to manage hypersensitivity reactions included antihistamines, corticosteroids and paracetamol. Some patients received corticosteroids as pre-medication for succeeding doses.

2.8 Please identify the main resource use to the NHS associated with the technology being appraised. Describe the location of care, staff usage, administration costs, monitoring and tests. Provide details of data sources used to inform resource estimates and values.

It is anticipated that belimumab will be prescribed as part of a routine tertiary care outpatient appointment.

An administration cost of £126 per infusion of belimumab has been calculated based on two hours of senior hospital staff nurse time (£63/hr) from PSSRU Unit Costs of Health and Social Care 2010. Two hours is considered appropriate due to one hour required for the actual infusion and another hour for patient preparation and monitoring post-infusion.

No additional tests or investigations are specified for selection and monitoring of patients on treatment with belimumab other than those employed currently in routine clinical practice for the treatment of SLE.

Whilst measures such as the ACR criteria and SELENA-SLEDAI were designed largely for use in clinical trials, most clinicians will be familiar with these measures and will be able to use these to guide their selection of suitable patients in clinical practice. Biomarkers such as ANA, anti-dsDNA and complement are routinely measured in clinical practice.

2.9 Does the technology require additional infrastructure to be put in place?

Given the proposed patient subgroup, and the fact that these patients are most likely already being managed in tertiary centres, we do not anticipate at this time that additional infrastructure will be required.

2.10 Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how it might improve the way that current need is met (is this a 'step-change' in the management of the condition)?

Belimumab is an innovative technology as it has been developed as a targeted therapy for a specific aspect of SLE pathology through genomic science. It is designed to act on the specific pathway (BLyS) associated with immune response in SLE. There has been little therapeutic innovation in treatments for SLE, with no evidence leading to the development of new licensed treatments for several decades.

Belimumab addresses an area of significant unmet need, i.e. the management of SLE patients who have severe highly active disease despite being managed on current standard of care. This is important in addressing the short-term impact on the morbidity of highly active disease (e.g. fatigue), the impact on patients developing organ damage in the long-term (e.g. renal or cardiovascular damage) and the significant burden of side effects experienced with currently used therapies; all of which are not fully captured in the current economic modelling. As SLE is a relapsing remitting disease with long-term consequences, the full clinical benefit of belimumab may not be identified in the studies available at the time of product launch.

2.11 Do you consider that the use of the technology can result in any potential significant and substantial health-related benefits that are unlikely to be included in the QALY calculation?

There are certain aspects of belimumab, SLE and the potential economic evaluation that need to be considered in relation to the QALY calculation.

In the Phase 3 clinical trials, belimumab plus standard of care was compared to standard of care alone. The current standard of care for the management of SLE consists of relatively ‘old’ non-specific treatments (antimalarials corticosteroids and immunosuppressants), with very low acquisition costs making the demonstration of cost-effectiveness particularly challenging within the current NICE cost per QALY methodology.

The standard of care includes corticosteroids, which when used chronically at high doses are associated with long-term adverse effects (osteoporosis,

diabetes and cardiovascular disease). Belimumab has demonstrated corticosteroid sparing effects (reduction in corticosteroid dose) during the Phase 3 trials. This occurred against a background of a blinded clinical trial in which clinicians may have been cautious to reduce/stop corticosteroids. Due to the limited data measured in the trials and the difficulty in extrapolating to the long-term adverse effects associated with corticosteroids, it will not be possible to account for these benefits in the economic model. There is the potential to significantly underestimate the potential benefit of avoiding chronic corticosteroid use.

It is also worth considering that there is no clear association between disease activity and quality of life (e.g. fatigue). Fatigue has been identified by patients as contributing significantly to the decrease in their quality of life, but this is currently not well reflected in the EQ-5D measure and therefore not captured in the cost effectiveness model.

EQ-5D may not be the most sensitive measure to assess the true impact of the disease on HRQL experienced by SLE patients. Patients may experience disease flares at any time and not necessarily at the time the EQ-5D was completed for the pre-defined time points of the clinical trials. In addition, certain relevant dimensions of health that are not directly included in the EQ-5D instrument, such as fatigue or sensory impairment, or where the disease course is characterised by flares of unpredictable symptom severity (Wailoo et al. 2010).

A critical aspect of the management of lupus is the impact of SLE on long-term organ damage. Although the Phase 3 clinical trials collected data on organ damage (SLICC scores), this is unlikely to be fully reflective of belimumab's impact on long-term damage, due to the design and duration of the trial. This will need to be modelled via belimumab’s effects on disease activity, with the inherent uncertainties of long-term modelling.

Given the demographic of patients suffering from SLE, age of onset 15 to 44 years (Danchenko et al. 2006), it has a substantial impact on employment,

with over half of patients no longer working 15 years after diagnosis (Yelin et al. 2007).

2.12 Please identify the nature of the data which you understand to be available to enable the Appraisal Committee to take account of these benefits.

As regards fatigue, the FACIT-Fatigue questionnaire was included in the clinical trials and the results will be presented as part of the submission evidence. In addition, SLE patients have reported the substantial impact this chronic symptom has on their quality of life.

Longitudinal data from the Johns Hopkins Lupus Cohort has been used to estimate natural history models that describe the progress of SLE outcomes over a long follow-up period. The Johns Hopkins Lupus Cohort reports data on a large population (2,047) of SLE patients from Baltimore, Maryland. The anticipated long-term effect of belimumab treatment has been modelled using the clinical events and outcomes recorded in this longitudinal database in order to estimate the long-term benefits belimumab may provide by reducing organ damage in severe SLE patients.

3 Equity and equality

NICE considers equity in terms of how the effects of a health technology may deliver differential benefits across the population. Evidence relevant to equity considerations may also take a variety of forms and come from different sources. These may include general-population-generated utility weightings applied in health economic analyses, societal values elicited through social survey and other methods, research into technology uptake in different population groups, evidence on differential treatment effects in different population groups, and epidemiological evidence on risks or incidence of the condition in different population groups.

3.1 Identification of equity and equalities issues

3.1.1 Please specify any issues relating to equity or equalities in NICE guidance, or protocols for the condition for which the technology is being used.

There are no NICE guidance or protocols in existence for the condition for which the technology is being used.

3.1.2 Are there any equity or equalities issues anticipated for the appraisal of this technology (consider issues relating to current legislation and any issues identified in the scope for the appraisal)?

As NICE has noted in the final scope, SLE is more prevalent in women and African-Caribbean, South Asian and Chinese populations than in European white populations. The demographic of SLE patients is likely to include a significant portion of women of child-bearing age.

SLE has a substantial impact on employment, with over half of patients no longer working 15 years after diagnosis (Yelin et al. 2007). Patients with SLE experience a reduced Health-Related Quality of Life (HRQL) compared with healthy individuals (Lau et al. 2009). Effects on HRQL are similar to that in patients with other chronic, debilitating diseases such as congestive heart failure and depression (Jolly 2005). Fatigue is one of the most prevalent

clinical manifestations of SLE and severely affects HRQL (Thumboo et al. 2007; Zonana-Nacach et al. 2000b).

3.1.3 How have the clinical and cost-effectiveness analyses addressed these issues?

The pooled data from the BLISS-52 and BLISS-76 studies was used in the cost-effectiveness model. This pooled dataset comprised 94% females; 9% of patients were of black African-American ethnicity and 21% of Asian ethnicity. A priori subgroup efficacy analyses for gender and ethnicity were conducted, and most subgroups demonstrated similar benefit in the primary outcomes with belimumab.

The efficacy results from the subgroup analyses of gender, age and race will be discussed in the clinical section of the submission document (see Section 5.5). No separate cost-effectiveness analyses have been conducted in these subgroups.

4 Statement of the decision problem

In this section the manufacturer or sponsor should specify the decision problem that the submission addresses. The decision problem should be derived from the final scope issued by NICE and should state the key parameters that the information in the evidence submission will address.

Table 4.1. Statement of the decision problem

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
Population	Adults with active autoantibody-positive systemic lupus erythematosus	**Phase 3 Trial Population** Adults with active autoantibody-positive systemic lupus erythematosus. **High Disease Activity Subgroup** Adults with active autoantibody-positive systemic lupus erythematosus with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥10.	Mindful of NHS resources, the proposed population of interest to this decision problem is a subgroup of the Phase 3 trial population which applies the additional criteria of evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI disease activity score of ≥10, This subgroup experienced an additional treatment effect to belimumab over and above the Phase 3 trial population and is aimed at identifying SLE patients at the greatest risk of experiencing long-term organ damage.
Intervention	Belimumab as an add on to standard therapy	Belimumab 10mg/kg administered as an intravenous infusion over a one hour period on days 0, 14 and 28, and at 4	

		week intervals thereafter in addition to standard therapy.	
Comparator(s)	• Standard therapy alone; For people in whom it is considered appropriate: • Rituximab plus standard therapy • Cyclophosphamide plus standard therapy	• Standard therapy which comprises (alone or in combination): antimalarials, NSAIDs, corticosteroids, or other immunosuppressants (azathioprine, methotrexate, and mycophenolate mofetil). • Rituximab plus standard therapy for the more severe SLE sub-population	Despite failing to meet primary or secondary outcomes in a Phase 2/3 SLE trial, rituximab, is used in the more severe patient population in addition to standard therapy. Therefore, rituximab plus standard therapy is a relevant comparator. The patient population and outcomes measured are not comparable to those in the belimumab trials. Therefore, conducting indirect comparisons of efficacy are problematic and have not been incorporated into the cost-effectiveness model. However, the benefits of belimumab compared with rituximab will be discussed in the written submission. Cyclophosphamide, whilst used in the more severe patient population, is largely reserved for the treatment of lupus nephritis. This is not the proposed target population for belimumab, therefore, cyclophosphamide plus standard therapy is not a relevant comparator. In addition, adverse effects associated with long-term exposure to cyclophosphamide including bladder cancer, bone marrow suppression, haematologic malignancies,

			infections, myelodysplasia, and infertility (Kalunian et al. 2009), limit the appropriateness of cyclophosphamide given that a high proportion of patients are women of childbearing age.
Outcomes	The outcome measures to be considered include: • disease activity • incidence and severity of flares • mortality • health-related quality of life, including fatigue • adverse effects of treatment	The outcome measures included in the cost-effectiveness model are: • Disease activity • Incidence and severity of flares • Mortality • Health-related quality of life • Disease progression in terms of long-term organ damage – As discussed at the scoping workshop, although not collected in the clinical trials, long-term organ damage will be considered in the assessment of cost-effectiveness based on modelled data from the Johns Hopkins Lupus Cohort. Additional endpoints discussed in the written submission and not included in the health economic model are: • Fatigue - In the Phase 3 trials this was	Adverse effects of treatment have not been included in the base case economic model as significant differences between treatments were not noted from the two pivotal Phase 3 trials. The side effect profile of belimumab will be discussed in the clinical section of the submission.

		measured using the FACIT-Fatigue instrument and was reported as the mean change in scale score at Weeks 12, 24, 52 and 76 (BLISS-76 only). • Adverse events of treatment	
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services	• Cost effectiveness will be expressed in terms of incremental cost per quality-adjusted life year. • The time horizon for the model will be lifetime. • Costs will be considered from an NHS and Personal Social Services perspective.	Not applicable.

	perspective.		
Subgroups to be considered	None outlined in scope.	See population section above.	See population section above.
Special considerations, including issues related to equity or equality	None outlined in scope.	It will be important to acknowledge the innovative nature of belimumab in the treatment of SLE. There is a limitation with the current cost per QALY methodology not able to capture all the benefits of belimumab (i.e. avoidance of corticosteroids, impact of fatigue and loss of productivity). SLE has a significantly greater impact on certain ethnic groups and is most prevalent in woman of childbearing age.	

Section B – Clinical and cost effectiveness

When estimating clinical and cost effectiveness, particular emphasis should be given to adhering to the 'reference case' (see the NICE document 'Guide to the methods of technology appraisal' – www.nice.org.uk). Reasons for deviating from the reference case should be clearly explained. Particularly important features of the reference case include those listed in the table below.

Element of health technology assessment	**Reference case**	**Section in 'Guide to the methods of technology appraisal'**
Defining the decision problem	The scope developed by NICE	5.2.5 and 5.2.6
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	5.2.5 and 5.2.6
Perspective costs	NHS and PSS	5.2.7 to 5.2.10
Perspective benefits	All health effects on individuals	5.2.7 to 5.2.10
Type of economic evaluation	Cost-effectiveness analysis	5.2.11 and 5.2.12
Synthesis of evidence on outcomes	Based on a systematic review	5.3
Measure of health effects	QALYs	5.4
Source of data for measurement of HRQL	Reported directly by patients and carers	5.4
Source of preference data for valuation of changes in HRQL	Representative sample of the public	5.4
Discount rate	An annual rate of 3.5% on both costs and health effects	5.6
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	5.12
HRQL, health-related quality of life; NHS, National Health Service; PSS, Personal Social Services; QALY(s), quality-adjusted life year(s)		

5 Clinical evidence

Manufacturers and sponsors are requested to present clinical evidence for their technology in the following sections. This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 3 and 5.3.1 to 5.3.8.

5.1 Identification of studies

5.1.1 Describe the strategies used to retrieve relevant clinical data, both from the published literature and from unpublished data that may be held by the manufacturer or sponsor. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used should be provided. Exact details of the search strategy used should be provided in section 9.2, appendix 2.

A systematic review of the published literature was conducted to identify all relevant published randomised controlled trials (RCTs) for belimumab and relevant comparators in SLE. Searches for non-RCTS for belimumab are described in Section 5.8 and Section 9.7.1, appendix 7. The following description of the search strategies includes searches for both belimumab and for comparator products according to the scope of the systematic review. The PICO method was used to develop the search strategies with reference to the decision problem and combined intervention search terms with terms for the specific disease area. Systematic searching in standard databases was supplemented with hand searches of reference lists and relevant conference proceedings. The inclusion and exclusion criteria were chosen to identify all relevant RCTs. Details of searched databases and results are presented in Section 9.2, appendix 2. The systematic review identified 3,774 citations in total and included 39 full publications and 4 conference proceedings.

5.2 Study selection

5.2.1 Describe the inclusion and exclusion selection criteria, language restrictions and the study selection process. A justification should be provided to ensure that the rationale is transparent. A suggested format is provided below.

The following inclusion and exclusion criteria were applied to identify only those studies relevant to the decision problem. Conference proceedings superseded by other conference proceedings or by full publications concerning the same trial were only included if they presented new data that had not been published in the later publications. See also Section 9.2.6. appendix 2.

Table 5.1. Eligibility criteria used in search strategy

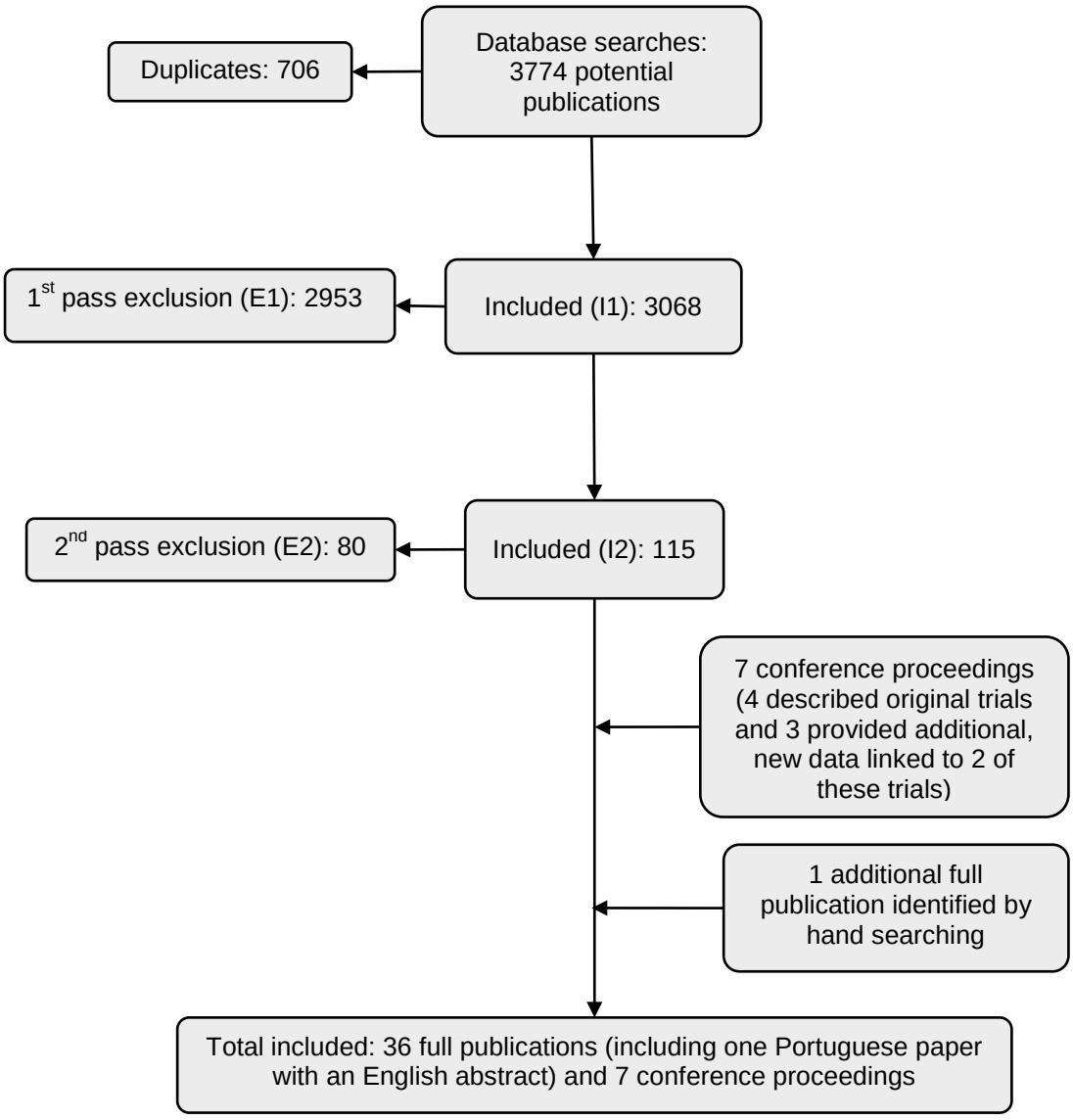
	Clinical effectiveness
Inclusion criteria	Population - Adults (≥ 18 years) with systemic lupus erythematosus (SLE); studies were also included for SLE patients with kidney involvement - Interventions o Belimumab o Rituximab o Mycophenolate mofetil o Prednisolone and other steroids o Hydroxychloroquine and other antimalarials o Azathioprine o Cyclophosphamide o Methotrexate Outcomes - Change in SELENA-SLEDAI score (Safety of Estrogens in Lupus Erythematosus National Assessment Systemic Lupus Erythematosus Disease Activity Index) - Change in BILAG score (British Isles Lupus Assessment Group) - Change in PGA (physician global assessment scale) - Change in SLICC score (Systemic Lupus International Collaborating Clinics) - Change in number/frequency of flares - Quality of life - Reduction in steroid use - Medical resource utilisation - Fatigue (e.g. FACIT, Functional Assessment of Chronic Illness Therapy score) - Adverse events including: o Incidence and severity (grade) of all adverse

	events (AEs) reported o Withdrawals due to AEs o Mortality o SAEs Study design - RCT, both cross-over and parallel, blinded and open-label designs Language restrictions - Only English publications (if only the abstract was in English, this would be included)
Exclusion criteria	Population - Studies enrolling patients with only active lupus nephritis were excluded Interventions - Non-specified Outcomes - Non-specified Study design - Designs other than RCT Language restrictions - Publications in languages other than English

5.2.2 A flow diagram of the numbers of studies included and excluded at each stage should be provided using a validated statement for reporting systematic reviews and meta-analyses such as the QUOROM statement flow diagram (www.consort-statement.org/?o=1065). The total number of studies in the statement should equal the total number of studies listed in section 5.2.4.

The QUORUM diagram below presents the results of the systematic review which had as its scope RCTs of both the intervention and the competitors. The number of included publications was 43 (36 full publications plus seven abstracts), including eight publications (of four trials) of belimumab and 35 publications of other interventions.

Figure 5.1. Flow diagram of included and excluded studies



5.2.3 When data from a single RCT have been drawn from more than one source (for example, a poster and a published report) and/or when trials are linked (for example, an open-label extension to an RCT), this should be made clear.

As stated in the QUORUM diagram in Section 5.2.2, three additional conference proceedings were identified, presenting new data from included studies. The manner in which the conference proceedings of the Phase 3

belimumab trials are linked is presented in Table 1 below. Two other linked publications of the Phase 2 belimumab trial LBSL02 are also presented. Table 5.2 presents linked publications of competitor drugs that were also included in the systematic review.

Table 5.2. Linked publications on belimumab

Study	Linked to:	Trial name
(Tanasescu et al. 2010) (abstract)	(Navarra et al. 2010) (abstract)	C1057 (BLISS-52)
(D'Cruz et al. 2010) (abstract)	(Navarra et al. 2010) (abstract)	C1057 (BLISS-52)
(Petri et al. 2010) (abstract)	(Furie et al. 2010) (abstract)	C1056 (BLISS-76)
(Wallace et al. 2009) (full publication)	(Furie et al. 2009) (full publication)	LBSL02

Table 5.3. Linked publications on competitor products

Study	Linked to:	Trial name
(Bykerk et al. 1991)	(Tsakonas et al. 1998)	The Canadian Hydroxychloroquine Study Group

Where the systematic review identified publications based on GlaxoSmithKline studies, we have augmented information in this submission with unpublished data.

Complete list of relevant RCTs

5.2.4 Provide details of all RCTs that compare the intervention with other therapies (including placebo) in the relevant patient group. The list must be complete and will be validated by independent searches conducted by the Evidence Review Group. This should be presented in tabular form. A suggested format is presented below.

Table 5.4. List of relevant RCTs

Trial no. (acronym)	Intervention	Comparator	Population	Primary study ref.
LBSL01 Phase 1	Standard of care plus belimumab 1 mg/kg or belimumab 4 mg/kg or belimumab 10 mg/kg or belimumab 20 mg/kg	Standard of care plus placebo	Adult patients (age ≥18 years) fulfilling the American College of Rheumatology criteria for SLE were enrolled in the trial. Eligible patients had stable SLE disease activity, as clinically judged by the principal investigator, for at least 2 months before screening and were either maintained with no medication or with a stable treatment regimen of low-dose (≤ 15 mg) prednisone, antimalarials, non-steroidal anti-inflammatory drugs, methotrexate, azathioprine, or mycophenolate mofetil. Patients were required to have a history of measurable anti-dsDNA, anti-Sm, anti-RNP, anti-cardiolipin, anti-Ro, or anti-La autoantibodies.	(Furie et al. 2008)
LBSL02 Phase 2	Standard of care plus belimumab 1 mg/kg or belimumab 4 mg/kg or belimumab 10 mg/kg	Standard of care plus placebo	Adult patients (age ≥18 years) fulfilling the American College of Rheumatology criteria for SLE who had active disease as defined by a SELENA–SLEDAI score of ≥4 at screening were eligible for enrolment. Inclusion criteria mandated a history of measurable autoantibodies (including any of the following: antinuclear antibodies [ANAs], anti-dsDNA, anti-Sm, anti-RNP, anti-Ro, anti-La, or anticardiolipin), but they did not have to be present at screening. In addition, adult patients were required to be receiving a stable regimen of prednisone (5–40 mg/day),	(Wallace et al. 2009)

			antimalarials, or immunosuppressant agents for at least 60 days prior to day 0 (first dose).	
C1057 (BLISS-52) Phase 3	Standard of care plus belimumab 1 mg/kg or belimumab 10 mg/kg	Standard of care plus placebo	Adult patients (aged ≥18 years) fulfilling the ACR criteria for SLE who had active disease as defined by a SELENA-SLEDAI of score ≥6 at screening were eligible for enrolment. Other inclusion criteria were unequivocally positive ANA (titre ≥1:80) or anti-dsDNA antibody (≥30 IU/mL), and a stable treatment regimen with fixed doses of prednisone (0–40 mg/day), or non-steroidal anti-inflammatory, antimalarial, or immunosuppressant drugs for at least 30 days before the first study dose.	(Navarra et al. 2011)
C1056 (BLISS-76) Phase 3	As per BLISS-52.	As per BLISS-52.	As per BLISS-52.	(GlaxoSmithKline data on file 2011; van Vollenhoven et al. 2010)

5.2.5 Please highlight which of the RCTs identified above compares the intervention directly with the appropriate comparator(s) with reference to the decision problem. If there are none, please state this.

All of the Phase 2/3 trials directly compare the intervention, belimumab plus standard therapy, with placebo plus standard therapy. None of the above trials directly compare belimumab with rituximab plus standard therapy or cyclophosphamide plus standard therapy.

5.2.6 When studies identified above have been excluded from further discussion, a justification should be provided to ensure that the rationale for doing so is transparent. For example, when studies have been identified but there is no access to the level of trial data required, this should be indicated.

Both the Phase 1 and Phase 2 belimumab studies have been excluded from further discussion. The reasons for this are outlined below.

The Phase 1 study (LBSL01) was designed to evaluate the safety, tolerability, immunogenicity, pharmacokinetics and pharmacodynamics of 4 doses (1, 4, 10, 20 mg/kg) of belimumab or placebo administered as either a single IV infusion or 2 infusions 21 days apart. As this was a small (n=70) exploratory study of limited duration, designed primarily to demonstrate safety and tolerability in humans, it does not reflect the proposed clinical use of belimumab and therefore will be excluded from further discussion.

The Phase 2 study (LBSL02) was conducted in 449 subjects with SLE who were randomised to placebo or 1, 4, or 10 mg/kg belimumab administered by IV infusion on Days 0, 14 and 28 and every 28 days thereafter for 48 weeks with a final assessment at Week 52. The primary endpoints of this study were percent change in SELENA-SLEDAI score at Week 24 and time to flare as measured by the SLE flare index (SFI) over 52 weeks.

Whilst the primary endpoints of this study were not met, post-hoc analyses of the data from this trial identified a large subgroup of subjects (72%) with autoantibody (antinuclear antibody and/or anti-dsDNA antibody) positive disease, in whom belimumab appeared to offer benefit. This autoantibody-positive population was selected as the population for the Phase 3 studies. Furthermore, the data from this Phase 2 trial guided the development of a novel composite response endpoint, the SLE Response Index or SRI, which was selected as the primary efficacy endpoint (at Week 52) in the Phase 3 studies.

The results of the Phase 2 study therefore, although supportive, do not contribute substantially to the assessment of efficacy given differences in the patient population, primary efficacy endpoint used and lack of SLE background medication control compared with the Phase 3 trials. For these reasons we will exclude the Phase 2 study from further discussion and focus on the 2 pivotal Phase 3 studies. Please note that the Phase 2 study is included within the adverse events section (Section 5.9) for completeness of safety evaluation.

List of relevant non-RCTs

5.2.7 Please provide details of any non-RCTs (for example experimental and observational data) that are considered relevant to the decision problem and a justification for their inclusion. Full details should be provided in section 5.8 and key details should be presented in a table; the following is a suggested format.

Table 5.5. List of relevant non-RCTs

Trial no. (acronym)	Intervention	Population	Objectives	Primary study ref.	Justification for inclusion
LBSL99 (Phase 2 Continuation Study for Protocol LBSL02)	Standard of care plus belimumab 10 mg/kg	Adult patients (age ≥18 years) with SLE who completed LBSL02 and achieved a satisfactory response	• To provide continuing treatment to subjects with SLE who achieved a satisfactory response in LBSL02. • To evaluate the long-term safety of belimumab in SLE patients.	(GlaxoSmithKline data on file 2011)	This study provides long-term safety and efficacy results of belimumab in SLE patients.

5.3 Summary of methodology of relevant RCTs

5.3.1 As a minimum, the summary should include information on the RCT(s) under the subheadings listed in this section. Items 2 to 14 of the CONSORT checklist should be provided, as well as a CONSORT flow diagram of patient numbers (www.consort-statement.org). It is expected that all key aspects of methodology will be in the public domain; if a manufacturer or sponsor wishes to submit aspects of the methodology in confidence, prior agreement must be requested from NICE. When there is more than one RCT, **the information should be tabulated**.

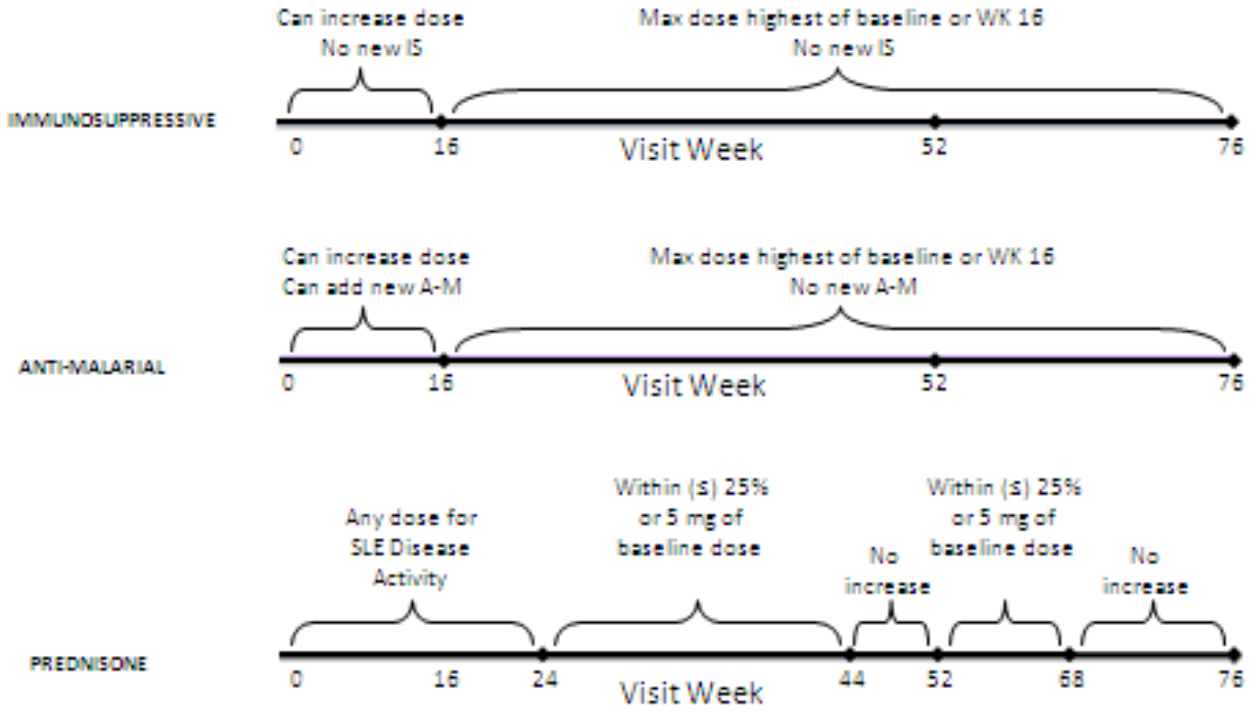
Methods

5.3.2 Describe the RCT(s) design (for example, duration, degree and method of blinding, and randomisation) and interventions. Include details of length of follow-up and timing of assessments. The following tables provide a suggested format for when there is more than one RCT.

Table 5.6. Comparative summary of methodology of the RCTs

Trial no. (acronym)	**C1057 (BLISS-52)**	**C1056 (BLISS-76)**
Location	90 centres in 13 countries in Latin America (Argentina, Brazil, Chile, Colombia and Peru), Asia-Pacific (Australia, Hong Kong, India, Korea, Philippines and Taiwan) and eastern Europe (Romania and Russia).	136 centres in 19 countries in North America (Canada, Costa Rica, Mexico, Puerto Rico and US) and Europe (Austria, Belgium, Czech Republic, France, Germany, Israel, Italy, The Netherlands, Poland, Romania, Slovakia, Spain, Sweden and UK).
Design	Randomised, double-blind, placebo-controlled, parallel-group study.	As per BLISS-52.
Duration of	52 weeks	76 weeks (primary end point at 52 weeks)

study		
Method of randomisation	Patients who underwent all screening procedures and met the entry criteria were enrolled in the study and assigned to treatment by use of a central interactive voice response system. Patients were randomised in a 1:1:1 ratio to placebo, or belimumab 1 mg/kg or 10 mg/kg. Randomisation was stratified according to the SELENA-SLEDAI score (6–9 vs ≥10), proteinuria concentration (<2 g/24 h vs ≥2 g/24 h) at screening, and ethnic origin (African descent or indigenous American [Alaska Native or American Indian from North, South, or Central America] vs other).	As per BLISS-52.
Method of blinding (care provider, patient and outcome assessor)	Patients, investigators, study coordinators, and sponsors were masked to treatment assignment during intravenous administration of the drug and assessment of the patients every 4 weeks during the trial until the database was locked. An unmasked pharmacist prepared unmarked infusion bags for administration. Belimumab and placebo were both prepared as sterile and lyophilised vials (5 mL for belimumab 1 mg/kg; 20 mL for belimumab 10 mg/kg and placebo), and contained the same formulations, except without the active drug for placebo.	As per BLISS-52.
Intervention(s) (n =) and comparator(s) (n =)	Standard of care plus belimumab 1mg/kg (n=288) or belimumab 10mg/kg (n=290) or placebo (n=287) administered by IV infusion on Days 0, 14 and 28 and every 28 days thereafter for 48 weeks. Standard of care consisted of the following (alone or in combination): antimalarials, NSAIDs, corticosteroids or other immunosuppressants (azathioprine, methotrexate, and mycophenolate mofetil).	Standard of care plus belimumab 1mg/kg (n=271) or belimumab 10mg/kg (n=273) or placebo (n=275) administered by IV infusion on Days 0, 14 and 28 and every 28 days thereafter for 72 weeks. Standard of care consisted of the following (alone or in combination): antimalarials, NSAIDs, corticosteroids or other immunosuppressants (azathioprine, methotrexate, and mycophenolate mofetil).
Progressive restrictions placed on	In both BLISS-52 and BLISS-76, progressive restrictions were placed on standard of care as the study progressed. These are outlined in the Figure 5.2 below.	

standard of care	**Figure 5.2. Progressive restrictions placed on standard of care**  	
Primary outcomes (including scoring methods and timings of assessments)	The primary efficacy endpoint was the response rate at week 52, assessed with SLE Responder Index (SRI). With the SRI criteria, a responder was defined as having a reduction of at least 4 points in the SELENA-SLEDAI score (defined as clinically meaningful) (Gladman et al. 2000), no new BILAG A organ domain score, no more than 1 new BILAG B organ domain score, and no worsening in PGA score (increase <0.3) at week 52 compared with baseline.	As per BLISS-52.

Secondary outcomes (including scoring methods and timings of assessments)	Major secondary endpoints: • Percent of subjects with ≥ 4-point reduction in SELENA-SLEDAI at Week 52. • Mean change in PGA at Week 24. • Percent of subjects with prednisone (equivalent) reduction ≥ 25% from baseline to ≤ 7.5 mg/day during Weeks 40 – 52 (in subjects whose prednisone equivalent dose was > 7.5 mg/day at baseline). • Mean change in SF-36 PCS at Week 24.	Major secondary endpoints: • As per BLISS-52. • Additionally, response rate (SRI) at Week 76.
Duration of follow-up	52 or 56 weeks dependent on participation in the continuation protocol.	76 or 80 weeks dependent on participation in the continuation protocol.

Participants

5.3.3 Provide details of the eligibility criteria (inclusion and exclusion) for the trial. The following table provides a suggested format for the eligibility criteria for when there is more than one RCT. Highlight any differences between the trials.

Table 5.7. Eligibility criteria in the RCTs

Trial no. (acronym)	Inclusion criteria	Exclusion criteria
C1057 (BLISS-52)	Adult patients (aged ≥18 years) who met the American College of Rheumatology criteria for systemic lupus erythematosus and had active disease (score ≥6 at screening on SELENA-SLEDAI) were eligible for enrolment. Other inclusion criteria were unequivocally positive ANA (titre ≥1:80) or anti-dsDNA antibody (≥30 IU/mL), and a stable treatment regimen with fixed doses of prednisone (0–40 mg/day), or non-steroidal anti-inflammatory, antimalarial, or immunosuppressant drugs for at least 30 days before the first study dose.	The main exclusion criteria were severe active lupus nephritis or CNS lupus; pregnancy; and previous treatment with any B-lymphocyte-targeted drug (including rituximab), intravenous cyclophosphamide within 6 months of enrolment, and intravenous Ig or prednisone (>100 mg/day) within 3 months.
C1056 (BLISS-76)	As per BLISS-52.	As per BLISS-52.
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee		

5.3.4 Describe the patient characteristics at baseline. Highlight any differences between study groups. The following table provides a suggested format for the presentation of baseline patient characteristics for when there is more than one RCT.

Table 5.8 shows selected demographic characteristics in the Phase 3 trials. The Phase 3 studies included mostly young females (74% ≤ 45 years of age; 94% female), a population that is representative of patients with SLE. The studies were performed in antibody-positive (positive antinuclear antibody titre ≥ 1:80 and/or anti-dsDNA antibodies ≥ 30 IU/mL) SLE patients with active SLE disease (SELENA-SLEDAI ≥ 6 at screening).

Across the 2 studies, 47% of patients were white, 23% American Indian, 21% Asian, and 8.8% black. There were differences in the racial profiles between the 2 studies that reflect the racial distributions in the geographic regions in which the trials were conducted. The BLISS-76 study was conducted primarily in North America and Europe and enrolled predominantly white patients (70%), while the BLISS-52 study was conducted primarily in South America, Asia-Pacific and Eastern Europe and enrolled predominantly Asian (38%) and American Indian (32%) patients. In BLISS-76, which included over 50% representation from the United States, approximately 14% of the population was black compared with 3.5% in BLISS-52. Patients of Hispanic descent comprised 49% and 21% of the population of BLISS-52 and BLISS-76, respectively.

Table 5.8. Selected demographic characteristics in Phase 3 trials

	BLISS-52				BLISS-76				Pooled Total Population			
	Placebo N = 287	1 mg/kg N = 288	10 mg/kg N = 290	All N = 865	Placebo N = 275	1 mg/kg N = 271	10 mg/kg N = 273	All N = 819	Placebo N = 562	1 mg/kg N = 559	10 mg/kg N = 563	All N = 1684
Sex												
Female	270 (94.1%)	271 (94.1%)	280 (96.6%)	821 (94.9%)	252 (91.6%)	253 (93.4%)	259 (94.9%)	764 (93.3%)	522 (92.9%)	524 (93.7%)	539 (95.7%)	1585 (94.1%)
Age (years) Mean ± SD	36.2 ± 11.8	35.0 ± 10.6	35.4 ± 10.8	35.5 ± 11.1	40.0 ± 11.9	40.0 ± 11.4	40.5 ± 11.1	40.2 ± 11.5	38.1 ± 12.0	37.4 ± 11.3	37.9 ± 11.3	37.8 ± 11.5
≤ 45	225 (78.4%)	236 (81.9%)	236 (81.4%)	697 (80.6%)	189 (68.7%)	184 (67.9%)	178 (65.2%)	551 (67.3%)	414 (73.7%)	420 (75.1%)	414 (73.5%)	1248 (74.1%)
Race[1]												
White	82 (28.6%)	76 (26.4%)	71 (24.5%)	229 (26.5%)	188 (68.4%)	192 (70.8%)	189 (69.2%)	569 (69.5%)	270 (48.0%)	268 (47.9%)	260 (46.2%)	798 (47.4%)
Asian	105 (36.6%)	106 (36.8%)	116 (40.0%)	327 (37.8%)	11 (4.0%)	6 (2.2%)	11 (4.0%)	28 (3.4%)	116 (20.6%)	112 (20.0%)	127 (22.6%)	355 (21.1%)
Black	11 (3.8%)	8 (2.8%)	11 (3.8%)	30 (3.5%)	39 (14.2%)	40 (14.8%)	39 (14.3%)	118 (14.4%)	50 (8.9%)	48 (8.6%)	50 (8.9%)	148 (8.8%)
Alaska Native or American Indian from North/Central/ South America	89 (31.0%)	98 (34.0%)	92 (31.7%)	279 (32.3%)	36 (13.1%)	33 (12.2%)	34 (12.5%)	103 (12.6%)	125 (22.2%)	131 (23.4%)	126 (22.4%)	382 (22.7%)
Hispanic or Latino origin	143 (49.8%)	141 (49.0%)	136 (46.9%)	420 (48.6%)	55 (20.0%)	62 (22.9%)	56 (20.5%)	173 (21.1%)	198 (35.2%)	203 (36.3%)	192 (34.1%)	593 (35.2%)

[1] Patients who checked more than 1 race category are counted under individual race category according to the minority rule as well as the multiracial category.

Table 5.9 shows selected baseline disease activity parameters in the Phase 3 trials. Patients in the Phase 3 trials had SLE for a mean duration of 6.4 years. The baseline SELENA-SLEDAI score in the Phase 3 trials was ≥ 10 for 52% of patients, ≤ 9 for 48% of patients, with a mean score of 9.71. Approximately 61% of patients had at least 1A or 2B BILAG organ domain scores and almost 16% had at least 1 BILAG A organ domain score. Over 80% of patients had baseline PGA scores of less than 2. The mean SLICC/ACR damage index was 0.77 at baseline, with over 50% of patients having a baseline score of 0. Approximately 6% of patients had proteinuria of 2 g/24h or more.

Disease characteristics of patients in BLISS-52 and BLISS-76 were similar across the 2 studies and generally balanced between treatment groups within the studies with a few exceptions. Patients in BLISS-52 had more proteinuria and more had at least 1A BILAG organ domain score. However, patients in BLISS-76 had longer disease duration and more organ damage as reflected by higher SLICC damage scores. Within trials, baseline disease activity was generally balanced across treatment groups.

Organ systems involved at baseline in greater than 10% of patients were: mucocutaneous (82%), immunology (80%), musculoskeletal (65%) and renal (16%). Greater than 50% of patients had 3 or more organ systems involved at baseline. Individual SELENA-SLEDAI manifestations present in greater than 5% of patients across both studies were increased DNA binding (69%), arthritis (65%), rash (63%), low complement (62%), alopecia (49%), mucosal ulcers (23%), proteinuria (13%), leukopenia (6.1%), vasculitis (6.6%) and pleurisy (5.2%). The most commonly involved organ systems at baseline based on BILAG were similar (data not shown). This profile of organ system involvement is typical of the general SLE population (Gordon et al. 2003; Hay et al. 1993).

Table 5.9. Selected baseline disease characteristics in Phase 3 trials

	BLISS-52				BLISS-76				Pooled Total Population			
	Placebo N = 287	1 mg/kg N = 288	10 mg/kg N = 290	All N = 865	Placebo N = 275	1 mg/kg N = 271	10 mg/kg N = 273	All N = 819	Placebo N = 562	1 mg/kg N = 559	10 mg/kg N = 563	All N = 1684
SLE Disease duration (yr)[1]												
Mean ± SD	5.93 ± 6.17	4.96 ± 4.58	5.03 ± 5.07	5.31 ± 5.32	7.42 ± 6.72	7.93 ± 7.13	7.20 ± 7.45	7.52 ± 7.10	6.66 ± 6.48	6.40 ± 6.13	6.08 ± 6.42	6.38 ± 6.35
SELENA SLEDAI score												
≥ 10	158 (55.1%)	139 (48.3%)	160 (55.2%)	457 (52.8%)	141 (51.3%)	144 (53.1%)	136 (49.8%)	421 (51.4%)	299 (53.2%)	283 (50.6%)	296 (52.6%)	878 (52.1%)
Mean ± SD	9.70 ± 3.62	9.56 ± 3.78	9.97 ± 3.88	9.75 ± 3.76	9.80 ± 3.97	9.70 ± 3.65	9.52 ± 3.64	9.67 ± 3.75	9.75 ± 3.79	9.63 ± 3.71	9.75 ± 3.77	9.71 ± 3.76
PGA score												
< 1	43 (15.0%)	38 (13.2%)	32 (11.0%)	113 (13.1%)	33 (12.0%)	39 (14.4%)	51 (18.7%)	123 (15.0%)	76 (13.5%)	77 (13.8%)	83 (14.7%)	236 (14.0%)
1 - < 2	195 (67.9%)	207 (71.9%)	212 (73.1%)	614 (71.0%)	196 (71.3%)	189 (69.7%)	175 (64.1%)	560 (68.4%)	391 (69.6%)	396 (70.8%)	387 (68.7%)	1174 (69.7%)
≥ 2	49 (17.1%)	43 (14.9%)	46 (15.9%)	138 (16.0%)	46 (16.7%)	43 (15.9%)	47 (17.2%)	136 (16.6%)	95 (16.9%)	86 (15.4%)	93 (16.5%)	274 (16.3%)
BILAG organ domain involvement												
at least 1A or 2B	166 (57.8%)	166 (57.6%)	172 (59.3%)	504 (58.3%)	187 (68.0%)	173 (63.8%)	160 (58.6%)	520 (63.5%)	353 (62.8%)	339 (60.6%)	332 (59.0%)	1024 (60.8%)
at least 1A	52 (18.1%)	58 (20.1%)	54 (18.6%)	164 (19.0%)	37 (13.5%)	38 (14.0%)	24 (8.8%)	99 (12.1%)	89 (15.8%)	96 (17.2%)	78 (13.9%)	263 (15.6%)
SLICC Damage Index score (Mean ± SD)	0.55 ± 0.93	0.60 ± 1.06	0.55 ± 1.00	0.57 ± 1.00	0.99 ± 1.45	1.04 ± 1.39	0.94 ± 1.38	0.99 ± 1.41	0.77 ± 1.23	0.81 ± 1.25	0.74 ± 1.21	0.77 ± 1.23
SLICC Damage Index score = 0	182 (63.4%)	190 (66.0%)	193 (66.6%)	565 (65.3%)	145 (52.7%)	125 (46.1%)	145 (53.1%)	415 (50.7%)	327 (58.2%)	315 (56.4%)	338 (60.0%)	980 (58.2%)
SLICC Damage Index score = 1	70 (24.4%)	56 (19.4%)	60 (20.7%)	186 (21.5%)	66 (24.0%)	76 (28.0%)	62 (22.7%)	204 (24.9%)	136 (24.2%)	132 (23.6%)	122 (21.7%)	390 (23.2%)
SLICC Damage Index score **≥ 2**	35 (12.2%)	42 (14.6%)	37 (12.8%)	114 (13.2%)	64 (23.3%)	69 (25.5%)	66 (24.2%)	199 (24.3%)	99 (17.6%)	111 (19.9%)	103 (18.3%)	313 (18.6%)

Proteinuria (g/24 hour)												
≥ 2	21 (7.3%)	26 (9.0%)	19 (6.6%)	66 (7.6%)	11 (4.0%)	7 (2.6%)	15 (5.5%)	33 (4.0%)	32 (5.7%)	33 (5.9%)	34 (6.0%)	99 (5.9%)
Mean ± SD	0.62 ± 1.15	0.63 ± 1.13	0.54 ± 0.91	0.60 ± 1.07	0.39 ± 0.81	0.33 ± 0.65	0.40 ± 0.73	0.37 ± 0.74	0.50 ± 1.00	0.48 ± 0.94	0.48 ± 0.83	0.49 ± 0.93

[1] Time elapsed between date of SLE diagnosis and the date of informed consent.

Baseline IgG, autoantibody, and complement for patients in the Phase 3 studies are summarised in Table 5.10. As expected, approximately 97% of patients were antinuclear antibody and/or anti-dsDNA positive ('autoantibody positive') at baseline (positivity during screening was an eligibility criterion). BLyS was detected in 98% of patients in Phase 3. Overall, patients in BLISS-52 had a higher degree of serological activity compared with patients in BLISS-76. The proportion of patients with anti-dsDNA antibodies was 75% in BLISS-52 vs 64% in BLISS-76. Low C3 and C4 were present in 45% and 56% of patients, respectively, with a somewhat greater proportion of patients in BLISS-52 presenting with low complement levels. Elevated IgG levels were present in 44% of patients (51% in BLISS-52 and 38% in BLISS-76). Within the Phase 3 studies, treatment groups were relatively balanced with regard to baseline autoantibodies, IgG and complement.

Table 5.10. Selected baseline serological characteristics in Phase 3 trials

	BLISS-52				BLISS-76				Pooled Total Population			
	Placebo N = 287	1 mg/kg N = 288	10 mg/kg N = 290	All N = 865	Placebo N = 275	1 mg/kg N = 271	10 mg/kg N = 273	All N = 819	Placebo N = 562	1 mg/kg N = 559	10 mg/kg N = 563	All N = 1684
Anti-dsDNA positive (≥ **30** IU/mL)	205 (71.4%)	221 (76.7%)	218 (75.2%)	644 (74.5%)	174 (63.3%)	171 (63.1%)	179 (65.6%)	524 (64.0%)	379 (67.4%)	392 (70.1%)	397 (70.5%)	1168 (69.4%)
Anti-Smith positive (≥ **15** U/mL)	101/287 (35.2%)	102/288 (35.4%)	105/287 (36.6%)	308/862 (35.7%)	72/269 (26.8%)	69/269 (25.7%)	75/265 (28.3%)	216/803 (26.9%)	173/556 (31.1%)	171/557 (30.7%)	180/552 (32.6%)	524/1665 (31.5%)
IgG >ULN (16.18 g/L)	146 (50.9%)	140 (48.6%)	151 (52.1%)	437 (50.5%)	108 (39.3%)	105 (38.7%)	94 (34.4%)	307 (37.5%)	254 (45.2%)	245 (43.8%)	245 (43.5%)	744 (44.2%)
Complement												
Normal/high C3 and C4	102 (35.5%)	100 (34.7%)	89 (30.7%)	291 (33.6%)	113 (41.1%)	122 (45.0%)	112 (41.0%)	347 (42.4%)	215 (38.3%)	222 (39.7%)	201 (35.7%)	638 (37.9%)
Low C3 or C4, but not both	78 (27.2%)	55 (19.1%)	75 (25.9%)	208 (24.0%)	65 (23.6%)	57 (21.0%)	60 (22.0%)	182 (22.2%)	143 (25.4%)	112 (20.0%)	135 (24.0%)	390 (23.2%)
Low C3 (< 900 mg/L)	132 (46.0%)	148 (51.4%)	147 (50.7%)	427 (49.4%)	116 (42.2%)	100 (36.9%)	115 (42.1%)	331 (40.4%)	248 (44.1%)	248 (44.4%)	262 (46.5%)	758 (45.0%)
Low C4 (< 16 mg/dL)	160 (55.7%)	173 (60.1%)	180 (62.1%)	513 (59.3%)	143 (52.0%)	141 (52.0%)	147 (53.8%)	431 (52.6%)	303 (53.9%)	314 (56.2%)	327 (58.1%)	944 (56.1%)
Low C3 and C4	107 (37.3%)	133 (46.2%)	126 (43.4%)	366 (42.3%)	97 (35.3%)	92 (33.9%)	101 (37.0%)	290 (35.4%)	204 (36.3%)	225 (40.3%)	227 (40.3%)	656 (39.0%)
BLyS (above LOQ, ≥ **0.5** ng/mL)	273/283 (96.5%)	273/285 (95.8%)	281/285 (98.6%)	827/853 (97.0%)	268/271 (98.9%)	267/270 (98.9%)	263/268 (98.1%)	798/809 (98.6%)	541/554 (97.7%)	540/555 (97.3%)	544/553 (98.4%)	1625/1662 (97.8%)

In the Phase 3 trials, 86% of patients were receiving corticosteroids at baseline, 58% at doses of > 7.5 mg/day prednisone equivalent (see Table 5.11). A difference in the proportion of patients using steroids at baseline between the Phase 3 trials was observed: 96% of patients in BLISS-52 and 76% of patients in BLISS-76. This difference is also reflected in the number of patients using higher doses of steroids (prednisone equivalent > 7.5 mg/day): 69% in BLISS-52 and 46% in Study BLISS-76. In contrast, more patients in BLISS-76 were receiving immunosuppressant agents at baseline (56%) compared with patients in BLISS-52 (42%). Azathioprine use was similar across the 2 trials (20-26%), while more patients in BLISS-76 were receiving methotrexate and mycophenolate (19% and 17%, respectively), than in BLISS-52 (9% and 6%, respectively). Almost all patients receiving immunosupressants at baseline were receiving only 1 immunosuppressant. In addition, more patients in BLISS-76 were using NSAIDs (41%) compared with patients in BLISS-52 (20%). Antimalarial use was relatively similar across trials (63-67%).

The majority of patients in each of the Phase 3 studies were receiving steroids and an antimalarial, with or without an immunosuppressant at baseline. In BLISS-52, there was greater use of steroids alone and in combination with antimalarials only, compared with BLISS-76. Patients in BLISS-76 more frequently were receiving immunosuppressants and antimalarials without steroids. The treatment groups within studies were reasonably well-balanced with regards to baseline therapies used.

Table 5.11. Selected baseline concomitant medications in Phase 3 trials

	BLISS-52				BLISS-76				Pooled Total Population			
	Placebo N = 287	1 mg/kg N = 288	10 mg/kg N = 290	All N = 865	Placebo N = 275	1 mg/kg N = 271	10 mg/kg N = 273	All N = 819	Placebo N = 562	1 mg/kg N = 559	10 mg/kg N = 563	All N = 1684
Total corticosteroid use	276 (96.2%)	276 (95.8%)	278 (95.9%)	830 (96.0%)	212 (77.1%)	211 (77.9%)	200 (73.3%)	623 (76.1%)	488 (86.8%)	487 (87.1%)	478 (84.9%)	1453 (86.3%)
Prednisone or equivalent												
> 0 to **≤ 7.5 mg/day**	84 (29.3%)	72 (25.0%)	74 (25.5%)	230 (26.6%)	86 (31.3%)	81 (29.9%)	80 (29.3%)	247 (30.2%)	170 (30.2%)	153 (27.4%)	154 (27.4%)	477 (28.3%)
> 7.5 to < 20 mg/day	136 (47.4%)	133 (46.2%)	131 (45.2%)	400 (46.2%)	76 (27.6%)	96 (35.4%)	81 (29.7%)	253 (30.9%)	212 (37.7%)	229 (41.0%)	212 (37.7%)	653 (38.8%)
≥ 20 mg/day	56 (19.5%)	71 (24.7%)	73 (25.2%)	200 (23.1%)	50 (18.2%)	34 (12.5%)	39 (14.3%)	123 (15.0%)	106 (18.9%)	105 (18.8%)	112 (19.9%)	323 (19.2%)
Antimalarials	201 (70.0%)	195 (67.7%)	185 (63.8%)	581 (67.2%)	180 (65.5%)	171 (63.1%)	168 (61.5%)	519 (63.4%)	381 (67.8%)	366 (65.5%)	353 (62.7%)	1100 (65.3%)
Other immunosuppressants	122 (42.5%)	120 (41.7%)	123 (42.4%)	365 (42.2%)	154 (56.0%)	153 (56.5%)	148 (54.2%)	455 (55.6%)	276 (49.1%)	273 (48.8%)	271 (48.1%)	820 (48.7%)
1 immunosupressant	111 (38.7%)	116 (40.3%)	118 (40.7%)	345 (39.8%)	140 (50.9%)	143 (52.8%)	140 (51.3%)	423 (51.6%)	251 (44.7%)	259 (46.3%)	258 (45.8%)	768 (45.6%)
2 immunosupressants	11 (3.8%)	4 (1.4%)	5 (1.7%)	20 (2.3%)	13 (4.7%)	10 (3.7%)	8 (2.9%)	31 (3.8%)	24 (4.3%)	14 (2.5%)	13 (2.3%)	51 (3.0%)
Azathioprine	67 (23.3%)	71 (24.7%)	84 (29.0%)	222 (25.7%)	57 (20.7%)	52 (19.2%)	58 (21.2%)	167 (20.4%)	124 (22.1%)	123 (22.0%)	142 (25.2%)	389 (23.1%)
Methotrexate	35 (12.2%)	24 (8.3%)	20 (6.9%)	79 (9.1%)	60 (21.8%)	53 (19.6%)	39 (14.3%)	152 (18.6%)	95 (16.9%)	77 (13.8%)	59 (10.5%)	231 (13.7%)
Mycophenolate	19 (6.6%)	16 (5.6%)	17 (5.9%)	52 (6.0%)	42 (15.3%)	45 (16.6%)	50 (18.3%)	137 (16.7%)	61 (10.9%)	61 (10.9%)	67 (11.9%)	189 (11.2%)
Cyclosporin	6 (2.1%)	5 (1.7%)	2 (0.7%)	13 (1.5%)	5 (1.8%)	4 (1.5%)	5 (1.8%)	14 (1.7%)	11 (2.0%)	9 (1.6%)	7 (1.2%)	27 (1.6%)
Leflunomide	2 (0.7%)	-	3 (1.0%)	5 (0.6%)	3 (1.1%)	7 (2.6%)	1 (0.4%)	11 (1.3%)	5 (0.9%)	7 (1.3%)	4 (0.7%)	16 (1.0%)
Cyclophosphamide	2 (0.7%)	3 (1.0%)	1 (0.3%)	6 (0.7%)	2 (0.7%)	2 (0.7%)	2 (0.7%)	6 (0.7%)	4 (0.7%)	5 (0.9%)	3 (0.5%)	12 (0.7%)
NSAIDs	59 (20.6%)	56 (19.4%)	58 (20.9%)	173 (20.0%)	119 (43.3%)	114 (42.1%)	101 (37.0%)	334 (40.8%)	178 (31.7%)	170 (30.4%)	159 (28.2%)	507 (30.1%)

Study populations were similar between studies and across treatment groups, except for some differences in baseline medication use, serological activity, as well as racial differences. Compared with patients in BLISS-52, patients in BLISS-76 appeared to have a lower baseline level of disease activity as evidenced by proteinuria, serological markers, and use of baseline corticosteroids, but had a higher level of pre-existing organ damage. A difference between the 2 Phase 3 trials was observed in the proportion of patients using steroids at baseline: 96% of patients in BLISS-52 and 76% of patients in BLISS-76.

Outcomes

5.3.5 Provide details of the outcomes investigated and the measures used to assess those outcomes. Indicate which outcomes were specified in the trial protocol as primary or secondary, and whether they are relevant with reference to the decision problem. This should include therapeutic outcomes, as well as patient-related outcomes such as assessment of health-related quality of life, and any arrangements to measure compliance. Data provided should be from pre-specified outcomes rather than post-hoc analyses. When appropriate, also provide evidence of reliability or validity, and current status of the measure (such as use within UK clinical practice). The following table provides a suggested format for presenting primary and secondary outcomes when there is more than one RCT.

Table 5.12. Primary and secondary outcomes of the RCTs

Trial no. (acronym)	Primary outcome(s) and measures	Reliability/validity/ current use in clinical practice	Secondary outcome(s) and measures	Reliability/validity/ current use in clinical practice
C1057 (BLISS-52)	The primary efficacy endpoint was the response rate at week 52, assessed by the SLE Responder Index (SRI). With the SRI criteria, a responder was defined as having a reduction of at least 4 points in the SELENA-SLEDAI score (defined as clinically meaningful) (Gladman et al. 2000), no new BILAG A organ domain score, no more than 1 new BILAG B organ domain score, and no worsening in Physician's Global Assessment (PGA) score (increase <0.3) at week 52 compared with baseline.	**SLE Responder Index** The primary efficacy endpoint is evidence-based and supported by the data from the Phase 2 SLE trial (LBSL02) (Wallace et al. 2009). It includes an objective measure of the reduction in global disease activity (reduction in SELENA-SLEDAI score ≥ 4) for efficacy and 2 measures to ensure that the improvement in disease activity (score) is not offset by worsening of the subject's condition overall (i.e. no worsening in the PGA) or worsening in any specific organ system (i.e. no new BILAG A or 2 new B flares). This primary efficacy endpoint was agreed with regulatory authorities prior to initiation of the Phase 3 trials (and was included in the SPA agreement with the FDA) and is consistent with the recommendations in the Draft Guidance for Industry: Systemic Lupus Erythematosus-Developing	Major secondary outcomes: • Percent of subjects with ≥ 4-point reduction in SELENA-SLEDAI at Week 52. • Mean change in PGA at Week 24. • Percent of subjects with prednisone (equivalent) reduction ≥ 25% from baseline to ≤ 7.5 mg/day during Weeks 40 – 52 (in subjects whose prednisone equivalent dose was > 7.5 mg/day at baseline). • Mean change in SF-36 PCS at Week 24. Other secondary outcomes relevant to the final scope: • Mean change in PGA at Week 52.	**SELENA-SLEDAI** See Primary outcomes. **PGA** See Primary outcomes. **Steroid reduction** Percent of subjects with prednisone (equivalent) reduction ≥ 25% from baseline to ≤ 7.5 mg/day during Weeks 40 – 52 (in subjects whose prednisone equivalent dose was > 7.5 mg/day at baseline) was included to assess the steroid sparing effects of belimumab. This reduction in steroids was felt to represent clinically meaningful steroid sparing in SLE subjects (Fortin et al. 2008). In clinical practice, clinicians will try to keep the steroid dose as low as possible in order to avoid side effects.

		Drugs for Treatment (March 2005). In addition, the primary efficacy endpoint and the Phase 3 clinical trial design are consistent with the recommendations of the Task Force on SLE of the EULAR Standing Committee for International Clinical Studies Including Therapeutics and the recently issued EMA Committee for Medicinal Products for Human Use (CHMP) concept paper on the need for a guideline on the clinical investigation of medicinal products intended for the treatment of SLE (Bertsias et al. 2008; Bertsias et al. 2009; European Medicines Agency 2009). The SRI is not used in clinical practice, however, the individual components may be used. **SELENA-SLEDAI** Safety of Estrogens in Lupus Erythematosus National Assessment trial – Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI) scores capture the subjects'	• Mean change in SF-36 PCS at Week 52. • Modified SELENA-SLEDAI SLE Flare Index (SFI) over the course of study. • SLICC/ACR damage index at Week 52. • FACIT-Fatigue scale over the course of the study. • EQ-5D over the course of the study.	**SF-36** SF-36: The Medical Outcomes 36-Item Short Form Health Survey (SF-36) was employed as a generic HRQL instrument, since it has been shown to reflect the impact of SLE on all domains of HRQL in observational cohort studies, as well as randomised trials (Alonso et al. 2004; Gladman et al. 1996b; Smolen et al. 1999; Strand et al. 2005; Thumboo et al. 1999; Thumboo et al. 2000; Ware, Jr. et al. 1992). SF-36 has also been widely tested and validated in a variety of chronic diseases including RA, osteoarthritis, etc. It is not used routinely in clinical practice. **SFI** SLE flare serves as an important indicator of SLE disease control. Flares increase with increases in disease activity (see Appendix 17). A reduction in flares is the goal of any SLE treatment. The assessment of flare and

		condition over the 10 days prior to the visit (see Appendix 14). Classification of increased disease activity using the SELENA-SLEDAI score has been described as an increase of 3 points or more (Petri et al. 1991; Petri et al. 1999), and a reduction of more than 3 points in SELENA-SLEDAI score has been defined as an improvement (Gladman et al. 2000). Consequently, a reduction ≥ 4 is considered evidence of improvement. Moreover, on the SELENA-SLEDAI, a reduction of 4 points equates to elimination (rather than mere improvement) of a disease manifestation and, as such, is a clear demonstration of clinical benefit. The SELENA-SLEDAI has been shown to be valid, reliable and sensitive to change. Disease activity can range from 0 to 105, with a score of 0 indicating no activity and scores of 20 and above indicating very high activity (Griffiths et al. 2005).		severe flare was conducted using the modified SELENA-SLEDAI SLE Flare Index (SFI). The SLE Flare Index categorises SLE flare as “mild or moderate” or “severe” based on 6 variables (Buyon et al. 2005; Petri et al. 1999; Petri et al. 2005): • Change in SELENA-SLEDAI score from the most recent assessment to current. • Change in signs or symptoms of disease activity. • Change in prednisone dosage. • Use of new medications for disease activity or hospitalisation. • Change in PGA score. • Hospitalisation for SLE activity (severe flare only). The SLE flare index was modified such that severe flares triggered only by an increase in SELENA-SLEDAI score to > 12 alone were excluded. One or more other items defining severe flare in the SFI needed to be present for a severe flare to be recorded. This modification was made because patients entering the trial with high disease activity (e.g. ≥ 11) could too easily trigger a severe

		SELENA-SLEDAI is unable to capture worsening of an already existing feature or detect partial improvements. Disease activity as measured by SLEDAI correlates significantly with organ damage by SLICC/ACR Damage Index in SLE patients (Swaak et al. 1999). The primary measure of efficacy in the endpoint is the reduction in disease activity measured by the SELENA-SLEDAI score. To be a responder, a subject must have a 4-point or greater reduction in her/his SELENA-SLEDAI score compared with her/his baseline value. **BILAG** The British Isles Lupus Assessment Group (BILAG) measures changes in disease activity over the past 28 days and was specifically developed to identify the need to alter a subject's treatment based on clinical signs and symptoms (see Appendix 15). It is the only activity index based on the evaluation of disease activity in individual organ		flare by minor increases in SELENA-SLEDAI score. **SLICC/ACR** SLICC/ACR damage index: A group of investigators interested in the clinical outcome of patients with SLE developed a clinical index of chronic damage: the Systemic Lupus International Collaborating Clinics Damage Index, which has been endorsed by the American College of Rheumatology, so that the full title of the index is the SLICC/ACR damage index (see Appendix 18). The index has 41 items covering 12 systems. It includes specific comorbidities associated with SLE. Manifestations should be recorded as damage only if they develop after the onset of lupus, provided they fulfil the definition in the glossary, and irrespective of attribution. To be scored as damage, items should persist continuously for 6 months or be associated with an immediate pathological scar

		systems. BILAG scores range from A (very active disease) to D (no current disease activity). A score of E means that the organ system has never been involved. An A or 2 B flare represents either an increase in disease activity thought to be sufficient to require alteration of therapy with steroids or immunosuppressants (A) or mild reversible problems in 2 organ systems (2 B) (Hay et al. 1993; Isenberg et al. 2000). BILAG is a reliable and valid instrument for measuring clinical disease activity in SLE (Griffiths et al. 2005). The choice of the BILAG score to evaluate worsening provided a sensitive measure of increase in disease activity. The definition of no worsening by the BILAG score was “no new BILAG A organ domain score or 2 BILAG B organ domain scores compared with baseline at the time of the assessment (i.e., Week 52)”.		indicative of damage (for example, a myocardial infarction). Some items can score two for recurrent events, such as repeated strokes and avascular necrosis at two sites. The maximum score is 47 but patients rarely score above 12 points. The SLICC/ACR damage index has been shown to be valid and reliable. It is distinct from disease activity but more common in patients with persistent or recurrent disease activity. Early accumulation of damage correlates with a poor prognosis, including an increased mortality. Damage is an important outcome measure in observational studies and clinical trials (Griffiths et al. 2005). **FACIT-Fatigue scale** FACIT-Fatigue scale: Fatigue was measured by the Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue scale, Version 4.0. The FACIT-Fatigue scale is a 13-item questionnaire

		PGA The Physician's Global Assessment (PGA) is a sensitive, semi quantitative test of the patient's condition. It uses a graduated 10 cm visual analogue scale from 0 to 3 on which the physician marks his/her assessment (see Appendix 16). A score of 1 corresponds to mild lupus disease activity, a score of 2 to 2.5 with moderate disease activity, and a score of 3 with severe disease activity. Consequently, a change of 1 unit on the PGA is associated with worsening of disease activity. An increase of ≥ 1 unit from the last assessment resulting in a PGA score ≤ 2.5 is considered a mild-moderate flare. If the increase in PGA is to > 2.5, it is considered a severe flare (Bombardier et al. 1992; Petri et al. 1991; Petri et al. 1999). The minimum clinically significant improvement was 6.2% or 0.62 cm on a patient global assessment 10 cm VAS scale in		designed to assess clinically-relevant problems associated with a chronic medical condition (http://www.facit.org) (see Appendix 19). The range of possible scores is 0 to 52 (0 is the worst possible score and 52 is the best). The FACIT-Fatigue has been validated for use with rheumatoid arthritis patients and has demonstrated good internal consistency, clinical sensitivity, and a high correlation with the SF-36 vitality domain (r = 0.73-0.84) (Cella et al. 2005). **EQ-5D** The EQ-5D is a standardised instrument used to measure health outcomes because it is applicable to a wide range of health conditions and treatments and it is designed to provide a single index value for health status (http://www.euroqol.org). The EQ-5D consists of 5 dimensions (mobility, self-care, usual activities, pain/discomfort,

		RA, and the minimum significant worsening was 16.3% or 1.63 cm (Wells et al. 1993). Therefore, an increase of ≥ 0.3 points (> 10% on the 3-point 10 cm VAS) from baseline was considered to be a conservative estimate for the minimum clinically significant worsening. The PGA is included in the combined response endpoint to ensure that improvement in disease activity measured by SELENA-SLEDAI is not achieved at the expense of worsening of the patient's overall condition.		and anxiety/depression) each of which includes 1 of 3 responses to describe a subject's health state: 1) no problems, 2) some or moderate problems, and 3) severe problems. The EQ-5D questionnaire also includes the EQ VAS, a standard vertical 20 cm visual analogue scale used to record a subject's rating for current health-related quality of life state based on a best imaginable health state at the top and worst imaginable health state at the bottom, having numeric values of 100 and 0, respectively.
C1056 (BLISS-76)	As per C1057 (BLISS-52).	As per C1057 (BLISS-52).	As per BLISS-52, with the following addition: • Response rate (SRI) at Week 76.	As per BLISS-52.

Statistical analysis and definition of study groups

5.3.6 State the primary hypothesis or hypotheses under consideration and the statistical analysis used for testing hypotheses. Also provide details of the power of the study and a description of sample size calculation, including rationale and assumptions. Provide details of how the analysis took account of patients who withdrew (for example, a description of the intention-to-treat analysis undertaken, including censoring methods; whether a per-protocol analysis was undertaken). The following table provides a suggested format for presenting the statistical analyses in the trials when there is more than one RCT.

Table 5.13. Summary of statistical analyses in RCTs

Trial no. (acronym)	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
C1057 (BLISS-52)	To assess the efficacy, safety, and tolerability of belimumab with standard of care in patients with seropositive systemic lupus erythematosus.	The response rate at week 52 (primary endpoint) was assessed with SRI in each belimumab group and was compared with the placebo group by use of a logistic regression model adjusted for baseline randomisation stratification factors. Analysis was done in a modified intention-to-treat population, defined as all randomly assigned patients who received a dose of the study drug. In the primary efficacy analyses, a step-down procedure was used to control the type 1 error (two-sided α=0·05) for comparison of	The sample size of 810 patients (270 per group) was calculated to provide 90% power at a significance level of 5% to detect a 14% absolute improvement in the SRI response rate at week 52 with belimumab 10 mg/kg relative to placebo. A standard deviation of 50% was used to account for the worst-case variability.	Patients who withdrew or required changes in background drugs for systemic lupus erythematosus that were other than those permitted by protocol were judged to be treatment failures.

		belimumab 10 mg/kg with placebo; if 10 mg/kg was better than placebo, belimumab 1 mg/kg was then compared with placebo. Binary efficacy variables were assessed with a logistic regression model, continuous variables were analysed with an analysis of covariance model, and time-to-flare variables were analysed by use of a Cox proportional hazards model. All analyses were adjusted for baseline randomisation factors.		
C1056 (BLISS-76)	As per BLISS-52.	As per BLISS-52.	As per BLISS-52.	As per BLISS-52.

5.3.7 Provide details of any subgroup analyses that were undertaken and specify the rationale and whether they were pre-planned or post-hoc.

A series of pre-specified and post-hoc subgroup analyses for efficacy data were conducted. A comparison between each belimumab treated group and the placebo group was performed by the following major subgroups which were prespecified in each Phase 3 analytical plan:

• Baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10); stratification factor.
• Race (African descent or indigenous-American descent vs other); stratification factor.
• Baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent); stratification factor.
• Baseline anti-dsDNA (≥ 30 IU/mL vs < 30 IU/mL).
• Baseline prednisone dose level (≤ 7.5 mg/day vs > 7.5 mg/day).
• Baseline C3 levels (normal/high vs low).
• Baseline C4 levels (normal/high vs low).
• Region (modified for the integrated analysis from protocol analytical plans as follows: US/Canada vs other; Western Europe (includes Austria, Belgium, Germany, Spain, France, UK, Italy, Netherlands, Sweden)/Australia/Israel vs other; Eastern Europe (includes Romania, Russia, Czech Republic, Poland, Slovakia) vs other; Americas (excluding US/Canada) (includes Argentina, Brazil, Chile, Colombia, Peru, Costa Rica, Mexico, Puerto Rico) vs other; Asia (includes Hong Kong, India, Korea, Philippines, Taiwan) vs other).

Additional exploratory subgroup analyses that were pre-specified in the individual analytical plans and that were evaluated in the pooled Phase 3 population are listed below:

• Age (≤ 45 years vs > 45; modified from 3 age groups since few subjects above 65).
• Gender (female vs male).
• Baseline medications (steroids > 7.5 mg/day vs not; other

immunosuppressant/immunomodulatory agents vs not).
• Baseline BILAG (at least 1A/2B score vs other).

Other subgroups as listed below were explored post-hoc:

• Baseline SELENA-SLEDAI score (≤ 12 vs ≥ 13).
• Race (white vs other; black vs other; Asian vs other; Alaska native/American Indian vs other).
• Baseline ANA (< 1:640 vs ≥ 1:640).
• Baseline anti-Sm (< 15 U/mL vs ≥ 15 U/mL).
• Baseline BLyS (≥ LOQ vs < LOQ).
• Baseline steroid use (yes vs no).

In order to identify baseline factors that are predictive of response at Week 52 irrespective of treatment received and to evaluate belimumab treatment effect adjusted for the predictive factors, a logistic regression main effects model was developed based on the pooled data from the Phase 3 studies. The model building process began with effects for treatment and study in the model, then a stepwise forward selection process was used to evaluate the group of eligible baseline characteristics for entry into the model using $\alpha = 0.05$ significance level.

The following baseline characteristics were identified as significant predictors of Week 52 response, irrespective of whether a subject was treated with belimumab or placebo (listed in order of significance):

- baseline SELENA SLEDAI (≤ 9, ≥ 10)
- complement (normal/high C3 and C4, low C3 or C4, low C3 and C4)
- immunosuppressant use (yes, no)
- region (US/Canada, Western Europe, Eastern Europe, Americas excluding US/Canada, and Asia
- SLICC/ACR Damage Index score (0, 1, ≥ 2)
- anti-dsDNA (< 30, ≥ 30 IU/mL).

These findings from multivariate analyses and bearing in mind real-world clinical considerations, support the observation that subjects with higher disease activity are more likely to respond to belimumab. Based on these findings, a number of clinically relevant characteristics were examined to identify high disease activity patients who benefit the most from belimumab treatment (see Figure 5.3).

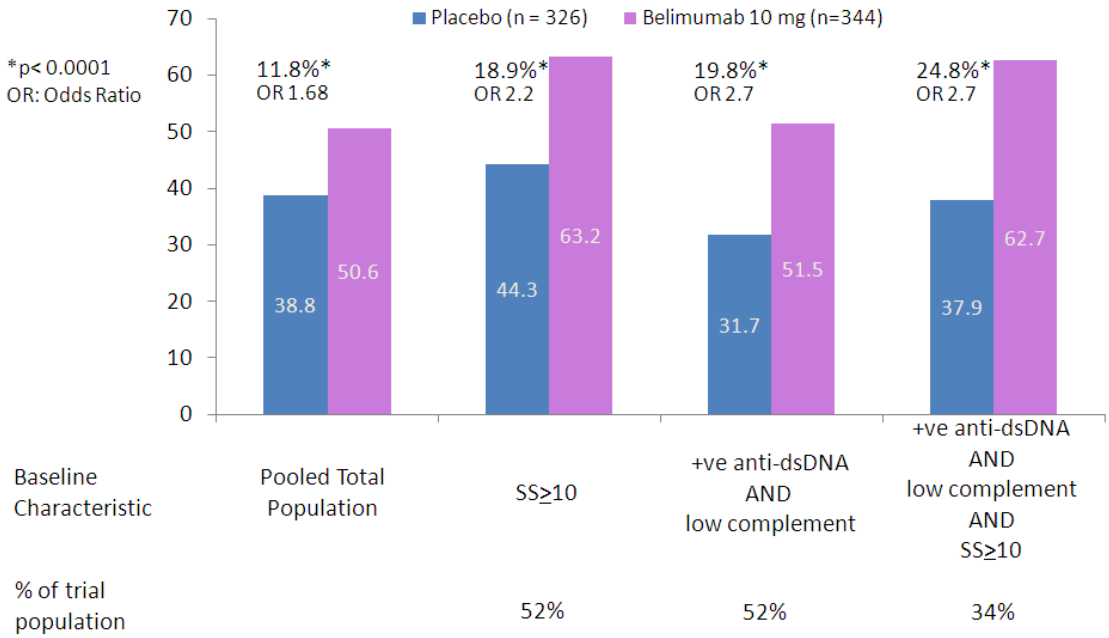
SELENA-SLEDAI was chosen. Whilst not used routinely in general practice, it is a direct measure of disease activity and was the most significant predictor of Week 52 response; which is not surprising as a 4 point reduction in SELENA-SLEDAI is the efficacy component of the SRI.

Secondly, anti-dsDNA and complement were chosen. These two are objective measures used routinely in SLE and accessible to physicians in general practice. They are widely considered important measures of disease activity. Patients with positive anti-dsDNA and low complement are immunologically active and at higher risk for flares and lupus nephritis (Petri et al. 2009; Tseng et al. 2006).

Immunosuppressants were not chosen. Whilst they are routinely used in SLE management, their use is relatively subjective depending on individual physician's experience and access to therapies. SLICC/ACR was not chosen as it is primarily designed to measure damage in clinical trials/observational studies and is distinct from disease activity.

Compared to the SRI response rate for the primary efficacy population (11.8%), the subgroup with SELENA-SLEDAI ≥10 and the subgroup with low complement and positive anti-dsDNA both showed a greater treatment effect (18.9% and 19.8% respectively) (see Figure X). The greater efficacy achieved with belimumab in a highly immunologically active subgroup of SLE patients (with low complement and positive anti-dsDNA) is consistent with the mechanism of action of belimumab. An even greater treatment effect of 24.8% was seen in patients who had SELENA-SLEDAI ≥10, low complement and positive anti-dsDNA at baseline (see Figure 5.3).

Figure 5.3. SRI response rate at Week 52 – Subgroups



With reference to the decision problem and our intention to explicitly identify patients who benefit the most, we will present efficacy data for the high disease activity subgroup of patients with evidence for serological disease activity (low complement and positive anti-dsDNA) and additionally have a SELENA-SLEDAI disease activity score ≥10 at baseline. This subgroup combines routinely used subjective laboratory measures with a clinical measure of disease activity; allowing clinicians to identify patients with significant disease activity.

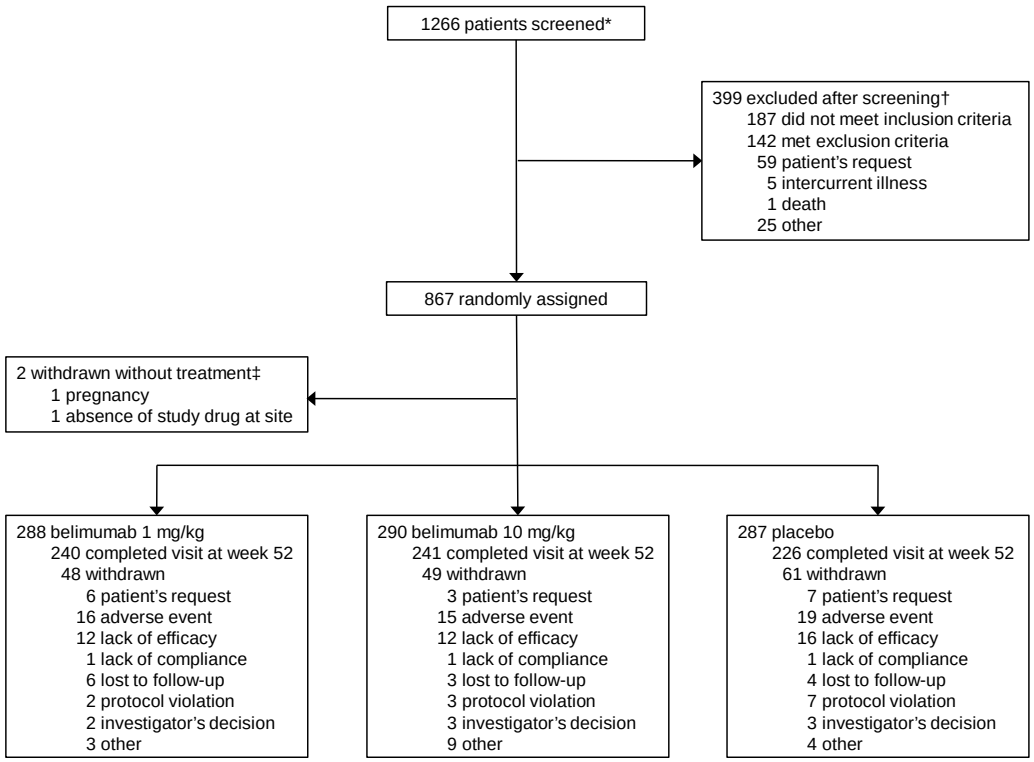
Patients in this subgroup experienced the greatest treatment effect over and above the primary efficacy population. Whilst this subgroup was not pre-specified, each of the individual components (low complement, positive anti-dsDNA and SELENA-SLEDAI ≥10) were pre-specified. In addition, all of these measures are widely considered important measures of disease activity. As disease activity correlates significantly with organ damage (Swaak et al. 1999), this subgroup of patients, whilst appearing to benefit most from belimumab are also the most likely to experience the worst long-term morbidity from SLE.

Efficacy data for this subgroup will be presented alongside data from the individual BLISS studies and pooled Phase 3 population (BLISS-52 and BLISS-76). Pooling is appropriate given that the trials were essentially identical in design and in the analysis of the primary endpoint, the p-values for the treatment-by-study interaction were not significant (interaction p-values > 0.5). See Section 5.5 for further details.

Participant flow

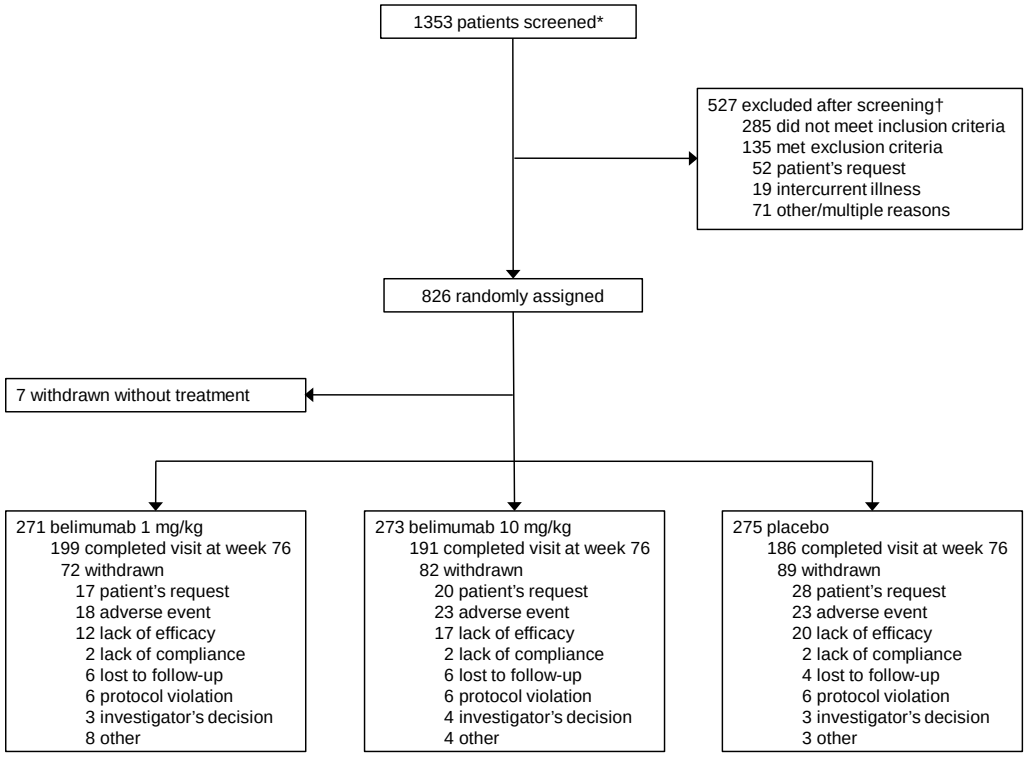
5.3.8 Provide details of the numbers of patients who were eligible to enter the RCT(s), randomised, and allocated to each treatment. Provide details of, and the rationale for, patients who crossed over treatment groups and/or were lost to follow-up or withdrew from the RCT. This information should be presented as a CONSORT flow chart.

Figure 5.4. Flow chart for BLISS-52



*Patients who were rescreened were counted more than once. †Patients could have more than one reason for being excluded. ‡One patient withdrawn from the belimumab 1mg/kg group (because of lack of study drug at site) and one from the placebo group (because of pregnancy).

Figure 5.5. Flow chart for BLISS-76



*Patients who were rescreened were counted more than once. †Multiple reasons for some patients.

5.4 Critical appraisal of relevant RCTs

5.4.1 The validity of the results of an individual study will depend on the robustness of its overall design and execution, and its relevance to the decision problem. Each study that meets the criteria for inclusion should therefore be critically appraised. Whenever possible, the criteria for assessing published studies should be used to assess the validity of unpublished and part-published studies. The critical appraisal will be validated by the ERG. The following are the minimum criteria for assessment of risk of bias in RCTs, but the list is not exhaustive.

- Was the method used to generate random allocations adequate?
- Was the allocation adequately concealed?
- Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?
- Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?
- Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?
- Is there any evidence to suggest that the authors measured more outcomes than they reported?
- Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?

5.4.2 Please provide as an appendix a complete quality assessment for each RCT. See section 9.3, appendix 3 for a suggested format.

5.4.3 If there is more than one RCT, tabulate a summary of the responses applied to each of the critical appraisal criteria. A

suggested format for the quality assessment results is shown below.

Table 5.14. Quality assessment results for RCTs

Trial no. (acronym)	C1057 (BLISS-52)	C1056 (BLISS-76)
Was randomisation carried out appropriately?	Yes	Yes
Was the concealment of treatment allocation adequate?	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes
Were there any unexpected imbalances in drop-outs between groups?	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes
Adapted from Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

5.5 Results of the relevant RCTs

5.5.1 Provide the results for all relevant outcome measure(s) pertinent to the decision problem. Data from intention-to-treat analyses should be presented whenever possible and a definition of the included patients provided. If patients have been excluded from the analysis, the rationale for this should be given. **If there is more than one RCT, tabulate the responses.**

5.5.2 The information may be presented graphically to supplement text and tabulated data. If appropriate, please present graphs such as Kaplan-Meier plots.

5.5.3 For each outcome for each included RCT, the following information should be provided.

- The unit of measurement.
- The size of the effect; for dichotomous outcomes, the results ideally should be expressed as both relative risks (or odds ratios) and risk (or rate) differences. For time-to-event analysis, the hazard ratio is an equivalent statistic. Both absolute and relative data should be presented.
- A 95% confidence interval.
- Number of participants in each group included in each analysis and whether the analysis was by ‘intention to treat’. State the results in absolute numbers when feasible.

- When interim RCT data are quoted, this should be clearly stated, along with the point at which data were taken and the time remaining until completion of that RCT. Analytical adjustments should be described to cater for the interim nature of the data.
- Other relevant data that may assist in interpretation of the results may be included, such as adherence to medication and/or study protocol.
- Discuss and justify definitions of any clinically important differences.
- Report any other analyses performed, including subgroup analysis and adjusted analyses, indicating those pre-specified and those exploratory.

Whilst a 1 mg/kg dose was examined in the Phase 3 studies, we will only present results for the 10 mg/kg belimumab dose as this is the dose submitted for Marketing Authorisation. When discussing the results from the Phase 3 studies, the belimumab group refers to belimumab plus standard of care and the placebo group refers to placebo plus standard of care.

Both Phase 3 trials were positive. The primary endpoint, SRI response at Week 52 was met in both Phase 3 trials for the 10 mg/kg dose of belimumab. These studies demonstrated substantial evidence of effectiveness of belimumab as measured by reductions in disease activity by the SRI. Reductions in risk of severe flare and steroid use were also observed, as were improvements in patient reported quality of life. Serological activity was reduced as measured by reductions in autoantibodies and normalisation of hypergammaglobulinemia and complement levels. B cells, including autoreactive B cells, were also reduced, but not severely depleted, consistent with BLyS inhibition.

The pooled total population comprises pooled data from the 2 randomised, placebo-controlled trials BLISS-52 and BLISS-76. The Phase 3 studies are identical in their design, with the only differences being that in addition BLISS-76 evaluates B cell subsets, has a vaccine substudy, has a different frequency of collection of autoantibodies and patient reported outcome (PRO) assessments, and is 76 weeks in duration (although both trials have the primary efficacy endpoint at Week 52). Therefore, it is appropriate that the data from these studies be pooled for the evaluation of efficacy.

Furthermore, pooled data provides increased statistical power for treatment comparisons within subgroups, where there is reduced sample size. Finally, the use of pooled data provides increased power for endpoints where there are relatively few events, such as severe flares.

The analyses defined in this document represent an integrated analysis performed by pooling subject-level data from two studies, BLISS-52 and BLISS-76, and conducting the analyses on the pooled data. Analyses of data after Week 52 and up to Week 76 are not meta-analyses as they pertain only to study BLISS-76. Meta-analyses methods in which summary statistics are combined across the individual studies in order to integrate the findings will not be employed.

Where the size and composition of the subpopulation allowed, analyses were performed controlling for baseline stratification factors and study; otherwise, some of the covariates were omitted or unadjusted analyses were performed.

With reference to the decision problem, we will present efficacy data for the high disease activity subgroup of patients (post-hoc) with evidence for serological disease activity (low complement and positive anti-dsDNA) and additionally have a SELENA-SLEDAI disease activity score ≥10. Patients in this subgroup experienced the greatest treatment effect over and above the pooled total population. Whilst this subgroup was not pre-specified, each of the individual components (low complement, positive anti-dsDNA

and SELENA-SLEDAI ≥10) were pre-specified. In addition, all of these measures are widely considered important measures of disease activity. As disease activity correlates significantly with organ damage (Swaak et al. 1999), this subgroup of patients, whilst appearing to benefit most from belimumab are also the most likely to experience the worst long-term morbidity from SLE (see Section 5.3.7). Data for this subgroup will be presented alongside data from the individual BLISS studies and the pooled Phase 3 population (BLISS-52 and BLISS-76).

Primary Efficacy Endpoint and Analyses

The primary efficacy endpoint of both Phase 3 studies was the SLE Responder Index (SRI) at Week 52. The SRI is evidence-based and supported by the data from the Phase 2 SLE study (LBSL02) (Wallace et al. 2009). It includes an objective measure of the reduction in global disease activity (reduction in SELENA SLEDAI score ≥ 4) for efficacy and 2 measures to ensure that the improvement in disease activity (score) is not offset by worsening of the subject's condition overall (i.e. no worsening in the PGA) or worsening in any specific organ system (i.e. no new BILAG A or 2 new B flares). The SRI is a robust responder index that measures clinically meaningful change and mitigates against the possibility of disease worsening in some organ systems while other organs improve (Stone 2011). In addition, its use was agreed with regulatory authorities prior to initiation of the Phase 3 studies.

In the both Phase 3 trials and the pooled total population, belimumab 10 mg/kg was shown to be superior to placebo as assessed by the SRI at Week 52 (see Table 5.15). The greatest response rate was seen with belimumab 10 mg/kg in the high disease activity subgroup (see Table 5.15). Lack of response was driven by failure to achieve a 4 point reduction in SELENA-SLEDAI followed by dropouts and medication failures.

Table 5.15. Primary efficacy endpoint (SRI) at Week 52 (dropout = failure)

	BLISS-52		BLISS-76		Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 287	10 mg/kg N = 290	Placebo N = 275	10 mg/kg N = 273	Placebo N = 562	10 mg/kg N = 563	Placebo N = 203	10 mg/kg N = 193
No. (%) Response	125 (43.6%)	167 (57.6%)	93 (33.8%)	118 (43.2%)	218 (38.8%)	285 (50.6%)	77 (37.9%)	121 (62.7%)
Observed difference vs placebo (%)	-	14.03	-	9.41	-	11.8	-	24.8
OR (95% CI)[1] vs placebo	-	1.83 (1.30, 2.59)	-	1.52 (1.07, 2.15)	-	1.68 (1.3, 2.2)	-	2.7 (1.8, 4.1)
P-value[1]	-	0.0006	-	0.0207	-	< 0.0001	-	< 0.0001
Treatment by study interation p-value[2]	N/A	N/A	N/A	N/A	-	0.5579	N/A	N/A

[1] Odds Ratio (95% confidence interval) and p-values were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, covariates include baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.
[2] Obtained from a regression by adding the treatment-by-study interaction to the above model.

SRI was also evaluated at each visit. For BLISS-52 and the pooled total population, belimumab 10 mg/kg was associated with an improved response over placebo that begins to become apparent at Week 8 and is statistically significant at each visit from Weeks 16-52 (see Figure 5.6 and Figure 5.8). For the high disease activity subgroup, a statistically significant improvement in response of belimumab 10 mg/kg over placebo is seen as early as Week 8 and maintained through to Week 52 (see Figure 5.9). For BLISS-76, a statistically significant improvement in response of belimumab 10 mg/kg over placebo was only seen at Week 52 (see Figure 5.7).

Figure 5.6. Response by Visit through Week 52 – BLISS-52

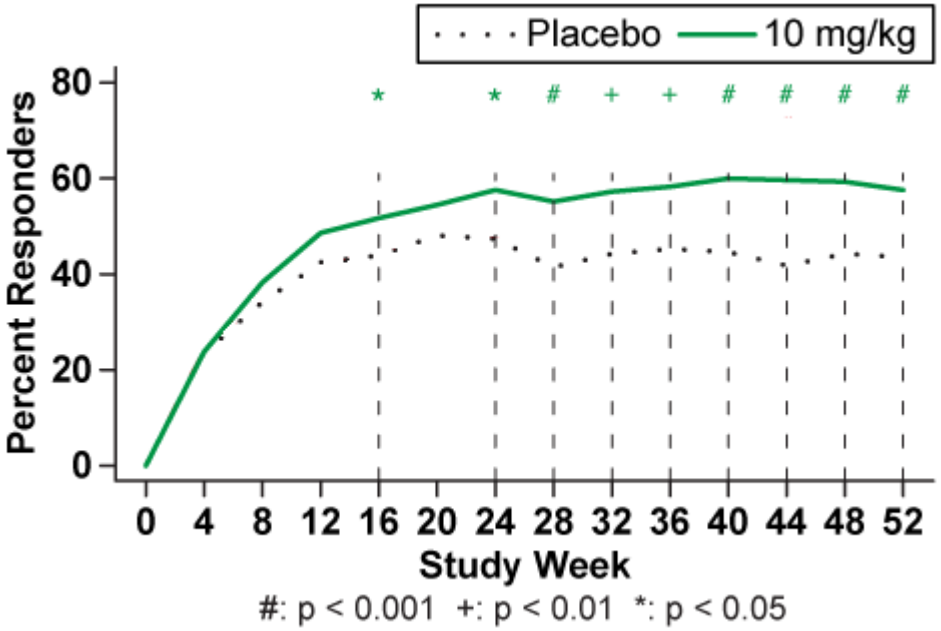


Figure 5.7. Response by Visit through Week 52 – BLISS-76

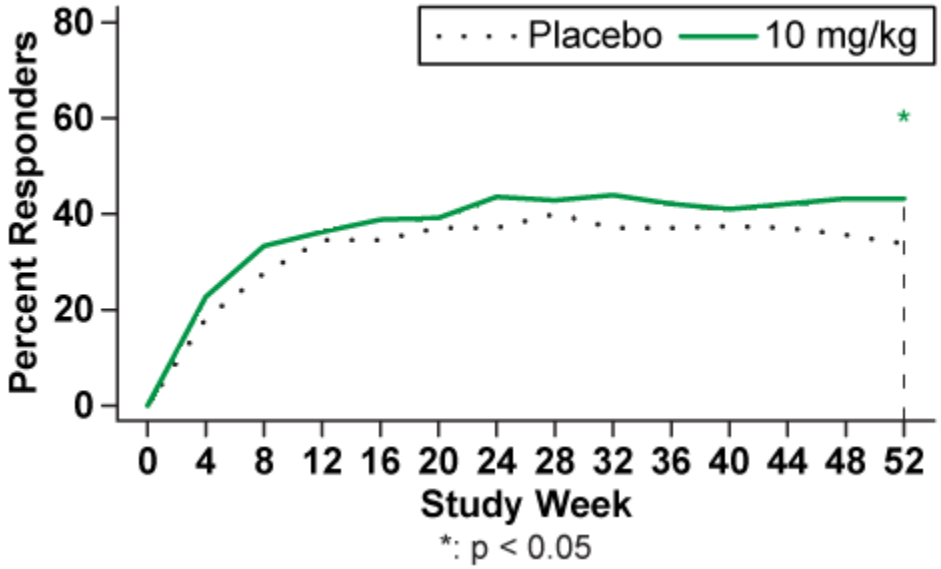


Figure 5.8. Response by Visit through Week 52 – Pooled Total Population

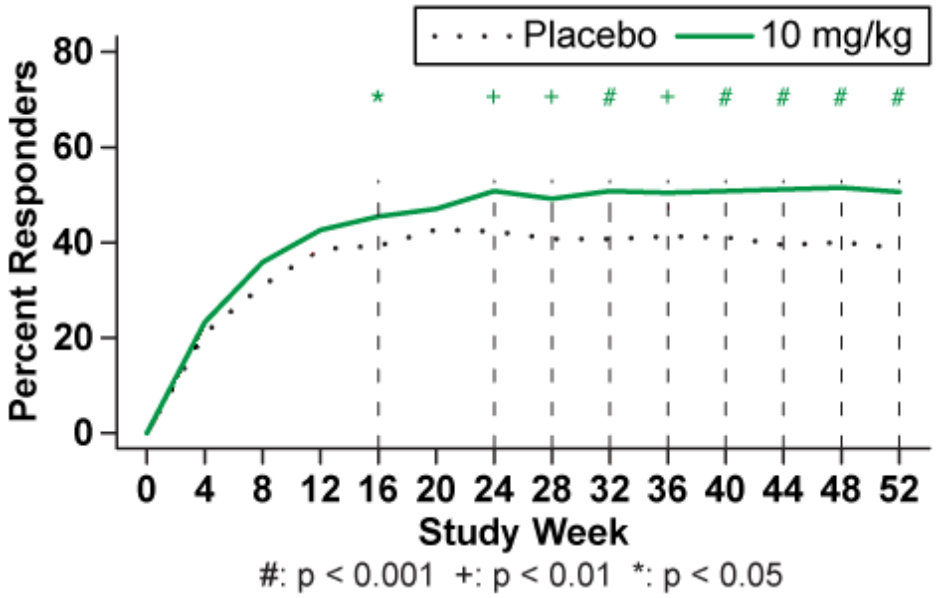
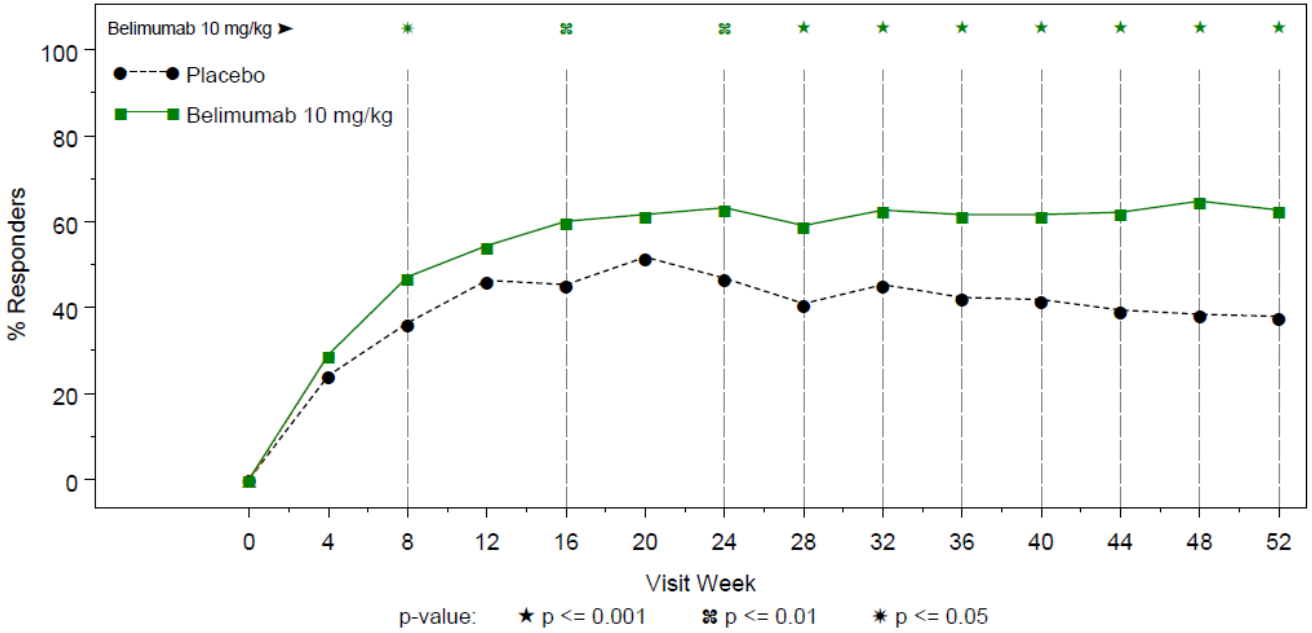


Figure 5.9. Response by Visit through Week 52 – High Disease Activity Subgroup



The Phase 3 study results for the components of the primary efficacy endpoint (SRI) are shown below (see Table 5.16). For BLISS-52, the pooled total population and the high disease activity subgroup belimumab 10 mg/kg was superior to placebo for each of the 3 components. For BLISS-76, a statistically significant response with belimumab 10 mg/kg over placebo was only seen with the 4-point reduction in SELENA-SLEDAI component.

Table 5.16. Components of the SRI at Week 52 (adjusted)

	BLISS-52		BLISS-76		Pooled Total Population[4]		High Disease Activity Subgroup	
	Placebo N = 287	10 mg/kg N = 290	Placebo N = 275	10 mg/kg N = 273	Placebo N = 562	10 mg/kg N = 563	Placebo N = 203	10 mg/kg N = 193
4-point reduction in SELENA-SLEDAI	132 (46.0%)	169 (58.3%)	98 (35.6%)	128 (46.9%)	230 (40.9%)	297 (52.8%)	84 (41.4%)	125 (64.8%)
Observed difference vs placebo (%)	-	12.3	-	11.3	-	11.9	-	23.4
OR (95% CI)[1] vs placebo	-	1.71 (1.21, 2.41)	-	1.63 (1.15, 2.32)	-	1.68 (1.3, 2.2)	-	2.6 (1.7, 3.9)
P-value[1]		0.0024		0.0062	-	< 0.0001	-	< 0.0001
No New 1A/2B BILAG domain scores	210 (73.2%)	236 (81.4%)	179 (65.1%)	189 (69.2%)	389 (69.2%)	425 (75.5%)	125 (61.6%)	145 (75.1%)
Observed difference vs placebo (%)	-	8.2	-	4.1	-	6.3	-	13.6
OR (95% CI)[1,2] vs placebo	-	1.62 (1.09, 2.42)	-	1.20 (0.84, 1.73)	-	1.4 (1.1, 1.8)	-	1.9 (1.2, 3.0)
P-value[1,2]		0.0181		0.3193	-	0.0190	-	0.0034
No worsening in PGA	199 (69.3%)	231 (79.7%)	173 (62.9%)	189 (69.2%)	372 (66.2%)	420 (74.6%)	119 (58.6%)	142 (73.6%)
Observed difference vs placebo (%)	-	10.4	-	6.3	-	8.4	-	15.0
OR (95% CI)[1,3] vs placebo	-	1.74 (1.18, 2.55)	-	1.32 (0.92, 1.90)	-	1.5 (1.2, 2.0)	-	2.0 (1.3, 3.1)
P-value[1,3]	-	0.0048	-	0.1258	-	0.0017	-	0.0015

[1] Odds Ratio (95% confidence interval) and p-value were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, covariates include baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.

[2] Additional covariate: baseline BILAG domain involvement (at least 1A/2B).

[3] Additional covariate: baseline PGA score.

[4] No significant treatment-by-study interactions were observed (all p > 0.287).

Modified SRI Response Removing Serology Components

On the SELENA-SLEDAI disease activity index, 2 points each are given for increased DNA binding and low complement items. Since the 4 point reduction in SELENA-SLEDAI required by the SRI can be achieved by normalisation of anti-dsDNA antibodies and complement, an analysis of a modified SRI response was undertaken in which the increased DNA binding and low complement items were removed from the SELENA-SLEDAI component of the SRI. This analysis was performed in patients that still had a SELENA SLEDAI score ≥ 4 at baseline after points for low complement and increased DNA binding were removed from the scale.

In the individual studies, the response rates with belimumab 10 mg/kg remained greater than placebo and the difference between 10 mg/kg and placebo was similar to that observed in the primary analysis. In BLISS-52, response rates were 48% and 61% in the placebo and 10 mg/kg groups respectively ($p = 0.0038$). In BLISS-76, the response rates were 36% and 45%, in the placebo and 10 mg/kg groups respectively ($p = 0.0604$). In the pooled total population, the response rates excluding the serology components were 42% and 53% in the placebo and belimumab 10 mg/kg groups, respectively, differences that reached statistical significance.

This result shows that improvement in serological activity, although considered an important clinical outcome, does not drive the SRI responses observed with belimumab and that patients receive benefit in clinical manifestations over and above serological changes.

SELENA-SLEDAI

Reductions in SELENA-SLEDAI score are clinically important because they represent resolution of individual manifestations of the patient's disease activity. Moreover, on the SELENA-SLEDAI, a reduction of 4 points equates to elimination (rather than mere improvement) of a disease manifestation and, as such, is a clear demonstration of clinical benefit. The SELENA-SLEDAI has been shown to be valid, reliable and sensitive to change (Griffiths et al. 2005). Disease activity as measured by SLEDAI correlates significantly with organ damage by SLICC/ACR Damage Index in SLE patients (Swaak et al. 1999).

A similar pattern of response as was seen for 4 point reduction in SELENA-SLEDAI was observed in the mean and mean percentage change from baseline. Statistically significant reductions in SELENA-SLEDAI score with belimumab 10 mg/kg compared with placebo were observed in the individual studies, pooled total population and high disease activity subgroup, with the greatest reductions seen in the high disease activity subgroup (see Table 5.17).

Table 5.17. SELENA-SLEDAI mean and mean percent reduction from baseline at Week 52

	BLISS-52		BLISS-76		Pooled Total Population[2]		High Disease Activity Subgroup	
Change from Baseline at Week 52	Placebo N= 287	10 mg/kg N= 290	Placebo N= 275	10 mg/kg N= 273	Placebo N = 562	10 mg/kg N = 563	Placebo N = 203	10 mg/kg N = 193
Mean change from baseline (± SE)	-3.57 ± 0.24	4.97 ± 0.27	-2.77 ± 0.25	-3.70 ± 0.27	-3.18 ± 0.18	-4.36 ± 0.19	-4.1 ± 0.3	-5.8 ± 0.3
P-value[1]	-	<0.0001	-	0.0063	-	<0.0001	-	0.0005
Mean % change (± SE)	-34.76 ± 2.50	-45.60 ± 2.45	-25.97 ± 2.72	-35.94 ± 2.80	-30.47 ± 1.85	-40.93 ± 1.86	-30.5 (2.3)	-45.5 (2.4)
P-value[1]	-	0.0018	-	0.0073	-	<0.0001	-	<0.0001

[1] ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for baseline SELENA SLEDAI score (≤ 9 vs. ≥ 10), baseline proteinuria level (< 2 g/24 hour vs. ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs. other). For pooled data analysis, study was also included as an additional covariate.

[2] No treatment-by-study interactions observed (all p-values > 0.367).

Physician's Global Assessment

The major secondary PGA endpoint was mean change from baseline in PGA at Week 24. The mean percent change and change from baseline at Week 24 and Week 52 are presented in the Table 5.18 below.

Belimumab 10 mg/kg was associated with a significant mean reduction (improvement) from baseline in PGA compared with placebo at Week 24 in BLISS-52, while this effect was not observed in BLISS-76. One factor that may partially explain the outcome in the 10 mg/kg dose group in BLISS-76, is the imbalance in baseline PGA scores between treatment groups: 19% of subjects in the 10 mg/kg group entered the study with low PGA scores (0-1) and little room for improvement, compared with only 12% in the placebo group and 14% with low PGA scores. In contrast, in BLISS-52, 15% and 11% entered the study with low PGA scores in the placebo and 10 mg/kg belimumab groups, respectively. Accordingly, only in BLISS-52 did belimumab demonstrate superiority over placebo for mean reduction in PGA at Week 24 with the 10 mg/kg belimumab dose ($p = 0.0003$), while in the 10 mg/kg belimumab group significant mean percent decreases were observed (37%; $p < 0.0001$) compared with placebo (22%). In the pooled total population, belimumab 10 mg/kg was superior to placebo for both PGA mean change and percent change from baseline.

By Week 52, statistically significant improvement was seen in PGA compared with placebo in the 10 mg/kg belimumab group in BLISS-52, while in BLISS-76, numerical improvements were observed but statistical significance was not achieved at Week 52.

For the pooled total population and high disease activity subgroup, belimumab 10 mg/kg was associated with significant improvements in PGA for both mean change and percent change compared with placebo by Week 24 that were sustained through Week 52.

Table 5.18. PGA mean percent change and change from baseline at Week 24 and Week 52

	BLISS-52		BLISS-76		Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 287	10 mg/kg N = 290	Placebo N = 275	10 mg/kg N = 273	Placebo N = 562	10 mg/kg N = 563	Placebo N = 203	10 mg/kg N = 193
Major secondary endpoint at Week 24								
Change (n)	287	290	275	273	562	563	203	193
Mean ± SE	-0.39 ± 0.03	-0.54 ± 0.03	-0.49 ± 0.04	-0.44 ± 0.03	-0.44 ± 0.02	-0.49 ± 0.02	-0.42 ± 0.04	-0.52 ± 0.04
LS Mean ± SE[1]	-0.35 ± 0.04	-0.50 ± 0.04	-0.49 ± 0.05	-0.48 ± 0.05	-0.40 ± 0.03	-0.48 ± 0.03	-0.41 ± 0.05	-0.53 ± 0.05
P-value[1]	-	0.0003	-	0.7987	-	0.0167	-	0.0268
Other secondary endpoints								
Week 24								
Percent change (n)	287	290	275	272	562	562	203	193
Mean ± SE	-22.44 ± 2.64	-36.75 ± 2.39	-26.18 ± 4.21	-27.57 ± 3.37	-24.27 ± 2.46	-32.30 ± 2.05	-22.55 ± 3.26	-30.78 ± 3.66
LS Mean ± SE[1]	-20.10 ± 3.29	-34.44 ± 3.31	-28.16 ± 6.17	-31.90 ± 6.04	-23.13 ± 3.24	-32.18 ± 3.22	-22.58 ± 4.00	-31.37 ± 3.94
P-value[1]	-	<0.0001	-	0.4682	-	0.0029	-	0.0453
Week 52								
Percent change (n)	287	290	275	272	562	562	203	193
Mean ± SE	-27.83 ± 3.45	-45.68 ± 2.66	-26.34 ± 3.16	-29.34 ± 3.97	-27.11 ± 2.34	-37.77 ± 2.38	-20.98 ± 3.82	-37.50 ± 4.09
LS Mean ± SE[1]	-22.04 ± 3.95	-39.89 ± 3.97	-29.67 ± 6.00	-34.65 ± 5.88	-23.52 ± 3.41	-35.24 ± 3.39	-18.73 ± 4.66	-35.63 ± 4.60
P-value[1]	-	<0.0001	-	0.3204	-	0.0002	-	0.0010
Change (n)	287	290	275	273	562	563	203	193
Mean ± SE	-0.48 ± 0.04	-0.67 ± 0.04	-0.46 ± 0.04	-0.49 ± 0.04	-0.47 ± 0.03	-0.58 ± 0.03	-0.41 ± 0.05	-0.62 ± 0.05
LS Mean ± SE[1]	-0.38 ± 0.05	-0.57 ± 0.05	-0.47 ± 0.06	-0.55 ± 0.06	-0.40 ± 0.04	-0.54 ± 0.04	-0.36 ± 0.06	-0.59 ± 0.06
P-value[1]	-	0.0001	-	0.1159	-	< 0.0001	-	0.0003

[1] All statistics, including the difference in LSM (least square means), were from ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for the baseline PGA score, baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.

SF-36 Physical Component Summary (PCS)

The major secondary endpoint mean change in SF-36 PCS at Week 24 was not achieved in either belimumab treatment group as shown in Table 5.19. The belimumab 10 mg/kg dose showed improvement at Week 52 only in BLISS-52 ($p = 0.0247$). For the pooled total population, belimumab 10 mg/kg was associated with significant improvements in mean PCS score over placebo at Week 52 ($p = 0.0374$). Whilst there were trends towards improvements in mean PCS score over placebo at Week 52 in the high disease activity subgroup, these did not reach significance, most likely due to smaller patient numbers (see Table 5.19).

Table 5.19 SF-36 PCS score change from baseline at Week 24 and Week 52 (LOCF)

	BLISS-52		BLISS-76		Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 287	10 mg/kg N = 290	Placebo N = 275	10 mg/kg N = 273	Placebo N = 562	10 mg/kg N = 563	Placebo N = 203	10 mg/kg N = 193
Week 24 (major secondary endpoint)								
n	286	284	274	269	560	553	203	187
Mean ± SE	3.64 ± 0.42	3.58 ± 0.46	3.36 ± 0.51	3.22 ± 0.43	3.50 ± 0.33	3.41 ± 0.32	4.69 ± 0.50	4.93 ± 0.63
Median	3.11	3.06	3.03	2.66	3.03	2.77	4.14	4.84
LS Mean ± SE[1]	3.26 ± 0.54	3.34 ± 0.55	5.63 ± 0.74	5.36 ± 0.72	4.05 ± 0.44	3.98 ± 0.44	5.02 ± 0.65	5.59 ± 0.65
Treatment differences (95% CI)[1] vs placebo	-	0.08 (-1.00, 1.15)	-	-0.27 (-1.48, 0.94)	-	-0.07 (-0.88, 0.74)	-	0.57 (-0.84, 1.97)
P-value[1]	-	0.8870	-	0.6601	-	0.8679	-	0.4276
Week 52								
n	286	284	274	269	560	553	203	187
Mean ± SE	2.96 ± 0.45	4.18 ± 0.48	2.85 ± 0.52	3.41 ± 0.47	2.91 ± 0.34	3.80 ± 0.34	4.21 ± 0.54	5.09 ± 0.68
Median (Min, Max)	2.60 (-22.20, 26.19)	3.74 (-24.41, 30.24)	2.62 (-27.48, 33.20)	2.87 (-21.53, 31.25)	2.62 (-27.48, 33.20)	3.41 (-24.41, 31.25)	3.88 (-20.90, 23.06)	5.80 (-24.41, 30.24)
LS Mean ± SE	2.84 ± 0.60	4.19 ± 0.60	4.60 ± 0.78	5.03 ± 0.77	3.38 ± 0.47	4.31 ± 0.47	4.37 ± 0.70	5.59 ± 0.69
Treatment differences (95% CI) vs placebo	-	1.35 (0.17, 2.54)	-	0.43 (-0.86, 1.71)	-	0.93 (0.05, 1.80)	-	1.22 (-0.29, 2.73)
P-value[2]	-	0.0247	-	0.5134	-	0.0374	-	0.1124

[1] All statistics, including the difference in LSM (least square means), were from ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for the baseline PCS score and baseline stratification factors.

Steroid Use

The total dose of systemic steroids could be adjusted as clinically required during the first 6 months of the trial, but had to return to within 25% or no more than 5 mg over the baseline dose by Study Day 168 (Week 24 visit). After the Day 168 visit, an increase in steroids of > 25% or > 5 mg over baseline (whichever is higher) for SLE activity would result in the subject being declared a non-responder. The restriction of steroid use for the remainder of the trial was to evaluate improvement in disease activity with belimumab (refer to Figure 5.2). Furthermore, since the primary objective of the Phase 3 studies was to evaluate the ability of belimumab to reduce SLE disease activity, steroid reductions were not mandated per protocol given the concern that premature or rapid reduction in steroids may confound interpretation of study results and could induce a flare. As such, steroid reductions were to be performed as clinically indicated, and it was recommended that steroid reduction should not be undertaken unless the subject had improving disease activity for at least 8 weeks.

A major secondary endpoint was the percentage of subjects whose average prednisone (equivalent) dose was reduced by ≥ 25% from baseline to ≤ 7.5 mg/day during Weeks 40-52 in the subgroup of subjects who were receiving > 7.5 mg/day of prednisone at baseline, since this reduction in steroids was felt to represent clinically meaningful steroid sparing in SLE subjects (Fortin et al. 2008). This endpoint measured steroid reduction from baseline dose that was sustained over the last 12 weeks of the study.

In BLISS-52, the percentage of subjects whose average prednisone dose was reduced by ≥ 25% from baseline to ≤ 7.5 mg/day during Weeks 40-52 was higher (but did not reach statistical significance) in the 10 mg/kg group vs placebo (19% vs 12%, p = 0.0526) (see Table 5.20). In BLISS-76, although a greater percentage of subjects in the belimumab 10 mg/kg group met this endpoint (4% more than placebo), this difference did not reach statistical significance (see Table 5.20). Fewer subjects had a

baseline prednisone dose > 7.5 mg/day in BLISS-76 (46% of subjects overall or 120-126 subjects per group) vs BLISS-52 (69% of subjects overall or 192-204 subjects per group). Although the magnitude of the steroid sparing effect was similar in both studies, the smaller subgroup size in BLISS-76 limited the ability to achieve statistical significance. However, for the pooled total population, both belimumab doses achieved significance compared with placebo for reduction of prednisone by ≥ 25% from baseline to ≤ 7.5 mg/day during Weeks 40-52 as shown (see Table 5.20) and the odds ratio for prednisone reduction was similar across the individual trials and pooled total population. In the high disease activity subgroup, belimumab 10 mg/kg achieved the largest treatment response compared with placebo for reduction of prednisone by ≥ 25% from baseline to ≤ 7.5 mg/day during Weeks 40-52 as shown (see Table 5.20).

Table 5.20. Prednisone reduction by ≥ 25% from baseline to ≤ 7.5mg/day during Weeks 40 through 52 – Phase 3 trials[1]

	BLISS-52		BLISS-76		Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 192	10 mg/kg N = 204	Placebo N = 126	10 mg/kg N = 120	Placebo N = 318	10 mg/kg N = 324	Placebo N = 126	10 mg/kg N = 126
No.(%) Response[2]	23 (12.0%)	38 (18.6%)	16 (12.7%)	20 (16.7%)	39 (12.3%)	58 (17.9%)	9 (7.1%)	20 (15.9%)
Observed difference vs Placebo	-	6.65	-	3.97	-	5.64	-	8.73
OR (95% CI)[3] vs placebo	-	1.75 (0.99, 3.08)	-	1.26 (0.61, 2.60)	-	1.57 (1.01, 2.45)	-	2.43 (1.05, 5.65)
P-value[3]	-	0.0526	-	0.5323	-	0.0451	-	0.0389
Treatment by study interaction p-value[4]	-	N/A	-	N/A	-	0.5177	-	N/A

[1] Includes only subjects with baseline prednisone > 7.5 mg/day.

[2] Any subject who withdrew from the study prior to the Day 364 (Week 52) visit, missed the Day 364 (Week 52) visit (± 28 day window allowed), and/or received a protocol-prohibited medication or a dose of allowable (but protocol-restricted) medication that resulted in treatment failure designation prior to the Day 364 (Week 52) visit was considered a treatment failure for prednisone reduction.

[3] Odds Ratio (95% confidence interval) and p-value were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, the covariates include baseline prednisone level, baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.

[4] Obtained from a logistic regression by adding study and the treatment-by-study interaction to the above model.

In summary, belimumab 10 mg/kg generally demonstrated greater reductions in steroid use compared with placebo. It should be noted that steroid taper was not mandated in the Phase 3 studies and was based upon the investigator's clinical judgement, and rapid steroid reduction was discouraged to prevent escape of disease control and possible SLE flare. Overall, the totality of the evidence supports the conclusion that belimumab treatment allows a reduction in steroids, particularly for subjects with a baseline prednisone equivalent dose greater than 7.5 mg/day.

Flare

SLE flares serve as an important surrogate for SLE disease control. Flares increase with increases in disease activity. A reduction in flares is the goal of any SLE treatment.

SLE flares were measured in the Phase 3 studies was using the modified SELENA-SLEDAI SLE Flare Index (SFI), where the modification excludes severe flares that are triggered *only* by an increase of SELENA-SLEDAI score to > 12 (i.e. at least one of the other severe flare criterion on the SFI must be present irrespective of the SELENA-SLEDAI score).

Time to first flare and time to first severe flare over 52 weeks is shown in Figures 5.10, 5.11, 5.12 and 5.13. In BLISS-76, a trend was noted in the 10 mg/kg group for time to first severe flare (HR = 0.72 [0.50, 1.05], p = 0.0867) (see Figure 5.11). In BLISS-52, the 10 mg/kg belimumab group was associated with a significantly lower risk of severe flare over 52 weeks (HR = 0.57 [0.39, 0.85], p = 0.0055) (see Figure 5.10). Subjects in the belimumab 10 mg/kg group had a statistically significantly lower risk for severe flare compared with placebo over 52 weeks in the pooled total population (HR = 0.64, [0.49, 0.84], p = 0.0011) (see Figure 5.12). Although there was no difference in the median time to first flare overall for mild, moderate or severe flares in BLISS-76 (see Figure 5.11), the belimumab 10 mg/kg group in BLISS-52 and the pooled total population had significantly lower flare rates (see Figure 5.10 and 5.12). In BLISS-52, the belimumab 10mg/kg group significantly reduced the risk of flare compared with placebo (HR = 0.76 [0.63, 0.91], p = 0.0036) (see Figure 5.10). In the pooled total population, the belimumab 10 mg/kg group significantly reduced the risk of flare overall (HR = 0.84 [0.74, 0.96], p = 0.0120) (see Figure 5.12).

For the high disease activity subgroup, belimumab 10 mg/kg significantly reduced the risk of flare overall (HR = 0.70 [0.56, 0.88], p = 0.0017) and severe flares (HR = 0.55 [0.37, 0.81], p = 0.0028) over 52 weeks (see Figure 5.13).

Figure 5.10. Time to SLE flare and severe SLE flare over 52 weeks by modified SLE flare index – BLISS-52

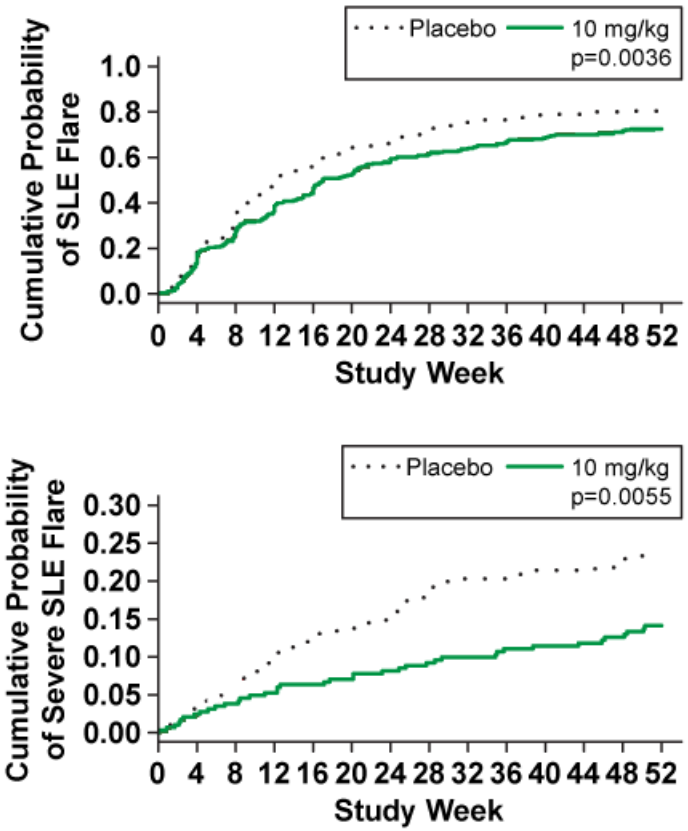


Figure 5.11. Time to SLE flare and severe SLE flare over 52 weeks by modified SLE flare index – BLISS-76

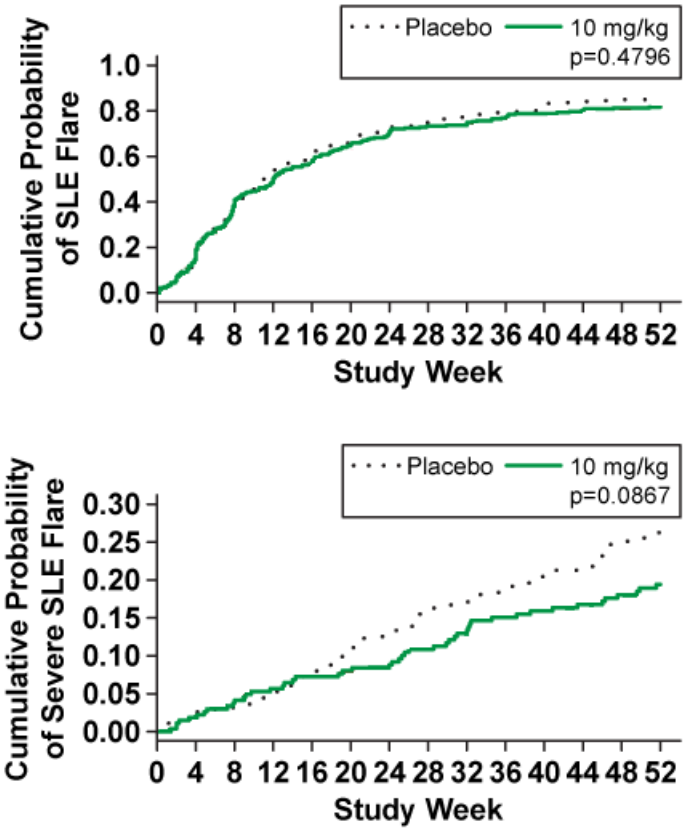


Figure 5.12. Time to SLE flare and severe SLE flare over 52 weeks by modified SLE flare index – Pooled Total Population

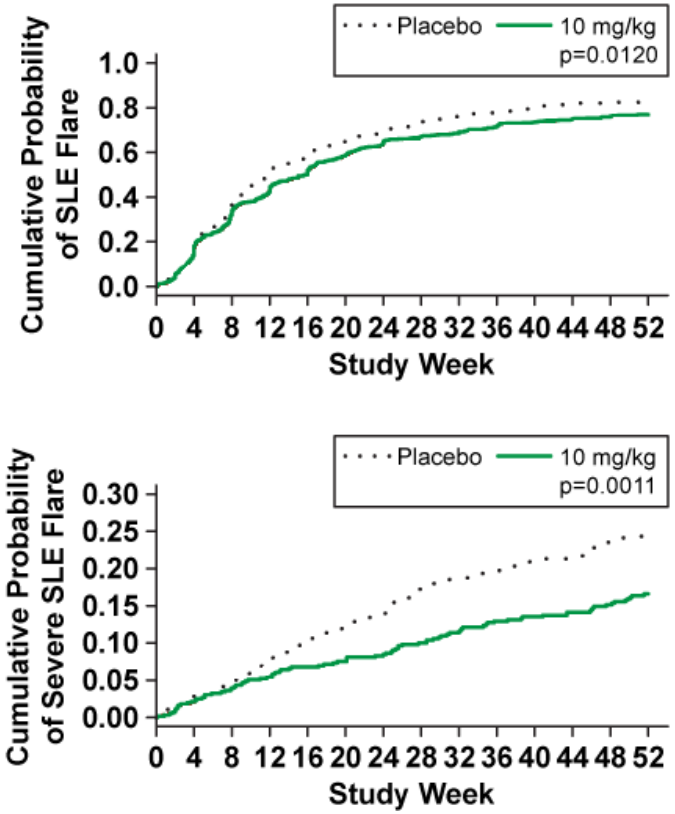
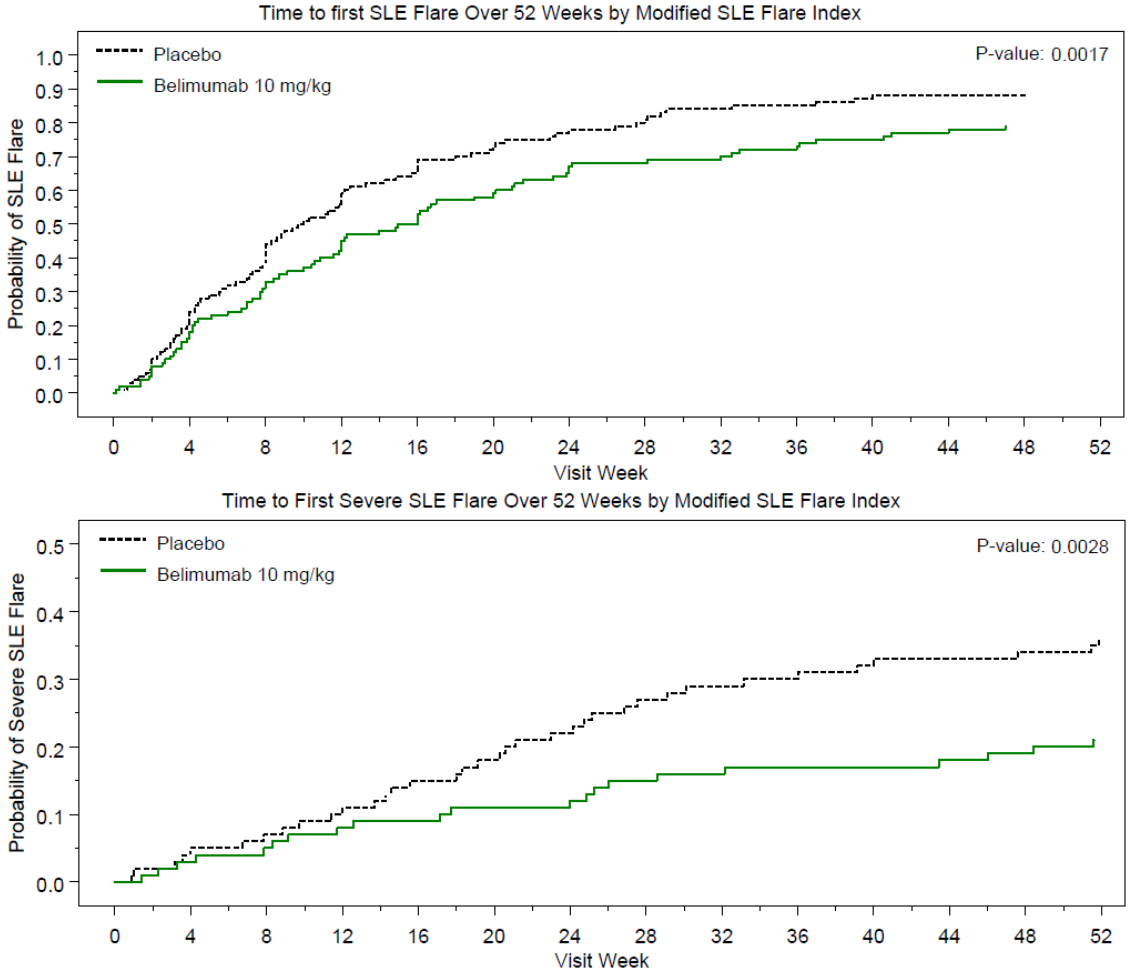


Figure 5.13. Time to SLE flare and severe SLE flare over 52 weeks by modified SLE flare index – High Disease Activity Subgroup



SLICC/ACR Damage Index

There was no difference between the belimumab and placebo groups in the change in SLICC/ACR Damage Index at Week 52 compared with baseline. Damage is an important outcome measure in observational studies and clinical trials. However, if new damage that has accrued during a clinical trial is to be recorded then the trial should be over 12 months duration, as some items of damage can be recorded only if they have been present for at least 6 months. The duration of the Phase 3 studies may have been inadequate to measure clinically detectable progression using the SLICC/ACR Damage Index. Nonetheless, all groups had small increases in mean scores (0.05-0.06 and 0.03-0.04 in the placebo and 10 mg/kg belimumab groups, respectively).

FACIT-Fatigue Scale

A composite fatigue score was created from the FACIT-Fatigue questionnaire and the mean change from baseline by visit was compared between each belimumab treatment group and placebo as shown in Figures 5.14, 5.15, 5.16 and 5.17.

There were numerical improvements in the FACIT-Fatigue score (although not significant) in the 10 mg/kg belimumab group over placebo in BLISS-76 (see Figure 5.15). In BLISS-52, significant improvement in FACIT-Fatigue score was observed with the 10 mg/kg belimumab dose at Week 52 (see Figure 5.14). A rapid onset of fatigue improvement in the belimumab group was observed by Week 8 and was generally sustained through Week 52. In the placebo group, improvement was also observed, but after Week 24, fatigue scores began to decline (i.e. worsening fatigue) likely due to the additional concomitant medication restrictions required after that visit. The pooled data from both studies showed that belimumab 10 mg/kg was associated with significantly improved fatigue scores compared with placebo at Weeks 8, 12, and 52 (see Figure 5.16). In the high disease activity subgroup, belimumab

10 mg/kg was associated with significantly improved fatigue scores compared with placebo at Weeks 8 and 12 and a clinically important but not statistically significant difference at Week 52 (see Figure 5.17).

Overall, belimumab demonstrated improvement in fatigue compared with placebo measured by the increase in FACIT-Fatigue scores from baseline. The onset of improvement in fatigue occurred by Week 8 and the benefit of reduction in fatigue was sustained through Week 52. Fatigue is an important HRQL measure that has meaningful impact upon SLE patients. This finding appears to reflect the improvements observed in other clinical and biomarker measures of SLE disease activity.

Figure 5.14. Mean change in FACIT-Fatigue score over time (LOCF) – BLISS-52

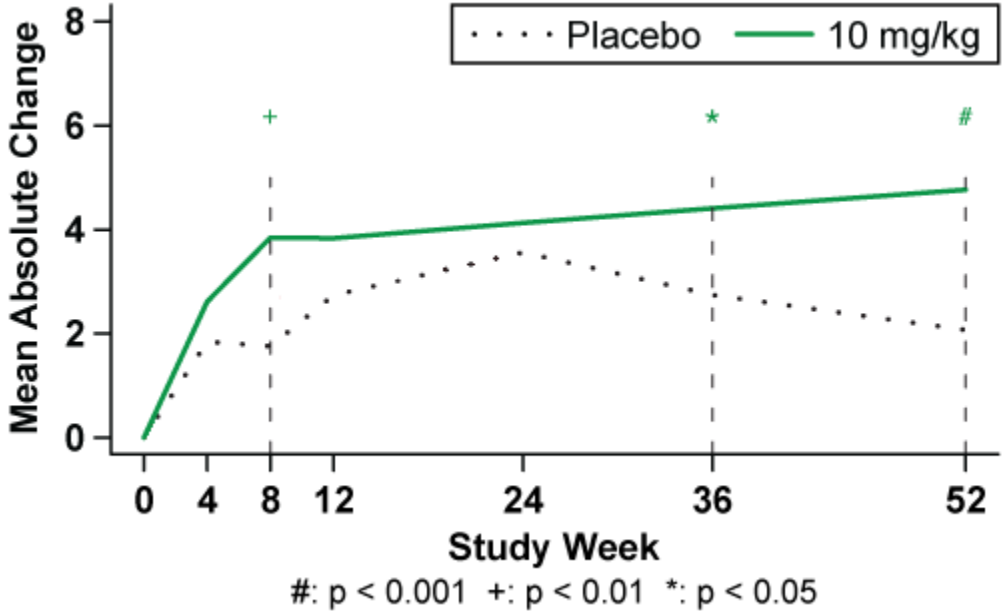


Figure 5.15. Mean change in FACIT-Fatigue score over time (LOCF) – BLISS-76

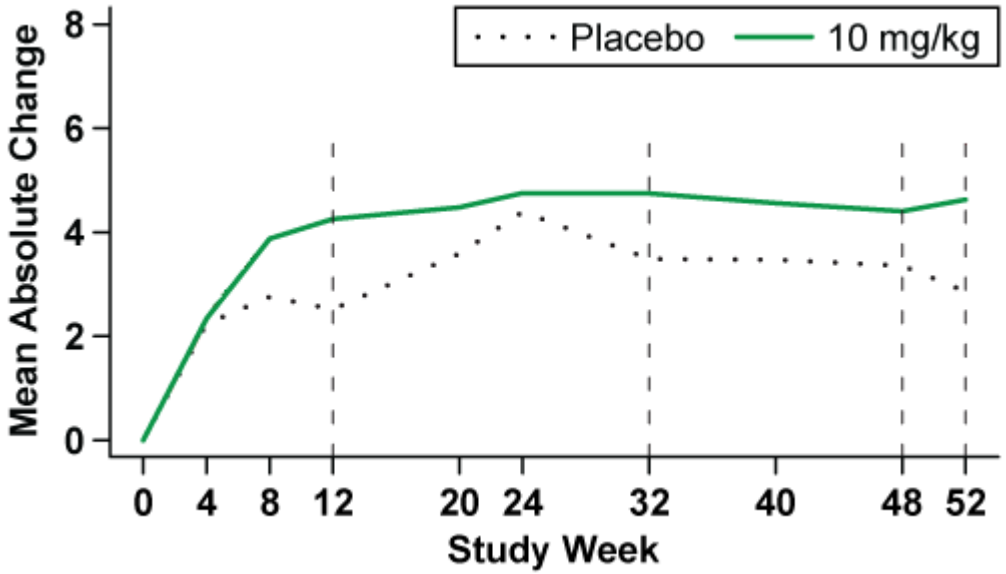


Figure 5.16. Mean change in FACIT-Fatigue score over time (LOCF) – Pooled Total Population

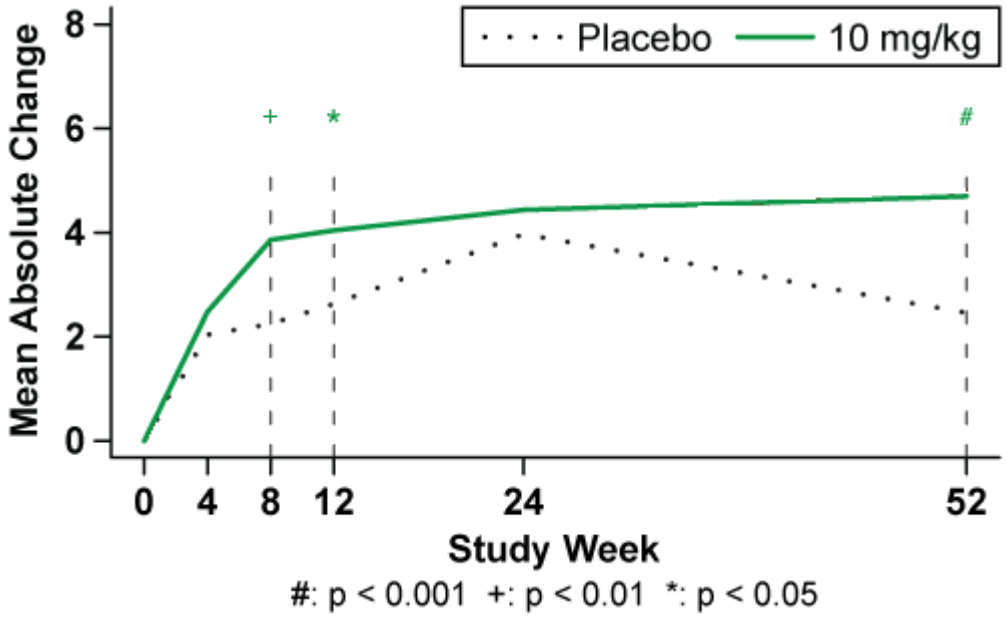
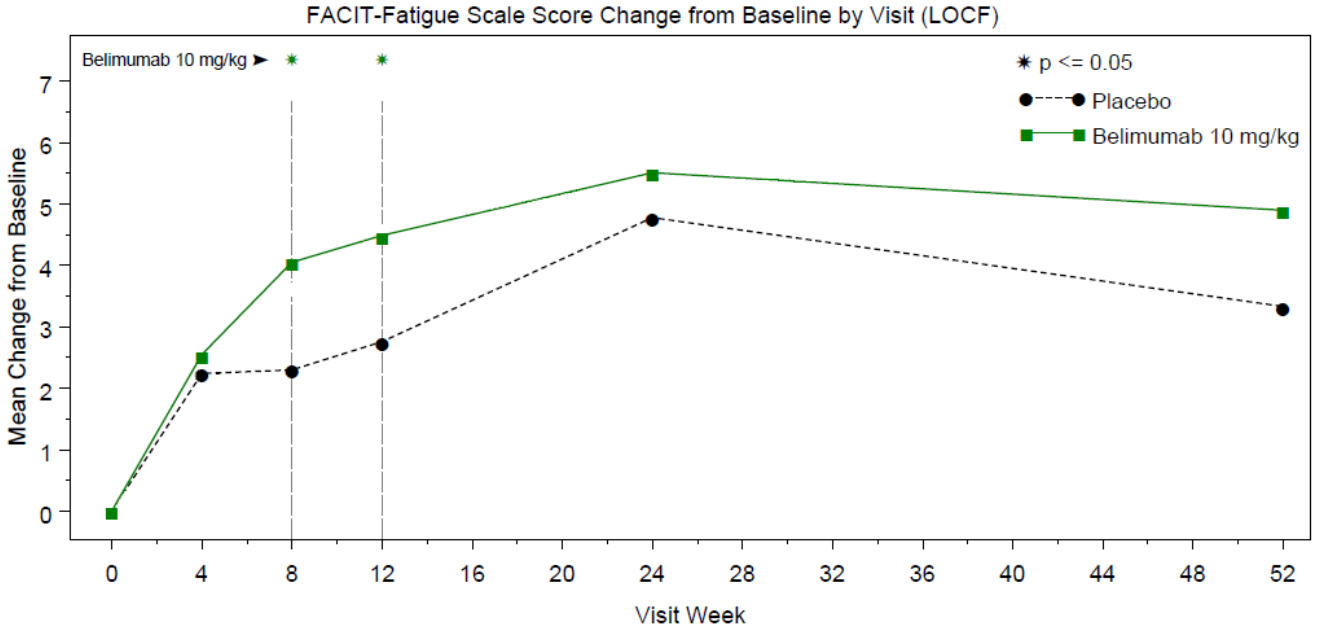


Figure 5.17. Mean change in FACIT-Fatigue score over time (LOCF) – High Disease Activity Subgroup



EQ-5D

The EQ-5D is a standardised instrument used to measure health outcomes because it is applicable to a wide range of health conditions and treatments and it is designed to provide a single index value for health status (http://www.euroqol.org). The EQ-5D consists of 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) each of which includes 1 of 3 responses to describe a subject's health state: 1) no problems, 2) some or moderate problems, and 3) severe problems.

In BLISS-52, the EQ-5D index scores were generally comparable across groups at baseline. There were no significant differences in the mean change from baseline EQ-5D utility index scores between the 10 mg/kg belimumab group compared with placebo at any time point. However, Figure 5.18 shows that from Week 4, numerical improvement in the EQ-5D utility index in the belimumab 10 mg/kg group was consistently greater than that of the placebo group.

In BLISS-76, the pooled total population and the high disease activity subgroup, baseline EQ-5D index scores were similar across treatment groups (see Figures 5.19 and 5.20). There were no significant differences in the mean change from baseline EQ-5D utility index scores between the 10mg/kg belimumab group compared with placebo at any timepoint, although there were numerical differences seen in the pooled total population and the high disease activity subgroup.

In the high disease activity subgroup, the was a difference in the mean change from baseline EQ-5D utility index scores between the 10mg/kg belimumab group compared with placebo which reached significance at Week 24 but was not maintained at Week 52.

EQ-5D may not be the most sensitive measure to assess the true impact of the disease on HRQL experienced by SLE patients. Patients may experience disease flares at any time and not necessarily at the time the EQ-5D was completed for the pre-defined time points of the clinical trials. In addition, certain relevant dimensions of health that are not directly included in the EQ-5D instrument, such as fatigue or sensory impairment, or where the disease course is characterised by flares of unpredictable symptom severity (Wailoo et al. 2010). This might explain the differences observed in the results for EQ-5D and FACIT-Fatigue score.

Figure 5.18. Change from baseline in EQ-5D utility index through week 52 – BLISS-52

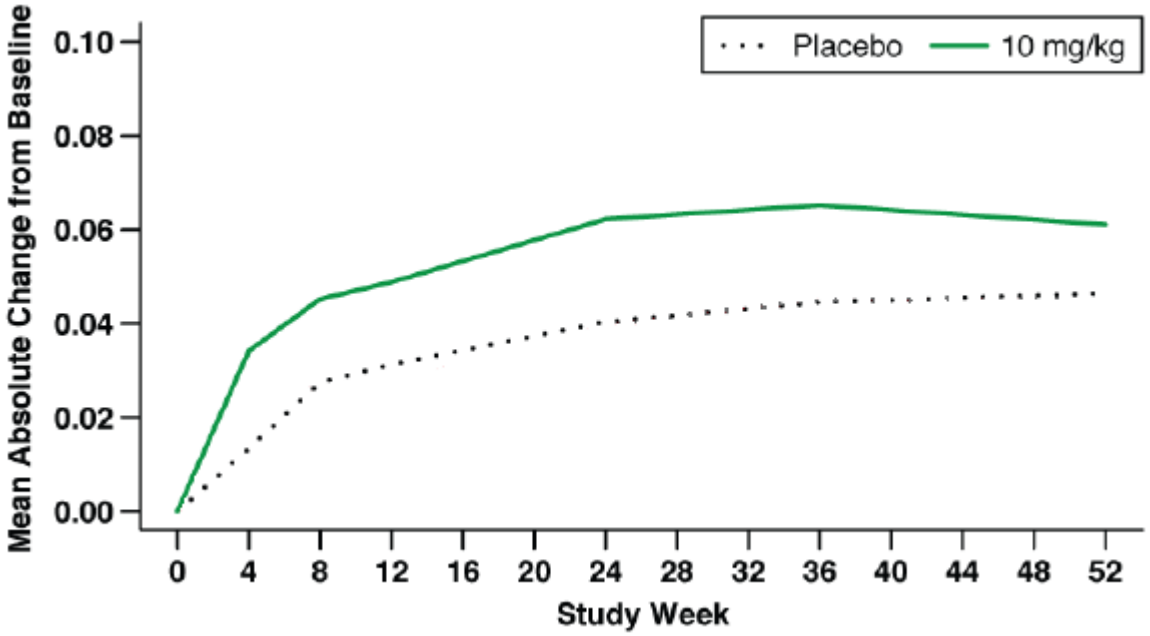


Figure 5.19. Change from baseline in EQ-5D utility index through week 52 – BLISS-76

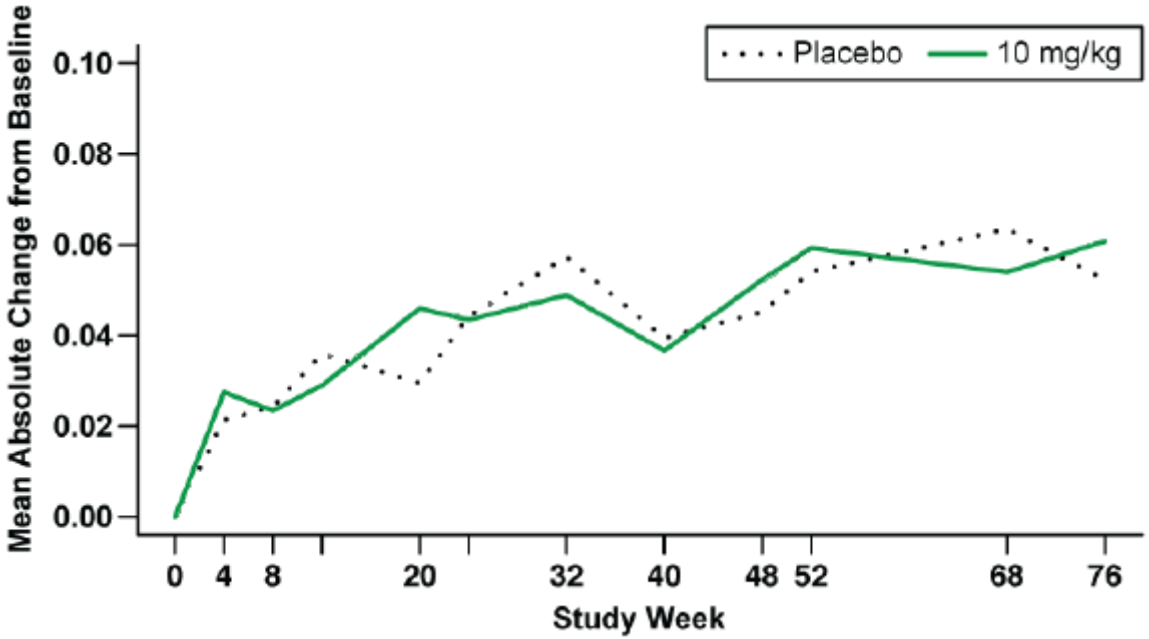


Figure 5.20. Change from baseline in EQ-5D utility index through week 52 – Pooled Total Population

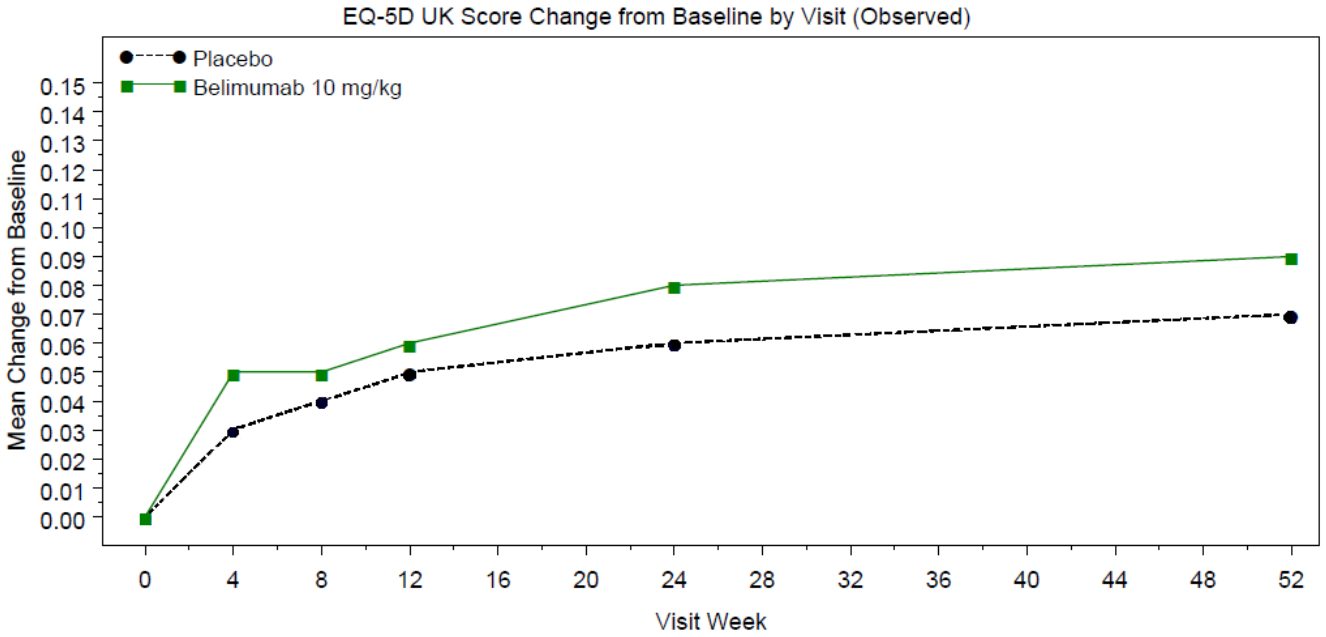
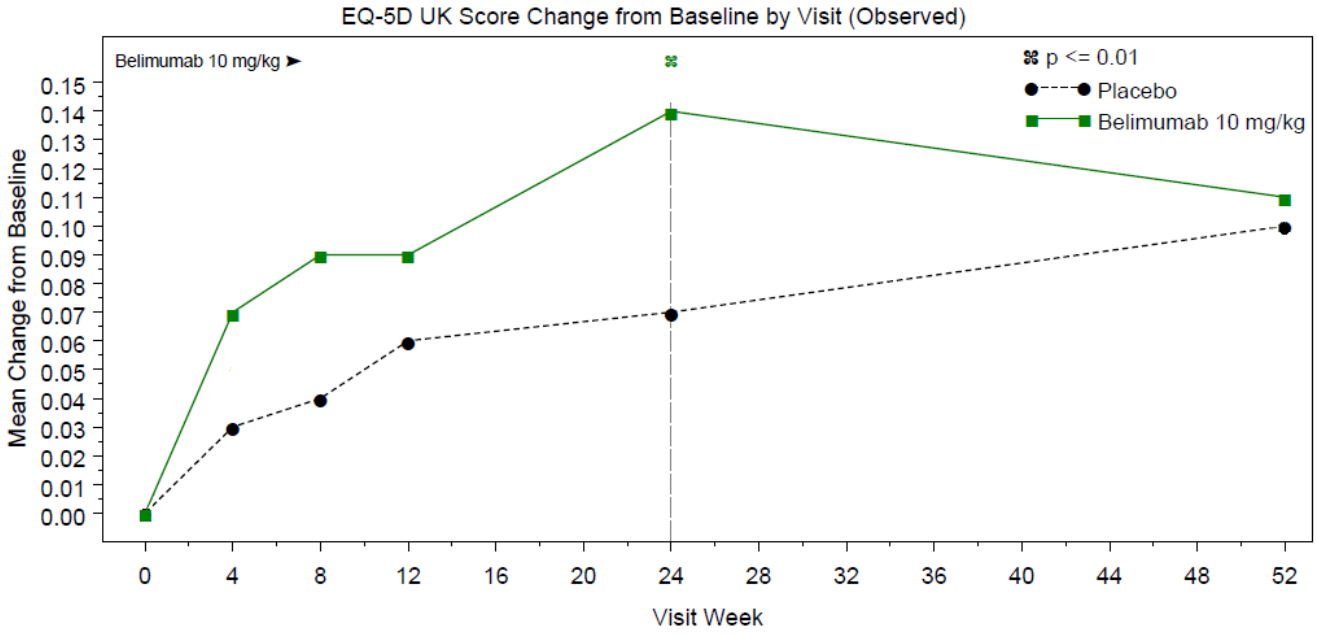


Figure 5.21. Change from baseline in EQ-5D utility index through week 52 – High Disease Activity Subgroup



Workplace Productivity Questionnaire

The Workplace Productivity Questionnaire consisted of the following three questions:

1. Have you been employed at one or more jobs in the past 4 weeks?
2. How many total days were you scheduled or planning to work at your paid jobs in the past 4 weeks?
3. How many days have you missed because of illness at your paid jobs in the past 4 weeks?

It was administered at baseline (Day 0) and every 4 weeks throughout the treatment and follow-up period.

BLISS-52

There were no significant difference between the 10 mg/kg belimumab group and placebo at any of the visits or at the 8-week follow-up in the percentage of subjects employed in the 4 weeks prior to each visit at any time point as well as those subjects who were unemployed at baseline.

BLISS-76

The percentage of subjects employed in the 4 weeks prior to each visit over 76 weeks was generally similar across groups in both subjects who were employed or unemployed at baseline. Among subjects employed at baseline, a significant difference favouring placebo over the 10 mg/kg belimumab group in percentage of subjects employed was observed at Weeks 36 and 56.

Workplace productivity is difficult to capture due to the short duration of the BLISS trials and the restrictions placed on patients within the RCT setting.

Gender and Age Subgroup

With reference to the decision problem, SLE largely affects women of child-bearing age. The majority of the patients in the Phase 3 studies were ≤ 45 years of age (74%) and female (94%); less than 2% of subjects were 65 years of age or older. When exploring the pre-specified subgroups of age (≤ 45 vs > 45) and gender, no interactions were observed, with belimumab treatment offering benefit in each subgroup relative to placebo.

Race Subgroups

Response with belimumab was evaluated by the race stratification used in the Phase 3 studies (African/indigenous-American (AIA) descent vs other), as well as more standard race classifications (i.e., white, black, Asian, Alaska native or American Indian, and Hawaiian or other Pacific islander (the latter group was not analysed given only 1 subject of this race was enrolled).

Race Stratification (African-American/Indigenous American Descent vs Other)

The race stratification when applied to the primary efficacy population, showed no treatment-by-subgroup interaction, although in the AIA subgroup the observed treatment differences in response rates for belimumab relative to placebo were not as large as those observed in the non-AIAs, with the lower magnitude of treatment effect among the AIA subgroup being driven by a high placebo response.

In AIA subjects, the placebo response rate was 48% compared with 35% in subjects not of this heritage (and 39% in the overall population), while responses with belimumab were similar across the subgroups and in the overall population groups (50-53% with

10 mg/kg); It is notable that the AIA subgroup behaved differently between BLISS-52 and BLISS-76. AIA subjects in BLISS-76 demonstrated a poorer response rate with belimumab relative to placebo. Additional exploration of this group was undertaken.

Within the AIA subgroup of subjects in BLISS-76, a greater number of subjects in the placebo group had high disease activity compared with the belimumab groups and fewer subjects in the 10 mg/kg groups were receiving steroids, antimalarials and immunosuppressants at baseline. As such, response in AIA subjects in BLISS-76 with higher baseline disease activity (SELENA-SLEDAI ≥ 10, n = 38-49/group) was explored, and belimumab response rates in this group improved to about 50% (from ~ 40%), similar to the rates observed with placebo (51%), although this result must be interpreted cautiously given the small sample size.

Notably, a similar effect was not observed in BLISS-52 in the AIA subgroup. It should be noted that in BLISS-76 53% of the subjects in the AIA subgroup were of black race, compared with ~ 10% of subjects in this group in BLISS-52.

Race

Among whites, Asians, and Alaska Native/American Indians, belimumab 10 mg/kg achieved a better response compared with placebo for all race subgroups; there were no significant treatment-by-subgroup interactions (or trends) observed (see Table 5.21).

Table 5.21. Primary response at Week 52 by race subgroup – Pooled Total Population

	Pooled Total Population	
	Placebo N = 562	10 mg/kg N = 563
Overall	218 (38.8%)	285 (50.6%)
Race		
White -Caucasian	94/270 (34.8%)	133/260 (51.2%)
Black - African American or African Heritage	22/50 (44.0%)	18/50 (36.0%)
Alaska Native or American Indian	61/125(48.8%)	75/126 (59.5%)
Asian	41/116 (35.3%)	59/127 (46.5%)

Among black subjects, there was a strong trend ($p < 0.06$) for a treatment-by-subgroup interaction in both belimumab groups vs placebo even though this group was small (~ 50 subjects/group; see Table 5.22). In black subjects (80% of whom participated in Study BLISS-76), the placebo response rate was higher (44%) than in the overall and non-black populations (~ 38%), while the belimumab response rates were lower (36% with 10 mg/kg belimumab) than those observed in the overall (51% with 10 mg/kg belimumab) and non-black (52% with 10 mg/kg belimumab) populations.

Table 5.22. Primary response at Week 52 by black race – Pooled Total Population

	Black		Others	
	Placebo N = 50	10 mg/kg N = 50	Placebo N = 512	10 mg/kg N = 513
No.(%) Response	22 (44.0%)	18 (36.0%)	196 (38.3%)	267 (52.0%)
Observed difference vs Placebo	-	-8.00	-	13.77
OR (95% CI)[1] vs placebo	-	0.76 (0.33, 1.75)	-	1.81 (1.40, 2.34)
P-value[2]	-	0.0521	-	NA

[1] From logistic regression for the comparison between each belimumab dose and placebo in pooled data. Independent variables will include treatment group, baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and study.
[2] For treatment by subgroup interaction effect from a logistic regression model by adding the subgroup and interaction effect to the above model.

5.6 Meta-analysis

When more than one study is available and the methodology is comparable, a meta-analysis should be undertaken. This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 5.3.9 to 5.3.12.

5.6.1 The following steps should be used as a minimum when presenting a meta-analysis.

- Perform a statistical assessment of heterogeneity. If the visual presentation and/or the statistical test indicate that the RCT results are heterogeneous, try to provide an explanation for the heterogeneity.
- Statistically combine (pool) the results for both relative risk reduction and absolute risk reduction using both the fixed effects and random effects models (giving four combinations in all).
- Provide an adequate description of the methods of statistical combination and justify their choice.
- Undertake sensitivity analysis when appropriate.
- Tabulate and/or graphically display the individual and combined results (such as through the use of forest plots).

Refer to pooled total population in Section 5.5.

5.6.2 If a meta-analysis is not considered appropriate, a rationale should be given and a qualitative overview provided. The overview should summarise the overall results of the individual studies with reference to their critical appraisal.

N/A

5.6.3 If any of the relevant RCTs listed in response to section 5.2.4 (Complete list of relevant RCTs) are excluded from the meta-analysis, the reasons for doing so should be explained. The impact

that each exclusion has on the overall meta-analysis should be explored.

See Section 5.2.6.

5.7 Indirect and mixed treatment comparisons

Data from head-to-head RCTs should be presented in the reference-case analysis, if available. If data from head-to-head RCTs are not available, indirect treatment comparison methods should be used. This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 5.3.13 to 5.3.22.

5.7.1 Describe the strategies used to retrieve relevant clinical data on the comparators and common references both from the published literature and from unpublished data. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used should be provided. Exact details of the search strategy used should be provided in section 9.4, appendix 4.

The search strategies of the systematic review as described in Section 5.1.1 covered the retrieval of clinical data for the comparators since the scope of the systematic review comprised all relevant comparators.

There are no studies directly comparing belimumab with rituximab. Differences in the end points considered and the patient populations preclude the conduct of any meaningful indirect and mixed treatment comparisons between belimumab and rituximab as outlined below.

The inclusion criteria of the published Phase 2/3 randomised, double-blind study of rituximab (two 1,000 mg IV doses given 14 days apart), required SLE patients to have significantly active disease at screening, defined as ≥ 1 organ system with a BILAG A score or ≥ 2 organ systems with a BILAG B score despite background immunosupressants at study entry (Merrill et al. 2010a). This is likely to correspond to a more severe patient population compared with both the pooled total population and the high disease activity subgroup, where 49% and 53% of patients respectively, were receiving an immunosuppressant

at baseline and 61% and 72% respectively, had ≥ 1 organ system with a BILAG A score or ≥ 2 organ systems with a BILAG B score.

The study did not collect data on changes in SELENA-SLEDAI, which as discussed previously is an important short-term outcome which can be linked to longer term impact on organ damage. So both the trial populations and the outcomes reported are different for the rituximab and Phase 3 belimumab trials, making any indirect comparisons of these technologies using this study problematic.

The primary endpoint was the effect of rituximab versus placebo in achieving and maintaining a major clinical response, a partial clinical response, or no clinical response at week 52 assessed using BILAG scores.

A major clinical response was defined as achieving BILAG C scores or better in all organs at week 24 without experiencing a severe flare (1 new domain with a BILAG A score or 2 new domains with a BILAG B score) from day 1 to week 24, and maintaining this response without a moderate or severe flare (≥1 new domains with a BILAG A or B score) to week 52.

A partial clinical response was defined as 1) achieving BILAG C scores or better at week 24 and maintaining this response without a new BILAG A or B score for 16 consecutive weeks, 2) achieving no more than 1 organ with a BILAG B score at week 24 without achieving ≥1 new BILAG A or B score to week 52, or 3) achieving a maximum of 2 BILAG B scores at week 24 without developing BILAG A or B scores in new domains until week 52 if the baseline BILAG score for the patient was 1 A score plus ≥2 B scores, ≥2 A scores, or ≥4 B scores.

No clinical response was defined as failure to meet the definition of a major clinical response or a partial clinical response. Patients who terminated the study early were scored as having no clinical response.

At week 52, no difference was noted in major clinical responses or partial clinical responses between the rituximab group (12.4% had a major clinical response, and 17.2% had a partial clinical response) and the placebo group (15.9% had a major clinical response, and 12.5% had a partial clinical response) relative to the overall response rate (29.6% versus 28.4%).

In addition, the rituximab trial demonstrated no difference in secondary endpoints between the rituximab group and the placebo group and over 52 weeks of treatment, in patients with moderate-to-severe SLE. Secondary end points are outlined below.

Secondary end points included:

1) the time-adjusted area under the curve minus baseline (AUCMB) of the BILAG score over 52 weeks

2) the proportion of patients who achieved a major clinical response (excluding a partial clinical response) and the proportion of patients with a partial clinical response (including a major clinical response) at week 52

3) the proportion of patients with a BILAG C score or better in all organs at week 24

4) the time to the first moderate or severe disease flare

5) improvement in quality of life as measured by the Lupus Quality of Life index

6) the proportion of patients who achieved a major clinical response with a prednisone dosage of <10 mg/day from week 24 to week 52.

The efficacy and safety of rituximab has also been investigated as part of an analysis of prospective data from the French AutoImmunity and Rituximab (AIR) registry (Terrier et al. 2010). One hundred thirty-six patients received treatment for SLE. The mean SELENA-SLEDAI score at baseline was 11.3 ± 8.9 which indicates a slightly less severe population than our high disease activity subgroup (12.7 ± 3.2).

Overall response (defined as a reduction in SELENA-SLEDAI of ≥3) was observed in 80 of 113 patients (71%) by the SELENA-SLEDAI assessment prior to rituximab infusion and 6 ± 3 months (mean ± SD) after the last rituximab infusion. Efficacy did not differ significantly between patients receiving rituximab monotherapy and those receiving rituximab and concomitant immunosuppressant agents (who had higher baseline disease activity).

Although this study appears to indicate some benefit for rituximab in a more real-world setting, it is limited in terms of its ability to make a formal comparison with belimumab. It may however suggest that the full clinical benefit of the use of biologics, like rituximab and belimumab is not reflected in a randomised clinical trial setting.

5.7.2 Please follow the instructions specified in sections 5.1 to 5.5 for the identification, selection and methodology of the trials, quality assessment and the presentation of results. Provide in section 9.5, appendix 5, a complete quality assessment for each comparator RCT identified.

N/A

5.7.3 Provide a summary of the trials used to conduct the indirect comparison. A suggested format is presented below. Network diagrams may be an additional valuable form of presentation.

N/A

5.7.4 For the selected trials, provide a summary of the data used in the analysis.

N/A

5.7.5 Please provide a clear description of the indirect/mixed treatment comparison methodology. Supply any programming language in a separate appendix.

N/A

5.7.6 Please present the results of the analysis.

N/A

5.7.7 Please provide the statistical assessment of heterogeneity undertaken. The degree of, and the reasons for, heterogeneity should be explored as fully as possible.

N/A

5.7.8 If there is doubt about the relevance of a particular trial, please present separate sensitivity analyses in which these trials are excluded.

N/A

5.7.9 Please discuss any heterogeneity between results of pairwise comparisons and inconsistencies between the direct and indirect evidence on the technologies.

N/A

5.8 Non-RCT evidence

Non-RCT, both experimental and observational, evidence will be required, not just for those situations in which RCTs are unavailable, but also to supplement information from RCTs when they are available. This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 3.2.8 to 3.2.10.

5.8.1 If non-RCT evidence is considered (see section 5.2.7), please repeat the instructions specified in sections 5.1 to 5.5 for the identification, selection and methodology of the trials, and the presentation of results. For the quality assessments of non-RCTs, use an appropriate and validated quality assessment instrument. Key aspects of quality to be considered can be found in 'Systematic reviews: CRD's guidance for undertaking reviews in health care' (www.york.ac.uk/inst/crd). Exact details of the search strategy used and a complete quality assessment for each trial should be provided in sections 9.6 and 9.7, appendices 6 and 7.

The searches for non-RCTs resulted in 14 hits. None of these citations met the inclusion criteria. Two conference proceedings were identified (Chatham et al. 2010; Merrill et al. 2010b), linked to the same trial (LBSL99). The study was an open-label extension of the Phase 2 trial LBSL02 included in Section 5.2.4. Details of the search strategies are Section 9.6, appendix 6.

Methodology

Table 5.23. Summary of methodology of the non-RCTs

Trial no. (acronym)	**LBSL99**
Location	58 centres (US, 57; Canada, 1)
Design	Open label continuation study for LBSL02
Duration of study	Ongoing
Method of randomisation	Open label
Method of blinding (care provider, patient and outcome assessor)	Open label
Intervention(s) (n =) and comparator(s) (n =)	Standard of care plus belimumab 10 mg/kg (n=296)
Primary outcomes (including scoring methods and timings of assessments)	Adverse events and serious adverse events from study start through 8 weeks after last belimumab dose.
Secondary outcomes (including scoring methods and timings of assessments)	Efficacy endpoints including: • Time-to-flare (SLE Flare Index and BILAG) • SELENA-SLEDAI • PGA • BILAG • Reduction in steroid use Assessed at Day 0, every 8-16 weeks, at the Exit visit, at any unscheduled visits and at the 8- and 24-week post-Exit follow-up
Duration of follow-up	8- and 24-week post-Exit follow-up

Participants

Table 5.24. Eligibility criteria in the non-RCTs

Trial no. (acronym)	Inclusion criteria	Exclusion criteria
LBSL99	Subjects were to be assessed for eligibility at the last visit of the 24-week extension period of protocol LBSL02 (i.e., Day 532). Only subjects who completed LBSL02 through Day 532 were eligible for this study and required an improvement in PGA at the Day 532 or Day 476 visit of the 24-week extension period of LBSL02 compared to their PGA score on the day of 1st belimumab dose (either Day 364 [Week 52] of the treatment period or Day 0 [baseline] in LBSL02).	Subjects were to be excluded if they had an SLE flare (mild/moderate or severe as defined by the SLE Flare Index) during the last 30 days of LBSL02 and through the 1st dose in LBSL99. However, subjects who experienced a mild/moderate flare during this period may have been enrolled if, in the opinion of the investigator, their overall response at Day 532 or Day 476 compared to either Day 364 or Day 0 in LBSL02 clearly outweighed the effect of the flare.
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee		

Baseline characteristics

Table 5.25. Characteristics of participants in the non-RCTs across randomised groups[1]

Trial no. (acronym) Baseline characteristic	Belimumab 10 mg/kg
Trial LBSL99 (n = 296)	**(n = 296)**
Sex	
Female	276 (93.2%)
Race	
White	213 (72.0%)
Asian	5 (1.7%)
Black or African American	65 (22.0%)
American Indian or Alaska Native	6 (2.0%)
Native Hawaiian or Other Pacific Islander	2 (0.7%)
Multiracial[2]	5 (1.7%)
Hispanic or Latino origin	
Yes	53 (17.9%)
Age (years)	
Mean ± SD	42.6 ± 11.5
Range	(20, 75)
Weight (kg)	
Mean ± SD	81.6 ± 21.2
Range	(40.5, 145.8)
BILAG organ domain involvement	
At least 1 A or 2 B	190 (64.2%)

At least 1 A At most 1 B	42 (14.2%) 106 (35.8%)
SELENA-SLEDAI Score Mean ± SD (Min, Max)	 9.2 ± 4.55 (2.0, 30.0)
SLE flare index At least 1 flare Severe flare	 226 (76.4%) 40 (13.5%)
PGA Mean ± SD (Min, Max)	 1.4 ± 0.51 (0.2, 2.6)
[1] Based on LBSL02 baseline data [2] More than 1 race indicated	
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee	

Outcomes

Table 5.26. Primary and secondary outcomes of the non-RCTs

Trial no. (acronym)	Primary outcome(s) and measures	Reliability/validity/ current use in clinical practice	Secondary outcome(s) and measures	Reliability/validity/ current use in clinical practice
LBSL99	Adverse events and serious adverse events from study start through 8 weeks after last belimumab dose.	Adverse events are of interest in all clinical trials and are highly relevant to clinical practice.	• Time-to-flare (SLE Flare Index and BILAG) • SELENA-SLEDAI • PGA • BILAG • Reduction in steroid use	Refer to Section 5.3.5.

Statistical methods

Data from this study were explored using descriptive statistics. Analyses were performed on the population of all subjects who received at least 1 belimumab dose in this continuation study. Because LBSL99 was not placebo-controlled, all data should be interpreted with caution.

Participant flow

A CONSORT flow chart is not available, however, the completion status of patients by protocol study year is provided below (see Table 5.27).

Table 5.27. Completion status in LBSL99, by protocol study year[1]

	Year 1	Year 2	Year 3	Year 4
Number of subjects starting interval	296	264	246	227
Number of subjects discontinued	34 (11.4%)	18 (6.8%)	19 (7.7%)	4 (1.8%)
Subject Request	15 (5.0%)	6 (2.3%)	7 (2.8%)	1 (0.4%)
AE	7 (2.3%)	5 (1.9%)	4 (1.6%)	1 (0.4%)
Disease Progression/Lack of Efficacy	5 (1.7%)	4 (1.5%)	2 (0.8%)	-
Entered 99 but not dosed[2]	2 (0.7%)	-	-	-
Lack of Compliance	2 (0.7%)	2 (0.8%)	-	1 (0.4%)
Other	2 (0.7%)	-	3 (1.2%)	1 (0.4%)
Lost to Follow-up	1 (0.3%)	1 (0.4%)	-	-
Investigator Decision	-	-	3 (1.2%)	-
Ongoing	264 (88.6%)	246 (93.2%)	227 (92.3%)	223 (98.2%)

Each study year = 48 weeks in LBSL99.
[1] Year 5 data are not available.
[2] Subject US016-005 discontinued due to AE. Subject US037-001 discontinued due to Subject Request.

Results

An interim report of data was prepared based on a data cut-off of 6 March 2009, at which time the mean number of doses received was 34 with a range of 1 to 49; 86% of subjects received at least 16 doses of study agent in this study over and above the exposure they had in LBSL02 (20 doses for belimumab subjects in LBSL02 and 6 for placebo subjects in LBSL02).

Sustained improvement in SLE disease activity was evident over 4.5 to 5 years, particularly in autoantibody-positive subjects treated with belimumab and standard of care. The frequency of overall and severe flares as measured by the SLE Flare Index decreased over the 5-year study period with new BILAG 1A or 2B flares stabilising. PGA scores also improved over the 5-year study period.

The incidence of AEs, serious AEs (including infections and malignancies), and laboratory abnormalities generally remained stable or declined over time.

5.9 Adverse events

This section should provide information on the adverse events experienced with the technology in relation to the decision problem. Evidence from comparative RCTs and regulatory summaries is preferred; however, findings from non-comparative trials may sometimes be relevant. For example, post-marketing surveillance data may demonstrate that the technology shows a relative lack of adverse events commonly associated with the comparator, or the occurrence of adverse events is not significantly associated with other treatments.

5.9.1 If any of the main trials are designed primarily to assess safety outcomes (for example, they are powered to detect significant differences between treatments with respect to the incidence of an adverse event), please repeat the instructions specified in sections 5.1 to 5.5 for the identification, selection, methodology and quality of the trials, and the presentation of results. Examples for search strategies for specific adverse effects and/or generic adverse-effect terms and key aspects of quality criteria for adverse-effects data can found in 'Systematic reviews: CRD's guidance for undertaking reviews in health care' (www.york.ac.uk/inst/crd). Exact details of the search strategy used and a complete quality assessment for each trial should be provided in sections 9.8 and 9.9, appendices 8 and 9.

None of the main belimumab trials were powered to specifically address safety outcomes and we did not identify any additional trials designed primarily to detect adverse events in relation to belimumab. Adverse events will be discussed using a primary safety population which includes the 3 randomised, placebo-controlled trials completed to date with belimumab administered IV in patients with SLE (LBSL02, BLISS-52 and BLISS-76, referred to as IV SLE CRD [controlled repeat dose]). In addition, long-term data from the open-label, uncontrolled, Phase 2 continuation trial in patients with SLE (LBSL99) will also be presented.

Safety analysis has been conducted on the subgroup of patients with evidence for serological disease activity (low complement and positive anti-dsDNA) only (52% of the Phase 3 trial population) and this was consistent with the overall Phase 3 population. Therefore, there is no reason to believe that the safety data for the high

disease activity subgroup (34% of the Phase 3 trial population) should be any different from the overall Phase 3 population.

Whilst results discussed in Section 5.5 and 5.8 related solely to the recommended 10 mg/kg dose when used in the Phase 3 studies, discussion of adverse events will include all doses used in Phase 2, Phase 3 and the Phase 2 continuation trial. This is the most robust way to examine adverse event data and is a requirement of the regulatory agencies.

Discussion of adverse events will focus primarily on the 1 and 10 mg/kg belimumab doses and how they compare with placebo in the primary safety population. The 4 mg/kg dose was only studied in Study LBSL02 and had a safety and tolerability profile comparable to the placebo group and other belimumab dose groups.

In the primary safety population, patients had active SLE and belimumab was administered in combination with a wide range of concomitant SLE therapies including corticosteroids, antimalarials (e.g. hydroxychloroquine), and immunosuppressants (e.g. azathioprine, mycophenolate mofetil, methotrexate, cyclophosphamide). Therefore, the safety profile of belimumab has been characterised in combination with standard of care for SLE and the data generated in this program are generalisable to the population for whom belimumab is intended to be prescribed.

Subgroups evaluated in the pooled analyses of safety include sex, age, race, baseline proteinuria level, baseline SELENA-SLEDAI score, geographical regions, baseline prednisone use and baseline immunosuppressant use. Overall, no important subgroup effects on the safety profile of belimumab were observed.

5.9.2 Please provide details of all important adverse events for each intervention group. For each group, give the number with the adverse event, the number in the group and the percentage with the event. Then present the relative risk and risk difference and associated 95% confidence intervals for each adverse event. A suggested format is shown below.

In the 3 placebo-controlled SLE studies, belimumab in combination with standard of care therapies for SLE had an overall safety profile that was similar to placebo plus standard of care with regard to frequency, severity, and types of adverse events (AEs). Over 90% of subjects in each group experienced at least 1 AE (see Table 5.28). Serious adverse events (SAEs) were experienced by 19% and 17% in the 1 mg and 10 mg belimumab dose groups respectively, compared with 16% in the placebo group. Severe adverse events occurred at similar rates across the treatment groups. Across the double-blind treatment periods to date, there have been 14 deaths, 3 (0.4%) in the placebo groups, 5 (0.7%) in the 1 mg/kg belimumab groups and 6 (0.9%) in the 10 mg/kg belimumab groups. Infections were the most frequent event leading to death in all treatment groups.

Table 5.28. Number of subjects with AEs (IV SLE CRD)

	Placebo N = 675	1 mg/kg N = 673	4 mg/kg N = 111	10 mg/kg N = 674
At least 1 AE	624 (92.4%)	626 (93.0%)	107 (96.4%)	625 (92.7%)
At least 1 serious AE	107 (15.9%)	125 (18.6%)	15 (13.5%)	117 (17.4%)
At least 1 severe AE	104 (15.4%)	104 (15.5%)	26 (23.4%)	103 (15.3%)
At least 1 serious and/severe AE	145 (21.5%)	155 (23.0%)	32 (28.8%)	152 (22.6%)
At least 1 AE resulting in study agent discontinuation	48 (7.1%)	42 (6.2%)	4 (3.6%)	45 (6.7%)
Deaths	3 (0.4%)	5 (0.7%)	-	6 (0.9%)

Studies LBSL02, BLISS-52 and BLISS-76. Severe refers to Grade 3 and Grade 4.

By preferred term, the most frequent (> 10% of subjects in placebo, 1 mg/kg or 10 mg/kg belimumab groups) events were headache, upper respiratory tract infection (URTI), arthralgia, nausea, urinary tract infection (UTI), diarrhoea, and fatigue (see Table 5.29). Of these events, only diarrhoea and nausea occurred slightly more frequently in the belimumab groups.

Table 5.29. Most common (> 10% in any treatment group) adverse events by MedDRA preferred term (IV SLE CRD)

Preferred Term	Placebo N = 675	1 mg/kg N = 673	10 mg/kg N = 674
Headache	140 (20.7%)	138 (20.5%)	142 (21.1%)
Upper respiratory tract infection	130 (19.3%)	128 (19.0%)	118 (17.5%)
Arthralgia	112 (16.6%)	100 (14.9%)	109 (16.2%)
Nausea	82 (12.1%)	88 (13.1%)	99 (14.7%)

Urinary tract infection	82 (12.1%)	92 (13.7%)	87 (12.9%)
Diarrhoea	62 (9.2%)	81 (12.0%)	80 (11.9%)
Fatigue	70 (10.4%)	71 (10.5%)	66 (9.8%)

In the long-term open-label extension of the Phase 2 trial (LBSL99), the incidence of AEs, severe AEs, SAEs, including infections, remained stable or declined over time through 5 years of exposure.

Deaths

Fourteen deaths occurred during the double-blind periods of the 3 randomised, placebo-controlled SLE trials: 3 (0.4%) in the placebo group, 5 (0.7%) in the 1 mg/kg group and 6 (0.9%) in the 10 mg/kg group. An additional death due to respiratory arrest was reported more than 3 months after the subject's participation in BLISS-52 (1 mg/kg group). Nine deaths occurred in BLISS-52 compared with 3 in BLISS-76 and 2 in Study LBSL02. There were 4 infection-related deaths, 1 with placebo, 1 with 1 mg/kg belimumab and 2 with 10 mg/kg belimumab, and infection may have contributed to the deaths of 2 additional subjects (1 mg/kg and 10 mg/kg). There were 2 suicides, both in subjects receiving belimumab (1 mg/kg and 10 mg/kg), and 1 cancer-related death in a subject receiving 1 mg/kg belimumab (ovarian, likely pre-existing condition).

The available data do not suggest any clustering in the causes of death nor any causes that are unexpected in an SLE population. Overall, there does not appear to be an increased risk of fatal outcomes associated with belimumab treatment.

Serious Adverse Events

Overall, the percentage of subjects experiencing at least 1 SAE was comparable between the placebo and belimumab groups, ranging from 16-19% (see Table 5.30).

By preferred term, pneumonia, pyrexia, UTI, cholelithiasis and cellulitis were the most common SAEs (see Table 5.30), which generally occurred at similar incidence across treatment groups, with slightly more reports of serious pyrexia in the 10 mg/kg belimumab group, and more reports of cellulitis in the 1 mg/kg group.

Table 5.30. Most frequent (≥ 1% in any treatment group) serious adverse events by preferred term (IV SLE CRD)

	Placebo N = 675	1 mg/kg N = 673	4 mg/kg N = 111	10 mg/kg N = 674
At least 1 serious AE	107 (15.9%)	125 (18.6%)	15 (13.5%)	117 (17.4%)
Pneumonia	10 (1.5%)	7 (1.0%)	1 (0.9%)	6 (0.9%)
Pyrexia	3 (0.4%)	5 (0.7%)	-	9 (1.3%)
Urinary tract infection	4 (0.6%)	7 (1.0%)	1 (0.9%)	5 (0.7%)
Cholelithiasis	4 (0.6%)	5 (0.7%)	2 (1.8%)	2 (0.3%)
Cellulitis	2 (0.3%)	7 (1.0%)	1 (0.9%)	1 (0.1%)

In the long-term uncontrolled SLE experience (LBSL99), SAEs with the highest incidence were cellulitis and transient ischemic attack with the highest rates during the randomised, double blind period (i.e. in the LBSL02 study), and with declining rates over time. The incidence and prevalence of SAEs remained stable or generally declined over time.

Discontinuation Due to Adverse Events

Approximately 6-7% of subjects in the randomised, controlled SLE trials discontinued treatment due to an AE, with similar rates across treatment groups (see Table 5.28). The most frequent AEs leading to discontinuation were renal/urinary disorders, infections, nervous system disorders, and skin and subcutaneous tissue disorders, with rates generally similar across groups. More subjects receiving belimumab discontinued treatment due to infusion/hypersensitivity reactions than subjects receiving placebo (see Table 5.31). Overall, there appears to be no clinically meaningful increase in events leading to discontinuation in subjects receiving belimumab compared with those receiving placebo.

In the long-term uncontrolled Phase 2 SLE experience (LBSL99), the incidence of AEs leading to discontinuation of belimumab was highest in the first year of exposure, and declined thereafter.

Adverse Events of Special Interest

Infusion reactions, including hypersensitivity reactions, have been reported with the administration of therapeutic proteins including monoclonal antibodies. Monoclonal antibodies directed against components of the immune system can also exert

immunomodulatory effects and as such may increase both the risk of infections and the risk of developing malignancies. Given these potential risks, a group of AEs of special interest were pre-specified: infusion-related reactions and hypersensitivity reactions (those occurring on the day of infusion), infections, and malignant neoplasms.

Infusion-Related and Hypersensitivity Reactions

Infusion reactions were defined as any of a predetermined set of preferred terms that occurred on the day of an infusion and resolved within 7 days plus all preferred terms indicative of a hypersensitivity reaction that occurred on the day of infusion irrespective of the resolution date. Hypersensitivity reactions were also analysed separately, including those occurring on the day of an infusion, with a separate analysis of hypersensitivity reactions irrespective of day of onset. A summary of infusion reactions and hypersensitivity reactions occurring on an infusion day is provided in Table 5.31.

Table 5.31. All infusion and hypersensitivity reactions[1] summary (IV SLE CRD)

	Placebo N = 675	1 mg/kg N = 673	4 mg/kg N = 111	10 mg/kg N = 674
At least 1 AE	99 (14.7%)	112 (16.6%)	26 (23.4%)	113 (16.8%)
Hypersensitivity reaction	1 (0.1%)	9 (1.3%)	2 (1.8%)	3 (0.4%)
At least 1 serious AE	3 (0.4%)	6 (0.9%)	-	6 (0.9%)
Hypersensitivity reaction[2]	-	2 (0.3%)	-	2 (0.3%)
At least 1 serious and/or severe[3] AE	4 (0.6%)	8 (1.2%)	-	8 (1.2%)
Hypersensitivity reaction[2]	-	2 (0.3%)	-	2 (0.3%)
At least 1 AE resulting in study agent discontinuation	2 (0.3%)	4 (0.6%)	1 (0.9%)	7 (1.0%)
Hypersensitivity reaction[2]	-	2 (0.3%)	1 (0.9%)	2 (0.3%)

Studies LBSL02, C1056 and C1057.
[1] Occurring on the day of an infusion.
[2] Preferred terms included anaphylactic reaction, angioedema, and drug hypersensitvitiy.
[3] Severe refers to Grade 3 and Grade 4.

The incidence of infusion reactions was similar in the placebo group (15%) and the belimumab groups (17%). Most infusion and hypersensitivity reactions were mild or moderate in severity. Serious infusion reactions occurred in < 1% of subjects, with slightly more reported with belimumab (0.9%) than placebo (0.4%). Infusion

reactions and hypersensitivity reactions occurred most frequently with the first 2 infusions with incidence declining over time.

Overall, the risk of serious infusion or hypersensitivity reactions with belimumab is low, and when reactions do occur they are most often seen with the first or second infusion. It is important to note that pre-medication was not required in the clinical protocols of belimumab, although it was recommended in subjects who had a history of allergies to other exogenously administered proteins, drugs, etc.

Infections

Infections are common sources of morbidity and mortality in patients with autoimmune diseases such as SLE and RA. In addition to common and chronic infections, opportunistic infections are also known to occur in patients with autoimmune diseases. The primary risk factor for infections, in addition to the disease itself, is the use of immunosuppressant agents. Given that the mechanism of action of belimumab in inhibiting BLyS resulting in the reduction in B cells and immunoglobulins may increase susceptibility to infection, special attention was paid to the incidence of infections in the clinical development program.

In the randomised-controlled SLE trials, infections occurred slightly more often in belimumab groups compared with the placebo group, although severe and serious events occurred at similar rates across the placebo and the belimumab groups (see Table 5.32).

The top 5 most frequent infections by preferred term were upper respiratory tract infection (URTI), urinary tract infection (UTI), nasopharyngitis, sinusitis, and bronchitis.There was a slightly higher incidence of nasopharyngitis and bronchitis in the belimumab groups compared with placebo, while rates of sinusitis were slightly lower in the belimumab groups relative to placebo.

Serious infections occurred in 5% to 7% of subjects across the treatment groups. The top 5 most frequent serious infections (≥ 3 subjects in any treatment group) were pneumonia, UTI, cellulitis, bronchitis, and pyelonephritis; these events generally occurred at similar rates between the placebo and the belimumab groups. There was no apparent treatment effect or belimumab dose relationship in the

incidence of the individual serious infections that occurred most frequently. Likewise, the incidence of severe infections was similar across treatment groups (3-4%). The infection-related deaths included a sepsis leading to cardiac arrest in the placebo group, cellulitis leading to sepsis in the 1 mg/kg belimumab group, cutaneous infection leading to sepsis in the 10 mg/kg group, and infectious diarrhoea in the 10 mg/kg group; all of which occurred in BLISS-52. In addition, 2 other deaths were reported (10 mg/kg, Study LBSL02; 1 mg/kg BLISS-52) in which infection may have been a contributing factor.

Table 5.32. All infection adverse event summary (IV SLE CRD)

	Placebo N = 675	1 mg/kg N = 673	4 mg/kg N = 111	10 mg/kg N = 674
At least 1 AE	450 (66.7%)	478 (71.0%)	88 (79.3%)	471 (69.9%)
At least 1 serious AE	35 (5.2%)	46 (6.8%)	7 (6.3%)	35 (5.2%)
At least 1 serious and/or severe[1] AE	45 (6.7%)	49 (7.3%)	9 (8.1%)	40 (5.9%)
At least 1 AE resulting in study agent discontinuation	7 (1.0%)	5 (0.7%)	1 (0.9%)	4 (0.6%)
Deaths	-[2]	1 (0.1%)	-	2 (0.3%)

Studies LBSL02, BLISS-52 and BLISS-76.

[1] Severe refers to Grade 3 and Grade 4.

[2] There was 1 additional infection-related death, cardiac arrest, preceded by sepsis in Study C1057.

There was no apparent treatment effect or belimumab dose relationship in the incidence of infection AEs of special interest (prespecified to include cellulitis, fungal infections, herpes viral infections, sepsis, respiratory tract infections, and opportunistic infections), including severe and serious events, although a slightly higher rate of lower respiratory tract infections (LRTIs) was reported in the belimumab groups compared with placebo (see Table 5.33). The increase in belimumab LRTIs was due to infections other than pneumonia which occurred at a similar incidence across groups, and rates of serious and severe LRTIs were also similar across groups.

There have been no reports of tuberculosis, serious invasive fungal infections, or hepatitis B reactivation among subjects enrolled in belimumab studies.

In the long-term uncontrolled SLE experience, the incidence of infections, including severe and serious infections, remained stable or declined over time.

Table 5.33. Infections of special interest by category (IV SLE CRD)

	Placebo N = 675	1 mg/kg N = 673	4 mg/kg N = 111	10 mg/kg N = 674
Cellulitis	43 (6.4%)	55 (8.2%)	9 (8.1%)	43 (6.4%)
Fungal infections	22 (3.3%)	20 (3.0%)	4 (3.6%)	17 (2.5%)
Herpes viral infections	54 (8.0%)	51 (7.6%)	5 (4.5%)	44 (6.5%)
Sepsis	3 (0.4%)	4 (0.6%)	1 (0.9%)	5 (0.7%)
All respiratory infections[1]	327 (48.4%)	342 (50.8%)	66 (59.5%)	350 (51.9%)
Upper respiratory infections	292 (43.3%)	294 (43.7%)	61 (55.0%)	302 (44.8%)
Lower respiratory infections	58 (8.6%)	76 (11.3%)	13 (11.7%)	81 (12.0%)
Pneumonia	17 (2.5%)	21 (3.1%)	2 (1.8%)	16 (2.4%)
Possible opportunistic infections	-	-	-	2 (0.3%)

Studies LBSL02, BLISS-52 and BLISS-76.

[1] Respiratory tract infections coded to MedDRA high-level term (HLT) Respiratory Tract Infections not elsewhere classified (NEC) are unspecified in terms of location (i.e. lower or upper) and therefore are counted under "All respiratory infections" only.

Overall, based on the totality of the available data, belimumab treatment does not appear to meaningfully increase the risk for infection in subjects receiving a range of concomitant therapies, including corticosteroids and immunosuppressants, even with long-term use. In addition, treatment with belimumab does not appear to significantly impact the ability to maintain a protective immune response to vaccinations received prior to initiation of treatment. These findings are perhaps not unexpected since belimumab was associated with decreases in, but not complete depletion of, B cells, while the memory B cell compartment is preserved. Furthermore, belimumab tends to normalise IgG levels rather than inducing abnormally low values, as evidenced by the fact that IgG levels remained within normal limits at all timepoints in approximately 94% of subjects receiving belimumab. Finally, the steroid-sparing effects associated with belimumab may be beneficial in terms of infection risk.

Malignancies

In total, 9 malignant neoplasms were reported in the controlled portions of the trials. The 5 solid organ malignancies were a stomach carcinoid (placebo, Day 202), breast cancer (1 mg/kg, Day 102), cervical cancer (1 mg/kg, Stage 0 in situ, Day 439), ovarian cancer (1 mg/kg, Day 21) and thyroid neoplasm (1 mg/kg, unknown benign/malignant, Day 378). The thyroid neoplasm was not confirmed to be malignant (was judged not to be serious and no action was taken with study agent).

The ovarian cancer was likely pre-existing and ultimately resulted in the death of the subject. Another subject was diagnosed with breast cancer approximately 2 months after the last dose of study agent (placebo). There were 4 non-melanoma skin neoplasms: 2 basal cell carcinoma and 2 squamous cell carcinoma of skin (1 in the placebo group, 3 in the 10 mg/kg belimumab group). No solid organ neoplasms occurred in the 10 mg/kg group. No hematological neoplasms were reported.

No pattern of malignancies or an increase in any particular type of malignancy was identified in subjects receiving belimumab.

Laboratory Abnormalities

Overall, the incidence of these abnormalities was similar across the placebo, 1 mg/kg and 10 mg/kg belimumab treatment groups. Lymphopenia was the most common laboratory abnormality, with 23%, 26% and 24% of subjects in the placebo, 1 mg/kg and 10 mg/kg groups, respectively, experiencing Grade 3 lymphopenia during the study and 2.8%, 1.8% and 3.0%, respectively, experiencing Grade 4. Grade 3/4 prolonged prothrombin time (PT) was observed at similar rates across the placebo, 1 mg/kg and 10 mg/kg belimumab groups and in most cases was associated with the use of warfarin and related agents. Few subjects had Grade 3 or 4 chemistry abnormalities, with no remarkable differences across treatment groups. In the long-term experience in SLE, there does not appear to be any increase in laboratory abnormalities over time.

Immunogenicity

The rate of immunogenicity with belimumab appears to be low, with < 5% of subjects in the Phase 3 trials having a persistent-positive immune response to belimumab. Of those that were able to be tested in the neutralisation assay, 3 of 12 subjects with a persistent-positive response had neutralising antibodies. In most cases, immunogenicity was not associated with clinically-relevant AEs or any obvious impacts on belimumab exposure.

Of the subjects with a persistent-positive immune response, 4 experienced an infusion reaction, all of which were mild to moderate in severity and non-serious. Two subjects experienced headaches, a 3rd experienced nausea and erythema

(arm), and the 4th, with the first infusion, experienced eyelid oedema, dyspnoea, erythematous rash, and pruritis.

Pregnancy

As of 31 December 2009, 47 pregnancies were reported in the IV SLE studies. The total foetal loss rate in subjects treated with belimumab was 31% (10/32 subjects) which is higher than the background estimated rate in patients with SLE (15-25%) (Andrade et al. 2008; Clowse et al. 2005; Rahman et al. 1998; Yasmeen et al. 2001), but lower than the rate in the placebo group (50%, 3/6 subjects). However, it should be noted that the number of pregnancies, particularly in the placebo group, was small.

5.9.3 Give a brief overview of the safety of the technology in relation to the decision problem.

- Belimumab was generally safe and well-tolerated.
- Against a background of standard SLE therapies, the safety profile of belimumab was favourable. The background therapies encompassed a wide range of standard SLE therapies, including steroids, immunosuppressants, and antimalarials.
- The incidence of infusion reactions was 15% in the placebo group and 17% in the belimumab groups. These were generally mild to moderate and managed with routine treatment. Serious hypersensitivity reactions occurred in 5 patients receiving belimumab, 4 on the first infusion and 1 on the 3rd infusion. All these cases resolved with routine therapy of antihistamine and steroids, and 1 required adrenaline.
- Belimumab treatment was not associated with an increased risk of serious infections.
- Belimumab did not appear to impact the ability to maintain a protective immune response to vaccinations received prior to initiation of treatment.

- The malignancy rate observed in the program is consistent with the background rate reported for patients with SLE.

- The most common cause of death was infection followed by cardiovascular disease, which is reflective of the most common causes of death in the general SLE population. Across the entire belimumab SLE development program the death rate per 100 patient-years was similar for placebo- and belimumab-treated patients, (0.43 and 0.55, respectively).

- There were a limited number of pregnancies observed in these studies. In this small dataset, adverse outcomes in patients receiving belimumab were not increased compared with placebo.

- Long-term safety data support the chronic use of belimumab. The incidence of adverse events was stable or decreased over time in patients who have been treated for at least 4 years.

5.10 Interpretation of clinical evidence

5.10.1 Please provide a statement of principal findings from the clinical evidence highlighting the clinical benefit and harms from the technology.

In two pivotal Phase 3 randomised placebo controlled trials belimumab, when used with standard of care, has demonstrated a favourable benefit/risk profile for the treatment of SLE. Belimumab plus standard of care showed a significant reduction in disease activity compared with standard of care alone, as measured by the SRI composite primary endpoint. This ensures that the clinically meaningful change captured for any organ system is not accompanied by disease worsening in other organ systems or any reduction in the overall well-being of the patient. In addition, belimumab demonstrated a reduction in the risk of severe flare, reduction in overall steroid use and improvements in fatigue.

The safety profile of belimumab plus standard of care as seen in the clinical trial program is favourable. Infusion reactions were slightly higher in the belimumab group than the placebo group, but these were generally mild to moderate and managed with routine treatment. There was no increase in risk of serious infections

and malignancy rate is consistent with the background rate for patients with SLE. Within the clinical trial program, the death rate per 100 patient-years was similar for placebo and belimumab treated patients.

The long-term data to at least four years suggests continued benefit of belimumab treatment with no apparent increase in the rate of adverse events, infections or malignancies over time.

The benefits demonstrated in the clinical trial program are even greater in patients with higher disease activity as seen in the high disease activity subgroup (those with evidence for serological disease activity (low complement and positive anti-dsDNA) and a SELENA-SLEDAI disease activity score ≥10).

5.10.2 Please provide a summary of the strengths and limitations of the clinical-evidence base of the intervention.

Belimumab is the first biologic specifically developed for the treatment of SLE to successfully complete a Phase 3 clinical trial program.

A key strength is the fact that the patients enrolled into the Phase 3 studies were patients with stable current disease activity, which allows reduction in disease activity to be measured if the treatment has efficacy, without the need to induce a response with high dose corticosteroids.

A further strength of the Phase 3 program is the use of a robust primary endpoint. The SRI combines a measure of disease activity (SELENA-SLEDAI score) with a sensitive measure of disease worsening (BILAG score) and the Physician's Global Assessment (to address features not captured in the disease activity scores). Being an SRI responder meant that a patient would have complete resolution of one or more disease manifestations with no worsening in another organ system, no decline in overall health and no need for rescue therapies. Thus in its entirety, the SRI is a robust and clinically meaningful endpoint (Stone 2011). The use of this endpoint is further strengthened by the fact that investigators received intensive testing and training on the use of the measures of disease activity and flare.

Finally, the way in which concomitant medication use was controlled allowed clinical care to continue appropriately, without masking the effects of belimumab treatment.

Key limitations include the fact that, given the nature of the disease and standard of care, the trial designs did not allow a direct comparison to any specific therapy or combination of therapies. The exclusion of patients with severe active lupus nephritis and severe active central nervous system involvement limits the information that can be obtained on use in these groups of patients. The trial duration is not long enough to capture all the potential benefits of belimumab, in particular those regarding the impact of belimumab therapy on organ damage. Until larger numbers of patients are treated with belimumab for longer durations, the incidence of rare, severe, or serious AEs will remain unknown.

5.10.3 Please provide a brief statement of the relevance of the evidence base to the decision problem. Include a discussion of the relevance of the outcomes assessed in clinical trials to the clinical benefits experienced by patients in practice.

The evidence base is relevant to the decision problem. The baseline demographics of the patients recruited in to each of the Phase 3 studies matches the population that NICE have described in the final scope. A high disease activity subgroup has been presented based on our intent to identify patients most likely to benefit from belimumab. The evidence base for this subgroup demonstrates greater efficacy versus the population outlined in the final scope.

We have already discussed the clinical relevance of the composite primary endpoint used in the Phase 3 studies in Section 5.3.5 in terms of the SRI's ability to measure improvement in disease activity, while at the same time accounting for potential effects on other aspects of the disease and on patient well-being. To illustrate the clinical benefits of a 4 point improvement in SELENA-SLEDAI disease activity score experienced by patients in practice, this would, for example, equate to complete resolution of pleurisy and pericarditis (each scoring 2 points) or complete resolution of myositis (scoring 4 points) or complete resolution of arthritis (scoring 4 points).

Fatigue is one of the most prevalent clinical manifestations of SLE and severely affects HRQL (Thumboo et al. 2007; Zonana-Nacach et al. 2000b). It is nearly always a major factor in the life of a patient with SLE, and can be very difficult to treat. For some people with lupus, fatigue is their main symptom and can be debilitating. The pooled data from both studies showed that belimumab 10 mg/kg was associated with significantly improved fatigue scores compared with placebo at Weeks 8, 12, and 52 ($p < 0.05$). In the high disease activity subgroup, belimumab 10 mg/kg was associated with significantly improved fatigue scores compared with placebo at Weeks 8 and 12 ($p < 0.05$) and a clinically important but not statistically significant difference at Week 52.

The use of corticosteroids is associated with considerable long-term adverse effects (infections, osteoporosis, diabetes and cardiovascular disease). In clinical practice, clinicians will try to keep the steroid dose as low as possible in order to avoid side effects. Belimumab demonstrated clinically meaningful steroid-sparing effects in both patients in the pooled total population and the high disease activity subgroup.

SLE is a relapsing-remitting disease. Disease activity is characterised by periods of lower disease activity that are punctuated by severe disease exacerbations or ‘flares’. Flares are thought to be symptomatic of an aggressive increase in disease activity which causes irreversible organ-specific damage that accumulates steadily over time and leads to serious comorbidities in later life (Gladman et al. 2003). Belimumab demonstrated a statistically significantly lower risk for flare and severe flare compared with placebo over 52 Weeks.

5.10.4 Identify any factors that may influence the external validity of study results to patients in routine clinical practice; for example, how the technology was used in the trial, issues relating to the conduct of the trial compared with clinical practice, or the choice of eligible patients. State any criteria that would be used in clinical practice to select patients for whom treatment would be suitable based on the evidence submitted. What proportion of the evidence base is for the dose(s) given in the SPC?

Patients enrolled in the two pivotal Phase 3 studies were aged ≥18 years, met the ACR criteria for SLE, had active disease (SELENA-SLEDAI score ≥6) and had either

positive ANA (titre ≥1:80) or anti-dsDNA antibody (≥30 IU/mL). Additionally, those in the high disease activity subgroup had evidence for serological disease activity (low complement and positive anti-dsDNA) and additionally had a SELENA-SLEDAI disease activity score ≥10.

Whilst measures such as the ACR criteria and SELENA-SLEDAI were designed largely for use in clinical trials, clinicians will be familiar with these measures and will be able to use these to guide their selection of suitable patients in clinical practice. Biomarkers such as ANA, anti-dsDNA and complement are routinely measured in clinical practice.

The efficacy evidence base provided in this submission is entirely reflective of the belimumab 10 mg/kg dose given in the Summary of Product Characteristics. Safety data provided in this submission includes information on other doses of belimumab.

6 Cost effectiveness

6.1 Published cost-effectiveness evaluations

Identification of studies

6.1.1 Describe the strategies used to retrieve relevant cost-effectiveness studies from the published literature and from unpublished data held by the manufacturer or sponsor. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used should be provided. The search strategy used should be provided as in section 9.10, appendix 10.

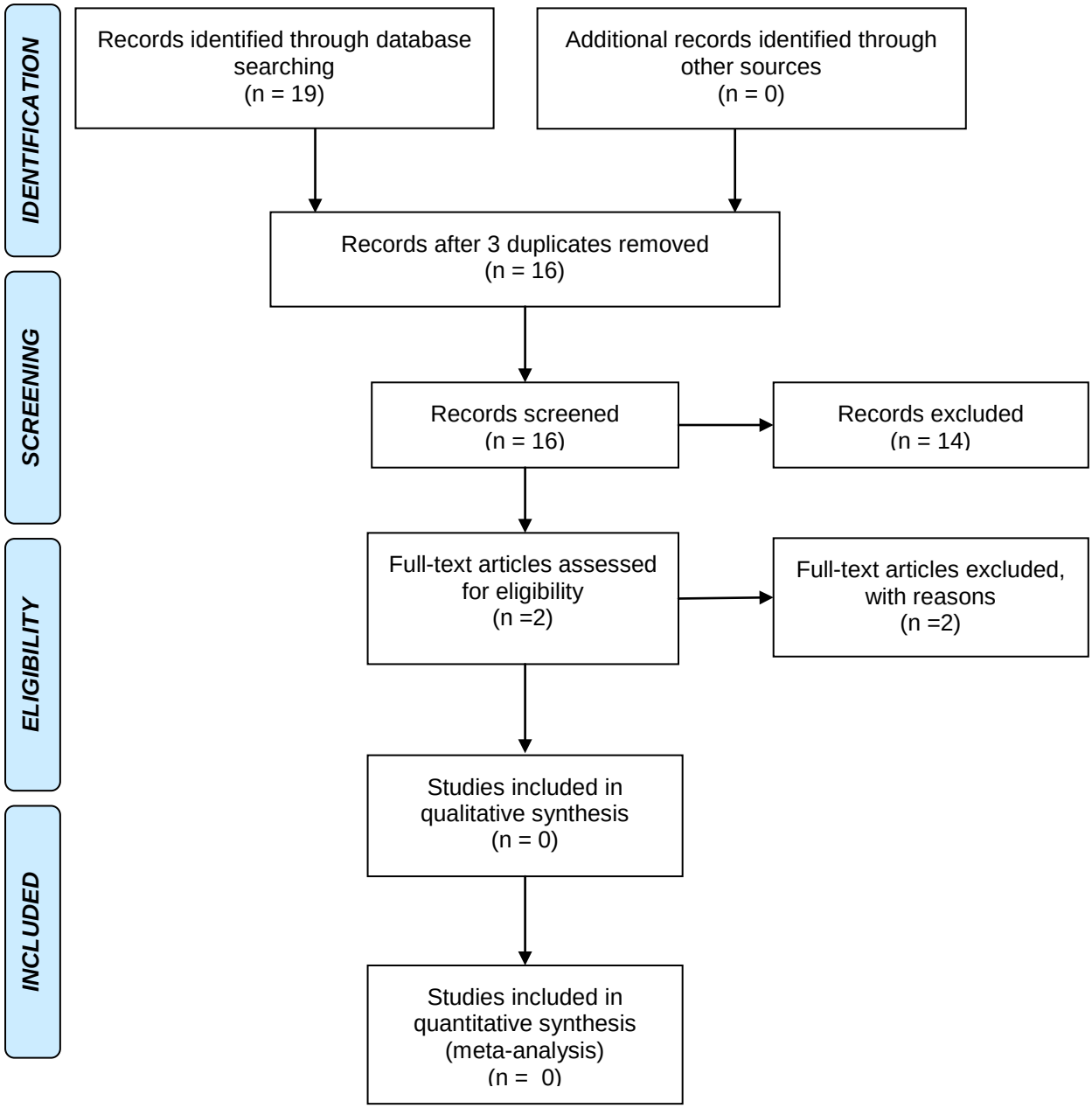
A range of databases indexing published research and other resources were searched for studies reporting the cost-effectiveness of belimumab for the treatment of SLE. The databases searched were MEDLINE, MEDLINE In-Process, EMBASE, HTA database, EconLit and the NHS Economic Evaluation database (NHS EED). In addition, searches were undertaken to identify additional reports in a clinical trials register (ClinicalTrials.gov), the websites of key rheumatology conferences and the websites of the US FDA and EMA. Full details of the databases and resources searched are provided in Section 9.10, Appendix 10.

The search strategy was structured to find records which contained three concepts: SLE, belimumab and economic evaluation. The Centre for Reviews and Dissemination (CRD) NHS EED search sensitive filter was used to find economic studies in MEDLINE and EMBASE. Economic search filters were not applied when searching economic databases such as NHS EED. No date or language limits were applied to the search. Full details of the search strategies are provided in section 9.10, Appendix 10. A flow diagram of the how records were searched and retrieved is provided in Figure 6.1 below.

There were only two full text articles assessed for eligibility. One article was a National Horizon Scanning Centre report (University of Birmingham) for belimumab (National Horizon Scanning Centre 2009). The second was a review article of the efficacy, safety, economic and therapeutic considerations of belimumab for New

Drug Developments (Wiglesworth et al. 2010). Neither of these articles contained any information on the cost-effectiveness of belimumab. The 14 excluded articles are listed in Section 9.10, Appendix 10.

Figure 6.1. Flow diagram of identification of records retrieved



Description of identified studies

6.1.2 Provide a brief overview of each study, stating the aims, methods, results and relevance to decision-making in England and Wales. Each study's results should be interpreted in light of a critical appraisal of its methodology. When studies have been identified and not included, justification for this should be provided. If more than one study is identified, please present in a table as suggested below.

Not appropriate as no relevant cost-effectiveness studies were identified.

6.1.3 Please provide a complete quality assessment for each cost-effectiveness study identified. Use an appropriate and validated instrument, such as those of Drummond and Jefferson (1996)[1] or Philips et al. (2004)[2]. For a suggested format based on Drummond and Jefferson (1996), please see section 9.11, appendix 11.

Not appropriate as no relevant cost-effectiveness studies were identified.

[1] Drummond MF, Jefferson TO (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. British Medical Journal 313 (7052): 275–83.

[2] Philips Z, Ginnelly L, Sculpher M, et al. (2004) Quality assessment in decision-analytic models: a suggested checklist (Appendix 3). In: Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technology Assessment 8: 36.

6.2 De novo analysis

Patients

6.2.1 What patient group(s) is (are) included in the economic evaluation? Do they reflect the licensed indication/CE marking or the population from the trials in sections 1.4 and 5.3.3, respectively? If not, how and why are there differences? What are the implications of this for the relevance of the evidence base to the specification of the decision problem? For example, the population in the economic model is more restrictive than that described in the (draft) SPC/IFU and included in the trials.

The base case for the economic evaluation comprises the total pooled SLE patient population recruited into the two Phase 3 clinical trials: BLISS-56 and BLISS-76, excluding he belimumab 1mg/kg treatment arm. This enables cost-effectiveness to be assessed consistent with the analysis and presentation of primary and key secondary endpoints in the study populations presented in Section 5. As discussed in the Clinical Section 5.5, the pooling of the trial data is considered appropriate given that the trials were essentially identical in design and in the analysis of the primary endpoint and its three separate components there were no evidence of a treatment-by-study interaction. Pooling the studies increased the sample size and provided more power for the statistical analyses. The proposed population of interest to this decision problem is a subgroup of the Phase 3 trial population which applies the additional criteria of evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI disease activity score of ≥10. The identification of this subgroup is discussed in Section 5.3.7; although it was not specifically pre-specified in the original data analysis plan, each of the individual components for patient selection (low complement, positive anti-dsDNA and SS score ≥10) were pre-specified for subgroup analysis. The aim of including the additional SS score criterion is to try and identify patients with the highest disease activity and who are considered most likely to experience the worst long-term morbidity from the disease; in the BLISS trials this subgroup of patients also demonstrated a greater benefit from belimumab. The results from the cost-effectiveness analysis for this subgroup population are presented in Section 6.9.

Model structure

6.2.2 Please provide a diagrammatical representation of the model you have chosen.

To be able to accurately reflect the heterogeneity and complexity of SLE, a micro-simulation model was built. The structure of the model was based on the result of a literature review and input from several clinical experts. The structure becomes complicated since SLE encompasses a range of variables that are correlated. The interaction between patient characteristics, disease activity, medication (corticosteroid use), risk of organ damage development and mortality, would make a cohort Markov model unsuitable due to the large number of health states required. To illustrate this complexity it is necessary to consider organ damage. An SLE patient could potentially develop damage in 12 different organs. To account for all combinations possible 2^{12} = 4096 distinct health states for organ damage alone would be required. Adding additional SLE features necessary to distinguish patients and their risk of future events would make the number of health states in a cohort Markov model too large to handle. Another reason for not choosing a Markov model is due to its lack of "memory". This is of major importance since average disease score over time is a major predictor of events. Due to this need for patient history, a micro-simulation was felt to be the most suitable model type.

The micro-simulation model simulates individual patients over a lifelong period. The patient population entering the model reflects 1) the pooled total population of the two RCTs: BLISS-52 and BLISS-76, and 2) a subgroup of this pooled population consistent with our target population for this decision problem. As both BLISS trials were of either 52 or 76 weeks duration, the effect of treatment on long-term disease outcomes could not be determined. Long-term outcomes, however, have a major effect on health-economic assessment, and as such, these outcomes were considered important to be included in a model that estimates the cost-effectiveness of belimumab treatment.

As discussed in Section 5.3.5, the primary composite outcome of the BLISS trials included SELENA-SLEDAI (SS) score, a measure of disease activity, which is linked to long-term outcomes (Swaak et al. 1999). High disease activity over time will

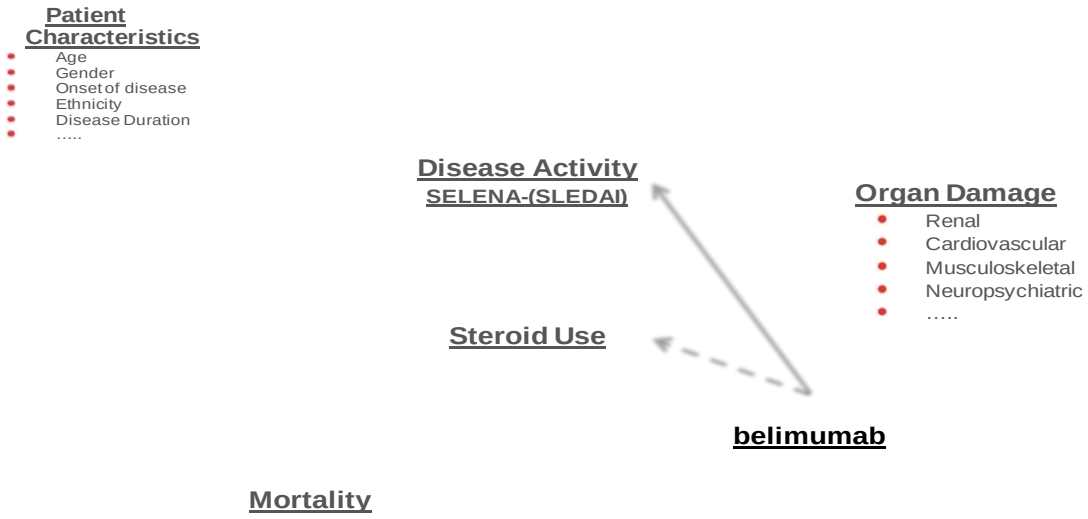
accrue organ damage. Therefore in order to address this in the model, the relationship between short-term and long-term outcomes was estimated based on a Lupus registry, the Johns Hopkins cohort (Kasitanon et al. 2006).

A literature review was conducted to identify all longitudinal databases of SLE patients. Section 9.20, Appendix 20 summarises all these databases. Of these cohorts only the Johns Hopkins (Kasitanon et al. 2006), Toronto (Ibanez et al. 2003; Ibanez et al. 2005), University College Hospital (UCH) (Stoll et al. 2004) and Tromso databases (Becker-Merok et al. 2006) had detailed information on disease activity, organ damage and mortality. The Tromso was not followed up as it was a small cohort. Analysis of the Toronto cohort was conducted only by the custodians of the database, no data was supplied to GSK, and so this limited the level of analysis that could be conducted. This database is therefore used to validate the modelling where sufficient results were available. The UCL cohort did not have any data on SELENA-SLEDAI score and so this database could not be used to link short-term with long-term outcomes. Only the Johns Hopkins database was available in full, was large in size, and contained all the data required to conduct the appropriate level of analysis and so this was used as the main database to use for the long-term modelling.

Based on the Johns Hopkins data, time to event (TTE) models were estimated that describe the relation between disease activity and other covariates on the risk of dying and on the risk of developing irreversible organ damage. Section 9.21, Appendix 21 contains the detailed report of these analyses, but a summary of the methods used is detailed in Section 6.3. The TTE models are implemented in the health-economic model to simulate a patient's future disease course based on the severity of the population and the short-term outcomes observed in the BLISS trials. Health-economic consequences (quality of life impairment and health-care costs) are assigned to each long-term outcome to translate clinical outcomes to health-economic outcomes. Together with the short-term health economic consequences this allows the cost-effectiveness of belimumab in addition to standard of care (SoC), hereafter referred to as belimumab, relative to SoC alone, to be assessed over a life-long period.

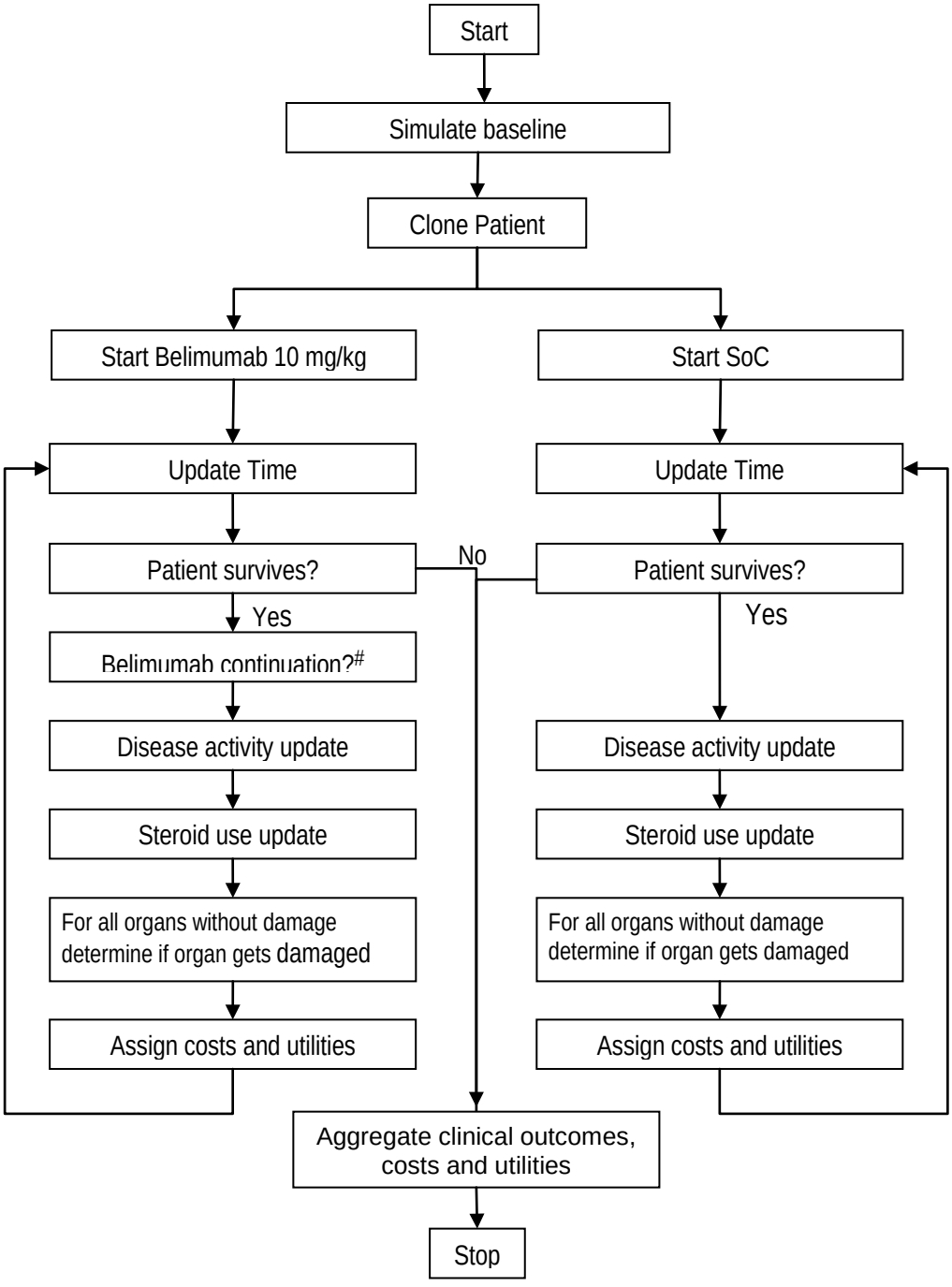
The interdependencies between SLE patient characteristics, SLE related outcomes and steroid treatment which are reflected in the simulation are presented in Figure 6.2 below.

Figure 6.2. Schematic overview of interdependencies between baseline characteristics, treatment and outcomes in the micro-simulation model



After simulating a patient's baseline characteristics they enter the model in which their remaining lifetime SLE history is simulated. A patient is "cloned" based on their baseline characteristics and enters both the belimumab and SoC treatment arms. Each year a patient's situation is re-established. The order in which this occurs for both treatment arms is demonstrated in the diagram in Figure 6.3. The patient's flow through the simulation model is outlined below. The methodology behind this simulation process is described in detail in Section 6.3.

Figure 6.3. Patient flow through the micro-simulation model



#If inadequate response to belimumab, the patient switches to SoC and continues through the model's yearly cycles on SoC until death.

Each year, it is first determined whether the patient survives. A patient who dies is assumed to die in the middle of the year to account for a continuity correction. For a "survivor" patient on belimumab, it is then established whether this patient continues belimumab medication. Belimumab treatment can be stopped due to three reasons:

1. **Natural discontinuation**. This reflects the natural pattern due to patient request, lack of efficacy, lack of compliance or an adverse event based on the percentages observed in the BLISS trials after one year of treatment.
2. **Insufficient response after the first six months**. This provides a choice to only continue belimumab treatment for individuals with a sufficient response to treatment assessed after six months of treatment. Natural discontinuation may still occur for these patients, but this probability is different, due to their "responder" status and are also based on the BLISS trial data.
3. **Maximum treatment duration reached**. The treatment duration for belimumab can be limited in the model. However this option is not considered appropriate for this decision problem, the reasons for which will be discussed in Section 6.3.

Once a patient's treatment for the current year is known, their average disease activity and steroid use is updated accordingly. For each organ system contained within the SLICC Damage Index (SDI) (see Section 9.18, Appendix 18), which is not damaged so far for that particular patient, the probability of damage during that year is calculated based on the patient's characteristics and disease activity at that time. A Bernoulli distribution is applied to simulate whether the patient develops damage in each specific undamaged organ system and the average SDI for that organ, based on Johns Hopkin's data is applied. Average costs and utilities calculated from regression analyses (detailed later) are assigned to a patient's "health state" for that particular year. Costs and utilities are then recorded together with clinical outcomes for that patient. In place of the drug costs, the model records belimumab usage in order to allow the user to change the belimumab price after the simulation if necessary. Time is then increased by one year and the process is repeated. These yearly cycles continue until a patient dies.

To obtain average outcomes as free of sampling errors as possible a sufficient amount of patients need to be simulated. The procedure to establish model convergence is detailed in Section 6.3, which details how the clinical data were implemented in the model and in Section 6.7, which refers to the validation of the model.

6.2.3 Please justify the chosen structure in line with the clinical pathway of care identified in section 2.4.

The clinical picture of SLE can be very complex, with multiple simultaneous manifestations and characterised by periods of relapse and episodes of remission. One of the main goals of SLE treatment is decreasing disease activity, which is reversible, and preventing irreversible organ damage. As discussed in Section 2.4 there are currently no clinical guidelines for SLE. Given the diversity of clinical manifestations, the clinical pathway of care for SLE varies according to the individual and disease severity. Therefore to be able to accurately reflect the heterogeneity and complexity of SLE, a micro-simulation model was considered to be the most appropriate structure; a Markov model would be too complex with the large number of required health states for organ damage. The micro-simulation model structure demonstrates how the disease progresses over time using a real-life longitudinal database to map the course of the disease and takes into consideration how patients with the severity of disease of interest to this decision problem are currently treated (SoC) in the UK and how some of these treatments, namely corticosteroids, can also impact on the occurrence of organ damage.

6.2.4 Please define what the health states in the model are meant to capture.

As the model is not a Markov model, disease activity, accrual of organ damage and mortality, rather than distinct health states, are discussed in Section 6.2.2

6.2.5 How does the model structure capture the main aspects of the condition for patients and clinicians as identified in section 2 (Context)? What was the underlying disease progression implemented in the model? Or what treatment was assumed to reflect underlying disease progression? Please cross-reference to section 2.1.

This has been discussed in Section 6.2.2.

6.2.6 Please provide a table containing the following information and any additional features of the model not previously reported. A suggested format is presented below.

Table 6.1. Key features of analysis

Factor	Chosen values	Justification
Time horizon	Lifetime	The economic evaluation estimates costs and health benefits over the full lifetime of each individual. This time horizon is necessary for the main health outcomes and resource use to be fully explored in this chronic disease.
Cycle length	Yearly	SLE is a long-term chronic disease. The changes in overall disease activity and the accumulation of organ damage are believed to be adequately captured with a yearly cycle over a lifetime horizon. However, if long-term data on the incidence and severity of flares had been available, a shorter cycle length may have been more appropriate to capture the pattern of flares over time.
Half-cycle correction	Not included	Not applicable
Were health effects measured in QALYs; if not, what was used?	Yes, health effects were measured in QALYs	This is consistent with the reference case
Discount of 3.5% for utilities and costs	Yes, discounting of 3.5% was used for both utilities and costs	This is consistent with the reference case
Perspective (NHS/PSS)	An NHS and PSS perspective was used	This is consistent with the reference case
NHS, National Health Service; PSS, Personal Social Services; QALYs, quality-adjusted life years		

Technology

6.2.7 Are the intervention and comparator(s) implemented in the model as per their marketing authorisations/CE marking and doses as stated in sections 1.3 and 1.5? If not, how and why are there differences? What are the implications of this for the relevance of the evidence base to the specified decision problem?

The intervention, belimumab, implemented in the model has the same dosing schedule (10 mg/kg) that will be used if the licence application is successful.

Belimumab is administered in addition to usual SoC treatment in the model and this represents how it is proposed that it will be prescribed in UK clinical practice.

The comparator in the model is usual standard of care alone, which as described in section 2.0, consists of the following drugs (alone or in combination): antimalarials, NSAIDs, corticosteroids or other immunosuppressants (azathioprine, methotrexate, and mycophenolate mofetil). Most of these drugs are not licensed specifically for treating SLE however they are frequently used to treat the disease in UK clinical practice.

Both rituximab and cyclophosphamide were identified as potential comparators to be considered as part of this decision problem.

The inclusion criteria of the published Phase 2/3 randomised, double-blind study of rituximab (two 1,000 mg IV doses given 14 days apart), required SLE patients to have significantly active disease at screening, defined as ≥ 1 organ system with a BILAG A score or ≥ 2 organ systems with a BILAG B score despite background immunosupressants at study entry (Merrill et al. 2010a). This is likely to correspond to a more severe patient population compared with both the pooled total population and the high disease activity subgroup, where 49% and 53% of patients respectively, were receiving an immunosuppressant at baseline and 61% and 72% respectively, had ≥ 1 organ system with a BILAG A score or ≥ 2 organ systems with a BILAG B score.

The study did not collect data on changes in SELENA-SLEDAI, which as discussed previously is an important short-term outcome which can be linked to longer term impact on organ damage. So both the trial populations and the outcomes reported are different for the rituximab and Phase 3 belimumab trials, making any indirect comparisons of these technologies using this study problematic.

The primary endpoint was the effect of rituximab versus placebo in achieving and maintaining a major clinical response, a partial clinical response, or no clinical response at week 52 assessed using BILAG scores.

A major clinical response was defined as achieving BILAG C scores or better in all organs at week 24 without experiencing a severe flare (1 new domain with a BILAG

A score or 2 new domains with a BILAG B score) from day 1 to week 24, and maintaining this response without a moderate or severe flare (≥1 new domains with a BILAG A or B score) to week 52.

A partial clinical response was defined as 1) achieving BILAG C scores or better at week 24 and maintaining this response without a new BILAG A or B score for 16 consecutive weeks, 2) achieving no more than 1 organ with a BILAG B score at week 24 without achieving ≥1 new BILAG A or B score to week 52, or 3) achieving a maximum of 2 BILAG B scores at week 24 without developing BILAG A or B scores in new domains until week 52 if the baseline BILAG score for the patient was 1 A score plus ≥2 B scores, ≥2 A scores, or ≥4 B scores.

No clinical response was defined as failure to meet the definition of a major clinical response or a partial clinical response. Patients who terminated the study early were scored as having no clinical response.

At week 52, no difference was noted in major clinical responses or partial clinical responses between the rituximab group (12.4% had a major clinical response, and 17.2% had a partial clinical response) and the placebo group (15.9% had a major clinical response, and 12.5% had a partial clinical response) relative to the overall response rate (29.6% versus 28.4%).

In addition, the rituximab trial demonstrated no difference in secondary endpoints between the rituximab group and the placebo group and over 52 weeks of treatment, in patients with moderate-to-severe SLE. Secondary end points are outlined below.

Secondary end points included:

1) the time-adjusted area under the curve minus baseline (AUCMB) of the BILAG score over 52 weeks

2) the proportion of patients who achieved a major clinical response (excluding a partial clinical response) and the proportion of patients with a partial clinical response (including a major clinical response) at week 52

3) the proportion of patients with a BILAG C score or better in all organs at week 24

4) the time to the first moderate or severe disease flare

5) improvement in quality of life as measured by the Lupus Quality of Life index

6) the proportion of patients who achieved a major clinical response with a prednisone dosage of <10 mg/day from week 24 to week 52.

The efficacy and safety of rituximab has also been investigated as part of an analysis of prospective data from the French AutoImmunity and Rituximab (AIR) registry (Terrier et al. 2010). One hundred thirty-six patients received treatment for SLE. The mean SELENA-SLEDAI score at baseline was 11.3 ± 8.9 which indicates a slightly less severe population than our high disease activity subgroup (12.7 ± 3.2).

Overall response (defined as a reduction in SELENA-SLEDAI of ≥3) was observed in 80 of 113 patients (71%) by the SELENA-SLEDAI assessment prior to rituximab infusion and 6 ± 3 months (mean ± SD) after the last rituximab infusion. Efficacy did not differ significantly between patients receiving rituximab monotherapy and those receiving rituximab and concomitant immunosuppressant agents (who had higher baseline disease activity).

Although this study appears to indicate some benefit for rituximab in a more real-world setting, it is limited in terms of its ability to make a formal comparison with belimumab. It may however suggest that the full clinical benefit of the use of biologics, like rituximab and belimumab is not reflected in a randomised clinical trial setting.

Cyclophosphamide, whilst used in the more severe patient population, is largely reserved for the treatment of lupus nephritis. This is not the proposed target population for belimumab, therefore, cyclophosphamide plus standard therapy is not a relevant comparator. In addition, adverse effects associated with long-term exposure to cyclophosphamide including bladder cancer, bone marrow suppression, haematologic malignancies, infections, myelodysplasia, and infertility (Kalunian et al. 2009), limit the appropriateness of cyclophosphamide given that a high proportion of patients are women of childbearing age.

6.2.8 Please note that the following question refers to clinical continuation rules and not patient access schemes. Has a treatment continuation rule been assumed? If the rule is not stated in the (draft) SPC/IFU, this should be presented as a separate scenario by considering it as an additional treatment strategy alongside the base-case interventions and comparators. Consideration should be given to the following.

- The costs and health consequences of factors as a result of implementing the continuation rule (for example, any additional monitoring required).
- The robustness and plausibility of the endpoint on which the rule is based.
- Whether the 'response' criteria defined in the rule can be reasonably achieved.
- The appropriateness and robustness of the time at which response is measured.
- Whether the rule can be incorporated into routine clinical practice.
- Whether the rule is likely to predict those patients for whom the technology is particularly cost effective.
- Issues with respect to withdrawal of treatment from non-responders and other equity considerations.

The current draft SPC for belimumab states in the "Posology and method of administration" section that *"The patient's condition should be evaluated continuously. Discontinuation of treatment with Benlysta should be considered if there is no improvement in disease control after 6 months of treatment."* This allows the assessment of adequate response to belimumab to be made on the basis of the physicians' clinical judgement after six months of treatment. Six months is identified as a suitable time period after which to assess response to treatment as it allows sufficient time for the drug's mode of action to have an impact on the clinical manifestations of the disease. Standardised disease activity/damage measures are recognised in clinical practice (e.g. SELENA-SLEDAI (SS), BILAG), however they are not currently used routinely to determine treatment outcomes. Generally, physicians will assess response based on the general well-being of the patient,

including frequency and severity of disease flares. However, for the health economic model a more objective assessment was required to determine whether belimumab should be continued or discontinued after six months treatment. The criterion for treatment continuation was based on a patient demonstrating a SS score decrease of 4 or greater. As well as being a pre-specified outcome of the BLISS trials, it is a validated, robust measure and relates to a clinically meaningful change in disease activity (Swaak et al. 1999). SS score is one of the three criteria included in the primary composite endpoint in the BLISS trials, the SLE Responder Index (SRI) and is the measure of efficacy. The SS score is used in the model rather than the SRI in order to be consistent with published evidence showing how SS score, representing disease activity, is related to long-term disease outcomes (Swaak et al. 1999); there is currently no evidence to show this relationship with SRI. As assessment of response will be carried out by physicians after the first six months of belimumab treatment, this continuation rule at six months has been included in the base case analysis. The assessment of cost-effectiveness excluding this responder rule and the inclusion of an alternative, more stringent responder rule, have been investigated as scenario analyses.

No additional monitoring is required specifically related to this continuation rule, as according to clinical experts, patients with high disease activity are likely to be seen at least every six months by their treating physician in routine clinical practice. If patients do not show an adequate response to belimumab after the first six months of treatment, there is the option for them to return to how they were managed previously, defined as the current standard of care. They should not experience any problems from the withdrawal of belimumab and there are no identified equity issues associated with withdrawal.

6.3 Clinical parameters and variables

When relevant, answers to the following questions should be derived from, and be consistent with, the clinical-evidence section of the submission (section 5). Cross-references should be provided. If alternative sources of evidence have been used, the method of identification, selection and synthesis should be provided as well as a justification for the approach.

6.3.1 Please demonstrate how the clinical data were implemented into the model.

Patient population

As discussed in Section 6.2.1 above, two patient populations are considered relevant to this decision problem: i) the pooled total BLISS study population hereafter labelled as “Pooled total population” and ii) a subgroup of this population identified as those patients with the highest disease activity due to having SELENA-SLEDAI (SS) scores of 10 or more at baseline (discussed in Section 5.3.7), hereafter referred to as the “high disease activity” subgroup. In order to be consistent with the key clinical evidence reported in the two Phase 3 BLISS studies, the results for the pooled total population is presented first, as the base case analysis, and the subgroup, although considered to be the most relevant population for this decision problem, is presented in Section 6.9 (Subgroups).

The primary efficacy endpoint in the two Phase 3 BLISS studies was the response rate at week 52, assessed by the SLE Responder Index (SRI). It includes an objective measure of the reduction in global disease activity (reduction in SELENA-SLEDAI score ≥ 4) for efficacy and 2 measures to ensure that the improvement in disease activity (score) is not offset by worsening of the subject’s condition overall (i.e. no worsening in the PGA) or worsening in any specific organ system (i.e. no new BILAG A or 2 new B flares).

Of the three criteria included in the primary composite endpoint SS is the measure of efficacy. The SS score is used in the model rather than the SRI in order to be consistent with published evidence showing how SS score, representing disease

activity, is related to long-term disease outcomes (Swaak et al. 1999); there is currently no evidence to show this relationship with SRI.

Key measures of SLE for disease activity and organ damage

1. Assessment of Disease Activity: The SELENA-SLEDAI

The main measurement of disease activity in the BLISS trials and in the long-term observational databases, discussed later, is the SELENA-SLEDAI score.
The Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) is a global index of SLE disease activity during the previous 10 days (See Section 9.14, Appendix 14). It includes 24 weighted variables, consisting of objective clinical and laboratory variables. Items are only marked if they are attributable to SLE. The disease activity is the sum of the marked items and ranges from 0 to 105, in which 0 is no activity and a theoretical maximum activity score of 105. The SLEDAI has shown to be valid, reliable and sensitive to change in disease activity(Gladman et al. 2000). Since the publication of the original SLEDAI several modifications of the SLEDAI have been made to how items are defined. The SLEDAI was adapted in the Safety of Estrogen in Lupus Erythematosus National Assessment (SELENA) trial in order to capture worsening disease activity. The SELENA-SLEDAI modifications excludes seizures that are caused by neuropsychiatric damage, visual disturbances were expanded to include scleritis and episcleritis, and the cranial nerve descriptor was expanded to include vertigo, the cerebrovascular accident excludes hypertensive causes, and other changes were made to pleurisy and pericarditis.
The main limitation of the original SLEDAI is that patients with persistent activity can have a decrease in SLEDAI score without an improvement in symptoms. This would suggest that the SLEDAI may underestimate improvements in disease activity, but this has not been tested in an observational cohort. It was not feasible to adjust the SLEDAI to better match the SELENA-SLEDAI. Both versions of the SLEDAI maintain the same 24 items and weighting system so they are very similar; however the classification of events is slightly different.

2. Systemic Lupus International Collaborating Clinics Damage Index (SDI or SLICC)

The SLICC/SDI measure (See Section 9.18, Appendix 18) was developed by a group of physicians specialising in SLE and endorsed by the American College of Rheumatology. The index contains 41 damage items in 12 systems that are specific comorbidities associated with SLE or damage due to toxicity of SLE treatment. Contrary to the assessment of disease activity, damage items are recorded irrespective of their attribution to SLE, however, damage items have to persist for a minimum of six months, or be associated with an immediate pathological scar indicative of damage. The total score is the sum of the marked scores and ranges from 0 to 47. Since damage is irreversible, items that are marked will stay marked for the lifetime of the patient (Gladman et al. 1997). The SLICC was expanded in 1996 with pancreatic insufficiency and ruptured tendon, and pulmonary resection was added as an alternative to pulmonary infarction (Gladman et al. 1997).

Baseline Characteristics of the study population

The baseline characteristics for the total population are presented in Tables 6.2, 6.3, and 6.4 below. The distribution and corresponding parameters used to simulate each characteristic are included.

It is important to note that despite the potential correlation between baseline characteristics, they are sampled independently. Bootstrapping (i.e. sampling from the trial data) was considered, but would underestimate the actual heterogeneity when simulating 50,000 patients. Due to the quantity of baseline characteristics and the different types of distributions, it was considered too complex to apply correlated sampling to the model. This is a limitation, but it is expected that this will not greatly influence the average results.

A multinomial distribution is used to generate an individual's age, as it is a discrete distribution and best reflects the age distribution in the trials (Johnson et al. 2002). The percentages of each age simulated will be equal to the percentages of each age observed in the trials. Gender and black ethnicity are binary variables that are simulated using a Bernoulli distribution. SLE disease duration is simulated from a

Geometric distribution with parameter *p* equal to the inverse of the mean of 6.4 (Wackerly et al. 1996). The variance of this distribution is 34.6, which is approximately equal to the variance observed in the trial. Instead of simulating a patient's total SDI score, the scores simulated for each individual item are summed to determine the total SDI score. These data are summarised in Table 6.2.

Table 6.2. Baseline patient demographics - Pooled total population

Patient demographics	Mean (sd) or %	Distribution	Parameter
Age (yrs)	37.9 (11.6)	Multinomial	Probability for each age
Gender (% females)	94.3%	Bernoulli	0.943
Black Ethnicity (%)	8.7%	Bernoulli	0.087
SLE Disease duration (yrs)	6.4 (6.5)	Geometric	0.157
SLICC Damage Index score (SDI)*	0.76 (1.23)	Multinomial*	NA

**Note that Instead of simulating a patient's total SDI score, the scores simulated for each individual item presented in Table 4 are summed to determine the total SDI.*

A multinomial distribution was used to simulate a patient's baseline SS score (Table 6.3) as this accurately reflects the (discrete) possible scores and probability for each score (Johnson et al. 2002). Besides the SELENA-SLEDAI score itself, the involvement of certain individual items in the SELENA-SLEDAI are simulated as baseline characteristics in the model. In estimating long-term SLE outcomes, certain individual SS items (or groups of items) have better (or additional) explanatory power than the SS score and this is the reason for initially including these parameters in the simulation. Daily steroid dose is simulated using a Gamma distribution that reflects the mean and standard deviation observed in the pooled BLISS trials at baseline (Wackerly et al. 1996).

Table 6.3. Baseline disease activity parameters based on BLISS trials – Pooled total population

SLE disease activity parameters	Mean (sd)	Distribution	Parameter
SELENA-SLEDAI score	9.74 (3.78)	Multinomial	Probability for each score
Increased DNA binding	69.0%	Bernoulli	0.690
Low Complement	62.9%	Bernoulli	0.629
Vasculitis	6.7%	Bernoulli	0.067
Neuropsychiatric involvement	2.1%	Bernoulli	0.021
Renal involvement	15.6%	Bernoulli	0.156
Serositis involvement	6.1%	Bernoulli	0.061
Haematological Involvement	7.3%	Bernoulli	0.073
Skin Involvement	82.0%	Bernoulli	0.820
Steroid dose (mg) at first visit (mean (SD))	10.8 (8.8)	Gamma	1.50; 7.17#

values for shape and scale for the Gamma distribution respectively

Instead of simulating a patient's total SLICC Damage Index (SDI) score, the scores simulated for each individual item presented in Table 6.4 are summed to determine the total SDI score. An individual organ damage item score was drawn from a multinomial distribution with each category having the probability as outlined in Table 6.4, which reflects the baseline SDI item occurrences observed in the pooled BLISS trials. The multinomial distribution accurately describes the (discrete) possible scores for each organ (Johnson et al. 2002). By separating out the individual items of the SDI, the subsequent long-term impact on that organ can be estimated through the damage observed in the Johns Hopkins database.

Table 6.4: Baseline individual SDI item scores simulated from the BLISS trials – Pooled total population

SLICC damage item	Score 0	Score 1	Score 2	Score 3	Score 4	Distribution
Cardiovascular	94.6%	4.8%	0.5%	0.1%	0.0%	Multinomial
Diabetes	97.1%	2.9%	0.0%	0.0%	0.0%	Multinomial
Gastrointestinal	95.4%	4.5%	0.1%	0.0%	0.0%	Multinomial
Malignancy	99.3%	0.7%	0.0%	0.0%	0.0%	Multinomial
Musculoskeletal	85.6%	10.0%	3.6%	0.6%	0.2%	Multinomial
Neuropsychiatric	87.9%	9.5%	2.3%	0.3%	0.0%	Multinomial
Ocular	93.3%	6.6%	0.1%	0.0%	0.0%	Multinomial
Peripheral vascular	95.7%	3.7%	0.3%	0.3%	0.0%	Multinomial
Premature gonadal failure	98.1%	1.9%	0.0%	0.0%	0.0%	Multinomial
Pulmonary	97.3%	2.4%	0.3%	0.0%	0.0%	Multinomial
Renal	97.8%	2.2%	0.0%	0.0%	0.0%	Multinomial
Skin	92.8%	6.7%	0.4%	0.2%	0.0%	Multinomial

After simulating a patient's baseline characteristics they enter the model in which their remaining lifetime SLE history is simulated.

Year one treatment effects

In the first year of the simulation, the effects on disease activity as observed in the pooled BLISS trials are applied. These can be divided into an effect on total SS score and an effect on the involvement of certain items in the SS score.

Effect on SELENA-SLEDAI (SS) score

The primary endpoint in the BLISS-52 and BLISS-76 trials was response in SLE Responder Index (SRI) at week 52, with response defined as:

i) a ≥ 4 point reduction from baseline in SELENA-SLEDAI (SS) score and
ii) no new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline. and
iii) no worsening (increase of < 0.30 points from baseline) in Physician's Global Assessment (PGA)

As discussed in the Section 5.5 of this document, belimumab 10 mg/kg demonstrated superiority to SoC for this endpoint in both the BLISS-52 and BLISS-76 trials. In this composite endpoint, SS score is the measure of efficacy in terms of

disease activity reduction whilst both BILAG and PGA are measured to ensure any observed improvement in SS score is not reported as a response if accompanied by a worsening of the disease in another organ system or in the general well-being of the patient. As stated previously, since there is no long-term cohort data in which all the three measures of the composite endpoint were recorded, determining the long-term effects of the SRI was not possible. The disease activity score itself (i.e. SS score) however, has been shown to be predictive of organ damage and mortality (Ibanez et al. 2003). As such, for the purpose of this health-economic model, the SS score alone was deemed more appropriate to link with long-term outcomes; it was part of the composite SRI endpoint; is the measure of efficacy within that endpoint; and is the primary driver of the SRI response in the BLISS trials. Belimumab 10 mg/kg showed a significantly higher percentage change from baseline in SS score after 52 weeks in BLISS-52 and BLISS-76 compared with SoC. The pooled average SS score from baseline to week 52 for SoC and belimumab 10mg/kg is shown in Figure 6.4 below for the total BLISS population.

Figure 6.4. Average SELENA-SLEDAI score from baseline to week 52 for SoC (placebo) and belimumab 10mg/kg – Pooled total population

To determine a patient's change in SS score at week 52 it is important to acknowledge the dependence with baseline score, the effect of treatment (whether a patient gets belimumab or SoC) and the difference between patients on belimumab with and without a response (defined as a reduction of ≥ 4 points SS at 24 weeks).

This can be achieved by fitting a linear regression on the pooled BLISS trial data that explains the difference between the SS score at baseline and week 52, depending on baseline SS score combined with a treatment indicator variable, and a "response" indicator variable identifying whether or not patients are classified as satisfying the treatment continuation rule at week 24 with belimumab. The results of the regression for estimating change in SS score at Week 52 for the pooled total population trial data are presented in Table 6.5 and show a good fit of the data. The relationship between baseline change in SS score at week 52 for the groups included in the model are shown in Figure 6.5.

Table 6.5. Linear regression explaining change in SELENA-SEDAI score at week 52 - Pooled total population

Parameter	Estimate	Std Error	p-value
SS_0 SoC	-0.390	0.016	<0.0001
SS_0 all belimumab	-0.285	0.028	<0.0001
SS_0 belimumab responders	-0.363	0.033	<0.0001
Adjusted R^2 =0.699			

Note "responders" are patients on belimumab who satisfy the treatment continuation rule.

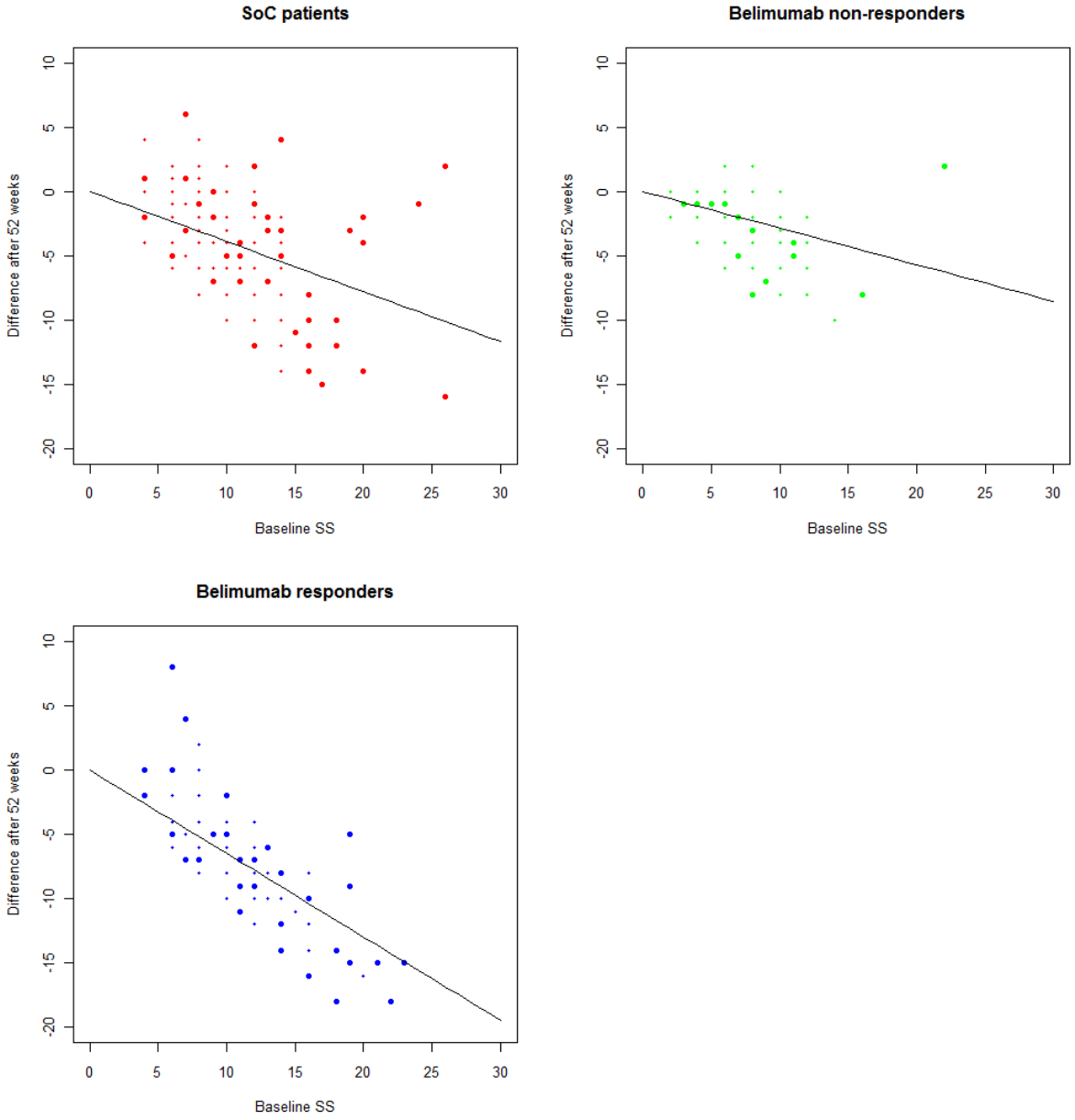
The coefficients in this table reflect the decrease in SS score (as a proportion of the baseline SS) at 52 weeks for SoC, belimumab and belimumab "responders". To determine a patient's score at week 52 (SS_{52}) having a baseline score SS_0, first the difference compared to SS_{52} is calculated as:

is 1 if on SoC; 0 if not
is 1 if on belimumab; 0 if not
is 1 if a responder on belimumab treatment;
0 if SoC or non-responder on belimumab.

For example a responder on belimumab with a baseline SS of 10 would have a difference after 52 weeks of:

This difference is added to SS_0 to get SS_{52} to give an SS score of 3.52 after 52 weeks. The higher the baseline SS score, the greater the decrease. Although the effect of SoC and belimumab is kept constant in this approach, due to different baseline SS scores, there will be variability in the simulated week 52 SS scores.

Figure 6.5. Plots of correlation between baseline SS and difference after 52 weeks for SoC patients, and belimumab responders and non-responders – Pooled total population.

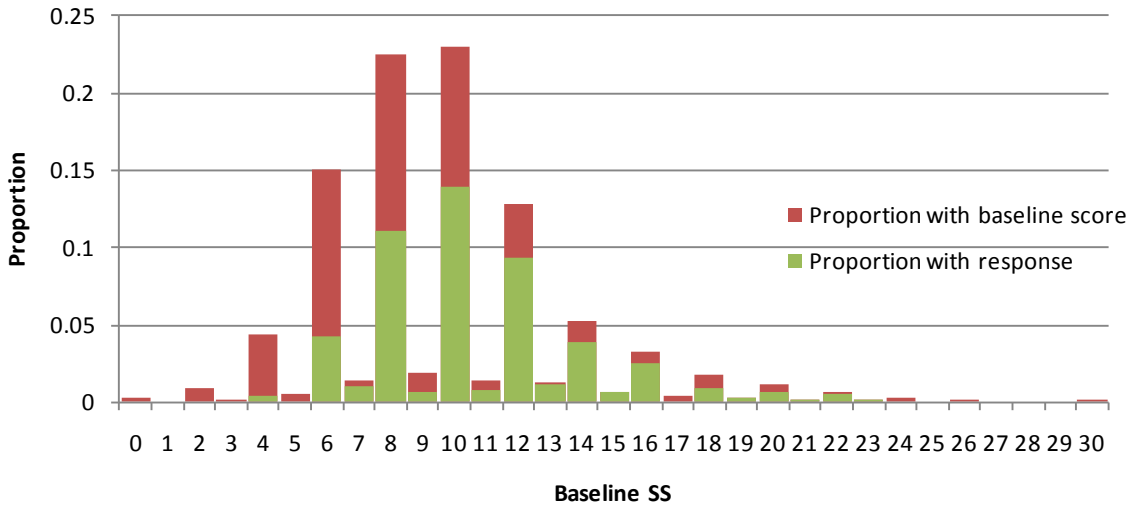


NOTE: point size represents number of patients. The top-right graph shows data for belimumab non-responders, whereas the regression in Table 6.5 contains the covariate “all belimumab”.

Treatment continuation rule

As shown in the previous section, for patients receiving belimumab, treatment is modeled to be continued only for patients with a reduction of ≥ 4 points in SS score at 24 weeks. As explained in Section 6.2.8, this time-point is chosen to be consistent with how the continuation rule for belimumab is likely to be implemented in clinical practice in line with the proposed SPC. This is different, however, to how the primary outcome of response was assessed in the clinical trials i.e. at 52 weeks. As shown in Table 6.6, the probability of treatment continuation after week 24 in the BLISS trials depends to a high degree on baseline SS score, so the probabilities of treatment continuation were stratified by baseline SS score. Figure 6.6 below presents a summary of the distribution of baseline SS scores and within each score the percentage of patients defined as a "responder". The spiky behaviour of baseline scores is related to the SELENA-SLEDAI score list, in which even scores were more likely. Only scores which were present in the BLISS data would be simulated i.e. as there were no baseline scores of 24, 25 or 26, none of these scores would be simulated.

Figure 6.6. Distribution of baseline scores and proportion with response in the BLISS trials – Pooled total population



Estimates of the probability of treatment continuation are based on the probability of response observed in the trials. In the model, treatment continuation is determined for belimumab patients by using a Bernouilli distribution with a probability corresponding to the values in the table.

Table 6.6. Probabilities of treatment continuation at 24 weeks for different baseline SS scores – Pooled total population

Baseline SELENA-SLEDAI	Probability of treatment continuation
0	0%
1	0%
2	0%
3	0%
4	11%
5	0%
6	28%
7	71%
8	49%
9	33%
10	61%
11	56%
12	73%
13	89%
14	75%
15	100%

Baseline SELENA-SLEDAI	Probability of treatment continuation
16	77%
17	0%
18	50%
19	100%
20	57%
21	100%
22	75%
23	100%
24	0%
25	0%
26	0%
27	0%
28	0%
29	0%
30	0%

Extrapolation to long-term SLE outcomes

The BLISS trials were not designed to capture long-term effects of belimumab due to their relatively short duration. However, high disease activity measured by SELENA-SLEDAI (SS) score is linked to the long-term accrual of organ damage and to mortality (Swaak et al. 1999). SS score was a major component of the primary endpoint the responder Index (SRI). Due to the lack of long-term data from the clinical trials and because in SLE long-term outcomes will have a major effect on the health-economic outcomes, other data sources were identified to help with assessing the likely effect of belimumab on organ damage and survival. Data from a longitudinal SLE database was used to estimate natural history models that describe the progress of SLE outcomes over a longer period of time.

The Johns Hopkins (JH) Lupus cohort reports data on a large population of SLE patients from Baltimore, Maryland. Patients in the JH cohort visit the clinic every 3 months from cohort entry. Data extracted from the database in early 2010 provided

a dataset of 2047 patients. The SLEDAI was originally developed in 1992 and so deriving a SLEDAI score prior to this time point would not have been straightforward and the analysis would not have been ready in time for this submission. In order to have a completed analysis to incorporate into the health economic model all patient observations that were conducted before 1992 were excluded from the analysis. However, an analysis is currently ongoing to incorporate the data prior to 1992 so that the results can be compared. The effect of the data exclusion reduced the sample size to 1985 patients. Seven hundred and three patients had follow-up durations of less than 24 months. It was assumed that patients with a short follow-up (defined as less than two years), although would increase the accuracy of estimates of short-term outcomes, would not contribute significantly to the estimation of long-term outcomes, and so they were excluded from the analysis, leaving a final sample size of 1282 patients. Time to event (TTE) models, discussed later in this section, are used to identify the relationship between disease activity (SLEDAI) and organ damage or mortality. SLEDAI scores over time are required in these TTE models so it seemed reasonable to exclude patients for whom a meaningful calculation of the score over time could not be estimated.

A second longitudinal SLE cohort was also examined, the Toronto SLE cohort, as an alternative database with which to estimate long-term effects. However, the patient level data was not available to GSK for analysis and it was not as complete as the JH database. The level of missing data resulted in a lack of robustness in models estimating mortality risk, a key outcome in the cost-effectiveness model, making interpretation of the results unreliable. However, there was some benefit in using this database to help with external validation of some of the long-term modelling and this is discussed in the section on validation later in this document.

The baseline patient characteristics of the subset of patients used from the JH registry are described in Table 6.7. The SS score to assess disease activity was used in the two BLISS trials, whereas the JH cohort used the original SLEDAI. The SELENA-SLEDAI modification alters the definitions of some symptoms to improve clarification and attribution of the items. Given that both indices use the same 24 items and weights, the small difference in classifications is thought unlikely to influence the model results.

Table 6.7. Baseline characteristics of Johns Hopkins cohort used for data analysis

Baseline characteristics	Summary statistics
Number of patients	1282
Females	1,190 (92.8%)
Black ethnicity	492 (38.4%)
Caucasian	672 (52.4%)
Age at diagnosis (mean (SD))	33.1 (13.0)
Age at cohort entry (mean (SD))	38.2 (12.8)
Disease duration at cohort entry (mean (SD))	5.15 (6.5)
SLEDAI score at first visit (mean (SD))	3.32 (3.7)
Steroid dose at first visit (mean (SD))	9.95 (15.3)
Past smoker (%)	38.9%
Hypertension (yearly risk)	15.8%
Anticardiolipid antibodies positive (%)	3.0%
Lupus anticoagulant positive (%)	9.6%

The baseline characteristics from the JH cohort are different from the pooled BLISS trials (Tables 6.2, 6.3); a higher proportion of patients of black ethnicity, a slightly lower disease duration and a considerably lower SLEDAI score observed in the JH cohort. This suggests that SLE in the JH cohort was on average less severe than that for the BLISS trials and therefore any associated impact on long-term organ damage could be underestimated. To account for these differences, baseline characteristics (e.g. ethnicity, age at diagnosis and disease duration) were added as potential confounders for the risk of mortality and developing organ damage. The analysis of the longitudinal JH data and the way this information was included in the HE model is explained in the next sections, covering long-term disease activity, steroid use, mortality and organ damage development. More detailed information on this cohort can be found in Section 9.21, Appendix 21.

Long-term SELENA-SLEDAI score

This section describes the considerations and results of the JH analysis to describe long-term disease activity and how these results were implemented in the health-economic model. Detailed information on the covariate selection can be obtained from Section 9.21, Appendix 21. A short overview of assumptions on SS score and available evidence is shown in Table 6.8 below.

Table 6.8. Overview of assumptions made in the model with respect to SS score

Assumption	Evidence
Disease activity reaching plateau	Johns Hopkins data (Figure 6.7) BLISS: 52 to 76 week data (Figure 6.9)
Adjustment to JH model	Phase 2 study (Figure 6.10)
Constant absolute effect of belimumab	BLISS: 52 to 76 week data (Figure 6.9) Phase 2 study (stabilising pattern in Figure 6.10)

In the model rather than using SS scores to reflect disease severity over time, the scores are used to calculate the adjusted mean SLEDAI (AMS) score. This AMS score was developed to measure disease severity over time (Ibanez et al. 2003) as opposed to the SS score which only reflects disease activity over the preceding 10 days. AMS is calculated as the area under the curve of disease activity measurements between two time-points. The area under the curve is then divided by time of follow-up to provide an average score over the period of interest. The disadvantage of using the AMS is that it effectively smoothes out the SLEDAI scores and so extreme highs and lows in disease activity present for short periods of time will not be represented as substantial changes in AMS. Consequently, if short-term flares in disease activity cause greater damage to patients than prolonged low levels of disease activity the model will not be able to distinguish between these two patterns. However, this concern has to some extent been addressed by Ibanez et al (2005) who investigated the effects of variability in SLEDAI score and found that AMS is able to predict poor outcomes in SLE independently of variability measures.

A “disease activity model” constructed from the JH data was included in the cost-effectiveness model in order to relate disease activity to risk of longer term organ damage. The time period of one year was chosen with which to provide estimates of AMS over time since the model uses a yearly cycle and 52 weeks was the primary endpoint of the BLISS studies. This “disease activity” model predicts the change in AMS score between two sequential one-year periods based on the time-dependant covariates listed in Table 6.9. These covariates were chosen based on a 10% level of statistical significance. This analysis provides an estimate of the average population AMS score (over ‘lifetime’, time in the model) weighted by time using panel data regression techniques. Random intercept models were adopted to allow

for dependence among the disease activity scores of patients within the cohort. The random effects model controls for repeated measures within patients and provides estimates of within–patient effects of covariates.

Table 6.9. Coefficient results for the linear regression model predicting change in mean SLEDAI – Johns Hopkins Cohort

Covariate	Coefficient	95% CI	
Mean SLEDAI score in previous period	-0.4163	-0.4396	-0.3929
Male gender	-0.0991	-0.2544	0.0562
Black ethnicity	0.3524	0.2566	0.4482
Log of age	-0.3586	-0.5072	-0.2100
Constant	2.0577	1.4855	2.6299
Sigma ui	0.4093		
Within R^2	0.3624		
Overall R^2	0.1668		

The regression model is used as follows in the simulation to predict a patient's SS score at year *t*, based on the SS score at year *(t-1)*:

where

The solid blue line in Figure 6.7 shows the extrapolation of disease activity of the average SS score for SoC in the BLISS trials using the above model specification. It is clear that after one year the predicted disease activity declines relatively fast to a level that is in the range of the mean SLEDAI levels in the JH cohort (Figure 6.8). This suggests that the average levels in the JH cohort are not representative of an average SLE population and particularly not of the BLISS populations.

Figure 6.7. Extrapolation of disease activity of the average patient SLEDAI score in the BLISS trials treated with SoC using the original Johns Hopkins model (blue line) and adjusted model (dashed red line)

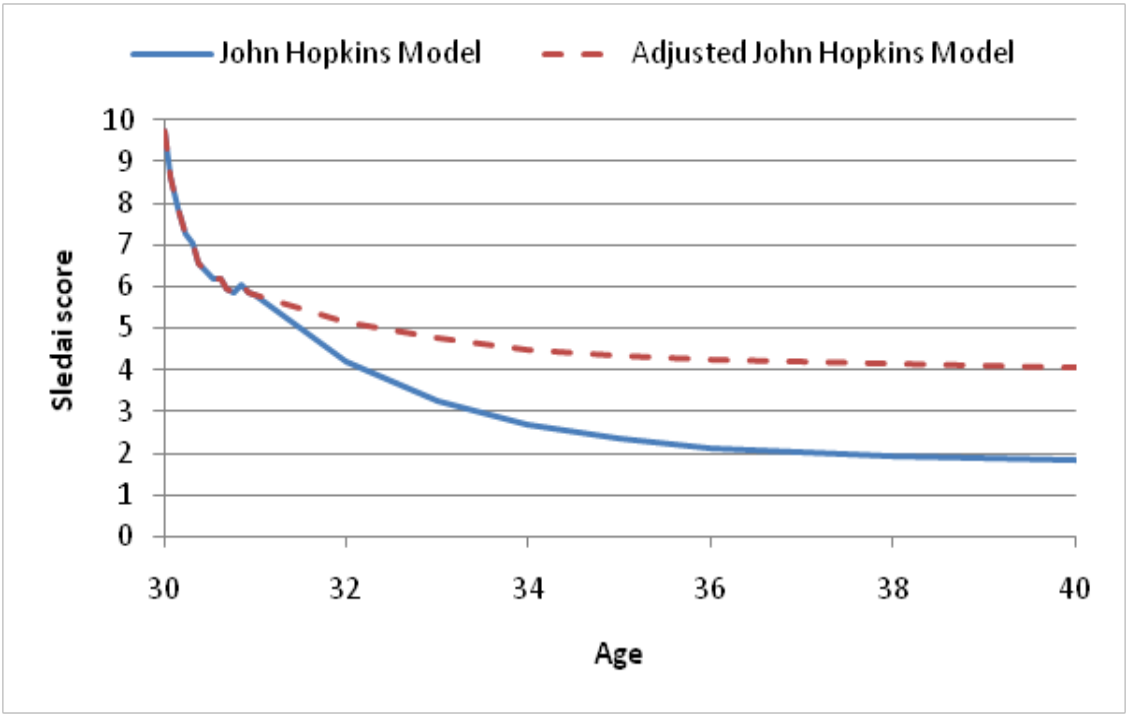


Figure 6.8. Mean Annual SLEDAI in Johns Hopkins Cohort

As illustrated in Figure 6.7, SS scores seem to level after around 36 weeks for SoC and remain relatively constant until week 52. It seems unlikely that after one year, the SS scores for the BLISS population will decline (improve) to JH levels as predicted by the long-term disease activity model. This is further substantiated by the additional data from the BLISS-76 trial (see Figure 6.9).

Figure 6.9. SELENA-SLEDAI score from week 52 to week 76 based on BLISS-76 study – Pooled total population

The disease activity model predicted from the JH cohort estimates the expected annual AMS score adjusted for a number of patient characteristics. However, the model has a tendency to predict average AMS scores for the JH cohort and these scores may not be applicable to other SLE populations. Table 6.10 details the annual AMS scores recorded in the JH cohort by severity. The table shows that the patients in the JH cohort most commonly have mild-moderate disease activity. Only a small proportion of observations(<5%) capture severe activity, AMS of 6 or more.

Table 6.10. The distribution of annual AMS scores in the Johns Hopkins cohort

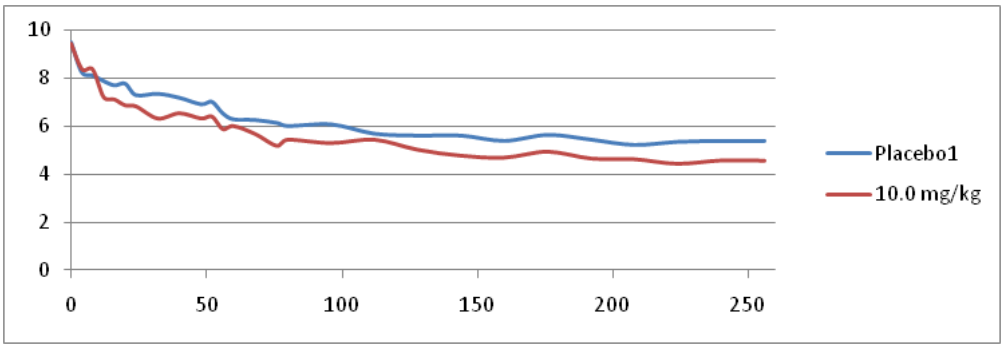
	Number of observations	% of observations
0-1 annual AMS	2272	24.2%
1-4 annual AMS	5228	55.5%
4-6 annual AMS	1454	15.5%
6-12 annual AMS	446	4.7%
>12 annual AMS	7	0.1%

The BLISS trials were designed to recruit patients with a SELENA-SLEDAI score greater than 6. It is likely that the JH cohort and their level of disease activity is unrepresentative of the BLISS patient population and may underestimate long-term

disease activity. To preserve the important effect of the covariates in the model presented in Table 6.9 and at the same time reflect long-term average disease scores more appropriate to the BLISS trial population, the model was adjusted by increasing the constant from 2.058 to 3.0, whilst keeping the other coefficients intact. The choice of the value of 3.0 was made from examining the range of SS scores in the Phase 2 study. A range of numbers were analysed to derive the adjusted constant, with a value of 3.0 providing a reasonable fit to these data. Adjusting the constant in this way is similar to the method in which the intercept is changed in the SCORE risk equation used in cardiovascular studies for the prevention of cardiovascular disease (van, I et al. 2010). In this SCORE risk equation the constant is changed to correct for different baseline risk in a certain country.

The dashed red line in Figure 6.7 shows the extrapolation of disease activity of the average SS score for SoC in the BLISS trials using the adjusted natural history model (with the constant of 3). It is clear, that the average SS scores over time stay much closer to the BLISS SoC week 52 score. There is still a decline in this curve representing a decreasing pattern as patients get older. When comparing the red curve with the data from the long-term phase 2 extension study (LBSL99) over the first 250 weeks (approximately 4.5 years) presented in Figure 6.10, it can be seen that the SS scores predicted with the adjusted JH model (SS between 4 and 5) better reflect the SS scores observed in the long-term extension study (SS around 5) than the unadjusted JH model (between 2 and 3). The patients enrolled in the LBSL99 study had similar SS scores at baseline to the patients enrolled in the two BLISS studies.

Figure 6.10. SELENA-SLEDAI score by visit (LOCF), autoantibody positive – LBSL99 extension study



The adjusted JH model was used to predict the SLEDAI score of a patient treated with SoC after one year. The model allows for the selection of the original and adjusted JH model. For the base case analysis, the adjusted JH model has been used.

Flares

Following personal correspondence from Professor Petri, a clinical consultant on the project from Johns Hopkins School of Medicine and Director of the Lupus Center, it was decided that the incidence and severity of flares would not be included in the disease activity and organ damage models. The JH cohort does not collect data on disease activity flares. Dr Petri did not consider it appropriate to estimate the SELENA Flare Index (SFI) (Section 9.17, Appendix 17), a validated measure of flare severity, from the JH dataset. It was suggested that using a change in SLEDAI score of 3 or 4 units could be used as an alternative measure of flare. However, this definition would not use any additional data compared with that used to estimate the AMS. The protocol for the JH cohort requires patients to visit the clinic every three months and patients often visit more frequently during periods of disease flare so most flares in disease will have been captured in the database and therefore in the AMS.

Although disease flares are not explicitly modelled in the JH analyses, disease activity at the time of organ damage or mortality is reflected in the individual system involvement covariates. Individual system involvement indicators were included in the baseline analysis of the JH cohort. These indicate the profile of disease activity

across organ systems at the time of damage. It was considered that these data would complement the AMS score by describing current disease activity and also indicating the type of disease activity.

Furthermore, incorporating disease flare into the cost-effectiveness analyses would have required assumptions regarding the relationship between flare, mean SLEDAI scores, and steroid dose. Therefore the approach was rejected because it would increase the complexity of the model, without substantial gain in the description of the disease.

Additional effect of belimumab compared to SoC on disease activity

There is a lack of data on the relative effect on disease activity reduction of belimumab 10 mg/kg beyond one year compared to SoC. The analysis on the pooled total population showed an increasing difference compared to SoC in SS score from 0.59 on average at week 28 to 1.2 on average at week 52 (see also Figure 6.4). The BLISS-76 study shows a stabilising pattern of the difference between belimumab 10 mg/kg compared to SoC over time (see Figure 6.9).

Also, Figure 6.10, the Phase 2 extension study shows a further decrease in SS score after 52 weeks for patients on belimumab, but no comparison with SoC can be established from this data as the “placebo1” group was switched from placebo to belimumab after one year.

In the simulation model, an assumption was made that the additional absolute effect of belimumab on disease activity reduction remains constant after one year. This is graphically illustrated by Figure 6.11. This is a key model assumption and was discussed with Professor Petri who has observed patients on belimumab in her clinic for a number of years as part of the Phase 2 open-label extension study.

Figure 6.11. Extrapolation of disease activity using the adjusted Johns Hopkins model of the average patient SLEDAI score in the BLISS trials assuming a constant additional effect of belimumab compared with SoC after Year 1

Long-term SLEDAI item involvement

There is no information on the natural history development of the SLEDAI item involvement in the JH data. The model could make the assumption that the percentage of involvement for the eight individual SLEDAI items seen in the BLISS trial data are kept constant over time after 52 weeks for belimumab and SoC. However, since this may cause inconsistency between item involvement and SS score combinations, the base case analysis uses statistical models for mortality and organ damage in which the effect of item involvement was removed and replaced by the AMS. This is discussed in more detail in Section 9.21, Appendix 21.

Steroid Use

Although 'steroid use was collected in the BLISS studies and showed some benefit of belimumab in reducing dose over time, the effect was not significant. However this is not unexpected within the constraints of the two Phase 3 RCTs as steroid tapering was not mandated in these studies and was based upon the investigator's clinical judgement; rapid steroid reduction was discouraged to prevent escape of disease control and possible SLE flare. On the basis that the results from the trial analysis are likely to have underestimated the benefit of belimumab in steroid reduction, it was deemed inappropriate to use these data to predict a patient's steroid dose at week 52 in the simulation. The JH database was favoured to look at the relationship between disease activity and steroid use and showed a clear

relationship between these two variables. In general, the higher the disease activity, the higher the dose of steroids used. Based on the JH data, a random effects model was fitted to estimate the linear relationship between disease activity (average SLEDAI score) and steroid dose whilst accounting for the unobserved individual patient characteristics that induce correlation between observations within a single patient (see Table 6.11). The model can be used to estimate mean steroid dose for each patient at any time conditional on their disease activity status. More detailed information on this analysis is provided in Section 9.21, Appendix 21.

Table 6.11. Linear regression model explaining average steroid dose per year (mg/day) based on SLEDAI score (model input) - Johns Hopkins cohort

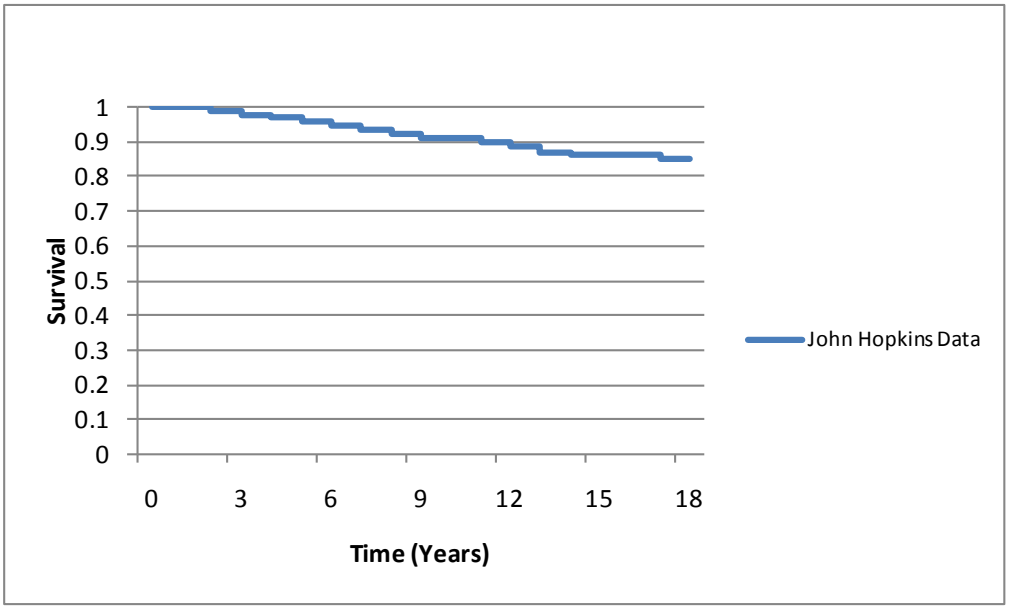
Regression parameter	Coefficient (95% CIs)	P-value
Average SLEDAI score during current year	0.7199 (0.617, 0.823)	<0.001
Constant	3.410 (3.073,3.747)	<0.001

The regression equation is implemented in the model to predict a patient's average steroid use during each year, depending on that same year's SLEDAI score. For example, if a patient had a mean SLEDAI score of 10 during Year 1, then the regression equation would estimate that an average steroid dose of 10.6 mg/day was taken by the patient.

Mortality

The JH data recorded a total of 82 deaths during the period of observation. Figure 6.12 below illustrates the Kaplan-Meier curve for mortality events.

Figure 6.12. Kaplan-Meier curve for mortality events



Individual univariate analyses using an exponential survival curve explaining the effect on the risk of death (exponential hazard) were conducted for covariates that *a priori* were thought to possibly affect the risk of death. The variables found to be statistically significant in the univariate analysis were included in a multivariate stepwise covariate selection process. The Exponential, Weibull, Gompertz and Log-logistic models were all tested and the Akaike Information Criterion (AIC) statistic suggested that the Weibull distribution had the best internal goodness of fit. The results of this model showing all the variables with a statistically significant effect on survival are presented in Table 6.12. Detailed information on this analysis and all other models tested are provided in Section 9.21, Appendix 21.

Table 6.12. Weibull survival model explaining risk of death with AMS included and item involvement effects removed - JH cohort

Covariates	Model coefficient
Constant	-10.366
Black ethnicity	0.7814
Age at diagnosis	0.0321
Cholesterol	0.0044
AMS over lifetime	0.2135
Cumulative Average Prednisone Dose (mg/month)	0.0012
Renal damage	0.652
Musculoskeletal damage at previous visit	0.415
Peripheral vascular damage at previous visit	0.9783
Gastrointestinal damage at previous visit	0.4684
Diabetes at previous visit	0.6764
Malignancy at previous visit	1.1489
Any infection at time of death at current visit	0.7409
Parametric distribution parameter for Weibull	1.6799

The above Weibull survival distribution was used to assess the risk of dying in the simulation. The mean age at entry for the JH cohort was 38 years, similar to that of the pooled BLISS study population, and there was a maximum follow-up of 18 years. Therefore, the incidence of mortality in elderly patients is unlikely to be captured in the above model even over a lifetime horizon. In order to avoid an underestimation of mortality in the model a correction was required to increase mortality risk at older ages using mortality estimates for the general population (Bernatsky et al. 2006). Bernatsky et al (2006) have shown that SLE patients have an increased risk of mortality compared with the general population. The increased mortality risk was reflected in the mortality correction.

The Weibull hazard function is a proportional hazards model and can be described as

where r is the parametric distribution parameter from Table 6.12 and where X are the covariates from Table 13 and β are the corresponding coefficients. The hazard ratio , can be calculated for an individual patient p relative to the average JH hazard by:

—

where for individual patient *p* and reflecting the average JH characteristics . is multiplied with the age specific standardised mortality ratio for a general SLE population based on Bernatsky et al (2006) (see Table 6.13). This gives an estimate of the hazard ratio of the individual patient *p* compared to the general population. Multiplying this with the age and gender related general population hazard would give an estimate of the average hazard of patient *p.*

Table 6.13. Standardised Mortality Ratios for SLE patients stratified by age groups according to Bernatsky et al (2006).

Age	Standardized Mortality Ratio	95% CI
16-24	19.2	14.7, 24.7
25-39	8.0	7.0, 9.1
40-59	3.7	3.3, 4
>60	1.4	1.3, 1.5

As an example, consider a 55 year old female with probability of death .
The steps below demonstrate how the average mortality hazard for a patient is calculated:

Step 1	Calculate the probability of death for a 55 year old female	p(death)=0.0012	(A)
Step 2	Calculate the probability of death for an average patient in the Johns Hopkins cohort	p(death)=0.00104	(B)
Step 3	Calculate the hazard ratio for the 55 year old female compared with the average patient	A/B=1.15	(C)
Step 4	Extract the relevant SMR for SLE in a 55 year old female (Table 6.13)	SMR=3.7	(D)
Step 5	Multiple the SMR by the individual patient hazard ratio	C*D=4.26	(E)
Step 6	Extract the general population risk of mortality from the relevant life tables*	life tables=0.0037	(F)
Step 7	Multiply the general population mortality rate by the individual SLE patient mortality ratio	F*E=0.0157	(G)
Step 8	Convert the rate of mortality into a probability	p(death)=0.0158	(H)
*(2007-2009 UK interim life tables, national office of statistics, [http://www.statistics.gov.uk/statbase/Product.asp?vlnk=14459]),			

Organ Damage Development

The JH data was also used to estimate the time to organ damage outcomes. The analysis identified risk factors for decreasing time for these outcomes occurring and quantified the effect each risk factor had on the probability of the outcome occurring at any stage of the disease. Univariate regression analyses were run on (time varying) covariates to identify suitable variables for inclusion in the multivariate analyses. All significant variables from the univariate analysis (using 10% level of statistical significance) were included in the multivariate regressions. Modifications to the model specification were made using a process of backward elimination to select a final set of statistically significant variables. A separate multivariate model was used for each individual organ system. Multivariate regression models were run using exponential, Weibull, Gompertz, and log-logistic distributions. Suitable parameter specification for each type of organ damage was selected using the AIC statistic.

The functional form and one year hazard rate for the exponential, Weibull, Gompertz, and Log-Logistic survival curve are outlined below.

Survival curve	**Functional Form**	**Yearly hazard**	**Par 1**	**Par 2**
Exponential				n/a
Weibull				
Gompertz				
Log-Logistic				

The models and corresponding covariates that were selected are summarised in Table 6.14. Any covariate that was significant in the model for a particular type of organ damage has a coefficient included in the corresponding organ damage column in the table. Detailed information on the methods and results from the model construction process are presented in Section 9.21, Appendix 21.

Table 6.14. Organ damage time to event models and corresponding covariates from Johns Hopkins cohort analysis

	CV	Diabetes	GI	Malignancy	MSK	NP	Ocular	PV	GF	Pulmonary	Renal	Skin
Survival model	Loglog	Exp	Exp	Exp	Loglog	Weibull	LogLog	Exp	Exp	Gompertz	Exp	LogLog
Covariates												
Male				0.4981								
Black		0.7805										
Age at diagnosis	-0.054			0.0229	-0.0354							
Past smoker								0.6066				-1.5658
Cholesterol				-0.0088		0.0047			0.005		0.008	
Hypertension	-1.089					0.5167		1.0051				
AAP										1.0132		
LAP								1.3705				
Log of age		2.2481				0.607	-2.97	1.1608		1.2316		
Log of disease duration	-0.741			0.3082	-0.6747							
AMS	-0.209		-0.0606		-0.0407	0.044	-0.045	0.1702		0.1388	0.3234	-0.0466
CAPD	-0.001	0.0019	0.0011		-0.0018		-0.002		0.0022			-0.0025
SLICC/ACR score				0.1467	-0.1448	0.0954				0.1039		
Renal damage	-0.834											
Diabetes at previous visit	-1.067											
Constant	10.123	-14.6564	-4.8419	-4.8106	7.0495	-7.3961	15.993	-11.695	-7.6433	-9.265	-8.293	9.651
Parametric par	1.2164				1.1421	0.8161	1.084			-0.0382		1.5938

CV = cardiovascular, MSK = musculoskeletal, NP = neuropsychatric, PV = peripheral vascular, GI = gastrointestinal, GF = Gonadal Failure, Loglog = loglogistic, Exp = exponential, AAP = Anticardiolipid antibodies, LAP = Lupus anticoagulant positive, AMS = average mean SLEDAI up to current time, CAPD = cumulative average prednisone dose up to current time, Seros = serositis, Paramteric par = additional parametric distribution parameter for non-exponential survival models.

Table 6.14 shows a significant relationship with the risk of various types of organ damage with the following risk factors: past smoker (yes/no), cholesterol level, hypertension (yes/no), anticardiolipid (ACL) antibodies positive (yes/no) and lupus anticoagulant positive (yes/no). These values for each patient were imputed based on the JH data as the data were not available in the BLISS dataset. As belimumab and SoC are not expected to behave differently on most of these risk factors, this will not substantially influence the incremental results. Belimumab may have an effect on ACL antibodies, however, as a conservative approach no effect is assumed in the model. The data used for imputing these values in the simulation is presented in Table 6.15

Table 6.15. Johns Hopkins characteristics imputed in simulation model

Characteristic	Summary value
Past smoker (%)	38.9%
Cholesterol level (mean (SD))	189.0 (48.91)
Hypertension (yearly risk)	15.8%
Anticardiolipid antibodies positive (%)	3.0%
Lupus anticoagulant positive (%)	9.6%

The organ damage time to event models were used to estimate a patient's risk (hazard) of developing that specific organ damage. The statistical models were analysed separately and the hazard for each organ system are estimated independently in the cost-effectiveness model. However, correlation in the risk of events across organ systems is captured using the covariates in the statistical models. For example, if a patient develops diabetes this immediately increases their risk of cardiovascular damage, or if a patient has high 'steroid exposure they will increase their risk of musculoskeletal damage and ocular damage. However, latent variables to correlate the time to an event in different organ systems have not been applied and estimates for these are not available. The hazards for each organ system were translated to a yearly probability that was used in a Bernoulli distribution from which a random number was sampled indicating whether organ damage would occur in that year. The statistical models were estimated for all damage events recorded in the observation data, such that patients remain in the analysis after their initial event. However, it was decided that only the first

damage event within a particular organ system would be modelled and progression of damage within that organ system would not be estimated. This simplification of disease progression was used to avoid model complexity. Average damage scores are applied at the first damage event to account for the additional costs and utility decrements of multiple organ damage within an organ system. A patient's SLICC score was increased with the average organ specific damage score obtained from the JH dataset (see Table 6.16). This reflects the fact that some SLICC item scores are more than 1 and that for certain organ systems, a patient may have more than one item involved. In reality though, a patient may first present only one item in a single organ and it would take some time before developing damage on a different item within that organ. However, to model organ damage at this level of detail would be too difficult and data consuming. The assumption of using the "basket" organ damage scores in Table 6.16 below may have two different impacts on the cost-effectiveness outcomes. On one hand, it may improve the cost-effectiveness outcomes for belimumab by overstating the burden of damage at the time of first damage event. However in contrast, it may also underestimate the benefit of belimumab because it does not allow belimumab to reduce the progression of organ damage events. For example, if a patient enters the model with renal damage, the model does not estimate the benefits of subsequent renal events avoided.

Table 6.16. Average SLICC scores per organ - based on all recordings in Johns Hopkins cohort

Organ	Score
Cardiovascular	1.42
Diabetes	1.00
Gastrointestinal	1.09
Malignancy	1.00
Musculoskeletal	1.41
Neuropsychiatric	1.37
Ocular	1.23
Peripheral vascular	1.21
Premature gonadal failure	1.00
Pulmonary	1.31
Renal	1.83
Skin	1.14

Belimumab discontinuation

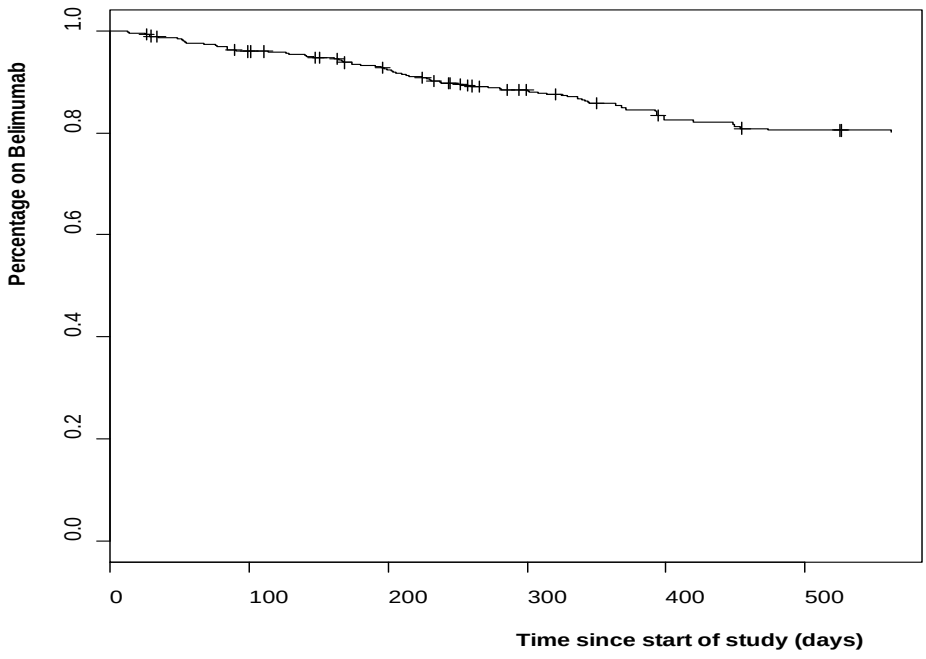
Belimumab treatment can be stopped due to three reasons.

1. Natural discontinuation
2. Insufficient response (model option, used in base case)
3. Maximum treatment duration

1. Natural discontinuation

This reflects the natural pattern due to patient request, lack of efficacy, lack of compliance or an adverse event as observed in the BLISS trials. The discontinuation probability after 1.5 years for belimumab in BLISS-76 was 17.1%. A time to discontinuation analysis was conducted. Patients discontinuing due to reasons other than those listed above, and the BLISS-52 study patients after 52 weeks, were censored in this analysis. This resulted in a yearly discontinuation probability of 12.1%, assuming a constant risk (see Figure 6.13).

Figure 6.13. Belimumab discontinuation over time (with 95% confidence intervals) - Pooled total population



However, it is likely that this probability would diminish over time due to the fact that patients with an insufficient response on belimumab would discontinue

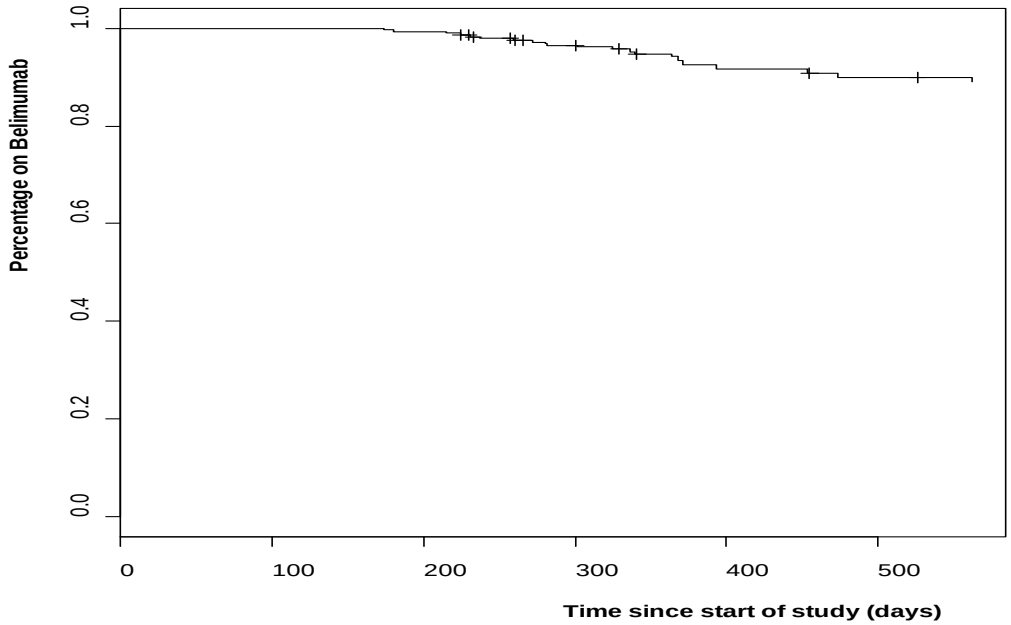
treatment early during the treatment; patients with a good response on treatment would have a lower discontinuation probability. To reflect this behaviour in the model, two different discontinuation probabilities are used; one for those satisfying the treatment continuation rule and one for those who did not. Figure 6.14 presents the probability of remaining on belimumab for patients satisfying the treatment continuation rule (i.e. "responders"). The definition of response is defined as ≥4 points decrease in SS score at 24 weeks compared to baseline, determined by the probability of treatment continuation presented in Table 6.6 for the different baseline SS scores. In the pooled BLISS trials, the discontinuation probability for responders in the total population was 10.9%, whereas patients with no response had a discontinuation probability of 21.4% (See Table 6.17 below). The percentage of patients satisfying the treatment continuation rule at 24 weeks was 52.4%. It was assumed that patients would discontinue treatment in the middle of the year. Treatment effect and costs for the year in which the patient discontinued were halved accordingly. Patients who did not satisfy the treatment continuation rule at week 24 were switched to SoC.

Table 6.17. Natural discontinuation and probability of treatment continuation for belimumab patients used in base case - Total pooled population

% belimumab patients satisfying treatment continuation rule at 24 weeks	52.4%	
Natural discontinuation	Patients satisfying treatment continuation at 24 weeks	Patients not satisfying treatment continuation at 24 weeks
Year 1	6.1%	21.4%
Subsequent years	10.9%	21.4%#

Note, only used if no continuation rule is applied in the model

Figure 6.14. Belimumab discontinuation over time for responders (with 95% confidence intervals). - Pooled total population.

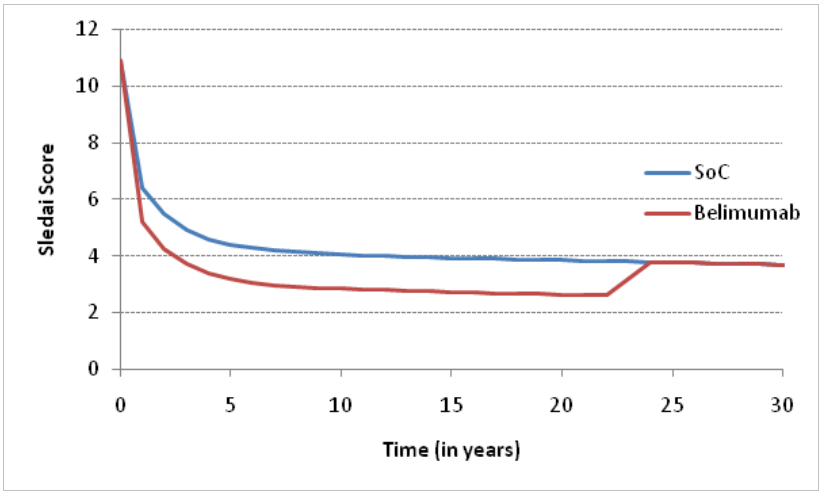


2. Insufficient response

For the base case in the model belimumab treatment is only continued for patients with a sufficient response at 24 weeks (defined as a reduction in SS of ≥4 points). Patients without a response are switched to SoC after 24 weeks. Responders do not discontinue in the first 24 weeks by definition, as response is measured after 24 weeks. After this period, discontinuation follows the pattern as described in Figure 6.13. A constant risk is applied throughout the model horizon, starting from week 24. This means that during the remainder of the first year, there is a probability of 6.1% of discontinuation and 10.9% for subsequent years.

After discontinuation a patient gets treated with SoC. This directly affects SLEDAI score and involvement parameters by applying the SoC effects. This is graphically illustrated for SLEDAI score in Figure 6.15

Figure 6.15, Example of SLEDAI score for a SOC patient and patient discontinuing belimumab treatment in Year 23.



NOTE: It is assumed that discontinuation takes place in the middle of the year. The red curve does not go up immediately due to the fact that mean SS scores are only measured at integer time points (years).

3. Maximum treatment duration reached.

It is possible to run the model using shorter than lifetime treatment durations for belimumab. However, this was not considered appropriate for this decision problem. For justification of this please see Section 6.6.1.

6.3.2 Demonstrate how the transition probabilities were calculated from the clinical data. If appropriate, provide the transition matrix, details of the transformation of clinical outcomes or other details here.

A Markov model has not been used for this decision problem. Probabilities for organ damage and death are based on linear regression and time to event models that are described in Section 6.3.1.

6.3.3 Is there evidence that (transition) probabilities should vary over time for the condition or disease? If so, has this been included in the evaluation? If there is evidence that this is the case, but it has not been included, provide an explanation of why it has been excluded.

There is evidence that the probabilities for organ damage and death change with time in this disease and this change is related to different factors such as oral ‘steroid use, previous organ damage and duration of organ damage. Because of

this, the natural history of the disease was studied in the JH cohort and the natural history models were developed, as discussed in Section 6.3.1. Time is included in several of these models as log of age (i.e. diabetes, gastrointestinal, malignancy, musculoskeletal, neuropsychiatric, ocular, peripheral vascular, gonadal failure, and pulmonary). The time effect is also included in the shape of the statistical models (e.g. Weibull model for mortality and neuropsychiatric).

6.3.4 Were intermediate outcome measures linked to final outcomes (for example, was a change in a surrogate outcome linked to a final clinical outcome)? If so, how was this relationship estimated, what sources of evidence were used, and what other evidence is there to support it?

Yes, adjusted mean SLEDAI (AMS) is linked to long-term organ damage in the model. This is discussed in detail in Section 6.3.1

6.3.5 If clinical experts assessed the applicability of values available or estimated any values, please provide the following details[3]:

- the criteria for selecting the experts
- the number of experts approached
- the number of experts who participated
- declaration of potential conflict(s) of interest from each expert or medical speciality whose opinion was sought
- the background information provided and its consistency with the totality of the evidence provided in the submission
- the method used to collect the opinions
- the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)
- the questions asked
- whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).

[3] Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee.

Advice from a number of clinical experts was obtained to check some of the assumptions in the model during its development. These were in the form of one to one consultations. In addition, Professor Petri of Johns Hopkins University Medical School, advised on the covariates of importance to consider in the construction of the long-term natural history models.

Summary of selected values

6.3.6 Please provide a list of all variables included in the cost-effectiveness analysis, detailing the values used, range (distribution) and source. Provide cross-references to other parts of the submission. Please present in a table, as suggested below.

Due to the very large number of parameters that are used in the model, this table was too large to summarise here. It is presented in Section 9.24, Appendix 24.

6.3.7 Are costs and clinical outcomes extrapolated beyond the trial follow-up period(s)? If so, what are the assumptions that underpin this extrapolation and how are they justified? In particular, what assumption was used about the longer term difference in effectiveness between the intervention and its comparator? For the extrapolation of clinical outcomes, please present graphs of any curve fittings to Kaplan-Meier plots.

Yes, both costs and clinical outcomes are extrapolated beyond the pooled clinical trial duration of 52 and 76 weeks for BLISS-56 and BLISS-76 respectively. The assumptions and methodology used have been summarised in Section 6.3.1 and more detailed methodology and modelling results are presented in Section 9.21, Appendix 21

6.3.8 Provide a list of all assumptions in the de novo economic model and a justification for each assumption.

All assumptions regarding statistical methodology used have been explained in Section 6.3.1. Only additional assumptions concerning the technology are presented in this section.

Exposure to belimumab

The average exposure to the trial product was assumed to be 100%. Fourteen infusions in the first year and 13 infusions in subsequent years are required. As this technology is not self-administered, patients are under specialist care and in a considerably poor state of health, it seems reasonable that compliance will be high while the physician perceives that the patient is receiving benefit from continuing this treatment. Level of compliance (i.e. exposure) can be changed in the model settings. However, this will affect only drug costs in the model; no adjustment of efficacy is made. This is because there is a lack of data to support what effect a reduced exposure would have on disease activity and longer term outcomes.

Vial Wastage

It is assumed in the base case that vial sharing between patients will not automatically occur, although a scenario has been included to look at the effect on the ICER if this were to be included. As the number of patients with moderate to severe SLE is relatively small, vial sharing may not be easy to manage in tertiary care units due to storage requirements. For rheumatoid arthritis patients, where vial sharing is often employed, this may be a more practical solution as the patient numbers are higher and so there will be less need to store opened vials for any considerable length of time.

6.4 Measurement and valuation of health effects

This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', section 5.4.

The HRQL impact of adverse events should still be explored regardless of whether they are included in cost-effectiveness analysis.

All parameters used to estimate cost effectiveness should be presented clearly in tabular form and include details of data sources. For continuous variables, mean values should be presented and used in the analyses. For all variables, measures of precision should be detailed.

Patient experience

6.4.1 Please outline the aspects of the condition that most affect patients' quality of life.

Systemic Lupus Erythematosus (SLE) is a chronic autoimmune disorder that may affect all organ systems, but most notably the skin, joints, kidneys, lungs, and nervous system. SLE often results in chronic debilitating ill health (Cervera et al. 2003; D'Cruz et al. 2007). Involvement of organs in the disease activity and chronic therapy with steroids and immunosuppressants lead to organ failure and increased morbidity (Gladman et al. 1996a; Zonana-Nacach et al. 2000a). The key aspects of SLE that most impact on patients' HRQL include the symptoms which occur with the presence of disease flares (such as joint and muscle pain, skin rash, and fever), the symptoms of chronic fatigue or malaise, and in more severe disease, the morbidity associated with organ damage and the side effects associated with repeated use of corticosteroid therapy. The inability to work and reliance on family and/or professional carers to help with normal everyday activities will have also have a major effect on a patient's mental wellbeing.

Within the evaluation, HRQL is captured through the EQ-5D which measures health status across five domains: mobility; self-care; usual activities; pain and discomfort; and anxiety and depression. However, this instrument may not be sensitive enough to reflect the impact of SLE disease flares especially if the instrument was completed while a patient's disease activity was in relative

quiescence. It is also likely to underestimate the impact of chronic fatigue, a very common symptom reported by SLE patients.

6.4.2 Please describe how a patient's HRQL is likely to change over the course of the condition.

For this decision problem we are specifically concerned with patients with moderate to severe SLE with high disease activity. These patients will already be experiencing poor HRQL as discussed in Section 6.4.1. Over time, with the cumulative effect of oral corticosteroid use and with the natural progression of organ damage inherent with this disease, the HRQL will deteriorate. The speed of this deterioration will depend on which organ systems are affected. Renal manifestations, neuropsychiatric disease and musculoskeletal disease are responsible for much of the morbidity directly related to SLE disease activity observed in the first 10 years (Chambers et al. 2009; Cooper et al. 2007). Patients with SLE have a 2.4-fold greater risk of mortality than the general population, with a higher risk of death due to cardiovascular disease, non-Hodgkin's lymphoma, lung cancer infections, and renal disease (Bernatsky et al. 2006).

HRQL data derived from clinical trials

6.4.3 If HRQL data were collected in the clinical trials identified in section 5 (Clinical evidence), please comment on whether the HRQL data are consistent with the reference case. The following are suggested elements for consideration, but the list is not exhaustive.

- Method of elicitation.
- Method of valuation.
- Point when measurements were made.
- Consistency with reference case.
- Appropriateness for cost-effectiveness analysis.
- Results with confidence intervals.

Both SF-36 and EQ-5D generic quality of life instruments were collected during the two Phase 3 studies. The latter instrument is consistent with the NICE reference case. However, this instrument may not be the most sensitive one to assess the true impact of the disease on HRQL experienced by SLE patients. This is because patients may experience disease flares at any time and not necessarily at the time the EQ-5D was completed for the pre-defined time points of the clinical trials. Consequently, the impact on HRQL is very likely to have been underestimated in the two BLISS studies.

The results of the EQ-5D instrument for the pooled BLISS dataset are presented in Figures 5.18 to Figures 5.21 in Section 5.5. These data are used to inform the baseline utilities used in the health economic model. The EQ-5D values were translated to utility values using the Dolan algorithm (Dolan 1997) to obtain UK general public related scores. The average baseline utility for patients in the BLISS trials was 0.70 (see Table 6.18 below).

Table 6.18. Descriptive statistics of EQ-5D data – Pooled total population

Descriptive statistic	Value
Mean EQ-5D (SD)	0.70 (0.26)
Number of EQ-5D observations	9,051
Mean Age (yrs) (range)	38 (18-73)

In addition, as fatigue is one of the most frequently cited symptoms by SLE patients, the FACIT-fatigue instrument (Section 9.19, Appendix 19) was also collected during the trials at 4, 8, 12, 24 and 52 week time-points and the results are also presented in Figures 5.14 to 5.17 in Section 5.5. The results demonstrate a significant improvement in fatigue scores with belimumab which was sustained over the trial period. This symptom will have a considerable impact on HRQL; if the EQ-5D is not sufficiently sensitive to detect the impact of this symptom, the overall utility benefit with belimumab will be underestimated in the cost-effectiveness model.

Mapping

6.4.4 If mapping was used to transform any of the utilities or quality-of-life data in clinical trials, please provide the following information.

- Which tool was mapped from and onto what other tool? For example, SF-36 to EQ-5D.
- Details of the methodology used.
- Details of validation of the mapping technique.

No mapping techniques were used to transform any of the utilities or quality-of-life data collected in the clinical trials**.**

HRQL studies

6.4.5 Please provide a systematic search of HRQL data. Consider published and unpublished studies, including any original research commissioned for this technology. Provide the rationale for terms used in the search strategy and any inclusion and exclusion criteria used. The search strategy used should be provided in section 9.12, appendix 12.

A formal systematic review for HRQL data was not conducted for several reasons. Firstly, the two Phase 3 BLISS studies are the first RCTs to provide EQ-5D data in the specific population of SLE patients of interest to this decision problem. Secondly, a significant impairment to HRQL comes from the long-term organ damage that SLE patients can experience. Due to the complexity of the disease, many organs can be affected. The SLICC score comprises 12 organ systems with a total of 41 damage items. Therefore the scope of a search on organ damage HRQL is almost unbounded due to the number and variety of damage items that are contained in the SLICC as well as the sometimes broad definition of damage. It was therefore not feasible to conduct a formal systematic review. In order to obtain utility weights for each type of organ damage a literature search was undertaken instead. Utility data was searched in Health Technology Assessments (HTAs) available on the NICE website. If the required information was unavailable from NICE, additional searches were carried out on Pubmed. A description of the search process conducted can be found in Section

9.12, Appendix 12. Additional information regarding the selection of utility weights is detailed in Section 9.25, Appendix 25.

6.4.6 Provide details of the studies in which HRQL is measured. Include the following, but note that the list is not exhaustive.

- Population in which health effects were measured.
- Information on recruitment.
- Interventions and comparators.
- Sample size.
- Response rates.
- Description of health states.
- Adverse events.
- Appropriateness of health states given condition and treatment pathway.
- Method of elicitation.
- Method of valuation.
- Mapping.
- Uncertainty around values.
- Consistency with reference case.
- Appropriateness for cost-effectiveness analysis.
- Results with confidence intervals.
- Appropriateness of the study for cost-effectiveness analysis.

As discussed above a formal systematic review was not conducted. A list of all HTA documents used to extract utilities related to organ damage are detailed in Section 9.25, Appendix 25.

6.4.7 Please highlight any key differences between the values derived from the literature search and those reported in or mapped from the clinical trials.

Short-term utility values were extracted directly from the pooled BLISS trials which provided the most appropriate HRQL data for short-term follow-up in the

population of SLE patients of interest to this decision problem. There are no other published studies in these patients reporting utility values. Utility weights relating to longer-term organ damage events were obtained from previous NICE HTAs or published literature; there is limited HRQL data from long-term clinical trials in SLE patients to enable a comparison.

Adverse events

6.4.8 Please describe how adverse events have an impact on HRQL.

As reported in the draft SPC, administration of Benlysta may result in hypersensitivity reactions and infusion reactions, although the incidence of serious infusion and hypersensitivity reactions (such as anaphylactic reaction, bradycardia, hypotension, angioedema, and dyspnea) seen in the clinical trials was generally low and were mild or moderate in severity. Infusion reactions occurred more frequently during the first two infusions and tended to decrease with subsequent infusions. If serious reactions do occur, the infusion is stopped and the patients are treated and monitored as appropriate. This will inevitably cause some discomfort and anxiety to the patient.

The mechanism of action of belimumab could also increase the potential risk for the development of infections, including opportunistic infections. The most common serious infections observed in the clinical trials comprised pneumonia, UTI, cellulitis, bronchitis, and pyelonephritis; however these events generally occurred at similar rates between the placebo and the belimumab groups. These infections will have an acute, but generally short-term negative impact on patient HRQL.

In the clinical trials, adverse reactions were reported in 93% of belimumab - treated patients and 92% of placebo-treated patients. The most frequently reported adverse reactions (≥10% of patient with SLE treated with belimumab plus SoC and at a rate ≥1% greater than placebo) were nausea, diarrhoea, and pyrexia. Although unpleasant, these adverse reactions will generally resolve after a few days and will require minimal intervention, thus will not have a significant effect on HRQL.

6.4.9 Was expert opinion used to estimate any clinical parameters Quality-of-life data used in cost-effectiveness analysis

No, expert opinion was not used to estimate any quality of life data used in the cost-effectiveness analysis.

6.4.10 Please summarise the values you have chosen for your cost-effectiveness analysis in the following table, referencing values obtained in sections 6.4.3 to 6.4.8. Justify the choice of utility values, giving consideration to the reference case.

Table 6.19 summarises the utility values used in the health economic model associated with disease activity (SS score) and long-term organ damage.

Table 6.19. Summary of quality-of-life values for the cost-effectiveness analysis

<table>
<tr><th rowspan="2">Organ Damage System</th><th colspan="3">Disutilities Year</th><th rowspan="2">SD</th><th rowspan="2">Reference</th><th rowspan="2">Assumption/justification</th></tr>
<tr><th>1</th><th>2</th><th>Subsequent</th></tr>
<tr><td>Cardio-vascular</td><td>0.72</td><td>0.76</td><td>Same as Y2</td><td>assumed 10%</td><td>Section 9, Appendices 25 and 26</td><td>Weighted average of:<table><tr><th>Item</th><th>utility Y1/Y2</th><th>weight</th></tr><tr><td>Angina or coronary artery bypass</td><td>0.77 /0.85</td><td>22%</td></tr><tr><td>Myocardial infarction 1</td><td>0.76/0.84</td><td>25%</td></tr><tr><td>Myocardial infarction 2</td><td>0.76/0.84</td><td>0%</td></tr><tr><td>Cardiomyopathy (ventricular dysfunction)</td><td>0.77 /0.77</td><td>25%</td></tr><tr><td>Valvular disease (diastolic or a systolic murmur > 3/6)</td><td>0.77 /0.77</td><td>18%</td></tr><tr><td>Pericarditis x 6 months or pericardiectomy</td><td>1 / 1</td><td>10%</td></tr></table></td></tr>
<tr><td>Diabetes</td><td>0.91</td><td>0.91</td><td>Same as Y2</td><td>assumed 10%</td><td>BLISS EQ-5D data analysis</td><td>Phase 3 BLISS trials</td></tr>
<tr><td>Gastro-intestinal</td><td>0.79</td><td>0.91</td><td>Same as Y2</td><td>assumed 10%</td><td>Section 9, Appendices 25 and 26</td><td>Weighted average of:<table><tr><th>Item</th><th>utility Y1/Y2</th><th>weight</th></tr><tr><td>Infarction or resection of bowel below duodenum, spleen, liver or gall bladder ever, for whatever cause (score 2 if > one site)</td><td>0.77/ 0.9</td><td>85%</td></tr><tr><td>resection > 1 site</td><td>0.77/ 0.9</td><td>1%</td></tr><tr><td>Mesenteric insufficiency</td><td>1 / 1</td><td>3%</td></tr></table></td></tr>
</table>

<table>
<tr><th>Organ Damage</th><th colspan="3">Disutilities Year</th><th>SD</th><th>Reference</th><th>Assumption/justification</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td>
<table>
<tr><td>Chronic peritonitis</td><td>1 / 1</td><td>3%</td></tr>
<tr><td>Stricture or upper gastrointestinal tract surgery ever</td><td>1 / 1</td><td>5%</td></tr>
<tr><td>Pancreatic insufficiency requiring enzyme replacement or with pseudocyst</td><td>1 / 1</td><td>3%</td></tr>
</table>
</td></tr>
<tr><td>Malignancy</td><td>0.92</td><td>0.92</td><td>Same as Y2</td><td>assumed 10%</td><td>Section 9, Appendices 25 and 26</td><td>Malignant tumours (excluding dysplasia) (Score 2 if > one site)</td></tr>
<tr><td>Musculo-skeletal</td><td>0.67</td><td>0.74</td><td>Inccreasing - See Appendix 9.26</td><td>assumed 10%</td><td>Section 9, Appendices 25 and 26</td><td>Weighted average of:
<table>
<tr><th>Item</th><th>utility Y1/Y2</th><th>weight</th></tr>
<tr><td>Muscle atrophy / weakness</td><td>1 / 1</td><td>8%</td></tr>
<tr><td>Deforming or erosive arthritis (including reducible deformities, excluding avascular necrosis</td><td>0.69 / 0.69</td><td>19%</td></tr>
<tr><td>Osteoporosis with fracture or vertebral collapse (excluding avascular necrosis</td><td>0.80 / 0.91</td><td>35%</td></tr>
<tr><td>Avascular necrosis</td><td>0.57/0.63</td><td>26%</td></tr>
<tr><td>Avascular necrosis 2</td><td>0.57/0.63</td><td>2%</td></tr>
<tr><td>Osteomyelitis</td><td>1 / 1</td><td>2%</td></tr>
<tr><td>Ruptured tendon</td><td>1 / 1</td><td>8%*</td></tr>
</table>
</td></tr>
<tr><td>Neuro-psychiatric</td><td>0.68</td><td>0.71</td><td>Same as Y2</td><td>assumed 10%</td><td>Section 9, Appendices 25 and 26</td><td>Weighted average of:
<table>
<tr><th>Item</th><th>utility Y1/Y2</th><th>weight</th></tr>
<tr><td>"Cognitive impairment OR major psychosis"</td><td>0.92 / 0.94</td><td>23%</td></tr>
<tr><td>Seizures requiring therapy for 6 months</td><td>0.78 / 0.78</td><td>14%</td></tr>
<tr><td>Cerebral vascular accident ever or resection (for causes other than malignancy)</td><td>0.63 / 0.69</td><td>28%</td></tr>
<tr><td>Cerebral vascular accident ever or resection >1</td><td>0.57 / 0.62</td><td>1%</td></tr>
<tr><td>Cranial or peripheral neuropathy</td><td>0.7 / 0.7</td><td>31%</td></tr>
<tr><td>Transverse myelitis</td><td>0.52 / 0.76</td><td>3%*</td></tr>
</table>
</td></tr>
<tr><td>Ocular</td><td>0.97</td><td>0.99</td><td>Same as Y2</td><td>assumed 10%</td><td>Section 9, Appendices 25 and 26</td><td>Weighted average of :
<table>
<tr><th>Item</th><th>utility Y1/Y2</th><th>weight</th></tr>
</table>
</td></tr>
</table>

Organ Damage	Disutilities Year			SD	Reference	Assumption/justification
						Cataract: 0.98 / 1: 78% Retinal damage / optic and trophy: 0.97/ 0.97: 22%*
Peripheral vascular	0.86	0.92	Same as Y2	assumed 10%	Section 9, Appendices 25 and 26	Weighted average of: Item: utility Y1/Y2: weight Claudication x 6 months: 0.79 / 1: 26% Minor tissue loss (pulp space): 1 /1: 12% Significant tissue loss ever (e.g. loss of digit or limb) (Score 2 if > one site): 0.64 / 0.64: 17% Significant tissue loss > 1 site: 1 / 1: 0% Venous thrombosis with swelling, ulceration or venous stasis: 0.99 / 0.99: 46%"
Premature gonadal failure	1	1	1		Section 9, Appendices 25 and 26	No disutility multiplier considered
Pulmonary	0.69	0.69	Same as Y2	assumed 10%	Section 9, Appendices 25 and 26	Weighted average of: Item: utility Y1/Y2: weight Pulmonary hypertension: 0.61 / 0.61: 33% Pulmonary fibrosis: 0.73 / 0.73: 42% Shrinking lung (on chest radiograph: 1 / 1: 2% Pleural fibrosis (on chest radiograph): 1 / 1: 20% Pulmonary infarction or resection: 0.94 / 0.94: 4%
Renal	0.97	0.96	Over time, the proportion ESRD increases. However, also proportion (successful) transplant increases	assumed 10%	Section 9, Appendices 25 and 26	Renal consisted of: Not in –ESRD: 1 Having ESRD Utility: Dialysis: 0.57 Graft transplant: 0.81 Functioning graft (immunosuppression): 0.81 Graft rejection: 0.57
Skin	0.94	0.94	Same as Y2	assumed 10%	Section 9, Appendices 25 and 26	Weighted average of: Item: utility Y1/Y2: weight

<table>
<tr><td>Organ Damage</td><td colspan="3">Disutilities Year</td><td>SD</td><td>Reference</td><td colspan="3">Assumption/justification</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td colspan="3"><table>
<tr><td>Scarring chronic alopecia</td><td>0.93</td><td>47%</td></tr>
<tr><td>Extensive scarring or panniculum other than scalp and pulp space</td><td>0.97</td><td>36%</td></tr>
<tr><td>Skin ulceration (not due to thrombosis) for more than 6 months</td><td>0.97</td><td>17%</td></tr>
</table></td></tr>
<tr><td>State</td><td colspan="3">Utility Value</td><td></td><td>Reference</td><td colspan="3">Assumption/justification</td></tr>
<tr><td>Baseline Utility</td><td colspan="3">0.63 (example A)
0.67 (example B)</td><td>-</td><td>Section 6.4.13 and 6.4.15</td><td colspan="3">For example: A: for a black African SLE patient, aged 40 years at entry with a SS score of 10

B: for a caucasian patient, aged 40 years at entry with a SS score of 10</td></tr>
</table>

** Exponentiated to the average number of damage items for patients with damage in that system.*

As described in Section 6.4.3, the EQ-5D values collected in the BLISS trials have been used in the health economic model to estimate impact of disease activity on HRQL and effect of belimumab treatment on these outcomes (see more detail in Section 6.4.16 below). However these utility values will not cover the impact from the accumulated effect of experiencing organ damage over the longer term which has a substantial impact on HRQL for SLE patients. In addition, due to the short duration of the trials the study utility weights cannot account for the premature mortality that SLE patients are subject to. Therefore it was necessary to identify other data sources in order to provide estimates of utility weights for long-term outcomes of the disease. The derivation of these utility weights for long-term outcomes is described in Section 6.4.18.

6.4.11 If clinical experts assessed the applicability of values available or estimated any values, please provide the following details[4]:

- the criteria for selecting the experts

[4] Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee.

- the number of experts approached
- the number of experts who participated
- declaration of potential conflict(s) of interest from each expert or medical speciality whose opinion was sought
- the background information provided and its consistency with the totality of the evidence provided in the submission
- the method used to collect the opinions
- the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)
- the questions asked
- whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).

No clinical experts provided estimates for utility values incorporated into the health economic model.

6.4.12 Please define what a patient experiences in the health states in terms of HRQL. Is it constant or does it cover potential variances?

This is discussed in Section 6.4.2. For the population of SLE patients with high disease activity of particular interest to this decision problem, HRQL is not constant over time; it varies depending on the presence or absence of disease flares and level of disease activity and is likely to deteriorate over time at a rate dependant on the frequency and type of organ damage experienced.

6.4.13 Were any health effects identified in the literature or clinical trials excluded from the analysis? If so, why were they excluded?

It is believed that the symptom of chronic fatigue, often cited as a debilitating symptom by SLE patients is not sufficiently captured by the EQ-5D. This has also been discussed by the NICE Decision Support Unit in their report ‘The incorporation of health benefits in cost utility Analysis using the EQ-5D’ (Wailoo et al. 2010). The FACIT-Fatigue questionnaire was collected in the BLISS studies and showed sustained improvement with belimumab in the FACIT-Fatigue score

from 8 to 52 weeks. However, there was no straightforward method of incorporating these data into the health economic model.

6.4.14 If appropriate, what was the baseline quality of life assumed in the analysis if different from health states? Were quality-of-life events taken from this baseline?

The baseline quality of life assumed in the cost-effectiveness analysis was determined by the following regression equation:

Where age = current age of patient, black is 1 if a patient is of black African ethnicity, or 0 otherwise, and SS = SELENA-SLEDAI score during the particular model yearly cycle. Derivation of this equation is detailed in Section 6.4.16. The baseline utility for each patient was adjusted for the impact of any organ damage experienced and the methodology for this is also discussed in Section 6.4.16.

6.4.15 Please clarify whether HRQL is assumed to be constant over time. If not, provide details of how HRQL changes with time.

HRQL is not assumed to be constant over time. This is due to the nature of SLE which comprises unpredictable flares of the disease, and which over time, for some of the more severe SLE patients with high disease activity (of interest to this decision problem), can lead to worsening morbidity due to accumulated organ damage. Therefore HRQL will be worse during flares of the disease and over time will deteriorate if associated with organ damage.

6.4.16 Have the values in sections 6.4.3 to 6.4.8 been amended? If so, please describe how and why they have been altered and the methodology.

A statistical model was estimated including all baseline variables (i.e. baseline characteristics, organ damage and organ involvement) that were included in the health economic model, listed in Table 6.20. This linear regression was made with the linear mixed effects package in R, correcting for the multiple observations per patient. This analysis included 1,125 patients with 9,051 EQ-5D

measurements from BLISS-52 and BLISS-76. Only those variables that had a p-value of less than 0.05 were selected for the final model. Based on this selection criterion the following variables were kept: age, black ethnicity, SS score and damage in ocular, neuropsychiatric (NP), musculoskeletal (MSK) and diabetes organs. Six out of eight involvement variables also reached significance. However, as a strong correlation was seen between SS score and involvement, the latter variables were excluded. This can also be seen for the coefficient of SS which decreases from -0.015 (full model) to -0.009 (reduced model).

Table 6.20. Linear regression explaining utility value based on pooled BLISS-52 and BLISS-76 studies (patient level trial analysis)

	Full model			Reduced Model (used in CE-model)	
	Coefficent	p-value		Coefficent	p-value
Constant	1.220	<0.001		1.297	<0.001
Log of age (yrs)	-0.114	<0.001		-0.145	<0.001
Sex (female)	-0.034	0.175			
Black Ethncitiy	-0.049	0.020		-0.054	0.012
SELENA-SLEDAI score	-0.015	<0.001		-0.009	0.001
Ocular	0.059	0.010		0.065	0.005
NP	-0.071	<0.000		-0.078	0.001
Renal	0.048	0.210			
Pulmonary	-0.020	0.565			
CV	-0.002	0.925			
PV	-0.011	0.691			
GI	-0.019	0.473			
MSK	-0.059	0.001		-0.062	<0.001
Skin	0.013	0.579			
GF	-0.042	0.301			
Diabetes	-0.084	0.008		-0.090	0.005
Malignancy	0.037	0.599			
Increased DNA-binding	0.046	<0.001			
Vasculitis	0.066	<0.001			
CNS	0.082	0.002			
Renal	0.050	<0.001			
Serositis	-0.045	0.001			
Haemo	0.005	0.658			
Other	-0.006	0.574			
Low Complement	0.047	<0.001			
Skin	0.020	0.003			

CV = cardiovascular, MSK = musculoskeletal, NP = neuropsychatric, PV = peripheral vascular, GI = gastrointestinal, GF = Gonadal Failure, Organ damage was included in the regression in order to estimate the utility of a patient without damage*

In the final model it can be seen that a person's utility value decreases, on average by 0.009 per SELENA-SLEDAI point. Ocular damage was observed to have a positive association with quality of life which is not a plausible finding. Also, for other damage items (i.e. renal, pulmonary, CV, PV, GI, skin, GF and malignancy) no statistically meaningful relationships with quality of life were found. However this does not imply that in reality there is no quality of life

impairment associated with these (sometimes severe) organ damage manifestations. The average SDI score for patients entering the BLISS trials was 0.76 and 59% of patients had no organ damage, 23% a SDI score of 1 and only 18% a SDI score more than 1. Taking into account that there are 12 organ systems, with a total of 43 different items it is difficult to establish quality of life impairments associated with organ damage based on the BLISS trial data as the patients on the whole did not report significant organ damage.

To reduce complexity in calculating utilities due to all types of organ damage, the regression analysis detailed above was used to determine a patient's 'clean' utility (U), i.e. free of damage items, using the following equation:

Disutilities associated with each type of organ damage were then applied to this "clean" utility in the model if a patient had developed organ damage in each model cycle. The preferred form of the literature utility data was the EQ-5D time trade off (TTO) methodology to be consistent with the NICE reference case wherever possible, indexed for the UK population. Since the model predicts damage in each of the 12 SLICC organ systems, weighted averages were constructed from the items for each of the organ systems (in which the weight was determined as the number of events for each item divided by the total number of events attributable to the organ system in the JH cohort data). To account for the fact that patients may have damage to more than one item of an organ system, the utility was exponentiated to the average number of affected items for that organ system.

For example, Ocular damage, Year 1:

Ocular damage consists of 78% cataract (utility 0.978) and 22% retinal damage (utility 0.974); however, some patients have both retinal damage and cataract. The average SLICC score for ocular damage, for patients with ocular damage, is 1.23. The calculated utility for ocular is therefore:

(0.78 * 0.978 + 0.22 * 0.974) ^1.23 = 0.9719

To acknowledge the decreased influence of added organ damage in patients with existing damage, utility multipliers were used instead of utility decrements. For conditions where only utility decrements were reported, the multiplier was calculated from a fraction with the baseline utility value as the denominator and the baseline minus the utility decrement as the numerator. The disutility multiplier for the first and second year with damage to an organ system are shown in Table 6.19 above, disutility multipliers after the second year can be found in Section 9.26, Appendix 9.26.

A patient's 'clean' utility predicted with the regression equation from Table 6.20 was multiplied with the lowest disutility multiplier from the organ damage systems a patient had developed. Using all disutility multipliers would underestimate a patient's utility value (i.e. attribute a lower utility value); our approach is conservative as it will overestimate the utility value.

6.5 Resource identification, measurement and valuation

This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', section 5.5.

All parameters used to estimate cost effectiveness should be presented clearly in a table and include details of data sources. For continuous variables, mean values should be presented and used in the analyses. For all variables, measures of precision should be detailed.

NHS costs

6.5.1 Please describe how the clinical management of the condition is currently costed in the NHS in terms of reference costs and the payment by results (PbR) tariff. Provide the relevant Healthcare Resource Groups (HRG) and PbR codes and justify their selection. Please consider in reference to section 2.

Costs in the analysis are limited to direct medical costs and are associated with disease activity and long-term organ damage. Costs related to disease activity were drawn from an analysis conducted in 2009 on the resource utilisation recorded in the one-year belimumab Phase 2 trial (LBSL02) in which 2005/06 NHS reference costs were used. The methods used to calculate the disease activity costs from this study are outlined below. These costs have been inflated to 2010 costs using an inflator based on the annual Consumer Retail Indices (CPI0) (OECD 2010a).

Costs related to disease activity

The following resource utilisation was collected in the Phase 2 study:

- Number of surgeries or procedures
- Number of Accident and Emergency attendances
- Number of days in a nursing home or rehabilitation centre
- Number of overnight hospitalisations
- Length of stay in hospital
- Number of visits to health professionals

• Number of tests or diagnostic procedures

The NHS reference unit costs for these types of resource are presented in Table 6.21.

Table 6.21. NHS Reference Costs - HRGs relevant to the management of SLE

Healthcare Costs		
Unit type	Cost per unit	Source
Accident and Emergency Visit	£75	NHS reference costs 2005-06 (Department of Health 2006) Cost inflated from 2006 to 2010 using CPI. (CPI 2006 = 102.334; CPI 2010 = 114.485) (OECD 2010a)
Surgeries	£96	
Inpatient stay (per night)	£220	PSSRU Unit Costs for Social Care 2007 (PSSRU 2007) and inflated to 2010 costs using CPI (OECD 2010a).
Nursing home (per night)	£84	
Diagnostic Tests and Procedures		
Unit type	Cost per unit	Source
X-rays	£46	NHS reference costs 2005-06 (Department of Health 2006) Cost inflated from 2006 to 2010 using the CPI (OECD 2010a).
Bone Scan	£49	
Ultrasound	£63	
EKG	£22	
MRI	£154	
Treadmill/Stress test	£61	
CT scan	£105	
Unit type	Cost per unit	Source
Blood tests	£2	
Urine Tests	£2	
Endoscopy	£274	
Colonoscopy	£274	
Sigmoidoscopy	£274	
Other Tests	£68	
Mammograms	£32	Obtained from Radiology Indicative Tariff (Top Talk for Radiography Leaders 2005).

In addition to the resource utilisation costs detailed above, Table 6.22 details all healthcare professional costs which have been included in the disease activity cost calculations.

Table 6.22. Professional unit costs

Health Care professionals	UK Cost per unit	Source
Internist	£27	PSSRU Unit Costs for Social Care 2007 (PSSRU 2007) and inflated to 2010 costs using the CPI (OECD 2010a).
General Practitioner (Per surgery consultation lasting 11.7 minutes)	£34	
Nurse Practitioner	£14	
Podiatrist	£27	
Other Health Care Workers	£14	
Nephrologist	£117	NHS reference costs 2005-06 (Department of Health 2006) and inflated to 2010 costs using the CPI (OECD 2010a).
Rheumatologist	£105	
Dermatologist	£57	
Gastroenterologist	£80	
Obstetrics and Gynaecology	£76	
Urologist	£79	
Chiropractor	£56	
Physical/occupational therapist	£56	
General or Orthopaedic Surgeon	£80	
Other Doctors	£96	

In the Phase 2 study, SELENA-SLEDAI (SS) score was reported every 28 days for the first six visits then either every 28 or 56 days for the remainder of the study. However, the resource use questionnaire was only recorded at baseline, day 168 and withdrawal from the study. Total resource use per patient was calculated over two periods between days 0-168 and days 168-365.

In order to estimate costs related to SS score it was necessary to define the severity of disease activity for the patient over the resource use period. The most severe point of disease activity over the 6 months was considered to be the most important determinant of resource use. The maximum SS score was identified as a proxy for disease activity over the period 0-168 days and 168-364 days. In the

analysis conducted in 2009 SS scores were grouped according to none, mild, moderate and severe categories. A linear regression analysis was conducted to explore the relationship between total resource use costs over a six-month period and the severity of disease activity using the SS score. A robust cluster model was employed to account for likely correlation in the observations of the same patient. The equation for the regression is:

$$= \beta + \beta \qquad\qquad + \varepsilon$$

where SLEDAISEVERITY was incorporated as a continuous variable taking the values of:

0 for SS score values of 0 (representing no disease activity)

1 for SS score values of 1 to 4 (representing mild disease activity)

2 for SS score values of 5 to 12 (representing moderate disease activity)

3 for SS score values of 12 or more (representing severe disease activity)

No other explanatory variables were included because it was only necessary to look at the relationship between disease activity and cost. The results from this regression analysis are presented in Table 6.23 below and suggest that there is a positive and statistically significant relationship between the SS score and resource use cost. The R^2 in the model is very low, however this is not surprising as many other variables such as age, co-morbidities, concomitant medications etc. will have a substantial impact on resource use.

Table 6.23. Results of the regression of costs on SS score

	Model Results	
Observations	457	P-value
SELENA-SLEDAI Severity Coefficient	149.09	0.018
Constant Coefficient	515.06	<0.001
R^2	0.01	

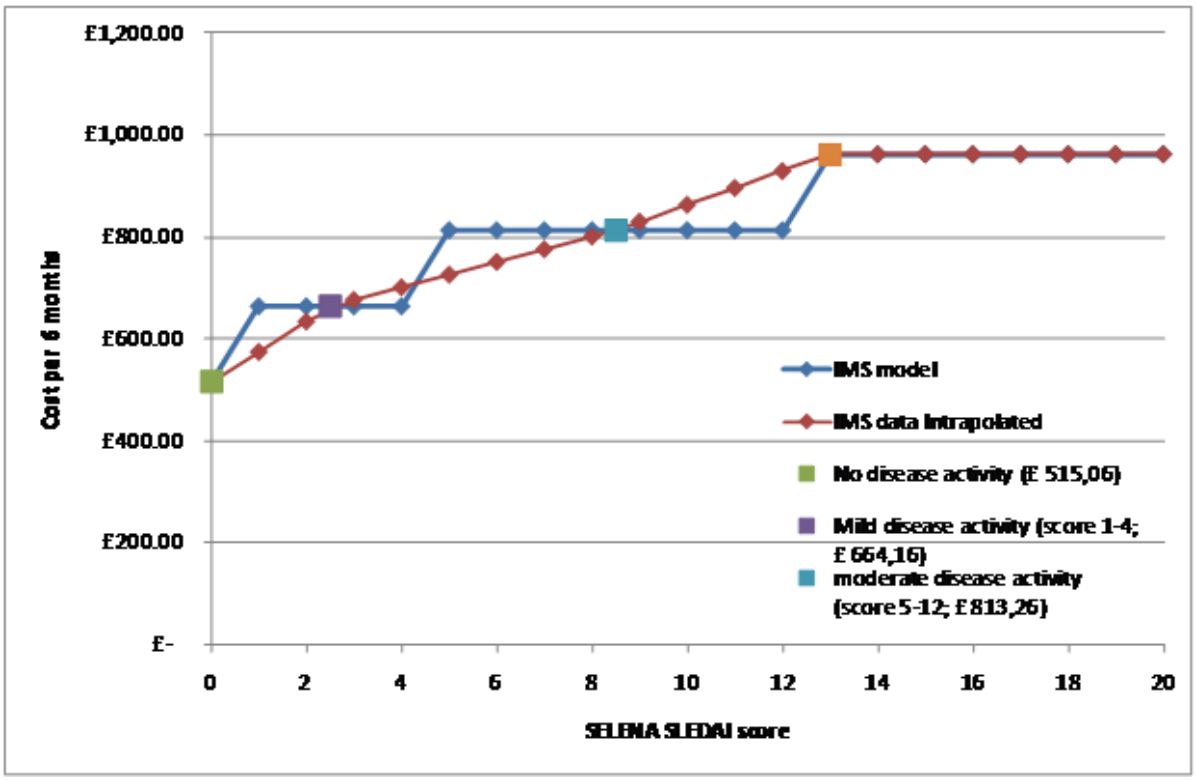
From this regression the estimated total six-month direct costs by disease activity level are summarised in Table 6.24 below.

Table 6.24. Summary of 6-month direct costs from the Phase 2 Study LBSL99

Disease Activity Level	Total Costs
No disease activity	£515.06
Mild disease activity (score 1-4;)	£664.16
Moderate disease activity (score 5-12;)	£813.26
Severe disease activity (score > 12)	£962.36

Since the model for our decision problem uses the SS score and not a SS severity category the cost values for each of the four SS score categories were interpolated in the following way (Figure 6.16)

Figure 6.16. Estimated costs by SS Score – Phase 2 Study LBSL99



These costs were then multiplied by 2 to scale up to one year, resulting in the scores presented in Table 6.25. These disease activity related costs were applied to each year the patient remained in the model based on their SS score for each year of the simulation.

Table 6.25. Association between SS score and yearly costs implemented in the model

SELENA-SLEDAI Score	Yearly Costs
0	£1,152.44
1	£1,285.87
2	£1,419.30
3	£1,513.84
4	£1,569.44
5	£1,625.04
6	£1,680.64
7	£1,736.23
8	£1,791.83
9	£1,856.72
10	£1,930.85
11	£2,004.98
12	£2,079.11
13	£2,153.26
14	£2,153.26
15	£2,153.26
16	£2,153.26
17	£2,153.26
18	£2,153.26
19	£2,153.26
20	£2,153.26

Organ damage costs

Organ damage has a potentially substantial effect on health care utilisation. Corresponding costs were obtained by conducting a literature search for each of the 41 damage items in the SLICC score. Similar to the utility search, the scope of this search was almost unbounded. It was unfeasible to conduct a formal systematic review. Cost data for organ damage was searched in all relevant HTAs published on the NICE website and publications with data for the organs of interest on Pubmed. The full search strategy for cost data is described in Section 9.13, Appendix 13. For conditions where no UK data were available, values were transformed to UK pounds using Purchasing Power Parities for health (OECD 2010b). Costs were inflated to 2010 costs with the CPI. Where costs consisted mostly of drug costs, the costs were not inflated. For the situation where drug costs consisted of a small part of the total costs, the total costs were inflated.

This could produce a minor and irrelevant increase of the total costs. However, since all costs were conservative estimates, this is not likely to produce a relevant overestimation of total costs.

Average costs per damage system

Since the health economic model predicts damage in a specific organ system, instead of predicting damage in individual items, average costs were calculated for each of the 12 organ systems. These average costs were calculated by taking the weighted average of the costs of the items (in which the weight was the number of events divided by the number of patients with damage to that organ system in the JH cohort). Reflecting the fact that patients can have damage in more than one item of the organ system, the sum of weights can exceed 100%. The costs for the first and second year after development of the organ damage are shown in Table 6.26. Costs for subsequent years can be found in Section 9.27, Appendix 9.27. The calculations used to derive these costs are detailed in Section 9.25, Appendix 25. Due to the high costs associated with end stage renal disease (ESRD) and the deteriorating course of renal damage, the expected costs and disutilities for renal damage increase over time. For skin damage and gonadal failure, although very limiting for the patient, it was assumed in the model that they do not incur any ongoing medical costs. Furthermore it was assumed that there was no beneficial effect of belimumab on gonadal failure. Most other organs have higher costs in the year of damage development and lower costs in subsequent years (follow-up costs).

Similar to the quality of life data, cost data were collected for the year of the event, and for the years following the event. For some organs, specific data were available for the development of costs for the years after the event or organ damage. For the other organ systems, the costs were kept fixed after the second year. For organ damage with ongoing influence on cost (e.g. diabetes), the costs were kept constant.

Table 6.26. Costs for organ damage in the first and second year after initial damage development

Organ damage Type	Year 1	Year 2
Cardiovascular	£3,440	£505
Diabetes	£2,338	£2,338
Gastrointestinal	£2,708	£0
Malignancy	£6,123	£0
Musculoskeletal	£5,431	£1,903
Neuropsychiatric	£3,660	£1,144
Ocular	£1,535	£17
Peripheral vascular	£2,988	£598
Premature gonadal failure	£0	£0
Pulmonary	£9,679	£9,603
Renal	£1,765	£2,453
Skin	£0	£0

6.5.2 Please describe whether NHS reference costs or PbR tariffs are appropriate for costing the intervention being appraised.

NHS reference costs have been used to determine an appropriate administration cost for delivering the belimumab infusion to the patient

Resource identification, measurement and valuation studies

6.5.3 Please provide a systematic search of relevant resource data for the UK. Include a search strategy and inclusion criteria, and consider published and unpublished studies. The search strategy used should be provided as in section 9.13, appendix 13. If the systematic search yields limited UK-specific data, the search strategy may be extended to capture data from non-UK sources. Please give the following details of included studies:

- country of study
- date of study
- applicability to UK clinical practice
- cost valuations used in study
- costs for use in economic analysis
- technology costs.

As explained for HRQL, a formal systematic review of resource use was not conducted. Instead a literature review of previous NICE HTAs was carried out. If the required information was unavailable from the NICE website, additional searches were carried out on Pubmed. A description of the search process conducted can be found in Section 9.13, Appendix 13.

6.5.4 If clinical experts assessed the applicability of values available or estimated any values, please provide the following details[5]:

- the criteria for selecting the experts
- the number of experts approached
- the number of experts who participated
- declaration of potential conflict(s) of interest from each expert or medical speciality whose opinion was sought
- the background information provided and its consistency with the totality of the evidence provided in the submission
- the method used to collect the opinions
- the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)
- the questions asked
- whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).

No clinical experts were used to estimate any values for resource use or costs

Intervention and comparators' costs

6.5.5 Please summarise the cost of each treatment in the following table. Cross-reference to other sections of the submission; for example, drugs costs should be cross-referenced to sections 1.10 and 1.11.

[5] Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee.

Provide a rationale for the choice of values used in the cost-effectiveness model discussed in section 6.2.2.

The average costs of standard of care treatment is substantially lower than the cost of belimumab treatment. As belimumab is given in addition to standard of care, it is assumed that the costs for standard of care treatments are negligible and will have little impact on the cost-effectiveness results, even taking into consideration any prolonged life expectancy predicted with belimumab. However, including a cost for standard of care has been investigated as a scenario analysis (see Section 6.6.1). Therefore the cost-effectiveness analysis for the base case only considers the additional acquisition costs for belimumab in the belimumab arm. Table 6.27 summarises all the costs associated with the administration of belimumab. The administration cost of £126 for belimumab was calculated based on two hours of senior hospital staff nurse time (£63/hr) from PSSRU Unit Costs of Health and Social Care 2010. Two hours is considered appropriate due to one hour required for the actual infusion and another hour for patient preparation and monitoring post-infusion. An alternative method of determining an infusion administration cost is to use the day case costs for “Inflammatory Spine, Joint or Connective Tissue Disorders without complications” (HRG=HD23C) from the NHS tariff costs 2009/10, which is £432 per day. Adjusting this cost to obtain an estimated cost for two hours gives £115 (i.e. £432 per day/7.5*2). The highest cost of these two methods has been used in the model for each administration of the infusion i.e. £126.

Table 6.27. Unit costs associated with the technology in the economic model

Items	Belimumab 10mg/kg	Ref. in submission
Technology cost belimumab 10 mg/kg		
Mean cost of technology treatment based on an average weight of 67.3 kg as seen in the pooled total population	Year 1 annual cost = £9,394 Year 2 Annual cost = £8,723	The model currently uses vial costs of £114.30 and £381 for 120 mg and 400 mg, respectively. For each weight, the optimal vial combination is chosen and costs for waste are added. Weight distribution according to the trials is used to determine average yearly belimumab costs.
Administration cost per infusion	£1764 (Year 1) £1638 (Year 2+)	£126 per infusion (14 in Year 1 and 13 in Year 2), see section 6.5.1
Monitoring and test cost	£0	No additional monitoring or tests are required for implementation of this technology
Total Year 1 costs	£11,158	
Total Subsequent Year costs	£10,361	
Mean cost of technology treatment based on an average weight of 65.4 kg as seen in the pooled BLISS study "high disease activity" subgroup	Year 1 annual cost = £9,154 Year 2 annual cost = £8,500	The model currently uses vial costs of £114.30 and £381 for 120 mg and 400 mg, respectively. For each weight, the optimal vial combination is chosen and costs for waste are added. Weight distribution according to the trials is used to determine average yearly belimumab costs.
Administration cost per infusion	£1764 (Year 1) £1638 (Year 2+)	£126 per infusion (14 in Year 1 and 13 in Year 2), see section 6.5.1
Monitoring and test costs	£0	No additional monitoring or tests are required for implementation of this technology
Total Year 1 costs for	£10,918	
Total Subsequent Year costs	£10,138	

Health-state costs

6.5.6 Please summarise, if appropriate, the costs included in each health state. Cross-reference to other sections of the submission for the resource costs. Provide a rationale for the choice of values used in the cost-effectiveness model. The health states should refer to the states in section 6.2.4.

As the model does not include health states, costs have been presented in terms of short-term disease activity related costs and long-term organ damage costs. Please see section 6.2.4.

Adverse-event costs

6.5.7 Please summarise the costs for each adverse event listed in section 5.9 (Adverse events). These should include the costs of therapies identified in section 2.7. Cross-reference to other sections of the submission for the resource costs. Provide a rationale for the choice of values used in the cost-effectiveness model discussed in section 6.2.2.

Adverse events (AEs) were not included in the health economic model. As discussed in Section 5.9.2, the Phase 2 and 3 studies did not find important differences in the incidence of all AEs and serious adverse events (SAEs) between the belimumab and placebo treatment groups. Importantly, the incidence of serious infections such as pneumonia, UTI, cellulitis, bronchitis, and pyelonephritis, which would require treatment in hospital and thus incur a significant cost to the NHS, was not significantly higher in the belimumab treatment arms compared with placebo. Therefore, by not including AEs in the model, it is expected that this would not have an important impact on the cost-effectiveness results.

Miscellaneous costs

6.5.8 Please describe any additional costs that have not been covered anywhere else (for example, PSS costs). If none, please state.

Due to the method of incorporating costs into the health economic model, PSS costs have been discussed in Section 6.5.1.

6.6 Sensitivity analysis

This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 5.1.11, 5.8, and 5.9.4 to 5.9.12.

Sensitivity analysis should be used to explore uncertainty around the structural assumptions used in the analysis. Analysis of a representative range of plausible scenarios should be presented and each alternative analysis should present separate results.

The uncertainty around the appropriate selection of data sources should be dealt with through sensitivity analysis. This will include uncertainty about the choice of sources for parameter values. Such sources of uncertainty should be explored through sensitivity analyses, preferably using probabilistic methods of analysis.

All inputs used in the analysis will be estimated with a degree of imprecision. Probabilistic sensitivity analysis (PSA) is preferred for translating the imprecision in all input variables into a measure of decision uncertainty in the cost effectiveness of the options being compared.

For technologies whose final price/acquisition cost has not been confirmed, sensitivity analysis should be conducted over a plausible range of prices.

6.6.1 Has the uncertainty around structural assumptions been investigated? Provide details of how this was investigated, including a description of the alternative scenarios in the analysis.

Uncertainty around structural assumptions has been examined using both one-way and probabilistic sensitivity analysis within the base population and sub-group population of interest to this decision problem.

To test the robustness of model assumptions and parameters, the effect of changing parameters in one-way sensitivity analyses was examined. Effects of varying individual parameters were examined using 95% confidence intervals. Sensitivity results for each input were ranked from most sensitive to least sensitive and those that had the greatest effect were plotted on tornado diagrams.

Analysed parameters, their base-case values, and ranges (upper and lower bounds) are presented in Section 9.24, Appendix 24.

Scenario analyses

A number of alternative scenario analyses among this population have also been conducted and these are detailed below.

- Inclusion of a more stringent responder criterion of reduction in SS score of ≥6. The American College of Rheumatology (ACR) Response Criteria for Systemic Lupus Erythematosus Clinical Trials (ACR Ad Hoc Committee on Systemic Lupus Erythematosus Response Criteria 2004), published in 2004, provides evidence that a reduction of 6 in SLEDAI represents a clinically meaningful improvement in a patient's SLE condition which is noticeable by the treating physician. In this study, 88 international experts rated vignettes of various patient clinical presentations with SLE as worsened, improved, or unchanged relative to the previous visit. These ratings were transformed by statistical procedures into performance characteristic curves that related a change on a particular SLE activity measure, one of which was the SLEDAI, to the physicians' agreement. The ACR committee members then voted on what level of expert agreement would be used to determine clinically meaningful change.

- The effect of excluding the treatment continuation rule in the model has been examined to demonstrate the requirement for a continuation rule for implementing this treatment in clinical practice in order to target the treatment to those that gain the greatest benefit and to reflect the wording in the draft SPC.

- As the vial price for belimumab has not been finalised, although the expected vial list price has been used in the base case analyses, a maximum expected vial price for both the 120mg and 400mg vials has also been investigated in a scenario analysis. These maximum prices are £127.80 for the 120mg vial and £426.00 for the 400mg vial.

- Use of unadjusted (original) natural history model for change in average yearly SS score. This model was adjusted in the base case to better reflect the average SS scores for the patient population of interest to this decision problem by increasing the constant value in the statistical model (see details under “Long-term SELENA-SLEDAI score” in Section 6.3 of this document). A separate analysis was run where this constant was unchanged and the model is identified as the “original natural history” model.

- Vial wastage: belimumab comes in two different vial sizes 120mg and 400mg and combinations of them will minimise wastage for most patients’ weights. As belimumab is an IV biologic drug, some wastage is expected for patients whose weight requires them to receive only a part of a vial. For example, a patient weighing 70kg will require 700 mg at the licensed dose of 10 mg/kg. The full dose can be given by combining either a) two 400mg vials giving a total of 800mg, resulting in 100mgs being wasted, or b) combining one 400mg vial with three 120mg vials, totaling 760mg, resulting in 60mgs being wasted, if no vial sharing is assumed. Consequently the annual average cost will be increased through wastage. The base case assumes that wastage will occur, to provide a conservative cost-effectiveness estimate. However if the optimum vial combination was chosen for each patients according to their weight in order to minimize drug wastage, there will be the opportunity for cost savings. This was explored in a scenario analysis.

- A different administration cost of £159 has been used in a scenario analysis, as this was suggested by the ERG who reviewed the STA appraisal for tocilizumab, a human monoclonal antibody for the treatment of rheumatoid arthritis which also requires administration over one hour.

- A scenario investigating the effect of increased standard of care costs on the belimumab arm compared with the SoC arm due to extended life expectancy with belimumab has been investigated. As mycophenolate mofetil (MMF) is the most expensive SoC treatment which can incur an annual cost of around £3,433 (MMF costs £87.33 for 50x500mg tablets (Wilson et al. 2007)), this

was used in the scenario analysis. The BLISS trials showed that 11.2% of the pooled total BLISS population was receiving this at study entry, so in the model the disease activity costs were increased by £385 (i.e. 11.2% of £3,433). This would enable the additional SoC cost from prolonged life expectancy with belimumab to be included in the total accrued costs.

Different maximum belimumab treatment durations were not considered relevant for scenario analysis as it is expected that patients with SLE comprising high disease activity would continue to take belimumab for as long as it was perceived by their treating physicians to offer them clinical benefit. Stopping belimumab would lead to the benefits of inhibiting the biological activity of BLyS also being curtailed and any beneficial reduction in disease activity. There is no current evidence to demonstrate whether limited durations of treatment, e.g. 5 or 7 years, would still result in clinically important long-term benefits on organ damage and survival.

6.6.2 Which variables were subject to deterministic sensitivity analysis? How were they varied and what was the rationale for this? If any parameters or variables listed in section 6.3.6 (Summary of selected values) were omitted from sensitivity analysis, please provide the rationale.

To understand the extent that changes in each parameter affected the incremental cost per QALY a univariate sensitivity analysis was performed on both the base case total BLISS population and the subgroup population**.** For the variables outlined below, each individual parameter was set to the corresponding lower and upper 95% confidence interval (CI) value (or +/- 20% of the mean where appropriate) and the simulation was run. Incremental QALYs, incremental costs and the incremental cost effectiveness ratios (ICERs) were calculated and presented in tornado diagrams. The corresponding lower and upper values for all parameters can be found in Section 9.24, Appendix 24.

1. **Coefficients of week 52 change in SS score regressions**
   Using a separate multivariate normal distribution with the covariance matrix obtained from the regression analysis (see Section 9.22, Appendix 22).

2. **Coefficients of natural history model for change in SELENA-SLEADAI score**

   Using a multivariate normal distribution with the covariance matrix obtained from the regressions (see Section 9.22, Appendix 9.22)

3. **Natural discontinuation**

   The yearly natural discontinuation probabilities are varied by simulating the yearly hazard from a normal distribution such that they reflect the uncertainty according to the Kaplan-Meier curve.

4. **Coefficients of natural history models for mortality and organ damage development**

   For each organ system and for mortality, the coefficients are drawn from a multivariate normal distribution with coefficients as presented in Table 6.14 and covariance matrices as presented in Section 9.22, Appendix 22.

5. **Standardised mortality ratios from Bernatsky**

   According to the uncertainty reported by Bernatsky et al, using a normal distribution with standard deviation equal to the maximum difference between mean and lower and upper value divided by 1.96.

6. **Coefficients for BLISS utility regression**

   Using a multivariate normal distribution with the covariance matrix (see Section 9.22, Appendix 22).

7. **Costs associated with each SS score**

   A factor is simulated from a gamma distribution with mean 1 and standard deviation 0.1 (alpha = 96.04, beta = 0.01). The same number is multiplied with the annual costs associated with each SS score (see Table 6.25). This way, the costs vary approximately between 80% and 120% of the mean.

8. **Organ Damage Costs**

   For each organ system, a factor is simulated from a gamma distribution with mean 1 and standard deviation 0.1 (alpha = 96.04, beta = 0.01). This number is multiplied with the annual costs associated with each organ damage item (see Table 6.26). This way, the costs vary approximately between 80% and 120% of the mean. For organ systems with different costs per year, the same number is used.

9. Organ Damage Disutility

To ensure that simulated disutilities do not become larger than 1 or smaller than 0, the log of the mean disutilities are multiplied with a non-negative factor and reverted back to a disutility by taking the exponent of the product. The factor is drawn for each organ system independently from a gamma distribution with mean 1 and standard deviation 0.1 (alpha = 96.04, beta = 0.01).

6.6.3 Was PSA undertaken? If not, why not? If it was, the distributions and their sources should be clearly stated if different from those in section 6.3.6, including the derivation and value of 'priors'. If any parameters or variables were omitted from sensitivity analysis, please provide the rationale for the omission(s).

Probabilistic sensitivity analyses were performed by varying the same sets of model parameters detailed for the univariate sensitivity analyses, simultaneously 1000 times to understand the impact on the cost per QALY results. There was a large amount of correlation between coefficients within each regression (i.e. parameters from the regressions numbered 1, 2, 4, and 6 for the sensitivity analyses in Section 6.6.2 above). To account for this correlation the covariance matrices were generated and from these a set of PSA inputs were used. This process uses a multivariate normal distribution; a normal distribution was therefore assigned to these regressions in the PSA. The standardised mortality reported by Bernatsky et al (2006) was assumed to follow a normal distribution. The costs (i.e. from regressions 7 and 8 identified in the sensitivity analysis section) were assigned a gamma distribution as recommended by (Briggs et al. 2006).

Results of probabilistic sensitivity analyses are presented in the form of a scatter plot and a cost-effectiveness acceptability curve. Parameters included in the probabilistic sensitivity analysis, their base case values, and their assumed distribution, are presented in Section 9.24, Appendix 24.

6.7 Results

Provide details of the results of the analysis. In particular, results should include, but are not limited to, the following.

- Link between clinical- and cost-effectiveness results.
- Costs, QALYs and incremental cost per QALY.
- Disaggregated results such as LYG, costs associated with treatment, costs associated with adverse events, and costs associated with follow-up/subsequent treatment.
- A statement as to whether the results are based on a PSA.
- Cost-effectiveness acceptability curves, including a representation of the cost-effectiveness acceptability frontier.
- Scatter plots on cost-effectiveness quadrants.
- A tabulation of the mean results (costs, QALYs, ICERs), the probability that the treatment is cost effective at thresholds of £20,000–£30,000 per QALY gained and the error probability.

Clinical outcomes from the model

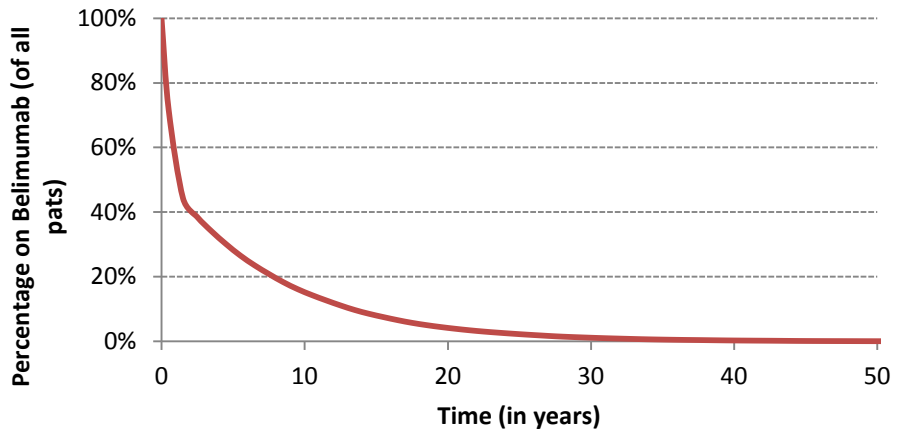
6.7.1 For the outcomes highlighted in the decision problem (see section 4), please provide the corresponding outcomes from the model and compare them with clinically important outcomes such as those reported in clinical trials. Discuss reasons for any differences between modelled and observed results (for example, adjustment for cross-over). Please use the following table format for each comparator with relevant outcomes included.

Base case analysis

For the base case analysis, the pooled total population has been modeled (excluding the belimumab 1mg arm) and the results for this population are presented in this section. This, however, is not considered the key population that we believe is most relevant for treatment with belimumab, which is the subgroup of SLE patients with the highest disease activity. The results for this high disease activity subgroup are presented in Section 6.9.

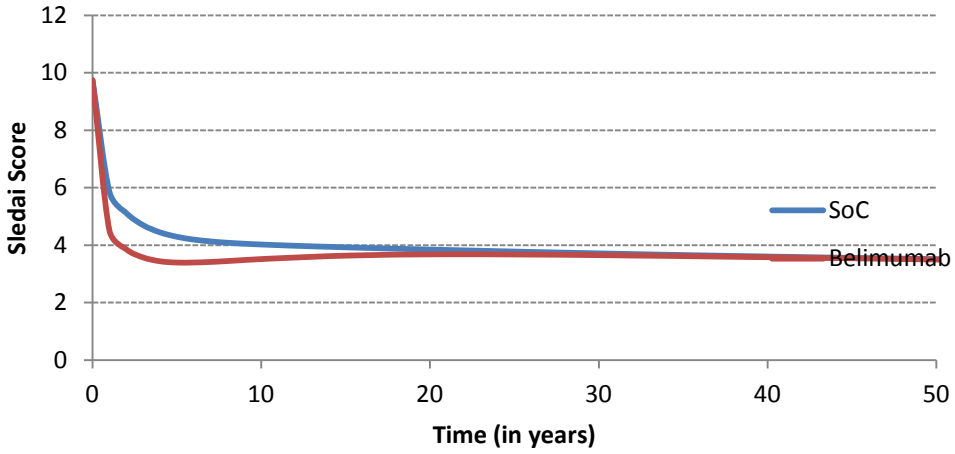
Figure 6.17 below presents the percentage of patients remaining on belimumab over time. The relatively steep drop observed during the first year is caused by non-responders discontinuing the drug**.** As a constant rate of discontinuation is applied for responders over time, and due to mortality, it is estimated in the model that approximately 15% of patients will be taking belimumab after 10 years. The constant rate of discontinuation is considered a conservative approach as it is likely that certain responders will remain responders over the longer term and continue to receive the benefit of being on belimumab.

Figure 6.17. Discontinuation from belimumab (includes death) – Total pooled population.



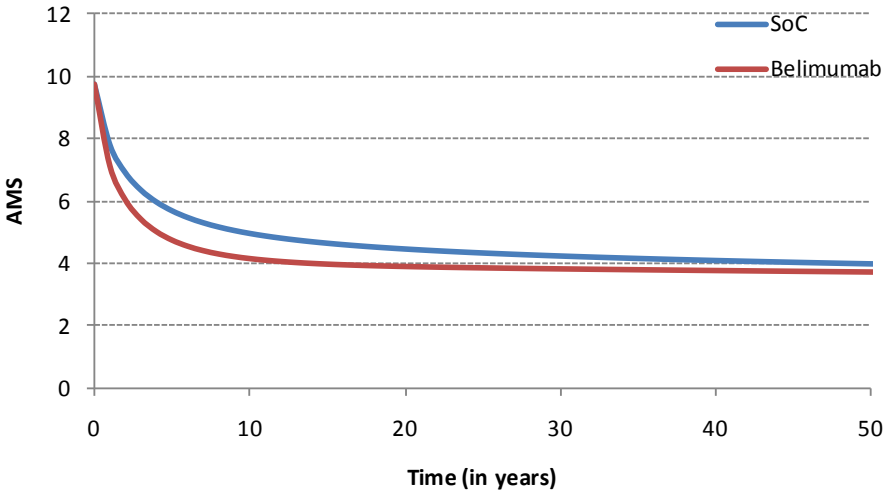
The most important SLE disease variable which we studied is the SS score. The average SLEDAI score (AMS) for 50,000 simulated patients is shown in Figure 6.18 for those patients who are still alive. It is clear from the graph that patients who are treated with belimumab (on top of SoC) have a greater reduction in SS score than patients who are treated with SoC alone. Over time, the difference between the SS score of belimumab and SoC patients declines because patients discontinue belimumab and subsequently lose its beneficial effect on disease activity.

Figure 6.18. SLEDAI Score over time (AMS) for 50,000 patients analysis – Total pooled population.



Although the activity returns to SoC levels (Figure 6.18), a beneficial effect is kept through a decreased average disease activity over time (Figure 6.19). The adjusted mean SLEDAI (AMS) is an important predictor of organ damage in the cardiovascular, renal and peripheral vascular systems (Table 6.14).

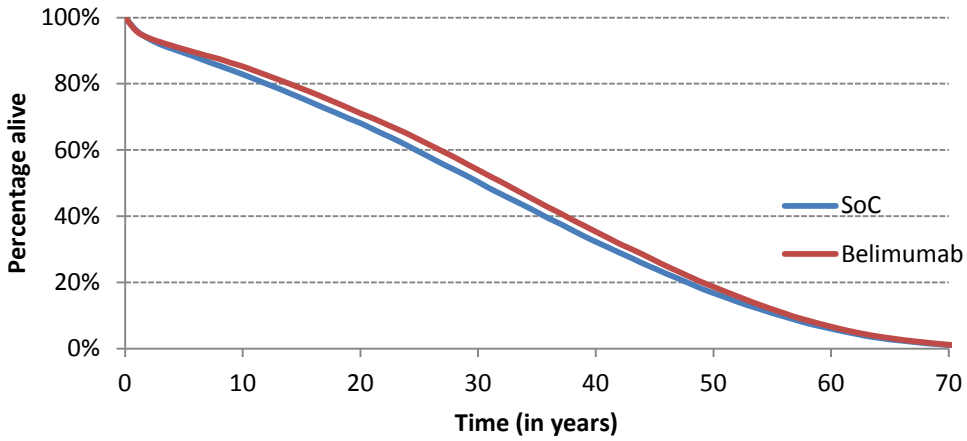
Figure 6.19. Adjusted Mean SLEDAI (AMS) over time censored for death



The lower disease activity for belimumab patients will lead to a decreased steroid dose and a decreased risk for organ damage. The average disease activity over time, cumulative average prednisone dose and organ damage, contribute to the

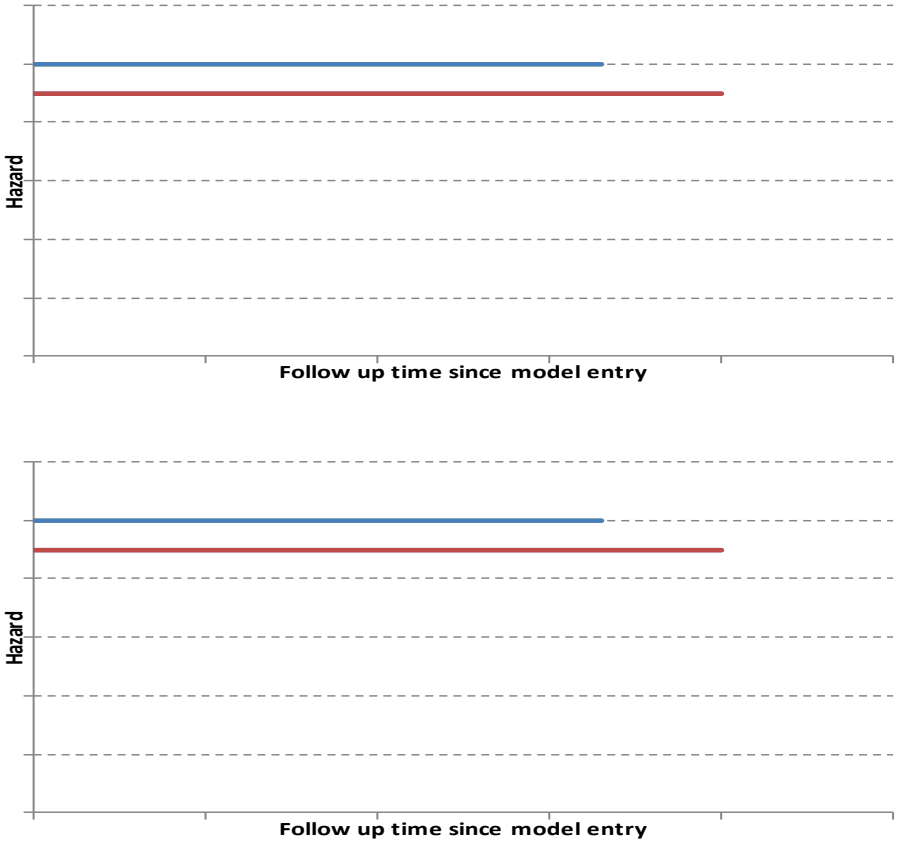
mortality risk (Table 6.12). The survival over time is therefore higher for belimumab patients than for patients on SoC (Figure 6.20). The relatively steep decline in survival in the first years for both arms is caused by the relatively high standardised mortality ratio for patients younger than 24 (see Table 6.13). The survival curve for belimumab shows a slightly less steep decline than for SoC, which is caused by the lower average disease activity scores.

Figure 6.20. Survival of patients over time – Pooled total population.



Belimumab influences the accrual of organ damage through reducing the average disease activity, and through also reducing steroid dose. However, since the modelling estimates that belimumab patients live longer, their exposure to the risk of organ damage is increased. The effect of belimumab on organ damage occurrences is therefore a balance between the prolonged exposure and the decreased risk through improved disease control. This is illustrated in the example in Figure 6.21 where the red line displays belimumab and the blue SoC. Although the hazard for developing organ damage with belimumab is lower in this picture, the area under the curve that defines lifetime risk of developing organ damage is higher due to prolonged exposure.

Figure 6.21. Hazard times exposure determines percentage of patients with organ damage – Total pooled population.



The balance between these two factors causes lower organ damage occurrences for some organ systems (e.g. cardiovascular and pulmonary) and higher occurrences for others (e.g. gastrointestinal and malignancy) (Table 6.28).

Table 6.28. Organ damage occurrence for SLE patients until death – Pooled total population.

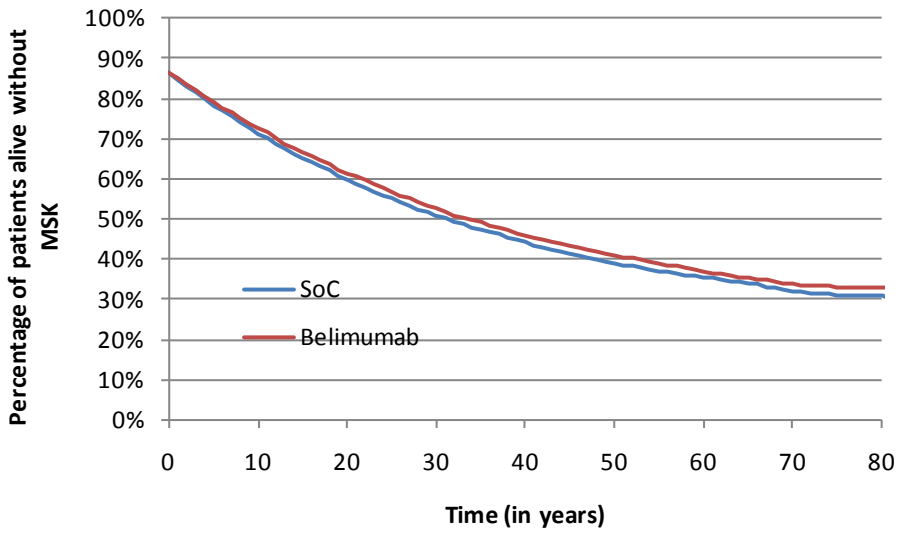
	SoC	Belimumab	Difference
Cardiovascular	23.5%	21.9%	-1.6%
Diabetes	19.2%	19.7%	0.5%
Gastrointestinal	23.3%	25.2%	1.9%
Malignancy	32.3%	33.8%	1.5%
Musculoskeletal	49.7%	50.3%	0.6%
Neuropsychiatric	45.3%	46.0%	0.7%
Ocular	36.7%	37.4%	0.7%
Peripheral vascular	19.8%	19.1%	-0.7%
Premature gonadal failure	7.5%	7.6%	0.1%
Pulmonary	38.2%	36.6%	-1.6%

Renal	18.0%	15.4%	-2.6%
Skin	7.5%	7.4%	0.0%

Fewer patients on belimumab develop damage for cardiovascular, peripheral vascular, pulmonary and renal organs, compared with SoC. This is explained by the lower average disease activity for patients on belimumab compared with SoC and the dependence of damage risk on disease activity (see Table 6.14). However, for diabetes, gastrointestinal, malignancy, musculoskeletal, neuropsychiatric, and ocular more belimumab patients develop damage than those on SoC; this is mainly caused by the estimated increased life expectancy for patients on belimumab.

It is notable that even though some organ damage depends on AMS e.g. gastrointestinal and ocular damage (see Table 6.12), and belimumab reduces disease activity (AMS), a higher total percentage of belimumab patients develop damage in these systems. This is likely due to the estimated extended life expectancy. However, the risk of developing damage in other organs is lower with belimumab, as illustrated in the Kaplan Meier plot for musculoskeletal development (see Figure 6.22). Kaplan-Meier curves for the other organs are shown in Section 9.23, Appendix 23.

Figure 6.22. Kaplan-Meier plot of proportion of patients alive without musculoskeletal damage - John Hopkins Cohort



Since the modelled results suggest that belimumab reduces the risk for organ damage (since AMS is a predictor of damage) for most organs, organ damage will occur later in belimumab patients. Organ damage is irreversible and lasts until death. The duration of the organ damage therefore depends on the remaining lifespan of the patient. The effect of belimumab on the duration of organ damage is thus a product of the decreased risk and extended onset of organ damage and the prolonged life of these patients. Although a decreased duration of damage is shown for organs on which belimumab has a big effect (cardiovascular, pulmonary and renal), the duration of the damage in other organ systems is increased due to the estimated prolonged life-expectancy (Table 6.29).

Table 6.29. Average duration (yrs) of organ damage for all patients – Pooled total population.

	SoC	Belimumab	Difference
Cardiovascular	5.14	4.86	-0.28
Diabetes	2.85	2.96	0.11
Gastrointestinal	4.83	5.36	0.53
Malignancy	4.28	4.63	0.35
Musculoskeletal	10.99	11.59	0.59
Neuropsychiatric	10.90	11.43	0.52
Ocular	7.85	8.20	0.35
Peripheral vascular	3.16	3.14	-0.02
Premature gonadal failure	1.81	1.87	0.06
Pulmonary	8.73	8.52	-0.20
Renal	3.61	3.16	-0.45
Skin	2.26	2.34	0.08

6.7.2 Please provide (if appropriate) the proportion of the cohort in the health state over time (Markov trace) for each state, supplying one for each comparator.

Not appropriate as a micro-simulation model was used in this decision problem.

6.7.3 Please provide details of how the model assumes QALYs accrued over time. For example, Markov traces can be used to demonstrate QALYs accrued in each health state over time.

For every yearly cycle a patient is alive, the utility is determined and added to their cumulative utility from the previous year. The yearly utility is calculated with a regression based on age, disease activity and baseline characteristics and multiplied with the lowest utility multiplier of applicable organ damage.

6.7.4 Please indicate the life years and QALYs accrued for each clinical outcome listed for each comparator. For outcomes that are a combination of other states, please present disaggregated results.

These data cannot be easily obtained from the model.

6.7.5 Please provide details of the disaggregated incremental QALYs and costs by health state, and of resource use predicted by the model by category of cost. Suggested formats are presented below.

As shown previously in Figure 6.20, belimumab patients have an increased life-expectancy. The belimumab-treated patients on average live 1.5 years longer, have a reduction in the average mean SLEDAI, and similar total organ damage score as measured by the SLICC compared to SoC patients (Table 6.30). Treatment with belimumab adds 0.43 (discounted) QALYs.

Table 6.30. Summary of effects – Pooled total population.

	SoC	Belimumab	Difference
Age at Death	68.4	69.9	1.5
SLICC at Death	4.0	4.0	0.0
AMS	4.8	4.33	-0.5
Average Monthly Steroid	214.2	203.2	-11.0
Life Years (undiscounted)	30.47	31.97	1.5
Life Years (discounted)	16.74	17.33	0.59
QALYs (undiscounted)	16.46	17.38	0.9
QALYs (discounted)	9.55	9.98	0.43

The total costs for patients consist of disease activity costs, belimumab costs, and costs incurred by organ damage. For both groups of patients, the organ damage costs are the highest expense (Table 6.31). These costs are influenced by the duration of the organ damage (Table 6.29), the onset of organ damage through the discount rate and the progression of costs over time. Since the first year costs are often higher than the costs for subsequent years (see Table 6.26), the occurrence is also a factor in the organ damage costs.

For the cardiovascular, pulmonary and renal organs, the costs are lower since the duration was shorter. Although the duration of peripheral vascular damage is slightly longer, the costs for this organ were slightly lower for belimumab treated patients. This is a result of the extended onset of the damage, which affects discounting. In total, the organ damage costs are lower for belimumab-treated patients due to the benefits on the pulmonary and renal systems. The costs related to disease activity are very slightly higher for belimumab treated patients. Although these patients have lower disease activity and therefore lower associated direct costs per year, the costs increase due to the increased life expectancy. The main difference in costs is caused by the belimumab treatment, comprising £37,638 (90.7%) of the total absolute cost difference of £41,492 (Table 6.31).

Table 6.31. Summary of (discounted) costs – Pooled total population

Discounted	SoC	Belimumab	Difference	Absolute difference	% absolute difference
Disease activity related costs	£27,004	£27,265	£262	£262	0.6%
Belimumab drug acquisition	£0	£31,687	£31,687	£31,687	76.4%
Belimumab administration	£0	£5,950	£5,950	£5,950	14.3%
Organ damage costs					
Cardiovascular	£1,760	£1,635	-£125	£125	0.3%
Diabetes	£2,948	£3,003	£55	£55	0.1%
Gastrointestinal	£413	£442	£29	£29	0.1%
Malignancy	£1,059	£1,092	£33	£33	0.1%
Musculoskeletal	£10,043	£10,371	£328	£328	0.8%
Neuropsychiatric	£6,635	£6,825	£190	£190	0.5%
Ocular	£431	£435	£3	£3	0.0%
Peripheral vascular	£1,225	£1,187	-£39	£39	0.1%
Premature gonadal failure	£0	£0	£0	£0	0.0%
Pulmonary	£38,796	£36,966	-£1,830	£1,830	4.4%
Renal	£7,268	£6,308	-£960	£960	2.3%
Skin	£0	£0	£0	£0	0.0%
Sum of organ damage costs	£70,579	£68,264	-£2,315		
Total direct costs	£97,583	£133,167	£35,584	£41,492	100%

Base-case analysis

6.7.6 Please present your results in the following table. List interventions and comparator(s) from least to most expensive and present ICERs in comparison with baseline (usually standard care) and then incremental analysis ranking technologies in terms of dominance and extended dominance.

Table 6.32. Base-case results – Pooled total population

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
SoC	£97,583	16.74	9.55	-	-	-	
Belimumab	£133,167	17.33	9.98	£35,584	0.59	0.43	£82,909
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years							

The results summarised in Table 6.32 show that belimumab-treated patients are estimated to live longer, however due to their increased life expectancy and acquisition costs for belimumab, overall costs are higher than for SLE patients treated with SoC. The incremental costs amount to £35,584, resulting in 0.59 added life years or 0.43 added QALYs (discounted). This results in an incremental cost effectiveness ratio (ICER) of £82,909 per QALY.

Sensitivity analyses

6.7.7 Please present results of deterministic sensitivity analysis. Consider the use of tornado diagrams.

Tornado diagrams for the ICER, incremental QALYs and incremental costs are presented in Figures 6.23, 6.24, 6.25 respectively.

Figure 6.23. Tornado diagram for univariate sensitivity analyses on the ICER – Pooled total population

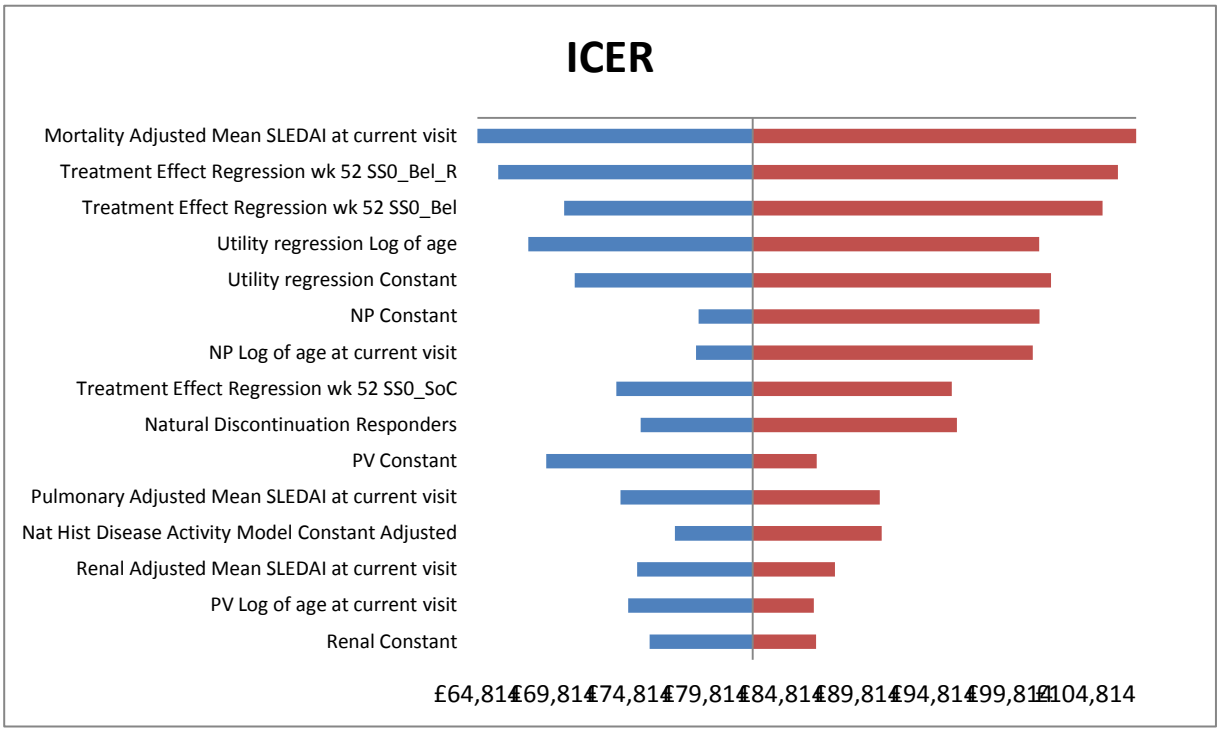


Figure 6.24. Tornado diagram for univariate sensitivity analyses on the incremental QALYs (delta E) – Pooled total population

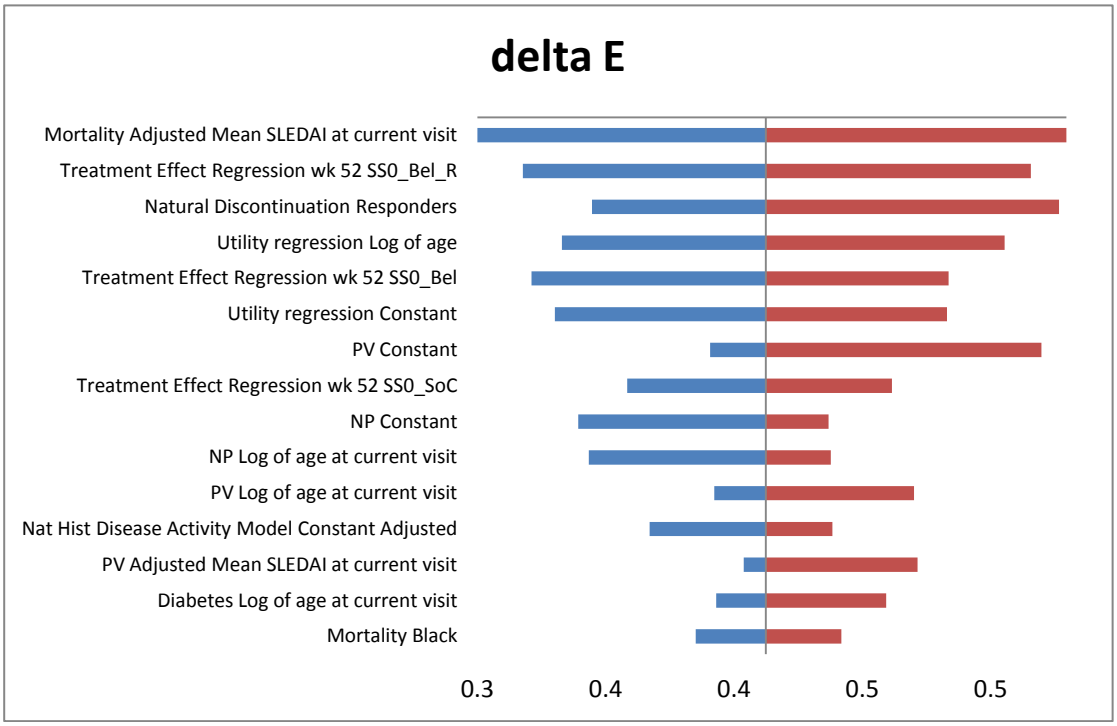
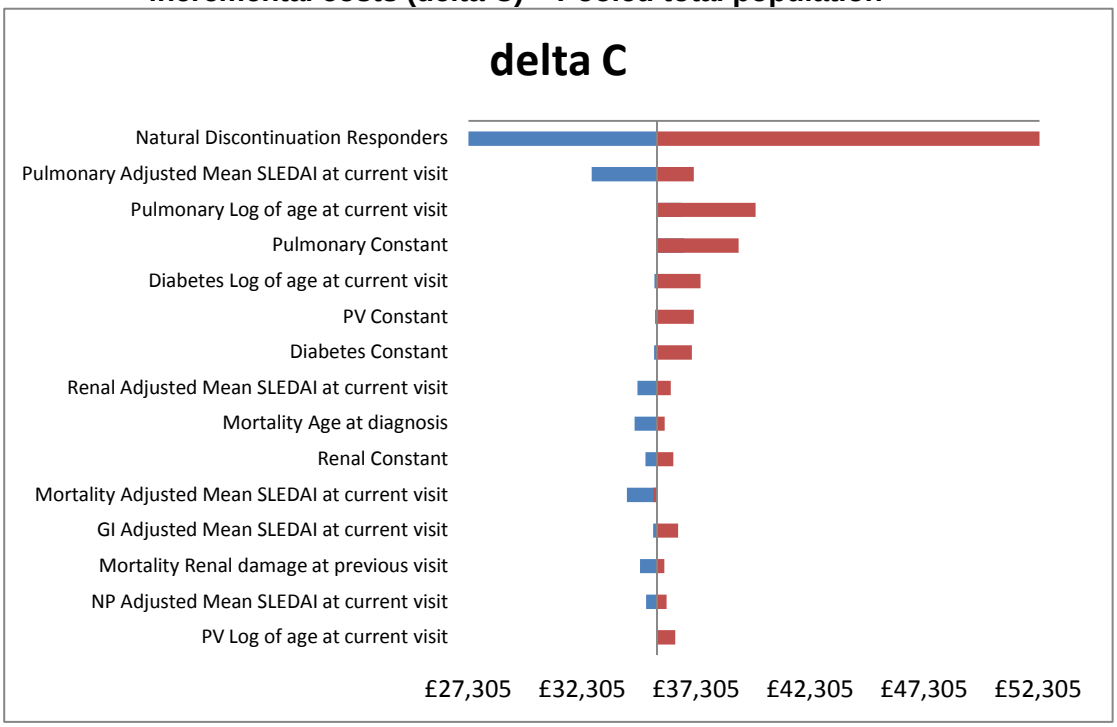


Figure 6.25. Tornado diagram for univariate sensitivity analyses on the incremental costs (delta C) – Pooled total population



6.7.8 Please present the results of a PSA, and include scatter plots and cost-effectiveness acceptability curves.

The scatter plot and acceptability curve based on the PSA are presented in Figures 6.26 and 6.27.

Figure 6.26: Scatter plot of the PSA – Pooled Population

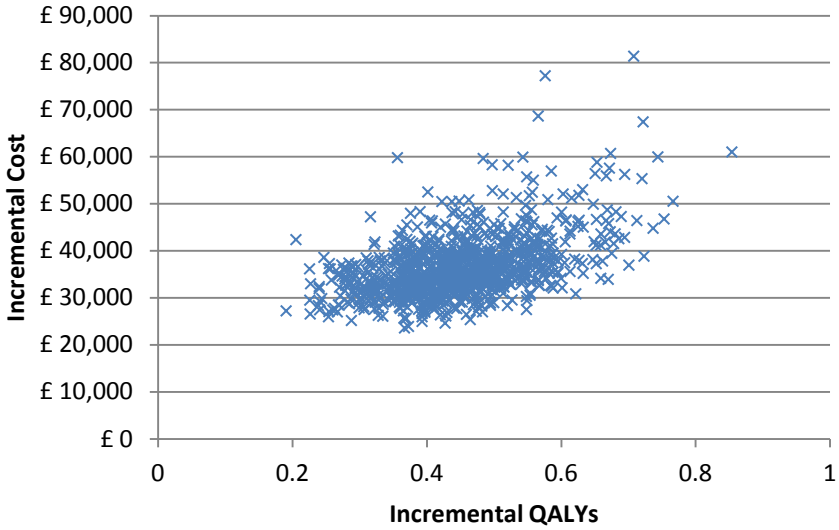
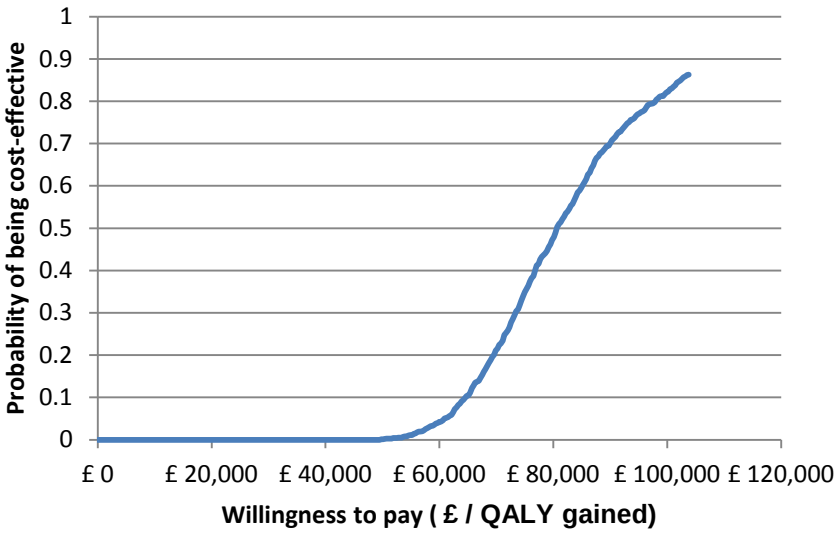


Figure 6.27: PSA Acceptability Curve –Pooled Population



6.7.9 Please present the results of scenario analysis. Include details of structural sensitivity analysis.

Table 6.33 below summarises the results of all the scenario analyses for the base case for the pooled total population**.**

Table 6.33. Summary of scenario analyses for the base case – Pooled total population

Description of Scenario	Scenario Details	Incremental Cost Belimumab	Incremental LYs Belimumab	Incremental QALYs Belimumab	Incremental Cost per QALY
Base Case for Pooled total population	Time horizon = lifetime, lifetime max effect of belimumab; responder rule of SS reduction ≥4 at week 24; adjusted natural history model and AMS included but item involvement removed; no vial sharing included (i.e. vial waste included).	£35,584	0.59	0.43	£82,909
Responder rule excluded	As base case but with responder rule at 24 weeks excluded	£50,496	0.44	0.33	£151,936
Alternative Responder rule	As base case but with responder rule of SS reduction of ≥6 at week 24;	£24,140	0.48	0.36	£68,074
Original natural history model	As base case but with original natural history model chosen	£35,253	0.48	0.38	£93,654
With vial sharing	As base case but with vial sharing (without waste option) chosen	£34,157	0.59	0.43	£79,582
Higher drug administration cost	As base case but with a drug administration cost of £159 as recommended by ERG as a sensitivity analysis for the tocilizumab appraisal for RA	£37,143	0.59	0.43	£86,540
Increased vial price	As base case but with vial price increased (120mg=£127.80 400mg=£426)	£39,327	0.59	0.43	£91,629
Inclusion of SoC costs	As base case but including an additional cost of SoC treatment to account for any extended life expectancy with belimumab	£35,815	0.59	0.43	£83,445

6.7.10 What were the main findings of each of the sensitivity analyses?

Univariate Sensitivity Analysis Results

The most important drivers of the model are the effect of the adjusted mean SLEDAI (AMS) on mortality, the treatment effect regression to estimate the effect of belimumab after 52 weeks and the constant value and log of age coefficient from the baseline utility regression. Higher AMS values will be linked to greater opportunity for the benefits of belimumab treatment to be seen in increasing life expectancy compared with SoC, as a consequence there will be more of an increase in QALYs leading to lower ICERs. As regards the treatment effect regression, the greater the benefit seen with belimumab compared with SoC, the higher the incremental QALYs and the lower the ICER.

The univariate analysis suggests that the greater the coefficient for $\log_e$ (age) and the smaller the constant value in the regression equation for estimating baseline utility, the higher the ICER will be. However for these particular parameters, a univariate analysis is conditional on keeping the other parameter values fixed, which in this case is not appropriate due to the dependence between both coefficients. There is substantial negative correlation between the constant and the effect of $\log_e$ (age) in the utility regression (see Section 9.22, Appendix 22). As such, changing one parameter to the upper limit implies that the other parameter would likely be lower thereby (partly) canceling each other out. This also applies to the effect of $\log_e$ age and the constant in the neuropsychiatric and pulmonary models. In conclusion, caution should be applied with interpreting the univariate results due to the correlation between several model parameters. The PSA however, acknowledges this correlation by drawing from multivariate normal distributions with covariance matrices.

Discontinuation probabilities for patients satisfying the six-month treatment continuation rule affect both incremental benefits and costs and thereby the ICER. For example, lower probabilities for natural discontinuation lead to higher incremental QALYs with belimumab compared with the base case value but significantly increased drug costs resulting in higher ICERs.

The effect of all varied model parameters can be found in (see Section 9.29, Appendix 29).

PSA results

The PSA results show that at a willingness to pay (WTP) of £30,000 per QALY gained, there is a 0% probability that belimumab is cost-effective compared to SoC. With a WTP of £60,000 per QALY gained, there is still only a 4% probability that belimumab is cost-effective compared to SoC. This increases to 50% at a WTP of £80,700 per QALY gained.

Scenario Analysis Results

Excluding the treatment continuation rule from the cost-effectiveness analysis had a major impact on the ICER, almost doubling the base case ICER of £82,909 to £151,936 per QALY. This is to be expected as continuing to include patients receiving belimumab in the model who incur considerable drug acquisition costs without also demonstrating important benefits in reducing disease activity, organ damage and mortality will lead to much higher ICERs. In contrast, using an alternative more stringent treatment continuation rule of a decrease in SS score ≥6 after 24 weeks treatment, rather than SS score ≥4 used in the base case, resulted in an ICER of £68,074 per QALY, which was nearly £15,000 per QALY lower than the base case ICER. This is to be expected as the sample of patients experiencing this more stringent criterion for response will be considerably reduced. While this smaller responder subgroup also shows fewer QALYs gained compared with the base case (as the total benefit seen is averaged across responders and non-responders) the reduction in the incremental drug costs with belimumab, compared with the base case, outweighs this reduced overall benefit, resulting in a more favourable ICER.

The scenario which included the option for vial sharing resulted in reducing the base case ICER by approximately £3,300 per QALY.

The scenario considering a higher administration cost of belimumab of £159 as suggested by the ERG reviewing the tocilizumab STA, compared with the value of £126 used in the base case, had the effect of increasing the ICER by £3631 per QALY. However, we believe that the value used in our base case of £126 is a fair assessment of administration costs as the infusion is only of one hour duration and is straightforward to administer.

The scenario examining the ICER using the original natural history model rather than the adjusted model, used in the base case, led to an increased ICER of £93,654 per QALY. This is due to less benefit being observed from reducing organ damage compared with the base case due to performing long-term modeling on patients with less disease activity i.e. less severe than the patients recruited into the BLISS studies.

Increasing the vial price to the expected maximum price limit, led to increasing the base case ICER by approximately £8,700 per QALY. However the vial prices used in the base case are our best estimate of what the final vial price will be.

Adding additional standard of care costs to account for any additional life expectancy had minimal effect on the ICER, increasing it by £536 per QALY.

6.7.11 What are the key drivers of the cost-effectiveness results?

The inclusion of a responder rule and how this rule is defined is a key driver of the ICER; a more stringent responder rule seems to indicate improved overall cost-effectiveness. The amount of benefit seen with belimumab in the BLISS studies in reducing disease activity (SELENA-SLEADAI score) after 52 weeks treatment and the benefit of increased life expectancy with belimumab are also important drivers, however in the cost-effectiveness analysis this latter benefit is weighed against the increased time spent with organ damage. In addition, the acquisition costs of the drug have a major influence on the ICER.

6.8 Validation

6.8.1 Please describe the methods used to validate and quality assure the model. Provide references to the results produced and cross-reference to evidence identified in the clinical, quality of life and resources sections.

Model convergence measures

For a micro-simulation it is important that the model results are stable. The results are conditional on the specific group of patients simulated and their history. In the average results of a group of patients there is always a certain degree of randomness involved due to sampling error. Sampling error can be reduced by simulating more patients.

Two approaches to test convergence were considered. The first approach was to construct a convergence graph outlining average model outcome versus an increasing number of patients simulated. With a relatively small number of patients simulated, the variability in average model outcomes is likely to follow a ‘spiky’ behaviour with high and low peaks. The more patients are simulated to obtain the average value of the outcome of interest, the more ‘stable’ the course of this curve will become. The curve should converge to the actual value of the outcome. Convergence graphs are drawn for the incremental QALYs and costs, see Figures 6.28 and 6.29 below for the pooled total population and Figures 6.30

and 6.31 for the high disease activity subgroup. The graphs show good convergence for incremental costs for the pooled total population and for both incremental costs and QALYs for the high disease activity subgroup. The convergence graph for the incremental QALYs for the pooled total population is a little less stable. All results are not affected by sampling errors.

Figure 6.28. Convergence graph for incremental QALYs – Pooled total population

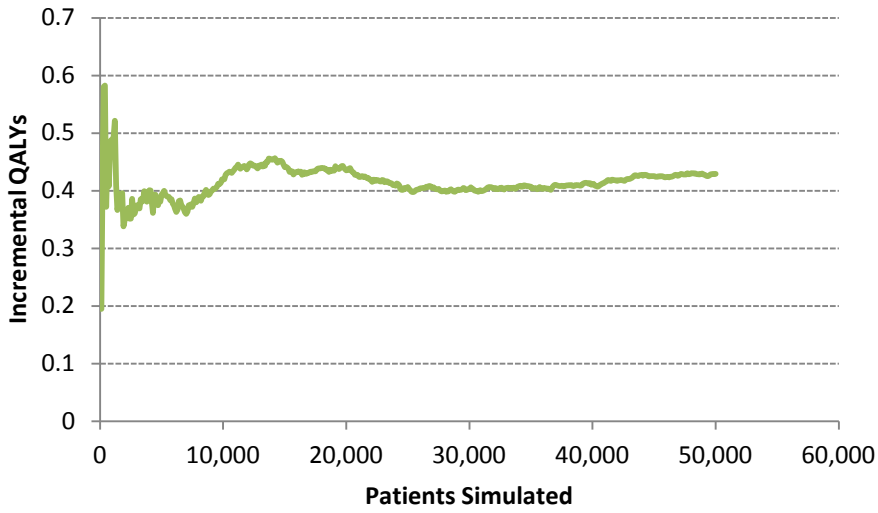


Figure 6.29. Convergence graph for incremental Costs – Pooled total population

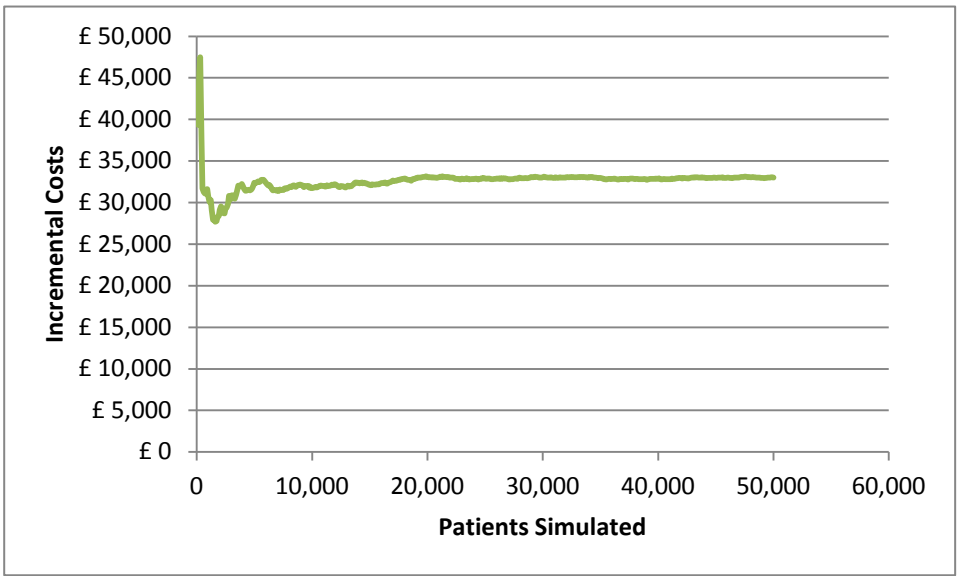


Figure 6.30. Convergence graph for incremental QALYs – High disease activity subgroup

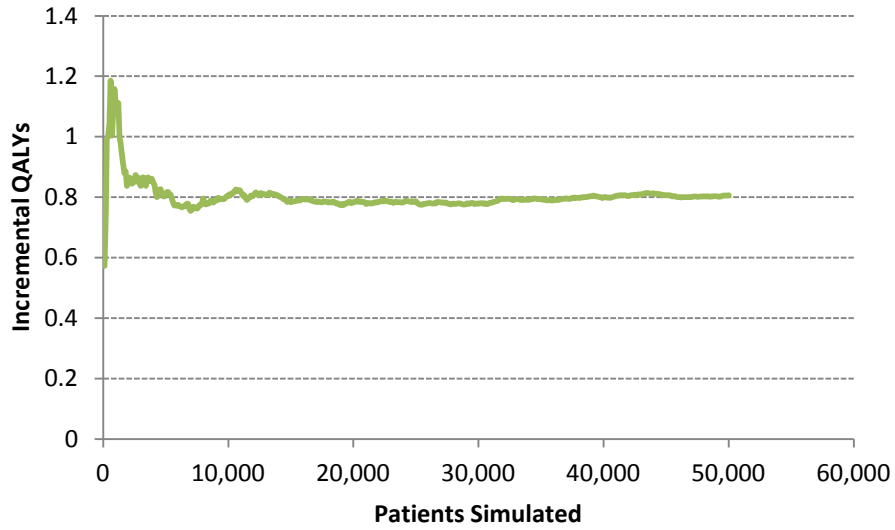
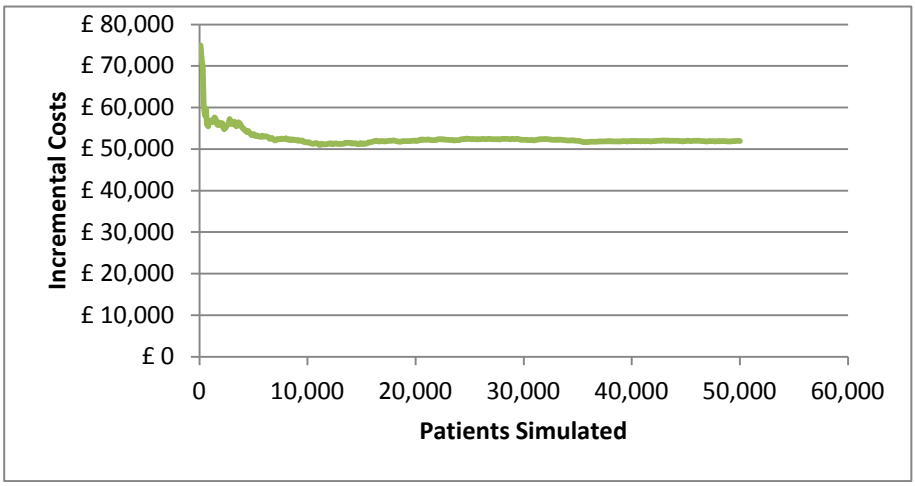


Figure 6.31. Convergence graph for incremental Costs – High disease activity subgroup



The second approach considers multiple runs of 50,000 patients with different random number seeds, thereby using a different sequence of random numbers per run. The results of each different run should lead to similar values (apart from small deviations due to sampling error). The results of this test demonstrated that the results from runs using 5 different random number seeds did lead to similar values being generated for incremental QALYs and costs (see Table 6.34 below).

Table 6.34. Summary of incremental benefit, costs and ICERs for different seed values – Pooled total population and high disease activity subgroup

Population	Seed	Δ QALYs	Δ Costs	ICER
Pooled total population	1	0.429	£35,584	£82,909
	2	0.446	£36,078	£80,918
	3	0.455	£36,291	£79,704
	4	0.436	£36,185	£83,003
	5	0.442	£36,171	£81,789
High disease activity subgroup	1	0.712	£45,944	£64,545
	2	0.732	£46,505	£63,563
	3	0.732	£46,408	£63,369
	4	0.721	£46,695	£64,795
	5	0.728	£46,622	£64,021

Internal and external validation of long-term outcome predictions

In addition, as regards the accuracy of predicting long-term outcomes and mortality the internal validation of the models was performed on the JH cohort. The natural history model for SLE was developed following an extensive analysis of the JH cohort and detailed examination of the relationships between various risk factors and organ damage and mortality. This comprised a backward stepwise approach to covariate selection and the use of goodness of fit statistics to select the function form of the parametric survival distribution. The process of model selection is described in full in the JH Data Analysis report (Section 9.21, Appendix 21). In order to validate the natural history model of SLE, a patient level simulation (10,000 simulations) was constructed to reproduce a SLE cohort of patients. In the first simulation a patient cohort with baseline characteristics and loss to follow-up similar to the JH lupus cohort were used. For an external validation of the predictive modeling, a second simulation was conducted using the Toronto dataset, with the aim of determining whether the models provided estimates that are representative of an SLE patient with the severity of disease of interest to this decision problem. As there wasn't any patient level data available for this cohort, the baseline data and loss to follow-up were taken from summary statistics from the Toronto Lupus Cohort.

Table 6.35 reports the incidence of events in the real cohorts and the simulation to illustrate the model's accuracy in predicting long-term outcomes.

Table 6.35. Number of events observed in simulation for each longitudinal SLE cohort

	Simulated Johns Hopkins (N=50000)	Johns Hopkins cohort (N=1282)	Simulated Toronto (N=50000)	Toronto cohort (N=967)
Mortality	5.4%	6.4%	7.8%	11.6%
Cardiovascular damage	8.5%	9.2%	12.4%	12.4%
Renal damage	3.3%	3.5%	8.2%	9.5%
Musculoskeletal damage	22.6%	22.4%	21.1%	26.8%
Neuropsychiatric damage	14.2%	14.4%	15.1%	16.1%
Pulmonary damage	10.6%	11.8%	9.1%	3.7%
Peripheral vascular damage	2.2%	3.4%	3.9%	5.7%
Gastrointestinal damage	6.7%	6.1%	8.2%	4.9%
Ocular damage	11.8%	12.6%	8.3%	21.5%
Skin damage	1.5%	2.3%	1.1%	9.7%

The simulation suggests that the SLE natural history model accurately predicts the outcomes of the JH cohort. This is not surprising given that the models are based on this data, but it does illustrate that combining the independent estimates of disease activity, steroids, and long-term events does not lead to biased estimates of the incidence of outcomes.

The simulation of the Toronto Lupus cohort highlights some organ systems where the JH data may have weak external validity in predicting long-term outcomes. Mortality, musculoskeletal damage, ocular damage, and skin damage are underestimated. The robustness of the Toronto mortality model was affected by a considerable amount of death events excluded from the analysis due to missing data on important covariates. Pulmonary and gastrointestinal damage are overestimated. However, the overall fit of the simulation is reasonable.

Unfortunately, access to the Toronto data has been restricted by the data custodians and patient level data was unavailable, which has limited the extent of the analysis of this cohort. It was not feasible to run an iterative analysis of the Toronto data to select covariates and parametric distributions. However, it was possible to replicate the JH time to event models using the Toronto data. For this

analysis the same covariate and functional forms were used to generate parametric survival analyses for the time to mortality and time to damage in each organ system. The intention of this validation exercise was to identify differences in the magnitude and statistical significance of covariates in the JH time to event models. It is particularly interesting to observe the differences in the coefficient for AMS in all models between the two cohorts as this parameter estimates the long-term benefits of treatment. However, using models on the Toronto database which were derived from the JH database has limitations as they may not be the most appropriately constructed models for that particular dataset, but it was not possible to check this.

Table 6.36 details the differences in the hazard ratios and time ratios estimated in the two SLE cohorts. The risk of mortality is increased by approximately 20% for a unit increase in AMS. The impact of AMS on renal damage and cardiovascular damage is slightly less in the Toronto cohort but statistically significant. The risk of neuropsychiatric damage for a unit change in AMS is statistically significant in the Toronto, but not in the Hopkins cohort.

Table 6.36. Comparison of the Hazard ratios and Time ratios for AMS in the two longitudinal SLE cohorts

Model	Johns Hopkins Cohort		Toronto Cohort	
	AMS coefficient	p-value	AMS coefficient	p-value
Mortality	1.238	0.000	1.203	0.000
Cardiovascular damage*	0.811	0.002	0.925	0.001
Renal damage	1.242	0.000	0.171	0.000
Musculoskeletal damage*	0.966	0.530	0.996	0.821
Neuropsychiatric damage	1.045	0.270	1.114	0.000
Pulmonary damage*	1.149	0.000	1.112	0.020
Peripheral Vascular damage	1.186	0.025	1.041	0.303
Gastrointestinal damage	0.941	0.391	0.934	0.284
Ocular damage*	0.956	0.423	0.975	0.224
Skin damage*	0.955	0.769	0.941	0.109
* Time ratios for loglogistic distributions				

Detailed analyses of covariates, and alternative parametric functional forms were not undertaken for the Toronto data therefore there is a risk of model misspecification in the Toronto analyses. The mortality model underestimates the

incidence of mortality and the cardiovascular damage model over-estimates the incidence of cardiovascular damage in the Toronto cohort. At this time the JH natural history model is considered the most robust method to predict long-term outcomes in SLE. However, the results of the Toronto simulation and Toronto statistical analyses are informative in identifying differences between the cohorts and suggesting alternative parameter estimates for sensitivity analyses.

In addition to the above validation, the model was checked for errors in formula and functionality both by a separate team employed by the supplier who constructed the model and also by Dr Liz Fenwick, Lecturer in Health Economics, Glasgow University. No important errors in the model formulae and functionality were identified.

Two independent academic health economists conducted reviews of the model suitability to address the decision problem and provided advice on how to improve the explanation of statistical methodology and assumptions used which have been incorporated.

6.9 Subgroup analysis

For many technologies, the capacity to benefit from treatment will differ for patients with differing characteristics. This should be explored as part of the reference-case analysis by providing separate estimates of clinical and cost effectiveness for each relevant subgroup of patients.

This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', section 5.10.

Types of subgroups that are not considered relevant are those based solely on the following factors.

- Individual utilities for health states and patient preference.
- Subgroups based solely on differential treatment costs for individuals according to their social characteristics.
- Subgroups specified in relation to the costs of providing treatment in different geographical locations within the UK (for example, when the costs of facilities available for providing the technology vary according to location).

6.9.1 Please specify whether analysis of subgroups was undertaken and how these subgroups were identified. Were they identified on the basis of an a priori expectation of differential clinical or cost effectiveness due to known, biologically plausible, mechanisms, social characteristics or other clearly justified factors? Cross-reference the response to section 5.3.7.

The subgroup considered here is defined as SLE patients with positive anti-dsDNA, low complement and a baseline SS score greater or equal to 10. It comprises a total of 574 (34%) patients out of the total 1684 patients randomised into the two BLISS studies. For the assessment of efficacy and cost-effectiveness however only the 10 mg/kg dose of belimumab has been considered to be consistent with the proposed licensed dose. This provides a total subgroup size of 396 patients. This subgroup will be referred to in the remainder of this section as the "high disease activity" subgroup. Although this subgroup was not defined *a priori* it is felt to be a very relevant subgroup for this decision problem as it attempts to identify the SLE patients who, in addition to

experiencing a high degree of disease activity, are also most likely to progress to experience long-term organ damage and the greatest morbidity (see Section 5.3.7). There still exists a significant amount of unmet need in these patients, who experience a high amount of disease activity despite being managed on high dose corticosteroids or immunosuppresants as part of current standard of care. It is in these patients that we believe belimumab has an important role in helping to manage both the short-term impact on patient quality of life by reducing disease activity, with the potential to rationalise their current standard of care, while also contributing in the longer term to the prevention of organ damage and consequently the impact on both morbidity and mortality.

6.9.2 Please clearly define the characteristics of patients in the subgroup.

The baseline characteristics for the "high disease activity" subgroup used in the model is summarised below in Tables 6.37, 6.38, and 6.39. The proportion of females, the disease duration and the SLICC Damage Index score are very similar between this subgroup and the pooled total population. However, this subgroup is on average 3.6 years younger than the pooled total population, and there is a slightly lower proportion of SLE patients of black African or black Caribbean ethnicity (6.8%) compared with the pooled total population (8.7%) (Table 6.2).

Table 6.37. Baseline patient demographics – High disease activity subgroup

Patient demographics	Mean or %	Distribution	Parameter
Age (years)	34.3	Multinomial	Probability for each age
Gender (% females)	94.2%	Bernoulli	0.9419
Black Ethnicity (%)	6.8%	Bernoulli	0.0682
SLE Disease duration (yrs)	6.7	Geometric	0.1488
SLICC Damage Index score*	0.64	NA	NA

*Note that Instead of simulating a patient's total SDI score, the scores simulated for each individual item presented in Table 4 are summed to determine the total SDI.

Table 6.38 presents baseline disease activity parameters and steroid use. The mean SELENA-SLEDAI (SS) score at baseline for this subgroup was 12.7 compared with the pooled total population mean SS score of 9.7. The majority of patients in the subgroup had skin involvement (92%). The other main types of

disease activity observed in this subgroup were related to renal involvement (28% patients) and vasculitis (12.4% patients). These percentages were lower in the pooled total population with 82%, 16% and 7% of patients showing skin involvement, renal involvement and vasculitis respectively.

Table 6.38. Baseline disease activity parameters and steroid use – High disease activity subgroup

SLE disease activity parameters	Mean (sd)	Distribution	Parameter
SELENA-SLEDAI score	12.72	Multinomial	Probability for each score
Increased DNA binding	100.0%	Bernoulli	1.000
Low Complement	99.7%	Bernoulli	0.997
Vasculitis	12.4%	Bernoulli	0.124
Neuropsychiatric involvement	2.5%	Bernoulli	0.025
Renal involvement	28.0%	Bernoulli	0.280
Serositis involvement	6.8%	Bernoulli	0.068
Haematological Involvement	8.6%	Bernoulli	0.086
Skin Involvement	91.7%	Bernoulli	0.917
Daily steroid dose (mg/day) – mean (SD)	11.9 (9.12)	Gamma	1.72; 6.96[#]

values for shape and scale for the Gamma distribution respectively

Table 6.39 presents the baseline SLICC Damage Index (SDI) item occurrences observed in the high disease activity subgroup. These scores are similar to those seen for the pooled total population.

Table 6.39. Baseline individual SDI item scores – High disease activity subgroup

SLICC damage item	Score 0	Score 1	Score 2	Score 3	Score 4	Distribution
Cardiovascular	95.5%	4.0%	0.5%	0.0%	0.0%	Multinomial
Diabetes	98.2%	1.8%	0.0%	0.0%	0.0%	Multinomial
Gastrointestinal	97.5%	2.3%	0.3%	0.0%	0.0%	Multinomial
Malignancy	99.7%	0.3%	0.0%	0.0%	0.0%	Multinomial
Musculoskeletal	88.1%	8.8%	2.5%	0.3%	0.3%	Multinomial
Neuropsychiatric	90.4%	7.3%	1.8%	0.5%	0.0%	Multinomial
Ocular	94.7%	5.1%	0.3%	0.0%	0.0%	Multinomial
Peripheral vascular	95.2%	4.3%	0.5%	0.0%	0.0%	Multinomial
Premature gonadal failure	99.0%	1.0%	0.0%	0.0%	0.0%	Multinomial
Pulmonary	97.5%	2.3%	0.3%	0.0%	0.0%	Multinomial
Renal	97.2%	2.8%	0.0%	0.0%	0.0%	Multinomial
Skin	92.4%	6.8%	0.5%	0.3%	0.0%	Multinomial

6.9.3 Please describe how the statistical analysis was undertaken.

The same methodology was used to analyse the subgroup population as described in Section 6.2.2 for the pooled total population. The results from the various regression analyses to derive the inputs to the model are presented below for the subgroup population.

Treatment continuation probabilities with belimumab and natural discontinuation probabilities

As described for the pooled total population, patients who did not satisfy the treatment continuation rule at Week 24 (as measured by a minimal decrease of 4 points on the SS) were switched to the SoC arm in the model. Table 6.40 presents the percentage of patients continuing treatment with belimumab and the discontinuation rates observed in the clinical trials and used in the model, separately for those that satisfied or who did not satisfy the treatment continuation rule on belimumab. In this subgroup there were 67% of patients who satisfied the treatment continuation rule compared with 51% in the pooled total population. The withdrawal rate for the patients who did not continue belimumab after 24 weeks and thus switched to SoC was much higher over the model horizon for the subgroup (37.4%) compared with the pooled total population (21.4%).

Table 6.40. Summary of natural discontinuation and probability of treatment continuation after 24 weeks for belimumab patients – High disease activity subgroup

Patients satisfying the treatment continuation rule	66.8%	
Natural discontinuation (Withdrawal)	Patients who continue belimumab after 24 weeks	Patients who discontinue belimumab after 24 weeks
Year 1	4.4%	37.4%
Subsequent years	8.0%	37.4%

After simulating a patient's baseline characteristics they enter the model in which their remaining lifetime SLE history is simulated.

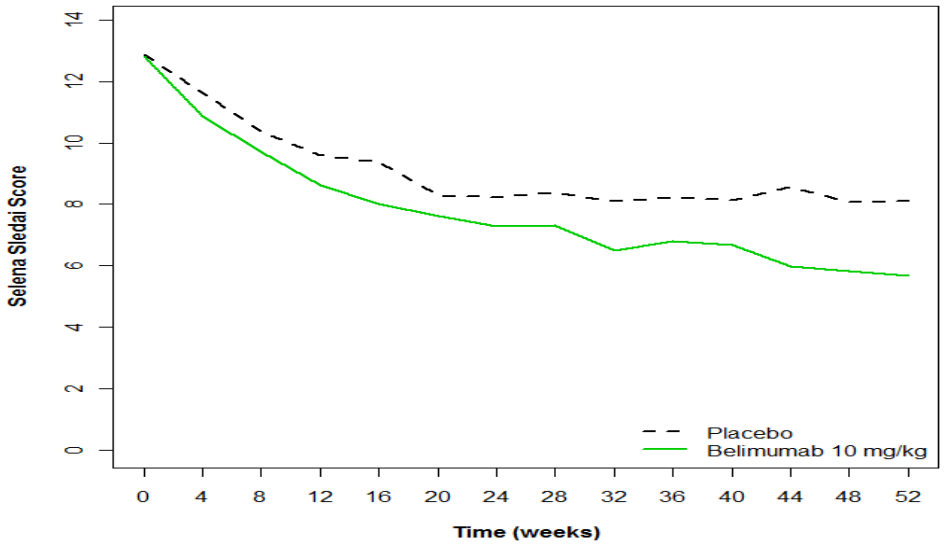
Year one treatment effects

In the first year of the simulation, the effects on disease activity as observed in the pooled BLISS trials for the subgroup are applied.

Effect on SELENA-SLEDAI (SS) score

The pooled average SS score from baseline to week 52 for SoC and belimumab 10 mg/kg is shown in Figure 6.32 for the subgroup. It demonstrates a very similar pattern of decline in SS score over time to the pooled total population.

Figure 6.32. Pooled average SELENA-SLEDAI score from baseline to week 52 for SoC (placebo) and belimumab 10 mg/kg – High disease activity subgroup



The results of the linear regression analysis of change in SS score at Week 52 based on baseline score (SS_0) for SoC, SS score for belimumab treated patients and an additional effect of belimumab "responders" for the subgroup are presented in Table 6.41.

Table 6.41. Linear regression explaining change in SELENA-SLEDAI (SS) score at week 52 – High disease activity subgroup

Parameter	Estimate	SE	p-value
SS_0 SoC	-0.349	0.022	<0.001
SS_0 all belimumab	-0.343	0.046	<0.001
SS_0 belimumab responders	-0.280	0.052	<0.001

Treatment continuation rule

Table 6.42 presents the probability of treatment continuation based on baseline SS score. These results are similar to those observed for the pooled total population and show that probability of treatment continuation is very dependent on baseline SS score.

Table 6.42. Probabilities of treatment continuation at 24 weeks for different baseline SS – High disease activity subgroup

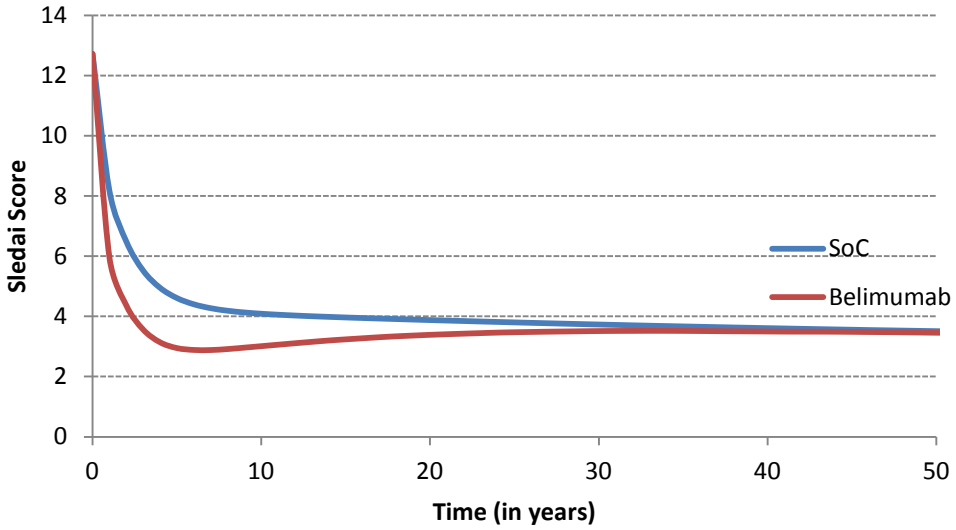
Baseline SELENA-SLEDAI	Probability of treatment continuation
0	0%
1	0%
2	0%
3	0%
4	0%
5	0%
6	0%
7	0%
8	0%
9	0%
10	60%
11	33%
12	73%
13	86%
14	84%
15	100%
16	63%
17	0%
18	50%
19	100%
20	80%
21	0%
22	75%
23	100%
24	0%
25	0%
26	0%
27	0%
28	0%
29	0%
30	0%

6.9.4 What were the results of the subgroup analysis/analyses, if conducted? Please present results in a similar table as in section 6.7.6 (Base-case analysis).

The average SLEDAI score (AMS) for 50,000 simulated patients is shown in Figure 6.33 for those patients who are still alive. It is clear from the graph that patients who are treated with belimumab (in addition to SoC) have a larger reduction in SS than patients who are treated with SoC alone over the first 15 years. Over time, the difference between the SS of belimumab and SoC patients declines because patients discontinue belimumab and subsequently lose its

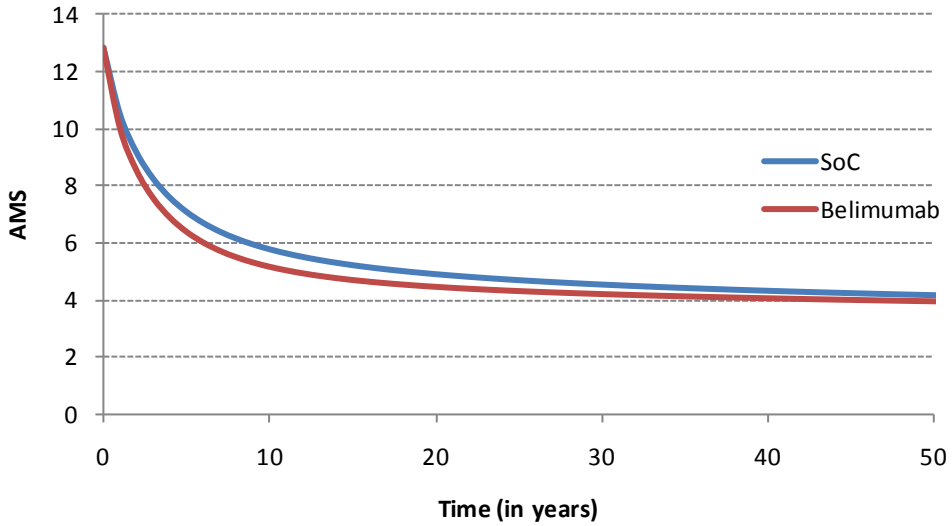
beneficial effect on disease activity.

Figure 6.33. SLEDAI Score over time (AMS) for 50,000 patients simulated - High disease activity subgroup.



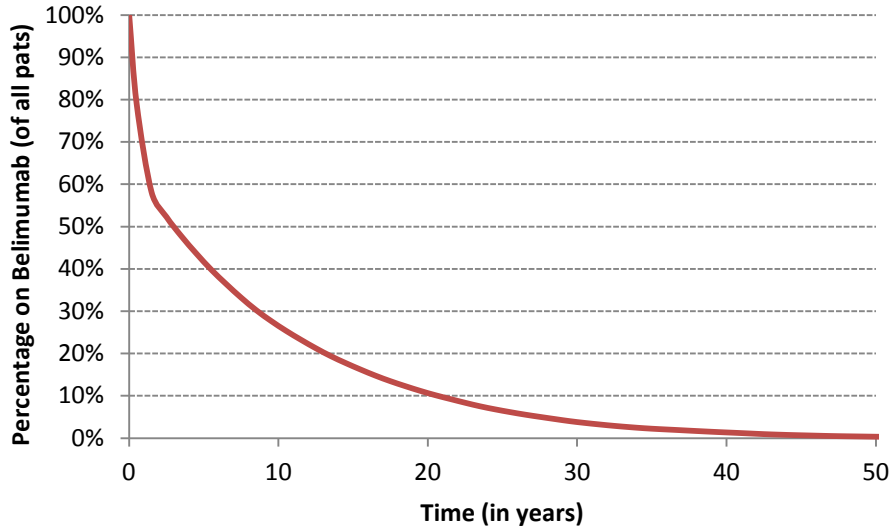
Although the disease activity of these patients returns to SoC levels in the long-term, a beneficial effect is kept through a decreased average mean SLEDAI (AMS) score over time (Figure 6.34).

Figure 6.34. Adjusted Mean SLEDAI over time censored for death - High disease activity subgroup.



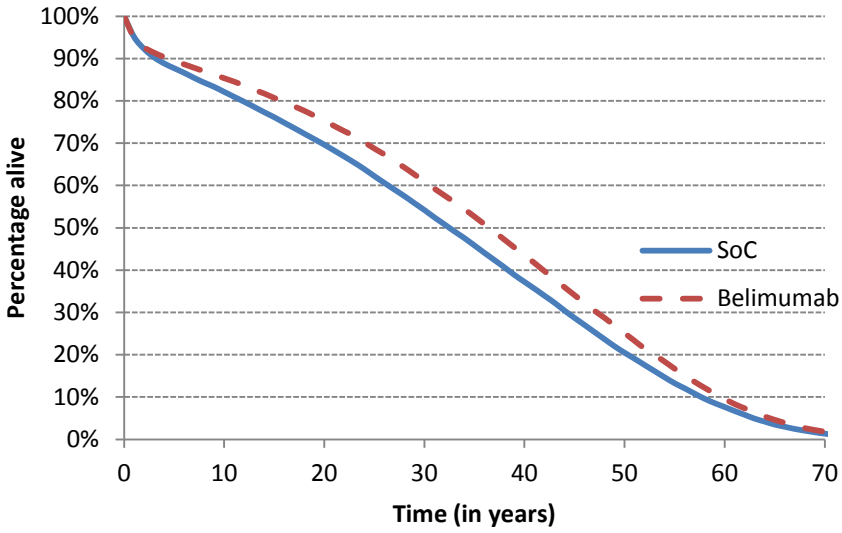
The discontinuation of patients on belimumab is shown in Figure 6.35. The large fall in patients continuing with belimumab in the first year is caused by patients not satisfying the criterion for treatment continuation at 24 weeks moving to SoC in the model**.** It can be seen that after 10 years only approximately 25% of patients are estimated to still be receiving belimumab. As stated for the pooled total population, assuming a constant rate of discontinuation is considered a conservative approach as it is likely that certain responders will remain responders over the longer term and continue to receive the benefit of being on belimumab.

Figure 6.35. Discontinuation from belimumab (includes death) – High disease activity subgroup.



The lower disease activity for belimumab patients will lead to a decreased steroid dose and a decreased risk for organ damage. The average disease activity over time, cumulative average prednisone dose and organ damage, contribute to the mortality risk (Table 6.12). The survival over time is therefore higher for belimumab patients than for patients on SoC (Figure 6.36). The relatively steep decline in survival in the first year for both arms is caused by the relatively high standardised mortality ratio for patients younger than 24 years (see Table 6.13). The survival curve for belimumab shows a less steep decline than for SoC, which is caused by the lower average disease activity scores.

Figure 6.36. Survival of patients over time – High disease activity subgroup



As discussed for the pooled total population, because belimumab patients have an estimated longer life expectancy, the exposure to the risk of organ damage is increased for belimumab patients, hence why for six of the organs (diabetes, gastrointestinal, malignancy, musculoskeletal, neuropsychiatric and ocular) the percentage of occurrence is similar or higher than for SoC (see Table 6.43). However, for cardiovascular, peripheral vascular, premature gonadal failure, pulmonary and renal systems, fewer patients on belimumab develop damage compared to SoC. This is due to the dependence of damage risk on disease activity which is lower for patients receiving belimumab.

Table 6.43. Organ damage occurrence for SLE patients until death – High disease activity subgroup

	SoC	Belimumab	Difference
Cardiovascular	23.9%	21.3%	-2.6%
Diabetes	17.9%	19.2%	1.3%
Gastrointestinal	22.1%	25.0%	3.0%
Malignancy	32.0%	34.1%	2.2%
Musculoskeletal	48.5%	48.9%	0.4%
Neuropsychiatric	44.7%	45.8%	1.1%
Ocular	35.1%	36.0%	0.8%
Peripheral vascular	21.5%	20.8%	-0.7%
Premature gonadal failure	7.2%	7.2%	0.0%
Pulmonary	39.9%	36.8%	-3.1%
Renal	24.3%	19.2%	-5.1%
Skin	7.9%	7.9%	0.0%

Although damage from some organ systems depend on AMS (see Table 6.14), a higher total percentage of belimumab patients develop damage in these systems. However, as discussed for the pooled total population, for the patients still alive, the risk of developing organ damage is slightly lower for cardiovascular, pulmonary and renal systems.

Since belimumab reduces the risk for organ damage for three of the organs, this damage will occur later in belimumab patients. Organ damage is irreversible and lasts until death. The duration of the organ damage therefore depends on the remaining lifespan of the patient. The effect of belimumab on the duration of organ damage is thus a product of the decreased risk, delayed onset of organ damage and the prolonged life of these patients. Although a decreased duration of damage is shown for organs on which belimumab has a large effect (cardiovascular, pulmonary and renal), the duration of damage for the other organ systems is increased due to the prolonged life-expectancy (Table T44).

Table 6.44. Average duration (yrs) of organ damage – High disease activity subgroup

	SoC	Belimumab	Difference
Cardiovascular	5.60	5.22	-0.38
Diabetes	2.64	3.01	0.37
Gastrointestinal	4.62	5.65	1.03
Malignancy	4.39	5.07	0.69
Musculoskeletal	11.24	12.14	0.90
Neuropsychiatric	11.17	12.08	0.92
Ocular	7.88	8.48	0.60
Peripheral vascular	3.66	3.75	0.08
Premature gonadal failure	1.77	1.86	0.09
Pulmonary	9.87	9.50	-0.37
Renal	5.38	4.46	-0.92
Skin	2.47	2.68	0.21

As shown previously in Figure 6.36, belimumab patients have an increased life-expectancy. The model predicts that belimumab-treated patients, in the subgroup with high disease activity, live on average 2.9 years longer (compared with 1.5 years for the pooled total population), have a reduction in average mean SLEDAI score, and similar total SLICC organ damage score compared with SoC patients (Table 6.45). Treatment with belimumab in this high disease activity subgroup provides an estimated additional 1.1 life years and 0.8 QALYs (discounted).

Table 6.45. Summary of health economic outcomes – High disease activity subgroup

	SoC	Belimumab	Difference
Age at Death	66.2	69.1	2.9
SLICC at Death	4.1	4.0	-0.1
AMS	5.5	4.55	-0.9
Average monthly steroid cumulative dose	228.1	207.9	-20.2
Life Years (undiscounted)	31.93	34.87	2.9
Life Years (discounted)	17.05	18.11	1.1
QALYs (undiscounted)	17.31	19.17	1.9
QALYs (discounted)	9.81	10.61	0.8

The total costs for patients consist of resource costs related to disease activity, belimumab acquisition and administration costs, and longer-term costs incurred by organ damage. For both treatment groups, the organ damage costs are the highest expense (Table 6.46). These costs are influenced by the duration of the organ damage shown in Table 6.44, the onset of organ damage through the discount rate, and the progression of costs over time.

For the cardiovascular, pulmonary and renal organs, the costs are lower since the estimated duration was shorter. Although the duration of peripheral vascular damage is slightly higher, the costs for this organ were slightly lower for belimumab treated patients. This is a result of the extended onset of the damage, which affects discounting. In total, the organ damage costs are lower for belimumab-treated patients due to the benefits on the pulmonary and renal systems. The costs related to disease activity are similar in the two treatment arms.

Although belimumab patients have less disease activity and consequently lower direct costs per year, the costs increase due to the estimated increased life expectancy. Overall, the main difference in costs is caused by belimumab acquisition and administration, amounting to £56,067 (89.6%) of the total absolute cost difference of £62,610 (Table 6.46).

.

Table 6.46. Summary of (discounted) costs over life time - High disease activity subgroup

Discounted	SoC	Belimumab	Difference	Absolute difference	% absolute difference
Disease activity related costs	£27,882	£28,130	£248	£248	0.4%
Belimumab drug acquisition	£0	£47,008	£47,008	£47,008	75.1%
Belimumab administration	£0	£9,059	£9,059	£9,059	14.5%
Organ damage costs					
Cardiovascular	£1,838	£1,633	-£205	£205	0.3%
Diabetes	£2,493	£2,731	£238	£238	0.4%
Gastrointestinal	£359	£399	£40	£40	0.1%
Malignancy	£998	£1,031	£33	£33	0.1%
Musculoskeletal	£9,758	£10,114	£356	£356	0.6%
Neuropsychiatric	£6,434	£6,719	£286	£286	0.5%
Ocular	£392	£391	-£1	£1	0.0%
Peripheral vascular	£1,380	£1,339	-£41	£41	0.1%
Premature gonadal failure	£0	£0	£0	£0	0.0%
Pulmonary	£42,692	£39,652	-£3,040	£3,040	4.9%
Renal	£11,139	£9,083	-£2,056	£2,056	3.3%
Skin	£0	£0	£0	£0	0.0%
Sum of organ damage costs	£77,483	£73,093	-£4,390	-	
Total direct costs	£105,366	£157,291	£51,925	£62,610	100.0%

Belimumab-treated patients are estimated to live longer, however, due to their increased life expectancy and due to belimumab treatment, costs are higher than for SoC patients. The incremental costs are £51,925, resulting in 1.05 added life years (discounted) or 0.806 added QALYs. This results in an incremental cost effectiveness ratio (ICER) of £64,410 per life year gained (Table 6.47).

Table 6.47. Base-case results – High disease activity subgroup

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus baseline (QALYs)	ICER (£) incremental (QALYs)
SoC	£105,366	17.05	9.81	-				
Belimumab	£157,291	18.11	10.61	£51,925	1.05	0.806	£64,410	£64,410
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years								

Sensitivity Analysis

Univariate Sensitivity Analysis

Tornado diagrams for the ICER, incremental QALYs and incremental costs are presented in Figures 6.37, 6.38, 6.39 respectively.

Similar drivers of the ICER results were observed in the univariate sensitivity analyses as seen for the pooled total population. The most important model drivers are the treatment effect regression to estimate the effect of belimumab after 52 weeks and the natural discontinuation probability. Clearly, the smaller the benefit seen with belimumab compared to SoC, the lower the incremental QALYs and the higher the ICER. Discontinuation probability affects both incremental QALYs and costs and thereby the ICER. For example, lower probabilities for natural discontinuation lead to higher incremental QALYs with belimumab compared with the base case value but significantly increased drug costs resulting in higher ICERs.

In addition, the effect of the AMS on mortality is an important driver of the model results. The greater the reduction in AMS with belimumab, the greater the increase in life expectancy with belimumab compared with SoC and consequently the higher the QALY gain leading to more favourable ICERs.

The constant and effect of log age in the utility regression also have an important effect on the incremental effects and the ICER. However for these particular parameters, a univariate analysis is conditional on keeping the other parameters fixed, which in this case is not very likely due to the dependence between both coefficients. There is substantial negative correlation between the constant and the effect of log age in the utility regression (see Section 9.22 Appendix 22). As such, changing one parameter to the upper limit implies that the other parameter would likely be lower. As such they will (partly) cancel each other out. This also applies to the effect of log age and the constant in the neuropsychiatric and pulmonary models. This is probably why the lower values for the latter analyses

are above the base case value. In summary, caution should be used when interpreting the univariate results due to the correlation between several model parameters. The PSA however, acknowledges this correlation by drawing from multivariate normal distributions with covariance matrices. The effect of all varied model parameters can be found in (see Section 9.29, Appendix 29).

Figure 6.37. Tornado diagram for univariate sensitivity analyses on the ICER – High disease activity subgroup

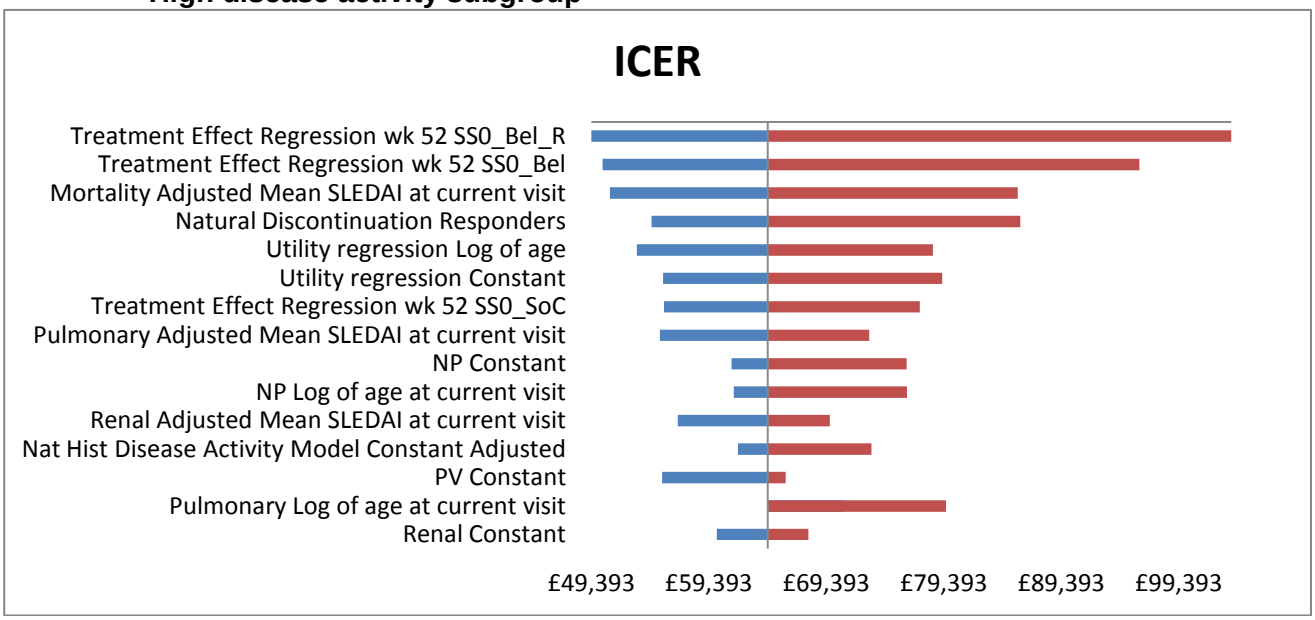


Figure 6.38. Tornado diagram for univariate sensitivity analyses on the incremental QALYs (delta E) – High disease activity subgroup

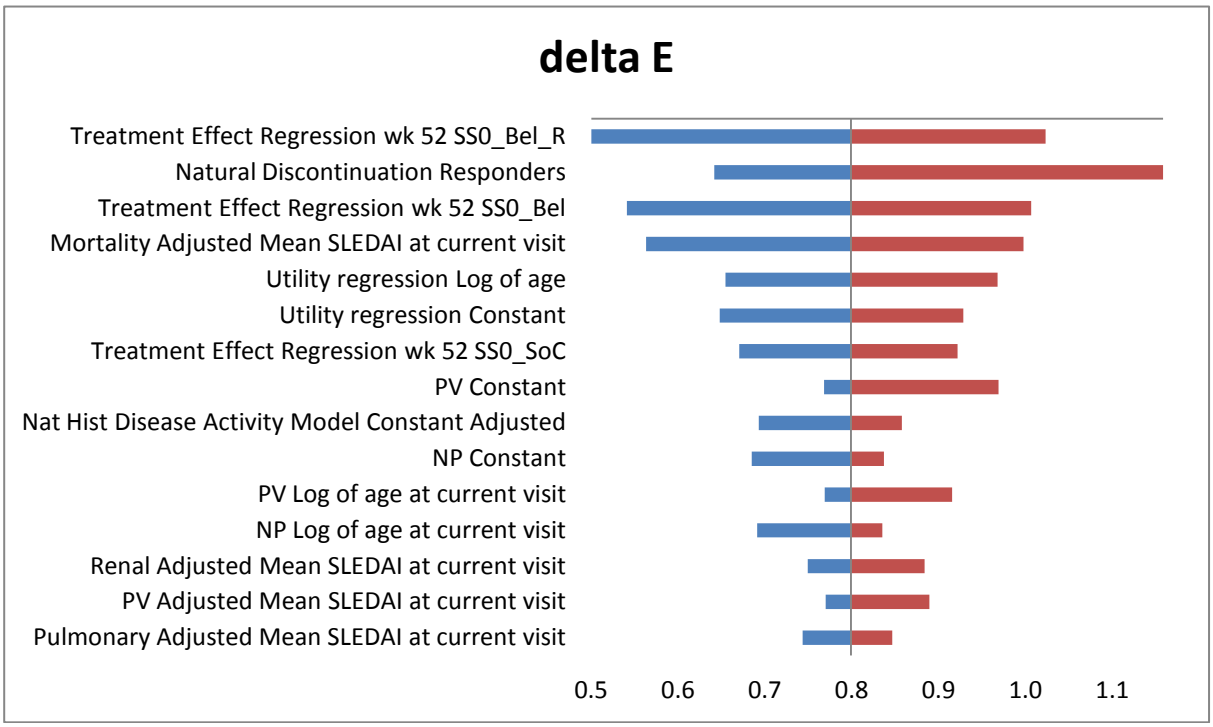
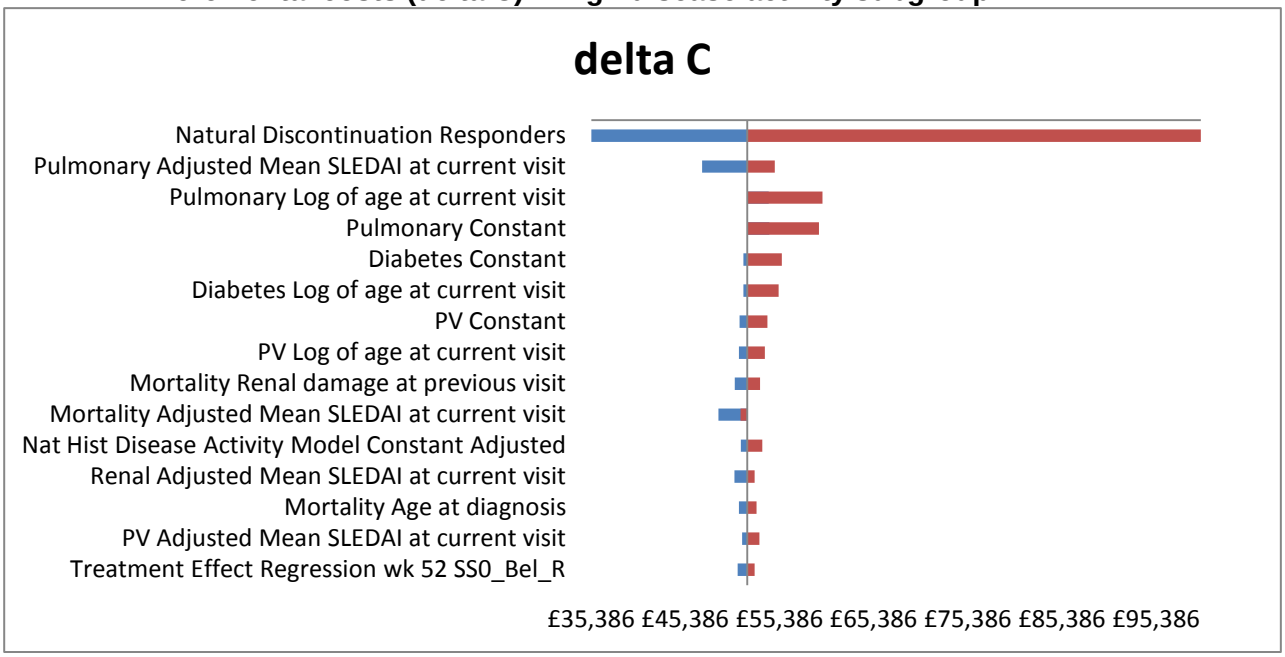


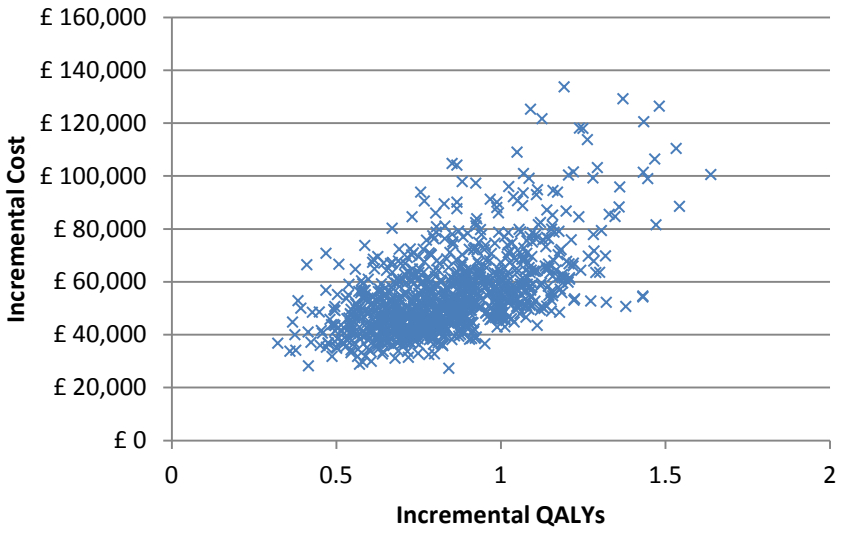
Figure 6.39. Tornado diagram for univariate sensitivity analyses on the incremental costs (delta C) – High disease activity subgroup



Probabilistic Sensitivity Analysis

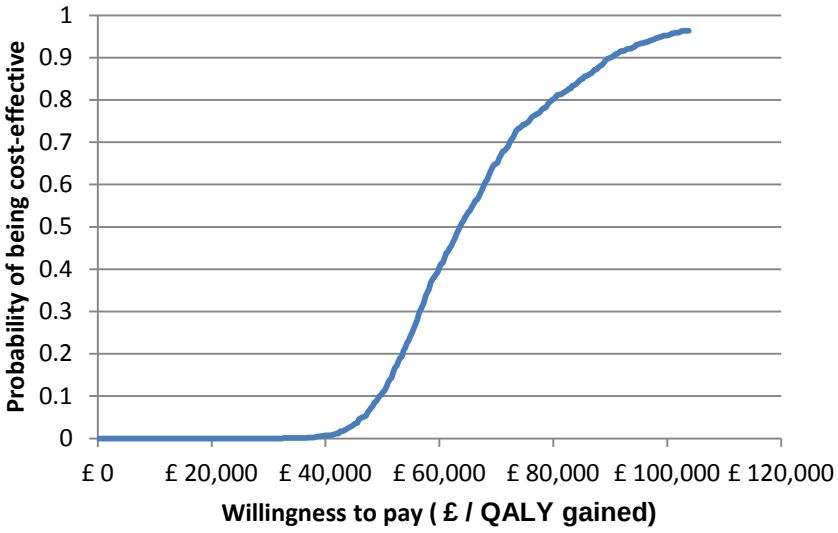
The scatter plot and acceptability curve based on the PSA are presented in Figure 6.40 and Figure 6.41 respectively.

Figure 6.40. Scatter plot of the PSA - High disease activity subgroup



The PSA results show that at a willingness to pay (WTP) of £30,000 per QALY gained, there is a 0% probability that belimumab is cost-effective compared to SoC. With a willingness to pay of £60,000 per QALY gained, there is a 35% probability that belimumab is cost-effective compared to SoC.

Figure 6.41. Acceptability curve of PSA - High disease activity subgroup



Scenario Analysis Results

The results of the scenario analyses are presented in Table 6.48 and present a very similar picture to those described for the pooled total population.

Table 6.48. Summary of Scenario Results - High disease activity subgroup

Description of Scenario	Scenario Details	Incremental Cost Belimumab	Incremental LYs Belimumab	Incremental QALYs Belimumab	Incremental Cost per QALY
Base Case for High Disease Activity Subgroup	Time horizon = lifetime, lifetime max effect of belimumab; treatment continuation rule of SS reduction ≥4 at week 24; adjusted natural history model; no vial sharing	£51,925	1.05	0.806	£64,410
Treatment continuation rule excluded	As base case but with Treatment continuation rule at 24 weeks excluded	£56,631	1.01	0.784	£72,207
Alternative treatment continuation rule	As base case but with treatment continuation rule of SS reduction of ≥6 at week 24;	£30,760	0.81	0.614	£50,114
Increased vial price	As base case but with vial price increased (120mg=£127.80; 400mg=£426)	£57,478	1.05	0.806	£71,297
Original natural history model	As base case but with original natural history model chosen	£51,227	0.82	0.659	£77,707
With vial sharing	As base case	£49,717	1.05	0.806	£61,671
Higher drug administration cost	As base case but with a drug administration cost of £159 as recommended by ERG as a sensitivity analysis for the tocilizumab appraisal for rheumatoid arthritis	£54,298	1.05	0.806	£67,353

Consistent with that seen for the total population, excluding a responder rule from the cost-effectiveness analysis again had a major impact of the ICER, increasing the base case ICER of £64,410 to £72,207 per QALY, approximately £7,797 per QALY higher. As discussed in the pooled total population results, this is to be

expected due to continuing to include drug acquisition costs for patients on belimumab who do not show benefits in significantly reducing disease activity and organ damage thus not providing benefits in terms of additional life years gained.

In contrast however, using an alternative more stringent responder rule comprising a decrease in SS score of ≥6 resulted in a significantly reduced ICER of £50,114 per QALY, just over £14,000 per QALY less than the subgroup base case ICER. Again, as explained for the total BLISS population, this is an expected finding due to a smaller sample of patients who would satisfy the more stringent response criterion continuing on belimumab, resulting in much lower mean total drug costs. While this smaller responder subgroup also shows fewer QALYs gained compared with the base case (as the total benefit seen is averaged across responders and non-responders) this reduced overall benefit does not outweigh the reduced incremental drug costs calculated for the belimumab arm, resulting in a more favourable ICER.

The scenario examining the impact of using the original natural history model, rather than the adjusted model, led to an ICER increase of £13,297 per QALY compared with the subgroup base case. This is due to less benefit being observed from reducing organ damage compared with the base case due to modeling on patients will less disease activity than those recruited into the BLISS studies.

The scenario which included the option for vial sharing, decreased the base case ICER by £2,739 to £61,671 per QALY.

The scenario considering a higher administration cost of belimumab of £159 compared with the value of £126 used in the base case, had the effect of increasing the ICER by just under £3,000 per QALY.

Increasing the vial price slightly to a maximum expected vial price limit led to increasing the base case ICER by £6,887 to £71,297 per QALY.

6.9.5 Were any obvious subgroups not considered? If so, which ones, and why were they not considered? Please refer to the subgroups identified in the decision problem in section 4.

As discussed in Section 4, although the evidence shows that SLE is more common in females, the black African American, South Asian and Chinese populations, we have not specifically run cost-effectiveness analyses on these subgroups. The total pooled BLISS dataset comprised 94% females; 9% patients of black African-American ethnicity; and 21% of Asian ethnicity. A priori subgroup efficacy analyses for gender and ethnicity were conducted, and most subgroups demonstrated similar benefit in the primary outcomes with belimumab. It was observed that the primary response rate with belimumab in subjects of black race was lower than that observed for other races. However, this was not considered to be a robust finding due to the small numbers of patients of this race recruited in the studies and noteworthy imbalances in baseline characteristics between the belimumab and placebo groups which may have influenced outcomes and therefore makes interpretation difficult.

One other subgroup of the pooled total population was also investigated to try and identify a group that may show greater benefit than the total population. This subgroup comprises patients with positive anti-dsDNA and who also had low complement C3 or C4. This is the current proposed licensed population for belimumab. The criteria for this sub-population are indicative of SLE patients with high disease activity. However the level of disease activity for our proposed target subgroup for this decision problem is assessed as being higher still due to patients having to satisfy the additional criterion of an SS score of 10 or more. This alternative subgroup (anticipated licence population) will be larger in size than our selected subgroup, comprising 52% of the pooled total population. The key results from running the simulation in the cost-effectiveness model for this subgroup is summarised in Table 6.49 showing the regression estimated efficacy in terms of change in SS score at Week 52 and the percentage of patients continuing treatment after 24 weeks based on a decrease in SS score of 4 or more. These results show that the probability of treatment continuation at week 24 (55.9%) is lower than that seen in our selected “high disease activity”

subgroup (60.9%), however it was still higher than the probability seen for the pooled total population (52.4%). The ICER incorporating the base case assumptions for this alternative subgroup was also slightly higher compared with our selected subgroup, showing £66,170 per QALY. The patient selection criteria for this alternative subgroup was therefore assessed to be less appropriate if the aim is to target patients that will potentially gain the greatest benefit from belimumab and at the same time being mindful of limited NHS resource.

Table 6.49. Summary of key endpoints for an additional subgroup with high disease activity – Pooled BLISS-52 and BLISS-76 studies

	High Disease Activity Subgroup	Subgroup A: Anti-dsDNA +ve AND low (C3 or C4)
Week 52 SS score regression		
SS0 SoC	-0.349	-0.319
SS0 all belimumab	-0.343	-0.300
SS0 belimumab responders	-0.280	-0.306
Probability of continuing treatment after Week 24	60.9%	55.9%
Base case ICER	£64,410	£66,170

6.10 Interpretation of economic evidence

6.10.1 Are the results from this economic evaluation consistent with the published economic literature? If not, why do the results from this evaluation differ, and why should the results in the submission be given more credence than those in the published literature?

There are no published economic studies in SLE patients with high disease activity using treatment with belimumab or other human monoclonal antibody treatments. Therefore the results from this cost-effectiveness analysis cannot be compared with other studies.

The cost outcomes in the model can be validated by comparing the SoC arm to literature values. The annual direct costs of SLE in the UK have been reported by Sutcliffe et al. as £2,613 (1996 price level), corresponding to £3,357 converted to 2010 price level[6]. The model predicts £196,505 (undiscounted) direct medical costs over 30.47 undiscounted life-years resulting in £6,448 per life-year. This difference in costs can probably be explained by the differences in populations and study duration: 105 SLE patients were studied by Sutcliffe et al. with an average SLICC score of 1.1 over a duration of one year; the model studies patients over a lifetime horizon, and the mean SLICC at death is estimated as being 4.0. Since £4,875 per year is as a result of organ damage, the difference in SLICC score may explain the higher medical costs seen in our model.

For this decision problem, the use of belimumab in addition to usual standard of care treatments was compared to standard of care alone in patients with moderate to severe SLE with high disease activity. In the clinical trials belimumab significantly reduced disease activity after one year (measured by a composite endpoint which included SELENA-SLEDAI score). In the health economic model, the estimated increased life expectancy and QALYs seen for patients on belimumab were explained by lower disease activity scores associated with a decreased mortality risk and a higher quality of life. Lower

[6] Using the Consumer Price Index from OECD (http://stats.oecd.org/Index.aspx), accessed November 2, 2010

disease activity was also associated with decreased risk of organ damage, resulting in fewer occurrences of cardiovascular, peripheral vascular, pulmonary, renal and skin damage. However, due to prolonged life-expectancy and therefore prolonged exposure risk, occurrences were higher for all other organs.

An ICER of approximately £83,000 per QALY was observed for the pooled total population and a lower ICER of £64,410 per QALY was obtained for a subgroup of these patients, defined as experiencing particularly high disease activity, and whom we feel are most likely to benefit from this innovative medicine. Due to limited long-term outcomes data and lack of long-term evidence with belimumab in a population of SLE patients with the disease severity of interest for this decision problem, the inclusion of some assumptions on benefit used in the health economic model may be conservative. This may result in these stated ICERs being conservative.

6.10.2 Is the economic evaluation relevant to all groups of patients who could potentially use the technology as identified in the decision problem in section 4?

There is no evidence to suggest that any SLE patients identified in the decision problem should not benefit from treatment with belimumab. The evidence does suggest that the greatest benefit may be in the more severe SLE patients, with a high degree of disease activity and who are judged likely by clinicians to experience long-term complications from the disease. Clinicians will be assessing response to the treatment after around six months and so any patients who do not seem to be benefiting from belimumab will be withdrawn from this treatment.

6.10.3 What are the main strengths and weaknesses of the evaluation? How might these affect the interpretation of the results?

The main strengths of this evaluation comprise:

- Short-term clinical efficacy was based on two well-designed RCTs.

- It was observed that the model closely predicted the effect on SS score over the duration of the studies. This helps to validate the short-term accuracy of the model.
- The natural history model for SLE was developed following an extensive analysis of the JH cohort, a large, long-term, observational dataset. The model is therefore able to accurately predict the long-term course of the disease and captures the heterogeneity and complexity of SLE, where a patient's history plays an important role in the future disease course. It enabled detailed examination of the relationships between various risk factors, organ damage and mortality.
- The predictions of organ damage events over time and mortality were validated with a second longitudinal SLE database (Toronto cohort) and showed good predictive accuracy for most disease organ systems and mortality.
- A conservative assumption was made with respect to long-term effect of belimumab on disease activity levels (SS score). The difference in SS score may in fact increase over time whereas the assumption used in the analyses is that the difference observed at 52 weeks remains constant over time. As a result, beneficial effect on HRQL on long-term outcomes may be underestimated in the model compared with what may be observed over the long-term in UK clinical practice, thus the ICER may be conservative.
- Comprehensive sensitivity and scenario analyses have been performed using all available data.

The main weaknesses of this evaluation comprise:

- The lack of published evidence and access to data on disease progression on a sufficiently large (UK) SLE cohort with the severity of disease of particular interest to this decision problem is an important limitation. This would have considerably improved the robustness of the modeling of effects of disease activity on long-term outcomes. Comprehensive attempts were made to adjust the JH cohort to resemble a cohort more in line with the proposed target population for belimumab, however this cohort is likely to have recorded less organ damage and death events than a more severe SLE

population, thus the benefits of belimumab over the long-term may be underestimated and consequently our results may be conservative.

- Linked with this is the lack of long-term data to examine the effect of belimumab on disease activity, organ damage and long-term survival as the BLISS trials were of relatively short duration. Duration of survival is a key outcome for patients diagnosed with SLE with high disease activity.
- SLE flares, a common symptom in SLE were not simulated in the model. Measures of flare were considered for inclusion in the disease activity model however the SLEDAI Flare Index instrument was not collected in the JH database. An alternative measure of flare could have been used however this may have caused problems due to the correlation between flare and SLEDAI score. Due to the fact that the model uses the adjusted mean SLEDAI, disease activity is 'smoothed' over time, and a flare or relapse of activity cannot be shown. A decrease in frequency of flares due to belimumab will however also decrease the AMS over the treatment period. Therefore, although flares are not directly simulated in the model, some effect of decreasing flares is incorporated. However this "smoothed" effect may lead to underestimating the benefit of belimumab.
- During the internal validation exercise it was seen that the predicted incidence of mortality was slightly underestimated. The reason for the lower incidence of death may arise because the organ systems were modeled independently. Solutions to this problem were explored, however the complexity of statistical modeling required to account for this is considerable and would not have been possible within the timelines of this project. This may lead to a conservative estimate of cost-effectiveness.
- There are limitations to the mortality model in that it does not describe the rate of mortality for patients aged >65 years. As a consequence an adjustment is made in the cost-effectiveness model to allow the risk of mortality to increase in line with the general population at ages not represented in the JH data.
- Costs related to disease activity were modelled independently of the costs associated with organ damage. There is a risk that this approach leads to double counting of some costs. It is not likely that this will have an impact on

the results since the difference in disease activity costs is minor. In addition, the cost of SoC was not included in this model. It could be argued that this approach might underestimate the cost difference between the treatment arms because of the estimated additional survival for belimumab patients. However as showed in the sections relating to additional scenario analyses this did not have a major impact on the ICER.

- The EQ-5D may not be the most sensitive generic instrument to detect all aspects of SLE on patient HRQL. For example, the benefit of belimumab on reducing chronic fatigue, a very common debilitating symptom associated with this disease, is very likely to have been underestimated in the cost-effectiveness assessment as the impact is not directly captured within the EQ-5D.
- Some less tangible aspects of the disease considered important have not been included in the cost-effectiveness assessment. Particularly for the more severe SLE patients, their inability to work and their reliance on carers, carries both a financial burden and will impact significantly on their mental wellbeing.

6.10.4 What further analyses could be undertaken to enhance the robustness/completeness of the results?

Overall, we have run a number of sensitivity and scenario analyses. The main challenge has been the limited evidence available on long-term progression of the disease, particularly in a subgroup of more severe SLE patients of particular interest to this decision problem. Considerable effort has been made to make the models used linking short-term to long-tem outcomes as robust and reliable as possible. It seems unlikely that additional analyses would contribute greatly to the decision making process.

Section C – Implementation

7 Assessment of factors relevant to the NHS and other parties

The purpose of this section is to provide an analysis of any factors relevant to the NHS and other parties that may fall outside the remit of the assessments of clinical effectiveness and cost effectiveness. This will allow the subsequent evaluation of the budget impact analysis. Such factors might include issues relating to service organisation and provision, resource allocation and equity, societal or ethical issues, plus any impact on patients or carers.

7.1 How many patients are eligible for treatment in England and Wales? Present results for the full marketing authorisation/CE marking and for any subgroups considered. Also present results for the subsequent 5 years.

Based on the proposed licensed population for belimumab, it is estimated that 6,593 patients across England and Wales will be eligible for belimumab. However, we propose that belimumab would be used in a subgroup of SLE patients. These patients have evidence for serological disease activity (low complement and positive anti-dsDNA) and additionally have high disease activity as indicated by a SELENA-SLEDAI disease activity score ≥10. Patients in this subgroup experienced an additional treatment effect to belimumab over and above the licensed population (see section 5.3.7). This equates to 4,150 patients across England and Wales (see Table 2.2).

Table 7.1. Eligible patient population and proposed subgroup

		Year 1	Year 2	Year 3	Year 4	Year 5
		2011	2012	2013	2014	2015
Population England and Wales (Office for National Statistics 2009)		55,601,320	55,993,805	56,387,650	56,781,482	57,175,519
		27,412,029	27,624,990	27,837,782	28,049,645	28,260,544
		28,189,291	28,368,815	28,549,868	28,731,837	28,914,975
Number of patients with SLE (prevalence: 71 patients per 100,000 females; 10 patients per 100,000 males. Incidence: 4 patients per 100,000) (Nightingale et al. 2007)		22,756	24,995	27,251	29,522	31,809
		2,741				
		20,014				
Number of patients with active disease (58%)# (Caseload Data 2010)		13,198	14,497	15,805	17,123	18,449
Proposed licensed population Patients with a high degree of disease activity (e.g. positive anti-dsDNA, low complement) (52% of Phase 3 trial population)* (GlaxoSmithKline data on file 2011)		6,593	7,241	7,895	8,553	9,215
Subgroup Patients with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥10 (34% of Phase 3 trial population)* (GlaxoSmithKline data on file 2011)		4,151	4,842	5,114	5,388	5,663

includes all ages; * Includes only patients aged 18 years or over

7.2 What assumption(s) were made about current treatment options and uptake of technologies?

As discussed above, there are approximately 4000 patients in the proposed high disease activity subgroup (low complement, positive anti-dsDNA and SELENA-

SLEDAI ≥10). Currently the usage of rituximab is estimated to be approximately 600-700 patients in the United Kingdom. However this is likely to be a significant underestimate as rituximab has a variety of licenses (oncology and rheumatoid arthritis). Rheumatoid arthritis shares many of the same symptoms as SLE and may be managed by similar clinicians. Rituximab also has positive NICE guidance for the treatment of rheumatoid arthritis (TAR 195).
We assume that of the eligible patient population only 50% will receive belimumab, therefore approximately 2000 patients in Year 1. This takes into account both the estimate based on UK epidemiology and belimumab Phase 3 trials distribution as well as the current rituximab usage.

Table 7.2. Eligible patient population and anticipated usage

	Year 1	Year 2	Year 3	Year 4	Year 5
	2011	2012	2013	2014	2015
Total eligible population	4151	4842	5114	5388	5663
Likely usage (50%)	2075	2421	2557	2694	2832

7.3 What assumption(s) were made about market share (when relevant)?

We have assumed belimumab would be given to the anticipated population described in Question 7.2, as this already includes an assumption about likely usage. As belimumab will be the only licensed biologic we present the base case where belimumab is used in 100% of patients and an alternate scenario where belimumab is compared to the cost of those patients receiving rituximab.

7.4 In addition to technology costs, please consider other significant costs associated with treatment that may be of interest to commissioners (for example, procedure codes and programme budget planning).

The administration cost of £126 for belimumab was calculated based on two hours of senior hospital staff nurse time (£63/hr) from PSSRU Unit Costs of Health and Social Care 2010. Two hours is considered appropriate due to one hour required for the actual infusion and another hour for patient preparation and monitoring post-infusion. An alternative method of determining an infusion administration cost is to use the day case costs for "Inflammatory Spine, Joint or Connective Tissue Disorders without complications" (HRG=HD23C) from the NHS tariff costs 2009/10, which is £432 per day. Adjusting this cost to obtain an estimated cost for two hours gives £115 (i.e. £432 per day/7.5*2). The highest cost of these two methods has been used in the model for each administration of the infusion i.e. £126. Belimumab is administered on days 0, 14 and 28, and at every 4-week interval thereafter. Therefore a patient will receive 14 infusions in year 1 and 13 infusions in year 2.

The recommended method of administration for rituximab for the first infusion is an initial rate for infusion is 50 mg/hr; after the first 30 minutes, it can be escalated in 50 mg/hr increments every 30 minutes, to a maximum of 400 mg/hr. For subsequent infusions rituximab can be infused at an initial rate of 100 mg/hr, and increased by 100 mg/hr increments at 30 minutes intervals, to a maximum of 400 mg/hr. This results in between 4-5hours infusion time. If you consider patient preparation and monitoring post-infusion (recommended in SPC), the total time of administration could be between 5-6hours, we have used a conservative assumption of 5 hours infusion and proportion of the HRG cost £288 (i.e. £432 per day/7.5*5). Rituximab was administered on days 1, 15, 168 and 182 in the Phase 3 clinical trial. Therefore patients receive 4 infusions per year.

7.5 What unit costs were assumed? How were these calculated? If unit costs used in health economic modelling were not based on national reference costs or the PbR tariff, which HRGs reflected activity?

See section 6.5. In order to keep the budget impact straightforward and conservative, it will focus on drug acquisition costs and administration, see question 7.4.

7.6 Were there any estimates of resource savings? If so, what were they?

N/A.

7.7 What is the estimated annual budget impact for the NHS in England and Wales?

Table 7.3. Estimated annual budget impact for the NHS in England and Wales

	Year 1	Year 2	Year 3	Year 4	Year 5
	2011	2012	2013	2014	2015
Total eligible population	4151	4842	5114	5388	5663
Likely usage (50%)	2075	2421	2557	2694	2832
Total cost of belimumab	£21,302,157	£24,686,898	£26,066,236	£27,454,704	£28,852,348
Total cost of belimumab (PAS)	£ [redacted]	£ [redacted]	£ [redacted]	£ [redacted]	£ [redacted]
Total cost of rituximab	£16,888,144	£19,701,522	£20,807,842	£21,921,491	£23,042,503
Difference between belimumab (PAS) and rituximab	£ [redacted]	£ [redacted]	£ [redacted]	£ [redacted]	£ [redacted]

7.8 Are there any other opportunities for resource savings or redirection of resources that it has not been possible to quantify?

Although it may be difficult to calculate any resources savings due to uncertainty, if belimumab is used in a similar patient population to those patients receiving rituximab, it is likely to represent a [redacted]

8 References

See Section 9 at end of document.

9 Appendices

9.1 Appendix 1

9.1.1 SPC/IFU, scientific discussion or drafts.

9.2 Appendix 2: Search strategy for section 5.1 (Identification of studies)

The following information should be provided.

9.2.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- The Cochrane Library.

Embase, Medline & Medline (R) In-Process and the Cochrane library (incorporating the Cochrane Database of Systematic Reviews (Cochrane Reviews), the Database of Abstracts of Reviews of Effects (DARE), the Cochrane Central Register of Controlled Trials (CENTRAL), the Health Technology Assessment Database (HTA), the NHS Economic Evaluation Database (NHS EED).

9.2.2 The date on which the search was conducted.

All searches were conducted on 8 December 2010. Publications published after this end date were excluded.

9.2.3 The date span of the search.

Medline & Medline (R) In-Process: 1950 to present day
Embase: 1980 to present day
No date restrictions were imposed on other databases

9.2.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

All the following searches were combined and inclusion/exclusion criteria applied.

Embase searched 8th December 2010

Medline & Medline (R) In-Process searched 8th December 2010

Cochrane library searched 8th December 2010

ID	Search	Hits
#1	MeSH descriptor **Lupus Erythematosus, Systemic** explode all trees	409
#2	SLE or lupus	912
#3	(#1 OR #2)	912
#4	belimumab or benlysta	5
#5	rituximab or rituxan or mabthera	559
#6	MeSH descriptor **Mycophenolic Acid** explode all trees	620
#7	cellcept or myfortic or mycophenolate mofetil	1111
#8	MeSH descriptor **Prednisolone** explode all trees	3119
#9	MeSH descriptor **Prednisone** explode all trees	2563
#10	prednisolone or prednisone	7655
#11	MeSH descriptor **Hydroxychloroquine** explode all trees	118
#12	plaquenil or hydroxychloroquine	218
#13	MeSH descriptor **Azathioprine** explode all trees	982
#14	Azasan or Imuran or Azamun or Imurel or azathioprine	1903
#15	MeSH descriptor **Cyclophosphamide** explode all trees	3574
#16	Endoxan or Cytoxan or Neosar or Procytox or Revimmune or cyclophosphamide	6390
#17	MeSH descriptor **Methotrexate** explode all trees	2463
#18	amethopterin or methotrexate	4472
#19	(#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR	17973

	#17 OR #18)	
#20	(#3 AND #19)	384

Two Cochrane groups were identified and excluded.

9.2.5 Details of any additional searches, such as searches of company databases (include a description of each database).

Additional studies were identified by hand searching the following resources:

- Reference lists of previous trials and systematic reviews
    - The reference lists of relevant studies retrieved for full review were manually checked for any additional references that had not been identified by the search strategies
    - The reference lists of systematic reviews and relevant qualitative reviews were also checked for additional references that had not been identified by the search strategies
- Conference proceedings (2006–2010):
    - American College of Rheumatology (ACR)
    - European League Against Rheumatism (EULAR)
    - British Society of Rheumatology
- Clinical trial registries
    - clinicaltrials.gov
    - ISRCTN Register
    - UK Clinical Trials Gateway
    - metaRegister (mRCT) of Controlled Trials

Where the systematic review identified publications based on GlaxoSmithKline studies, we have augmented information in this submission with unpublished data.

9.2.6 The inclusion and exclusion criteria.

Inclusion and exclusion criteria were applied as follows.

Inclusion parameters:

- Patient population

  Adults (≥ 18 years) with SLE. Studies were also included if SLE patients had kidney involvement

- Efficacy and safety outcomes

  - Change in SELENA-SLEDAI score
  - Change in BILAG score
  - Change in PGA (physician global assessment scale)
  - Change in SLICC score
  - Change in number/frequency of flares
  - Quality of life
  - Mortality
  - Reduction in steroid use
  - Medical resource utilisation
  - Fatigue (e.g. FACIT score)
  - Adverse events including:
    - Incidence and severity (grade) of all adverse events (AEs) reported
    - Withdrawals due to AEs
    - SAEs

- Interventions/treatments

  - Belimumab
  - Rituximab
  - Mycophenolate mofetil

- Prednisolone and other steroids
- Hydroxychloroquine and other antimalarials
- Azathioprine
- Cyclophosphamide
- Methotrexate

- Study design

  Randomised controlled trials

- Language restrictions

  English language only

Exclusion parameters:

- Population
  - Studies enrolling patients with only active lupus nephritis were excluded
- Interventions
  - Non-specified
- Outcomes
  - Non-specified
- Study design
  - Designs other than RCT
- Language restrictions
  - Publications in languages other than English

9.2.7 The data abstraction strategy.

A pre-determined data extraction table (DET) was designed in Microsoft Excel®.

Two reviewers extracted data in parallel from included publications.

Discrepancies in extraction were reviewed and resolved by a third party.

9.3 Appendix 3: Quality assessment of RCT(s) (section 5.4)

9.3.1 A suggested format for the quality assessment of RCT(s) is shown below.

C1057 (BLISS-52)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Patients who underwent all screening procedures and met the entry criteria were enrolled in the study and assigned to treatment by use of a central interactive voice response system. Patients were randomised in a 1:1:1 ratio to placebo, or belimumab 1 mg/kg or 10 mg/kg. Randomisation was stratified according to the SELENA-SLEDAI score (6–9 vs ≥10), proteinuria concentration (<2 g/24 h vs ≥2 g/24 h) at screening, and ethnic origin (African descent or indigenous American [Alaska Native or American Indian from North, South, or Central America] vs other).	Yes
Was the concealment of treatment allocation adequate?	An unmasked pharmacist prepared unmarked infusion bags for administration. Belimumab and placebo were both prepared as sterile and lyophilised vials (5 mL for belimumab 1 mg/kg; 20 mL for belimumab 10 mg/kg and placebo), and contained the same formulations, except without the active drug for placebo.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The three groups did not differ in any of the main baseline characteristics.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Patients, investigators, study coordinators, and sponsors were masked to treatment assignment during intravenous administration of the drug and assessment of the patients every 4 weeks during the 52-week trial until the database was locked.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	The three groups did not differ in reasons for discontinuation of treatment.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	The authors reported outcomes as specified in the study protocol.	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Analysis was done in a modified intention-to-treat population, defined as all randomly assigned patients who received a dose of the study drug. This was appropriate and appropriate methods for handling missing data were outlined in the clinical study report.	Yes

C1056 (BLISS-76)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Patients who underwent all screening procedures and met the entry criteria were enrolled in the study and assigned to treatment by use of a central interactive voice response system. Patients were randomised in a 1:1:1 ratio to placebo, or belimumab 1 mg/kg or 10 mg/kg. Randomisation was stratified according to the SELENA-SLEDAI score (6–9 vs ≥10), proteinuria concentration (<2 g/24 h vs ≥2 g/24 h) at screening, and ethnic origin (African descent or indigenous American [Alaska Native or American Indian from North, South, or Central America] vs other).	Yes
Was the concealment of treatment allocation adequate?	An unmasked pharmacist prepared unmarked infusion bags for administration. Belimumab and placebo were both prepared as sterile and lyophilised vials (5 mL for belimumab 1 mg/kg; 20 mL for belimumab 10 mg/kg and placebo), and contained the same formulations, except without the active drug for placebo.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Baseline demographics, SLE disease characteristics, and medications were generally well balanced across treatment groups.	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Patients, investigators, study coordinators, and sponsors were masked to treatment assignment during intravenous administration of the drug and assessment of the patients every 4 weeks during the 52-week trial until the database was locked.	Yes

Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	There were no differences among groups in discontinuation rates.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	The authors reported outcomes as specified in the study protocol.	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Analysis was done in a modified intention-to-treat population, defined as all randomly assigned patients who received a dose of the study drug. This was appropriate and appropriate methods for handling missing data were outlined in the clinical study report.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

9.4 Appendix 4: Search strategy for section 5.7 (Indirect and mixed treatment comparisons)

The following information should be provided.

9.4.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- The Cochrane Library.

A specific search was not conducted. Identification of comparator clinical data was done by selecting the relevant papers from the identified studies from the search of the systematic review. See Section 9.2.1.

9.4.2 The date on which the search was conducted.

See Section 9.2.2.

9.4.3 The date span of the search.

See Section 9.2.3.

9.4.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

See Section 9.2.4.

9.4.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

See Section 9.2.5.

9.4.6 The inclusion and exclusion criteria.

See Section 9.2.6.

9.4.7 The data abstraction strategy.

See Section 9.2.7.

9.5 Appendix 5: Quality assessment of comparator RCT(s) in section 5.7 (Indirect and mixed treatment comparisons)

9.5.1 A suggested format for the quality assessment of RCT(s) is shown below.

Study ID or acronym		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?		
Was the concealment of treatment allocation adequate?		
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?		
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?		
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?		
Is there any evidence to suggest that the authors measured more outcomes than they reported?		
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?		
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

9.6 Appendix 6: Search strategy for section 5.8 (Non-RCT evidence)

The following information should be provided.

9.6.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- The Cochrane Library.

Embase, Medline & Medline (R) In-Process and the Cochrane Central Register of Controlled Trials (CENTRAL).

9.6.2 The date on which the search was conducted.

All searches were conducted on 3 March 2011.

9.6.3 The date span of the search.

Medline & Medline (R) In-Process: 1950 to present day
Embase: 1980 to present day
No date restrictions were imposed on other databases

9.6.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

All the following searches were combined and inclusion/exclusion criteria applied.

Embase searched 3rd March 2011

# ▲	Searches	Results
1	lupus erythematosus.mp. or exp lupus erythematosus/	64432
2	(SLE or lupus).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer]	74437
3	1 or 2	74437
4	belimumab.mp. or BELIMUMAB/	395
5	Benlysta.mp. or benlysta/	20
6	4 or 5	396
7	Clinical study/	29967
8	Case control study/	50980
9	Family study/	9010
10	Longitudinal study/	42858
11	Retrospective study/	223904
12	Prospective study/	164022
13	Randomized controlled trials/	289969
14	12 not 13	141465
15	Cohort analysis/	94187
16	(Cohort adj (study or studies)).mp.	60725
17	(Case control adj (study or studies)).tw.	50604
18	(follow up adj (study or studies)).tw.	34776
19	(observational adj (study or studies)).tw.	32925
20	(epidemiologic$ adj (study or studies)).tw.	56470
21	(cross sectional adj (study or studies)).tw.	46753
22	or/7-11,14-21	724476

23	3 and 6	298
24	22 and 23	9

Medline & Medline (R) In-Process searched 3rd March 2011

# ▲	Searches	Results
1	exp Lupus Erythematosus, Systemic/ or lupus erythematosus.mp.	50896
2	(SLE or lupus).mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier]	59509
3	1 or 2	59509
4	(Benlysta or belimumab).mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier]	66
5	Epidemiologic studies/	4903
6	exp case control studies/	491775
7	exp cohort studies/	1072750
8	Case control.tw.	55280
9	(cohort adj (study or studies)).tw.	52767
10	Cohort analy$.tw.	2460
11	(Follow up adj (study or studies)).tw.	31644
12	(observational adj (study or studies)).tw.	26837
13	Longitudinal.tw.	103583
14	Retrospective.tw.	196312
15	Cross sectional.tw.	110959
16	Cross-sectional studies/	119713

17	or/5-16	1446724
18	3 and 4	46
19	17 and 18	1

Cochrane library searched 3rd March 2010

ID	Search	Hits
#1	MeSH descriptor **Lupus Erythematosus, Systemic** explode all trees	412
#2	SLE or lupus	928
#3	(#1 OR #2)	928
#4	belimumab or benlysta	6
#5	(#3 AND #4)	4

9.6.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

Additional studies were identified by handsearching the following resources:

- Reference lists of previous trials and systematic reviews
    - The reference lists of relevant studies retrieved for full review were manually checked for any additional references that had not been identified by the search strategies
    - The reference lists of systematic reviews and relevant qualitative reviews were also checked for additional references that had not been identified by the search strategies
- Conference proceedings (2006–2010):
    - American College of Rheumatology (ACR)
    - European League Against Rheumatism (EULAR)
    - British Society of Rheumatology
- Clinical trial registries
    - clinicaltrials.gov

- UK Clinical Trials Gateway

9.6.6 The inclusion and exclusion criteria.

Inclusion and exclusion criteria were applied as follows.

Inclusion parameters:

- Patient population

  Adults (≥ 18 years) with SLE. Studies were also included if SLE patients had kidney involvement

- Efficacy and safety outcomes

  - Change in SELENA-SLEDAI score
  - Change in BILAG score
  - Change in PGA (physician global assessment scale)
  - Change in SLICC score
  - Change in number/frequency of flares
  - Quality of life
  - Mortality
  - Reduction in steroid use
  - Medical resource utilisation
  - Fatigue (e.g. FACIT score)
  - Adverse events including:
    - Incidence and severity (grade) of all adverse events (AEs) reported
    - Withdrawals due to AEs
    - SAEs
- Interventions/treatments

- Belimumab was the intervention of interest versus any of the following comparators:
    - Rituximab
    - Mycophenolate mofetil
    - Prednisolone and other steroids
    - Hydroxychloroquine and other antimalarials
    - Azathioprine
    - Cyclophosphamide
    - Methotrexate

- Study design

  Non-RCTs (including, but not limited to: observational studies and experimental studies)

- Language restrictions

  English language only

Exclusion parameters:

- Population
    - Studies enrolling patients with only active lupus nephritis were excluded

- Interventions
    - Non-specified

- Outcomes
    - Non-specified

- Study design
    - RCT

- Language restrictions
    - Publications in languages other than English

9.6.7 The data abstraction strategy.

A pre-determined data extraction table (DET) was designed in Microsoft Excel®. Two reviewers extracted data in parallel from included publications. Discrepancies in extraction were reviewed and resolved by a third party.

9.7 Appendix 7: Quality assessment of non-RCT(s) in section 5.8 (Non-RCT evidence)

9.7.1 Please tabulate the quality assessment of each of the non-RCTs identified.

It is difficult to assess the quality of non-RCTs due to the lack of validated checklists; therefore we conducted qualitative appraisal for the on-going study LBLS99, based on the interim clinical study report. Study treatments are described in Section 5.2.7.

Summary

Study LBSL99 is a multi-centre, open-label, continuation trial of belimumab in patients with SLE who received a satisfactory response in the Phase 2 trial (LBSL02). The study aimed to evaluate the long-term safety of belimumab in subjects with SLE.

Key features

Patient recruitment: a total of 449 subjects were originally enrolled (in LBSL02) and of the 321 completers, 298 were enrolled and 296 were treated in the LBSL99 study. The selection/eligibility criteria were adequately described. All subjects were analysed for both efficacy and safety.

Patient characteristics: mean age was 42.6 ±11.5 years and 93.2% was female. 72% was White, 1.7% Asian, 22% Black/African American, 2% American Indian or Alaska Native, 0.7% was Native Hawaiian or other pacific

islander, 1.7% indicated more than 1 race. All patients were seropositive. Mean SELENA-SLEDAI score was 9.2 ± 4.55; mean PGA score 1.4 ±0.51. Mean duration of treatment exposure was 138.96 ± 52.14 weeks for the 296 participants.

Withdrawals and dropouts: Withdrawals and dropouts were adequately reported. The discontinuation rate was 3-9% per year.

Analyses: The frequency and severity of AEs, SAEs, and discontinuation from the study due to AEs were reported. Other outcomes included the SRI (SLE responder index) rate; frequency of 1 new BILAG A or 2 new BILAG B flares; new SFI (SLE flare index) flares; and autoantibody levels.

Completeness of reporting: Pre-specified outcomes were adequately reported. Patients were evaluated for tolerability and safety at 6-month intervals for the entire follow-up.

9.8 Appendix 8: Search strategy for section 5.9 (Adverse events)

The following information should be provided.

9.8.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- The Cochrane Library.

See Section 9.2.1.

9.8.2 The date on which the search was conducted.

See Section 9.2.2.

9.8.3 The date span of the search.

See Section 9.2.3.

9.8.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

See Section 9.2.4.

9.8.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

See Section 9.2.5.

9.8.6 The inclusion and exclusion criteria.

See Section 9.2.6.

9.8.7 The data abstraction strategy.

See Section 9.2.7.

9.9 Appendix 9: Quality assessment of adverse event data in section 5.9 (Adverse events)

9.9.1 Please tabulate the quality assessment of each of the non-RCTs identified.

No new studies were identified. Quality assessments of the studies identified are provided in Sections 9.3.1 and 9.5.1. LBSL02 has previously been identified. However, a quality assessment has not previously been performed due to the reasons outlined in Section 5.2.6. A quality assessment is provided below.

LBSL02		
Study question	How is the question addressed in the study?	Grade (yes/no/not clear/N/A)
Was randomisation carried out appropriately?	After subjects had undergone all screening procedures and had been determined to be eligible for the study, they were randomly assigned (via a centralised interactive voice-response system [IVRS]) to 1 of the following 4 treatment groups: • 1 mg/kg belimumab • 4 mg/kg belimumab • 10 mg/kg belimumab • placebo	Yes
Was the concealment of treatment allocation adequate?	An unmasked pharmacist prepared unmarked infusion bags for administration.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The four groups did not differ in any of the main baseline characteristics.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	During the course of the study, only the site pharmacist or designee (responsible for receiving and dispensing study agent) was unblinded to a subject's specific treatment assignment.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	The four groups did not differ in reasons for discontinuation of treatment.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	The authors reported outcomes as specified in the study protocol.	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Analysis was done in a modified intention-to-treat population, defined as all randomly assigned patients who received a dose of the study drug. This was appropriate and appropriate methods for handling missing data were outlined in the clinical study report.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

9.10 Appendix 10: Search strategy for cost-effectiveness studies (section 6.1)

The following information should be provided.

9.10.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- EconLIT
- NHS EED.

The following databases were searched:

- PubMed (MEDLINE, MEDLINE In-Process)
- Embase
- EconLit
- CRD Interface
  - (NHS Economic Evaluation Database (NHS EED):
  - Health Technology Assessment Database (HTA)
  - Database of Abstracts of Reviews of Effects (DARE)
- RePEc (Research papers in Economics) (http://repec.org/)
- ClinicalTrials.gov (http://clinicaltrials.gov/)
- American College of Rheumatology (ACR)
- US Food and Drug Administration (http://www.fda.gov/)
- European Medicines Agency (http://www.ema.europa.eu/)

9.10.2 The date on which the search was conducted.

All searches were conducted on 18th March 2011.

9.10.3 The date span of the search.

The resources were searched over the following time periods or for all records available to be searched at a specific point in time:

- PubMed (MEDLINE) - 18th March 2011
- Embase – 18th March 2011
- EconLit – 18th March 2011
- CRD Interface – 18th March 2011
    - (NHS Economic Evaluation Database (NHS EED):
    - Health Technology Assessment Database (HTA)
    - Database of Abstracts of Reviews of Effects (DARE)
- RePEc (Research papers in Economics) (http://repec.org/) – 18th March 2011
- ClinicalTrials.gov (http://clinicaltrials.gov/) – 18th March 2011
- American College of Rheumatology (ACR) – 18th March 2011
- US Food and Drug Administration (http://www.fda.gov/) -18th March 2011
- European Medicines Agency (http://www.ema.europa.eu/) – 18th March 2011

9.10.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

A search was carried out on the 18th March 2011 for any literature referring to the cost-effectiveness of belimumab (Benlysta) for treatment of systemic lupus erythemeatosus. The complete strategies used are presented below. Economic search filters were not applied when searching economic databases such as NHS EED. No date or language limits were applied to the search.

PUBMED

#1 lupus (61727)
#2 lupus[Title/Abstract] (49910)
#3 SLE[Title/Abstract] (19816)
#4 #1 OR #2 OR #3 (63300)
#5 belimumab[Title/Abstract] (55)
#6 ((benlysta[Title/Abstract]) OR HGS 1006[Title/Abstract]) OR lymphostat-B[Title/Abstract] (9)
#7 #5 OR #6 (60)
#8 #4 AND #7 (41)

#9 economics (524161)
#10 costs AND “cost analysis” (39803)
#11 economics AND hospital (133135)
#12 economics AND medical (132474)
#13 economics AND nursing (39662)
#14 economics AND pharmaceutical (23157)
#15 (((((((economic$[Title/Abstract]) OR cost[Title/Abstract]) OR costs[Title/Abstract]) OR costly[Title/Abstract]) OR costing[Title/Abstract]) OR price[Title/Abstract]) OR prices[Title/Abstract]) OR pricing[Title/Abstract]) OR pharmacoeconomic$[Title/Abstract] (333780)
#16 budget$ (21545)
#17 #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 (726270)
#18 #8 AND #17 (2)
#19 animals NOT humans (3445557)
#20 #18 NOT #19 (2)

EMBASE 18/03/2011

#1 exp AND ('lupus'/exp OR lupus) OR 'sle'/exp OR sle (58059)
#2 lupus:ab,ti (56558)
#3 sle:ab,ti (23537)
#4 #1 OR #2 OR #3 (74676)
#5 'belimumab'/exp OR belimumab (406)
#6 belimumab:rn,ab,ti (75)
#7 benlysta:rn,ab,ti OR hgs1006:rn,ab,ti OR 'lymphostat b':rn,ab,ti (10)
#8 #5 OR #6 OR #7 (408)
#9 #4 AND #8 (304)
#10 'health'/exp OR health AND ('economics'/exp OR economics) (580757)
#11 exp AND economic AND ('evaluation'/exp OR evaluation) (977)
#12 exp AND ('health'/exp OR health) AND care AND ('cost'/exp OR cost) (1259)
#13 exp AND ('pharmacoeconomics'/exp OR pharmacoeconomics) (567)
#14 #10 OR #11 OR #12 OR #13 (581232)
#15 econom$:ab,ti OR cost:ab,ti OR costs:ab,ti OR costly:ab,ti OR costing:ab,ti OR price:ab,ti OR prices:ab,ti OR pricing:ab,ti OR pharmacoeconomic$:ab,ti (332710)
#16 expenditure$:ab,ti (25936)
#17 (value NEAR/2 money):ab,ti (905)
#18 budget$:ab,ti (12794)
#19 #15 OR #16 OR #17 OR #18 (351581)
#20 #14 OR #19 (771475)
#21 editorial:pt (0)
#22 note:pt (0)
#23 letter:pt (0)
#24 #21 OR #22 OR #23 (0)
#25 #20 NOT #24 (771,475)

#26 rat:ab,ti OR rats:ab,ti OR mouse:ab,ti OR mice:ab,ti OR hamster:ab,ti OR hamsters:ab,ti OR animal:ab,ti OR animals:ab,ti OR dogs:ab,ti OR dog:ab,ti OR cats:ab,ti OR bovine:ab,ti OR sheep:ab,ti (2779978)
#27 exp AND ('animal'/exp OR animal) (207030)
#28 'nonhuman'/exp OR nonhuman (15213516)
#29 #26 OR #27 OR #28 (16222170)
#30 exp AND ('human'/exp OR human) (293118)
#31 exp AND ('human'/exp OR human) AND ('experiment'/exp OR experiment) (82191)
#32 #30 OR #31 (293118)
#33 #29 NOT (#29 AND #32) (15935989)
#34 #25 NOT #33 (264353)
#35 #9 AND #34 (1)

EconLit via AEA – 18/03/2011

#1 belimumab OR benlysta OR HGS1006 OR Lymphostat-B (0)

CRD - 18/03/2011

#1 belimumab (1)
#2 benlysta OR HGS 1006 OR lymphostat B (1)
#3 #1 or #2 (1)

RePEc – 18/03/2011

#1 belimumab OR benlysta OR HGS1006 OR lymphostat B (0)

ClinicalTrials.gov – 18/03/2011

#1 (belimumab OR benlysta OR HGS1006 OR lymphostat B) AND (cost* or economic* or pharmacoeconomic*) (0)

American College of Rheumatology -18/03/2011

#1 Belimumab (5)
#2 Benlysta (0)
#3 HGS1006 (0)
#4 lymphostat B (1)

US Food and Drug Administration

belimumab OR benlysta OR HGS1006 OR lymphostat B (6)

European Medicines Agency – 18/03/2011

belimumab OR benlysta OR HGS1006 OR lymphostat B (3)

Excluded Records Bibliography

The following reports were assessed for relevance and excluded from the review:

1. National Horizon Scanning Centre (NHSC). Belimumab (Benlysta) for active systemic lupus erythematosus 2009.
2. Wiglesworth, A.K., Ennis, K. M., and Kockler, D. R. Belimumab: A BLyS-Specific Inhibitor for Systemic Lupus Erythematosus. Annals of Pharmacotherapy. 2010; 44: 12 (1955-1961)

9.10.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

A search was conducted on the GlaxoSmithKline internal study tracking database for any relevant cost-effectiveness studies including belimumab and none were found.

9.11 Appendix 11: Quality assessment of cost-effectiveness studies (section 6.1)

Not applicable as no cost-effictiveness studies were identified.

9.12 Appendix 12: Search strategy for section 6.4 (Measurement and valuation of health effects)

The following information should be provided.

9.12.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- NHS Economic Evaluation Database (NHS EED)
- EconLIT.

Utility data was searched in the following databases:

- Health Technology Assessments (NICE, http://www.nice.org.uk/)
- Medline and Medline (R) In-Process through Pubmed (http://www.ncbi.nlm.nih.gov/pubmed)

9.12.2 The date on which the search was conducted.

11th November 2010 to 14th December 2010.

9.12.3 The date span of the search.

All publication dates were included (1950-2010), but recent articles (2000-2010) were preferred.

9.12.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

Utility data for organ damage was searched in Health Technology Assessments, using free text searches with the search strings detailed below.

If no useful data was found in Health Technology Assessments, Pubmed searches of Medline were performed with the following search terms:

Item	Pubmed search strategy
Cataract	("Cataract"[Mesh] OR "Cataract Extraction"[Mesh]) AND "Quality of Life"[Mesh]
Retinal damage / optic atrophy – Initial search terms*	Retinal change and "Lupus Erythematosus, Systemic"[Mesh] AND "Quality of Life"[Mesh]
Retinal damage / optic atrophy	"Diabetic Retinopathy"[Mesh] ; "Optic Nerve Diseases"[Mesh] AND "Quality of Life"[Mesh]
Cognitive impairment	cognitive impairment; "Psychotic Disorders"[Mesh] AND "Quality of Life"[Mesh]
OR major psychosis	"Psychotic Disorders"[Mesh] AND "Quality of Life"[Mesh]
Seizures requiring therapy for 6 months	"Seizures"[Mesh] ; "Epilepsy"[Mesh] AND "Quality of Life"[Mesh]
Cerebral vascular accident ever or resection (for causes other than malignancy)	NICE report
Cranial or peripheral neuropathy	cranial neuropathy ; peripheral neuropathy AND "Quality of Life"[Mesh]
Transverse myelitis – Initial search terms*	"Myelitis, Transverse"[Mesh] AND "Quality of Life"[Mesh]
Transverse myelitis	"Multiple Sclerosis"[Mesh] and UK AND "Quality of Life"[Mesh]
Glomerular filtration rate < 50%	"Lupus Nephritis"[Mesh] ; "Kidney Failure, Chronic"[Mesh] AND "Quality of Life"[Mesh] additional data from http://www.nhsbt.nhs.uk/ and Health Technology Assessments
Proteinuria > 3.5 gm / 24 h	
End-stage renal disease	
Pulmonary hypertension	"Hypertension, Pulmonary"[Mesh] AND "Quality of Life"[Mesh]
Pulmonary fibrosis	"Pulmonary Fibrosis"[Mesh] AND "Quality of Life"[Mesh]
Shrinking lung (on chest radiograph)	Shrinking lung AND "Quality of Life"[Mesh]
Pleural fibrosis (on chest radiograph)	pleural fibrosis AND "Quality of Life"[Mesh]
Pulmonary infarction or resection	"Pulmonary Infarction"[Mesh] ; "Pulmonary Embolism"[Mesh] AND "Quality of Life"[Mesh]
Angina or	All data from Health Technology Assessments, search terms: Angina, coronary artery bypass, Myocardial infarction
coronary artery bypass	
Myocardial infarction 1	
Myocardial infarction 2	
Cardiomyopathy (ventricular dysfunction)	"Cardiomyopathies"[Mesh] AND "Quality of Life"[Mesh]
Valvular disease (diastolic murmur, or a systolic murmur > 3/6)	"Heart Valve Diseases"[Mesh] AND "Quality of Life"[Mesh]

Item	Pubmed search strategy
Pericarditis x 6 months or pericardiectomy	“Pericarditis”[Mesh] AND “Quality of Life”[Mesh]
Claudication x 6 months	“Intermittent Claudication”[Mesh] AND “Quality of Life”[Mesh]
Minor tissue loss (pulp space)	(“tissue loss” OR “pulp space”) AND “Quality of Life”[Mesh]
Significant tissue loss ever (e.g. loss of digit or limb) (Score 2 if > one site)	(“Amputation”[Mesh] OR “ tissue loss”) AND “Quality of Life”[Mesh]
Venous thrombosis with swelling, ulceration or venous stasis	“Venous Thrombosis”[Mesh] AND “Quality of Life”[Mesh]
Infarction or resection of bowel below duodenum, spleen, liver or gall bladder ever, for whatever cause (score 2 if > one site)	(“bowel resection” OR “spleen resection” OR “liver resection” OR “gall bladder resection”) AND “Quality of Life”[Mesh]
Mesenteric insufficiency	“mesenteric insufficiency “ AND “Quality of Life”[Mesh]
Chronic peritonitis	“Peritonitis”[Mesh] AND “Quality of Life”[Mesh]
Stricture or upper gastrointestinal tract surgery ever	gastrointestinal tract surgery AND “Quality of Life”[Mesh]
Pancreatic insufficiency requiring enzyme replacement or with pseudocyst	pancreatic insufficiency AND “Quality of Life”[Mesh]
Muscle atrophy / weakness	Muscle atrophy AND “Quality of Life”[Mesh]
Deforming or erosive arthritis (including reducible deformities, excluding avascular necrosis)	“Arthritis, Rheumatoid”[Mesh] AND UK[All Fields] AND “Quality of Life”[Mesh]
Osteoporosis with fracture or vertebral collapse (excluding avascular necrosis)	“Osteoporosis”[Mesh] AND “Quality of Life”[Mesh]
Avascular necrosis	“Osteonecrosis”[Mesh] avascular necrosis AND “Quality of Life”[Mesh]
Osteomyelitis	“Osteomyelitis”[Mesh] AND “Quality of Life”[Mesh]
Scarring chronic alopecia	scarring alopecia AND “Quality of Life”[Mesh]
Extensive scarring or panniculum other than scalp and pulp space	scarring AND “Quality of Life”[Mesh]
Skin ulceration (not due to trombosis) for more than 6 months	“Skin Ulcer”[Mesh] AND “Quality of Life”[Mesh]

Item	Pubmed search strategy
Diabetes mellitus sufficient to regard some manner of intervention	"Great Britain"[Mesh and "Diabetes Mellitus"[Mesh] AND "Quality of Life"[Mesh]
Malignant tumors (excluding dysplasia) (Score 2 if > one site)	"Neoplasms"[Mesh] ; Malignancy AND "Quality of Life"[Mesh]

* When the initial search terms did not produce useful results, additional searches were performed with broader search terms

The following NICE reports were used to source utility data:

Organ	Item	Source
Ocular	Cataract	Sullivan PW, Ghushchyan V. Preference-Based EQ-5D index scores for chronic conditions in the United States. 2006;26(4):410-420.
Ocular	Retinal change	Black C, Cummins E, Royle P, Philip S, Waugh N. The clinical effectiveness and cost-effectiveness of inhaled insulin in diabetes mellitus: a systematic review and economic evaluation. 2007;11(33):1-126.
Neuropsychiatric (NP)	Psychosis	Barton GR, Hodgekins J, Mugford M, Jones PB, Croudace T, Fowler D. Measuring the benefits of treatment for psychosis: validity and responsiveness of the EQ-5D. 2009;195(2):170-177.
NP	Seizures	Wilby J, Kainth A, Hawkins N et al. Clinical effectiveness, tolerability and cost-effectiveness of newer drugs for epilepsy in adults: a systematic review and economic evaluation. 2005;9(15):1-iv.
NP	Stroke	Ara R, Pandor A, Stevens J, Rees A, Rafia R. Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation. 2009;13(34):1-118.
NP	Neuropathy	O'Connor AB. Neuropathic pain: quality-of-life impact, costs and cost effectiveness of therapy. 2009;27(2):95-112.
NP	Transverse myelitis	McCrone P, Heslin M, Knapp M, Bull P, Thompson A. Multiple sclerosis in the UK: service use, costs, quality of life and disability. 2008;26(10):847-860.
Renal	Renal	Liem YS, Bosch JL, Hunink MG. Preference-based quality of life of patients on renal replacement therapy: a systematic review and meta-analysis. 2008;11(4):733-741.
Pulmonary	Pulmonary hypertension	Zisman DA, Schwarz M, Anstrom KJ, Collard HR, Flaherty KR, Hunninghake GW. A controlled trial of sildenafil in advanced idiopathic pulmonary fibrosis. 2010;363(7):620-628.
Pulmonary	Pulmonary infarction	Simpson EL, Stevenson MD, Rawdin A, Papaioannou D. Thrombophilia testing in people with venous thromboembolism: systematic review and cost-effectiveness analysis. 2009;13(2):iii, ix-91.
Cardiovascular (CV)	Angina	Ara R, Pandor A, Stevens J, Rees A, Rafia R. Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation. 2009;13(34):1-118.
CV	MI	Ara R, Pandor A, Stevens J, Rees A, Rafia R. Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation. 2009;13(34):1-118.
CV	Cardiomyopathy	Clegg AJ, Scott DA, Loveman E et al. The clinical and cost-effectiveness of left ventricular assist devices for end-stage heart

Organ	Item	Source
		failure: a systematic review and economic evaluation. 2005;9(45):1-iv.
CV	Valvular disease	Clegg AJ, Scott DA, Loveman E et al. The clinical and cost-effectiveness of left ventricular assist devices for end-stage heart failure: a systematic review and economic evaluation. 2005;9(45):1-iv.
Peropheral Vascular (PV)	Claudication	Michaels J, Brazier J, Palfreyman S, Shackley P, Slack R. Cost and outcome implications of the organisation of vascular services. 2000;4(11):i-191.
PV	Significant tissue loss	Clarke P, Gray A, Holman R. Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). 2002;22(4):340-349.
PV	Thrombosis	Simpson EL, Stevenson MD, Rawdin A, Papaioannou D. Thrombophilia testing in people with venous thromboembolism: systematic review and cost-effectiveness analysis. 2009;13(2):iii, ix-91.
Gastrointestinal (GI)	Resection of gall bladder	Wilson E, Gurusamy K, Gluud C, Davidson BR. Cost-utility and value-of-information analysis of early versus delayed laparoscopic cholecystectomy for acute cholecystitis. 2010;97(2):210-219.
Musculoskeletal (MSK)	Arthritis	Chen YF, Jobanputra P, Barton P et al. Cyclooxygenase-2 selective non-steroidal anti-inflammatory drugs (etodolac, meloxicam, celecoxib, rofecoxib, etoricoxib, valdecoxib and lumiracoxib) for osteoarthritis and rheumatoid arthritis: a systematic review and economic evaluation. 2008;12(11):1-278, iii.
MSK	Arthritis	Maetzel A, Krahn M, Naglie G. The cost effectiveness of rofecoxib and celecoxib in patients with osteoarthritis or rheumatoid arthritis. 2003;49(3):283-292.
MSK	Avascular necrosis	Vale L, Wyness L, McCormack K, McKenzie L, Brazzelli M, Stearns SC. A systematic review of the effectiveness and cost-effectiveness of metal-on-metal hip resurfacing arthroplasty for treatment of hip disease. 2002;6(15):1-109.
MSK	Osteoporosis	Kanis JA, Stevenson M, McCloskey EV, Davis S, Lloyd-Jones M. Glucocorticoid-induced osteoporosis: a systematic review and cost-utility analysis. 2007;11(7):iii-xi, 1.
MSK	Osteoporosis	Stevenson M, Jones ML, De NE, Brewer N, Davis S, Oakley J. A systematic review and economic evaluation of alendronate, etidronate, risedronate, raloxifene and teriparatide for the prevention and treatment of postmenopausal osteoporosis. 2005;9(22):1-160.
Skin	Alopecia	Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. 2008;6:84.
Skin	Scarring	Chen CL, Kuppermann M, Caughey AB, Zane LT. A community-based study of acne-related health preferences in adolescents. 2008;144(8):988-994.
Skin	Ulceration	Sullivan PW, Ghushchyan V. Preference-Based EQ-5D index scores for chronic conditions in the United States. 2006;26(4):410-420.
Malignancy	Malignancy	Sullivan PW, Sculpher MJ, Ghushchyan VH, Slejko JF. Catalogue of EQ-5D scores for the UK. 2009;12(7):A398.

9.12.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

No additional searches were conducted.

9.12.6 The inclusion and exclusion criteria.

Inclusion criteria:

- Articles reporting utility
- Time Trade Off (TTO) value method, societal perspective
- UK data or value set

Exclusion criteria:

- Non-UK data unless UK data was unavailable

9.12.7 The data abstraction strategy.

Not applicable as a literature search rather than a formal systematic review was conducted for utilities on organ damage.

9.13 Appendix 13: Resource identification, measurement and valuation (section 6.5)

The following information should be provided.

9.13.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- NHS EED
- EconLIT.

Cost data was searched in the following databases:

- NICE Health Technology Assessments (http://www.nice.org.uk)
- Medline and Medline (R) In-Process through Pubmed (http://www.ncbi.nlm.nih.gov/pubmed)

9.13.2 The date on which the search was conducted.

11th November 2010 to 14th December 2010.

9.13.3 The date span of the search.

All publication dates were included (1950-2010), but recent articles (2000-2010) were preferred.

9.13.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

Cost data was searched in Health Technology Assessments, using free text searches with the search strings below. If no useful data was found in Health

Technology Assessments, Pubmed searches in Medline were performed with the following search terms:

General search terms: ("Costs and Cost Analysis"[Mesh] OR "Economics"[Mesh] OR "economics "[Subheading] OR "Cost-Benefit Analysis"[Mesh] OR "Health Care Costs"[Mesh] OR "Hospital Costs"[Mesh] OR "Drug Costs"[Mesh])

Combined with the following condition specific terms: (combined with AND)

Medical condition	Search terms
Cataract	
Retinal damage / optic atrophy - Initial search terms*	
Retinal damage / optic atrophy	
Cognitive impairment	
OR major psychosis	
Seizures requiring therapy for 6 months	
Cerebral vascular accident ever or resection (for causes other than malignancy)	NICE report
Cranial or peripheral neuropathy	
Transverse myelitis - Initial search terms*	
Transverse myelitis	"Multiple Sclerosis"[Mesh] and UK
Glomerular filtration rate < 50%	additional data from http://www.nhsbt.nhs.uk/ and NICE
Proteinuria > 3.5 gm / 24 h	
End-stage renal disease	
Pulmonary hypertension	
Pulmonary fibrosis	"Pulmonary Fibrosis"[Mesh]
Shrinking lung (on chest radiograph)	
Pleural fibrosis (on chest radiograph)	
Pulmonary infarction or resection	
Angina or	All data from Health Technology Assessments, search terms: Angina, coronary artery bypass, Myocardial infarction
coronary artery bypass	
Myocardial infarction primary	
Myocardial infarction subsequent	
Cardiomyopathy (ventricular dysfunction)	
Valvular disease (diastolic murmur, or a systolic murmur > 3/6)	
Pericarditis x 6 months or pericardiectomy	
Claudication x 6 months	
Minor tissue loss (pulp space)	tissue loss OR pulp space
Significant tissue loss ever (e.g. loss of digit or limb) (Score 2 if > one site)	
Venous thrombosis with swelling, ulceration or venous stasis	
Infarction or resection of bowel below duodenum, spleen, liver or gall bladder ever, for whatever cause (score 2 if > one site)	

Medical condition	Search terms
Mesenteric insufficiency	
Chronic peritonitis	
Stricture or upper gastrointestinal tract surgery ever	
Pancreatic insufficiency requiring enzyme replacement or with pseudocyst	
Muscle atrophy / weakness	
Deforming or erosive arthritis (including reducible deformities, excluding avascular necrosis)	"Arthritis, Rheumatoid"[Mesh] AND UK[All Fields]
Osteoporosis with fracture or vertebral collapse (excluding avascular necrosis)	NICE report "Osteoporosis"[Mesh]
Avascular necrosis	
Osteomyelitis	
Scarring chronic alopecia	
Extensive scarring or panniculum other than scalp and pulp space	
Skin ulceration (not due to trombosis) for more than 6 months	
Diabetes mellitus sufficient to regard some manner of intervention	
Malignant tumors (excluding dysplasia) (Score 2 if > one site)	

* *When the initial search terms did not produce useful results, additional searches were performed with broader search terms*

The following Health Technology assessments were used to source cost data:

Organ	Item	Source
Ocular	Cataract	Sach TH, Foss AJ, Gregson RM et al. Falls and health status in elderly women following first eye cataract surgery: an economic evaluation conducted alongside a randomised controlled trial. 2007;91(12):1675-1679. Sach TH, Foss AJ, Gregson RM et al. Second-eye cataract surgery in elderly women: a cost-utility analysis conducted alongside a randomized controlled trial. 2010;24(2):276-283.
Ocular	Retinal change	Black C, Cummins E, Royle P, Philip S, Waugh N. The clinical effectiveness and cost-effectiveness of inhaled insulin in diabetes mellitus: a systematic review and economic evaluation. 2007;11(33):1-126.
NP	Psychosis	Albon E, Tsourapas A, Frew E et al. Structural neuroimaging in psychosis: a systematic review and economic evaluation. 2008;12(18):iii-163.
NP	Seizures	Cockerell OC, Hart YM, Sander JW, Shorvon SD. The cost of epilepsy in the United Kingdom: an estimation based on the results of two population-based studies. 1994;18(3):249-260.
NP	Stroke	Ara R, Pandor A, Stevens J, Rees A, Rafia R. Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation. 2009;13(34):1-118.
NP	Transverse myelitis	McCrone P, Heslin M, Knapp M, Bull P, Thompson A. Multiple sclerosis in the UK: service use, costs, quality of life and disability. 2008;26(10):847-860.
Renal	Renal	Woodroffe R, Yao GL, Meads C et al. Clinical and cost-effectiveness of newer immunosuppressive regimens in renal transplantation: a systematic review and modelling study. 2005;9(21):1-iv.

Organ	Item	Source
Pulmonary	Pulmonary hypertension	Chen YF, Jowett S, Barton P et al. Clinical and cost-effectiveness of epoprostenol, iloprost, bosentan, sitaxentan and sildenafil for pulmonary arterial hypertension within their licensed indications: a systematic review and economic evaluation. 2009;13(49):1-320.
Pulmonary	Pulmonary infarction	Simpson EL, Stevenson MD, Rawdin A, Papaioannou D. Thrombophilia testing in people with venous thromboembolism: systematic review and cost-effectiveness analysis. 2009;13(2):iii, ix-91.
CV	Angina	Ara R, Pandor A, Stevens J, Rees A, Rafia R. Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation. 2009;13(34):1-118.
CV	MI	Ara R, Pandor A, Stevens J, Rees A, Rafia R. Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation. 2009;13(34):1-118.
CV	Cardiomyopathy	Hessel FP, Wegner C, Muller J, Glaveris C, Wasem J. Economic evaluation and survival analysis of immunoglobulin adsorption in patients with idiopathic dilated cardiomyopathy. 2004;5(1):58-63.
CV	Pericarditis	Ortega-Sanchez IR, Sniadack MM, Mootrey GT. Economics of cardiac adverse events after smallpox vaccination: lessons from the 2003 US Vaccination Program. 2008;46 Suppl 3:S168-S178.
PV	Significant tissue loss	Clarke P, Gray A, Legood R, Briggs A, Holman R. The impact of diabetes-related complications on healthcare costs: results from the United Kingdom Prospective Diabetes Study (UKPDS Study No. 65). 2003;20(6):442-450.
PV	Thrombosis	Simpson EL, Stevenson MD, Rawdin A, Papaioannou D. Thrombophilia testing in people with venous thromboembolism: systematic review and cost-effectiveness analysis. 2009;13(2):iii, ix-91.
GI	Resection of gall bladder	Wilson E, Gurusamy K, Gluud C, Davidson BR. Cost-utility and value-of-information analysis of early versus delayed laparoscopic cholecystectomy for acute cholecystitis. 2010;97(2):210-219.
MSK	Arthritis	Callaghan R, Prabu A, Allan RB et al. Direct healthcare costs and predictors of costs in patients with primary Sjogren's syndrome. 2007;46(1):105-111.
MSK	Avascular necrosis	Vale L, Wyness L, McCormack K, McKenzie L, Brazzelli M, Stearns SC. A systematic review of the effectiveness and cost-effectiveness of metal-on-metal hip resurfacing arthroplasty for treatment of hip disease. 2002;6(15):1-109.
MSK	Osteoporosis	Kanis JA, Stevenson M, McCloskey EV, Davis S, Lloyd-Jones M. Glucocorticoid-induced osteoporosis: a systematic review and cost-utility analysis. 2007;11(7):iii-xi, 1.
MSK	Osteoporosis	Stevenson M, Jones ML, De NE, Brewer N, Davis S, Oakley J. A systematic review and economic evaluation of alendronate, etidronate, risedronate, raloxifene and teriparatide for the prevention and treatment of postmenopausal osteoporosis. 2005;9(22):1-160.
Diabetes	Diabetes	Jonsson B. Revealing the cost of Type II diabetes in Europe. 2002;45(7):S5-12.
Malignancy	Malignancy	Hind D, Ward S, De NE, Simpson E, Carroll C, Wyld L. Hormonal therapies for early breast cancer: systematic review and economic evaluation. 2007;11(26):iii-xi, 1.

9.13.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

No additional searches were conducted.

9.13.6 The inclusion and exclusion criteria.

Inclusion criteria:

- Articles reporting yearly costs for condition
- UK data if possible

Exclusion criteria:

- Not reporting yearly costs, or reporting lifetime costs
- Non-UK data where UK data is available

9.13.7 The data abstraction strategy.

Not applicable as a literature search rather than a formal systematic review was conducted for costs on organ damage.

10 Related procedures for evidence submission

10.1 Cost-effectiveness models

NICE accepts executable economic models using standard software – that is, Excel, TreeAge Pro, R or WinBUGs. If you plan to submit a model in a non-standard package, NICE should be informed in advance. NICE, in association with the ERG, will investigate whether the requested software is acceptable, and establish if you need to provide NICE and the ERG with temporary licences for the non-standard software for the duration of the appraisal. NICE reserves the right to reject economic models in non-standard software. A fully executable electronic copy of the model must be submitted to NICE with full access to the programming code. Care should be taken to ensure that the submitted versions of the model program and the written content of the evidence submission match.

NICE will need to distribute an executable version of the model to consultees and commentators because it will be used by the Appraisal Committee to assist their decision-making. On distribution of the appraisal consultation document (ACD) or final appraisal determination (FAD), and the evaluation report produced after the first committee meeting, NICE will advise consultees and commentators by letter that the manufacturer or sponsor has developed a model as part of their evidence submission for this technology appraisal. The letter asks consultees to inform NICE if they wish to receive an electronic copy of the model. If a request is received, NICE will release the model as long as it does not contain information that was designated confidential by the model owner, or the confidential material can be redacted by the model owner without producing severe limitations on the functionality of the model. The letter to consultees indicates clearly that NICE will distribute an executable copy, that the model is protected by intellectual property rights, and can be used only for the purposes of commenting on the model's reliability and informing a response to the ACD or FAD.

Manufacturers and sponsors must ensure that all relevant material pertinent to the decision problem has been disclosed to NICE at the time of submission.

There will be no subsequent opportunity to submit information unless it has been specifically requested by NICE.

When making a submission, manufacturers and sponsors should check that:

- an electronic copy of the submission has been given to NICE with all confidential information highlighted and underlined
- an executable electronic copy of the economic model has been submitted
- the checklist of confidential information (provided by NICE along with invitation to submit) has been completed and submitted.

10.2 Disclosure of information

To ensure that the appraisal process is as transparent as possible, NICE considers it highly desirable that evidence pivotal to the Appraisal Committee's decisions should be publicly available. NICE recognises that because the appraisal is being undertaken close to the time of regulatory decisions, the status of information may change during the STA process. However, at the point of issuing the FAD or ACD to consultees and commentators, all the evidence seen by the Committee should be available to all consultees and commentators.

Under exceptional circumstances, unpublished evidence is accepted under agreement of confidentiality. Such evidence includes 'commercial in confidence' information and data that are awaiting publication ('academic in confidence'). Further instructions on the specification of confidential information, and its acceptability, can be found in the agreement between the Association of the British Pharmaceutical Industry (ABPI) and NICE (www.nice.org.uk).

When data are 'commercial in confidence' or 'academic in confidence', it is the manufacturer's or sponsor's responsibility to highlight such data clearly, and to provide reasons why they are confidential and the timescale within which they will remain confidential. The checklist of confidential information should be completed: if it is not provided, NICE will assume that there is no confidential

information in the submission. It is the responsibility of the manufacturer or sponsor to ensure that the confidential information checklist is kept up to date.

The manufacturer or sponsor must ensure that any confidential information in their evidence submission is clearly underlined and highlighted. NICE is assured that information marked 'academic in confidence' can be presented and discussed during the public part of the Appraisal Committee meeting. NICE is confident that such public presentation does not affect the subsequent publication of the information, which is the prerequisite allowing for the marking of information as 'academic in confidence'.

Please therefore underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise and information submitted under 'academic in confidence' in yellow.

The manufacturer or sponsor will be asked to supply a second version of the submission with any information that is to remain confidential removed. The confidential information should be 'blacked out' from this version, taking care to retain the original formatting as far as possible so that it is clear which data have been removed and where from. For further details on how the document should be redacted/stripped, see the checklist of confidential information.

The last opportunity to review the confidential status of information in an STA, before publication by NICE as part of the consultation on the ACD, is 2 weeks before the Appraisal Committee meeting; particularly in terms of 'academic in confidence' information. The 'stripped' version will be issued to consultees and commentators along with the ACD or FAD, and made available on NICE's website 5 days later.

It is the responsibility of the manufacturer or sponsor to ensure that the 'stripped' version of the submission does not contain any confidential information. NICE will ask manufacturers and sponsors to reconsider restrictions on the release of data if there appears to be no obvious reason for the restrictions, or if such restrictions would make it difficult or impossible for NICE to show the evidential basis for its guidance. Information that has been

put into the public domain, anywhere in the world, cannot be marked as confidential.

Confidential information submitted will be made available for review by the ERG and the Appraisal Committee. Confidential information may be distributed to all consultees with the permission of the manufacturer or sponsor. NICE will at all times seek to protect the confidentiality of the information submitted, but nothing will restrict the disclosure of information by NICE that is required by law (including in particular, but without limitation, the Freedom of Information Act 2000).

The Freedom of Information Act 2000, which came into force on 1 January 2005, enables any person to obtain information from public authorities such as NICE. The Act obliges NICE to respond to requests about the recorded information it holds, and it gives people a right of access to that information. This obligation extends to submissions made to NICE. Information that is designated as ‘commercial in confidence’ may be exempt under the Act. On receipt of a request for information, the NICE secretariat will make every effort to contact the designated company representative to confirm the status of any information previously deemed ‘commercial in confidence’ before making any decision on disclosure.

10.3 Equity and equality

NICE is committed to promoting equality and eliminating unlawful discrimination, including paying particular attention to groups protected by equalities legislation. The scoping process is designed to identify groups who are relevant to the appraisal and reflect the diversity of the population. NICE consults on whether there are any issues relevant to equalities within the scope of the appraisal, or if there is information that could be included in the evidence presented to the Appraisal Committee to enable them to take account of equalities issues when developing guidance.

Evidence submitters are asked to consider whether the chosen decision problem could be impacted by NICE’s responsibility in this respect, including

when considering subgroups and access to recommendations that use a clinical or biological criterion.

For further information, please see the NICE website (www.nice.org.uk/aboutnice/howwework/NICEEqualityScheme.jsp).

10. References

Abu-Shakra, M., Urowitz, M. B., Gladman, D. D. et al. Mortality studies in systemic lupus erythematosus. Results from a single center. I. Causes of death. Journal of Rheumatology 22[7], 1259-1264. 1995.

ACR Ad Hoc Committee on Systemic Lupus Erythematosus Guidelines. Guidelines for referral and management of systemic lupus erythematosus in adults. American College of Rheumatology Ad Hoc Committee on Systemic Lupus Erythematosus Guidelines. Arthritis and Rheumatism 42[9], 1785-1796. 1999.

ACR Ad Hoc Committee on Systemic Lupus Erythematosus Response Criteria. The American College of Rheumatology response criteria for systemic lupus erythematosus clinical trials: measures of overall disease activity. Arthritis Rheum 50[11], 3418-3426. 2004.

Alarcon, G. S., McGwin, G., Jr., Bastian, H. M. et al. Systemic lupus erythematosus in three ethnic groups. VII [correction of VIII]. Predictors of early mortality in the LUMINA cohort. LUMINA Study Group. Arthritis and Rheumatism 45[2], 191-202. 2001.

Alonso, J., Ferrer, M., Gandek, B. et al. Health-related quality of life associated with chronic conditions in eight countries: results from the International Quality of Life Assessment (IQOLA) Project. Quality of Life Research 13[2], 283-298. 2004.

Andrade, R., Sanchez, M. L., Alarcon, G. S. et al. Adverse pregnancy outcomes in women with systemic lupus erythematosus from a multiethnic US cohort: LUMINA (LVI) [corrected]. Clin Exp.Rheumatol. 26[2], 268-274. 2008.

Baker, K. P., Edwards, B. M., Main, S. H. et al. Generation and characterization of LymphoStat-B, a human monoclonal antibody that antagonizes the bioactivities of B lymphocyte stimulator. Arthritis and Rheumatism 48[11], 3253-3265. 2003.

Becker-Merok, A. & Nossent, H. C. Damage accumulation in systemic lupus erythematosus and its relation to disease activity and mortality. J Rheumatol. 33[8], 1570-1577. 2006.

Bernatsky, S., Boivin, J. F., Joseph, L. et al. Mortality in systemic lupus erythematosus. Arthritis and Rheumatism 54[8], 2550-2557. 2006.

Bernatsky, S., Boivin, J. F., Joseph, L. et al. An international cohort study of cancer in systemic lupus erythematosus. Arthritis and Rheumatism 52[5], 1481-1490. 2005.

Bernatsky, S., Ramsey-Goldman, R., Isenberg, D. et al. Hodgkin's lymphoma in systemic lupus erythematosus. Rheumatology.(Oxford) 46[5], 830-832. 2007.

Bertsias, G., Ioannidis, J. P., Boletis, J. et al. EULAR recommendations for the management of systemic lupus erythematosus. Report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics. Annals of the Rheumatic Diseases 67[2], 195-205. 2008.

Bertsias, G. K., Ioannidis, J. P., Boletis, J. et al. EULAR points to consider for conducting clinical trials in systemic lupus erythematosus: literature based evidence for the selection of endpoints. Annals of the Rheumatic Diseases 68[4], 477-483. 2009.

Bombardier, C., Gladman, D. D., Urowitz, M. B. et al. Derivation of the SLEDAI. A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. Arthritis and Rheumatism 35[6], 630-640. 1992.

Briggs, A., Claxton, K., & Sculpher, M. 2006, *Decision Modelling for Health Economic Evaluation* Oxford University Press.

Buyon, J. P., Petri, M. A., Kim, M. Y. et al. The effect of combined estrogen and progesterone hormone replacement therapy on disease activity in systemic lupus erythematosus: a randomized trial. Ann Intern.Med 142[12 Pt 1], 953-962. 21-6-2005.

Bykerk, V., Sampalis, J., & Esdaile, J. M. A randomized study of the effect of withdrawing hydroxychloroquine sulfate in systemic lupus erythematosus. The Canadian Hydroxychloroquine Study Group. N.Engl.J Med 324[3], 150-154. 17-1-1991.

Campbell, R., Jr., Cooper, G. S., & Gilkeson, G. S. Two aspects of the clinical and humanistic burden of systemic lupus erythematosus: mortality risk and quality of life early in the course of disease. Arthritis and Rheumatism 59[4], 458-464. 15-4-2008.

Cancro, M. P., D'Cruz, D. P., & Khamashta, M. A. The role of B lymphocyte stimulator (BLyS) in systemic lupus erythematosus. J.Clin.Invest 119[5], 1066-1073. 2009.

Caseload Data. Market Insight. 2010.

Cella, D., Yount, S., Sorensen, M. et al. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. J Rheumatol. 32[5], 811-819. 2005.

Cervera, R., Khamashta, M. A., Font, J. et al. Morbidity and mortality in systemic lupus erythematosus during a 10-year period: a comparison of early and late manifestations in a cohort of 1,000 patients. Medicine (Baltimore) 82[5], 299-308. 2003.

Chambers, S. A., Allen, E., Rahman, A. et al. Damage and mortality in a group of British patients with systemic lupus erythematosus followed up for over 10 years. Rheumatology.(Oxford) 48[6], 673-675. 2009.

Chatham, W., Weinstein, A., & Petri, M. Five-year safety and efficacy experience with belimumab, a BLyS-specific inhibitor, in patients with systemic lupus erythematosus (SLE). Annals of the Rheumatic Diseases 69[Suppl 3], 147. 2010.

Cheema, G. S., Roschke, V., Hilbert, D. M. et al. Elevated serum B lymphocyte stimulator levels in patients with systemic immune-based rheumatic diseases. Arthritis and Rheumatism 44[6], 1313-1319. 2001.

Clowse, M. E., Magder, L. S., Witter, F. et al. The impact of increased lupus activity on obstetric outcomes. Arthritis Rheum 52[2], 514-521. 2005.

Cooper, G. S., Treadwell, E. L., St Clair, E. W. et al. Sociodemographic associations with early disease damage in patients with systemic lupus erythematosus. Arthritis and Rheumatism 57[6], 993-999. 15-8-2007.

D'Cruz, D., Tanasescu, C., & Navarra, S. Belimumab, a BLyS-specific inhibitor, reduced disease activity, flares and prednisone use in patients with active seropositive SLE: phase 3 BLISS-52 study. Rheumatology 49[Suppl 1]. 2010.

D'Cruz, D. P., Khamashta, M. A., & Hughes, G. R. Systemic lupus erythematosus. Lancet 369[9561], 587-596. 17-2-2007.

Danchenko, N., Satia, J. A., & Anthony, M. S. Epidemiology of systemic lupus erythematosus: a comparison of worldwide disease burden. Lupus 15[5], 308-318. 2006.

Department of Health. NHS reference costs 2005-06. http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_062884 . 7-12-2006. 9-4-2011.

Dolan, P. Modeling valuations for EuroQol health states. Med Care 35[11], 1095-1108. 1997.

European Medicines Agency 2009, *Concept paper on the need for a guideline on the clinical investigation of medicinal products intended for treatment of systemic and cutaneous lupus erythematosus*.

Fortin, P. R., Abrahamowicz, M., Ferland, D. et al. Steroid-sparing effects of methotrexate in systemic lupus erythematosus: a double-blind, randomized, placebo-controlled trial. Arthritis and Rheumatism 59[12], 1796-1804. 15-12-2008.

Furie, R., Stohl, W., Ginzler, E. M. et al. Biologic activity and safety of belimumab, a neutralizing anti-B-lymphocyte stimulator (BLyS) monoclonal antibody: a phase I trial in patients with systemic lupus erythematosus. Arthritis Res.Ther. 10[5], R109. 2008.

Furie, R., Zamani, O., & Wallace, D. Belimumab, a BLyS-Specific Inhibitor, Reduced Disease Activity and Severe Flares in Seropositive SLE Patients: BLISS-76 Study Results through Wk 76. Arthritis Rheum 62[10], S606. 2010.

Furie, R. A., Petri, M. A., Wallace, D. J. et al. Novel evidence-based systemic lupus erythematosus responder index. Arthritis Rheum 61[9], 1143-1151. 15-9-2009.

Gill, J. M., Quisel, A. M., Rocca, P. V. et al. Diagnosis of systemic lupus erythematosus. Am Fam.Physician 68[11], 2179-2186. 1-12-2003.

Gladman, D., Ginzler, E., Goldsmith, C. et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. Arthritis Rheum 39[3], 363-369. 1996a.

Gladman, D., Urowitz, M., Fortin, P. et al. Systemic Lupus International Collaborating Clinics conference on assessment of lupus flare and quality of life measures in SLE. Systemic Lupus International Collaborating Clinics Group. Journal of Rheumatology 23[11], 1953-1955. 1996b.

Gladman, D. D., Urowitz, M. B., Goldsmith, C. H. et al. The reliability of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index in patients with systemic lupus erythematosus. Arthritis Rheum 40[5], 809-813. 1997.

Gladman, D. D., Urowitz, M. B., Kagal, A. et al. Accurately describing changes in disease activity in Systemic Lupus Erythematosus. Journal of Rheumatology 27[2], 377-379. 2000.

Gladman, D. D., Urowitz, M. B., Rahman, P. et al. Accrual of organ damage over time in patients with systemic lupus erythematosus. J Rheumatol. 30[9], 1955-1959. 2003.

GlaxoSmithKline data on file. 2010.

GlaxoSmithKline data on file. 2011.

Goldblatt, F., Chambers, S., Rahman, A. et al. Serious infections in British patients with systemic lupus erythematosus: hospitalisations and mortality. Lupus 18[8], 682-689. 2009.

Gordon, C., Sutcliffe, N., Skan, J. et al. Definition and treatment of lupus flares measured by the BILAG index. Rheumatology (Oxford) 42[11], 1372-1379. 2003.

Griffiths, B., Mosca, M., & Gordon, C. Assessment of patients with systemic lupus erythematosus and the use of lupus disease activity indices. Best.Pract.Res.Clin Rheumatol. 19[5], 685-708. 2005.

Hay, E. M., Bacon, P. A., Gordon, C. et al. The BILAG index: a reliable and valid instrument for measuring clinical disease activity in systemic lupus erythematosus. Q.J.Med 86[7], 447-458. 1993.

Ibanez, D., Gladman, D. D., & Urowitz, M. B. Adjusted mean Systemic Lupus Erythematosus Disease Activity Index-2K is a predictor of outcome in SLE. J Rheumatol. 32[5], 824-827. 2005.

Ibanez, D., Urowitz, M. B., & Gladman, D. D. Summarizing disease features over time: I. Adjusted mean SLEDAI derivation and application to an index of disease activity in lupus. J Rheumatol. 30[9], 1977-1982. 2003.

Isenberg, D. A. & Gordon, C. From BILAG to BLIPS--disease activity assessment in lupus past, present and future. Lupus 9[9], 651-654. 2000.

Johnson, R. A. & Wichern, D. W. 2002, *Applied multivariate statistical analysis* Prentice Hall.

Jolly, M. How does quality of life of patients with systemic lupus erythematosus compare with that of other common chronic illnesses? Journal of Rheumatology 32[9], 1706-1708. 2005.

Kalunian, K. & Joan, T. M. New directions in the treatment of systemic lupus erythematosus. Current Medical Research and Opinion 25[6], 1501-1514. 2009.

Kasitanon, N., Magder, L. S., & Petri, M. Predictors of survival in systemic lupus erythematosus. Medicine (Baltimore) 85[3], 147-156. 2006.

Kotzin, B. L. Systemic lupus erythematosus. Cell 85[3], 303-306. 3-5-1996.

Lau, C. S. & Mak, A. The socioeconomic burden of SLE. Nat.Rev.Rheumatol. 5[7], 400-404. 2009.

Manson, J. J. & Rahman, A. Systemic lupus erythematosus. Orphanet.J.Rare.Dis. 1, 6. 2006.

Manzi, S. Lupus update: perspective and clinical pearls. Cleveland Clinic Journal of Medicine 76[2], 137-142. 2009.

Merrill, J. T., Neuwelt, C. M., Wallace, D. J. et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis Rheum 62[1], 222-233. 2010a.

Merrill, J. T., Wallace, D. J., & Furie, R. A. Five-Year Experience with Belimumab, a BLyS-Specific Inhibitor, in Patients with Systemic Lupus Erythematosus (SLE). Arthritis and Rheumatism 62[10], S608. 2010b.

Mok, C. C. Biomarkers for lupus nephritis: a critical appraisal. J.Biomed.Biotechnol. 2010, 638413. 2010.

Molad, Y., Borkowski, T., Monselise, A. et al. Maternal and fetal outcome of lupus pregnancy: a prospective study of 29 pregnancies. Lupus 14[2], 145-151. 2005.

Monthly Index of Medical Specialities (MIMS) 2011, Haymarket Medical Media.

National Horizon Scanning Centre. Belimumab (Benlysta) for active systemic lupus erythematosus. http://www.haps.bham.ac.uk/publichealth/horizon/outputs/documents/2009/sept-dec/Belimumab.pdf . 2009.

Navarra, S., Ilianova, E., & Bae, S. C. Belimumab, a BLyS-specific inhibitor, reduced disease activity, flares and steroid use in patients with seropositive systemic lupus erythematosus (SLE): BLISS-52 study. Ann Rheum Dis 69[Suppl 3], 555. 2010.

Navarra, S. V., Guzman, R. M., Gallacher, A. E. et al. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. Lancet 377[9767], 721-731. 26-2-2011.

Nightingale, A. L., Farmer, R. D., & de Vries, C. S. Systemic lupus erythematosus prevalence in the UK: methodological issues when using the General Practice Research Database to estimate frequency of chronic relapsing-remitting disease. Pharmacoepidemiol.Drug Saf 16[2], 144-151. 2007.

OECD. Annual Consumer Retail Indices (CPI0). http://stats.oecd.org. 2010a. 5-11-2010a.

OECD. Purchasing Power Parities for health. http://stats.oecd.org . 2010b. 5-11-2010b.

Office for National Statistics. 2008-based National Population Projections. http://www.statistics.gov.uk/downloads/theme_population/NPP2008/wUK08cc.xls . 21-10-2009.

Parikh-Patel, A., White, R. H., Allen, M. et al. Cancer risk in a cohort of patients with systemic lupus erythematosus (SLE) in California. Cancer Causes and Control 19[8], 887-894. 2008.

Petri, M., Buyon, J., & Kim, M. Classification and definition of major flares in SLE clinical trials. Lupus 8[8], 685-691. 1999.

Petri, M., Genovese, M., Engle, E. et al. Definition, incidence, and clinical description of flare in systemic lupus erythematosus. A prospective cohort study. Arthritis and Rheumatism 34[8], 937-944. 1991.

Petri, M., Kim, M. Y., Kalunian, K. C. et al. Combined oral contraceptives in women with systemic lupus erythematosus. N.Engl.J Med 353[24], 2550-2558. 15-12-2005.

Petri, M., Singh, S., Tesfasyone, H. et al. Prevalence of flare and influence of demographic and serologic factors on flare risk in systemic lupus erythematosus: a prospective study. J Rheumatol. 36[11], 2476-2480. 2009.

Petri, M., Stohl, W., Chatham, W. et al. Association of plasma B lymphocyte stimulator levels and disease activity in systemic lupus erythematosus. Arthritis and Rheumatism 58[8], 2453-2459. 2008.

Petri, M., Van Vollenhoven, R. F., & Zamani, O. Belimumab, a BLyS-Specific Inhibitor, Reduces Disease Activity and Severe Flares in Seropositive Systemic Lupus Erythematosus (SLE) Patients: BLISS-76 Study. Int J Rheum Dis 13[Suppl s1]. 2010.

PSSRU. Unit Costs of Health & Social Care 2007. http://www.pssru.ac.uk/pdf/uc/uc2007/uc2007.pdf . 2007. 9-4-2011.

Rahman, A. & Isenberg, D. A. Systemic lupus erythematosus. New England Journal of Medicine 358[9], 929-939. 28-2-2008.

Rahman, P., Gladman, D. D., & Urowitz, M. B. Clinical predictors of fetal outcome in systemic lupus erythematosus. J Rheumatol. 25[8], 1526-1530. 1998.

Roman, M. J., Shanker, B. A., Davis, A. et al. Prevalence and correlates of accelerated atherosclerosis in systemic lupus erythematosus. New England Journal of Medicine 349[25], 2399-2406. 18-12-2003.

Smolen, J. S., Strand, V., Cardiel, M. et al. Randomized clinical trials and longitudinal observational studies in systemic lupus erythematosus: consensus on a preliminary core set of outcome domains. Journal of Rheumatology 26[2], 504-507. 1999.

Stoll, T., Sutcliffe, N., Mach, J. et al. Analysis of the relationship between disease activity and damage in patients with systemic lupus erythematosus--a 5-yr prospective study. Rheumatology.(Oxford) 43[8], 1039-1044. 2004.

Stone, J. H. BLISS! Lupus learns its lessons. Lancet 377[9767], 693-694. 26-2-2011.

Strand, V. & Crawford, B. Improvement in health-related quality of life in patients with SLE following sustained reductions in anti-dsDNA antibodies. Expert.Rev.Pharmacoecon.Outcomes.Res. 5[3], 317-326. 2005.

Swaak, A. J., van den Brink, H. G., Smeenk, R. J. et al. Systemic lupus erythematosus: clinical features in patients with a disease duration of over 10 years, first evaluation. Rheumatology.(Oxford) 38[10], 953-958. 1999.

Tanasescu, C., Gallacher, A., & Garcia, M. Belimumab, a BLyS-specific inhibitor, significantly improved physical functioning, fatigue, and other health-related quality of life (HRQOL) measures in patients with seropositive systemic lupus erythematosus (SLE): BLISS-52 study. Ann Rheum Dis 69[Suppl 3], 556. 2010.

Terrier, B., Amoura, Z., Ravaud, P. et al. Safety and efficacy of rituximab in systemic lupus erythematosus: results from 136 patients from the French AutoImmunity and Rituximab registry. Arthritis Rheum 62[8], 2458-2466. 2010.

Thumboo, J., Fong, K. Y., Chan, S. P. et al. A prospective study of factors affecting quality of life in systemic lupus erythematosus. Journal of Rheumatology 27[6], 1414-1420. 2000.

Thumboo, J., Fong, K. Y., Ng, T. P. et al. Validation of the MOS SF-36 for quality of life assessment of patients with systemic lupus erythematosus in Singapore. Journal of Rheumatology 26[1], 97-102. 1999.

Thumboo, J. & Strand, V. Health-related quality of life in patients with systemic lupus erythematosus: an update. Ann.Acad.Med Singapore 36[2], 115-122. 2007.

Top Talk for Radiography Leaders. Radiology Indicative Tariff. Society of Radiography . 2005. 9-4-2011.

Tsakonas, E., Joseph, L., Esdaile, J. M. et al. A long-term study of hydroxychloroquine withdrawal on exacerbations in systemic lupus erythematosus. The Canadian Hydroxychloroquine Study Group. Lupus 7[2], 80-85. 1998.

Tseng, C. E., Buyon, J. P., Kim, M. et al. The effect of moderate-dose corticosteroids in preventing severe flares in patients with serologically active, but clinically stable, systemic lupus erythematosus: findings of a prospective, randomized, double-blind, placebo-controlled trial. Arthritis Rheum 54[11], 3623-3632. 2006.

Urowitz, M. B., Gladman, D., Ibanez, D. et al. Accumulation of coronary artery disease risk factors over three years: data from an international inception cohort. Arthritis and Rheumatism 59[2], 176-180. 15-2-2008.

van Vollenhoven, R., Zamini, O., & Wallace, D. Belimumab, a BLyS-specific inhibitor, reduces disease activity and severe flares in seropositive SLE patients: BLISS-76 study. Annals of the Rheumatic Diseases 69(Suppl 3), 74. 2010.

van, D., I, Kromhout, D., Geleijnse, J. M. et al. Evaluation of cardiovascular risk predicted by different SCORE equations: the Netherlands as an example. Eur.J Cardiovasc.Prev.Rehabil. 17[2], 244-249. 2010.

Wackerly, D. D., Mendenhall, W., & Schaeffer, R. L. 1996, *Mathematical statistics with applications* Duxbury Press.

Wailoo, A., Davis, S., & Tosh, J. The incorporation of health benefits in cost utility analysis using the EQ-5D. 15-11-2010.

Wallace, D. J., Stohl, W., Furie, R. A. et al. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. Arthritis and Rheumatism 61[9], 1168-1178. 15-9-2009.

Ware, J. E., Jr. & Sherbourne, C. D. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 30[6], 473-483. 1992.

Wells, G. A., Tugwell, P., Kraag, G. R. et al. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. Journal of Rheumatology 20[3], 557-560. 1993.

Wiglesworth, A. K., Ennis, K. M., & Kockler, D. R. Belimumab: a BLyS-specific inhibitor for systemic lupus erythematosus. Ann Pharmacother. 44[12], 1955-1961. 2010.

Wilson, E. C., Jayne, D. R., Dellow, E. et al. The cost-effectiveness of mycophenolate mofetil as firstline therapy in active lupus nephritis. Rheumatology (Oxford) 46[7], 1096-1101. 2007.

Yasmeen, S., Wilkins, E. E., Field, N. T. et al. Pregnancy outcomes in women with systemic lupus erythematosus. J Matern.Fetal Med 10[2], 91-96. 2001.

Yelin, E., Trupin, L., Katz, P. et al. Work dynamics among persons with systemic lupus erythematosus. Arthritis and Rheumatism 57[1], 56-63. 15-2-2007.

Zhang, J., Roschke, V., Baker, K. P. et al. Cutting edge: a role for B lymphocyte stimulator in systemic lupus erythematosus. Journal of Immunology 166[1], 6-10. 1-1-2001.

Zonana-Nacach, A., Barr, S. G., Magder, L. S. et al. Damage in systemic lupus erythematosus and its association with corticosteroids. Arthritis and Rheumatism 43[8], 1801-1808. 2000a.

Zonana-Nacach, A., Roseman, J. M., McGwin, G., Jr. et al. Systemic lupus erythematosus in three ethnic groups. VI: Factors associated with fatigue within 5 years of criteria diagnosis. LUMINA Study Group. LUpus in MInority populations: NAture vs Nurture. Lupus 9[2], 101-109. 2000b.

CONFIDENTIAL

10th June 2011

Dr Carole Longson
Director
Health Technology Evaluation Centre
National Institute for Health and Clinical Excellence
Level 1A, City Tower
Piccadilly Plaza
Manchester
M1 6NA

Quarry House
Quarry Hill
Leeds
LS2 7UE

Tel: 0113 254 5000

Dear Carole

Patient Access Scheme – Benlysta – systemic lupus erythematosus

I am writing to confirm the Department of Health's position on the Patient Access Scheme (PAS) arrangements that have been proposed by GlaxoSmithKline (GSK) for Benlysta (belimumab) in the treatment of systemic lupus erythematosus.

I understand that GSK have proposed the simple discount PAS on the condition that the level of discount offered through the scheme should remain confidential and should not be published in final NICE guidance.

GSK is yet to confirm with us the exact level of discount. We will reiterate to GSK that, whilst it was not necessary for DH to know the exact level of discount at this stage, they will need to inform us and, in particular, NICE, of the precise discounted price to the NHS for the appraisal to proceed.

The Department is content in this case for the confidential PAS proposal to be considered in the appraisal of Benlysta. NICE must of course be satisfied that sufficient information can be communicated to stakeholders to explain an appraisal recommendation. In this regard, what constitutes a sufficient level of transparency is a matter for the Institute to determine in developing its guidance. In addition, the NHS must have access to the discount price when final NICE guidance is made available, so Trusts and commissioners are able to properly account for the PAS.

Yours sincerely

Dr Helen Knight
Associate Director - Appraisals
National Institute for Health and Clinical Excellence
City Tower
Piccadilly Plaza
Manchester
M1 4BD

GlaxoSmithKline UK Ltd
Stockley Park West
Uxbridge
Middlesex
UB11 1BT

Tel. +44 (0)20 8990 9000
Fax. +44 (0)20 8990 4321
www.gsk.com

19th September 2011

Dear Dr Knight,

Re: NICE STA – Belimumab for the treatment of patients with SLE

Please accept this letter as confirmation that the final list vial prices for belimumab, which have been ratified through the PPRS procedure, are as follows:

120mcg vial = £121.50
400mcg vial = £405.00

We have submitted a patient access scheme (PAS) with our STA submission, which has been accepted by the Department of Health. This PAS offers a straight discount of approximately [redacted] on the belimumab vial acquisition cost. Incorporation of the PAS discount for our proposed high disease activity target population, based on the weight distribution seen for these patients in the two Phase 3 BLISS randomised controlled trials, provides the following vial prices:

120mcg vial = £[redacted]
400mcg vial = £[redacted]

Kind Regards

[redacted]

[redacted]
Health Outcomes Department
GlaxoSmithKline UK Ltd
Stockley Park West
Uxbridge
UB11 1BT
Tel; [redacted]

Section 9.29, Appendix 29

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Technology appraisals

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus - Patient access scheme submission template

13 April 2010

1 Introduction

The 2009 Pharmaceutical Price Regulation Scheme (PPRS) (www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpriceregulationscheme/2009PPRS) is a non-contractual scheme between the Department of Health and the Association of the British Pharmaceutical Industry. The purpose of the 2009 PPRS is to ensure that safe and cost-effective medicines are available on reasonable terms to the NHS in England and Wales. One of the features of the 2009 PPRS is to improve patients' access to medicines at prices that better reflect their value through patient access schemes.

Patient access schemes are arrangements which may be used on an exceptional basis for the acquisition of medicines for the NHS in England and Wales. Patient access schemes propose either a discount or rebate that may be linked to the number, type or response of patients, or a change in the list price of a medicine linked to the collection of new evidence (outcomes). These schemes help to improve the cost effectiveness of a medicine and therefore allow the National Institute for Health and Clinical Excellence (NICE) to recommend treatments which it would otherwise not have found to be cost effective. More information on the framework for patient access schemes is provided in the 2009 PPRS (www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpriceregulationscheme/2009PPRS.

Patient access schemes are proposed by a pharmaceutical company and agreed with the Department of Health, with input from the Patient Access Schemes Liaison Unit (PASLU) within the Centre for Health Technology Evaluation at NICE.

2 Instructions for manufacturers and sponsors

This document is the patient access scheme submission template for technology appraisals. If manufacturers and sponsors want the National Institute for Health and Clinical Excellence (NICE) to consider a patient access scheme as part of a technology appraisal, they should use this template. NICE can only consider a patient access scheme after formal referral from the Department of Health.

The template contains the information NICE requires to assess the impact of a patient access scheme on the clinical and cost effectiveness of a technology, in the context of a technology appraisal, and explains the way in which background information (evidence) should be presented. If you are unable to follow this format, you must state your reasons clearly. You should insert 'N/A' against sections that you do not consider relevant, and give a reason for this response.

Please refer to the following documents when completing the template:

- 'Guide to the methods of technology appraisal' (www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/guidetothemethodsoftechnologyappraisal.jsp)
- 'Specification for manufacturer/sponsor submission of evidence' (http://www.nice.org.uk/aboutnice/howwework/devnicetech/singletechnologyappraisalsubmissiontemplates.jsp) and
- Pharmaceutical Price Regulation Scheme 2009 (www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpriceregulationscheme/2009PPRS).

For further details on the technology appraisal process, please see NICE's 'Guide to the single technology appraisal (STA) process' and 'Guide to the multiple technology appraisal (MTA) process' (http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/technology_appraisal_process_guides.jsp). The

‘Specification for manufacturer/sponsor submission of evidence’ provides details on disclosure of information and equality issues.

Make the submission as brief and informative as possible. Only mark information as confidential when absolutely necessary. Sufficient information must be publicly available for stakeholders to comment on the full content of the technology appraisal, including details of the proposed patient access scheme. Send submissions electronically to NICE in Word or a compatible format, not as a PDF file.

Appendices may be used to include additional information that is considered relevant to the submission. Do not include information in the appendices that has been requested in the template. Appendices should be clearly referenced in the main submission.

When making a patient access scheme submission, include:

- an updated version of the checklist of confidential information, if necessary
- an economic model with the patient access scheme incorporated, in accordance with the ‘Guide to the methods of technology appraisal’ (www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/guidetothemethodsoftechnologyappraisal.jsp).

If you are submitting the patient access scheme at the end of the appraisal process, you should update the economic model to reflect the assumptions that the Appraisal Committee considered to be most plausible. No other changes should be made to the model.

3 Details of the patient access scheme

3.1 Please give the name of the technology and the disease area to which the patient access scheme applies.

Belimumab (Benlysta®) – Systematic Lupus Erythematosus (SLE)

3.2 Please outline the rationale for developing the patient access scheme.

SLE is a relapsing and remitting disease. It is a chronic condition associated with significant morbidity and mortality. Many patients with SLE experience general symptoms including fatigue, malaise, fever, anorexia, weight loss, skin rash and muscle and joint pain. SLE can lead to arthritis, kidney failure, heart and lung inflammation, neuropsychiatric disease, vasculitis, severe skin rash and blood dyscrasias such as anaemia, leucopoenia and thrombocytopenia (Manson et al. 2006). These all contribute towards a decrease in their quality of life.

SLE also has a substantial impact on employment, with over half of patients no longer working 15 years after diagnosis. It is associated with a 2.4-fold greater risk of mortality than the general population (Bernatsky et al. 2006).

SLE is more prevalent in women and African-Caribbean, South Asian and Chinese populations than in European white populations (Danchenko et al. 2006; Manzi 2009). The demographic of SLE patients is likely to include a significant portion of women of child-bearing age.

Patients with SLE are currently managed by a range of treatments (NSAIDs, corticosteroids, immunosuppressants and anti-malarials); variously used alone or in different combinations constitutes standard or care (SoC).

Current standard of care may be associated with undesirable effects, either from chronic use of steroids (osteoporosis, diabetes and cardiovascular disease) or side effects associated with immunosuppressants (infection, toxicity and infertility). Many of these treatments used are unlicensed and a significant number of patients with advanced SLE do not respond to current

treatments even at high doses. Those patients with more severe, active SLE are managed in tertiary centres and many routinely receive rituximab (MabThera®), an unlicensed biologic, which although appearing to have some benefit in clinical practice, failed to demonstrate efficacy in Phase 2/3 trials.

There has been little therapeutic innovation in treatments for SLE, with no evidence leading to the development of new licensed treatments for several decades.

Belimumab, a human IgG1λ monoclonal antibody that binds to soluble human B-lymphocyte stimulator (BLyS) and inhibits its biological activity, has been specifically developed for the treatment of SLE and demonstrated efficacy in two Phase 3 clinical trials, showing a significant degree of innovation in addressing an area of unmet need. As SLE is a relapsing remitting disease with long term consequences, the full clinical benefit may not be identified in the studies available at launch.

In order to make the PAS competitive while still reflecting the innovation and value GSK believe belumimab delivers, the proposed PAS would involve a straight discount from the NHS list price.

3.3 Please describe the type of patient access scheme, as defined by the PPRS.

The patient access scheme is a straight discount from the NHS list price of belimumab.

3.4 Please provide specific details of the patient population to which the patient access scheme applies. Does the scheme apply to the whole licensed population or only to a specific subgroup (for example, type of tumour, location of tumour)? If so:

- How is the subgroup defined?
- If certain criteria have been used to select patients, why have these have been chosen?
- How are the criteria measured and why have the measures been chosen?

As the patient access scheme is a straight discount and in order to remove any administrative burden on the NHS, the availability of the patient access scheme applies to the whole licensed population. However, being mindful of NHS resources, GSK proposes NICE consider issuing guidance on a specific subgroup (high disease activity subgroup).

The proposed subgroup is for adults with active autoantibody-positive systemic lupus erythematosus with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥10; this comprises 34% of the overall belimumab trial population, and we refer to this as the high disease activity subgroup.

The two serological markers, low complement and positive anti-dsDNA are objective measures used routinely in SLE and accessible to physicians in general practice. They are widely considered important measures of disease activity. Patients with positive anti-dsDNA and low complement are immunologically active and at higher risk for flares and lupus nephritis.

SELENA-SLEDAI (SS) is the efficacy component of the composite endpoint of the Phase 3 trials and measures disease activity. A score of ≥10 is likely to indicate a patient with highly active disease.

Details of this subgroup are presented in the main submission (Section 5.3.7 and Section 6.8).

3.5 Please provide details of when the scheme will apply to the population specified in 3.4. Is the scheme dependent on certain criteria, for example, degree of response, response by a certain time point, number of injections? If so:

- Why have the criteria been chosen?
- How are the criteria measured and why have the measures been chosen.

The scheme will apply to the licensed population outlined in 3.4 from the time they are deemed eligible for treatment with belimumab, i.e. they are receiving standard therapy and still require additional reduction in their disease activity.

Patients would have to be serologically active (low complement and positive anti-dsDNA), however given the more severe nature of these patients they would be being managed in tertiary care settings for a large portion of their care and clinicians have indicated that the serological status of patients is measured on a routine basis.

We are proposing that patients would have to have a SELENA-SLEDAI ≥10, an indication of highly active disease, to be eligible for treatment. The use of SELENA-SLEDAI is usually confined to clinical trials; however the majority of patients would have treatment initiated by a specialist in a tertiary setting. Most of these clinicians have indicated that they have familiarity with this measure and would be prepared to use it if it was a requirement to gain access to an innovative treatment for their SLE patients.

The current draft SPC for beluminab states in the "Posology and method of administration" section that *"The patient's condition should be evaluated continuously. Discontinuation of treatment with Benlysta should be considered if there is no improvement in disease control after 6 months of treatment."* This allows the assessment of adequate response to belimumab to be made on the basis of the physicians' clinical judgement after six months treatment. Six months is identified as a suitable time period after which to assess response to treatment as it allows sufficient time for the drugs mode of action to have an impact on the clinical manifestations of the disease. As mentioned above, physicians do not routinely measure SELENA-SLEDAI for disease activity or SLICC scores for organ damage in the clinical management of their SLE patients; they will assess response based on the general wellbeing of the patient and on how many disease flares they have experienced and of what severity.

In order to reflect the wording of the SPC, and the concept of "responders" or "non-responders" , a more objective assessment as to whether belimumab should be continued or discontinued after six months treatment was used. Therefore, the criterion of a SS disease activity score increase of 4 or greater, indicating a 'response' has been used in the economic modelling. This is discussed in more detail in Section 6.2.8.

3.6 What proportion of the patient population (specified in 3.4) is expected to meet the scheme criteria (specified in 3.5)?

The proposed subgroup is for adults with active autoantibody-positive systemic lupus erythematosus with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥10; this comprises 34% of the overall belimumab trial population.

The proportion of patients specified in 3.4 should be equivalent to those specified in 3.5.

As the patient access scheme is a straight discount and in order to remove any administrative burden on the NHS, the availability of the patient access scheme applies to the whole licensed population. However, being mindful of NHS resources, GSK proposes NICE consider issuing guidance on a specific subgroup (high disease activity subgroup). Therefore, while the scheme in principle would be available to the entire licensed population any NICE guidance is likely to be followed to ensure usage is in line with the proposed subgroup.

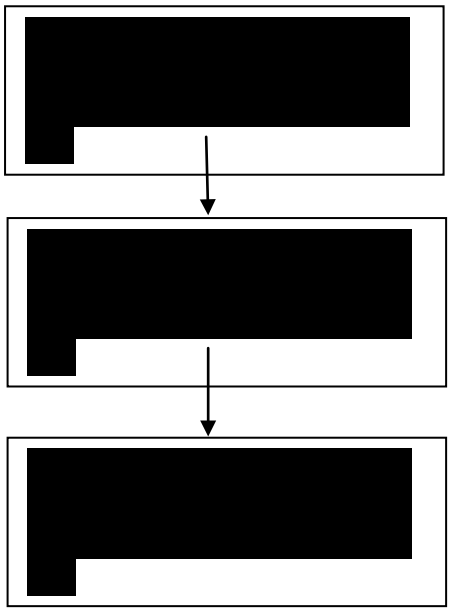
3.7 Please explain in detail the financial aspects of the scheme. How will any rebates be calculated and paid?

[redacted]

3.8 Please provide details of how the scheme will be administered. Please specify whether any additional information will need to be collected, explaining when this will be done and by whom.

[redacted]

3.9 Please provide a flow diagram that clearly shows how the scheme will operate. Any funding flows must be clearly demonstrated (CiC).



3.10 Please provide details of the duration of the scheme.

[redacted]

[redacted]

3.11 Are there any equity or equalities issues relating to the scheme, taking into account current legislation and, if applicable, any concerns identified during the course of the appraisal? If so, how have these been addressed?

SLE is more prevalent in women and African-Caribbean, South Asian and Chinese populations than in European white populations (Danchenko et al. 2006; Manzi 2009), and the demographic of SLE patients is likely to include a significant portion of women of child-bearing age (Danchenko et al. 2006).

The patient access scheme does not seek to specifically address any equity or equality issues; however the availability of belimumab, through a positive NICE recommendation and the corresponding patient access scheme, will allow these patients access to an innovative treatment in an area of significant unmet need.

3.12 If available, please list any scheme agreement forms, patient registration forms, pharmacy claim forms/rebate forms, guides for pharmacists and physicians and patient information documents. Please include copies in the appendices.

N/A. [redacted]

3.13 In the exceptional case that you are submitting an outcome-based scheme, as defined by the PPRS, please also refer to appendix B.

N/A

4 Cost effectiveness

4.1 If the population to whom the scheme applies (as described in sections 3.4 and 3.5) has not been presented in the main manufacturer/sponsor submission of evidence for the technology appraisal (for example, the population is different as there has been a change in clinical outcomes or a new continuation rule), please (re-)submit the relevant sections from the 'Specification for manufacturer/sponsor submission of evidence' (particularly sections 5.5, 6.7 and 6.9). You should complete those sections both with and without the patient access scheme. You must also complete the rest of this template.

The patient access scheme relates to a specific subgroup of the overall trial and licensed population. Being mindful of NHS resources, the proposed subgroup aims to identify those individuals who are likely to benefit the most from belimumab.

The subgroup to which the PAS relates has been presented as part of the main submission (Section 5 and Section 6.8). The results presented here relate to the updated ICERs based on the discounted drug acquisition cost.

4.2 If you are submitting the patient access scheme at the end of the technology appraisal process, you should update the economic model to reflect the assumptions that the Appraisal Committee considered to be most plausible. No other changes should be made to the model.

N/A – The proposed patient access scheme is being submitted for consideration during the technology appraisal process.

4.3 Please provide details of how the patient access scheme has been incorporated into the economic model. If applicable, please also provide details of any changes made to the model to reflect the assumptions that the Appraisal Committee considered most plausible.

The patient access scheme involves a straight discount, therefore the updated economic model and results reflect only the discounted drug acquisition cost and does not impact on the clinical outcomes for the subgroup under consideration.

4.4 Please provide the clinical effectiveness data resulting from the evidence synthesis and used in the economic model which includes the patient access scheme.

The clinical evidence for the subgroup to which the economic model relates has been presented as part of the main submission (Section 5.3.7 and Section 5.5).

4.5 Please list any costs associated with the implementation and operation of the patient access scheme (for example, additional pharmacy time for stock management or rebate calculations). A suggested format is presented in table 1. Please give the reference source of these costs. Please refer to section 6.5 of the 'Specification for manufacturer/sponsor submission of evidence'.

N/A

[redacted]

4.6 Please provide details of any additional treatment-related costs incurred by implementing the patient access scheme. A suggested format is presented in table 2. The costs should be provided for the intervention both with and without the patient access scheme. Please give the reference source of these costs.

N/A

[redacted]

Summary results

Base-case analysis

4.7 Please present in separate tables the cost-effectiveness results as follows.[1]

- the results for the intervention without the patient access scheme
- the results for the intervention with the patient access scheme.

A suggested format is shown below (table 3).

Table 3.1: Base-case results – Pooled Total Population

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
SoC	£97,583	16.74	9.55	-	-	-	
Belimumab	£133,167	17.33	9.98	£35,584	0.59	0.43	£82,909
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years							

Table 3.2: Base-case results – High Disease Activity Subgroup

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
SoC	£105,366	17.05	9.81	-			
Belimumab	£157,291	18.11	10.61	£51,925	1.05	0.806	£64,410
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years							

In the high disease activity subgroup, belimumab-treated patients are estimated to live longer, however, due to their increased life expectancy and due to belimumab treatment, costs are higher than for SoC patients. The incremental costs are £51,925, resulting in 1.05 added life years (discounted)

[1] For outcome-based schemes, please see section 5.2.8 in appendix B.

or 0.806 added QALYs. This results in an incremental cost effectiveness ratio (ICER) of £64,410 per QALY gained.

Table 3.3: High Disease Activity Subgroup – including PAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
SoC	£105,366	17.05	9.81	-			
Belimumab	£[redacted]	18.11	10.61	£[redacted]	1.05	0.806	£[redacted]
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years							

When the patient access scheme discount is considered for the high activity subgroup, the total costs for the belimumab-treated patients are estimated to be £[redacted] and the incremental costs are £[redacted] while the incremental LYG and incremental QALYs remain the same as the presented previously, at 1.05 and 0.806, added life years (discounted) and added QALYs, respectively. This results in an incremental cost effectiveness ratio (ICER) of £[redacted] per QALY gained.

4.8 Please present in separate tables the incremental results as follows. [2]

- the results for the intervention without the patient access scheme
- the results for the intervention with the patient access scheme.

List the interventions and comparator(s) from least to most expensive. Present the incremental cost-effectiveness ratios (ICERs) in comparison with baseline (usually standard care), and the incremental analysis ranking technologies in terms of dominance and extended dominance. A suggested format is presented in table 4.

Please see results above (Question 4.8).

[2] For outcome-based schemes, please see section 5.2.9 in appendix B.

Sensitivity analyses

4.9 Please present deterministic sensitivity analysis results as described for the main manufacturer/sponsor submission of evidence for the technology appraisal. Consider using tornado diagrams.

Tornado diagrams for the ICER, incremental QALYs and incremental costs are presented in Figures F1, F2, F3 respectively. Full details of sensitivity analysis can been found in Section 6.6 of the full submission.

Figure F1: Tornado diagram for univariate sensitivity analyses on the ICER – High Disease Activity Subgroup including PAS

Figure removed as commercial in confidence data.

Figure F2: Tornado diagram for univariate sensitivity analyses on the incremental QALYs (delta E) – High Disease Activity Subgroup including PAS

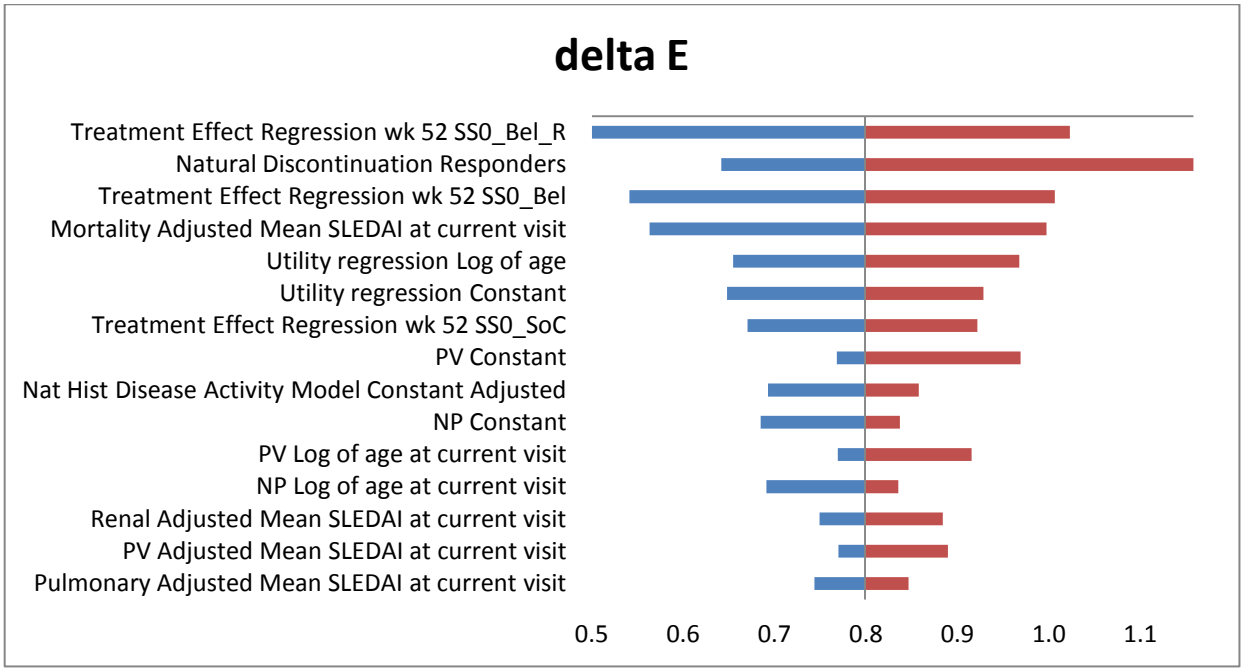


Figure F3: Tornado diagram for univariate sensitivity analyses on the incremental costs (delta C) – High Disease Activity Subgroup including PAS

Figure removed as commercial in confidence data.

4.10 Please present any probabilistic sensitivity analysis results, and include scatter plots and cost-effectiveness acceptability curves.

Figure F4: Scatter plot of the PSA – High Disease Activity Subgroup including PAS

Figure removed as commercial in confidence data.

Figure F5: PSA Acceptability Curve – High Disease Activity Subgroup including PAS

Figure removed as commercial in confidence data.

In the high disease activity subgroup, there is a [redacted]% probability that belimumab is cost-effective compared to standard of care at a willingness to pay (WTP) of

£ 30,000 per QALY gained. [redacted]

4.11 Please present scenario analysis results as described for the main manufacturer/sponsor submission of evidence for the technology appraisal.

Table T4 : Summary of Scenario Results - High Disease Activity Subgroup (without PAS)

Description of Scenario	Scenario Details	Incremental Cost Belimumab	Incremental LYs Belimumab	Incremental QALYs Belimumab	Incremental Cost per QALY
Base Case for High Disease Activity Subgroup	Time horizon = lifetime, lifetime max effect of belimumab; responder rule of SS reduction ≥4 at week 24; adjusted natural history model; no vial sharing	£51,925	1.05	0.806	£64,410
Responder rule excluded	As base case but with responder rule at 24 weeks excluded	£56,631	1.01	0.784	£72,207
Alternative Responder rule	As base case but with responder rule of SS reduction of ≥6 at week 24;	£30,760	0.81	0.614	£50,114
Increased vial price	As base case but with vial price increased (120mg=£127.80; 400mg=£426)	£57,478	1.05	0.806	£71,297
Original natural history model	As base case but with original natural history model chosen	£51,227	0.82	0.659	£77,707
With vial sharing	As base case	£49,717	1.05	0.806	£61,671
Higher drug administration cost	As base case but with a drug administration cost of £159 as recommended by ERG as a sensitivity analysis for the tocilizumab appraisal for rheumatoid arthritis	£54,298	1.05	0.806	£67,353

4.12 If any of the criteria on which the patient access scheme depends are clinical variable (for example, choice of response measure,

level of response, duration of treatment), sensitivity analyses around the individual criteria should be provided, so that the Appraisal Committee can determine which criteria are the most appropriate to use.

N/A

Impact of patient access scheme on ICERs

4.13 For financially based schemes, please present the results showing the impact of the patient access scheme on the ICERs for the base-case and any scenario analyses. A suggested format is shown below (see table 5). If you are submitting the patient access scheme at the end of the appraisal process, you must include the scenario with the assumptions that the Appraisal Committee considered to be most plausible.

Table T5 : Summary of Scenario Results - High Disease Activity Subgroup (including PAS)

Description of Scenario	Scenario Details	Incremental Cost Belimumab	Incremental LYs Belimumab	Incremental QALYs Belimumab	Incremental Cost per QALY
Base Case for High Disease Activity Subgroup	Time horizon = lifetime, lifetime max effect of belumimab; responder rule of SS reduction ≥4 at week 24; adjusted natural history model; no vial sharing	[redacted]	1.05	0.806	[redacted]
Responder rule excluded	As base case but with responder rule at 24 weeks excluded	[redacted]	1.01	0.784	[redacted]
Alternative Responder rule	As base case but with responder rule of SS reduction of ≥6 at week 24;	[redacted]	0.81	0.614	[redacted]
Original natural history model	As base case but with original natural history model chosen	[redacted]	0.95	0.699	[redacted]
With vial sharing	As base case	[redacted]	1.05	0.806	[redacted]
Higher drug administration cost	As base case but with a drug administration cost of £159 as recommended by ERG as a sensitivity analysis for the tocilizumab appraisal for rheumatoid arthritis	[redacted]	1.05	0.806	[redacted]

5 Appendices

5.1 Appendix A: Additional documents

5.1.1 If available, please include copies of patient access scheme agreement forms, patient registration forms, pharmacy claim forms/rebate forms, guides for pharmacists and physicians, patient information documents.

N/A

5.2 Appendix B: Details of outcome-based schemes

5.2.1 If you are submitting a proven value: price increase scheme, as defined in the PPRS, please provide the following information:

- the current price of the intervention
- the proposed higher price of the intervention, which will be supported by the collection of new evidence
- a suggested date for when NICE should consider the additional evidence.

N/A

5.2.2 If you are submitting an expected value: rebate scheme, as defined in the PPRS, please provide the following details:

- the current price of the intervention (the price that will be supported by the collection of new evidence)
- the planned lower price of the intervention in the event that the additional evidence does not support the current price
- a suggested date for when NICE should consider the additional evidence.

N/A

5.2.3 If you are submitting a risk-sharing scheme, as defined in the PPRS, please provide the following details:

- the current price of the intervention (the price that will be supported by the collection of new evidence)
- the proposed relationship between future price changes and the evidence to be collected.

N/A

5.2.4 For outcome-based schemes, as defined in the PPRS, please provide the full details of the new information (evidence) planned to be collected, who will collect it and who will carry the cost associated with this planned data collection. Details of the new information (evidence) may include:

- design of the new study
- patient population of the new study
- outcomes of the new study
- expected duration of data collection
- planned statistical analysis, definition of study groups and reporting (including uncertainty)
- expected results of the new study
- planned evidence synthesis/pooling of data (if applicable)
- expected results of the evidence synthesis/pooling of data (if applicable).

N/A

5.2.5 If you are submitting a risk-sharing scheme, please specify the period between the time points when the additional evidence will be considered.

N/A

5.2.6 Please provide the clinical effectiveness data resulting from the evidence synthesis and used in the economic modelling of the patient access scheme at the different time points when the additional evidence is to be considered.

N/A

5.2.7 Please provide the other data used in the economic modelling of the patient access scheme at the different time points when the additional evidence is to be considered. These data could include cost/resource use, health-related quality of life and utilities.

N/A

5.2.8 Please present the cost-effectiveness results as follows.

- For proven value: price increase schemes, please summarise in separate tables:
  - the results based on current evidence and current price
  - the anticipated results based on the expected new evidence and the proposed higher price.
- For expected value: rebate schemes, please summarise in separate tables:
  - the results based on the expected new evidence and the current price (which will be supported by the additional evidence collection)
  - the results based on the current evidence and the lower price (if the new evidence is not forthcoming).
- For risk-sharing schemes, please summarise in separate tables:
  - the results based on current evidence and current price
  - the results based on the expected new evidence and the current price (which will be supported by the additional evidence collection)
  - the results based on the current evidence and the lower price (if the new evidence is not forthcoming)
  - the anticipated results based on the expected new evidence and the proposed higher price.

A suggested format is shown in table 3, section 4.7.

5.2.9 Please present in separate tables the incremental results for the different scenarios as described above in section 5.2.8 for the type of outcome-based scheme being submitted.

List the interventions and comparator(s) from least to most expensive. Present the incremental cost-effectiveness ratios (ICERs) in comparison with baseline (usually standard care), and the incremental analysis ranking technologies in terms of dominance and extended dominance. A suggested format is presented in table 4, section 4.8.

N/A

18 May 2011

National Institute for Health and Clinical Excellence

NICE
Midcity Place
71 High Holborn
London
WC1V 6NA

Tel: 0161 870 3154
Fax: 020 7061 9792

Email: Kate.Moore@nice.org.uk

www.nice.org.uk

Dear [redacted],

Re: Single Technology Appraisal –
Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

The Evidence Review Group, Warwick Evidence and the technical team at NICE have now had an opportunity to take a look at submission received on the 13 April 2011 by GlaxoSmithKline. In general terms they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification relating to the clinical and cost effectiveness data.

Both the ERG and the technical team at NICE will be addressing these issues in their reports.

We request you to provide a written response to this letter to the Institute by **12:00, 2 June 2011**. Two versions of this written response should be submitted; one with academic/commercial in confidence information clearly marked and one from which this information is removed.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, and all information submitted under 'academic in confidence' in yellow.

If you present data that is not already referenced in the main body of your submission and that data is seen to be academic/commercial in confidence information, please complete the attached checklist for in confidence information.

Please do not 'embed' documents (i.e. PDFs, spreadsheets) within your response as this may result in your information being displaced or unreadable. Any supporting documents should be emailed to us separately as attachments, or sent on a CD.

If you have any further queries on the technical issues raised in this letter then please contact Helen Starkie – Technical Lead (Helen.Starkie@nice.org.uk) Any procedural questions should be addressed to Kate Moore – Project Manager (Kate.Moore@nice.org.uk) in the first instance.

Yours sincerely

Helen Knight
Associate Director – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for in confidence information

General considerations when addressing the clarification questions

Licence: the submission presents clinical effectiveness and economic analysis results for two main populations: one corresponding to the whole population in the BLISS 52 and BLISS 76 trials and a second corresponding to a high disease activity subgroup from these trials characterised by SLEDAI ≥10, raised anti dsDNA, and low complement (C3 C4).

The draft Summary of Product Characteristics (SPC) states:

"4.1 Therapeutic indications
Benlysta is indicated for reducing disease activity in adult patients with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g positive anti-dsDNA, low complement) despite standard therapy (see section 5.1)."

This implies high disease activity is indicated only by raised anti dsDNA and low complement only, and makes no reference to SLEDAI score.

In the clarification letter the following BLISS subgroup: anti-dsDNA +ve and low C3 or C4 [n=287+305] **is defined as marketing authorisation population**. The following BLISS subgroup: anti-dsDNA +ve and low C3 or C4 and SLEDAI ≥ 10 [n=203+193] **is defined as Target Population**.

Section A: Clarification on effectiveness data

Priority Questions

A 1 Please confirm that the interpretation of marketing authorisation population and the target Population is correct, and that the submission is based on a subgroup of the expected licensed population. If this is the case, NICE will be unable to make a recommendation for the expected licensed population. Please confirm that this is indeed your approach. If this is not your intention then please submit additional data as requested throughout on the 'marketing authorisation' population.

A 2 The cost effectiveness argument for belimumab against rituximab has been undertaken qualitatively, for reasons explained in

the submission. Please provide further discussion on the decision that the comparison could not be undertaken quantitatively. Please provide any approach/analyses undertaken to attempt to compare the treatments quantitatively.

A 3 Please supply demographic characteristics and baseline disease characteristics for marketing authorisation population and Target Population by trial and treatment. Please fill in tables with patient numbers etc.

Demographic characteristics and baseline disease characteristics for **marketing authorisation population**

		BLISS 52		**BLISS 76**		**Combined BLISS**	
		SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
	Total enrolled						
Gender	Male						
	Female						
Race	Caucasian						
	Asian						
	Black/African Am.						
	Alaskan/Native Am						
	Hawaiian/Pacific Isl						
	Multiracial						
	Hispanic origin						
Age	Years						
	n ≤ 45 yrs						
	n > 45 to < 65						
	n ≥ 65 to < 75						
Weight (kg)	Mean (SD)						
	Range						
Region& country	USA/Canada						
	W Europe/Israel						
	E Europe						
	America excluding USA/Canada						
	Latin America						
	Asia						
	Australia						

.

Demographic characteristics and baseline disease characteristics for **Target Population**

		BLISS 52		**BLISS 76**		**Combined BLISS**	
		SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
	Total enrolled						
Gender	Male						
	Female						
Race	Caucasian						
	Asian						
	Black/African Am.						
	Alaskan/Native Am						

	Hawaiian/Pacific Isl						
	Multiracial						
	Hispanic origin						
Age	Years						
	n ≤ 45 yrs						
	n > 45 to < 65						
	n ≥ 65 to < 75						
Weight (kg)	Mean (SD)						
	Range						
Region& country	USA/Canada						
	W Europe/Israel						
	E Europe						
	America excluding USA/Canada						
	Latin America						
	Asia						
	Australia						

Demographic characteristics and baseline disease characteristics for **marketing authorisation population**

		BLISS 52		**BLISS 76**		**Combined BLISS**	
		SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
	Total enrolled						
SLE duration yrs; mean (SD)							
BILAG organ involvement	At least 1A or 2B						
	At least 1A						
	At least 1A or1B						
	No A or B						
SELENA SLEDAI mean (SD)							
PGA category	0 to 1						
	>1 to 2.5						
	2.5 to 3						
SLICC Damage index; mean (SD)							
Prednisone or equivalent dose	0 mg/day						
	>0 - ≤7.5 mg/day						
	> 7.5 mg/day						
Average prednisone or equivalent dose; mean (SD) mg/day							

Demographic characteristics and baseline disease characteristics for **Target Population**

		BLISS 52		BLISS 76		Combined BLISS	
		SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
	Total enrolled						
SLE duration yrs; mean (SD)							
BILAG organ involvement	At least 1A or 2B						
	At least 1A						
	At least 1A or1B						
	No A or B						
SELENA SLEDAI mean (SD)							
PGA category	0 to 1						
	>1 to 2.5						
	2.5 to 3						
SLICC Damage index; mean (SD)							
Prednisone or equivalent dose	0 mg/day						
	>0 - ≤7.5 mg/day						
	> 7.5 mg/day						
Average prednisone or equivalent dose; mean (SD) mg/day							

A 4 What was the distribution of SLE manifestations at baseline (as in SLEDAI) in the marketing authorisation population AND Target Population?

Please fill in table with patient numbers **marketing authorisation population**

	BLISS 52		BLISS 76		Combined BLISS	
	SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
Total						
Neuropsychiatric						
Vascular						
Musculo skeletal						
Renal						
Skin						
Serositis						
Other						
Haemo						

Please fill in table with patient numbers **Target Population**

	BLISS 52		BLISS 76		Combined BLISS	
	SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
Total						
Neuropsychiatric						
Vascular						
Musculo skeletal						
Renal						
Skin						
Serositis						
Other						
Haemo						

A 5 What was the distribution of SLE improvements by organ system in the marketing authorisation population and Target Population at 52 weeks? Please fill in table with patient numbers.

Distribution of SLE improvements by organ system in the **marketing authorisation population**

		BLISS 52		**BLISS 76**		**Combined BLISS**	
		SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
	Total enrolled						
Mucocutaneous	Involved baseline						
	Improved week 52						
Musculoskeletal	Involved baseline						
	Improved week 52						
Renal	Involved baseline						
	Improved week 52						
Vascular	Involved baseline						
	Improved week 52						
Serositis	Involved baseline						
	Improved week 52						
Haemo	Involved baseline						
	Improved week 52						
Neuropsychiatric	Involved baseline						
	Improved week 52						
Other (specify)	Involved baseline						
	Improved week 52						
Other (specify)	Involved baseline						
	Improved week 52						
Other (specify)	Involved baseline						
	Improved week 52						

Distribution of SLE improvements by organ system in the **Target Population**

		BLISS 52		**BLISS 76**		**Combined BLISS**	
		SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
	Total enrolled						
Mucocutaneous	Involved baseline						
	Improved week 52						
Musculoskeletal	Involved baseline						
	Improved week 52						
Renal	Involved baseline						
	Improved week 52						
Vascular	Involved baseline						
	Improved week 52						
Serositis	Involved baseline						
	Improved week 52						
Haemo	Involved baseline						
	Improved week 52						
Neuropsychiatric	Involved baseline						
	Improved week 52						
Other (specify)	Involved baseline						
	Improved week 52						
Other (specify)	Involved baseline						
	Improved week 52						
Other (specify)	Involved baseline						
	Improved week 52						

A 6 What are the results of the efficacy end points for marketing authorisation population and Target Population? Please complete the tables below.

Efficacy end points for **marketing authorisation population**

		BLISS 52		**BLISS 76**		**Combined BLISS**	
		SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
	Total enrolled						
SRI at wk 52	n(%) responders						
	OR (95%CI) **Δ**						
SRI at wk 76	n(%) responders						
	OR (95%CI) **Δ**						
Modified SRI wk 52	Number at risk						
	n(%) responders						
	OR (95%CI) **Δ**						
SLEDAI (≥4 reduction)	n(%) responders						
	OR (95%CI) **Δ**						
SLEDAI change in mean score by wk 52							
No new BILAG 1A/2B	n(%) responders						
No worsening in PGA	n(%) responders						
Steroid use	Number at risk**¥**						
	n(%) responders						

Efficacy end points for **Target Population**

		BLISS 52		**BLISS 76**		**Combined BLISS**	
		SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
	Total enrolled						
SRI at wk 52	n(%) responders						
	OR (95%CI) **Δ**						
SRI at wk 76	n(%) responders						
	OR (95%CI) **Δ**						
Modified SRI wk 52	Number at risk						
	n(%) responders						
	OR (95%CI) **Δ**						
SLEDAI (≥4 reduction)	n(%) responders						
	OR (95%CI) **Δ**						
SLEDAI change in mean score by wk 52							
No new BILAG 1A/2B	n(%) responders						
No worsening in PGA	n(%) responders						

Steroid use	Number at risk¥						
	n(%) responders						

A 7 Please complete the tables below for mean of the change in SLEDAI score:

Marketing authorisation population - Mean of the change in SLEDAI score

SLEDAI (≥4 reduction) at week 24 (wk 24 responders)	**BLISS 52**		**BLISS 76**		**Combined BLISS**	
	SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
Among wk 24 responders:						
Mean of the change in SLEDAI score by wk 24						
Mean of the change in SLEDAI score by wk 52						
Among wk 24 non-responders:						
Mean of the change in SLEDAI score by wk 24						
Mean of the change in SLEDAI score by wk 52						
Overall:						
Mean of the change in SLEDAI score by wk 24						
Mean of the change in SLEDAI score by wk 52						

Target Population - Mean of the change in SLEDAI score

SLEDAI (≥4 reduction) at week 24 (wk 24 responders)	**BLISS 52**		**BLISS 76**		**Combined BLISS**	
	SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
Among wk 24 responders:						
Mean of the change in SLEDAI score by wk 24						
Mean of the change in SLEDAI score by wk 52						
Among wk 24 non-responders:						
Mean of the change in SLEDAI score by wk 24						
Mean of the change in SLEDAI score by wk 52						
Overall:						
Mean of the change in SLEDAI score by wk 24						
Mean of the change in SLEDAI score by wk 52						

A 8 Please complete the tables below on steroid use:

Marketing authorisation population - Defining wk 24 responders as SLEDAI (≥4 reduction) at week 24

	BLISS 52		BLISS 76		Combined BLISS	
	SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
Among wk 24 responders:	mg/day wk 24	mg/day wk 24	mg/day wk 24	mg/day wk 24	mg/day wk 24	mg/day wk 24
Average steroid dose at baseline						
Average steroid dose at wk 24						
Average steroid dose at wk 52						
Among wk 24 non-responders:						
Average steroid dose at baseline						
Average steroid dose at wk 24						
Average steroid dose at wk 52						
Overall:						
Average steroid dose at baseline						
Average steroid dose at wk 24						
Average steroid dose at wk 52						

Target Population - Defining wk 24 responders as SLEDAI (≥4 reduction) at week 24

	BLISS 52		BLISS 76		Combined BLISS	
	SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
Among wk 24 responders:	mg/day wk 24	mg/day wk 24	mg/day wk 24	mg/day wk 24	mg/day wk 24	mg/day wk 24
Average steroid dose at baseline						
Average steroid dose at wk 24						
Average steroid dose at wk 52						
Among wk 24 non-responders:						
Average steroid dose at baseline						
Average steroid dose at wk 24						
Average steroid dose at wk 52						
Overall:						
Average steroid dose at baseline						
Average steroid dose at wk 24						
Average steroid dose at wk 52						

A 9 The clinical effectiveness section notes that region was pre-specified as a subgroup in the data analysis protocol. It further notes that the major subgroups as predictor of response rate pooled between arms in order of significance were: baseline SS, complement, immunosuppressant use, region, SDI and anti-dsDNA. Please present the results of this analysis.

Related to the above question, did the major subgroup analysis include an analysis of them as predictors of the odds ratio of response for belimumab 10mg/kg versus SoC? If this, or a similar analysis of relative efficacy, has been undertaken please present the results of this analysis.

A 10 Please provide marketing authorisation population and Target Population major outcome results (SRI, SLEDAI [4 point reduction], and reduction in steroid use weeks 40 to 52) for the pooled 10mg Belimumab and SoC groups by geographical region (e.g. W. Europe, North America (USA/Canada), America (excluding USA Canada), E Europe, Asia, Australia).

A 11 Please clarify the reasons for non-responder status (SRI) by week 52 in the marketing authorisation population and Target Population. Please complete the following tables.

Reasons for non-responder status at week 52 - marketing authorisation population

		BLISS 52		**BLISS 76**		**Combined BLISS**	
		SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
Non-responders at wk 52 n (%)							
Lack of SRI							
Non-compliance	*Adverse events*						
	Lack of efficacy						
	Other						
	Loss to follow up						
Treatment violation (by drug class / type)	*Steroid*						
	Anti-malarial						
	Immunosuppressant						
	Lipid lowering						
	Hypertension						
	Other (specify)						
	Other (specify)						

Reasons for non-responder status at week 52 - Target Population

		BLISS 52		**BLISS 76**		**Combined BLISS**	
		SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
Non-responders at wk 52 n (%)							
Lack of SRI							
Non-compliance	*Adverse events*						
	Lack of efficacy						
	Other						
	Loss to follow up						
Treatment violation (by drug class / type)	*Steroid*						
	Anti-malarial						
	Immunosuppressant						
	Lipid lowering						
	Hypertension						
	Other (specify)						

	Other (specify)						

A 12 Please clarify how the standard of care (SoC) in the BLISS trials relates to that current in the UK for the appropriate high disease activity population. Please clarify differences between trial centres regarding SoC.

Non-priority Questions

A 13 Fig 5.9 shows a decline in SRI in the placebo group from wk24 to wk52. Is this attributable to patients receiving non-permitted steroid use during this period?"?

A 14 For table 5.15 on page 106 of the submission, please clarify why the trial interaction is not applicable (N/A) for the Target Population pooled result while it is for the pooled result for the whole population (this appears at odds with the figure footnote)

A 15 Page 119 of the submission states: "However, for the pooled total population, both belimumab doses achieved significance compared with placebo for reduction of prednisone by ≥ 25% from baseline to ≤ 7.5 mg/day during." Please clarify whether this refers to results for the 1mg/kg dose regimen in the phase 3 trials. If so please provide results.

A 16 The systematic review searches yielded 36 publications. These are not listed in the submitted documents. Please send a list of publications with reasons for exclusion.

A 17 Page 97 of the submission states: Pooling is appropriate given that the trials were essentially identical in design and in the analysis of the primary endpoint, the p-values for the treatment-by-study interaction were not significant (interaction p-values > 0.5).

However we note systematic differences between the trials with regard to demographic characteristics (Table 5.8 age and race), SLICC organ damage and proteinuria (Table 5.9), anti-dsDNA positivity, raised IgG, and abnormal complement (Table 5.10) so that pooling might be

considered inappropriate: Please provide further justification for pooling results for Target Population.

A 18 In Figure 5.6 Pg. 106 of the submission and figure 5.9, Pg. 109. Please clarify whether the vertical dashed lines represent clinic visit times.

A 19 Page 95 of the submission refers to “Fig X”. Please clarify which figure this text is referencing.

Section B: Clarification on cost-effectiveness data

Priority Questions

B 1 Please provide a table listing all the model assumptions (with justifications).

B 2 Setting the baseline distribution of ages to be a single value within the Subgroup BLISS Data worksheet by setting Q7:Q62=0 then setting one value within this array equal to Q64, appears to suggest that as the baseline age increases the ICER falls: e.g. for the Target Population age 30 the ICER is £65,498 per QALY while keeping everything else constant for the Target Population age 50 the ICER is £55,439 per QALY. Is this correct? If so this is the opposite of that what is often observed within economic modelling. Please provide a clarification and explanation of this result.

B 3 Please clarify what has been assumed for belimumab non-responders in terms of the evolution of their SS score and steroid dose. This might be most easily explained through a comparison of a hypothetical single High Disease Activity patient as a:

- belimumab non-responder, and
- SoC patient

and for each of these presenting:

- the common baseline characteristics,

- the calculation of the week 24 SS score and week 24 steroid dose, and
- the calculation of the week 52 SS score and week 52 steroid dose.

B 4 Please present the patient numbers (for both marketing authorisation population and Target Population) and the numbers continuing with treatment underlying *Table 6.42* (page 289 of the submission) at a greater degree of disaggregation, where responder status is defined by a change in the SS score at week 24 ≥ 4, N is the number of patients within the relevant category, and N Cont. is the number of patients continuing with belimumab treatment at week 24 within the relevant category.

Marketing authorisation population

	BLISS 52				**BLISS 76**			
	Wk24 Responder		Wk24 NonResponder		Wk24 Responder		Wk24 NonResponder	
SS t=0	N	N Cont.	N	N Cont.	N	N Cont.	N	N Cont.
4								
5								
etc								

Target Population

	BLISS 52				**BLISS 76**			
	Wk24 Responder		Wk24 NonResponder		Wk24 Responder		Wk24 NonResponder	
SS t=0	N	N Cont.	N	N Cont.	N	N Cont.	N	N Cont.
4								
5								
etc								

B 5 Please confirm labelling of axes for figures 6.9 and 6.11 and present the equivalent of Figures 6.18 and 6.19 for the Target Population.

Within this please also separately present a split of the belimumab arm into those responding at week 24 and those not responding at week 24. It is recognised that this latter might require another two runs of the model. Please can the data underlying these figures be retained and submitted within an excel workbook.

Non-priority Questions

In the light of the manufacturer strongly linking belimumab to the Target Population [n=203+193], the following clarification questions relate only to the Target Population of BLISS 52, BLISS 76 and pooled between these two trials. Depending on the response to question A1, please provide the data requested for the marketing authorisation population [n=287+305] patient populations of BLISS 52, BLISS 76 and pooled between these two trials.

B 6 Please clarify the extent to which the probabilities of treatment continuation at week 24 by SS score in *Table 6.42* differ from the probabilities of response at week 24. Are the values in *6.42* the addition of (A) non responders as defined by a change in SS score of less than 4; **PLUS** (B) responders who have discontinued by week 24?

Related to the above question, what role if any do the discontinuation rates for belimumab weeks 24 non-responders in year 1 of 37.4% and in subsequent years of 37.4% have upon the model structure?

B 7 Much of the modelling uses 2006 or 2007 unit costs and uprates these with the CPI. Please clarify the reasoning for the use of the dated unit costs when more recent reference costs and PSSRU costs of health care are available online.

B 8 Tornado diagrams: could the variables be explained / labelled more explicitly; for example the meaning of, and the distinction between "Treatment Effect Regression wk 52 SSO_Bel_R" and "Treatment Effect Regression wk 52 SSO_Bel" may not be immediately obvious to committee members. Please tabulate the values underlying the tornado diagrams of *Figures 6.37, 6.38 and 6.39* along the following lines:

Variable	Base value	Low value	ICER	High value	ICER

B 9 The user guide provided is relatively brief and provides limited background as to the model programming and structure within excel. If the manufacturer has been provided with a more detailed written account of the electronic model structure and VBA programming within the model, please could this be supplied.

B 10 In terms of the implementation of the probabilistic modelling this seems to work down the 1,000 sets of clinical effectiveness estimates in the *PSA Inputs* worksheet by setting the active set of clinical effectiveness estimates in row 9 equal to one row of estimates below through the use of the INDEX function. Presumably for each set of clinical effectiveness estimates the 50,000 patient simulations are run to yield ***the*** central estimate for costs and for QALYs for that set of clinical effectiveness estimates to give ***one*** point on the cost effectiveness plane. 1,000 sets of clinical effectiveness estimates yield 1,000 points on the cost effectiveness plane, from which the CEAC is generated. Is this a correct interpretation of the generation of the CEAC within the modelling?

B 11 The *Treatment Effect* worksheet cells *Q12:R19* outline a net change at week 52 in SS items differentiated by SoC and Belimumab responders, as drawn from the *Subgroup BLISS data* worksheet which in turn references the *PSA Inputs* worksheet. Are these used as parameter inputs to the model? If they are, how are they derived?

B 12 In the electronic copy of the model worksheet *Subgroup BLISS data* cells *BN7:BP7* relate to the week 24 evolution of SS scores. Are these used within the modelling? If they are, please re-estimate this function separately for BLISS 52 and BLISS 76 Target Population patients.

B 13 Within the *Subgroup BLISS data* worksheet there is a number of logically separate arrays of data. Which if any, of the following are superfluous to the current model implementation, assuming that only the responder rule of SS change ≥ 4 is of interest? (Superfluous in this

context is not to say that the data is not used elsewhere to estimate functional forms for the model, only that the running of the model does not directly draw on these data elements or the source of these data elements within the model if this is from referencing another worksheet within the model as applies to e.g. O217:Q219)

Group	From	To	Purpose ("none" if superfluous)
1	O7	Q64	
2	O66	Q71	
3	O73	Q83	
4	O85	Q183	
5	O185	Q215	
6	O217	Q219	
7	O221	Q223	
8	O225	Q231	
9	O233	Q235	
10	O237	Q243	
11	O245	T256	
12	O258	Q263	
13	O266	Q274	
14	O276	Q276	
15	O277	Q296	
16	AQ9	AS39	
17	AQ233	AS245	
18	AQ259	AS263	
19	AQ265	AS274	
20	AQ276	AS276	
21	AQ277	AS296	

B 14 In the electronic copy of the model worksheet *PSA Data* of cells *BN7:CR7* what data within this is used within the modelling and what is superfluous to the current model implementation?

B 15 In the electronic copy of the model worksheet PSA Data of cells FV7:GR7 what data within this is used within the modelling and what is superfluous to the current model implementation?

B 16 Within the electronic model what do V1 and V2 refer to?

B 17 Please clarify the observed distributions of SLEDAI for the Johns Hopkins cohort and for the Target Population. For example, please supply a diagram like that in Fig 6.6 for the Target Population and for the Johns Hopkins cohort at entry, and for Johns Hopkins cohort at last follow up.

B 18 Please also provide the SLEDAI score distribution of excluded John Hopkins patients.

B 19 The model (Appendix 21, Pg. 21) returns huge mortality risk up to 250% for haematological involvement and infection. Please clarify whether this reflects the instability of the model (small sample size in this group) rather than the real effect size.

B 20 Appendix 21, 4.4. Time to event Analysis, the discussion section (Pg. 24 paragraph 3): The section states, “The clinical interpretation of a relationship between cardiovascular and cerebrovascular damage and mortality is stronger than that for musculoskeletal damage.” Please clarify whether this claim is made following a causal analysis?

B 21 Table 6.7 (page 198 of the submission) indicates a high mean steroid use in the John Hopkins population despite relatively low mean SLEDAI score; might this indicate different use of steroid in SoC in the 1990s relative to the present decade? Please clarify.

B 22 Please clarify the rationale for the choice of sensitivity analyses undertaken (e.g. as illustrated on page 269 of the submission).

B 23 On page 236 of the submission the calculated example provides a utility of 0.9719 for ocular organ damage for year one. Does this mean that ocular damage experienced in year 1 by patient A (in table 6.19) incurs a disutility of 1-0.9719 for that patient resulting in a utility for patient A of 0.63 – 0.0281 = 0.6019 ?

B 24 Fig 6.17 on page 259 of the submission represents the proportion of the “total population” remaining on belimumab through time and similarly Fig 6.35 (Page 291) for the Target Population. Please compare these on a single graph and clarify the reason(s) for the difference. Please comment on the fact that patients getting more benefit appear more likely to discontinue.

B 25 Appendix 23: The Figures are stated to be K-M plots, but they look like parametric fits. Please clarify.

B 26 Footnote to Table 6.37 Pg. 285 refers incorrectly to Table 4 (which is on Pg 28). Please clarify.

Section C: Clarification on other issues

Priority Questions

C 1 Figure 6.7 on page 201 of the submission shows the adjustment to the Johns Hopkins model (dashed red line) which involves raising the constant from 2.058 to 3.0; it is stated that "*A range of numbers were analysed to derive the adjusted constant, with a value of 3.0 providing a reasonable fit to these data".* How was the fit tested and does this refer to a fit of all the data shown in *Fig 6.10* or to weeks 52 to 250 or other? Please outline what other values were considered.

Related to the above, was the assumed value of 3.0 for the intercept of the regression analysis of the long term change in SS score retained for the Target Population?

C 2 The regression analysis for the average steroid dose related to the SS/AMS score as estimated from the John Hopkins cohort is accepted uncritically within the modelling. This is despite the arguments around the unrepresentativeness of the John Hopkins cohort for the BLISS trials in estimating the regression analysis of the long term change in SS score. To what extent were similar considerations around steroid use explored; e.g. validation through varying the constant and aligning with BLISS baseline steroid use and SS scores?

C 3 For the patient access scheme, please supply a correct version of Figure F5 (Pg. 18, Appendix 29).

Non-Priority Questions

C 4 Pg 157 states: *"In the long-term open-label extension of the Phase 2 trial (LBSL99), the incidence of AEs, severe AEs, SAEs, including infections, remained stable or declined over time through 5 years of exposure".* No data appears to have been presented to support this statement. Please clarify the source of the data to support this statement, providing further results if required.

C 5 What other forms for the regression analysis for the average steroid dose related to the SS/AMS score as estimated from the John Hopkins cohort were explored: e.g. change in steroid use being dependent upon change in SS/AMS score? What were the results of these analyses?

C 6 Please present estimates of the linear regression of the Target Population of table 6.41 (page 288 of the submission) separately for BLISS 52 and BLISS 76.

C 7 Random effects model(s) (Appendix 21). Please justify the use of previous mean SLEDAI score; please clarify whether this might be affected by regression to mean?

C 8 Please clarify why the usual Inter-class correlation (ICC) coefficient was not used to assess the validity of random effects model.

General considerations when addressing the clarification questions

Licence: the submission presents clinical effectiveness and economic analysis results for two main populations: one corresponding to the whole population in the BLISS 52 and BLISS 76 trials and a second corresponding to a high disease activity subgroup from these trials characterised by SLEDAI ≥10, raised anti dsDNA, and low complement (C3 C4).

The draft Summary of Product Characteristics (SPC) states:

"4.1 Therapeutic indications
Benlysta is indicated for reducing disease activity in adult patients with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g positive anti-dsDNA, low complement) despite standard therapy (see section 5.1)."

This implies high disease activity is indicated only by raised anti dsDNA and low complement only, and makes no reference to SLEDAI score.

In the clarification letter the following BLISS subgroup: anti-dsDNA +ve and low C3 or C4 [n=287+305] **is defined as marketing authorisation population**. The following BLISS subgroup: anti-dsDNA +ve and low C3 or C4 and SLEDAI ≥ 10 [n=203+193] **is defined as Target Population**.

Section A: Clarification on effectiveness data

Priority Questions

A 1 Please confirm that the interpretation of marketing authorisation population and the target Population is correct, and that the submission is based on a subgroup of the expected licensed population. If this is the case, NICE will be unable to make a recommendation for the expected licensed population. Please confirm that this is indeed your approach. If this is not your intention then please submit additional data as requested throughout on the 'marketing authorisation' population.

The interpretation of the target population is correct. Whilst the marketing authorisation indication specifies patients with systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g. positive anti-dsDNA, low complement), it should be noted that high disease activity can be further defined clinically based on other factors besides positive anti-dsDNA and low complement, for example, SELENA-SLEDAI.

Our submission is based on a high disease activity subgroup of the marketing authorisation population, defined as the Target Population. We acknowledge that NICE will be unable to make a recommendation for the whole of the expected licensed population (marketing authorisation population), but are aware that our Target Population falls within the expected licensed population. Mindful of the need to make the most efficient use of NHS resources, this subgroup allows a

targeted approach to selecting the patients who are most likely to get the greatest benefit from treatment. We therefore will not be submitting any data on the whole of the expected licensed population in our response to this clarification letter. However key results for this marketing authorisation population are available in the documents provided as part of our European marketing authorisation application and can be made available if required by the ERG.

A 2 The cost effectiveness argument for belimumab against rituximab has been undertaken qualitatively, for reasons explained in the submission. Please provide further discussion on the decision that the comparison could not be undertaken quantitatively. Please provide any approach/analyses undertaken to attempt to compare the treatments quantitatively.

As detailed in the submission document, there are no studies directly comparing the efficacy of belimumab with rituximab. There is one published Phase 2/3 RCT of rituximab in moderately-to-severely active SLE (Merrill et al. 2010). The primary endpoint in this study was based on BILAG scores, SELENA-SLEDAI was not reported in the trial publication; SELENA-SLEDAI was the main component of the primary endpoint in the BLISS trials.

It would be possible to analyse the BILAG data collected in the BLISS trials in the same way as that presented in the Merrill *et al* publication and then compare the efficacy results between trials using indirect comparison methodology. However we have not conducted a comparison in this way as we do not believe it is appropriate. The main reason for this decision relates to important differences in patient selection and consequently the treatment management protocols employed in the studies, described in more detail below. In addition, it is the SELENA-SLEDAI score and not the BILAG score, as a marker of disease activity, which has robust published evidence to show a relationship with long-term morbidity (organ damage and mortality) from SLE and hence is at the core of our economic model. As a result, even if an indirect comparison were conducted based on BILAG data, it would not be possible to explore the implications of this on cost effectiveness using our economic model due to the importance of examining effects on long-term outcomes of this disease.

The patients in the rituximab Phase 2/3 trial had significant and acute disease activity at entry to the study; 53% had at least one BILAG A score (severe disease activity) and a further 28% had at least 3 BILAG B scores (please note that although a BILAG B score represents moderate disease activity, the presence of 3 BILAG B scores in some organs indicates more severe disease activity). Initially, patients were receiving very high daily doses of prednisone (mean 45.9 mg ±16.4 mg) to treat the significant level of disease activity and this dose was to be tapered where possible during the trial. In addition, all patients were receiving one immunosuppressant at study entry. In contrast, the patients in the BLISS studies were a broader population and not all patients were experiencing major disease flares at study entry requiring the very high doses of steroids seen in the rituximab trial. Even in the high disease activity subgroup (Target Population), only 19.3% had at least one BILAG A score at baseline, the average prednisone or equivalent dose was 12.3 mg ± 9.6mg and 53% were on an immunosuppressant. In particular, we believe that the differences in the use of steroids and immunosuppressants to manage disease activity between the rituximab and BLISS trials and consequently the differences in the type of response observed

in the placebo arms render the studies incomparable. One of the main reasons proposed for the failure of rituximab to demonstrate a clear clinical benefit over the placebo arm in the study was the very high doses of prednisone used which was believed to have had a strong effect on disease activity in the placebo arm.

In both BLISS trials, belimumab demonstrated a significant reduction in disease activity compared with placebo (standard of care alone) as measured by the SRI composite primary endpoint at Week 52.

Thus, for the reasons described above, a qualitative rather than a quantitative approach has been taken. It is worth noting that in the absence of a formal comparison we have assumed that the efficacy of rituximab is similar to that of belimumab, and this may be a conservative assumption in terms of relative effectiveness of belimumab compared with rituximab given that the clinical trial for rituximab did not achieve a statistically significant benefit compared to placebo.

A 3 Please supply demographic characteristics and baseline disease characteristics for marketing authorisation population and Target Population by trial and treatment. Please fill in tables with patient numbers etc.

Please note with reference to our response to question A1 data for all following questions will be provided for the Target Population only.

Table A3.1 Demographic characteristics and baseline disease characteristics - High disease activity subgroup (Target Population)

		BLISS-52		BLISS-76		Combined BLISS	
		SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
	Total enrolled	**n=107**	**n=112**	**n=96**	**n=81**	**n=203**	**n=193**
Gender	Male	10 (9.3%)	3 (2.7%)	6 (6.3%)	4 (4.9%)	16 (7.9%)	7 (3.6%)
	Female	97 (90.7%)	109 (97.3%)	90 (93.8%)	77 (95.1%)	187 (92.1%)	186 (96.4%)
Race	Caucasian	29 (27.1%)	23 (20.5%)	61 (63.5%)	54 (66.7%)	90 (44.3%)	77 (39.9%)
	Asian	40 (37.4%)	53 (47.3%)	5 (5.2%)	4 (4.9%)	45 (22.2%)	57 (29.5%)
	Black/African American	1 (0.9%)	6 (5.4%)	13 (13.5%)	7 (8.6%)	14 (6.9%)	13 (6.7%)
	Alaskan/Native American	37 (34.6%)	30 (26.8%)	17 (17.7%)	16 (19.8%)	54 (26.6%)	46 (23.8%)
	Hawaiian/Pacific Islander	0	0	0	0	0	0
	Multiracial	1 (0.9%)	0	0	0	1 (0.5%)	0
	Hispanic origin	55 (51.4%)	46 (41.1%)	28 (29.2%)	25 (30.9%)	83 (40.9%)	71 (36.8%)
Age	Years						
	n ≤ 45 yrs	93 (86.9%)	100 (89.3%)	78 (81.3%)	65 (80.2%)	171 (84.2%)	165 (85.5%)
	n > 45 to < 65	14 (13.1%)	12 (10.7%)	17 (17.7%)	16 (19.8%)	31 (15.3%)	28 (14.5%)
	n ≥ 65 to < 75	0	0	1 (1.0%)	0	1 (0.5%)	0

		BLISS-52		BLISS-76		Combined BLISS	
		SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
	Total enrolled	**n=107**	**n=112**	**n=96**	**n=81**	**n=203**	**n=193**
Weight (kg)	Mean (SD)	62.1 ± 13.9	61.4 ± 14.1	68.8 ± 13.7	70.0 ± 16.7	65.2 ± 14.2	65.0 ± 15.8
	Range	34.7-127.6	39.5-128.5	45.4-108.6	47.0-131.7	34.7-127.6	39.5-131.7
Region & country	USA/Canada	0	0	45 (46.9%)	24 (29.6%)	45 (22.2%)	24 (12.4%)
	W Europe/Israel	0	0	24 (25.0%)	30 (37.0%)	24 (11.8%)	30 (15.5%)
	E Europe	10 (9.3%)	11 (9.8%)	12 (12.5%)	12 (14.8%)	22 (10.8%)	23 (11.9%)
	America excluding USA/Canada	56 (52.3%)	48 (42.9%)	15 (15.6%)	15 (18.5%)	71 (35.0%)	63 (32.6%)
	Asia	39 (36.4%)	53 (47.3%)	0	0	39 (19.2%)	53 (27.5%)
	Australia	2 (1.9%)	0	0	0	2 (1.0%)	0

Table A3.1 Demographic characteristics and baseline disease characteristics (continued) – High disease activity subgroup (Target Population)

		BLISS-52		BLISS-76		Combined BLISS	
		SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
	Total enrolled	**n=107**	**n=112**	**n=96**	**n=81**	**n=203**	**n=193**
SLE duration yrs; mean (SD)		6.70 ± 6.96	5.26 ± 4.99	7.42 ± 6.40	7.94 ± 7.47	7.04 ± 6.69	6.38 ± 6.28
BILAG organ involvement	At least 1A or 2B	65 (60.7%)	78 (69.6%)	78 (81.3%)	58 (71.6%)	143 (70.4%)	136 (70.5%)
	At least 1A	18 (16.8%)	25 (22.3%)	21 (21.9%)	7 (8.6%)	39 (19.2%)	32 (16.6%)
	At least 1A or 1B	99 (92.5%)	103 (92.0%)	94 (97.9%)	78 (96.3%)	193 (95.1%)	181 (93.8%)
	No A or B	8 (7.5%)	9 (8.0%)	2 (2.1%)	3 (3.7%)	10 (4.9%)	12 (6.2%)
SELENA-SLEDAI mean (SD)		12.6 ± 3.0	12.8 ± 3.6	13.0 ± 3.5	12.4 ± 2.9	12.8 ± 3.3	12.6 ± 3.3
PGA category	0 to 1	15 (14.0%)	13 (11.6%)	8 (8.3%)	8 (9.9%)	23 (11.3%)	21 (10.9%)
	>1 to 2.5	91 (85.0%)	97 (86.6%)	86 (89.6%)	71 (87.7%)	177 (87.2%)	168 (87.0%)
	>2.5 to 3	1 (0.9%)	2 (1.8%)	2 (2.1%)	2 (2.5%)	3 (1.5%)	4 (2.1%)
SLICC Damage index; mean (SD)		0.6 ± 1.0	0.5 ± 0.9	0.8 ± 1.4	0.8 ± 1.2	0.7 ± 1.2	0.6 ± 1.0
Prednisone or equivalent dose	0 mg/day	5 (4.7%)	4 (3.6%)	15 (15.6%)	12 (14.8%)	20 (9.9%)	16 (8.3%)
	>0 - ≤7.5 mg/day	26 (24.3%)	27 (24.1%)	31 (32.3%)	24 (29.6%)	57 (28.1%)	51 (26.4%)
	> 7.5 mg/day	76 (71.0%)	81 (72.3%)	50 (52.1%)	45 (55.6%)	126 (62.1%)	126 (65.3%)
Average prednisone or equivalent dose; mean (SD) mg/day		12.8 ± 8.4	13.7 ± 10.4	10.3 ± 8.8	10.4 ± 8.1	11.6 ± 8.6	12.3 ± 9.6

A 4 What was the distribution of SLE manifestations at baseline (as in SLEDAI) in the marketing authorisation population AND Target Population?

Table A4.1 Selected Baseline SELENA-SLEDAI Scores – High disease activity subgroup (Target Population)

	BLISS-52		BLISS-76		Combined BLISS	
	SoC	**10 mg/kg**	**SoC**	**10 mg/kg**	**SoC**	**10 mg/kg**
Total enrolled	**n=107**	**n=112**	**n=96**	**n=81**	**n=203**	**n=193**
Organic Brain Syndrome (8)	0	0	1 (1.0%)	0	1 (0.5%)	0
Lupus Headache (8)	1 (0.9%)	3 (2.7%)	0	2 (2.5%)	1 (0.5%)	5 (2.6%)
Vasculitis (8)	15 (14.0%)	19 (17.0%)	10 (10.4%)	5 (6.2%)	25 (12.3%)	24 (12.4%)
Arthritis (4)	65 (60.7%)	76 (67.9%)	83 (86.5%)	63 (77.8%)	148 (72.9%)	139 (72.0%)
Haematuria (4)	9 (8.4%)	7 (6.3%)	3 (3.1%)	6 (7.4%)	12 (5.9%)	13 (6.7%)
Proteinuria (4)	31 (29.0%)	28 (25.0%)	17 (17.7%)	21 (25.9%)	48 (23.6%)	49 (25.4%)
Rash (2)	74 (69.2%)	75 (67.0%)	72 (75.0%)	52 (64.2%)	146 (71.9%)	127 (65.8%)
Alopecia (2)	66 (61.7%)	69 (61.6%)	50 (52.1%)	38 (46.9%)	116 (57.1%)	107 (55.4%)
Mucosal Ulcers (2)	28 (26.2%)	20 (17.9%)	30 (31.3%)	22 (27.2%)	58 (28.6%)	42 (21.8%)
Low Complement (2)	107 (100.0%)	112 (100.0%)	96 (100.0%)	80 (98.8%)*	203 (100.0%)	192 (99.5%)
Increased DNA Binding (2)	107 (100.0%)	112 (100.0%)	96 (100.0%)	81 (100.0%)	203 (100.0%)	193 (100.0%)
Leukopenia (1)	6 (5.6%)	4 (3.6%)	7 (7.3%)	10 (12.3%)	13 (6.4%)	14 (7.3%)

*One patient had low complement at baseline based on laboratory data, however, investigator did not check 'low complement' on the SELENA-SLEDAI assessment.

A 5 What was the distribution of SLE improvements by organ system in the marketing authorisation population and Target Population at 52 weeks?

Table A5.1 Distribution of SLE improvements by organ system (SELENA-SLEDAI) – High disease activity subgroup (Target Population)

		BLISS-52		BLISS-76		Combined BLISS	
		SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
	Total enrolled	n=107	n=112	n=96	n=81	n=203	n=193
Mucocutaneous	Baseline involvement	100	104	88	71	188	175
	Number (%) improved at Week 52	38 (38%)	58 (56%)	36 (41%)	37 (52%)	74 (39%)	95 (54%)
	P-value[1]	-	0.0120	-	0.2004	-	0.0046
Immunology	Baseline involvement	107	112	96	81	203	193
	Number (%) improved at Week 52	13 (12%)	30 (27%)	12 (13%)	25 (31%)	25 (12%)	55 (28%)
	P-value[1]	-	0.0068	-	0.0031	-	<0.0001
Musculoskeletal	Baseline involvement	65	77	83	64	148	141
	Number (%) improved at Week 52	31 (48%)	53 (69%)	29 (35%)	36 (56%)	60 (41%)	89 (63%)
	P-value[1]	-	0.0161	-	0.0122	-	0.0002
Central nervous system	Baseline involvement	1	4	3	2	4	6
	Number (%) improved at Week 52	0 (0%)	2 (50%)	0 (0%)	2 (100%)	0 (0%)	4 (67%)
	P-value[1]	-	1.0000	-	0.1000	-	0.0762
Cardiovascular and respiratory	Baseline involvement	4	3	11	9	15	12
	Number (%) improved at Week 52	0	1 (33%)	6 (55%)	5 (56%)	6 (40%)	6 (50%)
	P-value[1]	-	0.4286	-	1.0	-	0.7068
Vascular	Baseline involvement	15	19	10	5	25	24
	Number (%) improved at Week 52	5 (33%)	14 (74%)	3 (30%)	3 (60%)	8 (32%)	17 (71%)
	P-value[1]	-	0.0359	-	0.3287	-	0.0101
Haematological and fever	Baseline involvement	7	7	12	15	19	22
	Number (%) improved at Week 52	4 (57%)	3 (43%)	3 (25%)	3 (20%)	7 (37%)	6 (27%)
	P-value[1]	-	1.0000	-	1.0000	-	0.7374
Renal	Baseline involvement	39	32	17	23	56	55
	Number (%) improved at Week 52	15 (38%)	15 (47%)	5 (29%)	11 (48%)	20 (36%)	26 (47%)
	P-value[1]	-	0.6296	-	0.3322	-	0.2505

[1]P-values are based on Fisher's exact test

A 6 What are the results of the efficacy end points for marketing authorisation population and Target Population?

Table A6.1 Selected efficacy endpoints – High disease activity subgroup (Target Population)

		BLISS-52		BLISS-76		Combined BLISS	
		SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
	Total enrolled	**n=107**	**n=112**	**n=96**	**n=81**	**n=203**	**n=193**
SRI at Week 52	n(%) responders	44 (41.1%)	75 (67.0%)	33 (34.4%)	46 (56.8%)	77 (37.9%)	121 (62.7%)
	OR (95% CI)[1]	-	3.0 (1.7, 5.2)	-	2.5 (1.3, 4.6)	-	2.7 (1.8, 4.1)
	P-value[1]	-	0.0001	-	0.0045	-	<0.0001
SRI at Week 76	n(%) responders	-	-	30 (31.3%)	40 (49.4%)	-	-
	OR (95% CI)[1]	-	-	-	2.1 (1.1, 3.9)	-	-
	P-value[1]	-	-	-	0.0188	-	-
Modified SRI at Week 52[4]	n(%) responders	42 (39.3%)	73 (65.2%)	29 (30.2%)	43 (53.1%)	71 (35.0%)	116 (60.1%)
	OR (95% CI)[1]	-	3.0 (1.7, 5.2)	-	2.5 (1.4, 4.8)	-	2.8 (1.8, 4.2)
	P-value[1]	-	0.0001	-	0.0036	-	<0.0001
SLEDAI (≥4 reduction)	n(%) responders	47 (43.9%)	76 (67.9%)	37 (38.5%)	49 (60.5%)	84 (41.4%)	125 (64.8%)
	OR (95% CI)[1]	-	2.8 (1.6, 4.8)	-	2.4 (1.3, 4.4)	-	2.6 (1.7, 3.9)
	P-value[1]	-	0.0004	-	0.0063	-	<0.0001
SLEDAI change in mean score by Week 52	Mean ± SE	-4.1 ± 0.4	-6.3 ± 0.5	-4.0 ± 0.5	-5.2 ± 0.5	-4.1 ± 0.3	-5.8 ± 0.3
	P-value[2]	-	0.0008	-	0.1705	-	0.0005
No new BILAG 1A/2B	n(%) responders	68 (63.6%)	88 (78.6%)	57 (59.4%)	57 (70.4%)	125 (61.6%)	145 (75.1%)
	OR (95% CI)[3]	-	2.3 (1.2, 4.2)	-	1.6 (0.9, 3.1)	-	1.9 (1.2, 3.0)
	P-value[3]	-	0.0099	-	0.1297	-	0.0034
No worsening in PGA	n(%) responders	64 (59.8%)	86 (76.8%)	55 (57.3%)	56 (69.1%)	119 (58.6%)	142 (73.6%)
	OR (95% CI)[1]	-	2.3 (1.3, 4.2)	-	1.6 (0.9, 3.0)	-	2.0 (1.3, 3.1)
	P-value[1]	-	0.0063	-	0.1312	-	0.0015
Prednisone reduction by ≥ 25% from baseline to ≤ 7.5mg/day during Weeks 40 through 52	Number at risk	n=76	n=81	n=50	n=45	n=126	n=126
	n(%) responders	4 (5.3%)	15 (18.5%)	5 (10.0%)	5 (11.1%)	9 (7.1%)	20 (15.9%)
	OR (95% CI)[1]	-	4.11 (1.29, 13.2)	-	0.88 (0.21, 3.60)	-	2.43 (1.05, 5.65)
	P-value[1]	-	0.0171	-	0.8586		0.0389

[1]Odds Ratios and 95% confidence intervals were from a logistic regression for the comparison between each belimumab dose and placebo with covariates, including baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs. other).

[2]All statistics were from ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for the baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs. other).

[3]Odds Ratio (95% confidence interval) and p-value were from a logistic regression for the comparison between each belimumab dose and placebo with covariates, including baseline BILAG domain involvement (at least 1A/2B versus at most 1B), baseline proteinuria level (<2 g/24 hour vs. >=2 g/24 hour equivalent) and race African descent or indigenous-American descent vs. other).

[4]Defined as SRI response with serology components (increased DNA binding and low complement items) removed.

A 7 Please complete the table below for mean of the change in SLEDAI score:

Table A7.1 Mean of the change in SELENA-SLEDAI score – High disease activity subgroup (Target Population)

Defining Week 24 responders as SELENA-SLEDAI (≥4 reduction) at Week 24	BLISS-52		BLISS-76		Combined BLISS	
	SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
Among Week 24 responders:	**n=63**	**n=80**	**n=42**	**n=50**	**n=105**	**n=130**
Mean of the change in SELENA-SLEDAI score by Week 24	-6.7 ± 0.3	-7.4 ± 0.4	-7.1 ± 0.5	-7.0 ± 0.4	-6.9 ± 0.3	-7.3 ± 0.3
P-value[1]	-	0.1786	-	0.8767	-	0.2571
Mean of the change in SELENA-SLEDAI score by Week 52	-5.4 ± 0.5	-7.5 ± 0.4	-6.1 ± 0.6	-6.8 ± 0.5	-5.7 ± 0.4	-7.2 ± 0.3
P-value[1]	-	0.0018	-	0.4409	-	0.0028
Among Week 24 non-responders:	**n=44**	**n=32**	**n=54**	**n=31**	**n=98**	**n=63**
Mean of the change in SELENA-SLEDAI score by Week 24	-0.7 ± 0.6	-0.8 ± 1.1	-1.3 ± 0.4	-1.0 ± 0.5	-1.1 ± 0.4	-0.9 ± 0.6
P-value[1]	-	0.5078	-	0.2581	-	0.2747
Mean of the change in SELENA-SLEDAI score by Week 52	-2.3 ± 0.7	-3.3 ± 1.1	-2.3 ± 0.6	-2.5 ± 0.8	-2.3 ± 0.4	-2.9 ± 0.7
P-value[1]	-	0.5860	-	0.9651	-	0.6822
Overall:	**n=107**	**n=112**	**n=96**	**n=81**	**n=203**	**n=193**
Mean of the change in SELENA-SLEDAI score by Week 24	-4.2 ± 0.4	-5.5 ± 0.5	-3.9 ± 0.4	-4.7 ± 0.4	-4.1 ± 0.3	-5.2 ± 0.3
P-value[1]	-	0.0508	-	0.3043	-	0.0238
Mean of the change in SELENA-SLEDAI score by Week 52	-4.1 ± 0.4	-6.3 ± 0.5	-4.0 ± 0.5	-5.2 ± 0.5	-4.1 ± 0.3	-5.8 ± 0.3
P-value[1]	-	0.0008	-	0.1705	-	0.0005

[1]All statistics were from ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for the baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs. other).

A 8 Please complete the table below on steroid use:

Table A8.1 Steroid use – High disease activity subgroup (Target Population)

Defining Week 24 responders as SELENA-SLEDAI (≥4 reduction) at Week 24	BLISS-52		BLISS-76		Combined BLISS	
	SoC	10 mg/kg	SoC	10 mg/kg	SoC	10 mg/kg
Among Week 24 responders:	**n=63**	**n=80**	**n=42**	**n=50**	**n=105**	**n=130**
Average steroid dose at baseline, mg	12.4 ± 8.6	13.7 ± 10.8	8.9 ± 8.2	10.9 ± 7.6	11.0 ± 8.6	12.6 ± 9.8
Average steroid dose at Week 24, mg	14.4 ± 11.7	11.5 ± 8.0	9.1 ± 6.7	11.4 ± 8.4	12.3 ± 10.3	11.5 ± 8.1
Average steroid dose at Week 52, mg	10.5 ± 6.3	8.6 ± 5.9	7.5 ± 6.2	11.9 ± 21.9	9.3 ± 6.4	9.8 ± 14.1
Among Week 24 non-responders:	**n=44**	**n=32**	**n=54**	**n=31**	**n=98**	**n=63**
Average steroid dose at baseline, mg	13.4 ± 8.1	13.6 ± 9.3	11.3 ± 9.2	9.5 ± 8.9	12.2 ± 8.7	11.6 ± 9.3
Average steroid dose at Week 24, mg	17.8 ± 30.2	14.1 ± 10.1	20.5 ± 54.8	12.5 ± 8.8	19.2 ± 44.8	13.3 ± 9.4
Average steroid dose at Week 52, mg	13.3 ± 7.0	12.1 ± 9.3	9.9 ± 9.2	9.9 ± 8.3	11.4 ± 8.3	11.0 ± 8.8
Overall:	**n=107**	**n=112**	**n=96**	**n=81**	**n=203**	**n=193**
Average steroid dose at baseline, mg	12.8 ± 8.4	13.7 ± 10.4	10.3 ± 8.8	10.4 ± 8.1	11.6 ± 8.6	12.3 ± 9.6
Average steroid dose at Week 24, mg	15.7 ± 20.6	12.1 ± 8.6	14.9 ± 39.5	11.8 ± 8.5	15.3 ± 30.6	12.0 ± 8.5
Average steroid dose at Week 52, mg	11.4 ± 6.6	9.4 ± 6.9	8.6 ± 7.8	11.2 ± 18.6	10.2 ± 7.3	10.1 ± 12.9

Please note that corticosteroid taper during the study was determined strictly at the investigators' discretion. There were no protocol mandates regarding dose reduction. The total dose of corticosteroids could be adjusted as clinically required during the first 24 weeks of the study; corticosteroid use beyond pre-specified dose limits resulted in the patient being designated as a non-responder.

A 9 The clinical effectiveness section notes that region was pre-specified as a subgroup in the data analysis protocol. It further notes that the major subgroups as predictor of response rate pooled between arms in order of significance were: baseline SS, complement, immunosuppressant use, region, SDI and anti-dsDNA. Please present the results of this analysis.

Table A9.1 Predictors of response at Week 52 – Pooled Total Population

	n	Observed Response Rate	AOR[1]	95% CI		P-value
Placebo	562	38.79%	-	-	-	-
1 mg/kg	559	46.15%	1.494	1.161	1.924	**0.0018**
10 mg/kg	563	50.62%	1.773	1.378	2.281	**<0.0001**
BLISS-76	819	39.19%	-	-	-	-
BLISS-52	865	50.87%	1.118	0.782	1.598	0.5419
SELENA-SLEDAI ≤ 9	806	34.49%	-	-	-	-
SELENA-SLEDAI ≥ 10	878	55.01%	3.018	2.411	3.779	**<0.0001**
USA/Canada	436	34.86%	-	-	-	-
Western Europe	217	38.25%	1.127	0.787	1.614	0.5128
Eastern Europe	191	51.31%	1.812	1.203	2.730	**0.0045**
America excluding USA/Canada	516	56.20%	2.086	1.384	3.146	**0.0004**
Asia	324	42.59%	1.357	0.835	2.205	0.2182
Anti-dsDNA < 30 U/mL	516	47.29%	-	-	-	-
Anti-dsDNA ≥ 30 U/mL	1168	44.26%	0.757	0.588	0.976	**0.0319**
C3 and C4 normal	638	50.00%	-	-	-	-
C3 or C4 low, but not both	390	45.64%	0.689	0.518	0.915	**0.0101**
C3 and C4 low	656	40.24%	0.488	0.373	0.637	**<0.0001**
Immunosuppressive: No	864	50.69%	-	-	-	-
Immunosuppressive: Yes	820	39.39%	0.680	0.552	0.839	**0.0003**
SLICC = 0	980	47.76%	-	-	-	-
SLICC = 1	390	45.90%	0.948	0.737	1.220	0.6799
SLICC ≥ 2	313	36.42%	0.622	0.468	0.827	**0.0011**

[1]AOR: Adjusted Odds Ratio; Multivariate analysis

Please note that the statistics in the table are for main effects which indicate independent predictors of response. In order to understand whether the treatment response (belimumab vs placebo) varied across different categories within each predictor (subgroup) the treatment-by-subgroup interactions were also examined. Significant treatment-by-subgroup interactions ($p < 0.05$) were observed for baseline complement levels, with a trend observed for baseline SELENA SLEDAI. Subjects responded better with belimumab vs. placebo if they entered the study with low complement (C3 and/or C4). Subjects with baseline SELENA SLEDAI ≥ 10 had a trend ($p=0.0610$) for a better response for belimumab vs placebo.

There were no significant treatment by subgroup interactions for study ($p = 0.9505$), anti-dsDNA ($p = 0.3259$), immunosuppressant use ($p = 0.7699$); SLICC Damage Index score ($p = 0.2100$), or region ($p = 0.2435$).

Related to the above question, did the major subgroup analysis include an analysis of them as predictors of the odds ratio of response for belimumab 10mg/kg versus SoC? If this, or a similar analysis of relative efficacy, has been undertaken please present the results of this analysis.

The results of the main effects model demonstrated that belimumab treatment significantly increased the odds of a Week 52 response compared with placebo while controlling for predictive baseline characteristics, with adjusted odds ratios of 1.8 ($p < 0.0001$) for belimumab 10 mg/kg. These results showed that controlling for the additional predictive characteristics slightly increased the estimated odds ratios compared with those estimated using the primary analysis model (OR = 1.8 vs 1.7 for belimumab 10 mg/kg).

We were unable to perform analyses of interaction terms in our Target Population as this is a combination of the high performing subgroups (Anti-dsDNA+/Low Complement, Low Complement, SS>=10, prednisone use/Low Complement). In order to perform such analyses, a portion of subjects would have to fall in the normal/high complement, SS≤9, and Anti-dsDNA- categories.

A 10 Please provide marketing authorisation population and Target Population major outcome results (SRI, SLEDAI [4 point reduction], and reduction in steroid use weeks 40 to 52) for the pooled 10mg Belimumab and SoC groups by geographical region (e.g. W. Europe, North America (USA/Canada), America (excluding USA Canada), E Europe, Asia, Australia).

Please note as there were only 2 patients from Australia, these have been combined with Western Europe/Israel.

Table A10.1 Primary efficacy endpoint (SRI) at Week 52, ≥ 4-point reduction in SELENA-SLEDAI at Week 52, prednisone reduction by ≥ 25% from baseline to ≤ 7.5mg/day during Weeks 40 through 52 – High disease activity subgroup (Target Population)

		Combined BLISS	
		SoC	**10 mg/kg**
		n=45	**n=24**
USA/Canada	**Primary efficacy endpoint (SRI) at Week 52**		
	· No.(%) Response	13 (28.9%)	12 (50.0%)
	· Observed difference vs Placebo	-	21.1
	· OR (95% CI) vs Placebo[1]	-	2.5 (0.9, 6.9)
	· P-value[1]	-	0.0858
	≥ 4-point reduction in SELENA-SLEDAI at Week 52		
	· No.(%) Response	16 (35.6%)	13 (54.2%)
	· Observed difference vs placebo (%)	-	18.6
	· OR (95% CI) vs placebo[1]	-	2.1 (0.8, 5.9)
	· P-value[1]	-	0.1388
		n=22	**n=12**
	Prednisone reduction by ≥ 25% from baseline to ≤		

		Combined BLISS	
		SoC	**10 mg/kg**
	7.5mg/day during Weeks 40 through 52 · No.(%) Response · Observed difference vs Placebo · OR (95% CI) vs placebo[2] · P-value[2]	 3 (13.6%) - - -	 2 (16.7%) 3.03 1.00 (0.13, 7.35) 0.9968
		n=71	**n=63**
America excluding USA/Canada	**Primary efficacy endpoint (SRI) at Week 52** · No.(%) Response · Observed difference vs Placebo · OR (95% CI) vs Placebo[1] · P-value[1]	 32 (45.1%) - - -	 39 (61.9%) 16.8 2.0 (1.0, 4.0) 0.0525
	≥ 4-point reduction in SELENA-SLEDAI at Week 52 · No.(%) Response · Observed difference vs placebo (%) · OR (95% CI) vs placebo[1] · P-value[1]	 34 (47.9%) - - -	 40 (63.5%) 15.6 1.9 (0.9, 3.8) 0.0711
		n=51	**n=47**
	Prednisone reduction by ≥ 25% from baseline to ≤ 7.5mg/day during Weeks 40 through 52 · No.(%) Response · Observed difference vs Placebo · OR (95% CI) vs placebo[2] · P-value[2]	 1 (2.0%) - - -	 6 (12.8%) 10.81 6.46 (0.72, 57.8) 0.0949
		n=26	**n=30**
Western Europe/Australia/ Israel	**Primary efficacy endpoint (SRI) at Week 52** · No.(%) Response · Observed difference vs Placebo · OR (95% CI) vs Placebo[1] · P-value[1]	 8 (30.8%) - - -	 17 (56.7%) 25.9 2.9 (1.0, 8.9) 0.0550
	≥ 4-point reduction in SELENA-SLEDAI at Week 52 · No.(%) Response · Observed difference vs placebo (%) · OR (95% CI) vs placebo[1] · P-value[1]	 8 (30.8%) - - -	 18 (60.0%) 29.2 3.4 (1.1, 10.2) 0.0314
		n=14	**n=14**
	Prednisone reduction by ≥ 25% from baseline to ≤ 7.5mg/day during Weeks 40 through 52 · No.(%) Response · Observed difference vs Placebo · OR (95% CI) vs placebo[2] · P-value[2]	 2 (14.3%) - - -	 2 (14.3%) 0.00 1.27 (0.14, 11.1) 0.8314
		n=22	**n=23**
Eastern Europe	**Primary efficacy endpoint (SRI) at Week 52** · No.(%) Response · Observed difference vs Placebo	 8 (36.4%) -	 17 (73.9%) 37.5

		Combined BLISS	
		SoC	**10 mg/kg**
	· OR (95% CI) vs Placebo[1]	-	5.0 (1.4, 17.7)
	· P-value[1]	-	0.0137
	≥ 4-point reduction in SELENA-SLEDAI at Week 52		
	· No.(%) Response	9 (40.9%)	18 (78.3%)
	· Observed difference vs placebo (%)	-	37.4
	· OR (95% CI) vs placebo[1]	-	5.2 (1.4, 19.2)
	· P-value[1]	-	0.0133
		n=13	**n=19**
	Prednisone reduction by ≥ 25% from baseline to ≤ 7.5mg/day during Weeks 40 through 52		
	· No.(%) Response	0 (0.0%)	2 (10.5%)
	· Observed difference vs Placebo	-	10.53
	· OR (95% CI) vs placebo[2]	-	-[3]
	· P-value[2]	-	-[3]
		n=39	**n=53**
Asia	**Primary efficacy endpoint (SRI) at Week 52**		
	· No.(%) Response	16 (41.0%)	36 (67.9%)
	· Observed difference vs Placebo	-	26.9
	· OR (95% CI) vs Placebo[1]	-	3.0 (1.3, 7.2)
	· P-value[1]	-	0.0112
	≥ 4-point reduction in SELENA-SLEDAI at Week 52		
	· No.(%) Response	17 (43.6%)	36 (67.9%)
	· Observed difference vs placebo (%)	-	24.3
	· OR (95% CI) vs placebo[1]	-	2.7 (1.2, 6.5)
	· P-value[1]	-	0.0210
		n=26	**n=34**
	Prednisone reduction by ≥ 25% from baseline to ≤ 7.5mg/day during Weeks 40 through 52		
	· No.(%) Response	3 (11.5%)	8 (23.5%)
	· Observed difference vs Placebo	-	11.99
	· OR (95% CI) vs placebo[2]	-	2.43 (0.57, 10.4)
	· P-value[2]	-	0.2321

[1]Odds Ratios and 95% confidence intervals were from a logistic regression for the comparison between each belimumab dose and placebo.

[2]Odds Ratio (95% confidence interval) and p-value from a logistic regression for the comparison between each belimumab dose and placebo adjusted for baseline prednisone dose.

[3]Model was not fit if any treatment arm had 0 responders.

A 11 Please clarify the reasons for non-responder status (SRI) by week 52 in the Target Population. Please complete the following table.

Table A11.1 Reasons for non-responder status at Week 52 – High disease activity subgroup (Target Population)

		BLISS-52		BLISS-76		Combined BLISS	
		SoC (n=107)	**10 mg/kg (n=112)**	**SoC (n=96)**	**10 mg/kg (n=81)**	**SoC (n=203)**	**10 mg/kg (n=193)**
Non-responders at wk 52 n (%)		63 (58.88%)	37 (33.04%)	63 (65.63%)	35 (43.21%)	126 (62.07%)	72 (37.31%)
Non-responders - Lack of SRI		32 (50.79%)	14 (37.84%)	31 (49.21%)	15 (42.86%)	63 (50.0%)	29 (40.28%)
Non-responders - Patient withdrawals	All withdrawals	31 (49.21%)	23 (62.16%)	32 (50.79%)	20 (57.14%)	63 (50.0%)	43 (59.72%)
	Subject request	0	2 (5.41%)	12 (19.05%)	1 (2.86%)	12 (9.52%)	3 (4.17%)
	Adverse event	12 (19.05%)	6 (16.22%)	9 (14.29%)	9 (25.71%)	21 (16.67%)	15 (20.83%)
	Lack of efficacy	11 (17.46%)	7 (18.92%)	7 (11.11%)	6 (17.14%)	18 (14.29%)	13 (18.06%)
	Lack of compliance	1 (1.59%)	1 (2.70%)	0	0	1 (0.79%)	1 (1.39%)
	Lost to follow-up	0	2 (5.41%)	0	1 (2.86%)	0	3 (4.17%)
	Protocol violation	5 (7.94%)	2 (5.41%)	2 (3.17%)	2 (5.71%)	7 (5.56%)	4 (5.56%)
	Investigator decision	2 (3.17%)	1 (2.70%)	1 (1.59%)	1 (2.86%)	3 (2.38%)	2 (2.78%)
	Other	0	2 (5.41%)	1 (1.59%)	0	1 (0.79%)	2 (2.78%)
Treatment failure by drug type[#]	Any drug type	16 (25.40%)	10 (27.03%)	16 (25.40%)	9 (25.71%)	32 (25.40%)	19 (26.39%)
	Steroid	10 (15.87%)	10 (27.03%)	11 (17.46%)	8 (22.86%)	21 (16.67%)	18 (25.00%)
	Antimalarial	1 (1.59%)	0	1 (1.59%)	0	2 (1.59%)	0
	Immunosuppressant	3 (4.76%)	0	4 (6.35%)	1 (2.86%)	7 (5.56%)	1 (1.39%)
	Lipid lowering	1 (1.59%)	0	1 (1.59%)	0	2 (1.59%)	0
	Hypertension	2 (3.17%)	0	2 (3.17%)	0	4 (3.17%)	0
	Other (NSAID)	1	0	0	0	1	0

		(1.59%)				(0.79%)	

#Treatment failures can come from either non-responders due to a lack of SRI response or non-responders due to patient withdrawals.. Four subjects on placebo reported multiple medications (two subjects each per study).

A 12 Please clarify how the standard of care (SoC) in the BLISS trials relates to that current in the UK for the appropriate high disease activity population. Please clarify differences between trial centres regarding SoC.

As mentioned in our original submission, belimumab will be administered alongside standard therapy for SLE, which in the clinical trials has included antimalarials, non-steroidal anti-inflammatory drugs (NSAIDs), corticosteroids and other immunosuppressants. This is in line with current treatment practice in the UK, where the choice of treatment is dependant upon the organ systems involved and individual patient factors. Other biologics (e.g. rituximab), although used in UK clinical practice for a select group of patients, were prohibited at any time during the BLISS trials to ensure patients were not receiving two biologics (safety concerns) and so as not to confound the trial results.

In the high disease activity population, standard of care consisted largely of corticosteroids with/without antimalarials and/or immunosuppressants (see Table A12.1). This is largely reflective of UK clinical practice as outlined previously in the clinical pathway of care (Figure 1.1, Page 46 of submission document).

Presenting differences in SOC between trial centres would have resulted in very small patient numbers which would be difficult to interpret. Instead we have presented regional data (see Table A12.1). Whilst regional differences exist, corticosteroid use in addition to antimalarials and/or immunosuppressants remains universal in the high disease activity population.

Table A12.1 Steroid, antimalarial and immunosuppressant use at baseline – High disease activity subgroup (Target Population)

		Combined BLISS	
		SoC	**10 mg/kg**
		n=203	**n=193**
All countries/regions	Corticosteroid only	18 (9%)	30 (16%)
	Immunosuppressant only	3 (1%)	7 (4%)
	Antimalarial only	6 (3%)	8 (4%)
	Steroid and immunosuppressant only	33 (16%)	42 (22%)
	Steroid and antimalarial only	68 (33%)	54 (28%)
	Immunosupressant and antimalarial only	8 (4%)	0
	Steroid and immunosuppressant and antimalarial	64 (32%)	51 (26%)
		n=45	**n=24**
USA/Canada	Corticosteroid only	3 (7%)	3 (13%)
	Immunosuppressant only	1 (2%)	2 (8%)

		Combined BLISS	
		SoC	**10 mg/kg**
	Antimalarial only	3 (7%)	4 (17%)
	Steroid and immunosuppressant only	8 (18%)	5 (21%)
	Steroid and antimalarial only	9 (20%)	3 (13%)
	Immunosupressant and antimalarial only	5 (11%)	0
	Steroid and immunosuppressant and antimalarial	15 (33%)	7 (29%)
		n=71	**n=63**
America excluding USA/Canada	Corticosteroid only	7 (10%)	6 (10%)
	Immunosuppressant only	1 (1%)	0
	Antimalarial only	1 (1%)	2 (3%)
	Steroid and immunosuppressant only	12 (17%)	15 (24%)
	Steroid and antimalarial only	28 (39%)	23 (37%)
	Immunosupressant and antimalarial only	2 (3%)	0
	Steroid and immunosuppressant and antimalarial	19 (27%)	17 (27%)
		n=26	**n=30**
Western Europe/Australia/ Israel	Corticosteroid only	1 (4%)	2 (7%)
	Immunosuppressant only	1 (4%)	4 (13%)
	Antimalarial only	2 (8%)	2 (7%)
	Steroid and immunosuppressant only	6 (23%)	7 (23%)
	Steroid and antimalarial only	5 (19%)	2 (7%)
	Immunosupressant and antimalarial only	1 (4%)	0
	Steroid and immunosuppressant and antimalarial	9 (35%)	13 (43%)
		n=22	**n=23**
Eastern Europe	Corticosteroid only	3 (14%)	7 (30%)
	Immunosuppressant only	0	0
	Antimalarial only	0	0
	Steroid and immunosuppressant only	6 (27%)	7 (30%)
	Steroid and antimalarial only	9 (41%)	8 (35%)
	Immunosupressant and antimalarial only	0	0
	Steroid and immunosuppressant and antimalarial	4 (18%)	1 (4%)
		n=39	**n=53**
Asia	Corticosteroid only	4 (10%)	12 (23%)
	Immunosuppressant only	0	1 (2%)
	Antimalarial only	0	0
	Steroid and immunosuppressant only	1 (3%)	8 (15%)
	Steroid and antimalarial only	17 (44%)	18 (34%)
	Immunosupressant and antimalarial only	0	0
	Steroid and immunosuppressant and antimalarial	17 (44%)	13 (25%)

Non-priority Questions

A 13 Fig 5.9 shows a decline in SRI in the placebo group from wk24 to wk52. Is this attributable to patients receiving non-permitted steroid use during this period?

From Weeks 24 to 52, there were 105 (51.7%) non-responders in the placebo group (n=203). Of the non-responders (n=105), 63 (60%) were deemed non-responders due to a lack of SRI response and 42 (40%) due to patient withdrawals. Amongst non-responders, 30 were non-responders due to treatment failure of which 21 were due to non-permitted steroid use. Reasons for non-responder status from Weeks 24 to 52 are summarised in Table A13.1 below.

Table A13.1 Reasons for non-responder status from Week 24 to 52 – High disease activity subgroup (Target Population)

		Combined BLISS	
		SoC (n=203)	**10 mg/kg (n=193)**
Non-responders from Week 24 to Week 52 n (%)		105 (51.7%)	57 (29.5%)
Non-responders - Lack of SRI		63 (60.0%)	29 (50.9%)
Non-responders - Patient withdrawals	All withdrawals	42 (40.0%)	28 (49.1%)
	Subject request	4 (3.8%)	0
	Adverse event	12 (11.4%)	9 (15.8%)
	Lack of efficacy	17 (16.2%)	10 (17.5%)
	Lack of compliance	0	1 (1.8%)
	Lost to follow-up	0	1 (1.8%)
	Protocol violation	6 (5.7%)	4 (7.0%)
	Investigator decision	2 (1.9%)	1 (1.8%)
	Other	1 (1.0%)	2 (3.5%)
Treatment failure by drug type#	Any drug type	30 (28.6%)	19 (33.3%)
	Steroid	21 (20.0%)	18 (31.6%)
	Antimalarial	2 (1.9%)	0
	Immunosuppressant	5 (4.8%)	1 (1.8%)
	Lipid lowering	2 (1.9%)	0
	Hypertension	4 (3.8%)	0
	Other (NSAID)	1 (1.0%)	0

#Treatment failures can come from either non-responders due to a lack of SRI response or non-responders due to patient withdrawals. Four subjects on placebo reported multiple medications (two subjects each per study).

A 14 For table 5.15 on page 106 of the submission, please clarify why the trial interaction is not applicable (N/A) for the Target Population pooled result while it is for the pooled result for the whole population (this appears at odds with the figure footnote)

At the time of our original submission, we did not have a value for the treatment by study interation p-value for the high disease activity subgroup. The p-value for the

treatment by study interaction from the logistic regression model is 0.744 for the belimumab 10mg/kg vs placebo comparison.

A 15 Page 119 of the submission states: “However, for the pooled total population, both belimumab doses achieved significance compared with placebo for reduction of prednisone by ≥ 25% from baseline to ≤ 7.5 mg/day during.” Please clarify whether this refers to results for the 1mg/kg dose regimen in the phase 3 trials. If so please provide results.

This sentence should read ‘However, for the pooled total population, **belimumab 10 mg/kg** achieved significance compared with placebo for reduction of prednisone by ≥ 25% from baseline to ≤ 7.5 mg/day during…’.

Whilst a 1 mg/kg dose was examined in the Phase 3 studies, we will only present results for the 10 mg/kg belimumab dose as this is the dose submitted for Marketing Authorisation. Results for the 1 mg/kg belimumab dose are included in the FDA Arthritis Advisory Committee Meeting Briefing Document for Belimumab, 16th November 2010.

A 16 The systematic review searches yielded 36 publications. These are not listed in the submitted documents. Please send a list of publications with reasons for exclusion.

The number of included publications was 43 (36 full publications plus seven abstracts), including eight publications (of four trials) of belimumab and 35 publications of other interventions. Table A16.1 below lists all the publications and the reasons for their exclusion. None of the excluded publications directly compare belimumab with the appropriate comparators.

Table A16.1 Summary of publications of RCTs reviewed and their reasons for exclusion

Publication	Reason for exclusion
1. Wallace DJ, Stohl W, Furie RA, Lisse JR, McKay JD, Merrill JT, *et al*. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. Arthritis Care and Research. 2009 15;61 (9):1168-78.	Included (LBSL02).
2. Furie RA, Petri MA, Wallace DJ, Ginzler EM, Merrill JT, Stohl W, *et al*. Novel evidence-based systemic lupus erythematosus responder index. Arthritis & Rheumatism. [Research Support, N.I.H., Extramural]. 2009 Sep 15;61(9):1143-51.	Included. Linked to LBSL02.
3. Furie R, Stohl W, Ginzler EM, Becker M, Mishra N, Chatham W, *et al*. Biologic activity and safety of belimumab, a neutralizing anti-B-lymphocyte stimulator (BLyS) monoclonal antibody: a phase I trial in patients with systemic lupus erythematosus. Arthritis research & therapy. 2008;10 (5):R109.	Included (LBSL01).
4. Navarra S, Ilianova E, Bae SC, Guzman R, *et al*. Belimumab, a BLYS-specific inhibitor, reduced disease activity, flares, and steroid use in patients with seropositive systemic lupus erythematosus (SLE): BLISS-52 study. EULAR. 2010:Abstract SAT0204.	Included (BLISS-52).
5. Tanasescu C, Gallacher A, Garcia M, Littlejohn G, Saaibi D, *et al*. Belimumab, a BLYS-specific inhibitor, significantly improved physical functioning, fatigue, and other health-related quality of life (HRQOL) measures in patients with seropositive systemic lupus erythematosus (SLE): BLISS-52 study. EULAR. 2010:abstract SAT0206.	Included (BLISS-52).
6. D'Cruz D, Tanasescu C, Navarra S, Guzman R, *et al*. Belimumab, a BLYS-specific inhibitor, reduced disease activity, flares and prednisone use in patients with active seropositive SLE: Phase 3 BLISS-52 study. BSR. 2010:abstract OP3.	Included (BLISS-52).
7. Furie R, Zamani O, Wallace D, Tegzova D, *et al*. Belimumab, a BLyS-Specific Inhibitor, Reduced Disease Activity and Severe Flares in Seropositive SLE Patients: BLISS-76 Study Results through Wk 76 ACR. 2010:Abstract 1454.	Included (BLISS-76).
8. Petri M, Van Vollenhoven RF, Zamani O, Furie RA, *et al*. Belimumab, a BLyS-Specific Inhibitor, Reduces Disease Activity and Severe Flares in Seropositive Systemic Lupus Erythematosus (SLE) Patients: BLISS-76 Study. International Journal of Rheumatic Diseases; Asia Pacific League of Associations of Rheumatology 2010;13:suppl. 1: 110-5, abstract 0281.	Included (BLISS-76).
9. Wallace DJ, Kalunian KC, Petri MA, Strand CV, *et al*. Epratuzumab Demonstrates Clinically Meaningful Improvements in Patients with Moderate to Severe Systemic Lupus Erythematosus (SLE): Results from EMBLEM™, a Phase IIb Study ACR. 2010:Abstract 1452.	Investigational drug. Not yet available in the UK.
10. Carneiro JRM, Sato EI. Randomized double-blind clinical study with methotrexate in systemic lupus erythematosus. [Portuguese]. Revista Brasileira de Reumatologia. 1999;39 (4):203-10.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
11. Islam N, Hossain M, Atiqul Haq S, Noor Alam M, *et al*. Efficacy and safety of methotrexate (MTX) in articular and	No requirement for patients to have active autoantibody-positive systemic lupus

Publication	Reason for exclusion
cutaneous manifestations of systemic lupus erythematosus. EULAR. 2006:Abstract THU0273.	erythematosus. Focus on articular and cutaneous manifestations only.
12. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, *et al*. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: The randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis and Rheumatism. 2010 January;62 (1):222-33.	Included.
13. Andrade-Ortega L, Irazoque-Palazuelos F, Lopez-Villanueva R, Barragan-Navarro Y, Bourget-Pietrasanta F, Diaz-Ceballos MDLT, *et al*. Efficacy of rituximab versus cyclophosphamide in lupus patients with severe manifestations. A randomized and multicentre study. [Spanish]. Reumatologia Clinica. 2010 September;6 (5):250-5.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Excluded patients on other immunosuppressants (except antimalarials). Cyclophosphamide is not a relevant comparator.
14. Fortin PR, Abrahamowicz M, Ferland D, Lacaille D, Smith CD, Zummer M. Steroid-sparing effects of methotrexate in systemic lupus erythematosus: A double-blind, randomized, placebo-controlled trial. Arthritis Care and Research. 2008 15;59 (12):1796-804.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Excluded patients taking azathioprine.
15. Carneiro JRM, Sato EI. Double blind, randomized, placebo controlled clinical trial of methotrexate in systemic lupus erythematosus. Journal of Rheumatology. 1999;26 (6):1275-9.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
16. Barile-Fabris L, Ariza-Andraca R, Olguin-Ortega L, Jara LJ, Fraga-Mouret A, Miranda-Limon JM, *et al*. Controlled clinical trial of IV cyclophosphamide versus IV methylprednisolone in severe neurological manifestations in systemic lupus erythematosus. Annals of the Rheumatic Diseases. 2005 Apr;64 (4):620-5.	Included patients with severe neurological involvement. Cyclophosphamide is not a relevant comparator.
17. Fries JF, Sharp GC, McDevitt HO, Holman HR. Cyclophosphamide therapy in systemic lupus erythematosus and polymyositis. Arthritis and Rheumatism. 1973 1973;16 (2):154-62.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Included patients with polymyositis. Cyclophosphamide is not a relevant comparator.
18. Dussán KB, Magder L, Brodsky RA, Jones RJ, Petri M. High dose cyclophosphamide performs better than monthly dose cyclophosphamide in quality of life measures. Lupus. 2008(12):1079-85.	Cyclophosphamide is not a relevant comparator.
19. Gonzalez-Lopez L, Cardona-Munoz EG, Celis A, Garcia-De La Torre I, Orozco-Barocio G, Salazar-Paramo M, *et al*. Therapy with intermittent pulse cyclophosphamide for pulmonary hypertension associated with systemic lupus erythematosus. Lupus. 2004;13 (2):105-12.	Cyclophosphamide is not a relevant comparator. Included patients with CNS lupus and lupus nephritis.
20. Petri M, Brodsky RA, Jones RJ, Gladstone D, Fillius M, Magder LS. High-dose cyclophosphamide versus monthly intravenous cyclophosphamide for systemic lupus erythematosus a prospective randomized trial. Arthritis and Rheumatism. 2010 May;62 (5):1487-93.	Cyclophosphamide is not a relevant comparator. Included patients with CNS lupus and lupus nephritis.
21. Bykerk V, Sampalis J, Esdaile JM, Choquette D, Senecal JL, Danoff D, *et al*. A randomized study of the effect of withdrawing hydroxychloroquine sulfate in systemic lupus erythematosus. New England Journal of Medicine. 1991;324 (3):150-4.	Withdrawal study in patients with stable SLE.
22. Tsakonas E, Joseph L, Esdaile JM, Choquette D, Senecal JL, Cividino A, *et al*. A long-term study of	Withdrawal study in patients with stable SLE.

Publication	Reason for exclusion
hydroxychloroquine withdrawal on exacerbations in systemic lupus erythematosus. Lupus. 1998;7 (2):80-5.	
23. Levy RA, Vilela VS, Cataldo MJ, Ramos RC, Duarte JLMB, Tura BR, *et al*. Hydroxychloroquine (HCQ) in lupus pregnancy: Double-blind and placebo-controlled study. Lupus. 2001;10 (6):401-4.	Study in pregnant patients.
24. Bezerra EL, Vilar MJ, da Trindade Neto PB, Sato EI. Double-blind, randomized, controlled clinical trial of clofazimine compared with chloroquine in patients with systemic lupus erythematosus. Arthritis and rheumatism. 2005(10):3073-8.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Clofazimine not available in the UK. Focus on cutaneous manifestations only.
25. Meinão IM, Sato EI, Andrade LE, Ferraz MB, Atra E. Controlled trial with chloroquine diphosphate in systemic lupus erythematosus. Lupus. 1996(3):237-41.	Chloroquine not available in the UK.
26. Danowski A, Magder L, Petri M. Flares in Lupus: Outcome Assessment Trial (FLOAT), a comparison between oral methylprednisolone and intramuscular triamcinolone. Journal of Rheumatology. 2006 January;33 (1):57-60.	Study in patients presenting with mild or moderate flare.
27. Dougados M, Job-Deslandre C, Amor B, Menkes CJ. Danazol therapy in systemic lupus erythematosus. A one-year prospective controlled trial on 40 female patients. Clinical Trials Journal. 1987;24 (2):191-200.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Danazol is not considered standard of care.
28. Bootsma H, Spronk P, Derksen R, De Boer G, Wolters-Dicke H, Hermans J, *et al*. Prevention of relapses in systemic lupus erythematosus. Lancet. 1995;345 (8965):1595-9.	Study designed to look at prevention of relapses in patients presenting with a rise in anti-dsDNA.
29. Edwards JC, Snaith ML, Isenberg DA. A double blind controlled trial of methylprednisolone infusions in systemic lupus erythematosus using individualised outcome assessment. Annals of the rheumatic diseases. 1987(10):773-6.	Study in patients with severe SLE presenting with an acute exacerbation.
30. Dammacco F, Della Casa Alberighi O, Ferraccioli G, Racanelli V, Casatta L, Bartoli E. Cyclosporine-A plus steroids versus steroids alone in the 12-month treatment of systemic lupus erythematosus. International Journal of Clinical and Laboratory Research. 2000;30 (2):67-73.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
31. Denburg SD, Carbotte RM, Denburg JA. Corticosteroids and neuropsychological functioning in patients with systemic lupus erythematosus. Arthritis and Rheumatism. 1994 Sep;37 (9):1311-20.	Study was designed to assess the effects of corticosteroids on nervous system functioning as well as disease-related symptoms in patients with mild SLE and mild neuropsychiatric symptoms.
32. Hahn BH, Kantor OS, Osterland CK. Azathioprine plus prednisone compared with prednisone alone in the treatment of systemic lupus erythematosus. Report of a prospective controlled trial in 24 patients. Annals of Internal Medicine. 1975 Nov;83(5):597-605.	Study in severe, life-threatening systemic lupus erythematosus.
33. Mackworth-Young CG, David J, Morgan SH, Hughes GR. A double blind, placebo controlled trial of intravenous methylprednisolone in systemic lupus erythematosus. Annals of the rheumatic diseases. 1988(6):496-502.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
34. Tseng CE, Buyon JP, Kim M, Belmont HM, Mackay M, Diamond B, *et al*. The effect of moderate-dose corticosteroids in preventing severe flares in patients with serologically active, but clinically stable, systemic lupus erythematosus: Findings of a prospective, randomized,	Included patients with inactive disease defined as a SLEDAI score ≤ 4.

Publication	Reason for exclusion
double-blind, placebo-controlled trial. Arthritis and Rheumatism. 2006 Nov;54 (11):3623-32.	
35. Mease PJ, Ginzler EM, Gluck OS, Schiff M, Goldman A, Greenwald M, *et al*. Effects of prasterone on bone mineral density in women with systemic lupus erythematosus receiving chronic glucocorticoid therapy. Journal of Rheumatology. 2005 Apr;32 (4):616-21.	Study designed to examine the effects of prasterone on bone mineral density in female patients with mild to moderate systemic lupus erythematosus.
36. Petri MA, Mease PJ, Merrill JT, Lahita RG, Iannini MJ, Yocum DE, *et al*. Effects of prasterone on disease activity and symptoms in women with active systemic lupus erythematosus: Results of a multicentre randomized, double-blind, placebo-controlled trial. Arthritis and Rheumatism. 2004 Sep;50 (9):2858-68.	Included patients with SLEDAI > 2.
37. Petri MA, Lahita RG, Van Vollenhoven RF, Merrill JT, Schiff M, Ginzler EM, *et al*. Effects of prasterone on corticosteroid requirements of women with systemic lupus erythematosus: A double-blind, randomized, placebo-controlled trial. Arthritis and Rheumatism. 2002;46 (7):1820-9.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
38. Sanchez-Guerrero J, Fragoso-Loyo HE, Neuwelt CM, Wallace DJ, Ginzler EM, Sherrer YRS, *et al*. Effects of prasterone on bone mineral density in women with active systemic lupus erythematosus receiving chronic glucocorticoid therapy. Journal of Rheumatology. 2008 August;35 (8):1567-75.	Study designed to examine the effects of prasterone on bone mineral density in female SLE patients.
39. Chang DM, *et al*. Dehydroepiandrosterone treatment of women with mild-to-moderate systemic lupus erythematosus. Arthritis & Rheumatism. 2002;46(11):2924-27.	Included patients with SLEDAI > 2.
40. Hartkamp A, Geenen R, Godaert GLR, Bijl M, Bijlsma JWJ, Derksen RHWM. Effects of dehydroepiandrosterone on fatigue and well-being in women with quiescent systemic lupus erythematosus: A randomised controlled trial. Annals of the Rheumatic Diseases. 2010 June;69 (6):1144-7.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
41. Nordmark G, Bengtsson C, Larsson A, Karlsson FA, Sturfelt G, Ronnblom L. Effects of dehydroepiandrosterone supplement on health-related quality of life in glucocorticoid treated female patients with systemic lupus erythematosus. Autoimmunity. 2005 Nov;38 (7):531-40.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
42. Van Vollenhoven RF, Engleman EG, McGuire JL. Dehydroepiandrosterone in systemic lupus erythematosus: Results of a double-blind, placebo-controlled, randomized clinical trial. Arthritis and Rheumatism. 1995 Dec;38 (12):1826-31.	Study in mild to moderate SLE. No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
43. Gordon C, Wallace DJ, Shinada S, Kalunian KC, Forbess L, Braunstein GD, *et al*. Testosterone patches in the management of patients with mild/moderate systemic lupus erythematosus. Rheumatology. 2008 Mar;47 (3):334-8.	Included patients with mild to moderate SLE defined by SELENA-SLEDAI ≥ 2.

A 17 Page 97 of the submission states: Pooling is appropriate given that the trials were essentially identical in design and in the analysis of the primary endpoint, the p-values for the treatment-by-study interaction were not significant (interaction p-values > 0.5).

However we note systematic differences between the trials with regard to demographic characteristics (Table 5.8 age and race), SLICC organ damage and proteinuria (Table 5.9), anti-dsDNA positivity, raised IgG, and abnormal complement (Table 5.10) so that pooling might be considered inappropriate: Please provide further justification for pooling results for Target Population.

Pooling studies under nearly identical protocols but with subjects of varying demographic and baseline characteristics can be justified by extension of the same principles outlined in ICH E9 (Statistical Principles of Clinical Trials) for which different centres in a single multicentre trial are pooled together, i.e. adherence to a common protocol that has been implemented in the same way at all centres using the same standardised procedures and evaluation criteria (as has been done in these studies), and a homogeneous treatment effect across studies (as is the case with these studies). In particular, when pooling the data across these studies, we considered study design, inclusion and exclusion criteria relative to disease severity, and whether the studies were run contemporarily such that the SoC treatment options were similar. These studies followed very similar protocols, were of nearly identical design, had identical inclusion and exclusion criteria, and were conducted over the same time period. Nevertheless, given the heterogeneous presentation of SLE disease and the fact that the Phase 3 program was run globally, one should expect to have variation in the patient population, both within the studies (e.g. between different centres) and between the studies (analogous to differences between centres within the same study).

Since it has been established that the conduct of the studies was effectively the same, one must then determine whether the relative treatment effect is different in one study compared with the other study when evaluating whether two studies are similar enough to pool. Each of the Phase 3 studies achieved statistical significance for belimumab 10mg/kg on the pre-specified primary endpoint of SRI response at Week 52; therefore, these nearly identical, studies provide independent replication of results. While pooling is not necessary to establish the effectiveness of belimumab, it was considered appropriate in order to evaluate treatment effects in high disease activity subgroups of interest, given that the individual studies were not designed to provide sufficient power to demonstrate effectiveness within subgroups. When the two Phase 3 studies were pooled a test for a treatment-by-study interaction was undertaken for the SRI analysis and the treatment-by-study interaction was >0.5. Likewise, for the Target Population of high disease activity, the treatment-by-study interaction was >0.7.

Additionally, a multivariate logistic regression model was developed in order to determine predicators of SRI response. Of the characteristics highlighted as being different between the two studies neither age, race, proteinuria, nor raised IgG were predictors of response. SLICC Damage Score and complement levels (and their interaction terms with treatment) were included in the final model and neither study (p=0.54) nor the treatment-by-study interaction (p=0.95) was a predictor of SRI response. This result further substantiates that the study is not a predictor of SRI response, thus we believe that is reasonable and valid to pool the two studies.

A 18 In Figure 5.6 Pg. 106 of the submission and figure 5.9, Pg. 109. Please clarify whether the vertical dashed lines represent clinic visit times.

The vertical dashed lines represent clinic visits at which a statistically significant difference in SRI response was observed between the belimumab 10 mg/kg group and placebo group.

A 19 Page 95 of the submission refers to “Fig X”. Please clarify which figure this text is referencing.

The text is referencing Figure 5.3 on page 95 of the submission document.

Section B: Clarification on cost-effectiveness data

Priority Questions

B 1 Please provide a table listing all the model assumptions (with justifications).

Assumptions	Justification/explanation
The difference between the SELENA-SLEDAI and SLEDAI instruments has no influence on the estimated efficacy and cost-effectiveness.	Both instruments contain the same items and weights with only a very slight change to definitions.
No impact of a difference between the pre and post 1996 version of SLICC.	There were only minor changes to the number of items included.
Adjusting the constant in natural disease model (JH).	There is a significant difference in baseline SS score between BLISS patients and JH cohort. To account for this the constant in the regression predicting disease activity over time was increased.
The absolute effect of belimumab on disease activity (SS score) remains constant after 1 year.	This assumption is supported by the Phase 2 study LBSL99 data where the benefit of belimumab was observed over 5 years of follow-up (see Section 6.3.1 of the submission document).
Disease flares were not directly included in the health economic model however AMS was incorporated and will account for disease flare activity to some degree although it is likely to have underestimated the predicted cost-effectiveness of belimumab.	The JH cohort database did not record data on disease activity flares so these data could not be modelled directly. Disease activity at time of organ damage is reflected in the individual system involvement covariates in the NHD models; these data would complement the AMS score by describing current disease activity and type of activity.
The probability of discontinuation remains constant over time (i.e. after 1st year) for belimumab responders.	Assumption supported by data from Phase 2 LBSL99 continuation study.
The SoC treatments used for the JH patients are similar to the SoC treatments used in the BLISS trials.	Assumed that best care has been given both to the JH cohort and the patients in the BLISS trials.
The JH natural history of disease (NHD) model can be applied to the BLISS population even thought the JH population may be less severe on average than the patients in the BLISS trials.	The NHD models are multivariate models that allow adjustment for differences in cohorts at baseline. A specific adjustment for difference in average disease activity was included.
The average exposure to the belimumab was assumed to be 100%.	As this technology is not self-administered, patients are under specialist care and in a considerably poor state of health, it seems

Assumptions	Justification/explanation
	reasonable that compliance will be high while the physician perceives that the patient is receiving benefit from continuing this treatment.
It is assumed in the base case that vial sharing between patients will not automatically occur.	As the number of patients with moderate to severe SLE is relatively small in UK, vial sharing may not be easy to manage in tertiary care units due to storage requirements. If vial sharing were to occur, its inclusion would only serve to improve cost effectiveness.

B 2 Setting the baseline distribution of ages to be a single value within the Subgroup BLISS Data worksheet by setting Q7:Q62=0 then setting one value within this array equal to Q64, appears to suggest that as the baseline age increases the ICER falls: e.g. for the Target Population age 30 the ICER is £65,498 per QALY while keeping everything else constant for the Target Population age 50 the ICER is £55,439 per QALY. Is this correct? If so this is the opposite of that what is often observed within economic modelling. Please provide a clarification and explanation of this result.

Yes this result is correct and the reason is related to mortality risk. Mortality risk increases over time, but is dependent on the adjusted mean SLEDAI (AMS) score, with higher AMS scores showing higher mortality risk. Belimumab is estimated to have a beneficial effect on survival.

This is best illustrated with an example. Table B2.1 below describes two patients, one aged 30 years and one aged 50 years, both of whom initiate and discontinue therapy with belimumab at the same time. As the average SS score decreases over time, the 50 year old patient will experience a greater mortality risk when 55 years of age (1.09%) compared with the 30 year old patient when 55 years of age (0.50%), due to the relationship with AMS. The lower the SS score (AMS) is, the lower the hazard ratio becomes. The hazard ratio is multiplied with the standardised mortality ratio, giving a greater mortality risk for the patient that initiated therapy at an older age. Belimumab has an estimated benefit on survival and therefore has a greater impact in older patients, as reducing a high mortality hazard is more cost-effective than reducing a low mortality hazard.

Table B2.1 Estimated effect on morality risk for two patients of different ages at start of belimumab treatment.

Baseline		
Age	30 yrs	50 yrs
SS	12	12
Discontinue therapy after	10 yrs	10 yrs
After 5 years		
Age(yrs)	35	55
SS	1.30	0.95
Mortality risk	0.0022	0.0110
After 25 years		
Age(yrs)	55	85
SS	3.78	3.50
Mortality risk	0.00505	0.0154

B 3 Please clarify what has been assumed for belimumab non-responders in terms of the evolution of their SS score and steroid dose. This might be most easily explained through a comparison of a hypothetical single High Disease Activity patient as a:

- belimumab non-responder, and
- SoC patient

and for each of these presenting:

- the common baseline characteristics,
- the calculation of the week 24 SS score and week 24 steroid dose, and
- the calculation of the week 52 SS score and week 52 steroid dose.

Belimumab non-responders are allocated to the standard of care (SoC) arm of the model at week 24 and will be assigned the same disease activity as a SoC patient. The steroid dose is calculated based on the Adjusted Mean SLEDAI (AMS). Steroid dose and SS score at week 24 are not considered in the model. The SS score and steroid dose calculation at week 52 is presented below.

<u>Hypothetical single high disease activity patient</u>

Start Age (years)	28.0
Gender	Female
Disease Duration (years)	6.0
Age At Diagnosis (years)	22.0
Black Ethnicity	No
Baseline SLICC	0.00

Start SLEDAI	10.0
Baseline steroid Dose (mg/day)	2.2

Linear regression explaining change in SLEDAI score after 52 weeks compared to baseline gives:

Variable in linear regression model	**Coefficient**
Baseline SS score (SS_0) for all SoC patients	-0.349
Baseline SS score (SS_0) for all belimumab patients	-0.343
Baseline SS score (SS_0) for all belimumab responder patients	-0.280

The JH steroid dose model used to calculate patients' mean steroid dose based on current year's average SLEDAI gives:

Average SLEDAI score	0.72
Constant	3.41

Responder

SS score after 52 weeks:
10 + 10 * (-0.343 - 0.280) = 3.77
AMS over first year = (10 + 3.77) / 2 = 6.885
Steroid dose after 1 year = 3.41 + 6.885 * 0.72 = 8.37 mg/day

Non responder (has discontinued at 24 weeks, now on SoC)

SS score after 52 weeks:
10 + 10 * -0.349 = 6.51
AMS over first year = (10 + 6.51) / 2 = 8.255
Steroid dose after 1 year = 3.41 + 8.255 * 0.72 = 9.35 mg/day

B 4 Please present the patient numbers (for both marketing authorisation population and Target Population) and the numbers continuing with treatment underlying *Table 6.42* (page 289 of the submission) at a greater degree of disaggregation, where responder status is defined by a change in the SS score at week 24 ≥ 4, N is the number of patients within the relevant category, and N Cont. is the number of patients continuing with belimumab treatment at week 24 within the relevant category.

The "probability of treatment continuation at week 24" is the same as the probability of response at week 24; only belimumab patients who show the defined level of response (i.e. SS score decrease of 4 or more at week 24) continue in the belimumab arm in the model. All other belimumab patients (non-responders) are switched to the standard of care (SoC) arm in the model after Week 24. Table 6.42 presented in our original submission document has been updated (see below) to

provide more information about the patient numbers used to derive the probabilities of response for each baseline SS score.

Table 6.42 Probabilities for treatment continuation (belimumab responders) at 24 weeks for different baseline SS scores – High disease activity subgroup (Target Population)

	BLISS (pooled)			BLISS-52			BLISS-76	
Baseline SS Score	Belimumab N	Responder N		Belimumab N	Responder N		Belimumab N	Responder N
0	0	0		0	0		0	0
1	0	0		0	0		0	0
2	0	0		0	0		0	0
3	0	0		0	0		0	0
4	0	0		0	0		0	0
5	0	0		0	0		0	0
6	0	0		0	0		0	0
7	0	0		0	0		0	0
8	0	0		0	0		0	0
9	0	0		0	0		0	0
10	75	45 (60%)		32	20 (63%)		43	25 (58%)
11	6	2 (33%)		5	1 (20%)		1	1 (100%)
12	48	35 (73%)		15	9 (60%)		33	26 (79%)
13	7	6 (86%)		5	4 (80%)		2	2 (100%)
14	19	16 (84%)		9	7 (78%)		10	9 (90%)
15	3	3 (100%)		2	2 (100%)		1	1 (100%)
16	16	10 (63%)		8	5 (63%)		8	5 (63%)
17	1	0		1	0		0	0
18	4	2 (50%)		1	0		3	2 (67%)
19	3	3 (100%)		0	0		3	3 (100%)
20	5	4 (80%)		1	1 (100%)		4	3 (75%)
21	0	0		0	0		0	0
22	4	3 (75%)		1	0		3	3 (100%)
23	1	1 (100%)		1	1 (100%)		0	0
24	0	0		0	0		0	0
25	0	0		0	0		0	0
26	0	0		0	0		0	0
27	0	0		0	0		0	0
28	0	0		0	0		0	0
29	0	0		0	0		0	0
30	1	0		0	0		1	0
Total	193	130 (67%)		81	50 (62%)		112	80 (71%)

B 5 Please confirm labelling of axes for figures 6.9 and 6.11 and present the equivalent of Figures 6.18 and 6.19 for the Target Population.

Within this please also separately present a split of the belimumab arm into those responding at week 24 and those not responding at week 24. It is recognised that this latter might require another two runs of the model.

The Y-axis for Figure 6.9 is correctly labelled as “SELENA-SLEDAI score”.

The title of Figure 6.9 should read “Figure 6.9. Mean SELENA-SLEDAI score from week 52 to week 76 – BLISS-76 study – **Total Population**” and not “Pooled Total Population”.

Please note the axis labels are missing for Figure 6.10. The X-axis is “Time (weeks)” and the Y-axis is “SELENA-SLEDAI score”.

For Figure 6.11 the X-axis is correctly labelled as age. This figure models the average SLE patient with an age of 30 years and average SS score. If time had been used for the X-axis it would have been 30=0yrs, 32=2yrs etc. The Y-axis is correctly labelled as “SLEDAI score”.

Please note the equivalent Figures 6.18 and 6.19 for the Target Population are presented in Figures 6.33 and 6.34 on page 290 of our submission document. These figures are provided here but with the belimumab arm split into responder and non-responders based on SS score at week 24 as requested. Please note that the pattern of SS score over time was very similar between patients on the SoC arm and the belimumab non-responder patients, hence the arms overlap on the figures, see Appendix 1 which presents the raw data used to create these figures and the spreadsheet “Question B5 Raw Data for Figures 6.33 and 6.34” which contains the data for these figures.

Figure 6.33. SELENA-SLEDAI score over time censored for death, with belimumab arm split into those responding and those not responding at 24 weeks – High disease activity subgroup (Target Population).

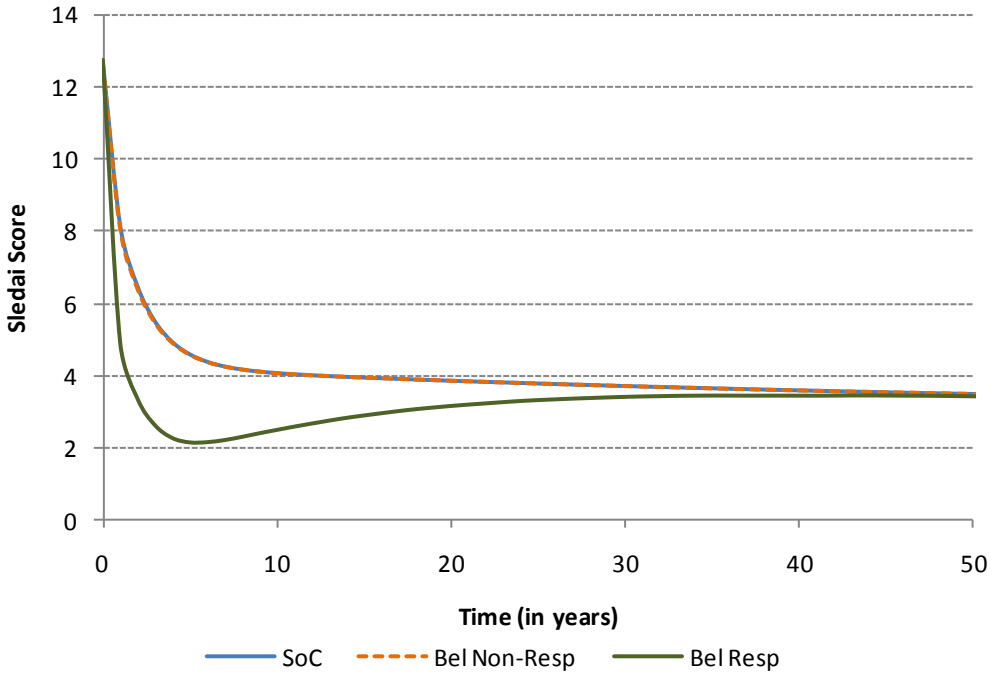
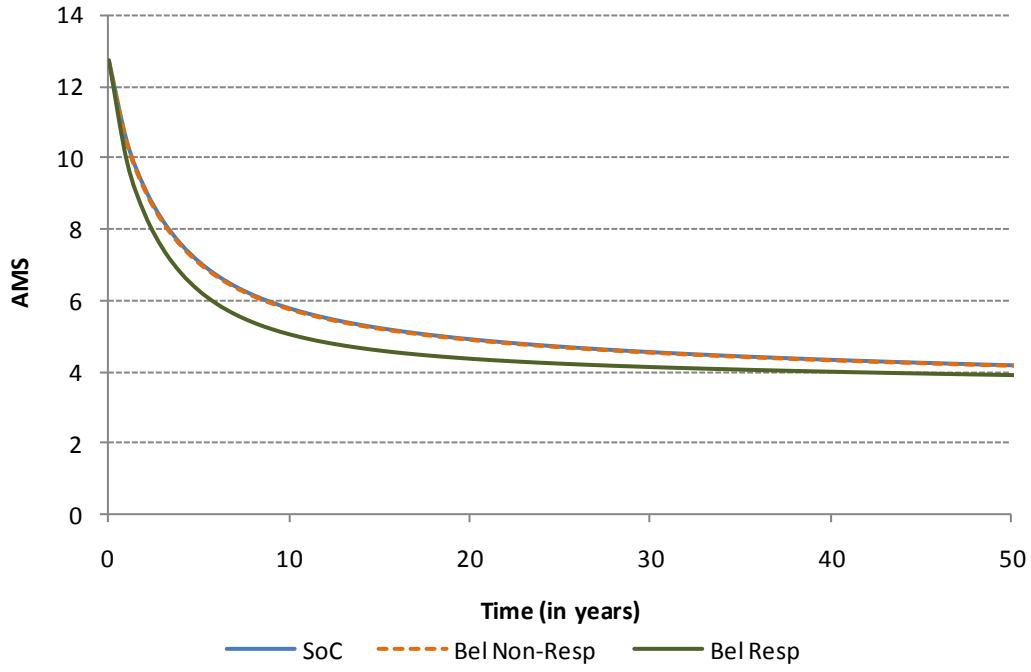


Figure 6.34. Adjusted Mean SLEDAI (AMS) over time censored for death, with belimumab arm split into those responding and those not responding at 24 weeks – High disease activity subgroup (Target Population).



Non-priority Questions

In light of the manufacturer strongly linking belimumab to the Target Population [n=203+193], the following clarification questions relate only to the Target Population of BLISS-52, BLISS-76 and pooled between these two trials. Depending on the response to question A1, please provide the data requested for the marketing authorisation population [n=287+305] patient populations of BLISS-52, BLISS-76 and pooled between these two trials.

B 6 Please clarify the extent to which the probabilities of treatment continuation at week 24 by SS score in *Table 6.42* differ from the probabilities of response at week 24. Are the values in *6.42* the addition of (A) non responders as defined by a change in SS score of less than 4; **PLUS** (B) responders who have discontinued by week 24?

Related to the above question, what role if any do the discontinuation rates for belimumab weeks 24 non-responders in year 1 of 37.4% and in subsequent years of 37.4% have upon the model structure?

The probabilities of treatment continuation (i.e. same as probability of response) at week 24 comprise all patients on belimumab who have a decrease in SS score of 4 or more at week 24 plus any patients who withdrew at or after week 20 and before week 24 and who have a decrease in SS score of 4 or more at week 20 i.e. for patients who withdrew after week 20, their week 20 value was carried forward to week 24.

For the base case where the responder rule (treatment continuation rule) is applied in the model the discontinuation rate for non-responders on belimumab is only used for the first six months. As these patients are switched to the SoC arm after 24 weeks and in this arm there is no discontinuation, no discontinuation probabilities are applied after 6 months for belimumab non-responders. However when the responder rule is ignored in the model the discontinuation probabilities are used for non-responders in Year 1 and subsequent years as these patients remain in the belimumab arm throughout the model horizon but will have a different rate to that seen for belimumab responders.

B 7 Much of the modelling uses 2006 or 2007 unit costs and uprates these with the CPI. Please clarify the reasoning for the use of the dated unit costs when more recent reference costs and PSSRU costs of health care are available online.

We acknowledge that this analysis did not use the most recent data. The regression analysis to obtain disease activity costs was performed earlier by a different health

economic agency using the Phase 2 belimumab study LSBL99 data and could not be easily updated with the most recent costs. However, these parameters were deemed non-critical to the outcomes of the model and this is demonstrated via univariate sensitivity analysis in which the results for the pooled total BLISS population change from £82,909 to £82,791 and £83,035 for outer limits of the distribution (2.5; 97.5).

B 8 Tornado diagrams: could the variables be explained / labelled more explicitly; for example the meaning of, and the distinction between "Treatment Effect Regression wk 52 SSO_Bel_R" and "Treatment Effect Regression wk 52 SSO_Bel" may not be immediately obvious to committee members. Please tabulate the values underlying the tornado diagrams of *Figures 6.37, 6.38 and 6.39* along the following lines:

Table B8.1: Univariate sensitivity analysis for effect on ICER – High disease activity subgroup (Target Population)

Variable	Base Value	Low Value	Low ICER	High Value	High ICER
Coefficient for belimumab responders from linear regression of change in SLEDAI score at 52 weeks	-0.280	-0.383	£49,393	-0.173	£103,840
Coefficient for **all** belimumab patients from linear regression of change in SLEDAI score at 52 weeks	-0.343	-0.437	£50,335	-0.251	£96,031
Adjusted Mean SLEDAI at current visit coefficient from the Mortality model	0.2135	0.085	£85,677	0.333	£50,962
Probability of remaining on belimumab over the first 364 days for belimumab responders	0.92	0.863	£54,518	0.981	£85,893
Coefficient of Log of age from the "clean utility" regression	-0.145	-0.180	£78,448	-0.103	£53,263
Constant coefficient in "clean utility" regression	1.297	1.146	£79,243	1.426	£55,493
Coefficient for all SoC patients from the linear regression of change in SLEDAI score at 52 weeks	-0.349	-0.394	£77,351	-0.307	£55,581
Adjusted Mean SLEDAI at current visit coefficient from the natural history Pulmonary model	0.1388	0.060	£73,044	0.216	£55,216
Constant coefficient from the natural history Neuropsychiatric model	-7.3961	-9.934	£61,333	-5.117	£76,231
Log of age at current visit coefficient in natural history Neuropsychiatric model	0.607	0.026	£61,514	1.226	£76,261
Adjusted Mean SLEDAI coefficient at current visit from the natural history Renal model	0.3234	0.228	£69,696	0.412	£56,744
Constant coefficient from the adjusted natural history disease activity model	3.0000	2.202	£73,226	3.934	£61,871
Constant coefficient from the natural history Peripheral Vascular model	-11.695	-16.475	£65,935	-6.806	£55,396
Log of age at current visit coefficient from the natural history Pulmonary model	1.2316	0.593	£70,841	1.916	£79,571
Constant coefficient from the natural history Renal model	-8.293	-9.010	£67,867	-7.560	£60,057

Table B8.2: Univariate sensitivity analysis for effect on incremental QALYs (Delta E) – High disease activity subgroup (Target Population)

Variable	Base Value	Low Value	Low QALY	High Value	High QALY
Coefficient for belimumab responders from linear regression of change in SLEDAI score at 52 weeks	-0.280	-0.383	1.030	-0.173	0.508
Probability of remaining on belimumab over 364 days for belimumab responders	0.92	0.863	0.649	0.981	1.165
Coefficient for **all** belimumab patients from linear regression of change in SLEDAI score at 52 weeks	-0.343	-0.437	1.014	-0.251	0.548
Adjusted Mean SLEDAI at current visit coefficient from the Mortality model	0.214	0.085	0.570	0.333	1.005
Coefficient of Log of age from the "clean utility" regression	-0.145	-0.180	0.662	-0.103	0.975
Constant coefficient in "clean utility" regression	1.297	1.146	0.655	1.426	0.936
Coefficient for all SoC patients from the linear regression of change in SLEDAI score at 52 weeks	-0.349	-0.394	0.678	-0.307	0.929
Constant coefficient from the natural history Peripheral Vascular model	-11.695	-16.475	0.775	-6.806	0.976
Constant coefficient from the adjusted natural history disease activity model	3.000	2.202	0.700	3.934	0.865
Constant coefficient from the natural history Neuropsychiatric model	-7.396	-9.934	0.844	-5.117	0.692
Log of age at current visit coefficient in natural history Neuropsychiatric model	0.607	0.026	0.842	1.226	0.698
Adjusted Mean SLEDAI coefficient at current visit from the natural history Renal model	0.323	0.228	0.756	0.412	0.891
Adjusted Mean SLEDAI at current visit coefficient from the natural history Pulmonary model	0.139	0.060	0.751	0.216	0.854

Table B8.3: Univariate sensitivity analysis for effect on incremental Costs (Delta C) – High disease activity subgroup (Target Population)

Variable	Base Value	Low Value	Low Cost	High Value	High Cost
Probability of remaining on belimumab over the first 364 days for belimumab responders	0.92	0.863	£35,386	0.981	£100,094
Adjusted Mean SLEDAI at current visit coefficient from the natural history Pulmonary model	0.1388	0.060	£54,840	0.216	£47,147
Log of age at current visit coefficient from the natural history Pulmonary model	1.2316	0.593	£54,194	1.916	£59,906
Constant coefficient from the natural history Pulmonary model	-9.265	-11.780	£54,265	-6.864	£59,554
Constant coefficient from the natural history Diabetes model	-14.656	-19.139	£51,525	-10.291	£55,610
Log of age at current visit coefficient in natural history Diabetes model	2.2481	1.162	£51,523	3.348	£55,251
Constant coefficient from the natural history Peripheral Vascular model	-11.695	-16.475	£51,120	-6.806	£54,076
Log of age at current visit coefficient from the natural history Peripheral Vascular model	1.161	0.431	£51,052	1.8926	£53,805
Renal damage at previous visit coefficient from the Mortality model	0.652	0.165	£50,600	1.189	£53,296
Adjusted Mean SLEDAI at current visit coefficient from the Mortality model	0.214	0.085	£48,876	0.333	£51,205
Constant coefficient from the adjusted natural history Disease Activity model	3.000	2.202	£51,261	3.934	£53,510
Adjusted Mean SLEDAI coefficient at current visit from the natural history Renal model	0.323	0.228	£52,718	0.412	£50,565
Age at diagnosis coefficient from the Mortality model	0.032	0.013	£52,912	0.050	£51,037
Adjusted Mean SLEDAI at current visit coefficient from the natural history Peripheral Vascular model	0.170	0.020	£51,389	0.313	£53,225
Coefficient for belimumab responders from linear regression of change in SLEDAI score at 52 weeks	-0.280	-0.383	£50,891	-0.173	£52,711

B 9 The user guide provided is relatively brief and provides limited background as to the model programming and structure within excel. If the manufacturer has been provided with a more detailed written account of the electronic model structure and VBA programming within the model, please could this be supplied.

We regret that we do not have a more detailed written description of the model functionality. However, Table B9.1 below contains a complete list of the items in the VBA code. If necessary a TC can be arranged with the health economic agency who produced the model if the ERG would like to discuss some of the details of the model functionality.

Table B9.1 Summary of VBA code used in the microsimulation model.

Modules	Subs / Functions	Short explanation
modDistributionsVBA	RandCat, RandCat2, RandWeib, RandNorm, RandPoisson, RandGeo, RandBern, RandExpon, RandGamma, RoundUp, RoundDown, Max, min	Contains the distributions used in the model.
modHazard	HExponential, HWeibull, HLogLog, HGompertz	Contain equations used to estimate the hazard.
modModelRun	Deterministic, PSA, Univariate, Simulate, Convergence, WriteBaseCaseResults, WritePSAResults, Tornado	Start of the model, either PSA or deterministic.
modScenarios	DifferentMaxDuration, AllScenarios, MultipleScenarios, MaxTimeScenarios,	Contain the different scenarios.
modTypes		Enumeration of variables, i.e. coding a string into a number.
Class Modules		
clsActivity	Init, Update, UpdSS, UpdSSFirstHalfYear, UpdSter, CheckConsistency	Used to calculate disease activity, for first and future year.
clsConvergence	Add	Used for creating convergence graphs.
clsCosts	AddShort, AddOrgan, AddMed, AddIndirect,	Used for cost calculations.
clsInputPars		Read in inputs.
clsOrgans	Init	Calculate SLICC damage.
clsOther	Init, Update, UpdateInfect	Other factors used for calculating organ damage and infections.
clsPatChars	Init, DrawAge, DrawGender, DrawDisDur, DrawEthn, Update	Draw certain patient characteristics.
clsPatient	Init, Simulate, MyFirstYear, UpdDeath, XBeta, Discontinuation, RecordBaseline, RecordEnd, Utilities, costs, CollectIndirect	Captures characteristics at baseline, during simulation and at end.
clsResults	AddChar, AddOrgan, AddSSInvolv, AddHistAv, AddSurvival, AddOrganKMStats, AddQALY	Write results to Excel.
clsStatusBarSLE		Used for status bar.

B 10 In terms of the implementation of the probabilistic modelling this seems to work down the 1,000 sets of clinical effectiveness estimates in the *PSA Inputs* worksheet by setting the active set of clinical effectiveness estimates in row 9 equal to one row of estimates below through the use of the INDEX function. Presumably for each set of clinical effectiveness estimates the 50,000 patient simulations are run to yield ***the*** central estimate for costs and for QALYs for that set of clinical effectiveness estimates to give ***one*** point on the cost effectiveness plane. 1,000 sets of clinical effectiveness estimates yield 1,000 points on the cost effectiveness plane, from which the CEAC is generated. Is this a correct interpretation of the generation of the CEAC within the modelling?

Yes, this interpretation is correct.

B 11 The *Treatment Effect* worksheet cells *Q12:R19* outline a net change at week 52 in SS items differentiated by SoC and Belimumab responders, as drawn from the *Subgroup BLISS data* worksheet which in turn references the *PSA Inputs* worksheet. Are these used as parameter inputs to the model? If they are, how are they derived?

Although the model calculates SS item involvement after 52 weeks, these are not used as the SS items are not included in the natural history disease models (JH – AMS forced in, involvement removed) chosen for the base case in the health economic model, see Section 6.3.1, Page 206 of the submission document.

The values are derived by calculating the difference between the percentage of patients with an SS item at baseline and 52 weeks for the placebo and belimumab arms from the BLISS trial data. This difference is then subtracted from the percentage simulated at baseline for combined SoC and belimumab arms.

B 12 In the electronic copy of the model worksheet *PSA Inputs* cells *BN7:BP7*. relate to the week 24 evolution of SS scores. Are these used within the modelling? If they are, please re-estimate this function separately for BLISS-52 and BLISS-76 Target Population patients.

The treatment effect regression at week 24 is not used in the model.

B 13 Within the *Subgroup BLISS data* worksheet there is a number of logically separate arrays of data. Which if any, of the following are superfluous to the current model implementation, assuming that only the responder rule of SS change ≥ 4 is of interest? (Superfluous in this context is not to say that the

data is not used elsewhere to estimate functional forms for the model, only that the running of the model does not directly draw on these data elements or the source of these data elements within the model if this is from referencing another worksheet within the model as applies to e.g. O217:Q219)

Group	From	To	Purpose ("none" if superfluous)
	O4	Q5	Number of patients in both arms (denominator in calculation of baseline SLICC damage)
1	O7	Q62	Baseline age distribution
	P63	Q64	Baseline gender
2	O66	Q71	Baseline Disease duration (P69 contains mean, other data is superfluous)
3	O73	Q83	Baseline ethnicity (B78 contains % black , other data is superfluous)
4	O85	Q183	Baseline weight (used for belimumab cost calculation)
5	O185	Q215	Baseline SS
6	O217	Q219	None
7	O221	Q223	Linear regression explaining change in SLEDAI score after 52 weeks compared to baseline
8	O225	Q231	Baseline steroid dose (P228 mean, P231 SE are used , other data is superfluous)
9	O233	Q235	None
10	O237	Q243	None
11	O245	T256	Baseline SLICC damage
12	O258	Q263	None
13	O266	Q274	None
14	O276	Q276	Discontinuation rate (Responders / Non-responder)
15	O277	Q296	None
	AQ4	AS5	None
16	AQ9	AS39	Probability of response based on baseline SS
	AQ216	AU219	None
	AQ220	AU223	Linear regression explaining change in SLEDAI score after 52 weeks compared to baseline
17	AQ233	AU235	None
18	AQ259	AS263	None
19	AQ265	AS274	None
20	AQ276	AS276	Discontinuation rate (Responders / Non-responder)
21	AQ277	AS296	None

Note: the rows without a group number are additional ranges not identified by ERG.

B 14 In the electronic copy of the model worksheet *PSA Data* of cells *BN7:CR7* what data within this is used within the modelling and what is superfluous to the current model implementation?

BN7:CR7 contains inputs for analysis without including the effect of a responder rule.

Label	From	To	Purpose ("none" if superfluous)
Treatment Effect Regression week 24	BN	BP	None
Treatment Effect Regression week 52	BR	BT	Linear regression explaining change in SLEDAI score after 52 weeks compared to baseline
Steroid Dose Regression	BV	BX	None
Natural Discontinuation Non-Responders	BZ	BZ	None
Wk 52 probabilities SS involvement	CB	CR	None

B 15 In the electronic copy of the model worksheet PSA Data of cells FV7:GR7 what data within this is used within the modelling and what is superfluous to the current model implementation?

Cells FV7:GR7 contains the subgroup inputs for the relevant subgroup.

Label	From	To	Purpose ("none" if superfluous)
Steroid Dose Regression	FV	FX	None
Natural Discontinuation Responders	FZ	FZ	Discontinuation for responding patients
Wk 52 probabilities SS involvement	GB	GR	None

B 16 Within the electronic model what do V1 and V2 refer to?

The electronic model uses V1 and V2 in "Subgroup BLISS data" rows 257 and 264. V1 refers to "Placebo", V2 refers to "belimumab 10 mg/kg". The corresponding table describes the percentage of patients with a SS item at baseline and week 52. The correct labelling of this table / data array should have been as detailed below:

		Placebo	Belimumab 10 mg/kg
Increased DNA binding	Baseline		
	52 weeks		
Low Complement	Baseline		
	52 weeks		
Vasculitis	Baseline		
	52 weeks		
NP involvement	Baseline		

		Placebo	Belimumab 10 mg/kg
	52 weeks		
Renal involvement	Baseline		
	52 weeks		
Serositis involvement	Baseline		
	52 weeks		
Haematological Involvement	Baseline		
	52 weeks		
Skin Involvement	Baseline		
	52 weeks		

B 17 Please clarify the observed distributions of SLEDAI for the Johns Hopkins cohort and for the Target Population. For example, please supply a diagram like that in Fig 6.6 for the Target Population and for the Johns Hopkins cohort at entry, and for Johns Hopkins cohort at last follow up.

Figure B17.1. Distribution of baseline scores in the pooled BLISS trials – High disease activity subgroup (Target Population)

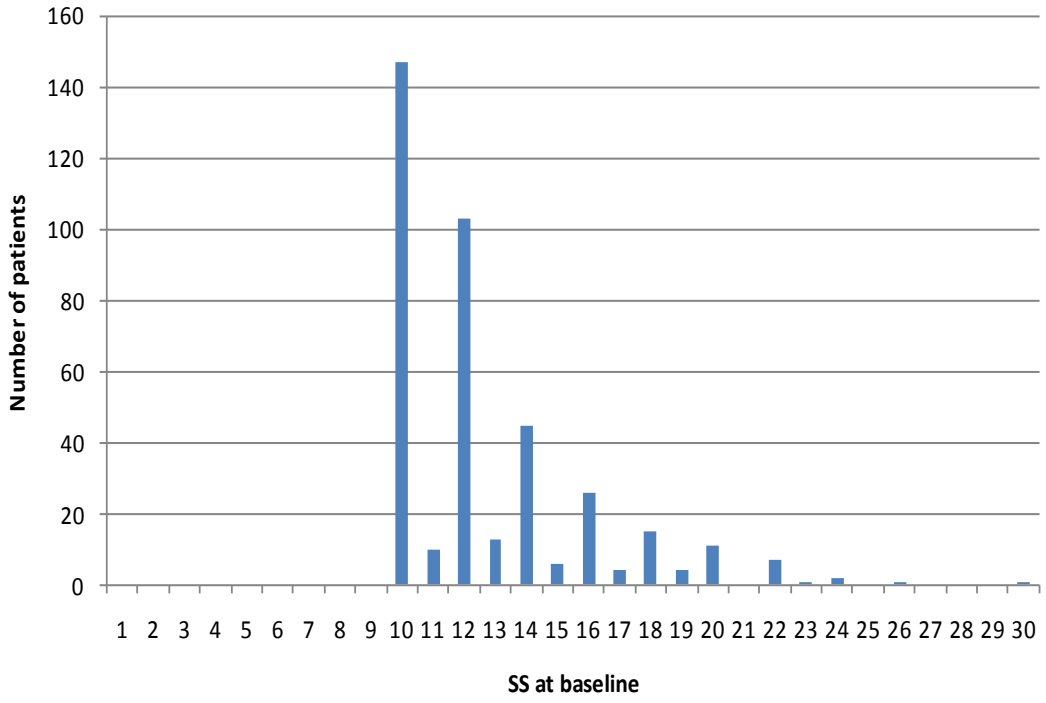
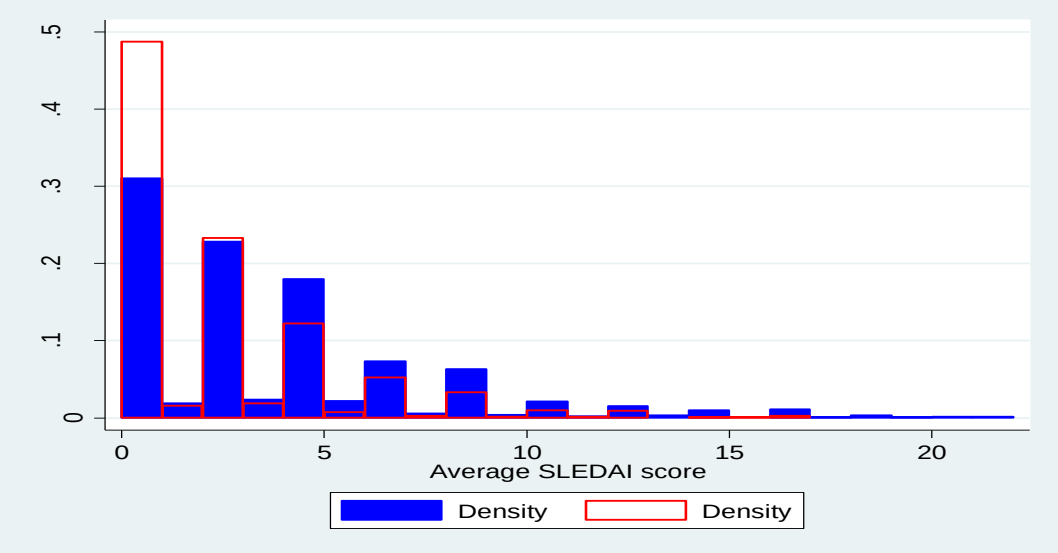


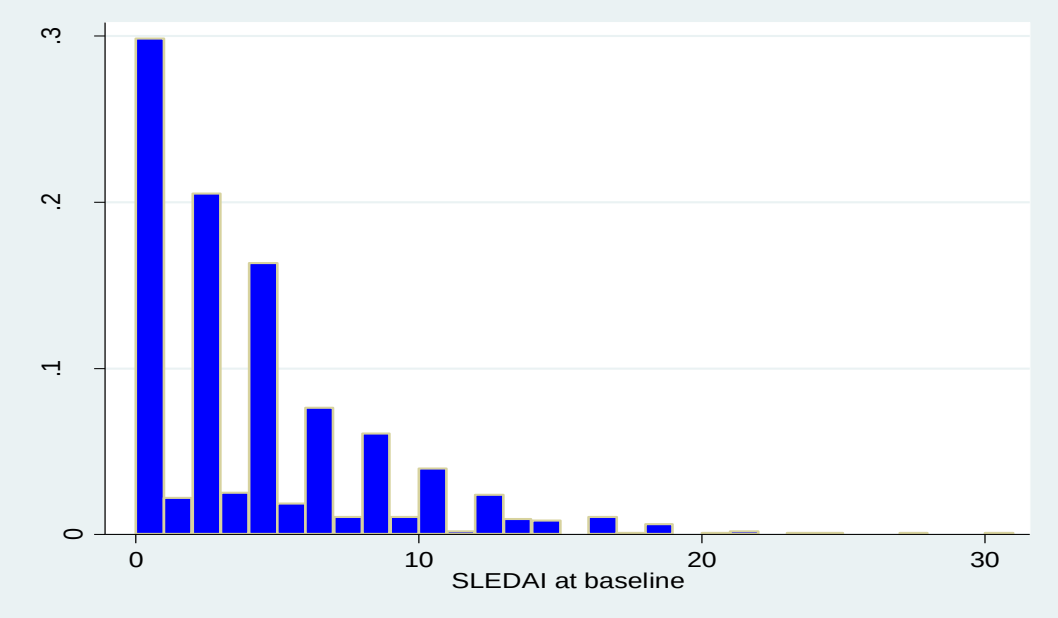
Figure B17.2. Distribution of Year 1 observation and last year observation SLEDAI scores for patients in the JH Cohort



Note: the Blue bars represent a histogram of the SLEDAI score at first visit. The Red outline bars represent a histogram of at the patient's last visit SLEDAI score in the cohort.

B 18 Please also provide the SLEDAI score distribution of excluded John Hopkins patients.

Figure B18.1 Distribution of the SLEDAI score at baseline of the patients excluded from the JH



The distribution of the SLEDAI score for the first visit for JH cohort patients excluded from the analysis is presented in Figure B18.1. This graph includes patients with observations collected before 1992 or whose duration of follow-up was less than 2 years. The mean SLEDAI score of patients excluded from the analysis at baseline was 3.79 (SD 4.25), slightly higher than the mean observed for the JH patients used to produce the natural history disease models (mean 3.32 (SD 3.7)).

B 19 The model (Appendix 21, Pg. 21) returns huge mortality risk up to 250% for haematological involvement and infection. Please clarify whether this reflects the instability of the model (small sample size in this group) rather than the real effect size.

It should be noted that the Weibull model for mortality risk chosen as the base case in the health economic model did not include the individual item involvements of the SLEDAI and therefore haematological involvement will not have influenced the simulated mortality risks in the base case.

However, we have investigated this question in three ways. First, we have generated two way tables illustrating the incidence of haematological involvement and infection in patients that died from the JH Cohort. Secondly, we present clinical evidence to support a strong association between haematological involvement and infection with mortality. Thirdly, we present the results of the mortality regression analyses with different model specifications to demonstrate the stability of the estimates.

1. A two-way table summarising the incidence of mortality and haematological involvement across all observations is reported below.

	No haematological involvement	**Haematological involvement**	**Total**
No death	35530	2656	38186
Death	69	13	82
Total	35599	2669	38268

The odds ratio = 2.52, indicating that patients with haematologic involvement are 2.5 times more likely to die than those without the involvement of this system.

A two-way table summarising the incidence of mortality and infection across all observations is reported below.

	No infection	**Infection**	**Total**
No death	36607	1579	38186
Death	74	8	82
Total	36681	1587	38268

The odds ratio = 2.51, indicating that patients with infections are 2.5 times more likely to die than those without infections.

2. The haematological involvement variable is a binary indicator of leucopenia and thrombocytopenia recorded in the SLEDAI. Thrombocytopenia has previously been associated with mortality in other longitudinal studies of SLE (Fernandez et al. 2007). It is possible that the effect of haematological involvement and mortality could have been inflated in these analyses if severe thrombocytopenia events are over-represented in this sample.

   Infection is common in SLE and has been identified as a causal factor for death in a large number of studies. The table below summarises the proportion of mortality events attributable to infection from these published studies. This suggests that the association between infection and mortality is supported in the literature.

Author name	**Date**	**Infection (%)**
Mok	2008	60.0
Nossent	2007	29.6
Bernatsky	2006	5.0
Doria	2006	23.0
Cervera	2003	25.0
Kasitanon	2002	51.9
Manger	2002	29.0
Alarcon	2001	32.4
Abu-Shakra	1995	32.0
Ward	1995	22.0

3. The table below reports the coefficients and p-values for the alternative statistical models for mortality. The results illustrate the stability of the estimates for haematological involvement and infection with alternative model specifications.

	Weibull	**Exponential distribution**	**Use SLICC index**	**Remove prednisone**
Black	0.806 (0.003)	0.924 (0.001)	0.587 (0.018)	0.905 (0.001)
Age at diagnosis	0.032 (0.000)	0.023 (0.009)	0.030 (0.001)	0.028 (0.002)

	Weibull	Exponential distribution	Use SLICC index	Remove prednisone
Cholesterol	0.005 (0.003)	0.005 (0.009)	0.005 (0.003)	0.006 (0.001)
AMS	0.209 (0.000)	0.156 (0.005)	0.234 (0.000)	0.237 (0.000)
Prednisone	0.001 (0.066)	0.003 (0.485)	0.001 (0.207)	
Haematological involvement	1.110 (0.000)	1.022 (0.001)	1.025 (0.001)	1.185 (0.000)
SLICC/ACR			0.294 (0.000)	
Renal damage	0.672 (0.009)	0.717 (0.006)		0.719 (0.005)
Musculoskeletal damage	0.414 (0.002)	0.537 (0.000)		0.453 (0.000)
Peripheral vascular damage	1.027 (0.000)	1.180 (0.000)		1.066 (0.000)
Gastrointestinal damage	0.506 (0.042)	0.530 (0.000)		0.519 (0.039)
Diabetes	0.635 (0.039)	0.739 (0.016)		0.614 (0.048)
Malignancy	1.090 (0.000)	1.252 (0.000)		1.079 (0.000)
Infection	0.783 (0.050)	0.882 (0.027)	0.745 (0.062)	0.817(0.040)
Constant	-10.656 (0.000)	-8.468 (0.000)	-10..43 (0.000)	-10.373 (0.000)
Weibull parameter	1.700 (0.000)		1.66 (0.000)	1.604 (0.000)

B 20 Appendix 21, 4.4. Time to event Analysis, the discussion section (Pg. 24 paragraph 3): The section states, “The clinical interpretation of a relationship between cardiovascular and cerebrovascular damage and mortality is stronger than that for musculoskeletal damage.” Please clarify whether this claim is made following a causal analysis?

This claim was not made following a causal analysis. The statement simply refers to an observation that cardiovascular and neuropsychiatric damage on the SLICC/ACR include events such as myocardial infarctions and cerebrovascular accident and are associated with a high mortality risk. Therefore, it is surprising that cardiovascular damage and neuropsychiatric damage were not statistically significant predictors of mortality in these analyses.

B 21 Table 6.7 (page 198 of the submission) indicates a high mean steroid use in the John Hopkins population despite relatively low mean SLEDAI score; might this indicate different use of steroid in SoC in the 1990s relative to the present decade? Please clarify.

The table below reports summary statistics relating the prednisone dose observed before and after 2000. The post 2000 group only includes patients entering the cohort after 2000 to reduce the confounding effects of length of follow-up. The mean prednisone dose at first visit in the cohort is not different for patients entering the cohort before 2000 and those entering the cohort after 2000. However, the overall observed prednisone in each time period suggests that prednisone doses may have decreased over time.

	Mean	Standard deviation	Median
Pre-2000			
Mean prednisone dose at first visit (SD)	10.01	14.4	5
Mean prednisone dose at all visits (SD)	9.24	10.77	5
Post 2000 (including only patients entering the cohort after 2000)			
Mean prednisone dose at first visit (SD)	9.86	16.21	5
Mean prednisone dose at all visits (SD)	5.62	10.53	0

The prednisone predictive model has been re-run with a dummy variable indicating observations made before Jan 2000. The variable is statistically significant and suggests that patients were prescribed higher doses of steroids before 2000. However, adjusting for date does not substantially impact on the coefficient for SLEDAI or the constant term.

	Original model	Revised model
SLEDAI score (p-value)	0.7199 (0.000)	0.6799 (0.000)
Pre-2000 observation (p-value)		1.433 (0.000)
Constant (p-value)	3.410 (0.000)	3.197 (0.000)

B 22 Please clarify the rationale for the choice of sensitivity analyses undertaken (e.g. as illustrated on page 269 of the submission).

The model includes a very large number of parameters so for the sensitivity analyses the parameters identified as likely to have an impact on the estimate of cost-effectiveness were investigated. The most critical parameters were identified as follows:

- Coefficients of week 52 change in SS score regressions
- Coefficients of natural history model for change in SS score
- Natural discontinuation
- Coefficients of natural history models for mortality and organ damage development
- Coefficients for BLISS utility regression
- Costs associated with each SS score
- Organ Damage Costs

- Organ Damage Disutility

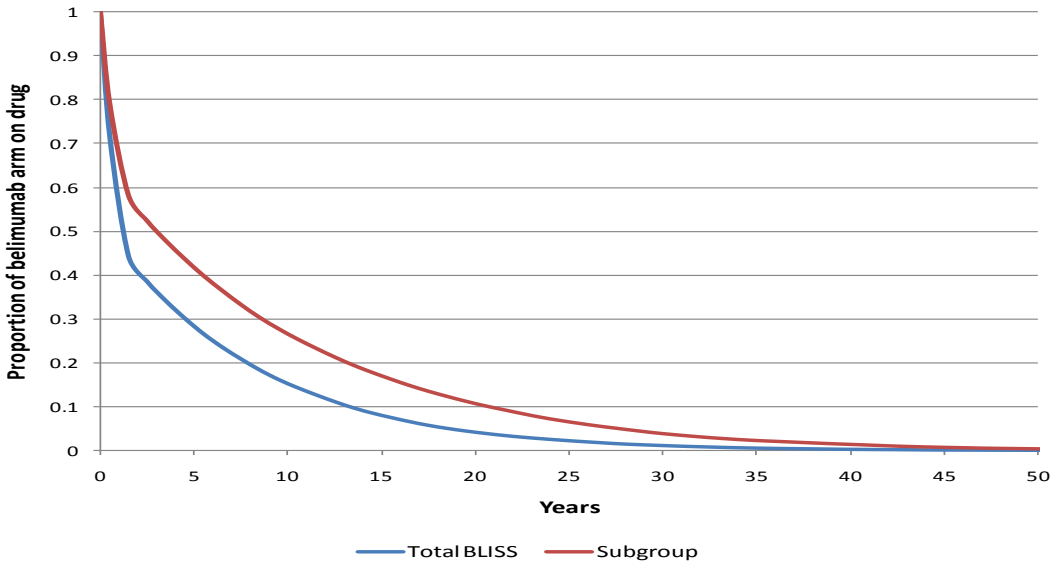
B 23 On page 236 of the submission the calculated example provides a utility of 0.9719 for ocular organ damage for year one. Does this mean that ocular damage experienced in year 1 by patient A (in table 6.19) incurs a disutility of 1-0.9719 for that patient resulting in a utility for patient A of 0.63 – 0.0281 = 0.6019 ?

No, utility multipliers were used, therefore the utility of patient A would become: 0.63*0.9719= 0.61

B 24 Fig 6.17 on page 259 of the submission represents the proportion of the "total population" remaining on belimumab through time and similarly Fig 6.35 (Page 291) for the Target Population. Please compare these on a single graph and clarify the reason(s) for the difference. Please comment on the fact that patients getting more benefit appear more likely to discontinue.

Please see Figure B24.1 below. The initial drop in the first year is caused by the proportion of responders. In the total population 52.4% patients respond, whereas in the subgroup, this percentage is 66.8%. The gradual decline over time is caused by natural discontinuation which is 6.1% and 4.4% in the first year and 10.9% and 8.0% in subsequent years, for the Total BLISS Population and subgroup, respectively. Mortality is the third factor contributing to termination of belimumab treatment. Mortality is slightly higher in the subgroup, caused by higher average disease activity, thereby decreasing the differences between the two groups over time.

Figure B24.1 Summary of belimumab patients remaining on treatment over time for both the Total BLISS Population and the subgroup (Target Population)



B 25 Appendix 23: The Figures are stated to be K-M plots, but they look like parametric fits. Please clarify.

The figures present Kaplan-Meier plots of the model outcome. Events, censored for death, were shown at yearly time points. As these plots were based on simulations with 50,000 patients, individual events or censoring can no longer be identified. Due to the number of patients the graph has a smoothed appearance instead of the regular 'staircase' appearance.

B 26 Footnote to Table 6.37 Pg. 285 refers incorrectly to Table 4 (which is on Pg 28). Please clarify.

The footnote should refer to Table 6.3 on page 190 of the submission document and not Table 4.

Additional Question from ERG provided to GSK on 24th May 2011

In discussions with their clinical advisor, the ERG has been alerted to "large errors" in two of the parameters that feed into the cost effectiveness model.

Referring to the two tables below from the submission:

Table 6.43: Organ damage occurrence (SLE until death)

	SoC	Belimumab	Difference
Pulmonary	39.9%	36.8%	-3.1%

Table 6.46: Discounted costs over life time.

Organ damage costs	SoC	Belimumab	Difference	Absolute difference	% absolute difference
Pulmonary	£42,692	£39,652	-£3,040	£3,040	4.9%

The clinician's experience is that pulmonary complications are not as common as implied in Table 6.43, especially if the main contributor is considered to be pulmonary hypertension; the clinician estimated an occurrence at near **1% only.**

In the clinician's opinion the cost of treatment of pulmonary complications is not as great as implied in Table 6.46 where it represents the biggest contributor to the difference between belimumab arm and SoC arm. The clinician questioned if the high cost is associated with the proposed use of expensive pulmonary hypertension drugs (e.g. bosentan, ambrisentan), which the clinician would very rarely use, opting for the much cheaper sildenafil. Also the clinician suggested if the high cost is contributed by the cost of transplant then this would be an almost vanishingly rare treatment.

GSK Response:

The prevalence of pulmonary damage observed in the Johns Hopkins Lupus cohort is substantially higher than that reported by the clinician and we are unable to account for the prevalence reported by the clinician. Currently the National Commissioning Guideline for Pulmonary Arterial Hypertension (PAH) have designated only 6 centres nationally to provide PAH services for adults and only these 6 centres can initiate reimbursed treatment (NHS National Specialised Commissioning Group 2009). The table below details the number of events recorded during the observation period of the analysis and the distribution of events among the SLICC/ACR damage index items within the JH cohort. A UK study of organ damage found that after 10 years of follow-up 3.4% of patients had pulmonary damage (Chambers et al. 2009), however a US study of the LUMINA SLE cohort (Bertoli et al. 2007) showed a cumulative rate of 11.6% over 10 years of follow-up. In contrast, in the JH cohort 16.6% of patients had recorded pulmonary damage over this time period. This demonstrates a significant amount of variability in the prevalence of pulmonary damage which could be due to several factors such as differences in disease severity or quality of recording of events, for example.

Breakdown of pulmonary damage events in the JH cohort

	Number of patients with pulmonary damage	**% of population with pulmonary damage**	**Mean SLICC score per patient with pulmonary damage**
Pulmonary damage at baseline	76	6%	1.12
Pulmonary damage at last visit	185	14%	1.31
	Events in Observation period*	**Event occurred before baseline observation**	**Proportion of total Pulmonary events**
Pulmonary hypertension	53	26	33%
Pulmonary fibrosis	65	36	42%
Shrinking lung	3	3	2%
Pleural fibrosis	38	10	20%
Pulmonary infarction	1	8	4%
*Events do not sum to total number of pulmonary damage because some events occur at the same time			

The costs for damage in ‘Pulmonary’ consists of 43% Pulmonary Hypertension. The weight was derived from 79 Pulmonary Arterial Hypertension (PAH) events (26 before baseline and 53 during the observation period) in 185 patients with pulmonary damage.

The yearly costs for a PAH patient was estimated based on two reports:

- A report by Condliffe *et al* was used to determine the treatment pattern of PAH patients (Condliffe et al. 2009). This report studied a (UK) national registry of patients with connective tissue disease associated PAH between 2001 and 2006. Of this cohort 8% had SLE and 76% had Systemic sclerosis (SSc). The treatment pattern of patients with SSc-PAH is described in this study and was adopted for our analysis (see table below).
- A NICE Health Technology Assessment report (Chen et al. 2009) which included treatment of PAH with bosentan was used for unit costs and other resource use.

Type of Treatment for PAH	**Usage (%)**
Monotherapy	62%
Endothelin receptor agonist	68%[1]
Prostanoid	17%[1]
PDE-5 inhibitor	15%[1]
Combination therapy	28%
No adv. Therapy	10%
Total	100%

[1]As percentage of all monotherapies

The costs in the table below were derived from the NICE HTA report (Chen et al. 2009):

Type of treatment	**Drug**	**Cost / 4 weeks**
Endothelin receptor agonist (ERA)	Bosentan	£ 1,540.00
Prostanoid	Iloprost	£ 2,773.40
PDE-5 inhibitor	Sildenafil	£ 348.60
Combination therapy (=average cost of the three monotherapies)		£ 1,570.97

It is possible that the treatment pattern has changed since the time period of the first study (2001-2006), but due to limited published data on this we are unable to identify anything more contemporary**.** However, the current National Commissioning Guideline for PAH (2008) (NHS National Specialised Commissioning Group 2009) recommends bosentan as first line therapy for PAH associated with Connective Tissue Disease (CTD), and recommends sildenafil as second line therapy or part of combination therapy. This is consistent with the treatment pattern shown above from Condliffe *et al*.

We ran two additional sensitivity analyses to investigate the impact of:

1) using the costs for sildenafil for all patients receiving treatment for PAH.
2) excluding all costs for pulmonary damage.

Both of these scenarios can be considered extreme cases

as it is highly unlikely that all patients would be treated in this way given current UK practice.

The results from these two additional analyses are provided in the table below and demonstrate only a fairly small increase in the ICERs compared with the base case ICER for the Target Population of £64,410.

Sensitivity analysis	Incremental cost Belimumab	Incremental LYs Belimumab	Incremental QALYs Belimumab	Incremental Cost per QALY Belimumab
All PAH treatment costs set to the price of sildenafil	£53,857	1.05	0.806	£66,807
All costs of pulmonary damage excluded	£54,966	1.05	0.806	£68,182

Additional Clarifications

1) Please note that the clean utility equation on page 236 of the submission should read:

2) The section in the health economic model on the treatment effect sheet cells J26 to P35 should be labelled as follows.

Belimumab continuation and discontinuation parameters. Used to determine whether or not Belimumab patient is a responder and if a patient continues with Belimumab treatment

Responders	52.4%	52.4%
Natural discontinuation	Resp	No Resp
Daily hazard of remaining on therapy	0.00032	0.00066
Year 1 probability of discontinuation (days 168 to 365)	6.1%	21.4%
Subsequent years probability of discontinuation	10.9%	21.4%

Section C: Clarification on other issues

Priority Questions

C 1 Figure 6.7 on page 201 of the submission shows the adjustment to the Johns Hopkins model (dashed red line) which involves raising the constant from 2.058 to 3.0; it is stated that "*A range of numbers were analysed to derive the adjusted constant, with a value of 3.0 providing a reasonable fit to these data".* How was the fit tested and does this refer to a fit of all the data shown in *Fig 6.10* or to weeks 52 to 250 or other? Please outline what other values were considered.

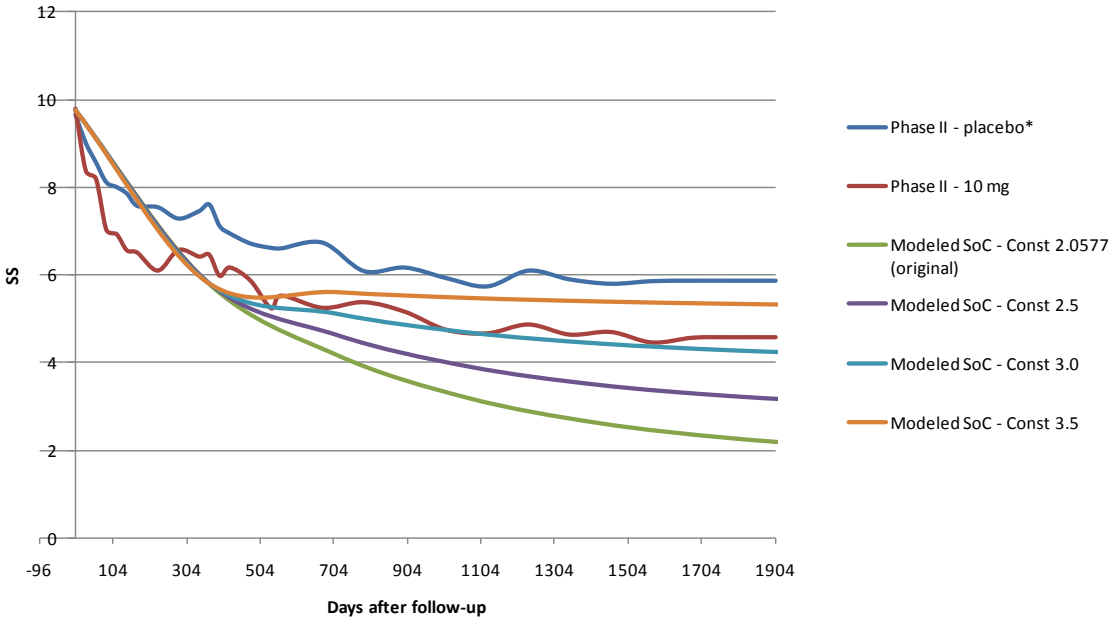
Related to the above, was the assumed value of 3.0 for the intercept of the regression analysis of the long term change in SS score retained for the Target Population?

A range of different constants were tested, from the original constant value of 2.0577 to a value of 3.5 to summarise SS score over time in the modelled Standard of Care (SoC) arm. The fit was checked visually compared to the SS score data from the Phase 2 trial (ANA positive patients, 5 years follow-up). No formal statistical test was used to determine the best fit. The constant of 3.0 was selected since it provided a good visual fit with the Phase 2 belimumab 10mg SS score data. However it may be a conservative estimate, as the constant could have been selected based on fit compared with the Phase 2 placebo arm, in which case a constant of 3.5 could have been justified. The ICER drops as the constant is increased. With a lower constant,

the SS score of patients will reduce rapidly, limiting the potential benefit from belimumab.

Life-time ICER for the Total BLISS Population when applying different constants.

Constant	ICER
2.0577	£93,654
2.5	£85,394
3.0	£82,909
3.5	£80,988



* Note that the placebo arm in the Phase 2 trial received belimumab after 532 days.

The value of 3.0 for the intercept of the regression analysis of the long term change in SS score was retained for the high disease activity subgroup (Target Population). The result of this will be to provide a conservative result due to the greater severity of the subgroup. As the constant adjustment itself, may be considered a little arbitrary and the subgroup is already a selection of patients within the BLISS trials we decided not to add a possible further benefit to this subgroup, particularly since there is no long-term follow-up data for patients with a high level of disease activity.

C 2 The regression analysis for the average steroid dose related to the SS/AMS score as estimated from the John Hopkins cohort is accepted uncritically within the modelling. This is despite the arguments around the unrepresentativeness of the John Hopkins cohort for the BLISS trials in estimating the regression analysis of the long term change in SS score. To what extent were similar considerations around steroid use explored; e.g. validation through varying the constant and aligning with BLISS baseline steroid use and SS scores?

This certainly could have been investigated, however we did not specifically look further into this because of the relatively small difference in steroid dosage used between the baseline dose of the BLISS trials and the JH cohort.

Table C2.1 Comparison of Steroid dose at baseline between the pooled BLISS study patients and the JH Cohort

	Steroid dosage (mg/day)		
	At first visit	**At 52 weeks**	**All observations**
BLISS (52 and 76 pooled)	10.78	8.74	10.78
JH	9.95	N/A	6.67

C 3 For the patient access scheme, please supply a correct version of Figure F5 (Pg. 18, Appendix 29).

The correct figure for the high disease activity subgroup is provided below.

Figure F5. Acceptability curve of PSA - high disease activity subgroup (Target Population)

Figure removed as shows commercial in confidence data.

Non-Priority Questions

C 4 Pg 157 states: *"In the long-term open-label extension of the Phase 2 trial (LBSL99), the incidence of AEs, severe AEs, SAEs, including infections, remained stable or declined over time through 5 years of exposure".* No data appears to have been presented to support this statement. Please clarify the source of the data to support this statement, providing further results if required.

Please see Table 1 below which summarises adverse event incidence from the Phase 2 trial LBSL99 to support this statement (Chatham et al. 2010).

Table 1: AE incidence (rate per 100 pt-yrs)[a]

	All patients treated with belimumab									
Interval Years	1 (0-0.5 yr)	2 (0.5-1 yr)	3 (1-1.5 yrs)	4 (1.5-2 yrs)	5 (2-2.5 yrs)	6 (2.5-3 yrs)	7 (3-3.5 yrs)	8 (3.5-4 yrs)	9 (4-4.5 yrs)	10 (4.5-5 yrs)
No. patients [pt-yrs]	424 [206]	398 [183]	353 [166]	314 [147]	284 [136]	261 [128]	252 [122]	240 [116]	227 [105]	188 [85]
Overall AEs	400 (194)	337 (184)	304 (184)	271 (184)	244 (179)	226 (177)	209 (171)	204 (176)	163 (156)	131 (155)
Serious AEs	43 (20.8)	33 (18.0)	28 (16.9)	25 (17.0)	24 (17.8)	26 (20.4)	15 (12.2)	19 (16.4)	16 (15.3)	12 (14.2)
Overall infections	246 (119)	197 (108)	170 (103)	165 (112)	145 (107)	130 (102)	108 (88)	121 (104)	81 (78)	67 (79)
Serious infections	14 (6.8)	10 (5.5)	6 (3.6)	8 (5.4)	4 (2.9)	5 (3.9)	2 (1.6)	6 (5.2)	1 (1.0)	3 (3.5)
Malignancies	0 (0)	1 (0.5)	3 (1.8)	2 (1.4)	0 (0)	1 (0.8)	1 (0.9)	3 (2.6)	1 (1.0)	2 (2.4)

[a] Interval 1 includes the placebo patients who initiated belimumab treatment at wk 56

C 5 What other forms for the regression analysis for the average steroid dose related to the SS/AMS score as estimated from the John Hopkins cohort were explored: e.g. change in steroid use being dependent upon change in SS/AMS score? What were the results of these analyses?

An alternative form for the steroid dose model has been generated and the results are detailed below. As the cost-effectiveness model is not sensitive to steroid dose in the univariate analyses, we would not expect that a different regression model would have a significant impact on the results.

	Baseline model	**Model including PGA**
Dependent variable	Coefficient for change in steroid	Coefficient for change in steroid
Change in SLEDAI score (p-value)	0.3582 (0.000)	0.1043 (0.050)
Change in Physician Global Assessment (PGA)	-	2.2596 (0.000)
Constant (p-value)	-0.3858 (0.000)	-0.3833 (0.000)

C 6 Please present estimates of the linear regression of the Target Population of table 6.41 (page 288 of the submission) separately for BLISS-52 and BLISS-76.

Tables 6.41a and 6.41b below summarise the regression models for change in SS score at week 52 for BLISS-52 and BLISS-76 respectively.

Table 6.41a Linear regression explaining change in SELENA-SLEDAI (SS) score at week 52 for BLISS 52 – High disease activity subgroup (Target Population)

Parameter	**Estimate**	**SE**	**p-value**
SS_0 SoC	-0.3629	0.0281	<0.001
SS_0 all belimumab	-0.3746	0.0619	<0.001
SS_0 belimumab responders	-0.2626	0.0680	<0.001

Table 6.41b Linear regression explaining change in SELENA-SLEDAI (SS) score at week 52 for BLISS 76 – High disease activity subgroup (Target Population)

Parameter	**Estimate**	**SE**	**p-value**
SS_0 SoC	-0.3341	0.0349	<0.001
SS_0 all belimumab	-0.3153	0.0688	<0.001
SS_0 belimumab responders	-0.2827	0.0821	<0.001

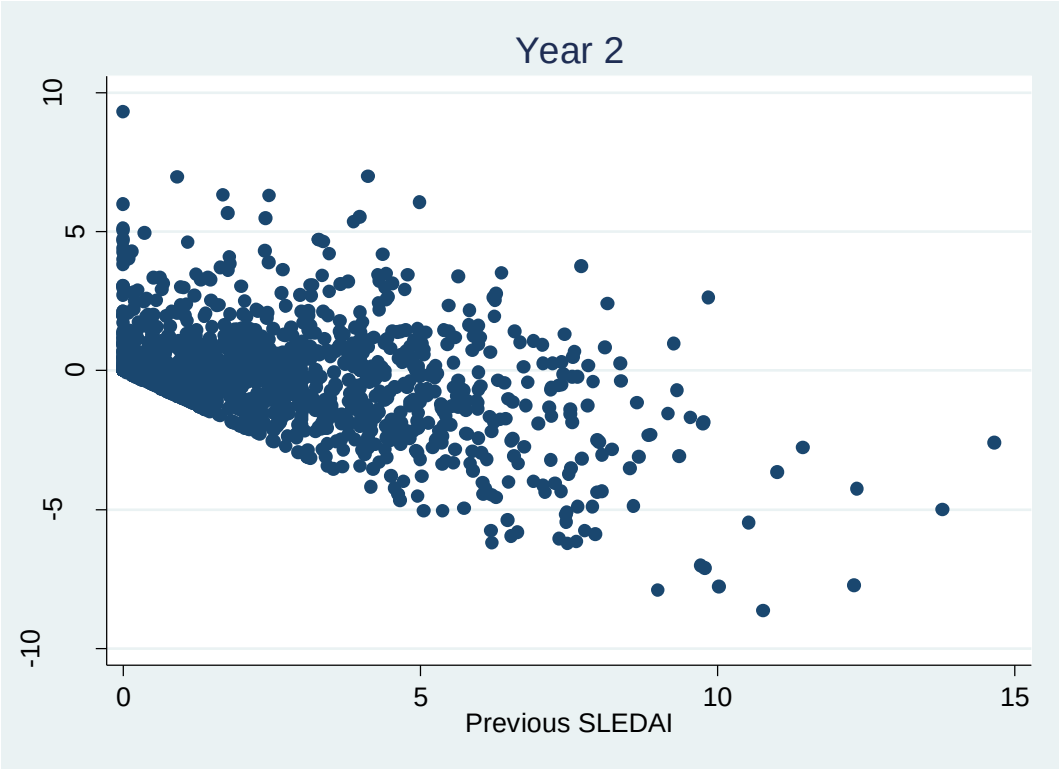
C 7 Random effects model(s) (Appendix 21). Please justify the use of previous mean SLEDAI score; please clarify whether this might be affected by regression to mean?

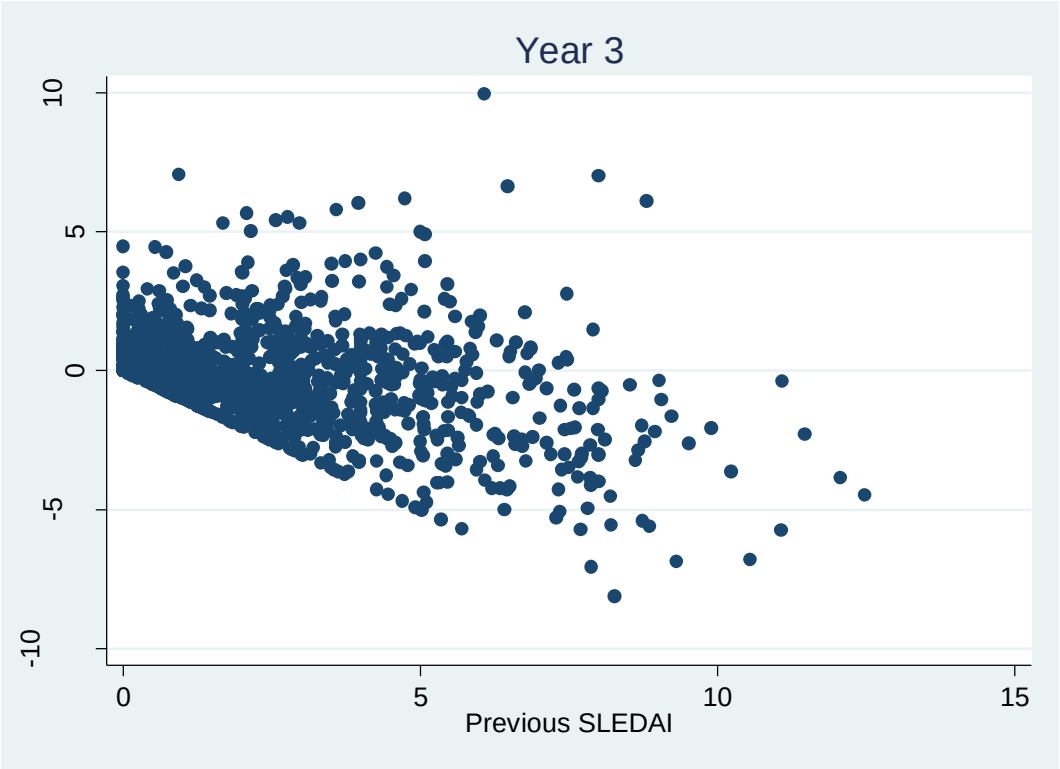
Previous SLEDAI score is negatively correlated with a decrease in SLEDAI score (r= -0.40). Patients with high SLEDAI scores are likely to experience large reductions in SLEDAI because the most severe items on the SLEDAI are assigned larger weights. Therefore, it was considered appropriate to include previous SLEDAI as a variable in the model to describe how patients will experience a larger reduction in SLEDAI if they have a previously high score.

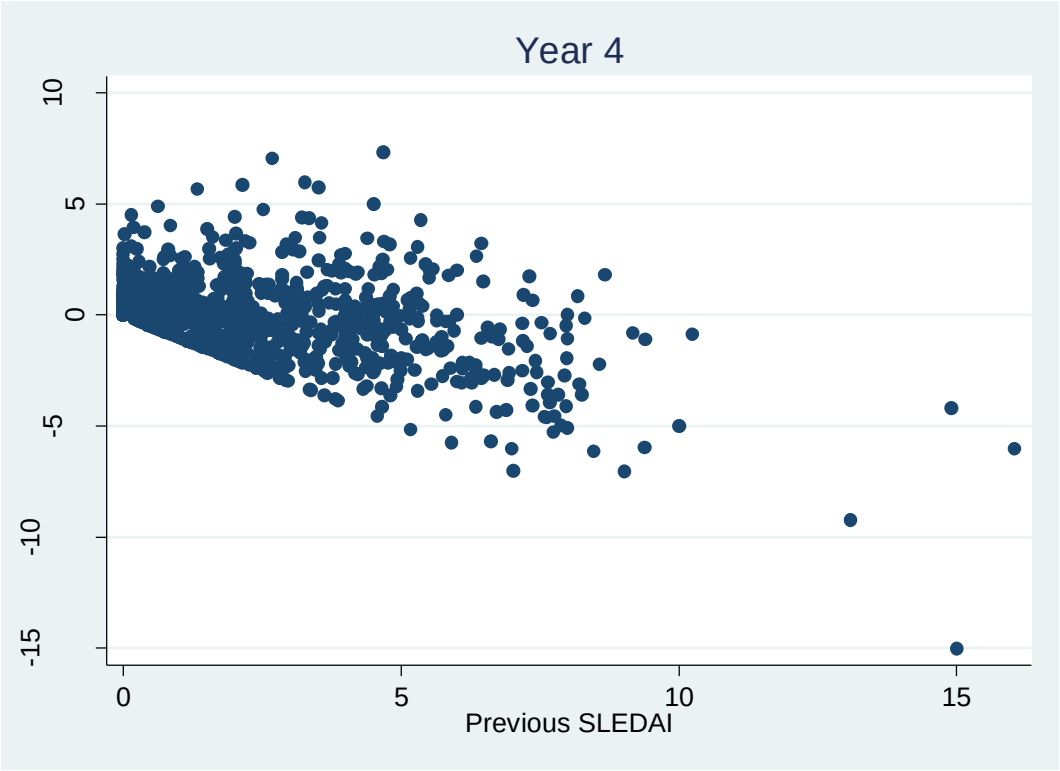
Regression to the mean may occur because patients are more likely to enter the JH cohort when they are experiencing a peak in their disease activity, which motivates their admittance to see a specialist lupus clinician. However, in the disease activity model observed SLEDAI scores are averaged to estimate the annual disease activity score for each annual interval. This process of smoothing the data will reduce the impact of regression to the mean. The model has been re-estimated to exclude the change in SLEDAI score observed from year 1 to 2. This exclusion does not substantially affect the coefficient estimates reported in a table below.

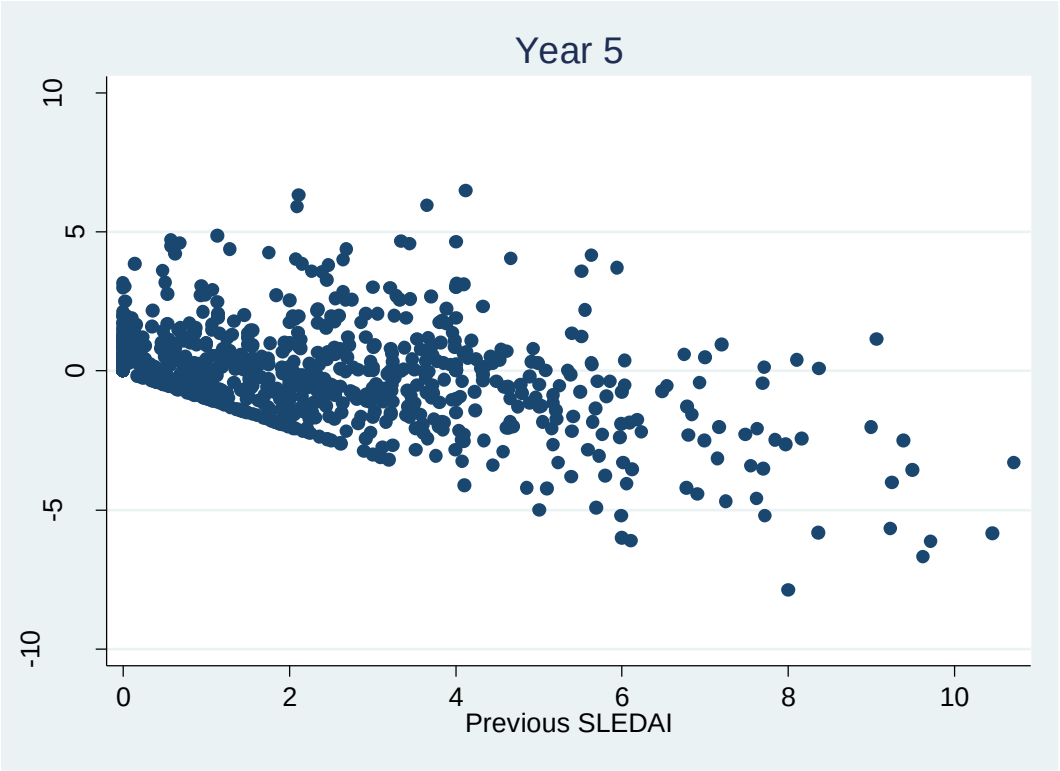
	Baseline Model	**Revised Model excluding the first year in the cohort**
SLEDAI score in previous year	-0.4163	-0.4682
Male gender	-0.0991	-0.1205
Black ethnicity	0.3524	0.3722
Log of age	-0.3586	-0.4087
Constant	2.0577	2.347

Scatter plots for the relationship between change in SLEDAI and previous SLEDAI score are reported by year of 5 years of follow-up. These show that the relationship between previous SLEDAI and change in SLEDAI is stable over time in the cohort.


Year 2
10
5
0
-5
-10
0
5
10
15
Previous SLEDAI

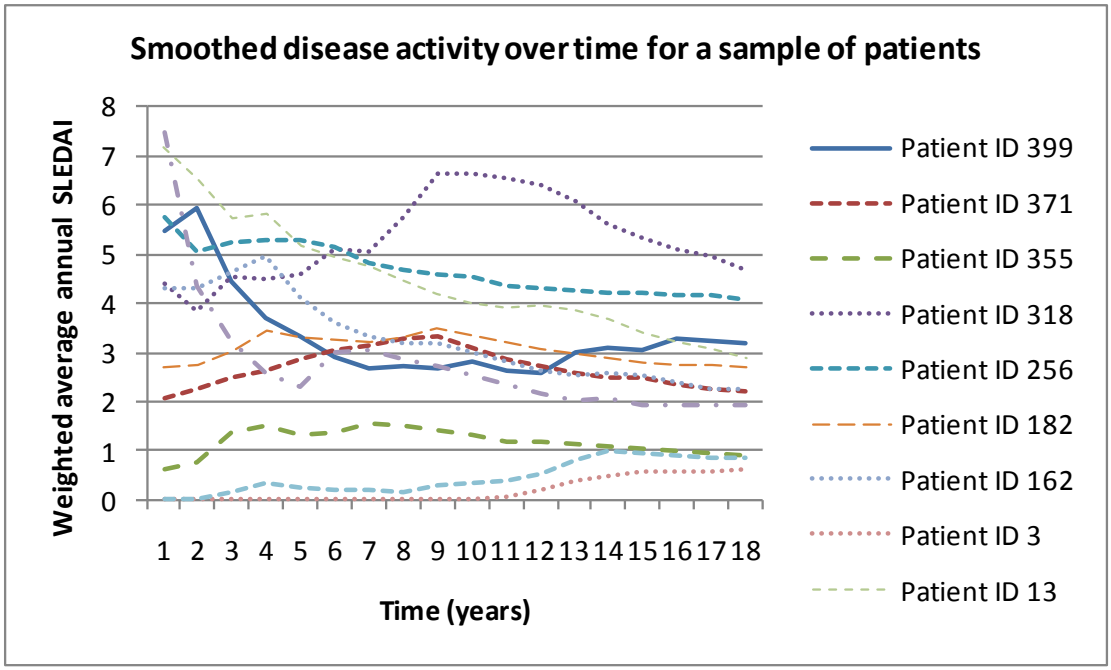

Year 3
10
5
0
-5
-10
0
5
10
15
Previous SLEDAI


Year 4
10
5
0
-5
-10
-15
0
5
10
15
Previous SLEDAI


Year 5
10
5
0
-5
-10
0
2
4
6
8
10
Previous SLEDAI

C 8 Please clarify why the usual Inter-class correlation (ICC) coefficient was not used to assess the validity of random effects model.

The ICC coefficient in the disease activity model was 0.072. This low estimate suggests that there are not significant individual patient effects in the data. Although it would have been possible to use pooled OLS to estimate the change in SLEDAI score the random effects method is useful when applying the estimates into a simulation model. The random effects variance estimate can be used in the micro-simulation to capture heterogeneity between SLE patients in their long term disease activity scores. The graph below shows the disease activity of a small randomly selected sample of patients who were observed for 18 years. The diagram illustrates that in the long term patients' SLEDAI scores converge slightly, but maintain a heterogeneous trajectory. Estimating individual patient random deviances in the disease activity model helps to maintain heterogeneity in disease activity scores in the long term.



References

Bertoli, A. M., Vila, L. M., Apte, M. et al. Systemic lupus erythematosus in a multiethnic US Cohort LUMINA XLVIII: factors predictive of pulmonary damage. Lupus 16[6], 410-417. 2007.

Chambers, S. A., Allen, E., Rahman, A. et al. Damage and mortality in a group of British patients with systemic lupus erythematosus followed up for over 10 years. Rheumatology.(Oxford) 48[6], 673-675. 2009.

Chatham, W., Weinstein, A., & Petri, M. Five-year safety and efficacy experience with belimumab, a BLyS-specific inhibitor, in patients with systemic lupus erythematosus (SLE). Annals of the Rheumatic Diseases 69[Suppl 3], 147. 2010.

Chen, Y. F., Jowett, S., Barton, P. et al. Clinical and cost-effectiveness of epoprostenol, iloprost, bosentan, sitaxentan and sildenafil for pulmonary arterial hypertension within their licensed indications: a systematic review and economic evaluation. Health Technology Assessment 13[49]. 2009.

Condliffe, R., Kiely, D. G., Peacock, A. J. et al. Connective tissue disease-associated pulmonary arterial hypertension in the modern treatment era. Am.J.Respir.Crit Care Med. 179[2], 151-157. 15-1-2009.

Fernandez, M., Alarcon, G. S., Apte, M. et al. Systemic lupus erythematosus in a multiethnic US cohort: XLIII. The significance of thrombocytopenia as a prognostic factor. Arthritis and Rheumatism 56[2], 614-621. 2007.

Merrill, J. T., Neuwelt, C. M., Wallace, D. J. et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis Rheum 62[1], 222-233. 2010.

NHS National Specialised Commissioning Group 2009, *Interim Commissioning Policy - Target therapies for the treatment of pulmonary arterial hypertension in Adults*.

Appendix 1:

Table 1.1 Disease activity (SS score) over time - High disease activity subgroup (Target Population)

Time (weeks)	SoC	Belimumab Responders	Belimumab Non-responders
0	12.718	12.665	12.745
1	8.166	8.068	4.908
2	6.486	6.411	3.361
3	5.516	5.473	2.631
4	4.946	4.921	2.277
5	4.606	4.592	2.157
6	4.398	4.391	2.164
7	4.269	4.266	2.223
8	4.185	4.183	2.311
9	4.127	4.127	2.411
10	4.085	4.085	2.503
11	4.052	4.053	2.593
12	4.025	4.026	2.678
13	4.001	4.002	2.760
14	3.981	3.981	2.836
15	3.961	3.962	2.902
16	3.943	3.943	2.964
17	3.925	3.926	3.024
18	3.907	3.908	3.077
19	3.890	3.891	3.123
20	3.874	3.874	3.166
21	3.858	3.858	3.203
22	3.842	3.843	3.237
23	3.827	3.827	3.271
24	3.812	3.812	3.303
25	3.797	3.797	3.328
26	3.782	3.782	3.349
27	3.768	3.768	3.369
28	3.754	3.755	3.386
29	3.740	3.740	3.403
30	3.727	3.727	3.418
31	3.713	3.714	3.429
32	3.700	3.702	3.437
33	3.687	3.689	3.446
34	3.675	3.676	3.450
35	3.663	3.662	3.454
36	3.652	3.649	3.453
37	3.640	3.638	3.452
38	3.628	3.626	3.451

Time (weeks)	SoC	Belimumab Responders	Belimumab Non-responders
39	3.617	3.615	3.452
40	3.606	3.604	3.451
41	3.595	3.594	3.450
42	3.584	3.584	3.453
43	3.574	3.574	3.456
44	3.564	3.565	3.457
45	3.554	3.555	3.455
46	3.543	3.546	3.451
47	3.534	3.536	3.449
48	3.525	3.527	3.445
49	3.515	3.516	3.440
50	3.506	3.508	3.433
51	3.497	3.500	3.429
52	3.486	3.490	3.424
53	3.476	3.480	3.419
54	3.467	3.470	3.417
55	3.458	3.461	3.414
56	3.447	3.452	3.407
57	3.438	3.442	3.400
58	3.429	3.432	3.393
59	3.420	3.423	3.387
60	3.410	3.416	3.385
61	3.399	3.410	3.375
62	3.391	3.402	3.367
63	3.384	3.394	3.356
64	3.375	3.385	3.345
65	3.366	3.375	3.338
66	3.359	3.367	3.329
67	3.350	3.357	3.321
68	3.341	3.348	3.308
69	3.333	3.342	3.289
70	3.323	3.335	3.282
71	3.314	3.322	3.272
72	3.306	3.312	3.278
73	3.296	3.306	3.271
74	3.283	3.299	3.250
75	3.269	3.290	3.240
76	3.266	3.280	3.234
77	3.261	3.271	3.251
78	3.251	3.266	3.255
79	3.244	3.261	3.249
80	3.234	3.246	3.238
81	3.232	3.235	3.229

Time (weeks)	SoC	Belimumab Responders	Belimumab Non-responders
82	3.233	3.210	3.221
83	3.223	3.196	3.212
84.00	3.212	3.212	3.207
85.00	3.204	3.209	3.197
86.00	3.191	3.201	3.186
87.00	3.178	3.193	3.178
88.00	3.166	3.185	3.170
89.00	3.158	.	3.162
90.00	3.150	.	3.149
91.00	3.143	.	3.149
92.00	3.141	.	3.142
93.00	3.130	.	3.125
94.00	3.122	.	3.134
95.00	3.130	.	3.126
96.00	.	.	3.118
97.00	.	.	3.107
98.00	.	.	3.100
99.00	.	.	.
100.00	.	.	.

Table 1.2 AMS over time - High disease activity subgroup (Target Population)

Time (weeks)	SoC	Belimumab Responders	Belimumab Non-responders
0	12.718	12.691	12.709
1	10.442	10.404	9.878
2	9.124	9.086	8.392
3	8.222	8.188	7.429
4	7.567	7.537	6.753
5	7.073	7.047	6.258
6	6.691	6.668	5.886
7	6.388	6.368	5.598
8	6.143	6.126	5.372
9	5.942	5.926	5.190
10	5.773	5.758	5.042
11	5.630	5.616	4.919
12	5.506	5.494	4.815
13	5.399	5.387	4.728
14	5.304	5.294	4.653
15	5.220	5.210	4.587
16	5.145	5.136	4.530
17	5.077	5.068	4.480

18	5.016	5.007	4.435
Time (weeks)	**SoC**	**Belimumab Responders**	**Belimumab Non-responders**
19	4.959	4.952	4.395
20	4.908	4.900	4.359
21	4.860	4.853	4.326
22	4.816	4.809	4.297
23	4.774	4.768	4.269
24	4.736	4.730	4.244
25	4.700	4.694	4.221
26	4.666	4.660	4.199
27	4.634	4.628	4.179
28	4.603	4.598	4.160
29	4.575	4.570	4.143
30	4.547	4.542	4.126
31	4.521	4.516	4.110
32	4.496	4.492	4.095
33	4.473	4.468	4.081
34	4.450	4.445	4.067
35	4.428	4.424	4.054
36	4.407	4.403	4.041
37	4.387	4.383	4.029
38	4.367	4.363	4.017
39	4.349	4.345	4.006
40	4.330	4.327	3.995
41	4.313	4.309	3.984
42	4.296	4.292	3.974
43	4.280	4.276	3.964
44	4.264	4.260	3.954
45	4.248	4.245	3.945
46	4.233	4.230	3.935
47	4.219	4.215	3.926
48	4.205	4.201	3.918
49	4.191	4.188	3.909
50	4.177	4.174	3.901
51	4.164	4.161	3.893
52	4.151	4.149	3.885
53	4.139	4.136	3.877
54	4.127	4.124	3.869
55	4.115	4.112	3.861
56	4.103	4.101	3.854
57	4.092	4.089	3.847
58	4.080	4.078	3.839
59	4.069	4.067	3.832
60	4.059	4.056	3.825

61	4.048	4.046	3.818
Time (weeks)	**SoC**	**Belimumab Responders**	**Belimumab Non-responders**
62	4.037	4.036	3.811
63	4.027	4.025	3.805
64	4.017	4.015	3.798
65	4.007	4.006	3.791
66	3.998	3.996	3.785
67	3.988	3.987	3.778
68	3.979	3.977	3.772
69	3.970	3.968	3.765
70	3.960	3.959	3.759
71	3.951	3.950	3.753
72	3.943	3.942	3.746
73	3.934	3.933	3.740
74	3.925	3.924	3.734
75	3.917	3.916	3.728
76	3.908	3.907	3.722
77	3.900	3.899	3.716
78	3.892	3.891	3.710
79	3.884	3.883	3.704
80	3.876	3.875	3.699
81	3.868	3.867	3.693
82	3.860	3.860	3.687
83	3.852	3.852	3.682
84	3.845	3.844	3.676
85	3.837	3.837	3.671
86	3.830	3.830	3.665
87	3.823	3.822	3.660
88	3.815	3.815	3.654
89	3.808	.	.
90	3.801	.	.
91	3.794	.	.
92	3.787	.	.
93	3.780	.	.
94	3.773	.	.
95	3.766	.	.
96	.	.	.
97	.	.	.
98	.	.	.
99	.	.	.
100	.	.	.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Thank you for agreeing to give us a statement on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: [redacted]

Name of your organisation: British Association of Dermatologists

Are you (tick all that apply):

- ✓ a specialist in the treatment of people with the condition for which NICE is considering this technology?

  As a dermatologist, I do see and treat patients with lupus erythematosus, but usually those with cutaneous disease.

- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)?

  No.

- ✓ an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)?

  As chair of the Therapy & Guidelines subcommittee, I represent the British Association of Dermatologists

- other? (please specify)

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

What is the expected place of the technology in current practice?

How the condition is currently treated in the NHS?

As indicated in Appendix A.

Is there significant geographical variation in current practice?

Not as far as we are aware.

Are there differences of opinion between professionals as to what current practice should be?

Not as far as we are aware.

What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

As indicated in Appendix A.

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient?

Gender and racial differences, as indicated in Appendix A.

Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

Not as far as we are aware.

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics?

Secondary care.

Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

As belimumab is administered intravenously, specialist nursing care will be required.

If the technology is already available, is there variation in how it is being used in the NHS?

Not to our knowledge.

Is it always used within its licensed indications? If not, under what circumstances does this occur?

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

We are not aware that belimumab is currently used for conditions other than lupus erythematosus.

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

We are not aware of any relevant clinical guidelines.

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK.

Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

We suspect that the practical implications for the use of belimumab will be similar to the other biological agents in current use.

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

We have no comment to make.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

We are not sufficiently familiar with the evidence base to comment.

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

We do not have sufficient experience in the use of belimumab to comment.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

We are not aware of any further relevant evidence.

Implementation issues

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

We assume that the delivery of this technology would be similar to the delivery of other biological agents, and, as such, would not require additional NHS resources.

Equality

Are there any issues that require special attention in light of the NICE's duties to have due regard to the need to eliminate unlawful discrimination and promote equality and foster good relations between people with a characteristic protected by the equalities legislation and others?

We are not aware of any.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Thank you for agreeing to give us a statement on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: [redacted]

Name of your organisation
British Health Professionals in Rheumatology

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology?
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)?
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)?
- other? (please specify)

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

SLE is currently treated both within specialist centres and DGH's; there is limited treatment currently available for what is a debilitating disease which has high mortality and morbidity.
This particular technology would be best placed in secondary care utilising specialist skills for both assessment of disease and monitoring the therapy.
The technology is not currently available within the NHS except for patients involved in the trials.

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements

for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

There are few treatments available for people with SLE, it is a rare disease but devastating for those who suffer with it. Available treatments are often not licensed for SLE but are utilised as there is some evidence for their efficacy. Belimumab is specially licensed for SLE and is effective for the reduction of disease activity in adult patients with active, autoantibody positive systemic lupus erythematosus who are receiving standard therapy. It is the first drug to be specifically approved for treating SLE in more than 50 years.

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

No

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Implementation issues

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

Nurses would not need extra training apart from knowledge on the side effect profile and infusion characteristics.

Equality

Are there any issues that require special attention in light of the NICE's duties to have due regard to the need to eliminate unlawful discrimination and promote equality and foster good relations between people with a characteristic protected by the equalities legislation and others?

no

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Thank you for agreeing to give us a statement on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you: A specialist in the treatment of patients with the condition for which NICE is considering a technology

Your name: [redacted]

Name of your organisation : The British Society for Rheumatology

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology?√
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)?
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)?
- other? (please specify)

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

_____

Systemic lupus erythematosus is a potential serious autoimmune rheumatic disease. It invariably involves the skin and joints, but more serious manifestations include involvement of the lungs, heart, central nervous system and the kidneys. Mild versions of the disease are invariably treated with combinations of hydroxychloroquine and non-steroid anti-inflammatory drugs often together with low doses of steroids. With more overt arthritis and pleuritic pain for example, moderate doses of corticosteroids (10-20mg per day, together with a drug like azathioprine and/or methotrexate are widely used). The more serious manifestations, particularly renal disease, are invariably treated with high doses of steroids (20mg+ and mycophenolate or intravenous cyclophosphamide). Benlysta is of potential value for patients with lupus who respond inadequately to this reasonably standardised set of drugs or where the drugs themselves cause an unacceptable side effect. In my experience (I've managed close to 600 patients with SLE over a 30-year period), around 10-15% of lupus patients come into this category.

Benlysta has met its primary end points in two large scale trials involving many hundreds of patients. It has principally been used to treat patients with mucocutaneous, musculoskeletal and respiratory problems. It has not yet been established how effective it will be in treating patients with renal or cerebral disease.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

As indicated previously, a small, but significant number of lupus patients fail to respond adequately to conventional drugs and for these patients at present, intravenous immunoglobulin (now relatively scarce and rather expensive) or rituximab (off label) have been widely used in the past ten years. There is an ongoing debate about the use of rituximab as this drug has failed to meet its end points in two large clinical trials in the United States, though >20 open label studies have indicated its utility. Major doubts have been expressed about the trial design used in the rituximab trials. Two major clinical trials of Benlysta in the United States using as their endpoint a composite activity index involving the SLEDAI global index (this had to be improved by 4 points); the BILAG index (no new BILAG A's or B's at the predefined time points were allowed) and a Physician's Global (this had to, remain the same or improve). Given that, to my mind, the bar was set high in these studies with fairly liberal amounts of steroids and other immunosuppressive drugs being co-prescribed although the differences between the patients given Benlysta and those given placebo were modest, the fact that the endpoints were achieved, is I think impressive.

I can see no reason why results obtained in the United States and elsewhere would not be applicable to the UK though I remain particularly anxious to discover whether

or not Benlysta is effective in renal and other aspects of lupus that have not been tested to date.

I'm unaware of any major side effects having emerged in the course of the two large Benlysta studies published so far.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

-

Implementation issues

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

There is no doubt that NICE approval of Benlysta would be in line with the FDA's approval of it (the first drug to be specifically approved for lupus in 52 years, incidentally). Clearly, there would be a "learning curve" required for staff in the NHS who we wish to use Benlysta but these in the first instance would almost certainly be highly specialised rheumatology units well used to using biologic agents.

Equality

Are there any issues that require special attention in light of the NICE's duties to have due regard to the need to eliminate unlawful discrimination and promote equality and foster good relations between people with a characteristic protected by the equalities legislation and others?

Not that I am aware of

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Thank you for agreeing to give us a statement on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: [redacted]

Name of your organisation: Primary Care Rheumatology Society

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? **YES**
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)?
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)?
- other? (please specify)

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS?
Drug treatments are generally tailored to either type of organ involvement and severity, and range from mild drugs for mild disease (e.g NSAIDS, Hydroxychloroquine) through to chemotherapeutic and biologic drugs for severe organ manifestations (e.g. Cyclophosphamide and off-licence use of Rituximab). However, disease activity and treatment response are unpredictable both within and between patients, and many also require treatment with long term oral steroids, with concomitant adverse effects on infection risks and cardiovascular disease.
Non-drug treatments include an essential role of Lupus nurse specialists in educating patients about self-management, avoidance of flares, counselling about drug treatment and providing rapid response to patients with active disease.

Is there significant geographical variation in current practice?
This is not currently known. However, given the multisystem nature of the disease, and it's rarity, patient management is improved when there is co-terminus access to other relevant specialists e.g rheumatology, renal, dermatology, fetomaternal, and this coordination of "secondary" care also enhances the ability of primary and secondary care to work as a team. This coordinated multidisciplinary care is likely to be better in large specialist (regional) centres, although this has not been formally assessed/evidenced. However, for patients with rare complex disease such as SLE, it is important for their GP to be able to access an appropriate "specialist" rather than "generalist" opinion, particularly when there is either diagnostic uncertainty or where complex intervention is needed.

Are there differences of opinion between professionals as to what current practice should be?
No substantial differences are likely to exist, particularly amongst specialist centres, but this has not been formally evaluated.

What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

This organisation does not yet know the precise positioning of the technology compared to other existing strategies.

Current alternatives are

1. IV cyclophosphamide – this has major disadvantages esp in the setting of a disease of young women, as the side effects include infection, marrow toxicity, and infertility.
2. IV Rituximab – unlicensed but in use in refractory disease. Disadvantages include the long half life (6 months), and lack of proven RCT benefit.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

3. Mycophenolate Mofetil – beneficial in renal disease but beneficial effects in other SLE manifestations not as clearly demonstrated compared to the proposed technology.
4. High dose steroids- disadvantages are many and include bone related (osteoporosis, AVN), endocrine (diabetes, wight gain), skin (striae, bruising), increased risk of infections, and increased cardiovascular risks. Thus any agent with proven steroid sparing/reduction capability in SLE is likely to have short and long term benefits

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient?
Yes, patients from the ethnic groups already mentioned, particularly who present with aggressive disease (e.g. renal) at diagnosis. And also – patients who present late with existing disease "damage" in whom the subtle and often insidious or silent (e.g. renal) symptoms have not been detected early in Primary Care.

Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?
I am not aware of particular subgroups of patients who are more likely to respond (other than those with disease activity meeting trial eligibility), or be at risk, unless this has been revealed by secondary analysis of existing data.

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics?
Secondary care, ideally specialists who are skilled in assessing and treating SLE, particularly if this would ensure appropriate use and assessment (and hence better use of NHS resource)

Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

There is likely to be additional specialist nursing support esp in terms of day case care.

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?
Not currently available

Please tell us about any relevant clinical guidelines and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

I am not aware of any guidelines on use of this technology. There are European (EULAR) guidelines on management principles of SLE.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?.

The use of IV biologic agents is well established in Rheumatology, and so from this perspective, there is likely to be existing staff and facility availability i.e. the logistics of delivering the drug to patients is not likely to pose difficulties.

The availability of a drug with RCT evidence of efficacy, compared to standard therapies, is likely to outweigh any issues of patient acceptability of monthly IV infusions.

There is no evidence that the technology will interact or conflict with existing (concomitant) drugs that are likely to be continued alongside (e.g Hydroxychloroquine, Steroids)

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

This is likely to benefit patients with active disease (indicative of B cell hyperactivity), who have persistent disease activity despite one or more standard drugs. This will need to be defined according to formal outcome measures, to enable continuation of therapy only if response met. It would be appropriate to consider whether outcomes other than a composite responder index (i.e the trial outcome measure) such as significant steroid reduction would also be appropriate in clinical practice. The published trial data appears to indicate a response signal at 16 weeks of therapy.

It is difficult to assess from the published data as to when/if treatment should be discontinued in patients who have responded.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes

The trial was predominately conducted in the Latin-American and Asian-Pacific regions, with only a 11% contribution from (Eastern) Europe.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

However, the trial entry criteria, reflecting active disease (i.e definite SLE disease activity) is likely to reflect the circumstances in which the drug would be used and are likely to be extrapolatable to UK practice. The outcome measure used were appropriate and indeed very stringent in terms of no-worsening of any features of disease.

What is the relative significance of any side effects or adverse reactions In what ways do these affect the management of the condition and the patient's quality of life?
There is a risk of infusion reactions (as with any IV biologic drug). This can be managed at the time of the infusion if it occurs. There is also a risk of reduced Immunoglobulin levels. Overall the risks compared to placebo across all side effects visible within the published BLISS 52 results does not appear to demonstrate any increased adverse events overall.

Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?
Not to my knowledge

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a

judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

I am not aware of any such information

Implementation issues

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

It is likely that any patients receiving Belimumab would need to only continue to do so if there was evidence of efficacy (i.e. analogous to DAS response criteria for anti-TNF therapy in RA). If the response criteria for clinical use is, for example, going to mirror the composite end-point response, then NHS staff may need training in using these

end points, unless the drug is only given in specialist units who are already familiar with this.
However, as a general comment, the adoption of more formal, systematic assessment of disease activity and damage is likely to have a secondary impact of improving the standard of care for these patients (by focusing clinicians on outcome), regardless of drug utilisation.
Additional day case resources may be required, depending on whether existing local biologic infusion facilities are at fully capacity or not.

Equality

Are there any issues that require special attention in light of the NICE's duties to have due regard to the need to eliminate unlawful discrimination and promote equality and foster good relations between people with a characteristic protected by the equalities legislation and others?

SLE does have a predilection for certain ethnic groups (African American, Asian) who are likely to be over-represented in SLE populations compared to the general population.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Thank you for agreeing to give us a statement on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: [redacted]

Name of your organisation – The Renal Association

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? Yes; run a large lupus nephritis clinic at the Imperial College Lupus Centre at the Hammersmith Hospital. With 3 other renal consultants and 2 rheumatologists we are responsible for ~400 patients.

- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)? Yes – was involved in the ALMS lupus nephritis trial, regularly consulted on new trials and have written reviews of the current evidence for treatment. Was not involved in the Belimumab trials but have participated in several advisory boards reviewing the data and planning future trials.

- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)? Not an employee but an Elected member of the National Executive of the Renal Association and Elected executive representative on the Renal Association Clinical Affairs Board.

- other? (please specify)

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

SLE is currently treated according the severity of the disease both globally and depending upon which organ systems are involved. Treatment ranges from symptomatic (antiinflammatories), to immunomodulatory (antimalarials, particularly effective for skin and joint manifestations but also effective at reducing major organ flares) to immunosuppressive. The latter group of drugs includes steroids, azathioprine, mycophenolate mofetil (MMF), Cyclophosphamide (CyP), calcineurin inhibitors (CNIs) and biologicals particularly anti CD20 antibodies. Standard of care for patients with moderately severe disease is likely to include steroids without or without a second agent. For patients without significant renal / haematological / cerebral involvement the second agent is likely to be azathioprine or possibly MMF (though this has an important cost implication and is much more costly than azathioprine). For significant renal involvement, the commonest agents added to steroids are MMF or CyP (usually given as IV pulses either monthly (the NIH regime) or lower doses every two weeks (the Eurolupus regime). Rituximab –an antiCD20 mab – is usually reserved for refractory cases; its use is limited both by cost and the lack of positive RCT data.

There are significant geographical differences in treatment – these are accounted for both by financial constraints and history. The high dose pulsed CyP regime was pioneered in the USA and tends to be used rather than the Eurolupus regimen – not least because the latter has only really been evaluated in a Caucasian population without severe renal involvement whereas the majority of patients treated in the USA are of African-American origin. Lupus is much more common in non-European white populations and these are overrepresented in urban areas of the UK. MMF is popular because it has been well trialled and does not cause infertility and is probably more effective than CyP in patients of African descent. However, it is costly and this has limited its use in regions where financial restrictions limit choice of medications. Importantly however, MMF is now off patent and generic versions are available making this a much more affordable drug. Many units in the UK are switching to the use of generic immunosuppressants with expected cost savings in the region of 50% which can amount to several £100K in large units. The use of Rituximab is limited by many PCTs because of cost (~£4000/course) and is essentially unobtainable in some regions of the UK.

There are professional disagreements:

a) CyP has been used for more than 3 decades – and there are long term outcome studies showing superiority over azathioprine for instance in reducing chronic kidney damage, long term safety (or not if cumulative doses >30-36g are used) and long term efficacy. It is also the drug that has been used most extensively in the most severe forms of lupus (cerebral lupus or rapidly progressive renal disease)

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

b) In contrast, trials using MMF have looked at relatively short term outcomes so that although MMF appears to be an effective induction agent, until recently there was little data about long term efficacy. There have been two recent maintenance trials – one from the ALMS group demonstrating clear superiority of MMF over azathioprine as a maintenance agent in a group of patients whose nephritis had responded to induction treatment with either IV CyP or oral MMF; and one from the Eurolupus group showing equivalence of MMF and azathioprine in maintenance treatment in patients randomised at the outset and having induction with low dose IV CyP.
c) Despite effective induction regimens and some data on maintenance, flare remains an important problem in lupus in general and lupus nephritis in particular. Increasing time to flare and reducing the severity of flares are important clinical endpoints as flares are associated with accrual of long term damage and increased likelihood of permanent renal damage and progression to end-stage kidney disease.
c) There is great debate about steroid usage. Steroids are used in all treatment regimens for lupus of any severity but are probably the cause of much of the long term damage accrued in patients with lupus and may well account for a significant part of the increased risk of premature cardiovascular disease in patients with lupus. Drugs that reduce the dose of steroid required and reduce the need for increases in dose of steroids are desirable – and the Belimumab trial appears to suggest that addition of Belimumab might allow reduction of steroid dosage and reduction in the need for increased doses. It should be noted that one unit (the Imperial College Lupus Centre) has pioneered the early use of Rituximab and MMF in lupus nephritis and developed a regimen that does not use oral steroids. However, this has yet to be trialled in a formal RCT and is not yet widely used.
d) Perhaps the biggest controversy at present is in the role of Rituximab. There is a great sense that Rituximab – an anti CD20 mAb – is an effective agent particularly for refractory disease. However, two RCTs (one in SLE without renal involvement, the Explorer study and one in lupus nephritis, the LUNAR study) using Rituximab as an additional agent, failed to show benefit for the addition of Rituximab. Explorer was hampered by requiring superiority with a large delta (20%) and including seronegative patients and allowing high doses of steroids to be used. LUNAR is not yet published but was probably underpowered and again allowed the use of quite high dose MMF and steroids which is clearly an effective combination to treat renal disease. It also only followed patients for one year and improvements are likely to be seen later. Again the data from the Imperial Lupus centre is of interest because they have pioneered the early use of Rituximab in order to minimise steroid exposure. Their data (unpublished) show high rates of complete and partial remission at one year without the use of oral steroids and suggest that early treatment may alter the course of lupus. In the current financial climate, prescribing Rituximab is likely to become more difficult in the face of negative (even if flawed) trials, however much clinicians believe it works. An agent that is licensed that might offer a route to reducing concomitant steroid usage is an attractive prospect.
e) The disadvantages of the current regimens – CyP can cause infertility in a dose and age dependent manner. This has made it increasingly unpopular though the Eurolupus regimen is unlikely to cause infertility (total dosage 3g). CyP is teratogenic. A key advantage of CyP is that it is an IV regimen and overcomes issues of non compliance. MMF in contrast is an oral agent and in trials has been used at quite high dosages. There is a growing literature on the value of dosing to

trough MPA levels which allows more tailored dosing and avoidance of side effects. A major issue with MMF is that it is almost certainly teratogenic so needs to be switched to azathioprine pre conception.
f) All the immunosuppressants carry the risk of infections - both classical bacterial and viral infections as well as those of the immunocompromised.

There is a real need for new agents in the treatment of lupus – no drug has been licensed for 53 yrs; indeed the only licensed drugs are prednisolone and hydroxychloroquine. Therapeutic strategies are needed that are targeted to the pathogenesis of lupus, that have an excellent safety profile (this is a relapsing remitting disease that requires treatment over a number of years often in young patients), and that spare the dose of steroids required. Such agents should reduce flare frequency and increase time to flare. On the basis of published data, it would seem that Belimumab might be just such an agent at least in those without marked renal or cerebral involvement. It is also important to note the suggestion that Belimumab might be particularly effective if used to treat lupus at an early stage though this needs to be confirmed in a trial.

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?
Different subgroups – certainly patients with nephritis or cerebral involvement would constitute the most severe groups of patients. However, the technology has not been specifically evaluated in these subgroups. There should be no group particularly put at risk by Belimumab or likely to be put at risk unless the cumulative effect of Belimumab on the background of a high burden of prior and concurrent immunosuppression proves to be associated with serious infections.

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?
Lupus is a complex multisystem disease that should ideally be managed in specialist multidisciplinary clinics by physicians with an interest in lupus and at the very least in centres of excellence where high quality rheumatology, renal and other specialist input is readily available. The use of new agents should initially be limited to such centres and Belimumab is very unlikely to be given / prescribed in a primary care setting. Administration of the drug in the current formulation requires day case stay on a monthly basis for an IV infusion. If a sub cutaneous formulation becomes available then an outpatient setting or self administration might be feasible. The added costs would be those of the drug and the need for monthly one day inpatient stays. However, this could be offset by reducing the costs due to fewer flares which would lead to a reduction in concomitant medications, and in admissions for investigation and treatment of flares and the long term complications of steroid use.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur? N/A

*Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations*.
There are no guidelines addressing the use of Belimumab. There are guidelines on lupus nephritis soon to be published by KDIGO – however these do not address the question of when should Belimumab be used. Those guidelines do not recommend the use of Rituximab at present in light of the absence of positive trial data. There are more general guidelines on the management of lupus which have recently published by EULAR (the European League against Rheumatism) – these are very wide ranging and offer advice on general management as well as specific therapeutic advice. Again, Belimumab has not been included in these guidelines.

The advantages and disadvantages of the technology
NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

As discussed earlier, the only biologic in fairly frequent use for the treatment of lupus in the UK is Rituximab. Introduction of Belimumab will pose problems for those units unused to giving day case infusions for the treatment of lupus but if limited to centres of excellence this is unlikely to be a practical problem. There is a great willingness among patients with lupus to use new technologies and all are keen to reduce the use of steroids. It is interesting to note that patients were very quick to note the news of Belimumab being effective in lupus and are lobbying quite actively for it to be available. It remains to be seen how acceptable it will be to have monthly infusions over a long period of time – there is good long term extension data available from the phase II and phase III trials but it is not yet clear what happens when the infusions are stopped – is the improvement maintained? Long term monthly infusions for large numbers of patients will be quite challenging to manage and costly – a full economic analysis will be needed to ensure that costs are saved long term not only terms of the health of the patient but in terms of ability to work.

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

It remains unclear exactly in which patient group to use the new technology in – the trials were very well designed but were very inclusive and included a heterogeneous group of patients – both an advantage and a disadvantage. It is likely that initially patients with grumbling active disease who would otherwise face an increase in

steroids or the addition of a new agent will be treated with Belimumab. There is too little data to recommend its use in lupus nephritis but it is very likely if available and approved by NICE it will be used instead of Rituximab in the units that find funding for Rituximab challenging. There is a thirst for new approved agents in this disease and it is likely that uptake of use will be rapid. It should be mandatory to report all usage to a biologics registry – this is a very effective way of picking up adverse events that might be missed in trials and to grow understanding about the most effective ways to monitor, start and stop treatment. The use of Belimumab in the first instance should be restricted to centres of excellence in the management of lupus – this will ensure patients being treated are well phenotyped, have appropriate measurements of disease activity undertaken and appropriately monitored. It is likely that treatment for one or two years will be used in the first instance.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

The trials were well conducted, undertaken in diverse geographical settings and in patients being treated with local standard of care. Hence they reflect current practice. There are two important areas where the trials differed from standard clinical practice – the use of concomitant medications and the use of the responder index to evaluate outcomes.
Two key groups of concomitant medications were limited within the trial – drugs that block the renin angiotensin system (ACE inhibitors or angiotensin receptor blockers) and statins to reduce cholesterol. The rationale was reasonable – RAS inhibitors reduce proteinuria and might bias (wrongly) towards interpreting patients as responders and statins can cause muscle problems and a rise in creatinine kinase which could be misinterpreted as a flare. However, these effects should have been balanced between the groups and in clinical practice all patients with proteinuria should be on RAS blockade drugs to control hypertension and reduce proteinuria as these classes of drugs are renoprotective. Patients were allowed to be on these drugs but the dose had to be stable at the time of admission to the trial. It is possible therefore that the real life changes in proteinuria may be lower when Belimumab is added in clinical practice as clinicians tend to titrate the dose of RAS blockers up to the maximum tolerated in order to reduce proteinuria. Similarly statins rarely cause overt myositis and cardiovascular risk factors should be aggressively managed in all lupus patients and especially those with chronic kidney disease. These drug issues would need to be addressed in any planned trial of the drug in lupus nephritis as nephrologists would be keen to be able to modify RAS blockade or statin dosage in the management of these patients.

With respect to clinical practice and outcome measures, in the trial outcomes were assessed by a composite responder index which is based on very logical principles but involves clinical scoring that is not routinely used in many excellent renal and

rheumatological units in the UK. For instance in lupus nephritis most clinicians monitor levels of proteinuria and serum creatinine. These variables can be incorporated readily into the scores used but the scoring system is not in routine practice. However, there is growing recognition that such scoring is required and the introduction of Belimumab will aid this process being implemented. The most important outcomes were the reduction in time to flare, the severity of the flare and the ability to reduce steroids and avoid the need for increase in steroids. These are significant and important outcomes in a chronic relapsing and remitting disease. Importantly the treatment appeared to be beneficial in all organ systems though the trials were not specifically powered or designed to look at lupus nephritis. Future trials need to be directed towards specific subgroups of patients especially those with lupus nephritis and any such trial should include the requirement for a baseline renal biopsy to ensure active disease is being treated and ideally follow up biopsy to ensure histological remission has been achieved.

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

The safety profile appears good and the drug has not been used in routine clinical practice.

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

N/A

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Implementation issues

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

As described earlier would need to be administered on a day case basis in centres of excellence. This would have a day case bed usage implication but does not require extra facilities or equipment. As with other biologics (and centres of excellence should be used to administering these), staff need to be trained in the use of appropriate premedication, the recognition and treatment of infusion reactions and the ability to deal with acute severe infusion reactions that are reportedly rare but can happen.

Equality

Are there any issues that require special attention in light of the NICE's duties to have due regard to the need to eliminate unlawful discrimination and promote equality and foster good relations between people with a characteristic protected by the equalities legislation and others?

Lupus is more common in non-northern European ethnic groups. Trials with other drugs have suggested ethnic variations in responses to certain drugs and it will be critically important to have data on ethnic specific responses lest an overall low rate of response masks a particularly good response in one ethnic group or another.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Thank you for agreeing to give us a statement on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: [redacted]

Name of your organisation: Royal College of Nursing (RCN Rheumatology Forum)

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? **X** [redacted]
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)?
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)? **X RCN Member**
- other? (please specify)

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?

*Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.*

SLE is a multi-system potentially life threatening disease and at present there are limited treatment options available to patients. Current practice is to use high dose corticosteroids and immunosuppressants such as Azathioprine, Methotrexate, Cyclophosphamide, Mycophenolate Mofetil and the anti malarial Hydroxychloroquine. Currently the only licensed treatments in the UK for SLE are Hydroxychloroquine and Prednisolone. Rituximab is used off licence for named patients only and has shown promising results especially in those with renal lupus. The new treatment with Belimumab offers healthcare professionals an opportunity to use a licensed treatment that has shown significant benefit for those refractory to current treatment options.

This technology would be most appropriate for those with moderate to severely active disease to protect organ systems and improve prognosis. Those with milder disease may not need this treatment although those remaining on high doses of corticosteroids for many years would also be a very useful group to consider due to the potential long term complications of steroids.

This should be available primarily through specialist units who have a large lupus cohort and experience, preferably in a secondary care setting. EULAR recommendations for the management of lupus (2008) include immunosuppression and corticosteroids for those with active disease, especially neuropsychiatric and renal manifestations.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

A number of issues will potentially present as this is an intravenous treatment. This will include the delivery in the secondary care setting, monitoring during infusions, appointments to hospitals and potential admission to hospital if an adverse event occurs.

This treatment should really be reserved for:

1. Those with moderate to severely active disease
2. Those not responding to conventional immunosuppressants
3. Those unable to reduce oral steroid dose particularly those who have been taking such doses for many years

This reviewer has not had experience of this drug in clinical trials settings.

Looking at the trial data, the frequency and type of adverse events and side effects were comparable with those on conventional treatment regimes and /or placebo. There is good evidence that this treatment reduces steroid dose and incidence of flares.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

We are not aware of additional sources of evidence at this stage.

Implementation issues

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

NICE guidance would enable healthcare professionals to consider this technology as a treatment option especially in those with active disease unresponsive to conventional treatments. It would also be a licensed product for use in the UK (currently awaiting decision, although approved through USA FDA).

Most rheumatology units have day case units where the treatment could be delivered.

There would be some training need for the nurses who will be delivering the drug especially related to any potential side effects, length of stay in the unit (including time to observe post treatment) and management of monitoring/adverse events.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Equality

Are there any issues that require special attention in light of the NICE's duties to have due regard to the need to eliminate unlawful discrimination and promote equality and foster good relations between people with a characteristic protected by the equalities legislation and others?

None that we are aware of specifically. We would, however, suggest that NICE actively seeks to involve relevant national community/patient organisations to get first hand information on how this disease affects people from various communities. A published record of an equality impact analysis would be helpful in this respect.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Thank you for agreeing to give us your views on the technology and the way it should be used in the NHS.

Primary Care Trusts (PCTs) provide a unique perspective on the technology, which is not typically available from the published literature. NICE believes it is important to involve NHS organisations that are responsible for commissioning and delivering care in the NHS in the process of making decisions about how technologies should be used in the NHS.

To help you give your views, we have provided a template. The questions are there as prompts to guide you. You do not have to answer every question. Short, focused answers, giving a PCT perspective on the issues you think the committee needs to consider, are what we need.

About you

Your name: [redacted]

Name of your organisation: **NHS Bolton**

Please indicate your position in the organisation:

- ✓ [redacted]
- ✓ [redacted]
- responsible for quality of service delivery in the PCT (e.g. medical director, public health director, director of nursing)?
- a specialist in the treatment of people with the condition for which NICE is considering this technology?
- a specialist in the clinical evidence base that is to support the technology (e.g. participation in clinical trials for the technology)?
- other (please specify)

[redacted]

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS?
Standard therapy would normally consist of NSAIDs, hydroxychloroquine and/or prednisolone (all licensed). Immunosuppressives may be added (methotrexate, azathioprine, ciclosporin, tacrolimus and cyclophosphamide) but these are all unlicensed. For more patients with more severe disease rituximab (unlicensed) has been used.

Is there significant geographical variation in current practice?
As many of the treatment choices for moderate to severe disease are unlicensed it is suspected there will be variation in practice due to prescriber's knowledge of the condition and past experiences with the various drugs. As rituximab is PbR excluded, the decision to fund this will be via IFR route as PCTs will determine if there is exceptionality for that specific patient. Funding decisions will vary across PCTs and may depend on the local patient population.

Are there differences in opinion between professionals as to what current practice should be?
For the reasons above it is suspected there will be a degree of variation, particularly as there is a lack of a licensed product for patients with severe disease.

What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Only those drugs that would be used in more severe cases of SLE have been considered. Standard treatments would be expected to have been tried (and may continue) as per the trial criteria.

Cyclophosphamide

Advantages:
Oral or intravenous forms
Induces remission
Has been used as common practice for several years
Experience of use
Adverse effects known (due to licensed use in other disease areas)

Disadvantages:
Unlicensed
Need to undertake regular blood monitoring (FBC, WCC)
Specialist needs to retain prescribing
Costs associated with follow-up appointments and frequency (whether prescribed oral or intravenous)
Long-term use limited due to associated toxicity
Risk of inducing bladder toxicity (mesna can prevent this with associated costs)
Risk of pneumocystis (co-trimoxazole concurrently can minimise this – again with associated costs and several drug interactions, particularly methotrexate)
Female patients require an annual smear for the first 3 years

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Risk of infertility in female and male patients

Rituximab

Advantages:
Infusions of only 2 injections (repeated after 6 months if flare-up of condition) – therefore maximum 4 admissions for injection per year (and hence admission costs).
Estimated 50% of patients respond to rituximab with a flare-up after approximately 18 months.
Drug has been used in other immunosuppressant type conditions (e.g. RA) for many years safely
Safety data known for this drug (due to licensed use in other disease areas)
Possible steroid-sparing effect (case reports)

Disadvantages:
Unlicensed
Most evidence for this drug in this condition is from case reports in tertiary and national centres where these patients are managed
High cost
Regular monitoring required

To what extent and in which population(s) is the technology being used in your local health economy?
The technology is not currently being used as it is unlicensed.

- is there variation in how it is being used in your local health economy?
Unknown

- is it always used within its licensed indications? If not, under what circumstances does this occur?
Not aware of any unlicensed usage locally.

- what is the impact of the current use of the technology on resources?
Not applicable

- what is the outcome of any evaluations or audits of the use of the technology?
Not applicable

- what is your opinion on the appropriate use of the technology?
The technology is likely to have a place in therapy for patients with more severe, active SLE with seropositive disease who have not responded to conventional (best-practice) oral treatments (currently unlicensed for this indication).

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Potential impact on the NHS if NICE recommends the technology

What impact would the guidance have on the delivery of care for patients with this condition?
If the technology were to be approved this would provide a licensed product for the management of patients with severe SLE, who may have previously been treated with an unlicensed product. Currently there are no licensed products for the management of severe cases of SLE.

If the technology is proven to reduce the dose of corticosteroids, this would impact on associated costs of patients who are taking corticosteroids long-term. Due to the adverse effects of long-term treatment with steroids, patients are at risk of putting on weight, developing osteoporosis, developing type 2 diabetes mellitus, having mood disorders and cataracts. The cost of managing these conditions could be extremely high including medical and social care management. These combined factors could increase a patients' mortality risk.

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics?
The technology should be used by specialists – (specifically tertiary centres) due to the small numbers of patients who would be affected, to ensure appropriate use and also experience of this technology.

Would there be any requirements for additional resources (for example, staff, support services, facilities or equipment)?
There would be requirements for staff to administer and monitor this treatment and undertake any follow-up appointments required.
Patients would need to be admitted for day case in order to have the drug administered and the follow-up monitoring undertaken.

Can you estimate the likely budget impact? If this is not possible, please comment on what factors should be considered (for example, costs, and epidemiological and clinical assumptions)
It is estimated that approximately 4 cases per 100,000 annually of patients with SLE would be eligible for belimumab each year. However, locally it is suspected there will be fewer patient numbers.
Presumably these patients will be those that would previously have received rituximab. The cost of one cycle (2 x rituximab injections) is estimated to be £4300, in addition to a day case admission tariff of £814 (HD23c) (ex. VAT).
The technology may replace rituximab in the management of SLE, therefore would be administered by the specialist services in tertiary centres.
Costs of each will need to be compared when the belimumab drug cost is available.

Would implementing this technology have resource implications for other services (for example, the trade-off between using funds to buy more diabetes nurses versus more insulin pumps, or the loss of funds to other programmes)?
This would depend on the cost of the belimumab. It is likely nurses to administer etc. would already be available with in the tertiary centre.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single Technology Appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosis

Would there be any need for education and training of NHS staff?
The staff administering the medication would need education and training as this is a newly licensed drug.
Prescribers would need to be aware of the specific groups of patients this drug should be used in and clear audit around the use should also be undertaken by the specialists.

Equality

Are there any issues that require special attention in light of the NICE's duties to have due regard to the need to eliminate unlawful discrimination and promote equality and foster good relations between people with a characteristic protected by the equalities legislation and others?

- **None identified**

Other Issues

Please include here any other issues you would like the Appraisal Committee to consider when appraising this technology.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Thank you for agreeing to give us your views on the technology and the way it should be used in the NHS.

Patients and patient advocates can provide a unique perspective on the technology, which is not typically available from the published literature.

To help you give your views, we have provided a template. The questions are there as prompts to guide you. You do not have to answer every question. Please do not exceed the 8-page limit.

About you

Your name:
Jane Dunnage

Name of your organisation:

LUPUS UK

Are you (tick all that apply):

- **Yes** a patient with the condition for which NICE is considering this technology?
- a carer of a patient with the condition for which NICE is considering this technology?
- an employee of a patient organisation that represents patients with the condition for which NICE is considering the technology? If so, give your position in the organisation where appropriate (e.g. policy officer, trustee, member, etc)
- other? (please specify)

What do patients and/or carers consider to be the advantages and disadvantages of the technology for the condition?

1. Advantages
(a) Please list the specific aspect(s) of the condition that you expect the technology to help with. For each aspect you list please describe, if possible, what difference you expect the technology to make.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

i) some suppression of the over-active immune system causing lupus
ii) reduction in the length of time of lupus flare
iii) increase in the length of time between flares
iv) reduction in the amount of other medications taken, esp steroids and immunosuppressive medications:

(b) Please list any short-term and/or long-term benefits that patients expect to gain from using the technology. These might include the effect of the technology on:

- the course and/or outcome of the condition
- physical symptoms
- pain
- level of disability
- mental health
- quality of life (lifestyle, work, social functioning etc.)
- other quality of life issues not listed above
- other people (for example family, friends, employers)
- other issues not listed above.

i) this is currently an incurable condition, but suppression of the over-active immune system should lead to fewer and/or less severe symptoms and the ability to lead a less restricted life.

ii) Lupus patients are normally on a number of life-long medications, some of which are toxic or have serious side effects and consequences. It should decrease the amount and number of medications necessary, but may not cut them out altogether. In particular, methotrexate and thalidomide can affect fertility; steroids have serious side effects such as osteoporosis, cataracts, weight gain; immunosuppressive medications lead to greater susceptibility to generalised infections. Serious damage to organs is caused by not only the aggressive nature of the illness itself, but by the burden of the medications, especially if treatment is necessary for a long duration, as is often the case in children and young people and those diagnosed in their 20s (LUPUS UK survey found that over 1/3rd of patients were diagnosed before the age of 35)

iii) If this treatment is licensed by NICE it could reduce the number of medications necessary, this would have a financial benefit to patients as there would be less prescription costs.

iv) Women with lupus are at highly increased risk of CHD. Although the reason for this is not currently clear, earlier and more effective treatment of lupus is likely to reduce the number of deaths and serious complications (*see Haque article*).

v) Pain and fatigue are the most widely experienced symptoms (92% and 86% of patients respectively reported in LUPUS UK Members' survey): treatments which may reduce these and other symptoms are to be welcomed.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

vi) Many lupus patients have found it difficult to keep their jobs because of the fluctuating nature of the illness and fatigue: Belimumab has shown both a reduction in the length of lupus flares and an increase in time between flares, and this would enable many patients to remain in employment without taking large amounts of sick leave, which inevitably leads to dismissal, or reduction in working hours and subsequent financial difficulties. (LUPUS UK members' survey (*see attached summary*) showed 16% were able to work full time, 19% part time: of the 46% who were retired, 50% had had to retire on health grounds. 42% are receiving some type of benefit.)

vii) Less quantifiable is the restriction on parenting and family life which the illness can cause: it can be extremely difficult to give full attention and care to family when one is hampered by pain, fatigue and the unpredictable nature of the illness. Many lupus patients look to other family members for help with childcare but this does not alleviate the sadness and frustration that the patient feels being unable to fully play the role of a parent. (*see Hale article*)

viii) Nearly a third of lupus patients have mobility problems: 29% receive DLA for mobility, 84% of which are at the high level (*LUPUS UK Members' survey)*. Whilst it is not clear whether Belimumab will have a direct effect on mobility, if it keeps patients more active for longer periods of time this has got to be of great value not just to the patient, but also to the Benefits system.

ix) Depression can be a problem for many lupus patients (it was listed as the 3rd most difficult symptom to live with in LUPUS UK Members' survey), sometimes as part of the illness itself, but for many the isolation, loss of work, breakdown of relationships, changed visual appearance, lack of belief in them as having an illness or understanding their symptoms by many including the medical profession, family members and employers, will cause depression. Patients also often experience the depression of a long-term illness itself. If Belimumab works to improve some of these situations it could lead to improvement in the mental health of patients as they are less likely to have to give up aspirations for careers, family life etc.

x) A reduction in time attending medical appointments would be a likely improvement from this treatment: lupus patients need to be monitored regularly and keep many hospital and doctors' appointments. When they are experiencing a lupus flare, visits to A&E departments are often necessary, where staff do not have detailed knowledge of lupus. Waiting times are very high, exposing lupus patients to the risk of infection and stress in a difficult and uncomfortable environment when they are already experiencing pain and fatigue and other serious, chronic symptoms.

xi) Another regular frustration for lupus patients is that they are unable to predict how the illness will affect them on a daily basis: this leads to problems within the family and with friends, but also within school, college and employment as the patient is not able to play a full part in these 'normal' activities and resentment or ostracism can result.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

xii) Compliance with medications (because of side effects, severity of medication, amount and number or just forgetfulness) can be a problem. I understand this treatment will be given intravenously at hospital and that would make it much easier for patients.

What do patients and/or carers consider to be the advantages and disadvantages of the technology for the condition? (continued)

2. Disadvantages

Please list any problems with or concerns you have about the technology. Disadvantages might include:

- aspects of the condition that the technology cannot help with or might make worse.
- difficulties in taking or using the technology
- side effects (please describe which side effects patients might be willing to accept or tolerate and which would be difficult to accept or tolerate)
- impact on others (for example family, friends, employers)
- financial impact on the patient and/or their family (for example cost of travel needed to access the technology, or the cost of paying a carer

i)The long term nature of the technology on the patient is unclear

ii) It could trigger a worse reaction within the immune system

iii) It may cause some side effects (patients may be willing to put up with these if they are temporary or not too painful/serious)

iv) Some patients may find it difficult to attend hospital for the length of time necessary to have the infusion, but if this results in less hospital visits (esp in emergency) then they are likely to accept that.

v) This treatment works on the B cells: if the person's lupus is not caused by this pathway, it will be unlikely that they will see an improvement, and that could cause concerns that they are not receiving the 'right' treatment

3. Are there differences in opinion between patients about the usefulness or otherwise of this technology? If so, please describe them.

Many lupus patients are thrilled to hear of a new treatment especially for lupus and have great hopes that it will improve treatment and the outcome of the illness, if not for them, at least for other patients.
There may be concern or even suspicion about new technologies in a small number of patients, but many lupus patients have a good relationship with their consultant and look to them for their advice on treatment.
There is some concern about the funding for this new treatment and whether it will be readily available in all areas of the country: presumably NICE will give guidance on this.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

4. Are there any groups of patients who might benefit more from the technology than others? Are there any groups of patients who might benefit less from the technology than others?

Yes to both, this depends on the nature and progression of the illness itself and the organ/systems and characteristics experienced. This drug has not yet been trialled on some lupus manifestations such as CNS, kidney and skin. Other patients have a more 'chronic' nature to their illness and experience fewer flares, so this drug may not help them. Fatigue and pain are the most commonly experienced symptoms, and there is not yet strong evidence about how Belimumab will affect these symptoms.

Comparing the technology with alternative available treatments or technologies

NICE is interested in your views on how the technology compares with with existing treatments for this condition in the UK.

(i) Please list any current standard practice (alternatives if any) used in the UK.
NSAID
Hydroxychloroquine
Steroids
Various immunosuppressive medications

(ii) If you think that the new technology has any **advantages** for patients over other current standard practice, please describe them. Advantages might include:

- improvement in the condition overall
- improvement in certain aspects of the condition
- ease of use (for example tablets rather than injection)
- where the technology has to be used (for example at home rather than in hospital)
- side effects (please describe nature and number of problems, frequency, duration, severity etc.)

The technology shows promise in reducing the length of time of flares and increasing the time in between flares.

If it reduces the need for steroids and immunosuppressive medications then this will be a big benefit esp with the side effects experienced from these serious medications and sometimes the need to take further medication to alleviate side effects from the more serious ones. (see earlier comments for more detail)

Some patients will find it easier to have IV injections rather than taking many medications on a daily basis, partly because of difficulties in remembering to take medications and the cost of prescriptions for the many medications necessary on a daily basis.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

Reduction in time attending medical appointments would also be an advantage.

(iii) If you think that the new technology has any **disadvantages** for patients compared with current standard practice, please describe them. Disadvantages might include:

- worsening of the condition overall
- worsening of specific aspects of the condition
- difficulty in use (for example injection rather than tablets)
- where the technology has to be used (for example in hospital rather than at home)
- side effects (for example nature or number of problems, how often, for how long, how severe).

It is difficult to predict how individual patients will react to the technology: there have been some side effects in the trials and there may be reactions to either the technology itself or the site of injections: patients will have to discuss with their consultant whether the side effects are more serious than the illness itself, or whether they are tolerable compared to the progress of the illness.

If the technology is administered within a hospital setting this will not make it difficult for the patient to use, if it eventually needs to be injected by the patient then this may be a problem for some. Hospital visits may be difficult for some if there is a long journey to be made, but if it reduces the number of hospital visits because it is more effective than existing treatments this will also be better news for patients.

Research evidence on patient or carer views of the technology

If you are familiar with the evidence base for the technology, please comment on whether patients' experience of using the technology as part of their routine NHS care reflects that observed under clinical trial conditions.

As far as I am aware it is not yet being used within the NHS. I know of several lupus patients who have received treatment with Rituximab, which I believe may have some similarities, and they have found that it has stabilised their condition and have been able to resume employment.

Are there any adverse effects that were not apparent in the clinical trials but have come to light since, during routine NHS care?

Not aware of its use in NHS currently.

Training needs to be given to those medical professionals who will monitor lupus patients during this treatment: the number of systems affected, fluctuating nature of the illness and the symptoms needs to be understood in order to monitor patient's response to the treatment. There are a number of indices which monitor lupus

damage and progress of the condition, but there will need to be training in use and interpretation.

NICE also needs to give clear guidance on its use especially on funding / commissioning decisions.

Are you aware of any research carried out on patient or carer views of the condition or existing treatments that is relevant to an appraisal of this technology? If yes, please provide references to the relevant studies.

LUPUS UK surveyed its members on their experience of the condition and a summary of the results is attached and have been referred to above. I have also attached and article by Liz Hale which raises important issues on quality of life, especially family life.

Availability of this technology to patients in the NHS

What key differences, if any, would it make to patients and/or carers if this technology was made available on the NHS?

I believe I have covered these earlier in the statement, particularly those listed under advantages. Whilst discussing her difficulties in living with lupus, one patient (among many) has said that the hospital has become her social life as she is there so often, and has no time or energy for social activities.

What implications would it have for patients and/or carers if the technology was **not** made available to patients on the NHS?

Many people would continue having their lives seriously curtailed by this condition, because of the severity of the illness itself and the serious side effects of current medications (many patients are on a cocktail of drugs to try to reduce the impact of the steroids which were the main treatment until around 10 years ago). This illness has cost many older lupus patients their jobs, their relationships and their other hard-earned plans for the future. Now that more younger people are being diagnosed with lupus it would be a great tragedy if they also had to cope with the damage due to the aggressive nature of the condition, alongside the current heavy burden of treatments which are known to cause serious side effects and restrictions.

Are there groups of patients that have difficulties using the technology?

As this is not yet available on the NHS, we are not able to answer this question. I am aware from the trials that some patients experienced skin irritation at the site of the infusion, but this would be a temporary difficulty, which most would tolerate if the treatment was beneficial in the long term.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

Other Issues

Please include here any other issues you would like the Appraisal Committee to consider when appraising this technology.

Training needs to be given to those medical professionals who will monitor lupus patients during this treatment: this is not an easy illness to cope with, and the fluctuating nature of the illness and symptoms needs to be understood and interpreted well in order to monitor patients' response to the treatment

Lupus affects people of all races, but a higher number of people from some ethnic minorities are affected, and often more seriously. This can lead to a variety of cultural issues in different races such as ability to have children, visual appearance (especially where steroids lead to excessive weight gain or permanent scarring leads to pigmentation differences) and lack of attraction leading to poor marriage prospects, suspicions about the effects of taking some medications, difficulties in explaining the illness to family members and employers.

LUPUS UK Members survey summary

Background and Aim

We at LUPUS UK, the national charity for people with lupus, wanted to gain a better picture of the nature of lupus and how it affects our members**.** This should inform the medical profession of the work of the charity and improve their knowledge of the illness.

During summer 2009 a questionnaire was posted to all 5700 members of LUPUS UK asking them to provide information of how lupus impacts on their lives. The immediate response was overwhelming. Over fifty percent of questionnaires were returned (3073 questionnaires) in 2 months. Excluding those incomplete questionnaires and any completed by members without lupus, 3017 questionnaires contributed to the survey results. [1]

Summary of findings

- It takes, on average, more than 7 years from symptom onset until SLE is diagnosed.
  - This average delay has not changed over a 20 year period
- More than 45% of patients are initially given a diagnosis other than SLE.
- Patients report that the most difficult symptoms to live with are fatigue, joint pain and depression.
- Two-thirds of patients report sunlight as a triggering factor
  - One quarter report that their lupus symptoms are made worse by fluorescent lighting
- On the day of the survey:
  - 87% had problems with pain/discomfort
  - 30% had problems with self-care
- 42% of all patients were receiving some kind of benefits
  - 24% were receiving higher level mobility allowance.
- Half of lupus patients who retire do so because of ill-health
- Rheumatology is the main speciality providing care for SLE in the UK

[1] Design and analysis of questionnaires completed by The University of Manchester Epidemiology Unit

Profile of respondents

Gender

Lupus affects around nine times as many women as men.

- 94% of respondents were female and 6% were male.

Ethnicity

Lupus appears in people of all races, but is more prevalent in people from black and ethnic minority races.

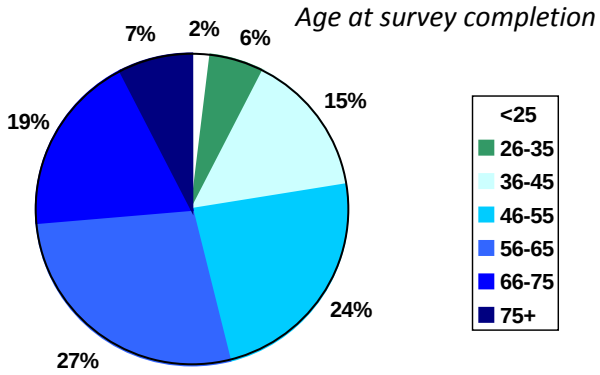
- 94% of respondents classified themselves as white, 98% of men and 94% of women[2].

Living status

- 78% of respondents live with their family or a partner, 88% are male and 77% female.
- Twice as many women (20%) compared with men (11%) are living alone.

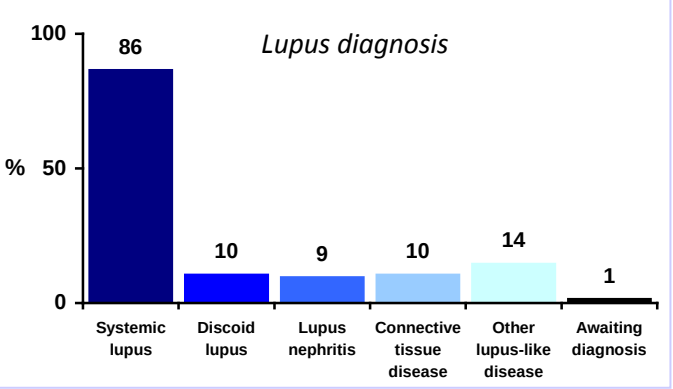
Age

- 50% of respondents were aged between 46 and 65 years at the time of the survey.



Details of diagnosis

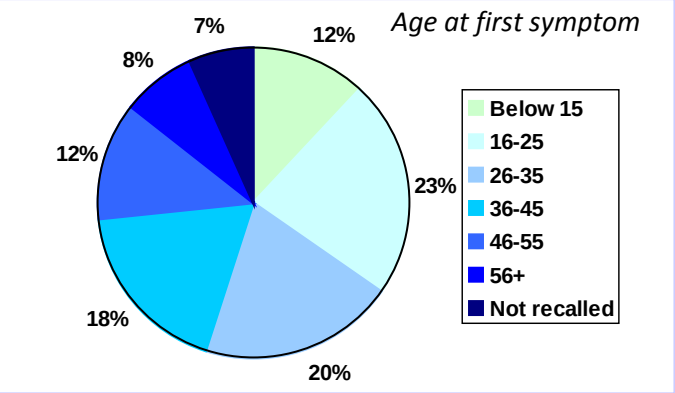
- 86% of respondents are diagnosed with Systemic Lupus Erythematosus (SLE)[3]



[2] Although many black and ethnic minorities were represented, the numbers were too small to represent the data by ethnic group
[3] Some patients will be diagnosed with more than one type of lupus

Age at first symptom

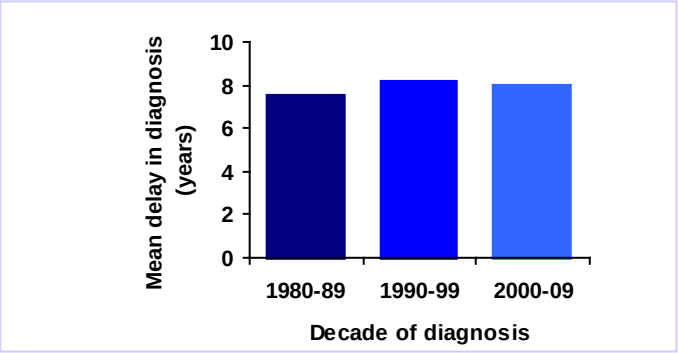
- More than 30% of respondents experienced their first symptom before the age of 26.



Timing of diagnosis

- The mean number of years between experiencing first symptom and receiving diagnosis is 7½ years.
  - men experience a shorter delay in diagnosis (5 years) than women (8 years).

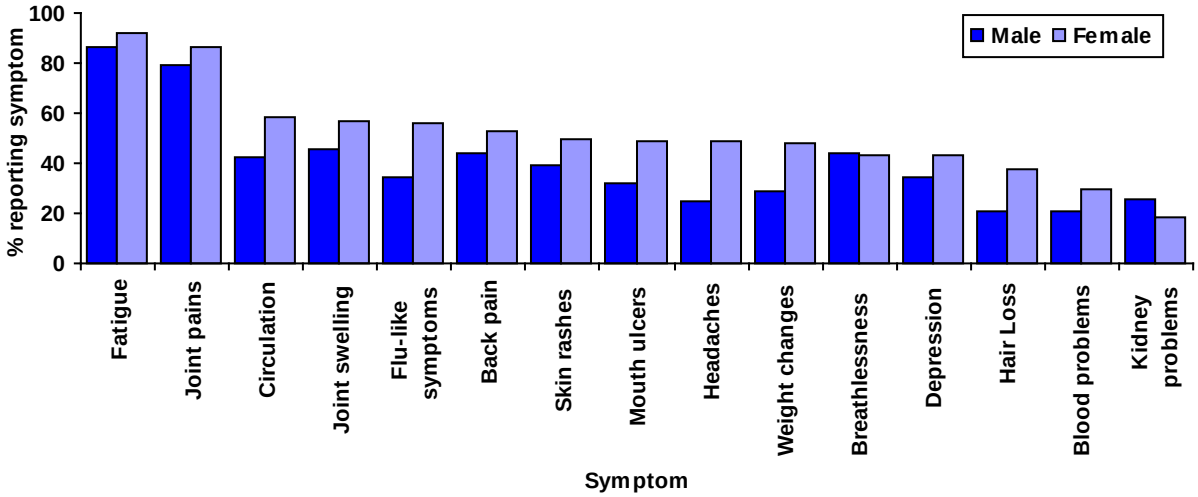
- The delay in diagnosis has stayed the same over a number of decades



- 45% of respondents were initially incorrectly diagnosed with another condition.
  - Over 30 different illnesses were cited.

Symptoms

- Respondents are regularly affected by many symptoms
  - Of a selection of 15 common symptoms/problems, 54% experience 6-10 symptoms and 20% have 11-15 symptoms on a regular basis



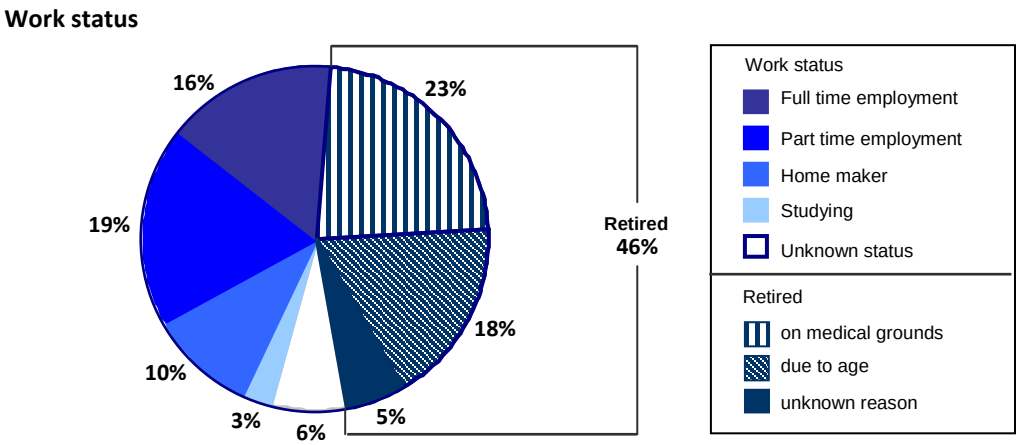
- When patients ranked which symptoms were most difficult to live with, fatigue (79%) and joint pain (63%) were most commonly listed.

Economic Status

Lupus has had a major impact on respondents' lives.

Work status

- Less than a fifth are working full time
- Approximately one fifth are in part-time employment
- Approximately a quarter of all respondents have had to retire on medical grounds.



- 42% of respondents receive benefits
- 28% are on mobility allowance of which 84% receive the higher level of mobility allowance
- Over a third are in receipt of Disabled Living Allowance
- Just under a fifth are receiving Incapacity Benefit

Further detailed information on other aspects such as triggers of lupus activity, helpful treatments, relations with medical professionals and other medical conditions experienced, can be obtained from the full survey, published by LUPUS UK, St James House, Eastern Avenue, Romford, Essex, RM1 3NH tel: 01708 731251 www.lupusuk.org.uk

NICE Patient statement attachment

LUPUS UK: National Survey 2010

Overview of Survey Results:

Survey Overview

5700 questionnaires were mailed to members of Lupus UK in Summer 2009. Overall 3017 (53%) returned a valid survey questionnaire. The mean (SD) age at survey date was 55.7 (13.8) yrs. 2832 (94%) respondents were female and 185 (6.1%) were male (Table 1). 2837 (94%) of respondents classified themselves as White, 96 (3.3%) were Black (Caribbean, African or other) and 42 (2%) were of Indo-Asian origin; 7.5% (218) of respondents were born outside the UK.

According to the 2001 census data, 92.1% of the UK population is white. SLE is more common in non-white ethnicities in the UK. This relatively lower proportion of non-white respondents may reflect either the membership of Lupus UK or perhaps the fact that this survey was only circulated in the English language.

Age of onset of SLE and Diagnostic Delay

The mean (SD) age at diagnosis was 40.6 (13.9) years while the mean (SD) age at first symptoms reported was 32.7 (18.1) years.

477 (16%) respondents were first diagnosed before the age of 25 and 1096 (36.3%) in total were less than 35 years old at diagnosis.

Men tended to be older at first symptom onset and diagnosis such that 41% of men and 56% of women had their first symptom before age 35 year old. In addition, 30% of men and 37% of women were diagnosed before age 35 years old.

Delay in diagnosis was calculated from time from 1st symptom to diagnosis. The mean (SD) overall lag time from first patient-reported symptoms to diagnosis was 7.7 (10.5) years. Women are more likely to experience a delay in diagnosis than men 7.8 (10.5) vs 5.4 (10.0) yrs (Table 5).

Patients with SLE and nephritis, experience the shortest delay in diagnosis, with overall mean delay of 4.8 (8.1) years; 2.6 (4.6) years in men and 5.0 (8.3) years in women, suggesting that in the SLE disease process, nephritis is often an earlier manifestation or leads to a quicker referral to consultant.

Table 8 shows the delay in diagnosis by the decade of when the diagnosis was made. Over the last 3 decades, the length of time taken from 1st symptom to diagnosis has not changed being still 7.9 (11.3) years in the period 2000-9.

Living and Work status:

Over 77% of respondents were living with partners or families. Twice as many women live alone 20% (605/2832) vs 11% (20/ 185) of men.

Of the respondents, 16% were working full time, 19% part-time, with more men working full-time (28% vs 15%) and more women working part-time (20% vs 7%) (Table 2). 81 (2.7%) respondents were students with two-thirds studying part time. 1380 (46%) were retired and of those who had retired, 50% had retired due to ill health. 1271 (42%) patients were receiving some kind of benefit and 1020 (34%) were claiming disabled living allowance (DLA). 865 (29%) were claiming DLA mobility allowance, of these 84% claiming the higher level of benefit. Therefore 727 (24%) respondents were on the higher level DLA mobility benefit. A higher proportion of men were claiming incapacity benefit*(as it was termed at the date of the survey) (22%) compared to women (17%).

Lupus and related diagnoses:

2577 (85%) respondents indicated Systemic Lupus Erythematosus (SLE) as their primary diagnosis. There were a number of other key diagnoses or overlapping diagnostic labels being used such as discoid lupus (DLE) in 315 (10.4%), 'connective tissue disease' in 314 (10.4%) and lupus nephritis in 267 (8.9%) respondents. In the entire group, 2065 (68.5%) had SLE as their only lupus-related diagnosis. 149 (5%) had DLE only, 114 (3.8%) had 'connective tissue disease' as their only diagnosis and in 41 (1.4%) lupus nephritis was their only lupus-related diagnosis. 522 (17.4%) patients had 2 or more lupus-related diagnoses.

Diagnoses reported by patients prior to their SLE diagnosis.

Before being diagnosed with SLE, 1353 (44.9%) patients had been given at least one other diagnosis. In 450 (33.5%), the diagnosis first given was rheumatoid arthritis. 30 (2.2%) had been originally diagnosed with a 'connective tissue disease' and 15 (1.1%) had a diagnosis of 'renal disease'. Overall approximately 38% of these patients had a diagnosis that could be related to 1-2 SLE criteria and may represent a disease in evolution. Importantly, 144 (4.4%) had a first diagnosis of fibromyalgia or chronic fatigue syndrome. Altogether 3.6% also had some diagnostic label related to 'psychological' or a hypochondriacal disorder and 34 (1.1%) were originally diagnosed with multiple sclerosis.

Co-morbidities and Current Symptoms and Triggering Factors:

Regarding other illnesses concurrent with SLE (co-morbidities) from those respondents, 282 (9%) also have thyroid disorders, 187 (6%) report concurrent RA and 174 (5.8%) also report being diagnosed with osteoporosis. In addition, 125 (4%) had cardiovascular disease, 107 (3.6%) had diabetes mellitus and 98 (3.3%) had a cancer.

Table 11 shows the symptoms most affecting the respondents. Fatigue affects 2759 (92%) patients and was the symptom most frequently reported; in addition 2596 (86.1%) report joint pain. Also at least half of all patients report poor circulation (n=1732 [57%]), flu-like symptoms (n=1661 [55.1%]), back pain (n=1583 [52.5%]) and rashes (n=1481 [49.1%]). Altogether, 1630 (54%) frequently experience 6-10 different symptoms and 576 (19.1%) suffer more that 10 different symptoms on a frequent basis.

In patients under 25 year old (n=55) a similar spread of symptoms is seen and 36 (69%) in this age group also report 6-15 frequently experienced symptoms.

We also asked patients to rank their 3 most debilitating symptoms. Table 13 shows the symptoms that have been ranked by the respondent as the top 3 symptoms that are most difficult to live with. Overall respondents rated fatigue (n=2378[79%]) and joint pain (n=1897 [63%]) to be the most difficult symptoms to live with and these two symptoms are also rated the highest amongst males and females. 24% (45/185) of men find breathlessness and 16% (465/2832) women also found their depression, to be the third most difficult symptom to live with. Similarly, in patients under 25 years old (n=55), 38(71%) felt fatigue was their most debilitating symptom followed by joint pain in 35(64%).

Table 15 reports those environmental conditions respondents indicated that makes their lupus worse. 1892 (63%) of respondents reported sunlight, along with 1540 (51%) reporting cold and 1270 (42%) heat make their Lupus worse. 819 (27.2%) also reported worsening of symptoms with fluorescent lighting.

The distribution across age groups was similar for the three conditions (Table 16). However, the over 75 year olds (n=225) reported heat (n=96 [42.7%]) to be the second most common condition to make their Lupus worse rather than the cold.

General Health Status:

Table 17 indicates the level of general health of the respondents on the day of the survey (using the General Health Questionnaire). 43% (1289) of respondents reported some problems with mobility. 918 (30%) had some problems with self-care including washing and dressing, 70% (2116) had some problems with performing usual activities and 87% (2616) experience problems with pain or discomfort of whom 789 (26%) had severe pain or discomfort at the time of survey. Regarding mental health status, 2170 (72%) had problems with feeling anxious or depressed.

When respondents rated symptoms on a scale from 0 to 100, the mean (SD) pain from lupus in the past week was 45.3 (28.4) and from lupus fatigue in the past week was 61.3 (28.3).

Care provision:

Table 18 shows the type of consultant respondents are currently seeing, with 2388 (79%) seeing a rheumatologist, 423 (14%) a dermatologist, 305 (10%) a nephrologist and 110 (4%) an immunologist. Only 504 (17%) respondents are seeing more than one consultant.

Table 19 indicates the proportion of consultants the respondents had previously seen. 64% have seen a rheumatologist, 36% a dermatologist, 13% a nephrologist and 11% an immunologist. 1063 (35%) respondents have previously seen more than one consultant.

We also asked respondents to rate, on a visual analogue scale (VAS) from 0-100 the level of support they receive from various social and medical contacts and we also categorised these into tertiles as low, medium and high. Respondents report receiving high levels of support from their partner: the mean (SD) for support was 79.4 (30.7). Females rate the level of support form their partner/spouse at a lower level than men. The mean (SD) level of support was 78.7 (31.1) reported by women and 89.7 (22.8) reported by men. 2071 (68.6%) respondents rated a high level of support from their consultant and 1842 (61.1%) have high levels of support from their family members. In addition, GP support was rated as high by 1604 (53.2%).

Medications and Complimentary and Alternative Medications (CAMs).

Table 23 indicates the current medication taken for lupus as recalled by the respondent. 2051 (68%) respondents are taking corticosteroids, 1820 (60%) anti-malarials and 1202 (40%) immunosuppressive agents.

506 (16.8%) respondents are also taking alternative therapy. This includes 498 (17.6%) women and 8 (4%) men. Acupuncture is the most frequently used CAM with 123 (4.1%) patients using this. Massage and reflexology were the next 2 most popular (n=93 [3.1%] and 79 [2.6%] respectively).

Table 1: Demographic data from 3017 individuals included in analysis

Demographic variable		**Men n=185 (6.1%)**	**Women n=2832 (93.9%)**	**Total n=3017**
Age at survey date (yrs) mean (SD)		59.1 (13.7)	55.5 (13.7)	55.7 (13.8)
Age group at survey n(%)	0-15	0	4 (0.01)	4 (0.1)
	16-25	8 (4.3)	43 (1.5)	51 (1.7)
	26-35	6 (3.2)	162 (5.7)	168 (5.6)
	36-45	14 (7.6)	440 (15.5)	454 (15.1)
	46-55	33 (17.8)	680 (24.0)	713 (23.6)
	56-65	56 (30.3)	779 (27.5)	835 (27.7)
	66-75	53 (28.7)	514 (18.2)	567 (18.8)
	75+	15 (8.1)	210 (7.4)	225 (7.5)
Age at 1st symptom(yrs) mean (SD)		38.7 (16.9)	32.3 (14.9)	32.7 (15.1)
Age group at 1st symptom n(%)	0-15	21 (11.4)	338 (11.9)	359 (11.9)
	16-25	24 (13.0)	664 (23.5)	688 (22.8)
	26-35	30 (16.2)	580 (20.5)	610 (20.2)
	36-45	39 (21.1)	518 (18.3)	557 (18.5)
	46-55	30 (16.2)	341 (12.0)	371 (12.3)
	56-65	23 (12.4)	148 (5.2)	171 (5.7)
	66-75	6 (3.2)	45 (1.6)	51 (1.7)
	75+	2 (1.1)	3 (0.1)	5 (0.2)
Age at 1st diagnosis (yrs) mean (SD)		44.3 (15.9)	40.4 (13.7)	40.6 (13.9)
Age group at 1st diagnosis n(%)	0-15	8 (4.3)	58 (2.1)	66 (2.2)
	16-25	20 (10.8)	391 (13.8)	411 (13.6)
	26-35	28 (15.1)	591 (20.9)	619 (20.5)
	36-45	25 (13.5)	735(26.0)	760 (25.2)
	46-55	53 (28.7)	616 (21.8)	669 (22.2)
	56-65	35 (18.9)	295 (10.4)	330 (10.9)
	66-75	10 (5.4)	91 (3.2)	101 (3.4)
	75+	3 (1.6)	8 (0.3)	11 (0.4)
Ethnicity n(%)	White	182 (98.4)	2655 (93.8)	2837 (94.0)
	Black-African	0	17 (0.6)	17 (0.6)
	Black-Caribbean	0	51 (1.8)	51 (1.7)
	Black-British	0	21 (0.7)	21 (0.7)
	Black-other	0	7 (0.3)	7 (0.3)
	Indian	2 (1.1)	33 (1.2)	35 (1.2)
	Pakistani	0	6 (0.2)	6 (0.2)
	Bangladeshi	0	1 (0.04)	1 (0.03)
	Chinese	0	13 (0.5)	13 (0.4)
	White Asian	1 (0.5)	10 (0.4)	11 (0.4)
	White Caribbean	0	2 (0.07)	2 (0.07)
	Other mixed race	0	6 (0.2)	6 (0.02)
	Other	0	1 (0.04)	1 (0.03)
	Did not indicate	0	9 (0.3)	9 (0.3)
Country of origin n(%)	East Africa	2 (1.1)	14 (0.5)	16 (0.5)
	South America	0	5 (0.2)	5 (0.2)

	Australia	0	12 (0.4)	12 (0.4)
	Caribbean	0	39 (1.4)	39 (1.4)
	UK	179 (96.8)	2620 (92.5)	2799 (92.3)
	Eastern Europe	0	5 (0.2)	5 (0.2)
	West Africa	0	7 (0.3)	7 (0.2)
	North America	1 (0.5)	10 (0.4)	11 (0.4)
	South East Asia	0	22 (0.8)	22 (0.7)
	Western Europe	1 (0.5)	24 (0.9)	25 (0.8)
	Scandinavia	0	6 (0.2)	6 (0.2)
	Middle East	0	3 (0.1)	3 (0.1)
	India	1 (0.5)	19 (0.7)	20 (0.7)
	Pakistan	0	3 (0.1)	3 (0.1)
	South Africa	0	9 (0.3)	9 (0.3)

Table 2: Work status and benefits by gender

Work & benefit status n(%)		**Men n=185**	**Women n=2832**	**Total n=3017**
Working full time paid		51 (27.6)	425 (15.0)	476 (15.8)
Working part-time paid		13 (7.0)	552 (19.5)	565 (18.7)
Working in the home		5 (2.7)	303 (10.7	308 (10.2)
Studying		4 (2.2)	77 (2.7)	81 (2.7)
Of those studying (n=81)	Full-time	3 (75.0)	26 (33.8)	29 (35.8)
	Part-time	1 (25.0)	48 (62.3)	49 (60.5)
Retired		107 (57.8)	1273 (45.0)	1380 (45.7)
Of those retired (n=1380)	Age	41 (38.3)	491 (38.6)	532 (38.6)
	Medical grounds	55 (51.4)	625 (49.1)	680 (49.3)
Receiving benefits		75 (40.5)	1196 (42.2)	1271 (42.1)
Claiming DLA		53 (28.7)	967 (34.2)	1020 (33.8)
Claiming DLA care		33 (17.8)	762 (26.9)	795 (26.4)
Of those claiming (n=795) Level:	Low	11 (33.3)	322 (42.3)	333 (41.9)
	Intermediate	12 (36.4)	248 (32.6)	260 (32.7)
	High	8 (24.2)	178 (23.4)	186 (23.4)
Claiming DLA mobility		49 (26.5)	816 (28.8)	865 (28.7)
Of those claiming (n=865) Level:	Low	5 (10.2)	91 (11.2)	96 (11.1)
	High	42 (85.7)	685 (84.0)	727 (84.1)
Claiming Incapacity benefit		41 (22.2)	476 (16.8)	517 (17.1)
Claiming Attendance allowance		15 (8.1)	140 (4.9)	155 (5.1)
Of those claiming (n=155) Level:	Low	6 (40.0)	65 (46.4)	71 (45.8)
	High	7 (46.7)	62 (44.3)	69 (44.5)

Table 5: Delay in diagnosis by gender

Delay in diagnosis	**Men n=185**	**Women n=2832**	**Total n=3017**
Delay in diagnosis (yrs) Mean (SD) (From 1st symptom diagnosis)	5.4 (10.0)	7.8 (10.5)	7.6 (10.5)

Table 8: Delay in diagnosis by decade of diagnosis

Delay in diagnosis (yrs) Mean (SD) (From 1st symptom	**1980-89 (n=654)**	**1990-99 (n=1040)**	**2000-09 (n=948)**	**Total (n=2642)**
	6.6 (8.8)	8.2 (10.8)	7.9 (11.3)	7.7 (10.5)

to diagnosis)	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
	5.6 (10.4)	6.7 (8.7)	4.8 (8.5)	8.4 (10.9)	6.0 (11.4)	8.1 (11.3)	5.4 (10.1)	7.9 (10.6)

Table 11: Reported symptom frequently suffering from by gender

Symptom n (%)		**Men n=185**	**Women n=2832**	**Total n=3017**
Fatigue		160 (86.5)	2599 (91.8)	2759 (91.5)
Joint pain		147 (79.5)	2449 (86.5)	2596 (86.1)
Poor circulation		78 (42.2)	1654 (58.4)	1732 (57.4)
Joint swelling		84 (45.4)	1602 (56.6)	1686 (55.9)
Flu-like symptoms		64 (34.6)	1597 (56.4)	1661 (55.1)
Back pain		82 (44.3)	1501 (53.0)	1583 (52.5)
Rash		73 (39.5)	1408 (49.7)	1481 (49.1)
Ulcers		60 (32.4)	1373 (48.5)	1433 (47.5)
Headache		46 (24.9)	1384 (48.9)	1430 (47.4)
Weight change		53 (28.7)	1354 (47.8)	1407 (46.6)
Breathlessness		81 (43.8)	1217 (43.0)	1298 (43.0)
Depression		64 (34.6)	1216 (42.9)	1280 (42.4)
Hair-loss		39 (21.1)	1057 (37.3)	1096 (36.3)
Blood		39 (21.1)	842 (29.7)	881 (29.2)
Kidney problems		47 (25.4)	527 (18.6)	574 (19.0)
Symptoms frequently suffering from	0	5 (2.7)	35 (1.2)	40 (1.3)
	1-5	82 (44.3)	689 (24.3)	771 (25.6)
	6-10	80 (43.2)	1550 (54.7)	1630 (54.0)
	11-15	18 (9.7)	558 (19.7)	576 (19.1)

Table 13: Symptom most difficult to live with ranked in the top 3 by individual by gender

Symptom ranked in top 3	**Men n=185**	**Women n=2832**	**Total n=3017**
Fatigue	125 (67.6)	2253 (79.6)	2378 (78.8)
Joint pain	107 (57.8)	1790 (63.2)	1897 (62.9)
Depression	23 (12.4)	465 (16.4)	488 (16.2)
Headache	14 (7.6)	420 (14.8)	434 (14.4)
Back pain	25 (13.5)	358 (12.6)	383 (12.7)
Joint swelling	22 (11.9)	361 (12.8)	383 (12.7)
Poor circulation	23 (12.4)	354 (12.5)	377 (12.5)
Flu-like symptoms	15 (8.1)	359 (12.7)	374 (12.4)
Breathlessness	45 (24.3)	311 (11.0)	356 (11.8)
Weight change	11 (6.0)	248 (8.8)	259 (8.6)
Rash	23 (12.4)	225 (7.9)	248 (8.2)
Ulcers	10 (5.4)	157 (5.5)	167 (5.5)
Hair-loss	4 (2.2)	135 (4.8)	139 (4.6)
Kidney problems	11 (6.0)	123 (4.3)	134 (4.4)
Haematological	7 (3.8)	81 (2.9)	88 (2.9)

NB 66 other symptoms ranked; 12 no symptoms at present

Table 17: General health by gender group

General health variable		Men n=185	Women n=2832	Total n=3017
Mobility	-No problems	103 (55.7)	1554 (54.8)	1657 (54.9)
	- Some problems	80 (43.2)	119 0 (42.0)	1270 (42.1)
	- Severe problems	1 (0.5)	18 (0.6)	19 (0.6)
Self-care	- No problems	138 (74.6)	1895 (66.9)	2033 (67.4)
	- Some problems	43 (23.2)	805 (28.4)	848 (28.1)
	- Severe problems	2 (1.1)	68 (2.4)	70 (2.3)
Carry out usual activities	- No problems	61 (33.0)	787 (27.8)	848 (28.1)
	- Some problems	95 (51.4)	1520 (53.7)	1615 (53.5)
	- Severe problems	28 (15.1)	473 (16.7)	501 (16.6)
Problems with pain/discomfort	- No problems	28 (15.1)	309 (10.9)	337 (11.2)
	- Some problems	121 (65.4)	1750 (61.8)	1871 (62.0)
	- Severe problems	32 (17.3)	713 (25.2)	745 (24.7)
Anxiety/depression	- No problems	66 (35.7)	723 (25.5)	789 (26.2)
	- Some problems	89 (48.1)	1476 (52.1)	1565 (51.9)
	- Severe problems	27 (14.6)	578 (20.4)	605 (20.1)
Pain from Lupus in past week (VAS 0-100) Mean (SD)		39.3 (29.7)	45.7 (28.3)	45.3 (28.4)
Fatigue from Lupus in past week (VAS 0-100) Mean (SD)		52.4 (30.5)	61.9 (28.0)	61.3 (28.3)

Lupus (2010) 19, 1118—1124.

Epidemiologic, socioeconomic and psychosocial aspects in lupus erythematosus

1. E. Aberer
    1. Department of Dermatology and Venerology, Medical University of Graz Auenbrugger Platz 8, A-8036 Graz, Austria, elisabeth.aberer@medunigraz.at

Abstract

Epidemiologic, socioeconomic, and psychosocial factors play an important role in health care and handling of patients with the various clinical forms of lupus erythematosus (LE). Patients with LE are mostly young women; adolescents and some ethnic groups are especially prone to a severe course of disease. The unpredictable and fluctuating flares of disease, the need for longterm treatment, and the side effects and damage caused by the disease itself severely reduce quality of life. Problems arise, involving family members, adherence to medical advice and therapy, communication and self management. Socioeconomically, patients are often unable to take regular employment and to pay for health insurance. Stress factors that arise have a negative impact on the course of disease, increasing both fatigue and the basic burden of illness. Healthcare professionals must pay careful attention to all these items, as they attempt to treat flares, minimize drug side effects, provide pain relief, arrange communication and exercise programs along with behavioral and psychosocial interventions in multidisciplinary cooperation, and also involve and support family members. Lupus (2010) 19, 1118—1124.

Rheumatology (2011) 50 (3): 431-433.

Meeting the needs of children who have parents with chronic inflammatory musculoskeletal diseases

Meeting the needs of children who have parents with chronic inflammatory musculoskeletal diseases

1. Elizabeth D. Hale[1],
2. Marianne Morris[2],
3. Sarah Hewlett[3,4],
4. Gareth J. Treharne[1,5] and

5. George D. Kitas[1]

+ Author Affiliations

1. [1]Department of Rheumatology, Dudley Group of Hospitals NHS Trust, Dudley, [2]Faculty of Health and Life Sciences, [3]Faculty of Health and Social Care, University of the West of England, [4]Department of Rheumatology, University Hospitals Bristol, Bristol, UK and [5]Department of Psychology, University of Otago, Dunedin, New Zealand

1. Correspondence to: Elizabeth D. Hale, Department of Rheumatology, Clinical Research Unit, North Block, Russells Hall Hospital, Dudley DY1 2HQ, UK. E-mail: liz.hale@dgoh.nhs.uk

Accepted May 20, 2010.

How can children be adequately supported?

Chronic inflammatory musculoskeletal diseases, such as RA, SLE and AS, affect millions of people worldwide, many of whom either are of parenting age or will already be parents at the time of disease onset. Living with a chronic illness involves adapting to a physically changing body while re-negotiating established roles (e.g. mother or father). Evidence-based information delivered at appropriate times during the course of illness is important to help such individuals cope with these changes. Moreover, a parent's rheumatic disease may have a major effect on the children (defined here as <18 years of age) and the family unit. Although we, and others, have examined the effects on adult patients in the family context [1, 2], no one has adequately established the effect from the children's perspective, nor is there any information or education resources to help children cope with their parent's illness. The type of information and format that would be useful to these children remain unknown, and many related research questions (such as the amount and degree of care that children and young people provide to patients and families) remain unanswered.

In studies about health-care experiences of women with SLE, it was apparent that the effect of the disease on the family unit was a major concern [3–5]. Children had to adapt to the fluctuating nature of the parent's disease and tolerate being cared for by people outside the immediate family unit. They also took care of themselves, their mothers and other family members.

Patients felt that no educational …

[Full Text of this Article]

Risk Factors for Clinical Coronary Heart Disease in Systemic Lupus Erythematosus: The Lupus and Atherosclerosis Evaluation of Risk (LASER) Study

1. SAHENA HAQUE,
2. CAROLINE GORDON,
3. DAVID ISENBERG,
4. ANISUR RAHMAN,
5. PETER LANYON,
6. AUBREY BELL,
7. PAUL EMERY,
8. NEIL McHUGH,
9. LEE SUAN TEH,
10. DAVID G.I. SCOTT,
11. MOHAMED AKIL,
12. SOPHIA NAZ,
13. JACQUELINE ANDREWS,
14. BRIDGET GRIFFITHS,
15. HELEN HARRIS,
16. HAZEM YOUSSEF,
17. JOHN McLAREN,
18. VERONICA TOESCU,
19. VINODH DEVAKUMAR,
20. JAMAL TEIR and
21. IAN N. BRUCE

+ Author Affiliations

1. From the arc Epidemiology Unit, School of Translational Medicine, University of Manchester; The Kellgren Centre for Rheumatology, Central Manchester and Manchester Children's University Hospitals NHS Trust, Manchester; Division of Immunity and Infection, University of Birmingham; Centre for Rheumatology, University College London; Department of Rheumatology, Queen's Medical Centre, Nottingham, UK; Department of Therapeutics and Pharmacology, Queen's University, Belfast, Ireland; Academic Unit of Musculoskeletal Disease, University of Leeds; Royal National Hospital for Rheumatic Diseases, Bath; Department of Rheumatology, Royal Blackburn Hospital; Department of Rheumatology, Norfolk and Norwich University Hospital; Sheffield Centre for Rheumatic Diseases, Royal Hallamshire Hospital, Sheffield; Department of Rheumatology, Northwick Park Hospital, The North West London Hospitals NHS Trust; Department of Rheumatology, Freeman Hospital, Newcastle; Department of Rheumatology, Cameron Hospital, Fife; Department of Rheumatology, Aberdeen Royal Infirmary; and Department of Rheumatology, Sunderland Royal Hospital.
2. S. Haque, MRCP, arc Epidemiology Unit, University of Manchester; C. Gordon, MD, Division of Immunity and Infection, University of Birmingham; D. Isenberg, PhD, Centre for Rheumatology, University College Hospital London; A. Rahman, Medicine, University College London; P. Lanyon, PhD, Rheumatology, Queen's Medical Centre; A. Bell, MD, Rheumatology, Queen's University; P. Emery, MD, Academic Unit of Musculoskeletal

Disease, Chapel Allerton Hospital; N. McHugh, MD, Rheumatology, Royal National Hospital for Rheumatic Diseases; L.S. Teh, MD, Rheumatology, Royal Blackburn Hospital; D.G.I. Scott, MD, Department of Rheumatology, Norfolk and Norwich University Hospital; M. Akil, MD; S. Naz, MD, Rheumatology, Royal Hallamshire Hospital; J. Andrews, MD, Rheumatology, Northwick Park Hospital; B. Griffiths, MD, Musculoskeletal Unit, Freeman Hospital; H. Harris, MD, Rheumatology, Cameron Hospital; H. Youssef, Rheumatology, Aberdeen Royal Infirmary; J. McLaren, MD, Rheumatology, Cameron Hospital; V. Toescu, MD, Rheumatology, Queen Elizabeth Hospital; V. Devakumar, MD, Rheumatology, Royal Blackburn Hospital; J. Teir, MD, Rheumatology, Sunderland Royal Hospital; I.N. Bruce, MD, FRCP, Professor and Honorary Consultant in Rheumatology, arc Epidemiology Unit, University of Manchester and Kellgren Centre for Rheumatology, Central Manchester and Manchester Children's University Hospitals NHS Trust

1. Address correspondence to Dr. I.N. Bruce, arc Epidemiology Unit, School of Translational Medicine, The University of Manchester, Manchester UK M13 9PT. E-mail: ian.bruce@manchester.ac.uk

Abstract

Objective. Accelerated atherosclerosis and premature coronary heart disease (CHD) are recognized complications of systemic lupus erythematosus (SLE), but the exact etiology remains unclear and is likely to be multifactorial. We hypothesized that SLE patients with CHD have increased exposure to traditional risk factors as well as differing disease phenotype and therapy-related factors compared to SLE patients free of CHD. Our aim was to examine risk factors for development of clinical CHD in SLE in the clinical setting.

Methods. In a UK-wide multicenter retrospective case-control study we recruited 53 SLE patients with verified clinical CHD (myocardial infarction or angina pectoris) and 96 SLE patients without clinical CHD. Controls were recruited from the same center as the case and matched by disease duration. Charts were reviewed up to time of event for cases, or the same "dummy-date" in controls.

Results. SLE patients with clinical CHD were older at the time of event [mean (SD) 53 (10) vs 42 (10) yrs; $p < 0.001$], more likely to be male [11 (20%) vs 3 (7%); $p < 0.001$], and had more exposure to all classic CHD risk factors compared to SLE patients without clinical CHD. They were also more likely to have been treated with corticosteroids (OR 2.46; 95% CI 1.03, 5.88) and azathioprine (OR 2.33; 95% CI 1.16, 4.67) and to have evidence of damage on the pre-event SLICC damage index (SDI) (OR 2.20; 95% CI 1.09, 4.44). There was no difference between groups with regard to clinical organ involvement or autoantibody profile.

Conclusion. Our study highlights the need for clinical vigilance to identify modifiable risk factors in the clinical setting and in particular with male patients. The pattern of organ involvement did not differ in SLE patients with CHD events. However, the higher pre-event SDI, azathioprine exposure, and pattern of damage items (disease-related rather than therapy-related) in cases suggests that a persistent active lupus phenotype contributes to CHD risk. In this regard,

corticosteroids and azathioprine may not control disease well enough to prevent CHD. Clinical trials are needed to determine whether classic risk factor modification will have a role in primary prevention of CHD in SLE patients and whether new therapies that control disease activity can better reduce CHD risk.

Footnotes

Supported by a grant from Lupus UK. Additional support was obtained from the arc Epidemiology Unit core support grant. Prof. Bruce received support from The Manchester Academic Health Science Centre and The Manchester NIHR Biomedical Research Centre.

Accepted for publication August 31, 2009.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Thank you for agreeing to give us your views on the technology and the way it should be used in the NHS.

Patients and patient advocates can provide a unique perspective on the technology, which is not typically available from the published literature.

To help you give your views, we have provided a template. The questions are there as prompts to guide you. You do not have to answer every question. Please do not exceed the 8-page limit.

About you

Your name: Chris Maker

Name of your organisation: LUPUS UK

Are you (tick all that apply):

- a patient with the condition for which NICE is considering this technology? **No**
- a carer of a patient with the condition for which NICE is considering this technology? **No**
- an employee of a patient organisation that represents patients with the condition for which NICE is considering the technology? If so, give your position in the organisation where appropriate (e.g. policy officer, trustee, member, etc) **Yes, Director**
- other? (please specify)

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

What do patients and/or carers consider to be the advantages and disadvantages of the technology for the condition?

1. **Advantages**

(a) Please list the specific aspect(s) of the condition that you expect the technology to help with. For each aspect you list please describe, if possible, what difference you expect the technology to make.

A reduction in the use of other medications with their side effects i.e. steroids.

Targeted suppression of the immune system to reduce the incidence and severity of lupus flares.

A reduction in medical appointments with the GP, Hospital and possibly emergency admissions.

(b) Please list any short-term and/or long-term benefits that patients expect to gain from using the technology. These might include the effect of the technology on:

- the course and/or outcome of the condition
- physical symptoms
- pain
- level of disability
- mental health
- quality of life (lifestyle, work, social functioning etc.)
- other quality of life issues not listed above
- other people (for example family, friends, employers)
- other issues not listed above.

Improved quality of life due to the reduction in disease activity (incidence and severity of flares) coupled with a reduction in the use of other therapies that often have difficult and often long term side effects, such as osteoporosis and weight gain, necessitating the use of other medications to counteract the side effects.

A reduction in the physical symptoms such as pain and fatigue will assist mobility and the ability to work, as well as improve mental health, as depression is a common feature with lupus patients. Pain and fatigue are the

Controlling this presently incurable disease it the key to overall wellbeing of the lupus patient. Lupus is largely invisible to family, friends and others as the symptoms and effects cannot be seen, other than when there is skin involvement, so the benefits of this technology can be seen to interlink improvements in the overall health of the lupus patient.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

What do patients and/or carers consider to be the advantages and disadvantages of the technology for the condition? (continued)

2. **Disadvantages**
Please list any problems with or concerns you have about the technology. Disadvantages might include:

- aspects of the condition that the technology cannot help with or might make worse.
- difficulties in taking or using the technology
- side effects (please describe which side effects patients might be willing to accept or tolerate and which would be difficult to accept or tolerate)
- impact on others (for example family, friends, employers)
- financial impact on the patient and/or their family (for example cost of travel needed to access the technology, or the cost of paying a carer).

All drugs have side effects and it is not known what these might be for this new technology.

Availability could be an issue if patients are required to travel long distances for treatment where there could be financial considerations regarding travel costs and also where a carer is involved. Some lupus patients have mobility problems and will always need assistance when travelling. Others may be unable to drive and will rely on family or friends for help as they may not be able to use public transport.

The potential of a 'postcode lottery' for this treatment.

3. Are there differences in opinion between patients about the usefulness or otherwise of this technology? If so, please describe them.

Most of our members are pleased to hear about this new technology that is specifically for lupus. Optimism appears high, perhaps unduly so as the treatment is not applicable to everyone who has lupus.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

4. Are there any groups of patients who might benefit **more** from the technology than others? Are there any groups of patients who might benefit **less** from the technology than others?

This depends on how lupus affects the individual patient and certainly those with more severe symptoms are likely to benefit more, when the new technology is prescribed.

Overall all lupus patients will benefit as a new lupus specific treatment will be available that will have a high profile leading to greater awareness within the medical profession.

Comparing the technology with alternative available treatments or technologies

NICE is interested in your views on how the technology compares with with existing treatments for this condition in the UK.

(i) Please list any current standard practice (alternatives if any) used in the UK.

NSAIDS
Anti-Malarials (hydroxychloroquine)
Steroids
Immunosuppressants
Biologics (Rituximab)

(ii) If you think that the new technology has any **advantages** for patients over other current standard practice, please describe them. Advantages might include:

- improvement in the condition overall
- improvement in certain aspects of the condition
- ease of use (for example tablets rather than injection)
- where the technology has to be used (for example at home rather than in hospital)
- side effects (please describe nature and number of problems, frequency, duration, severity etc.)

If it reduces the incidence and severity of flares along with a reduction in the level of other medications, currently being taken, it will be of considerable benefit to some lupus patients.

By reducing the level of other medications being taken there should be a commensurate reduction in side effects.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

(iii) If you think that the new technology has any **disadvantages** for patients compared with current standard practice, please describe them. Disadvantages might include:

- worsening of the condition overall
- worsening of specific aspects of the condition
- difficulty in use (for example injection rather than tablets)
- where the technology has to be used (for example in hospital rather than at home)
- side effects (for example nature or number of problems, how often, for how long, how severe).

This depends upon how individual patients react to the new technology. It will also depend on how it is administered.

Research evidence on patient or carer views of the technology

If you are familiar with the evidence base for the technology, please comment on whether patients' experience of using the technology as part of their routine NHS care reflects that observed under clinical trial conditions.

Are there any adverse effects that were not apparent in the clinical trials but have come to light since, during routine NHS care?

Not that I am aware of.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

Are you aware of any research carried out on patient or carer views of the condition or existing treatments that is relevant to an appraisal of this technology? If yes, please provide references to the relevant studies.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

Availability of this technology to patients in the NHS

What key differences, if any, would it make to patients and/or carers if this technology was made available on the NHS?

The benefits to patients of this new technology being made available are outlined above in 1b.

What implications would it have for patients and/or carers if the technology was **not** made available to patients on the NHS?

Loss of quality of life that would be gained as otherwise they would continue to rely on current medications along with all the side effects that these entail.

Are there groups of patients that have difficulties using the technology?

Provided the administering of the technology is managed by health professionals as appropriate this should not prove to be a major issue.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Single technology appraisal (STA)

Other Issues

Please include here any other issues you would like the Appraisal Committee to consider when appraising this technology.

Lupus is a multi symptom disease that is often difficult to diagnose and manage. The benefits of the new technology will greatly improve the outlook and quality of life for a significant number of lupus patients.

Evidence Review Group Report commissioned by the National Health Service Research & Development Programme on behalf of the National Institute for Health and Clinical Evidence

Title: Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Produced by: WARWICK EVIDENCE

Authors:
Martin Connock, Senior Research Fellow, Warwick Evidence
Ewen Cummins, Health Economist, McMDC Ltd.
Paul Sutcliffe, Senior Research Fellow, Warwick Evidence
Rachel Court, Information Specialist, Warwick Evidence
Rhona Johnson, Computer Programmer, McMDC Ltd
Amy Grove, Project Manager, Warwick Evidence
Kandala Ngianga-Bakwin, Principal Research Fellow, Warwick Evidence
Syed Mohiuddin, Research Fellow, Warwick Evidence
Aileen Clarke, Director for Warwick Evidence, Professor of Public Health & Health Services Research, Warwick Evidence

Correspondence to: *Dr Martin Connock, Warwick Evidence, Health Sciences Research Institute, Warwick Medical School, University of Warwick, Coventry CV4 7AL*

Date completed: *26.06.2011*

Source of funding:
This report was commissioned by the National Institute for Health Research Health Technology Assessment Programme as project number: 10/73/01

Declared conflicts of interest of authors:
None

Acknowledgements:
We gratefully acknowledge the expert clinical advice input from Professor C. Gordon (Professor of Rheumatology, University of Birmingham).

Responsibility for report:
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme or the NICE.

This report should be referenced as follows:
Connock M, Cummins E, Sutcliffe P, Court R, Johnson R, Grove A, Ngianga-Bakwin K, Mohiuddin S, and Clarke A; Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus: A Single Technology Appraisal. Warwick Evidence, 2011.

Contributions of authors:
Connock, Sutcliffe, Clarke – critical appraisal of clinical effectiveness. Cummins, Johnson, Ngianga Bakqin, Mohiuddin – critical appraisal of economic modelling. Ngianga-Bakwin – statistical appraisal. Court – searches. All authors – writing the report.

Table of Contents

Table of Tables

Table of Figures

Table of abbreviations

AC	Appraisal Committee
ACR	American College of Rheumatology
AE	Adverse Event
AMS	Adjusted Mean SLEDAI
ANCOVA	Analysis of Covariance
BILAG	British Isles Lupus Assessment Group
BLISS	Belimumab International SLE Study
BLyS	B-Lymphocyte Stimulator
BRAM	Birmingham Rheumatoid Arthritis Model
C	Complement
C3	Complement Component 3
C4	Complement Component 4
CAPD	Cumulative Average Prednisone Dose
CEAC	Cost Effectiveness Acceptability Curve
CI	Confidence Interval
CRD	Centre for Reviews and Dissemination
CVD	Cardiovascular Disease
EMA	European Medicines Agency
EQ-5D	EuroQoL 5 dimensions
ERG	Evidence Review Group
FACIT	Functional Assessment of Chronic Illness Therapy
FAD	Final Appraisal Determination
FDA	Food and Drug Administration
HDAS	High Disease Activity Subgroup
HGS	Human Genome Sciences
HRG	Health Research Group
HRQoL	Health-Related Quality of Life
ICER	Incremental Cost-Effectiveness Ratio
ISRCTN	International Standard Randomised Controlled Trial Number
IV	Intravenous
JHU	Johns Hopkins University
MeSH	Medical Subject Headings
MS	Manufacturer's Submission
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NSAIDs	Non-Steroidal Anti-Inflammatory Drugs
OR	Odds Ratio
PAS	Patient Access Scheme
PCS	Physical Component Summary
PGA	Physicians Global Assessment
PSA	Probabilistic Sensitivity Analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality Adjusted Life Year
QoL	Quality of Life
RCT	Randomised Controlled Trial
SELENA	Safety of Estrogen in Lupus National Assessment
SF-36	Short Form 36-Item Health Survey
SLE	Systemic Lupus Erythematosus

SLEDAI	Systemic Lupus Erythematosus Disease Activity Index
SLICC	Systemic Lupus International Collaborating Clinics
SoC	Standard of Care
SPC	Summary of Product Characteristics
SRI	SLE Responder Index
SS	SELENA-SLEDAI
STA	Single Technology Appraisal
UK	United Kingdom
USA	United States of America
URTI	Upper Respiratory Tract Infection
UTI	Urinary Tract Infection
VB	Visual Basic
Vs.	Versus
WTP	Willingness to Pay

Glossary of terms

POPULATION	SYNONYMS	DEFINITION	SOURCE
Auto-antibody positive active SLE	As BLISS trial populations	Active SLE disease, defined as a SELENA-SLEDAI (SELENA=Safety of Estrogens in Systemic Lupus Erythematosus National Assessment; SLEDAI=Systemic Lupus Erythematosus Disease Activity Index) score ≥6 and positive anti-nuclear antibody (ANA) test results (ANA titre ≥1:80 and/or a positive anti-dsDNA [≥30 units/ml]) at screening	Summary of Product Characteristics (Page 8) Manufacturer's clarification document (A1)
Proposed license population	Marketing authorisation population. High disease activity Population A	This subgroup comprises patients with positive anti-dsDNA and who also had low complement C3 or C4 Benlysta is indicated for reducing disease activity in adult patients with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g positive anti-dsDNA, low complement) despite standard therapy	Manufacturer's submission (Page 20) and clarification document (A1) Summary of Product Characteristics (Page 15)
Target population	High disease activity population that is the focus of submission	Adults with active autoantibody-positive systemic lupus erythematosus with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥10)	Manufacturer's submission (Page 20)

1 SUMMARY

1.1 Scope of the manufacturer submission

The manufacturer's scope encompasses the clinical effectiveness and cost effectiveness of belimumab plus Standard of Care (SoC) relative to SoC alone, for the treatment of adults with active auto-antibody positive Systemic Lupus Erythematosus (SLE) and also for a subgroup of these patients who exhibit signs of high disease activity. According to the manufacturer's scope and submission the population of greatest interest is the sub-group with high disease activity called the Target population. No subgroup is specified in the National Institute for Health and Clinical Excellence (NICE) scope. The Target population is a subgroup of the proposed licensed population; a decision on the manufacturer's license application is awaited.

The manufacturer's scope specifies that belimumab is delivered at 10mg/kg via a 1 hour intravenous (IV) infusion at 2 week intervals for the first 3 doses and every 4 weeks thereafter. The NICE scope merely states that belimumab is used as an add-on to SoC. The manufacturer's scope specifies SoC as the sole comparator, and considers there is no credible evidence to enable a statistical comparison of belimumab with other drugs, either rituximab or cyclophosphamide. These additional comparators are however specified in the NICE scope.

Outcomes listed in the manufacturer's scope include: disease activity, incidence and severity of flares, mortality, health-related quality of life (HRQoL), adverse effects of treatment, and fatigue; these coincide with those in the NICE scope.

The manufacturer's scope for economic analysis specifies a lifetime horizon, a National Health Service (NHS) and Personal Social Services (PSS) perspective, and a cost effectiveness analysis expressed in terms of incremental cost per quality adjusted life year; this corresponds to the NICE scope.

Special considerations raised in the manufacturer's scope include: 1) the innovative nature of belimumab for SLE; 2) the inability of the utility method to sensitively capture the quality of life (QoL) of SLE patients; and 3) the impact of SLE on particular ethnic groups and on women of childbearing age. There were no equity issues identified in the NICE scope.

The Food and Drug Administration (FDA) granted marketing authorisation in the United States of America (USA) in March 2011. Belimumab does not yet have marketing authorisation in Europe, the decision on an application is pending. It is therefore not yet certain if the manufacturer's scope will conform to directives from the European Licensing Authority.

Phase III trials have examined the effectiveness of belimumab at dosage regimens of 1mg/kg and 10mg/kg. The evidence relating to the 1mg/kg dose regimen has not been presented in the submission.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

A systematic literature search was undertaken by the manufacturer although the results of subsequent analysis of the retrieved studies were not clearly reported.

The submitted clinical effectiveness evidence is mainly derived from two international multicentre phase III randomised placebo controlled trials (Belimumab International SLE Study (BLISS)-52 and BLISS-76 lasting 52 and 76 weeks, respectively), which compared SoC plus belimumab vs. SoC plus placebo. In each trial patients were randomised (approximately 1:1:1) to one of three treatments: SoC + placebo, SoC + belimumab at 1mg/kg, or SoC + belimumab at 10mg/kg. BLISS-52 was undertaken mainly in Asia and South America while BLISS-76 patients mainly derived from North America and Europe.

There were 288 and 271 patients in the 1mg/kg arms of BLISS-52and BLISS-76 respectively, but results for effectiveness in these groups were not submitted. Of 865 patients in BLISS52 and 819 in BLISS-76, 577 and 548 were distributed almost equally into placebo and 10mg/kg belimumab groups, respectively.

Several sets of results were presented for placebo and 10mg/kg belimumab arms of the trials: [i] those for the whole population of trial patients, separately by trial and also after pooling patients from the two trials; [ii] those for the "Target population" after pooling data across the two trials; the Target population is a subgroup with high disease activity identified by post hoc analyses. The submission of multiple sets of results complicates any summary of the clinical effectiveness data.

For the Target population some outcome by-trial results became available during the clarification process, however the Evidence Review Group (ERG) was unable to comment on within trial comparison of belimumab vs. placebo for: the mean change in Short Form 36 Item Health Survey (SF-36) Physical Component Survey (PCS) score; time to first flare; SLICC/ACR (Systemic Lupus International Collaborating Clinics / American College of Rheumatology); and Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue Score. Box 1: Interpretation of marketing authorisation population and the Target populationprovides a summary of the manufacturer's response and reasons for focussing on the Target population.

Box 1: Interpretation of marketing authorisation population and the Target population

Our submission is based on a high disease activity subgroup of the marketing authorisation population, defined as the Target population. We acknowledge that NICE will be unable to make a recommendation for the whole of the expected licensed population (marketing authorisation population), but are aware that our Target population falls within the expected licensed population. Mindful of the need to make the most efficient use of NHS resources, this subgroup allows a Targeted approach to selecting the patients who are most likely to get the greatest benefit from treatment.

1.2.1 Primary outcome

The BLISS trials employed a novel composite outcome measure called the SLE Responder Index (SRI) which aimed to detect any improvement in SLE manifestation while guarding against the possibility that worsening involvement of other organ systems or a worsening in overall disease activity might be masked. To be classified as a "responder" a patient was required to satisfy specified minimum criteria in three measures of change in disease activity relative to baseline. The measures used were: [1] the Safety of Estrogen in Lupus Erythematosus National Assessment-SLE Disease Activity Index (SELENA SLEDAI) score, which detects an improvement in SLE manifestations, it scores disease activity over a range of 0 to 105 points and encompasses 24 weighted items scored as present or absent in the previous 10 days; clinically meaningful differences have been reported to be an improvement by a decrease of 6 points or a worsening by an increase of 8 points[1]; [2] the British Isles Lupus Activity Group (BILAG) index assesses organ system involvement over the preceding 28 days and is capable of detecting worsening of organ system involvement; it includes 86 items grouped into 8 organ systems, general (5 items), mucocutaneous (18 items), neurological (15 items), musculoskeletal (9 items), cardiorespiratory (12 items), vasculitis (8 items), renal (11 items), and hematological (8 items), (Isenberg and Gordon, 2000).[2] A score is calculated for each system depending on the SLE clinical manifestations (or signs and symptoms) present and whether they are new, worse, the same, improving, or not present in the last 4 weeks compared with the previous 4 weeks. BILAG uses classifications ranging from A to E as follows: A = worsening judged to require intensification of steroids or immunosuppressant treatments; B = worsening judged to require antimalarials, nonsteroidal anti-inflammatory drugs (NSAIDs), or low dose steroids; C = stable disease; D = improvement; E = system never involved; [3] a Physicians Global Assessment (PGA) score, employed to monitor for worsening in patient overall disease activity (scores can range between no disease = 0, and 3 = severe disease). The SRI criteria used to define a responder

were: an improved SLEDAI score by ≥ 4 points; a BILAG index showing no new grade A organ involvement or no two grade B organ involvements; a PGA score that has not increased by more than 0.3 points. The primary end point in both trials was the proportion of responders at week 52 relative to baseline according to the SRI.

In summary, the primary efficacy endpoint was the response rate at week 52, assessed with SLE Responder Index (SRI). A responder was defined as having a reduction of at least 4 points in the SELENA-SLEDAI score, no new BILAG A organ domain score, no more than 1 new BILAG B organ domain score, and no worsening in PGA score (increase < 0.3) at week 52 compared with baseline.

In both trials at 52 weeks SoC + 10mg/kg belimumab delivered a greater percentage of responders than did SoC + placebo. The difference in percentage of responders in the belimumab group relative to placebo group was 14% in BLISS-52and 9.4% in BLISS-76 and 11.8% for the whole population pooled across trials. The corresponding adjusted odds ratios for a response in BLISS-52 and in BLISS-76 were respectively 1.83 (95% CI: 1.30, 2.59; P = 0.0006) and 1.52 (95% CI: 1.07, 2.15; P = 0.027).

For the high disease activity subgroup (Target population) pooled across trials the difference in percentage of responders between the belimumab group and placebo group was 24.8% and the adjusted odds ratio was 2.7 (95% CI: 1.8, 4.1; P < 0.0001). For the Target population in BLISS-52 the difference in percentage of responders between the belimumab group and placebo group was 25.9% and the adjusted odds ratio was 3.0 (95% CI: 1.7, 5.2; P < 0.0001). For the Target population in BLISS-76 the difference in percentage of responders between the belimumab group and placebo group was 22.4% and the adjusted odds ratio was 2.5 (95% CI: 1.3, 4.6; P < 0.0045).

The percentage of responders was also reported at multiple follow up times. For the Target population pooled across trials and in BLISS-52, at many times, a significantly greater response was observed for the belimumab group relative to placebo group (significance tests uncorrected for multiple testing), however for BLISS-76 the only time a significantly (P < 0.05) greater response was observed for the belimumab group was at week 52.

1.2.2 Secondary Outcomes

The pre-specified major secondary outcomes were: the SRI response at week 76; the percentage of patients with a ≥ 4 point SLEDAI improvement at week 52; mean change in PGA score at week 24, percentage of patients with prednisone reductions ≥ 25% from baseline to ≤7.5 mg/day during weeks 40 to 52 (in subjects whose baseline dose was > 7.5 mg/day); mean change in SF-36 PCS at week 24.

The major secondary outcome of percentage of SRI responders at week 76 failed to reach statistical significance (odds ratio and P value not submitted; odds ratio 1.31, 95% CI: 0.92 – 1.87, P = 0.1323 by logistic regression, taken from the FDA briefing package).[3]

Mean change in PGA score at week 24 was defined as a major secondary outcome. For the whole population in BLISS-52 the change in PGA score (week 24 relative to baseline) for both groups indicated disease improvement and was greater in the belimumab group (-0.54) than placebo group (-0.39; P = 0.0003 in support of belimumab). For BLISS-76 the difference between groups was very small and in favour of placebo (-0.49 placebo and -0.48 belimumab) and did not reach statistical significance (P = 0.7987). For the Target population pooled across trials belimumab delivered a greater reduction in PGA score than placebo (P = 0.028 with mean changes of -0.42 and -0.52 for placebo and belimumab, respectively). Target population data was not been provided for the change in PGA score separately for the BLISS-52 and BLISS-76.

Components of the SRI at Week 52

A further major secondary outcome was the percentage of patients at week 52 that achieved a SLEDAI score reduction of ≥ 4 points. Both trials delivered a significantly greater percentage for belimumab than for placebo (P = 0.0024 and P = 0.0062 for BLISS-52 and BLISS-76, respectively). Similarly, the Target population data delivered a significantly greater percentage for belimumab (P = 0.0004 and P = 0.0063 for BLISS-52 and BLISS-76, respectively).

Results at week 52 for the other two SRI components (i.e. the BILAG index and PGA score) were submitted (non-major secondary outcomes). The percentage of patients in the whole population that satisfied BILAG and PGA criteria in BLISS-52 was greater for belimumab than placebo (significant at P = 0.0181 and P = 0.0048 for BILAG and PGA, respectively); however, for BLISS-76 the differences between belimumab and placebo were smaller and neither component reached statistical significance in favour of belimumab (P = 0.319 and P = 0.1258 for BILAG and PGA, respectively). Similarly, the percentage of patients in the Target population which satisfied BILAG and PGA criteria in BLISS-52 was greater for belimumab than placebo (P = 0.0099 and P = 0.0063 for BILAG and PGA, respectively); however, for BLISS-76 the differences between belimumab and placebo were far more modest and did not reach conventional statistical significance (P = 0.1297 and P = 0.1312 for BILAG and PGA, respectively).

In summary, in BLISS-52 for the total population and for the high disease activity subgroup, belimumab at 10mg/kg delivered significantly more responders at 52 weeks than placebo for

SLEDAI score reduction of ≥ 4 points, no worsening in PGA, and no worsening in BILAG. However, for BLISS-76 at 52 weeks total population and high disease activity subgroup, a significant response with belimumab 10mg/kg compared to placebo was only seen with the 4-point reduction in SELENA-SLEDAI component (difference between belimumab and placebo = 22%, odds ratio = 2.4 [95% CI: 1.3, 4.4; P < 0.0063]).

Reduction in steroid usage

Reduction in steroid use was specified as a major secondary outcome. In BLISS-52 at baseline 68.6% of patients were receiving ≥ 7.5 mg/day prednisone. The corresponding percentage for BLISS-76 was 44.9%. The percentage of these patients whose steroid use was reduced in weeks 40 to 52 by the pre-specified amount was greater in the belimumab arm than the placebo arm in both trials. The difference (belimumab vs. placebo) failed to reach statistical significance in either trial: 18.6% vs. 12.0% in BLISS-52 (P = 0.0526 from logistic regression including baseline covariates) and 16.7% vs. 12.7% in BLISS-76 (P = 0.5323). For the Target population pooled across trials 15.9% and 7.1% reduced steroid use in belimumab and placebo groups, respectively (P = 0.0389 from logistic regression). For the Target population in the BLISS-52 trial there was a large difference in reduced steroid use between belimumab and placebo groups (18.5% and 5.3% respectively; odds ratio = 4.11, 95% CI: 1.29, 13.2; P = 0.0171). For the Target population in the BLISS-76 trial there was no difference between groups (11.1% and 10% reduced steroid use in belimumab and placebo groups respectively; odds ratio = 0.88, 95% CI: 0.21, 3.60; P = 0.8586).

Quality of life

The mean change in SF 36 PCS scores was specified as, a major secondary outcome. At week 24 relative to baseline it showed little difference between belimumab and placebo groups in BLISS-52 (P = 0.8870), or in BLISS-76 (P = 0.6601), or in the Target population pooled across trials (P = 0.4276).

Change in SF 36 PCS scores at week 52 was also specified as a non-major secondary outcome. No significant improvement was observed for BLISS-76 or Target populations (P = 0.5134 and P = 0.1124, respectively) however in BLISS-52 there was a difference between belimumab and placebo arms (4.18 vs. 2.96) (P = 0.0247).

In BLISS-52 over the course of the study there was a statistically non-significant difference (P value not submitted) in favour of belimumab relative to placebo in the absolute change of EuroQoL 5 Dimensions (EQ-5D) score from baseline, however the results for belimumab and placebo groups in BLISS-76 were indistinguishable. For the pooled Target population the

difference between 10mg/kg and placebo groups reached statistical significance in favour of belimumab at week 24 ($P \leq 0.01$), but the difference almost completely faded by week 52.

SLEDAI flare index

Other specified non-major secondary efficacy outcomes for which results were submitted included: time to first SLE flare (assessed using the SLEDAI Flare Index which categorizes flares as "mild or moderate" or "severe" based on 6 variables (see Appendix 1); disease progression at week 52 relative to baseline assessed using the SLICC/ACR index; fatigue over the course of the study estimated using the FACIT-Fatigue scale4 which ranges from 0 to 52 (0 is the worst possible score and 52 is the best).

In BLISS-52 the time to first flare and time to first severe flare were delayed by belimumab relative to placebo (HR 0.76, 95% CI: 0.63 – 0.91, P = 0.0036; HR 0.57, 95% CI: 0.39 – 0.85, P = 0.0055, respectively). In BLISS-76 there was no difference between groups in time to first flare (P = 0.4796) but relative to placebo belimumab somewhat delayed time to first severe flare (HR 0.72, 95% CI: 0.50 – 1.05, P = 0.0867). For the Target population pooled across trials, relative to placebo, belimumab significantly delayed time to both first flare (P = 0.007) and to first severe flare (P = 0.0028).

SLICC/ACR organ damage

There was little difference between placebo and belimumab groups in terms of change in SLICC/ACR score at week 52; precise values by trial were not submitted. Data reported elsewhere[5] were: BLISS-52 score change 0.06 and 0.04 for placebo and belimumab groups respectively, P for difference 0.4222; BLISS-76 score change 0.05 and 0.03 for placebo and belimumab groups respectively, P for difference 0.3415.

Fatigue

At week 52 relative to baseline the belimumab group had greater improvement in FACIT-Fatigue score than the placebo group (4.8 belimumab and 2.1 placebo in BLISS-52, $P < 0.001$; 4.6 and 2.9 in BLISS-76, P = 0.05). For the Target population at weeks 8 and 12 the difference between groups was statistically in favour of belimumab ($P < 0.05$) however the difference between groups then diminished and at week 52 the difference no longer reached conventional statistical significance.

Modified SRI response

The results for a non-pre-specified secondary outcome, the "modified SRI" at week 52, were submitted. In the "modified SRI" serological improvements (2 points each for anti-dsDNA antibodies and for complement) were not allowed to contribute toward a ≥4 points reduction

in SLEDAI score. In BLISS-52 belimumab delivered a greater percentage of responders than did placebo (P = 0.0038); in BLISS-76 the difference in favour of belimumab failed to reach the conventional level of statistical significance (P = 0.064).

1.2.3 Safety

The submission pooled results from three randomised controlled trials (RCTs): BLISS-52, BLISS-76 and LBSL02, providing information on 675 patients who received placebo and 1458 who were exposed to belimumab. LBSL02 lasted 52 weeks, preceded the BLISS trials, was conducted in North America (98% patients from the USA) randomised 449 patients to one of four treatments: SoC + placebo, SoC + 1mg/kg belimumab, SoC + 4mg/kg belimumab, and SoC + 10mg/kg belimumab. Although all patients had a history of auto-immunity, at recruitment 30% currently lacked anti-nuclear antibodies. This trial did not employ the SRI composite outcome.

Deaths

There were 15 deaths during the controlled phase of the three trials; 3 in the placebo group (n=675), and 12 in the belimumab groups (n=1458) with 6 each in the 10mg/kg and 1mg/kg groups respectively. One death in the 1mg/kg belimumab group followed 15 weeks after the patient stopped belimumab treatment. The causes of death were various.

Adverse events

In all treatment groups > 90% of patients experienced at least 1 adverse event (AE). The most commonly occurring AEs were headache, upper respiratory tract infection, arthralgia, nausea, urinary tract infection (UTI), diarrhoea and fatigue.

The percentage of patients experiencing at least one serious AE and at least one serious AE was very similar between placebo and belimumab groups ranging from13.5% to 18.6%, with a very slight numerical excess in the belimumab group. The most frequent serious AEs (≥ 1% in any treatment group) were pneumonia, pyrexia, UTI, cholelithiasis, and cellulitis. The percentage of patients experiencing at least one severe AE was 15.4% for the placebo group and 16% across the belimumab groups; the most common severe adverse events were not identified.

Occurrence of infusion plus hypersensitivity reactions was similar between belimumab and placebo-treated patients (17% and 14.7%, respectively).

Infections

Infections occurred slightly more often in patients treated with belimumab compared to placebo. The most frequent infections were upper respiratory tract infection (URTI), UTI, nasopharyngitis, sinusitis, and bronchitis.

Malignancy

Five solid organ malignancies were reported across the trials: a stomach carcinoid (placebo group, day 202); a breast cancer (belimumab 1mg/kg, day 102); a cervical cancer (belimumab 1mg/kg, day 439); an ovarian cancer (belimumab 1mg/kg, day 21, patient died); and a thyroid cancer (belimumab 1mg/kg, day 378). There were four non-melanoma skin cancers: two basal cell carcinomas, and two squamous cell carcinomas (1 in the placebo group, 3 in the belimumab 1mg/kg group).

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The submission omitted results for the 1mg/kg groups from the two pivotal trials. Therefore from information in the manufactures submission (MS) alone, the consistency of results across the whole data set could not be fully assessed and it was not possible to gauge the evidence for a dose response relationship. However, data for the 1mg/kg groups is available in the public domain (FDA documents pertaining to the Human Genome Sciences (HGS) Briefing Document to the FDA[3,5]) and the ERG have considered this information in critiquing the submitted evidence.

Even without the 1mg/kg group results the MS provided clinical evidence for a large number of outcomes reported for six separate populations (whole population from BLISS-52, whole population from BLISS-76, pooled whole populations from BLISS-52 plus BLISS-76, pooled Target populations from BLISS-52 plus BLISS-76, and after the clarification process Target population from BLISS-52 and Target population from BLISS-76. Additionally, AEs for LBSL02 were included.[6]

The most noticeable aspect of the submitted results was the relative lack of evidence for clinical effectiveness of belimumab seen in the BLISS-76 trial. Although at week 52 for the pre-specified primary outcome measure the percent responders (SRI) reached statistical significance in favour of belimumab (P = 0.027), at no other time point did this outcome reach significance. Furthermore, all major and non-major secondary outcome results submitted, except for a $\geq$ 4 point SLEDAI improvement at week 52 which is a subcomponent of the SRI response, likewise failed to reach statistical significance including: PGA change at week 24 and 52, SRI responders at week 76, reduction in use of steroids week 40 to 52, SF-36 change

at week 24, time to first flare, time to first severe flare, change in SLICC/ACR organ damage score at week 52, fatigue status (FACIT change from baseline), and quality of life (EQ-5D change).

The SLE population in BLISS-76 is more likely to resemble that in the UK than that in the BLISS-52; therefore the BLISS-76 results are probably more relevant to the decision problem than those from BLISS-52. The results favourable for belimumab submitted for the whole population pooled across trials were largely driven by BLISS-52 results. For the Target population the results from the BLISS-52 trial were again more favourable to belimumab than those from BLISS-76 and additionally BLISS-52 provided more patients to the pooled Target population than BLISS-76 (55% vs. 45%), therefore results favourable to belimumab for the pooled Target population were again more strongly driven by the contribution from the BLISS-52 Target population.

Results in the public domain[3,5]for the 1mg/kg and 10mg/kg dose regimens in the BLISS-76 trial were not supportive of a dose response relationship. For many outcomes the results were as favourable for the low dose group as for the high dose group. These outcomes included: percentage of SRI responders across trial follow up, percentage of patients with no worsening in PGA at week 52, percentage of patients with no worsening in BILAG index at week 52, mean change in PGA score from baseline at week 52, reduction in steroid use weeks 40 to 52, time to first flare, time to first severe flare, and mean change in FACIT fatigue score at week 52.

1.4 Summary of cost effectiveness submitted evidence by the manufacturer

No published relevant economic evaluations were identified in the submission. The search strategy was poor but it appears unlikely that economic studies were missed.

The submitted cost-effectiveness work focuses entirely on a new model and economic evaluation undertaken by the manufacturer. This de novo individual patient micro-simulation model examined the cost-effectiveness of belimumab plus SoC versus SoC. This employed an annual cycle over a lifetime horizon and conformed to the NICE reference case.

In brief, the model was constructed using three main sources of data:

- The BLISS trials
- The John Hopkins University cohort
- Additional data drawn from the wider literature

The trial data determined distribution of patient characteristics at baseline. Regression analysis was also used to model patients' SLEDAI score at week 52. Those in the belimumab arm who responded at week 24 were modelled as remaining on belimumab and maintaining the modelled SLEDAI score at week 52. Those in the belimumab arm who did not respond at week 24 were assumed to stop treatment and were modelled as having the average SoC SLEDAI score at week 52. Discontinuation rates from week 24 were also drawn from the trial data.

Regression analysis from the trial data was also used to derive HRQoL and cost functions related to a patient's SLEDAI score.

Given these inputs, the bulk of the remainder of the model was derived from the Johns Hopkins University (JHU) cohort data. The survival function and the risks of developing each of the 12 organ involvements within the SLICC index were modelled on a range of covariates, these including the adjusted mean SLEDAI score to date and the average cumulative prednisolone dose. Steroid use was not drawn from trial data but was rather modelled using a function estimated from the JHU cohort relating steroid use to a patient's SLEDAI score. The evolution of the SLEDAI score subsequent to week 52 was also estimated from the JHU cohort data, with the manufacturer adjusting the constant of the functional form to better fit the belimumab phase II trial data.

The survival function estimated from the JHU cohort was amended by SMRs drawn from a paper within the literature. Additional data drawn from the literature was used to inform the HRQoL and cost impacts arising from involvement of the individual 12 organ involvements within the SLICC index.

Base case deterministic results were submitted by the manufacturer for three patient populations:

- The All BLISS patient population;
- The patient population within the BLISS trials that relates to the anticipated license of Anti-dsDNA+ve and low (C3 or C4);
- The Target population which restricts the patient population to the licensed patient population with an SS score at baseline of at least 10.

For the All BLISS population the central survival estimate was an additional 1.50 years survival from use of belimumab. The discounted net gains and costs were 0.43 QALYs at a

net cost of £35,584 to yield a cost effectiveness estimate of £82,909 per QALY. With the PAS the net cost fell to £[redacted] to yield a cost effectiveness estimate of £[redacted] per QALY.

Only the base case deterministic results were presented for the anticipated license population. The central survival estimate was an additional 2.13 years survival from use of belimumab. The discounted net gains and costs were 0.61 QALYs at a net cost of £40,303 to yield a cost effectiveness estimate of £66,170 per QALY. With the PAS the net cost fell to £[redacted] to yield a cost effectiveness estimate of £[redacted] per QALY.

For the Target population the central survival estimate was an additional 2.93 years survival from use of belimumab. The discounted net gains and costs were 0.81 QALYs at a net cost of £51,925, to yield a cost effectiveness estimate of £64,410 per QALY. With the PAS the net cost fell to £[redacted] to yield a cost effectiveness estimate of £[redacted] per QALY. The direct costs of belimumab and its administration accounted for 90% and 17% respectively of the total net costs. Lower costs from pulmonary involvement and from renal involvement provided costs offsets of around 6% and 4% respectively.

An ERG cross-check of the probabilistic modelling for the Target population resulted in a central estimate of £65,530 per QALY. Due in part to the results being reasonably linear and also the time inherent in running the model probabilistically all other results presented are from the deterministic model.

A range of sensitivity analyses were presented for the All BLISS patient population and the Target population. Restricting attention to the Target population, results were sensitive to the initial changes in the SLEDAI score that were modelled, the manufacturer adjustment to the long term SLEDAI score function, the impact of the adjusted mean SLEDAI score upon mortality and the natural history models for pulmonary and renal involvement.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

Assuming that belimumab week 24 non-responders will experience the average SLEDAI score within the SoC arm is likely to have over-estimated the average impact upon SS scores within the belimumab arm. The SLEDAI score drives the analysis and any error in its calculation will have a major impact on results. There may also be errors in the calculation of the adjusted mean SLEDAI score from not taking into account a patient's probable prior history, with this concern also applying to the calculation of the cumulative average steroid dose.

The maintenance through time of the absolute gain in SLEDAI score among those remaining on treatment compared to those on SoC may be optimistic, and at a minimum should have been explored in a scenario analysis. However, ERG expert opinion indicates that it may in some sense be possible to "reset" the immune system which may negate this concern.

Whether it is reasonable to extrapolate the 8% annual discontinuation rate for the Target population beyond the trial period is also unclear given the lack of detail around the figures underlying this rate. A high discontinuation rate from week 24 improves the cost effectiveness of belimumab.

Adjusting the JHU cohort survival model by SMRs from the literature may not be justified. The SMRs applied may also not be representative of the overall literature. This may have tended to exaggerate the impacts of the covariates within the JHU cohort survival model.

Costs as a function of the SLEDAI score may have been exaggerated by analysing the data on a six monthly basis rather than the annual basis on which the model is constructed. There are also some concerns that the separate estimation of costs for each organ involvement may have tended to double count the cost impacts of SLE.

There appear to be some discrepancies in the reported model outputs between the average durations of organ involvement, the annual costs of these and the discounted total costs of these organ involvements. There are as a consequence concerns around the calculation of the cost offsets from reduced organ involvement arising from belimumab.

With the exception of the last point the effects of which are currently ambiguous, the above suggest that the model may have tended to overestimate the impact of belimumab on the SLEDAI score and to have overestimated the likely impact of the SLEDAI score on the model outputs. There are few immediately obvious biases pulling in the opposite direction.

1.6 ERG commentary on the robustness of evidence submitted by the manufacturer

The submitted evidence concerned the clinical and cost effectiveness of 10mg/kg belimumab, used as an add-on to standard of care, compared to standard of care alone. Several SLE populations were considered. Evidence for clinical effectiveness came from two placebo-controlled phase III trials: BLISS-52[7] was conducted at 90 centres located in Pacific-Asia, South America and Eastern Europe (11 centres); BLISS-76 was conducted in 136 centres in

North America and Europe. As such the results from BLISS-76 are more likely to be generalisable to the UK.

Total populations in the BLISS trials (auto-antibody positive patients with active SLE disease sufficient to score ≥ 6 points on the SLEDAI scale) conformed to that in the NICE scope ,but the effectiveness of belimumab in the two trials was disproportionately greater in BLISS-52 than BLISS-76 and the evidence for effectiveness from BLISS-76 was not convincing. The manufacturer pooled data from the two BLISS trials, but the pooled analyses that favoured belimumab were almost exclusively driven by the effectiveness results from BLISS-52 and are arguably less applicable to the UK than the BLISS-76 results alone. The results for PGA of disease activity were noticeably disparate between trials. The ERG considered that inadequate allocation concealment of outcome assessors (physicians) in BLISS-52 might explain this discrepancy and may be a cause for concern since PGA is a component of the composite primary outcome.

The manufacturer submitted evidence for a high disease activity sub-population from the BLISS trials; this was called the "Target population". The Target population was not a pre-specified subgroup in the trial protocols; it was identified using post hoc analyses to seek out a more strongly responding subgroup of patients. The Target population was defined according to baseline disease activity score (≥ 10 SLEDAI points), level of antibodies to dsDNA, and low level of complement. Each of these three criteria defined a pre-specified subgroup from the BLISS trials, but the combination of all three was not pre-specified. The Target population represents a subpopulation of the proposed licensed population which in turn is a subpopulation of the total BLISS trial population. For the Target population only outcome results pooled across the trials were submitted and it was impossible to check for consistency of results between trials. This was a cause for concern because of the lack of convincing evidence of effectiveness for the whole population in BLISS-76 (see above).

Both BLISS trials had three randomised groups: placebo, 1mg/kg belimumab and 10mg/kg belimumab. The submission did not include results for the 1mg/kg arms of the trials; however these results are available in the public domain (FDA documents[3,5]). Results for the 1mg/kg dose regimen are relevant for checking consistency between trial results and in determining whether a dose response relationship exists. In this respect the most noticeable result was that in BLISS-76 there was essentially no evidence of any dose response relationship across the time span of the trial and no difference between the proportions of SRI responders to 1mg/kg and to 10mg/kg. For the SRI subcomponents PGA and BILAG in BLISS-76, the 1mg/kg regimen was more effective than 10mg/mg; similarly the 1mg/kg dose regimen appeared as effective as 10mg/kg in suppressing flares.

1.6.1 Strengths

The main strengths of the clinical effectiveness evidence were:

- The significant result ($P < 0.05$) for the pre-specified primary end point (52 week SRI) in both the two phase III RCTs;
- The fact that this primary outcome had been developed in consultation with a licensing authority (the FDA) and guarded against the possibility that improvement in some particular SLE manifestation or manifestations might mask deterioration in overall disease activity or involvement of new organ damage.

The main strengths of the cost effectiveness submission were:

- Provision of a well constructed model which conforms to the NICE reference case;
- An impressive attempt at modelling the longer term effects of SLE using extensive modelling of the JHU SLE cohort;
- The presentation of a simple and transparent PAS that allows easy implementation within the economic model.

1.6.2 Weaknesses and areas of uncertainty

There were a number of weaknesses and uncertainties; these include:

- The lack of convincing evidence from BLISS-76 that belimumab is superior to placebo in the total population. Only at week 52 did the proportion of SRI responders reach statistical significance in favour of belimumab. At other monitoring times significance was not reached and results for 1mg/kg (available from FDA reports[3,5]) and 10mg/kg belimumab are indistinguishable. Furthermore, of the five pre-specified major secondary end points in only one was belimumab favoured at a level that might not reasonably be accounted for by chance; MC this outcome, a reduction in the SLEDAI score of ≥ 4 points at week 52, represents the smallest disease activity improvement that can be considered clinically significant. None of the other major secondary outcomes in the BLISS-76 trial favoured belimumab at a level that strongly excluded the possibility of a chance result, including: PGA change at week 24 and 52, SRI responders at week 76, reduction in use of steroids week 40 to 52, SF-36 change at week 24. Furthermore, none of the other submitted secondary outcomes strongly

excluded chance from accounting for results in favour of belimumab, including: time to first flare, time to first severe flare, change in SLICC/ACR organ damage score at week 52, fatigue status (FACIT change from baseline), and quality of life (EQ-5D change). For some of these outcomes there was little distinction in effectiveness between 1mg/g and 10mg/kg belimumab dose regimens.

- There were considerable (and significant) geographical and racial differences between the BLISS trials (which may indicate potential differences in practice and in standard of care). The BLISS-76 population is more closely comparable to that of the UK than the BLISS-52 as also are the likely underlying care patterns. The submission pooled the two BLISS trials (52 and 76). The pooled results favourable for belimumab were almost exclusively driven by BLISS-52. The relevance of the pooled results for England and Wales is therefore uncertain. Similarly subgroup analysis (Manufacturer's clarification document) of the primary outcome according to geographical region (USA/Canada, Western Europe, Eastern Europe, America-excluding USA/Canada, Asia) indicated that response was strongest in America-excluding-USA/Canada and weakest in USA/Canada and Western Europe.

- The submission excluded results for the 1mg/kg arms of the two BLISS trials. The trial results available from the FDA indicated a lack of convincing evidence for an expected dose response relationship, with no consistent additional benefit from 10mg/kg compared to a 1mg/kg dose.

- The original submission only presented pooled effectiveness results for the Target population. The NICE submission template specifically requests separate results by trial when more than one trial is available. In the light of the relative lack of effectiveness displayed in BLISS-76 for the whole trial population, the lack of trial specific data for the Target population weakened the submission's case. Trial specific results for some outcomes for the Target population were made available during the clarification process. No data is available for the effectiveness of the 1mg/kg dose in the Target population.

- The ERG found outcome data common to BLISS and rituximab trials which was not explicit in the MS, so that the existence of data for an indirect comparison of interventions was not acknowledged. NICE request a rationale for not conducting meta-analysis when more than two RCTs are available. No rationale was provided in the MS.

- The BLISS populations exhibited a narrow range of SLE manifestations (mainly restricted to musculoskeletal and cutaneous problems). The BLISS trials were underpowered to estimate the effectiveness of belimumab with respect to manifestations in other domains. Also there was a lack of controlled evidence on the effectiveness of belimumab relative to SoC in the longer term beyond 76 weeks. Yet, in the economic analysis, effectiveness data from the pooled BLISS populations have been used in modelling belimumab's effect on a wide range of organ systems in SLE over a life-time horizon.

- The economic model generated a survival benefit for belimumab over SoC: an additional 2.93 undiscounted year's survival from belimumab within the Target population. There is no direct clinical evidence to support this. Actually, during the randomised phases of the belimumab trials, there were a greater number of deaths associated with belimumab than with placebo.

- The economic model generates better survival for patients with high disease activity than for those with low disease activity. This counterintuitive result appears to reflect the larger proportion of young patients in the Target population from the pooled BLISS trials. As such this will merely reflect the exigencies of trial recruitment and it is very uncertain whether this population is representative of high disease activity patients in England and Wales.

- In the economic model there may be an element of double counting in estimating the costs of complications, these costs being a function of the SLEDAI score with further costs being added for the individual components of the SLICC index.

- The economic model data from the JHU cohort to estimate a number of functions within the model: the long term evolution of the SLEDAI score, steroid use, mortality risks and the risk of developing organ involvement. The level of disease activity in this cohort is very much lower than that of the Target population and as a consequence the manufacturer made an informal adjustment to the SLEDAI score evolution function. This adjustment improved the estimate of the cost effectiveness of belimumab. Some informal justification for this adjustment has been provided within manufacturer responses to ERG clarification questions, but uncertainty remains because the reliability and validity of the adopted adjustment was not fully explored or robustly defended.

- The submission did not provide data about maintenance of response at the patient level. The SRI outcome was reported as a group response; the graph line showing percentage of responders across the duration of the trials rose and fell at various follow up times, thus an individual non-responder could later improve sufficiently to be classified as a responder.

- The economic model overestimated health benefit in the belimumab arm between weeks 24 and 52. This was because the estimate (i.e. the decrease in SLEDAI score) for non-responders in the belimumab arm (33% of patients in the belimumab arm) was calculated from observed changes in the whole SoC arm which included responders as well non-responders. In the pooled Target population the SoC arm consisted of 52% responders and 48% non-responders with average SLEDAI improvement of 6.9 points for responders and 1.1 points for non-responders; thus the improvement for the SoC arm as a whole was heavily weighted by the SoC responders. A more appropriate procedure would be to base the estimate of the health benefits for non-responders in the belimumab arm on the SLEDAI change observed for non-responders in the SoC arm (i.e. assuming that the two sets of non-responders experience similar disease trajectories as a result of their SLE). This weakness extends beyond week 52 because the manufacturer's model for the belimumab arm non-responders after week 52 continues to make an estimate of disease activity based on the whole SoC arm (made up of a mix of both responders and non-responders).

1.7 Summary of additional work undertaken by the ERG

The ERG undertook substantial additional work in the following areas:

1. Supplementing the MS with data in the public domain e.g. as available from the FDA.

2. Extensive clarifications required to understand the anticipated effectiveness in the different relevant sub-populations: Target; high disease activity and license.

3. Re-running the search strategies.

4. Running a probabilistic sensitivity analysis (PSA) to cross check the model.

1.8 Key issues

The proposed licensed population and the high disease activity "Target" population, the focus of the clinical effectiveness and economic submissions, were subgroups identified from post hoc analyses aimed at identifying patients with the greatest response to belimumab in the

pooled phase III trial populations. Although important as subgroup analyses, the results should be viewed with some caution and not assumed to represent the outcome of an independent randomised investigation of pre-defined "Target" populations.

The MS and clarification document presented results for multiple populations (whole population from two RCTs, pooled whole populations, pooled Target populations and Target population separately by trial). The results from the BLISS-76 trial, which were less supportive of belimumab than those from BLISS-52, are those more generalisable to the UK; the economic model employed results pooled across both trials and therefore may somewhat overestimate the cost effectiveness of belimumab for the UK population.

The submission did not present results for the 1mg/kg groups in the two pivotal trials. Data in the public domain,[3,5] although not formally tested statistically, indicated that for several outcomes, including the primary outcome in the BLISS-76 trial, there was little difference between the effect of belimumab in 1mg/kg and 10mg/kg groups.

The manufacturer's estimate of the number of Target population patients in the UK was based on the proportion of such patients at baseline in the BLISS trials; this will probably be an underestimate because SLE is a relapsing and remitting disease and the number of patients likely to reach Target population status at some stage in their disease will accumulate through time.

The manufacturer's economic model relied heavily on time to event analyses of the JHU SLE cohort. Based on SELENA SLEDAI scores there was a gross mismatch between JHU cohort patients and the populations modelled, the former had far less severe disease, especially in comparison to the Target population. To allow for this mismatch a major adjustment was required in modelling; the manufacturer's justification for the type of adjustment adopted was that a similar procedure had been explored in cardiovascular studies for the prevention of cardiovascular disease (CVD).[8] The robustness of the manufacturer's approach in this respect is difficult to gauge.

Participants in the 52 and 76 week pivotal RCTs experienced a relatively narrow range of SLE manifestations, predominantly in musculoskeletal, cutaneous and serological domains. The economic analysis used trial changes in SELENA SLEDAI scores for these patients in order to model long term accumulation of organ damage in many other systems. The reliability of this procedure is again difficult to gauge and was referenced in the MS with a single analysis published in 1999.[9]

2 BACKGROUND

2.1 Critique of manufacture's description of underlying health problem

The manufacturer provides an adequate description of the treatment pathways for patients with SLE. These are presented MS section 2 (Page 39).

The ERG is of the opinion that the manufacturer's summary of the disease context and available treatments for patients with SLE is reasonable.

2.2 Critique of manufacturer's overview of current service provision

The MS executive summary and MS section 2 adequately describe the aims and modes of treatment (Pages 17 to 18). The key points, taken from the MS, are as follows:

Treatments aim to: match treatment to an accurate diagnosis of the extent of organ involvement; maintain an appropriate level of therapy to control or halt the inflammatory disease activity while minimising side-effects and risk of infection; prevent further organ damage; maintain a patient's daily function and quality of life.

Currently a range of treatments (including NSAIDs, corticosteroids, immunosuppressants and antimalarials) are variously used either alone or in different combinations, constituting standard of care (SoC). The MS documents that current SoC may be associated with undesirable effects e.g. from chronic use of steroids or side effects associated with immunosuppressants. The MS points out that many treatments are not licensed for use in SLE and that a significant number of patients with advanced SLE do not respond to current treatments even at high doses". Patients with more severe, highly active SLE are usually managed in tertiary centres and may routinely receive rituximab.

Currently there is no accepted SLE treatment algorithm and no relevant NICE guideline exists. Agreeing on best practice poses a significant challenge owing to the heterogeneous nature of SLE.

The manufacturer has estimated the number of patients presenting with SLE in England and Wales who would be eligible for treatment with belimumab (see MS Table 2.2). Taking 13,198 as the number of patients with active SLE in England and Wales, the MS calculated that 92.5% of these are adults. The estimated number eligible for belimumab according to the "proposed license indication" was then based on the proportion of patients (52%) in the

pivotal phase III trials who fitted the criteria defining the "proposed license population" giving a total of 6,348 (i.e. 0.925 * 0.52 * 13,198). This proposed license population exhibited a higher level of serological disease than the total Phase III populations. However the MS further submitted that NICE should actually consider belimumab for a subgroup of the "proposed licence population". This population was a narrower population of high disease activity patients termed the "Target population", representing 34% of the patients in the Phase III trials, giving an estimated number of 4,151 patients in England and Wales (i.e. 0.925 * 0.34 * 13,198).

The manufacturer's calculations should be viewed with some caution because the Phase III trials upon which they are based were international studies in which UK patients were a very small minority, and because the actual proportions of "proposed license population" and "Target population" patients in these trials will reflect the vagaries of trial recruitment rather than the distribution of these patients in the countries from which they were selected.

The manufacturer's estimate of the cost to the NHS of treating the "Target population" in year one assumed 50% usage of belimumab (i.e. 2,075 patients) and came to [redacted] under the manufacturer's proposed patient access scheme (PAS). The manufacturer's estimate of the number of eligible patients rose by 346 in year two, and then by 137 for each of the next three years to year five in which the estimated cost to NHS was [redacted].

CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

Table 1 shows the MS decision problem with rationale for deviations from the NICE scope.

Table 1: Manufacturer's indicated scope (from MS Table 4.1)

	Final scope issued by NICE	Decision problem addressed in the submission	Manufacturer's Rationale if different from the scope
Population	Adults with active autoantibody-positive systemic lupus erythematosus	**Phase 3 Trial Population** Adults with active autoantibody-positive systemic lupus erythematosus. **High Disease Activity Subgroup** Adults with active autoantibody-positive systemic lupus erythematosus with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥ 10.	Mindful of NHS resources, the proposed population of interest to this decision problem is a subgroup of the Phase 3 trial population which applies the additional criteria of evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI disease activity score of ≥ 10 This subgroup experienced an additional treatment effect to belimumab over and above the Phase 3 trial population and is aimed at identifying SLE patients at the greatest risk of experiencing long-term organ damage.
Intervention	Belimumab as an add on to standard therapy	Belimumab 10mg/kg administered as an intravenous infusion over a one hour period on days 0, 14 and 28, and at 4 week intervals thereafter in addition to standard therapy.	

	Final scope issued by NICE	Decision problem addressed in the submission	Manufacturer's Rationale if different from the scope
Comparator(s)	• Standard therapy alone For people in whom it is considered appropriate: • Rituximab plus standard therapy • Cyclophosphamide plus standard therapy	• Standard therapy which comprises (alone or in combination): antimalarials, NSAIDs, corticosteroids, or other immunosuppressants (azathioprine, methotrexate, and mycophenolate mofetil) • Rituximab plus standard therapy for the more severe SLE sub-population	Despite failing to meet primary or secondary outcomes in a Phase 2/3 SLE trial, rituximab, is used in the more severe patient population in addition to standard therapy. Therefore, rituximab plus standard therapy is a relevant comparator. The patient population and outcomes measured are not comparable to those in the belimumab trials. Therefore, conducting indirect comparisons of efficacy are problematic and have not been incorporated into the cost-effectiveness model. However, the benefits of belimumab compared with rituximab will be discussed in the written submission. Cyclophosphamide, whilst used in the more severe patient population, is largely reserved for the treatment of lupus nephritis. This is not the proposed Target population for belimumab, therefore, cyclophosphamide plus standard therapy is not a relevant comparator. In addition, adverse effects associated with long-term exposure to cyclophosphamide including bladder cancer, bone marrow suppression, haematologic malignancies, infections, myelodysplasia, and infertility limit the appropriateness of cyclophosphamide given that a high proportion of patients are women of childbearing age.

	Final scope issued by NICE	Decision problem addressed in the submission	Manufacturer's Rationale if different from the scope
Outcomes	The outcome measures to be considered include: • disease activity • incidence and severity of flares • mortality • health-related quality of life, including fatigue • adverse effects of treatment	The outcome measures included in the cost-effectiveness model are: • Disease activity • Incidence and severity of flares • Mortality • Health-related quality of life • Disease progression in terms of long-term organ damage – As discussed at the scoping workshop, although not collected in the clinical trials, long-term organ damage will be considered in the assessment of cost-effectiveness based on modelled data from the Johns Hopkins Lupus Cohort Additional endpoints discussed in the written submission and not included in the health economic model are: • Fatigue - In the Phase 3 trials this was measured using the FACIT-Fatigue instrument and was reported as the mean change in scale score at Weeks 12, 24, 52 and 76 (BLISS-76 only) • Adverse events of treatment	Adverse effects of treatment have not been included in the base case economic model as significant differences between treatments were not noted from the two pivotal Phase 3 trials. The side effect profile of belimumab will be discussed in the clinical section of the submission.

	Final scope issued by NICE	Decision problem addressed in the submission	Manufacturer's Rationale if different from the scope
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	• Cost effectiveness will be expressed in terms of incremental cost per quality-adjusted life year • The time horizon for the model will be lifetime • Costs will be considered from an NHS and Personal Social Services perspective	Not applicable.
Subgroups to be considered	None outlined in scope.	See population section above.	See population section above.
Special considerations, including issues related to equity or equality	None outlined in scope.	It will be important to acknowledge the innovative nature of belimumab in the treatment of SLE. There is a limitation with the current cost per QALY methodology not able to capture all the benefits of belimumab (i.e. avoidance of corticosteroids, impact of fatigue and loss of productivity). SLE has a significantly greater impact on certain ethnic groups and is most prevalent in woman of childbearing age.	

2.3 Population

The manufacturer's scope specified two populations: the Phase III trial population (adults with active autoantibody-positive SLE), and a High Disease Activity Subgroup (HDAS).

The submitted clinical effectiveness evidence came from two multicentre international Phase III RCTs (BLISS-52 and BLISS-76). The geographical location of study centres differed considerably between trials. In BLISS-52 there were 90 centres: in 13 countries in Latin America there were 38 centres (Argentina, Brazil, Chile, Colombia and Peru), in Asia-Pacific there were 41 centres (Australia, Hong Kong, India, Korea, Philippines and Taiwan) and in Eastern Europe there were eleven centres (Romania and Russia). In BLISS-76 there were 136 centres in 19 countries in North America (Canada, Costa Rica, Mexico, Puerto Rico and USA) and Europe (Austria, Belgium, Czech Republic, France, Germany, Israel, Italy, The Netherlands, Poland, Romania, Slovakia, Spain, Sweden and UK); North America (65 centres) and Europe (62 centres) contributed 93% of the centres in BLISS-76. These geographical differences were reflected in racial differences between the populations in the two trials. Although both trials included adults with auto-antibody positive active SLE it appears clear that the population in BLISS-76 is more likely to be similar to that in England and Wales than that from BLISS-52. It is reasonable to assume that the results from BLISS-76 will be more generalisable to the UK. This would be of little consequence if the clinical results were consistent between trials; however this was not so for some outcomes and in general very little clinical benefit was observed in BLISS-76 compared to some benefits in BLISS-52.

The manufacturer's scope also specified a HDAS termed the "Target" population and described as the focus of the submission. The identification of the Target population, and the evidence for clinical effectiveness of belimumab in the Target population, arose from post hoc analyses of the two BLISS trials. The rationale for this deviation from NICE scope was largely on economic grounds in that cost effectiveness was more favourable. Because the BLISS-76 trial subpopulation is more likely to match high disease activity patients in the UK than the BLISS-52 subpopulation, it is arguable that the BLISS-76 Target population is the most appropriate.

The Target population was defined as: *"Adults with active autoantibody-positive systemic lupus erythematosus with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥ 10"* [MS Page 53]. There are undoubtedly patients in the UK who correspond to the Target population; however, according to expert clinical opinion, the SELENA SLEDAI is not commonly used to define high disease activity and it

may be difficult to estimate the number of patients in the UK who fit this definition. The manufacturer's estimate of 4,150 patients across England and Wales is presented on Page 310 of the submission.

The population proposed in the license application is a high disease activity subgroup, termed "population A" in the submission, and defined in the Summary of Product Characteristics (SPC) in Box 2: Proposed license population.

Box 2: Proposed license population

"...adult patients with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g. positive for anti dsDNA, low complement) despite standard therapy."

The submission presented very little evidence about the effectiveness of belimumab in the proposed licensed population (only one Figure was given (MS Figure 5.3; Page 96). The submission estimated that there are 6,348 "*population A*" patients in England and Wales. The Target population represents a subpopulation (~64.5%) of population A patients.

2.4 Intervention

The intervention described in the submission matches that in the NICE final scope

Belimumab is a human IgG1λ monoclonal antibody, expressed in a murine cell line that binds to soluble human B-lymphocyte stimulator (BLyS) and inhibits its biological activity. BLyS influences differentiation, survival and activation of B lymphocytes. In the proposed license Belimumab is delivered at 10mg/Kg by a one hour IV infusion. It is an add-on therapy to standard of care that commonly consists of a range of treatments (NSAIDs, corticosteroids, immunosuppressants and antimalarials) used alone or in various combinations.

Belimumab awaits marketing approval in Europe. Application was filed with the European Medicines Agency (EMA) on 4th June 2010 and is now under review via the Centralised procedure. CHMP opinion was expected in May 2011 followed by a Commission decision on European marketing authorisation in July 2011. [redacted]

In March 2011 the USA FDA approved belimumab for reducing disease activity in adult patients with autoantibody positive SLE. This is a wider population than that encompassed in the European license application according to the SPC document submitted by the manufacturer.

2.5 Comparators

Three comparators were identified in the NICE final scope: standard of care (SoC), rituximab, and cyclophosphamide. The clinical effectiveness and economic sections of the submission did not quantitatively consider rituximab or cyclophosphamide as comparators, only SoC was formally assessed.

The MS justifies the exclusion of rituximab as a comparator on the following grounds:

- There has been no head to head trial of rituximab versus belimumab;
- Outcome measures used in rituximab and belimumab trials have differed to the extent that there is little possibility of undertaking meaningful indirect comparison meta-analysis;
- Rituximab has not been shown to be effective versus SoC whereas belimumab has, therefore by implication belimumab is unlikely to be less effective than rituximab;

[redacted]

Regarding effectiveness, although the primary end point was not reached in the Phase II/III rituximab RCT (EXPLORER)[10] the ERG's clinical expert indicated that the EXPLORER end point was more stringent than that in the BLISS trials because it registered a sustained response (once a patient was classified as a non-responder they remained so classified for the remainder of the trial), whereas the primary end point in BLISS was a group response in which a non-responder could later become classified as a responder for the primary end point at week 52.

A literature search undertaken by ERG revealed published information on SLEDAI and SF 36 changes in the EXPLORER trial which might have been used for comparison with the BLISS trials. Furthermore, RCTs for both drugs recorded BILAG changes thus offering the potential for an indirect comparison to be undertaken[10,11]. For these reasons the ERG requested clarification regarding the manufacturer's justification for not considering rituximab as a comparator.

The manufacturer responded with further justification as shown in Box 3.

Box 3: Justification for no formal comparison of belimumab and rituximab

The main reason for this decision relates to important differences in patient selection and consequently the treatment management protocols employed in the studies, …….. The patients in the rituximab Phase 2/3 trial had significant and acute disease activity at entry to the study; 53% had at least one BILAG A score (severe disease activity) and a further 28% had at least 3 BILAG B scores (please note that although a BILAG B score represents moderate disease activity, the presence of 3 BILAG B scores in some organs indicates more severe disease activity). Initially, patients were receiving very high daily doses of prednisone (mean 45.9 mg ±16.4mg) to treat the significant level of disease activity and this dose was to be tapered where possible during the trial. In addition, all patients were receiving one immunosuppressant at study entry. In contrast, the patients in the BLISS studies were a broader population and not all patients were experiencing major disease flares at study entry requiring the very high doses of steroids seen in the rituximab trial. Even in the high disease activity subgroup (Target population), only 19.3% had at least one BILAG A score at baseline, the average prednisone or equivalent dose was 12.3 mg ± 9.6mg and 53% were on an immunosuppressant. In particular, we believe that the differences in the use of steroids and immunosuppressants to manage disease activity between the rituximab and BLISS trials and consequently the differences in the type of response observed in the placebo arms render the studies incomparable.

Justification for excluding cyclophosphamide as a comparator was stated as shown in Box 4.

Box 4: Justification for no formal comparison of belimumab and cyclophosphamide

Cyclophosphamide, whilst used in the more severe patient population, is largely reserved for the treatment of lupus nephritis. This is not the proposed Target population for belimumab".

The submitted SPC is shown in Box 5.

Box 5: From the submitted SPC document

There are no or insufficient data available on the effects of Benlysta in patients with severe active lupus nephritis or severe active central nervous system lupus. Therefore, Benlysta cannot be recommended to treat these conditions.

2.6 Outcomes

The NICE and manufacturer's scopes state that the outcome measures to be considered include: disease activity, incidence and severity of flares, mortality, HRQoL including fatigue, and adverse effects of treatment. All these are reported in the MS.

The primary outcome in the BLISS trials was the proportion of responders at week 52. To estimate the proportion of responders a novel composite outcome measure, the SLE Responder Index (SRI) was introduced. The SRI was developed in conjunction with the FDA to be used in the BLISS trials. The SRI outcome aims to detect improvement in disease activity in terms of resolution of an SLE manifestation or manifestations (estimated using the SELENA SLEDAI instrument) while guarding against the possibility that this improvement might mask detrimental involvement of new organ systems (estimated using the BILAG) index) or an overall deterioration in well being (estimated using a PGA).

These three components, SELENA SLEDAI and BILAG and PGA had to be satisfied according to pre-specified requirements before a patient could be classified as a responder. These requirements were as follows:

- A $\geq$ 4 point reduction in the SELENA-SLEDAI score compared to baseline;
- No worsening (an increase of no more than 0.3 points) in PGA score compared to baseline;
- No new BILAG A organ domain scores or no 2 new BILAG B organ domain scores at time of assessment compared to baseline.

The SELENA SLEDAI instrument detects the presence of a manifestation of SLE disease. It encompasses 24 individual SLE manifestations, each with a weighted score from 1 to 8 points. Assessment relates to the preceding 10 days. Each manifestation must be related to lupus. A summed score of $\geq$ 6 across manifestations is considered active disease. A decrease of 4 points relative to previous assessment is thought to equate to a clinically meaningful improvement. For most manifestations there is no intermediate score, the item is registered as presence or absent so that a SELENA SLEDAI item generally can only improve by its resolution. The tool is therefore a measure of improvement and is not designed to assess worsening of a manifestation once present. The SRI uses the BILAG and PGA as measures of worsening.

The BILAG CLASSIC instrument includes 86 items grouped into 8 organ systems, general (5 items), mucocutaneous (18 items), neurological (15 items), musculoskeletal (9 items), cardiorespiratory (12 items), vasculitis (8 items), renal (11 items), and hematological (8 items)[2]. A score is calculated for each system depending on the SLE clinical manifestations (or signs and symptoms) present and whether they are new, worse, the same, improving, or not present in the last 4 weeks compared with the previous 4 weeks. A BILAG A score is given for a disease manifestation considered sufficiently severe to normally require high-dose steroids (prednisolone > 20 mg/day or equivalent) and/or immunosuppressive / cytotoxic agents under normal circumstances. A more moderate manifestation, which it would be considered appropriate to treat with lower dose steroids, antimalarial drugs or NSAIDs, constitutes a BILAG B score. A mild symptomatic manifestation that would require only symptomatic therapy (e.g. analgesics and NSAIDs) constitutes a BILAG C score. If there are no current symptoms, but the system has previously been involved, then a BILAG D score is recorded, while if the system has never been involved, a BILAG E score is assigned.

2.7 Economic analysis

The manufacturer's economic analysis is in line with that stipulated in the NICE scope. The MS presented its economic assessment in terms of incremental cost per QALY and has modelled outcomes using a lifetime horizon. Costs are considered from an NHS and PSS perspective.

2.8 Other relevant factors

Special considerations and issues raised in the manufacturer's scope include: 1) the innovative nature of belimumab for SLE; 2) the inability of the utility method to capture the QoL of SLE patients sufficiently sensitively; and 3) the impact of SLE on particular ethnic groups and on women of childbearing age. The proposed SPC specifies that belimumab should not be administered to pregnant women or to those planning pregnancy and therefore the special consideration relating to women of childbearing age appears to be of marginal relevance.

There were no issues identified in the NICE scope.

3 CLINICAL EFFECTIVENESS

3.1 Critique of manufacturer's systematic review

The objective of the manufacturer's systematic review was stated in Box 6 (MS Page 60).

Box 6: Objective of systematic review

A systematic review of the published literature was conducted to identify all relevant published randomised controlled trials (RCTs) for belimumab and relevant comparators in SLE.

3.1.1 Description of manufactures search strategy

Two clinical literature searches are reported in the MS; one to identify RCTs and one to identify observational studies (MS Appendix 2). The purpose of the latter search was not explicitly stated. The search strategies were of good quality (a summary of the ERG's assessment is in Appendix 2). The ERG considers it unlikely that the search would have been missed relevant studies.

The searches undertaken by the manufacturer to identify all relevant RCTs were conducted on 8th December 2010. Four electronic databases were searched (Embase, Medline, Medline In-Process, The Cochrane Library). The search strategy utilised an appropriate combination of free-text and thesaurus terms to identify the patient group (systemic lupus erythematosus), the intervention (belimumab) and the comparators. A date limit and a search filter were applied to the Embase and Medline searches to limit them to studies published after 1970 and to a particular type of evidence (RCTs), which was appropriate. The search filter used in Medline closely resembles the SIGN RCT filter[12], but misses several lines relating to publication type indicating that an old version may have been used. No language restrictions appear to have been applied. In addition to database searches, hand searching was undertaken of reference lists; the proceedings of three relevant conferences between 2006-2010 and four clinical trial registers (Clinical Trials, International Standard Randomised Controlled Trial Number (ISRCTN) Register, UK Clinical Trials Gateway, metaRegister of Controlled Trials).

The searches undertaken by the manufacturer to identify non-RCT evidence were conducted on 3rd March 2011. Four electronic databases were searched (Embase, Medline, Medline In-Process, The Cochrane Library). The search strategy utilised an appropriate combination of free-text and thesaurus terms to identify the patient group (systemic lupus erythematosus) and the intervention (belimumab). Terms to identify comparators were not included. A search filter was applied to the Embase and Medline searches to limit them to a particular type of evidence (observational studies), which was not appropriate in Medline in light of the small

number of studies retrieved before the filter was applied (66). The search filter used was the SIGN observational study filter[12] No date restrictions appear to have been applied in the search strategies themselves, although this is unclear as MS Appendix 6 states that the date span of the search is "Medline & Medline In-Process: 1950 to present day and Embase: 1980 to present day". In addition to database searches, hand searching was undertaken of reference lists; the proceedings of three relevant conferences between 2006-2010 and two clinical trial registers (Clinical Trials, UK Clinical Trials Gateway). Whilst the ERG was not able to check the search results, the search strategies were of adequate quality.

The database search alone yielded 3774 references (MS Fig 5.1). It was not possible for the ERG to attempt to reproduce the manufacturer's study selection procedure because of the large number if publications retrieved and because the description of the manufacturer's selection procedure was unclear (see below). An independent selection of studies by the ERG, effectively a separate systematic review, was not within the ERG remit or feasible within time constraints for such a large number of references.

3.1.2 Inclusion and exclusion criteria used for study selection

MS Figure 5.1 provides a flow diagram for the selection of studies. With regard to selection of studies for inclusion, the MS Page 60 states as shown in Box 7.

Box 7: MS Page 60

The inclusion and exclusion criteria were chosen to identify all relevant RCTs

Details of these criteria were not clear; they may be those in MS Table 5.1 entitled "Eligibility criteria used in search strategy" and shown in Table 2: MS Table 5.1 Page 61 Eligibility criteria used in search strategy. However, since ERG could find little relationship between the criteria listed and the studies listed as included studies, the ERG considers it is possible that MS Table 5.1 (see Table 2) actually represents a summary of the search strategy, in which case a formal statement of inclusion criteria was not submitted.

Table 2: MS Table 5.1 Page 61 Eligibility criteria used in search strategy

	Clinical effectiveness
Inclusion criteria	Population - Adults (≥ 18 years) with systemic lupus erythematosus (SLE); studies were also included for SLE patients with kidney involvement - Interventions o Belimumab o Rituximab o Mycophenolate mofetil o Prednisolone and other steroids o Hydroxychloroquine and other antimalarials o Azathioprine o Cyclophosphamide o Methotrexate Outcomes - Change in SELENA-SLEDAI score (Safety of Estrogens in Lupus Erythematosus National Assessment Systemic Lupus Erythematosus Disease Activity Index) - Change in BILAG score (British Isles Lupus Assessment Group) - Change in PGA (physician global assessment scale) - Change in SLICC score (Systemic Lupus International Collaborating Clinics) - Change in number/frequency of flares - Quality of life - Reduction in steroid use - Medical resource utilisation - Fatigue (e.g. FACIT, Functional Assessment of Chronic Illness Therapy score) - Adverse events including: o Incidence and severity (grade) of all adverse events (AEs) reported o Withdrawals due to AEs o Mortality o SAEs Study design - RCT, both cross-over and parallel, blinded and open-label designs Language restrictions - Only English publications (if only the abstract was in English, this would be included)
Exclusion criteria	Population - Studies enrolling patients with only active lupus nephritis were excluded Interventions - Non-specified Outcomes - Non-specified Study design - Designs other than RCT Language restrictions - Publications in languages other than English

Thus the MS was unclear about how or if the criteria listed in Table 5.1 were actually applied to the publications retrieved from searching; for example, although a search for uncontrolled studies was undertaken one of the exclusion criteria stated in Table 5.1 is that non-RCT study

designs were to be excluded. After seeking clarification the ERG remain doubtful that the criteria from Table 5.1 were systematically applied because many studies were excluded on the basis that they lacked a requirement for patients to have active autoantibody-positive SLE or because patients were receiving azathioprine, yet active autoantibody disease is not a specified inclusion criterion in Table 5.1 and azathioprine is listed as an included intervention rather than an excluded one

In short the MS and the manufacturer's response to request for clarification fail to indicate clearly the criteria used for study inclusion and exclusion.

3.1.3 Studies included

The MS systematic review provided confusing information regarding which studies were included and which were excluded. MS Figure 5.1 and Page 62 of the submission state that 43 publications were included. These are presented in Box 8.

Box 8: Statement of the number of publications included

The number of included publications was 43 (36 full publications plus seven abstracts), including eight publications (of four trials) of belimumab and 35 publications of other interventions.

The MS provided details of only 11 publications, rather than 43, (submission Tables 5.2, 5.3 and 5.4). The ERG sought clarification and a full list of the 43 identified publications was supplied together with reasons for exclusion of excluded studies. This list is reproduced in Appendix 3. The clarification implies that of the 43 publications identified nine were classed as "included". Eight of these were publications on four industry sponsored belimumab studies (RCTs: LBSL02, BLISS-52, BLISS-76; and Phase I study LBSL01), while the ninth described an RCT of rituximab conducted in adults with moderate-to-severe active SLE[10] which was not listed in the MS as an included study.

The clarification list of 43 publications was unclear on the status of the two hydroxychloroquine publications shown in MS Table 5.3 and described therein as "linked publications of competitor drugs that were also included in the systematic review." In the clarification list as "Reason for exclusion" the entry for both studies reads "Withdrawal study in patients with stable SLE"; this may represent a reason for exclusion. The ERG searched the MS text for any further reference to these two hydroxychloroquine publications but could find none.

A publication describing the industry phase I study LBSL01 was listed as "included" but this study was not discussed in the MS. The manufacturer's stated reason for this is reproduced in Box 9.

Box 9: Reasons for not including study LBS01

As this was a small (n=70) exploratory study of limited duration, designed primarily to demonstrate safety and tolerability in humans, it does not reflect the proposed clinical use of belimumab and therefore will be excluded from further discussion.

3.1.3.1 Important included studies

No RCTs were found that compared belimumab with an alternative active intervention. The most important belimumab studies identified were three industry-sponsored RCTs conducted in adults comparing belimumab plus standard care with placebo plus standard care (trials: LBSL02[13], BLISS-52[7], and BLISS-76) together with an uncontrolled extension (LBSL99) of LBSL02. One rituximab RCT (EXPLORER trial[10]) was included in narrative discussion of potential comparators. Brief details of these studies are shown in Table 3.

Table 3: Important studies included in manufacturer's submission

ID Year ψ	Study type	Study duration	Patient Age, yr	Treatment Groups ¥	N (ITT)	Countries (% enrolled)
LBSL02 2006	Phase 2 Efficacy and Safety	52 wk	20 - 75	Bel 1mg/kg IV** Bel 4mg/kg IV* Bel 10mg/kg IV* Placebo**	114 111 111 113	USA (98%) Canada (2%)
BLISS-76 2009	Phase 3 Efficacy and Safety	76 wk	18 - 73	Bel 1mg/kg IV* Bel 10mg/kg IV* Placebo**	271 273 275	USA and Canada (53%) West Europe (25%) East Europe (11%) Latin America (11%)
BLISS-52 2009	Phase 3 Efficacy and Safety	52 wk	18 - 71	Bel 1mg/kg IV* Bel 10mg/kg IV* Placebo**	288 290 287	Latin America (50%) Asia (38%) East Europe and Australia (13%)
LBSL99 2006	Safety extension of L02			Bel 10mg/kg IV*	296	USA and Canada (100%)
EXPLORER	Phase 2/3 Efficacy and Safety	52 wk	16 - 75	Rit 1000mg# Placebo***	169 88	North America (100%)

ψ Year study subject enrolment ended
¥ All treatments were additional to standard care
* Bel = Belimumab 1, 4, or 10mg/kg administration by IV infusion on days 0, 14, 28, and every 28 days thereafter
** Placebo by IV infusion on days 0, 14, 28, and every 28 days thereafter
*** Placebo by IV infusion on days 0, 14, 167 and181
# Rit = Rituximab on days 0 and 14

3.1.4 Details of any relevant studies that were not included

The ERG conducted a systematic search for randomised controlled trials of belimumab and of rituximab; no relevant studies additional to those included in the MS were identified.

3.1.5 Summary statement on MS systematic review

The manufacturer's systematic review was confused. Although the search strategy was of good quality the use of the retrieved references to identify relevant studies for inclusion was not well described. The ERG remains unclear regarding the methods used and the list of included studies both in the MS, and the response to request for clarification was ambiguous. Despite these non-systematic aspects, studies relevant to the decision problem have been identified and the studies presenting evidence on belimumab appears complete, although a rigorous check would require a separate and independent systematic review.

3.2 Submitted clinical evidence results

3.2.1 Scope and synopsis of the studies providing clinical evidence

Belimumab was administered as additional therapy to "standard of care" and was compared to placebo plus "standard of care". No formal comparison of belimumab vs any other active intervention (rituximab) was attempted.

Four belimumab studies provided clinical evidence: three RCTs: LBSL02, BLISS-52, and BLISS-76, and an uncontrolled study (LBSL99) that was an extension of LBSL02. MS Tables 5.4 and 5.5 provide full details of these studies.

- A phase II RCT (LBSL02) with four patient groups receiving infusions of placebo (n=113), or 1mg (n=114) or 4mg (n=111) or 10mg (n=111) belimumab/kg. A peer reviewed full publication of trial LBSL02 appeared in 2009.[13]

- Two phase III RCTs, BLISS-52 (n=865) and BLISS-76 (n=819), conducted simultaneously each with three randomised groups receiving placebo or 1mg/kg or 10mg/kg belimumab infusions. A peer reviewed full publication of the BLISS-52 trial appeared in 2011,[7] BLISS-76 has yet to appear as a peer reviewed full publication.

- LBSL99, a Phase II Continuation Study of the phase II RCT LBSL02.

Table: 4 summarises the main features of the four studies. Further details of study design and patient demography are discussed in the following section of this report.

Table: 4 Belimumab studies for safety and effectiveness evidence

ID Year*	Study type	Study duration	Patient Age, yr	Treatment Groups ¥ #	N (ITT)	Countries (% enrolled)
LBSL02	Phase 2 Efficacy and Safety	52 wk	20-75	Bel 1mg/kg IV Bel 4mg/kg IV Bel 10mg/kg IV Placebo	114 111 111 113	USA (98%) Canada (2%)
BLISS-76	Phase 3 Efficacy and Safety	76 wk	18 - 73	Bel 1mg/kg IV Bel 10mg/kg IV Placebo	271 273 275	US and Canada (53%) West Europe (25%) East Europe, (11%) Latin America (11%)
BLISS-52	Phase 3 Efficacy and Safety	52 wk	18 - 71	Bel 1mg/kg IV Bel 10mg/kg IV Placebo	288 290 287	Latin America (50%) Asia (38%) East Europe and Australia (13%)
LBSL99	Safety extension of LBSL02			Bel 10mg/kg IV	296	USA and Canada (100%)

* Year study subject enrolment ended
¥ All treatments were additional to standard care
Bel = Belimumab 1, 4, or 10mg/kg administration by IV infusion on days 0, 14, 28, and every 28 days thereafter

For the assessment of safety, the submission pooled data from all belimumab arms of the three RCTs (LBSL02, BLISS-76, and BLISS-52 providing data for1458 patients) and from all placebo arms (providing 675 patients).

Although all patients in study LSBL02 had a history of anti DNA-antibodies, approximately 30% lacked positive auto-antibody status at recruitment. Consequently the MS excluded this study from the clinical effectiveness analyses and it was only included for assessment of safety. For the assessment of clinical effectiveness the submission presented results from BLISS-52 and BLISS-76 phase III trials for the placebo and 10mg/kg belimumab arms only. The explanation for excluding results for the 1mg/kg dose regimen was stated in Box 10 (MS Page 102).

Box 10: Manufacturer's reason for not including results for the BLISS 1mg/kg groups

Whilst a 1mg/kg dose was examined in the Phase III studies, we will only present results for the 10mg/kg belimumab dose as this is the dose submitted for Marketing Authorisation.

Since results for the 1mg/kg arms of the trials can provide information about consistency of response and the existence of a dose response relationship, when considered relevant the ERG have made use of public domain data provided in FDA documents pertaining to the USA licensing application for belimumab[3,5].

The submission compared clinical effectiveness of 10mg/kg belimumab vs placebo for the following populations:

- BLISS-52 alone
- BLISS-76 alone
- BLISS-52 plus BLISS-76 populations pooled across trials
- A HDAS (the "Target population") pooled across BLISS-52 and BLISS-76 trials

As results for the Target population in BLISS-52 and BLISS-76 were not supplied separately in the original submission, these results were requested and supplied during the clarification process. Results for Target population patients who received the 1mg/kg regimen are not in the public domain and the manufacturer stated that they were unable to supply these results within the time constraints of the clarification process because of the large amount of other information that was requested.

3.2.2 Description and critique of manufacturers approach to validity assessment

In the main text of the MS, validity assessment of the BLISS trials consisted only of a tabulated quality assessment checklist (MS Table 5.14 page 100); this is reproduced in the two left hand columns in Table 6 below. Further details were provided in MS Appendix 3 (BLISS trials) and in MS Appendix 9 (adverse event studies). The MS was not clear about how their assessment was conducted, or by who, or whether it was based upon the full HGS clinical trial reports or on other information.

A single ERG reviewer undertook an independent quality assessment of the Phase III trials. The MS provided insufficient information for full quality assessment and so additional information in FDA documents[3,5]and in the Navarra publication[14] of the BLISS-52 trial was also utilised. The assessment is summarised in Table 6 below.

Table 5: Quality assessment and ERG critique of BLISS-52 and BLISS-76 trials

Question	MS rating	ERG rating	ERG comment
Was randomisation carried out appropriately?	Yes	Yes	MS states randomisation was stratified and MS and Navarra[14]state that a computer generated randomisation schedule was created.
Was the concealment of treatment allocation adequate?	Yes	Yes	Unmasked pharmacist prepared unmarked treatment infusion bags; but MS not explicit whether the creation and ownership of the randomisation schedule was handled by a separate group who had no direct involvement in the study.
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Agree
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes	MS states: *patients, investigators, study coordinators, and sponsors were masked to treatment assignment during intravenous administration of the drug and assessment of the patients every 4 weeks during the 52-week trial*. But methods not described and adherence to blinding not investigated.
Were there any unexpected imbalances in drop-outs between groups?	No	No	Table 9 of the FDA briefing package[3]provides the relevant information
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	The submission did not report outcome results for the 1 mg/kg treatment arms of the trials. For the total population these are available in the public domain.
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	No	Analysis was done in a modified intention-to-treat (mITT) population of all randomly assigned patients who received a dose of belimumab. The mITT analysis was performed according to the treatment that a subject was randomized to receive, regardless of actual treatment received. This was appropriate.
Adapted from Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination			

Trial conduct

The BLISS trials were large, international, multi-centre, double-blind, placebo-controlled studies with three parallel groups, that employed a novel primary outcome measure which required proficiency training for assessors. According to the HGS FDA briefing document.[5]

"*they were conducted under Special Protocol Assessment agreement with the FDA with special agreement with respect to selected patient population, primary end point, sample size, stratification factors, statistical methods and concomitant medication controls*".

Randomisation

The MS Appendix 3 provided the following description of randomisation:

"*Patients who underwent all screening procedures and met the entry criteria were enrolled in the study and assigned to treatment by use of a central interactive voice response system. Patients were randomised in a 1:1:1 ratio to placebo, or belimumab 1 mg/kg or 10 mg/kg. Randomisation was stratified according to the SELENA-SLEDAI score (6–9 vs ≥10), proteinuria concentration (<2 g/24 h vs ≥2 g/24 h) at screening, and ethnic origin (African descent or indigenous American [Alaska Native or American Indian from North, South, or Central America] vs other).*"

ERG note that in the BLISS-52 trial 867 patients were randomised, and that 142 screened-patients who met inclusion criteria were not randomised; the corresponding numbers for BLISS-76 were: 826 randomised and 135 not randomised (data from MS Figures 5.4 and 5.5).

Allocation concealment

MS Appendix 3 states:

"An unmasked pharmacist prepared unmarked infusion bags for administration. Belimumab and placebo were both prepared as sterile and lyophilised vials (5 mL for belimumab 1 mg/kg; 20 mL for belimumab 10 mg/kg and placebo), and contained the same formulations, except without the active drug for placebo."

The ERG considers that the above provides some assurance that allocation concealment was maintained but notes the difficulties of maintaining concealment across large multi-centred studies.

Baseline balance between treatment groups

Data provided in the MS and in FDA documents[3,5] indicates that within each trial there was a reasonable balance between known and putative prognostic factor.

Blinding of treatment allocation

MS Appendix 3 states:

"Patients, investigators, study coordinators, and sponsors were masked to treatment assignment during intravenous administration of the drug and assessment of the patients every 4 weeks during the 52-week trial until the database was locked."

The above gives the manufacturer's description of blinding of care providers, patients and outcome assessors to treatment allocation. There is no mention of methods employed (e.g. all potential flares should be adjudicated by a data monitoring board blinded to treatment). The methods for, extent of, and any problems with, blinding were not described. In the ERG's opinion it is possible that BLISS-52 physician outcomes assessors might have been unblinded, thus explaining a more positive PGA in the intervention group in this study as compared to the PGA in the BLISS-76 study.

Imbalance of drop outs between groups

MS appendices 3 and 7 and state respectively:

"There were no differences among groups in discontinuation rates" "The three groups did not differ in reasons for discontinuation of treatment." and *"Withdrawals and dropouts were adequately reported".*

Drop outs were reasonably balanced between treatment arms. Infringement of concomitant medication rules was one reason for discontinuation of treatment, and this differed between treatment arms. According the FDA[3] analysis: *" unlike dropouts, 'medication failures' are not balanced across treatment groups (17%, 9%, and 10% for placebo, 1 mg/kg belimumab, and 10 mg/kg belimumab respectively in* BLISS-76 *study, and 11%, 7%, and 6% for* BLISS-52".

The MS did not provide relevant information about adherence of study medication (e.g. missed infusion due to missed clinic visits).

Intention to treat analysis

MS Appendix 3 states:

"Analysis was done in a modified intention-to-treat population, defined as all randomly assigned patients who received a dose of the study drug. This was appropriate and appropriate methods for handling missing data were outlined in the clinical study report."

The trials were analysed according to a modified intention treat (mITT) procedure. In BLISS-52 and in BLISS-76 respectively two of 867 randomised patients and 7 of 826 randomised patients withdrew before receiving medication. Outcome analyses were based on the remaining 865 and 819 patients according to their randomisation group. Thus the results of mITT analyses were unlikely to differ substantially from a full ITT.

Pooling of trials

The pooling of trial data across trials is considered in sections 4.2.4 and 4.2.6 in the current report.

Applicability to the UK and UK clinical practice

It is unclear how many of the 1684 patients recruited to the BLISS-52 and BLISS-76 trials were from UK centres. The ERG notes that patients in the trials were derived from other EU countries. The MS is unclear whether similar care pathways to the UK occur across all centres included in the trials.

3.2.3 Description and critique of manufactures outcome selection

The primary efficacy endpoint in both Phase III studies was the percentage of responders at week 52 estimated using the SRI. The SRI is a novel composite outcome which was developed in consultation with the FDA during protocol planning for the BLISS-52 and BLISS-76 trials. The manufacturer's submission to the FDA states that assessors received proficiency training in SRI outcome assessment at all the centres involved in the trials. SRI and other outcomes selected for reporting in the MS are listed in Table 6. The ERG considers these outcomes to be appropriate for the decision problem.

Table 6: Outcomes reported in MA

Outcome	Measure	Outcome specification
SLE Responder Index (SRI*)	% responders at wk 52	Specified primary outcome
Reduction in SLEDAI score by ≥ 4 points	% responders at wk 52	Specified major secondary outcome
Change in PGA score from baseline	Mean change at wk 24	Specified major secondary outcome
Steroid reduction weeks 40 to 52	% responders	Specified major secondary outcome
SF-36 Physical component summary score	Mean change at wk 24	Specified major secondary outcome
SLE Responder Index	% responders at week 76	Specified major secondary outcome
SLICC/ACR damage index	Mean change at wk 52	Specified secondary outcome
FACIT-fatigue scale mean change from baseline	Mean change at clinic visits	Specified secondary outcome
EQ-5D score	Mean change at clinic visits	Specified secondary outcome
Change in PGA score from baseline	Mean change at wk 52	Specified secondary outcome
SF-36 Physical component summary score	Mean change at wk 52	Specified secondary outcome
SLEDAI SLE flare index over 52 wks	Time to first flare	Specified secondary outcome
SLE Responder Index (SRI)	% responders at timed clinic visits	Other outcome reported
Modified SLE responder index	% responders at wk 52	Other outcome reported
No worsening in PGA score by ≥ 0.3	% responders at wk 52	Other outcome reported
No new BILAG 1A/2B domain scores	% responders at wk 52	Other outcome reported
Change in SLEDAI score from baseline	Mean change at week 52	Other outcome reported

* Composite outcome measure consisting of ≥ 4 points improvement in SLEDAI score, no worsening in PGA by ≥ 0.3 points and no new BILAG 1A or 2B domain scores
FACIT= Functional Assessment of Chronic Illness Therapy
EQ-5D = EuroQoL 5 dimensions
BILAG = British Isles Lupus Assessment Group
SLEDAI = Systemic Lupus Erythematosus Disease Activity Index
SF-36 = Short Form 36-Item Health Survey
SLICC = Systemic Lupus International Collaborating Clinics
ACR = American College of Rheumatology

3.2.4 Description and critique of the statistical approach used

The manufacturer's approach is described in Table 5.13 of the MS (Page 91).

Binary efficacy variables were assessed with a logistic regression model, continuous variables were analysed with an analysis of covariance model, and time-to-flare variables were analysed by use of a Cox proportional hazards model. All analyses were adjusted for baseline randomisation factors. In addition, the JHU observational cohort of patients was used to generate the analysis that was used in an SLE patient simulation.

The ERG reviewed the statistical approach submitted in the main report and notes that in general, the statistical methodologies proposed are suitable to these types of data. However, the ERG identified a number of concerns as shown below:

In order to identify baseline factors that were predictive of response at Week 52 irrespective of treatment received and to evaluate belimumab treatment effect adjusted for the predictive factors, a logistic regression main effects model was developed by the manufacturer based on the pooled data from the Phase III studies (BLISS-52 and BLISS-76).

The ERG notes that while the pooling of the two data sets might be considered appropriate, given that the trials were essentially identical in design and in the analysis of the primary endpoint, the approach used to account for potential between-study variability in the estimate of the baseline response or the uncertainty in the estimate of the population sampling variation was not appropriate (i.e. treatment-by-study interaction). It is not surprising that the P-values for the treatment-by-study interaction were not significant (interaction P-values > 0.5). This insignificant P-value is a reflection of the similarity between the two trials in terms of the primary endpoint and would not capture a real difference that might exist between the two trials.

A mixed model logistic regression would have been appropriate to account for the correlation structure between the two trials and any population sampling variation. Furthermore, a sensitivity analysis of the choice of correlation structure should have been conducted. Without taking into account the unobserved uncertainty or variability between the two trials, the ERG believes that the validity of pooling of data may have been overestimated.

The ERG also note that results of in the manufacturer's submission analysing time-to-flare variables did not take into account the time-varying effects of some of the covariates. A generalized mixed model with time-varying effects could have been considered to deal with the time-varying effect of covariates.

Sub-group analysis

The main submission indicates that a series of pre-specified and post-hoc subgroup analyses for efficacy data were conducted. A comparison between each belimumab treated group and the placebo group was performed by major subgroups which were pre-specified in each Phase III analytical plan.

With reference to the decision problem and the manufacturer's intention to explicitly identify patients who benefit the most, the ERG notes that some additional exploratory subgroup analyses which were not pre-specified in the individual analytical plans were evaluated using the pooled Phase III population Target or high disease activity sub-group. The subgroup of patients with evidence for serological disease activity (low complement and positive anti-dsDNA) and who additionally have a SELENA-SLEDAI disease activity score $\geq$ 10 at baseline. However, even though patients in this subgroup experienced the greatest treatment effect over and above the total pooled population, the ERG notes that this sub-group analysis was not pre-specified in the analytical plan. Therefore, the results of this sub-group analysis should not be regarded as definitive since the two trials were not powered to conduct this sub-group analysis.

The ERG notes that there was no attempt to summarise the studies by performing a meta-analysis or by conducting an incremental analysis.

3.2.5 Results from pivotal trials

The clinical effectiveness results in the MS are derived from the two BLISS trials.

3.2.5.1 BLISS trial study design and patient eligibility

Methodological details of the BLISS-52 and BLISS-76 trials were presented in Table 5.6 of the MS which is reproduced in Table 7.

Table 7: MS summary of BLISS trial methodology (from MS Table 5.6)

Trial no. (acronym)	**C1057 (BLISS-52)**	**C1056 (BLISS-76)**
Location	90 centres in 13 countries in Latin America (Argentina, Brazil, Chile, Colombia and Peru), Asia-Pacific (Australia, Hong Kong, India, Korea, Philippines and Taiwan) and eastern Europe (Romania and Russia).	136 centres in 19 countries in North America (Canada, Costa Rica, Mexico, Puerto Rico and US) and Europe (Austria, Belgium, Czech Republic, France, Germany, Israel, Italy, The Netherlands, Poland, Romania, Slovakia, Spain, Sweden and UK).
Design	Randomised, double-blind, placebo-controlled, parallel-group study.	As per BLISS-52.
Duration of study	52 weeks	76 weeks (primary end point at 52 weeks)
Method of randomisation	Patients who underwent all screening procedures and met the entry criteria were enrolled in the study and assigned to treatment by use of a central interactive voice response system. Patients were randomised in a 1:1:1 ratio to placebo, or belimumab 1 mg/kg or 10 mg/kg. Randomisation was stratified according to the SELENA-SLEDAI score (6–9 vs ≥10), proteinuria concentration (<2 g/24 h vs ≥2 g/24 h) at screening, and ethnic origin (African descent or indigenous American [Alaska Native or American Indian from North, South, or Central America] vs other).	As per BLISS-52.
Method of blinding (care provider, patient and outcome assessor)	Patients, investigators, study coordinators, and sponsors were masked to treatment assignment during intravenous administration of the drug and assessment of the patients every 4 weeks during the trial until the database was locked. An unmasked pharmacist prepared unmarked infusion bags for administration. Belimumab and placebo were both prepared as sterile and lyophilised vials (5 mL for belimumab 1 mg/kg; 20 mL for belimumab 10 mg/kg and placebo), and contained the same formulations, except without the active drug for placebo.	As per BLISS-52.
Intervention(s) (n=) and comparator(s) (n=)	Standard of care plus belimumab 1mg/kg (n=288) or belimumab 10mg/kg (n=290) or placebo (n=287) administered by IV infusion on Days 0, 14 and 28 and every 28 days thereafter for 48 weeks. Standard of care consisted of the following (alone or in combination): antimalarials, NSAIDs, corticosteroids or other immunosuppressants (azathioprine, methotrexate, and mycophenolate mofetil).	Standard of care plus belimumab 1mg/kg (n=271) or belimumab 10mg/kg (n=273) or placebo (n=275) administered by IV infusion on Days 0, 14 and 28 and every 28 days thereafter for 72 weeks. Standard of care consisted of the following (alone or in combination): antimalarials, NSAIDs, corticosteroids or other immunosuppressants (azathioprine, methotrexate, and mycophenolate mofetil).
Progressive restrictions placed on standard of care	In both BLISS-52 and BLISS-76, progressive restrictions were placed on standard of care as the study progressed. These are outlined in the Figure 5.2 below. (see following box)	
Primary outcomes (including scoring methods and timings of assessments)	The primary efficacy endpoint was the response rate at week 52, assessed with SLE Responder Index (SRI). With the SRI criteria, a responder was defined as having a reduction of at least 4 points in the SELENA-SLEDAI score (defined as clinically meaningful) [15], no new BILAG A organ domain score, no more than 1 new BILAG B organ domain score, and no worsening in PGA score (increase <0.3) at week 52 compared with baseline.	As per BLISS-52.
Secondary outcomes (including scoring methods and timings of assessments)	Major secondary endpoints: • Percent of subjects with ≥ 4-point reduction in SELENA-SLEDAI at Week 52. • Mean change in PGA at Week 24. • Percent of subjects with prednisone (equivalent) reduction ≥ 25% from baseline to ≤ 7.5 mg/day during Weeks 40 – 52 (in subjects whose prednisone equivalent dose was > 7.5 mg/day at baseline). • Mean change in SF-36 PCS at Week 24.	Major secondary endpoints: • As per BLISS-52. • Additionally, response rate (SRI) at Week 76.
Duration of follow-up	52 or 56 weeks dependent on participation in the continuation protocol.	76 or 80 weeks dependent on participation in the continuation protocol.

COMMENT

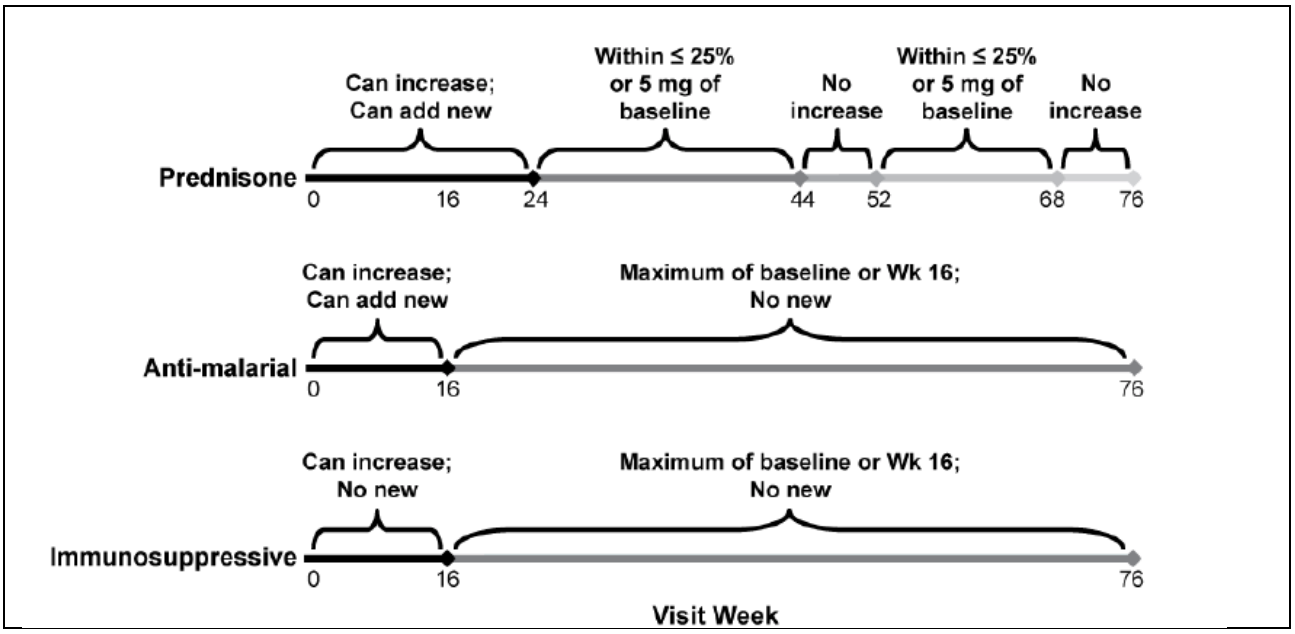
The two pivotal trials, BLISS-52 and BLISS-76, were international multicentre RCTs undertaken in different largely non-overlapping geographical regions. The geographical location of study centres differed considerably between trials. In BLISS-52 there were 90

centres: in 13 countries in Latin America there were 38 centres (Argentina, Brazil, Chile, Colombia and Peru), in Asia-Pacific there were 41 centres (Australia, Hong Kong, India, Korea, Philippines and Taiwan) and in Eastern Europe there were eleven centres (Romania and Russia). In BLISS-76 there were 136 centres in 19 countries in North America (Canada, Costa Rica, Mexico, Puerto Rico and US) and Europe (Austria, Belgium, Czech Republic, France, Germany, Israel, Italy, The Netherlands, Poland, Romania, Slovakia, Spain, Sweden and UK). North America (65 centres) and Europe (62 centres) contributed 93% of the centres in BLISS-76. These geographical differences were reflected in racial differences between the populations in the two trials. Although both trials included adults with auto-antibody positive active SLE it is arguable that the population in BLISS-76 is more likely to be similar to that in England and Wales than that from BLISS 52. It is reasonable to assume that the results from BLISS-76 are more generalisable to the UK.

Randomisation was stratified according to SELENA-SLEDAI score (6–9 vs ≥10), proteinuria concentration (<2 g/24 h vs ≥2 g/24 h) at screening, and ethnic origin.

Progressive constraints on standard care medications (immunosuppressives, anti-malarials and steroids) (see Figure 1) were imposed during the trials; these were implemented so as to increase the possibility of detecting improvement due to belimumab without interference from the effects of changing background standard care treatments.

Figure 1: Constraints on standard of care medications (MS Figure 5.2)



Patient eligibility for BLISS-52 and BLISS-76 was the same and summarised below in Table 8.

Table 8: Patient eligibility for BLISS-52 and BLISS-76 (From MS Table 5.7)

Trial no. (acronym)	Inclusion criteria	Exclusion criteria
C1057 **(BLISS-52)**	Adult patients (aged ≥ 18 years) who met the American College of Rheumatology criteria for systemic lupus erythematosus and had active disease (score ≥ 6 at screening on SELENA-SLEDAI) were eligible for enrolment. Other inclusion criteria were unequivocally positive ANA (titre ≥ 1:80) or anti-dsDNA antibody (≥ 30IU/mL), and a stable treatment regimen with fixed doses of prednisone (0–40mg/day), or non-steroidal anti-inflammatory, antimalarial, or immunosuppressant drugs for at least 30 days before the first study dose	The main exclusion criteria were severe active lupus nephritis or CNS lupus; pregnancy; and previous treatment with any B-lymphocyte-targeted drug (including rituximab), intravenous cyclophosphamide within 6 months of enrolment, and intravenous Ig or prednisone (>100 mg/day) within 3 months
C1056 **(BLISS-76)**	As per BLISS-52	As per BLISS-52
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee		

Relative to the whole trial population imbalance between treatment arms was more pronounced for the Target population in both trials, especially in BLISS-76 (see Appendix 4 of this report).

3.2.5.2 BLISS trials: demography of patients

Demographic characteristics of patients in BLISS-52 and BLISS-76 were presented in MS Tables 5.8 to 5.11 (see Appendix 4 of this report). Patients were mostly young females (74% ≤ 45 years of age; 94% female), a population which is representative of patients with SLE.

Selected characteristics for placebo and 10mg/kg groups taken from MS Table 5.8 are shown below in Table 9. Amongst all treatment arms pooled across the two studies 47% of patients were white, 23% American Indian, 21% Asian, and 8.8% black, however there were large differences in the racial makeup between the two studies reflecting the racial distributions in the geographic regions in which the trial centres were located. The substantial differences between trials in geographical and in racial distributions seen for the whole population were mirrored in the Target population Table 9 (Appendix 4).

Table 9: Demographic characteristics in the BLISS trials (adapted from MS Table 5.8)

Table 5.8. Selected demographic characteristics in Phase 3 trials

	BLISS-52			BLISS-76			Pooled Total Population		
Race[1]	**Placebo N = 287**	**10mg/kg N = 290**	**All N = 865**	**Placebo N = 275**	**10mg/kg N = 273**	**All N = 819**	**Placebo N = 562**	**10mg/kg N = 563**	**All N = 1684**
White	82 (28.6%)	71 (24.5%)	229 (26.5%)	188 (68.4%)	189 (69.2%)	569 (69.5)	270 (48.0%)	260 (46.2%)	798 (47.4%)
Asian	105 (36.6%)	116 (40.0%)	327 (37.8%)	11 (4.0%)	11 (4.0%)	28 (3.4%)	116 (20.6%)	127 (22.6%)	355 (21.1%)
Black	11 (3.8%)	11 (3.8%)	30 (3.5%)	39 (14.2%)	39 (14.3%)	118 (14.4%)	50 (8.9%)	50 (8.9%)	148 (8.8%)
Alaska Native or American Indian from North/Central/ South America	89 (31.0%)	92 (31.7%)	279 (32.3%)	36 (13.1%)	34 (12.5%)	103 (12.6%)	125 (22.2%)	126 (22.4%)	382 (22.7%)
Hispanic or Latino origin	143 (49.8%)	136 (46.9%)	420 (48.6%)	55 (20.0%)	56 (20.5%)	173 (21.1%)	198 (35.2%)	192 (34.1%)	593 (35.2%)

[1] Patients who checked more than 1 race category are counted under individual race category according to the minority rule as well as the multiracial category.

Differences also existed between studies in that BLISS-76 patients had longer disease duration and more organ damage (higher SLICC damage scores), and were using lower steroid dosages than BLISS-52 patients.

Both BLISS-52 and -76 populations presented a restricted range of SLE manifestations. The MS did not provide tabulated information for the frequency of SELENA SLEDAI manifestations at baseline, these are shown in Table 10 below based on the FDA discussion document3 for the whole BLISS populations, and in Table 11 for the target population. The majority of BLISS-76 participants had musculoskeletal and/or mucocutaneous manifestations of SLE as assessed by the SELENA SLEDAI disease activity index. Baseline disease involvement was generally well balanced within trial between the three treatment groups with the exception of rash. Higher proportions of placebo patients (68%) and patients in the 1mg/kg belimumab group (66%) had a rash at study entry as compare to patients in the 10mg/kg (56%). A similar pattern of SLE disease involvement at baseline was observed for subjects in BLISS-52, however, a lower rate of arthritis (59%) was reported compared to BLISS-76 (72%).

Table 10: Baseline SELENA SLEDAI involvement: whole population in BLISS trials

	BLISS-52				BLISS-76			
Condition (weight)	**Placebo (N=287)**	**Belimumab 1mg/kg (N=288)**	**Belimumab 10mg/kg (N=290)**	**Total (N=865)**	**Placebo (N=275)**	**Belimumab 1mg/kg (N=271)**	**Belimumab 10mg/kg (N=273)**	**Total (N=819)**
Organic Brain Syndrome (8)	0	2 (1%)	0	2 (1%)	1 (0%)	2 (1%)	3 (1%)	6 (1%)
Lupus HA (8)	4 (1%)	2 (1%)	4 (1%)	10(1%)	1 (0%)	4 (2%)	9 (3%)	14 (2%)
Vasculitis (8)	20 (7%)	16 (6%)	28 (10%)	64 (7%)	17 (6%)	20 (7%)	10 (4%)	47 (6%)
Arthritis (4)	165 (58%)	169 (59%)	173 (60%)	507 (59%)	206 (75%	193 (71%)	191 (70%	590 (72%)
Hematuria (4)	15 (5%)	16 (6%)	16 (6%)	47(5%)	5 (2%)	7 (3%)	8 (3%)	20 (2%)
Proteinuria (4)	50 (19%)	54(19%)	41 (14%)	145 (17%)	29 (11%)	23 (9%)	26 (10%)	78 (10%)
Rash (2)	176 (61%)	176 (61%)	182 (63%)	534 (62%)	87 (68%)	180 (66%)	154 (56%)	521 (64%)
Alopecia (2)	150 (52%)	138 (48%)	158 (55%)	446 (52%)	30 (47%	137 (51%)	116 (43%	383 (47%)
Mucosal Ulcers (2)	71 (25%)	52 (18%)	58 (20%)	181 (21%)	74 (27%)	57 (21%)	78 (29%)	209 (26%)
Low Complement (2)	183 (64%)	186 (65%)	198 (68%)	567 (66%)	160 (58%	149 (55%)	159 (58%	468 (57%)
Inc. DNA Binding (2)	205 (71%)	220 (76%)	218 (75%)	643 (74%)	175 (64%	168 (62%)	176 (65%	519 (63%)
Leukopenia (1)	18 (6%)	12 (4%)	9 (3%)	39 (5%)	16 (6%)	22 (8%)	23 (8%)	61 (7%)

Table 11: Baseline SELENA SLEDAI involvement: in the Target population in BLISS Trials

	BLISS-52		BLISS-76		Combined BLISS	
Condition (weight)	**Placebo (N=107)**	**Belimumab 10mg/kg (N=112)**	**Placebo (N=96)**	**Belimumab 10mg/kg (N=96)**	**Placebo (N=203)**	**Belimumab 10mg/kg (N=193)**
Organic Brain Syndrome (8)	0	0	1 (1.0%)	0	1 (0.5%)	0
Lupus HA (8)	1 (0.9%)	3 (2.7%)	0	2 (2.5%)	1 (0.5%)	5 (2.6%)
Vasculitis (8)	15 (14.0%)	19 (17.0%)	10 (10.4%)	5 (6.2%)	25 (12.3%)	24 (12.4%)
Arthritis (4)	65 (60.7%)	76 (67.9%)	83 (86.5%)	63 (77.8%)	148 (72.9%)	139 (72.0%)
Hematuria (4)	9 (8.4%)	7 (6.3%)	3 (3.1%)	6 (7.4%)	12 (5.9%)	13 (6.7%)
Proteinuria (4)	31 (29.0%)	28 (25.0%)	17 (17.7%)	21 (25.9%)	48 (23.6%)	49 (25.4%)
Rash (2)	74 (69.2%)	75 (67.0%)	72 (75.0%)	52 (64.2%)	146 (71.9%)	127 (65.8%)
Alopecia (2)	66 (61.7%)	69 (61.6%)	50 (52.1%)	38 (46.9%)	116 (57.1%)	107 (55.4%)
Mucosal Ulcers (2)	28 (26.2%)	20 (17.9%)	30 (31.3%)	22 (27.2%)	58 (28.6%)	42 (21.8%)
Low Complement (2)	107 (100.0%)	112 (100.0%)	96 (100.0%)	80 (98.8%)*	203 (100.0%)	192 (99.5%)
Inc. DNA Binding (2)	107 (100.0%)	112 (100.0%)	96 (100.0%)	81 (100.0%)	203 (100.0%)	193 (100.0%)
Leukopenia (1)	6 (5.6%)	4 (3.6%)	7 (7.3%)	10 (12.3%)	13 (6.4%)	14 (7.3%)

A specified major secondary outcome was the percentage of SRI responders at week 76. There was only a small difference between placebo and 10mg/kg belimumab (odds ratio and P value not submitted; odds ratio 1.31, 95% CI: 0.92 – 1.87, P = 0.1323 by logistic regression, taken from the FDA HGS briefing document.[5]

Relative to the whole trial population imbalance between treatment arms was more pronounced for the Target population in both trials, especially in BLISS-76 (Appendix 4).

Patients from BLISS-52 contributed more patients to the pooled Target population than did patients from BLISS-76 (55% and 45% respectively, and contributed a greater proportion of the patients receiving 10mg/kg belimumab (58% and 42% from each trial respectively); therefore effectiveness results pooled across trials will tend to reflect BLISS-52 outcomes more than BLISS-76.

3.2.5.3 BLISS trial results by outcome

Primary outcome: SRI at week 52

The pre-specified primary outcome in the BLISS trials was the proportion of responders at week 52 defined according to the composite SRI outcome measure. The results were provided in MS Table 5.15 and clarification Table A6.1 and summarised below in Table: 12.

Table: 12 Primary efficacy endpoint (SRI) at Week 52 (dropout-failure)

	BLISS-52		**BLISS-76**		**Pooled Total Population[4]**		**High Disease Activity Subgroup Pooled Total**		**High Disease Activity Subgroup BLISS-52**		**High Disease Activity Subgroup BLISS-76**	
SRI at Week 52	**Placebo N = 287**	**10mg/kg N = 290**	**Placebo N = 275**	**10mg/kg N = 273**	**Placebo N = 562**	**10mg/kg N = 563**	**Placebo N = 203**	**10mg/kg N = 193**	**Placebo N = 107**	**10mg/kg N = 112**	**Placebo N = 96**	**10 mg/kg N = 81**
No. (%) Response	125 (43.6%)	167 (57.6%)	93 (33.8%)	118 (43.2%)	218 (38.8%)	285 (50.6%)	77 (37.9%)	121 (62.7%)	44 (41.1%)	75 (67.0%)	33 (34.4%)	46 (56.8%)
Observed difference vs placebo (%)	-	14.03	-	9.41	-	11.8	-	24.8	-	25.9	-	22.4
OR (95% CI)[1] vs placebo	-	1.83 (1.30, 2.59)	-	1.52 (1.07, 2.15)	-	1.68 (1.3, 2.2)	-	2.7 (1.8, 4.1)	-	3.0 (1.7, 5.2)	-	2.5 (1.3, 4.6)
P-value[1]	-	0.0006	-	0.0207	-	< 0.0001	-	< 0.0001		0.0001	-	0.0045

[1] Odds Ratio (95% confidence interval) and p-values were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, covariates include baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.

In both trials SoC + 10mg/kg belimumab delivered a greater percentage of responders than SoC + placebo. The difference in percentage of responders in the belimumab group relative to placebo group for the whole population was 14% in BLISS-52 and 9.4% in BLISS-76. The corresponding adjusted odds ratios for a response in BLISS-52 and in BLISS-76 were respectively 1.83 (95% CI: 1.30, 2.59; P = 0.0006) and 1.52 (95% CI: 1.07, 2.15; P = 0.027). For the Target population pooled across trials the difference in percentage of responders in the belimumab group relative to placebo group was 24.8% and the adjusted odds ratio was 2.7 (95% CI: 1.8, 4.1; P < 0.0001). In BLISS-52 and BLISS-76 Target populations the difference

between groups was 25.9% and 22.4% respectively (odds ratio 3.0, 95% CI: 1.7, 5.2; P = 0.0001 for BLISS-52 and odds ratio 2.5, 95% CI: 1.3, 4.6; P = 0.0045 for BLISS-76).

Relative to the whole population the Target population generated results that were more supportive of belimumab. For the whole population and for the Target population BLISS-52 produced results more supportive of belimumab than did BLISS-76, however for the Target population the difference between trials was less than for the total population.

SRI at successive clinic visits and at week 76

The percentage of SRI responders was also reported at multiple follow up times (MS Figures 5.6, to 5.9 shown in Figure 2.

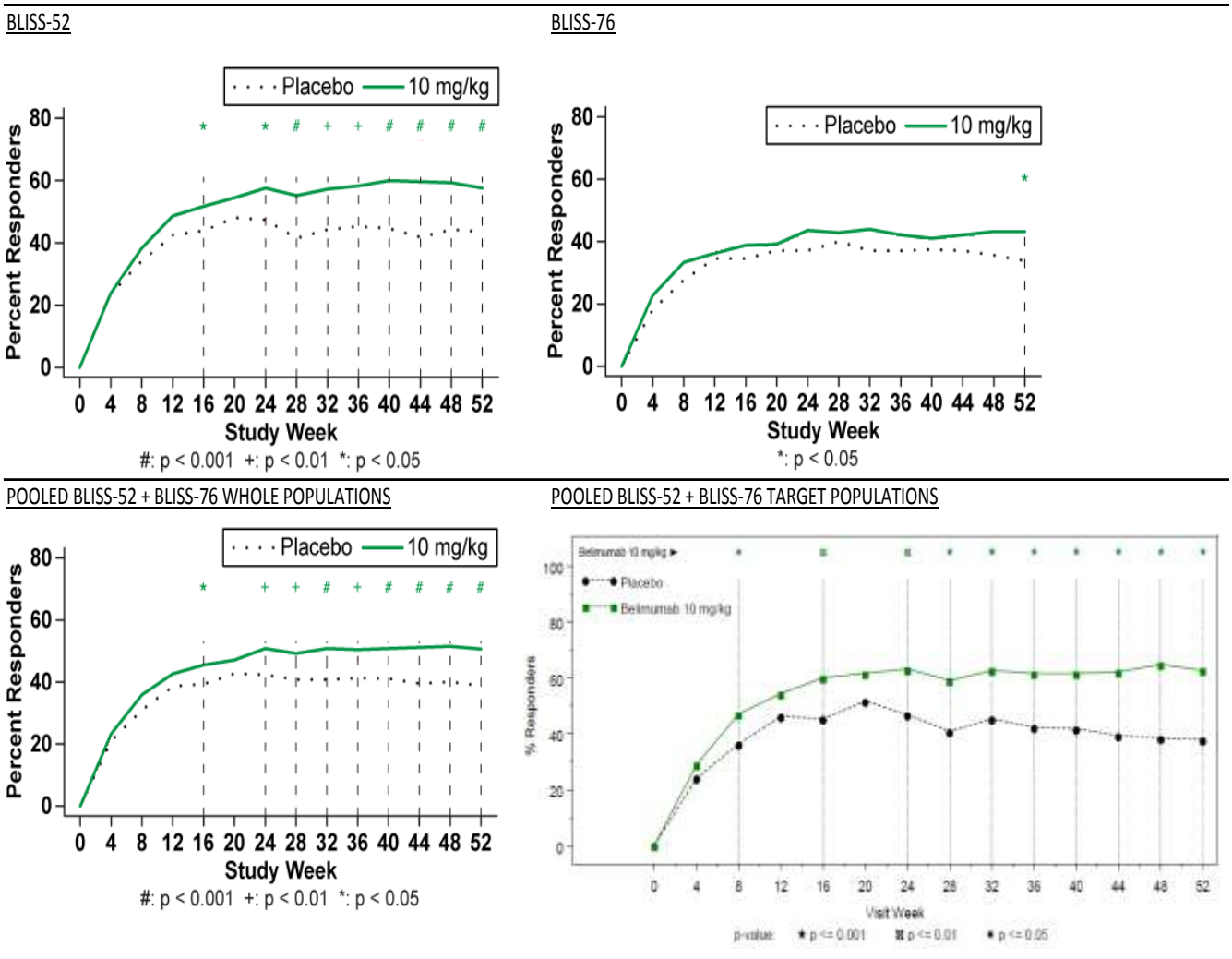


Figure 2: Percentage of SRI responders during follow up (from MS Figures 5.6 to 5.9)

For the Target population pooled across trials and in BLISS-52, at many times, a significantly greater response was observed for the belimumab group relative to placebo group

(significance tests uncorrected for multiple testing), however, for BLISS-76 the only time a significantly ($P < 0.05$) greater response was observed for the belimumab group was at week 52.

In the HGS/FDA[5]analysis there is little difference in response between 1mg/kg and 10mg/kg groups for BLISS-52.

A specified major secondary outcome was the percentage of SRI responders at week 76. There was only a small difference between placebo and 10mg/kg belimumab (odds ratio and P value not submitted; odds ratio 1.31, 95% CI: 0.92 – 1.87, P = 0.1323 by logistic regression, taken from the FDA discussion document[5]).

The HGS Briefing Document to the FDA[5] provided graphs for all three randomised groups (placebo, 1mg/kg belimumab and 10mg/kg belimumab) for the percentage of SRI responders observed at successive clinic visits up to 52 weeks for BLISS-52 and week 76 for BLISS-76. These graphs are in Figure 3. They indicate that in BLISS-76 there was a minimal difference in response between 1mg/kg and 10mg/kg groups. Baseline characteristics for the three groups (HGS Briefing Document Pages 87 to 100[5]) do not provide an obvious explanation for this result.

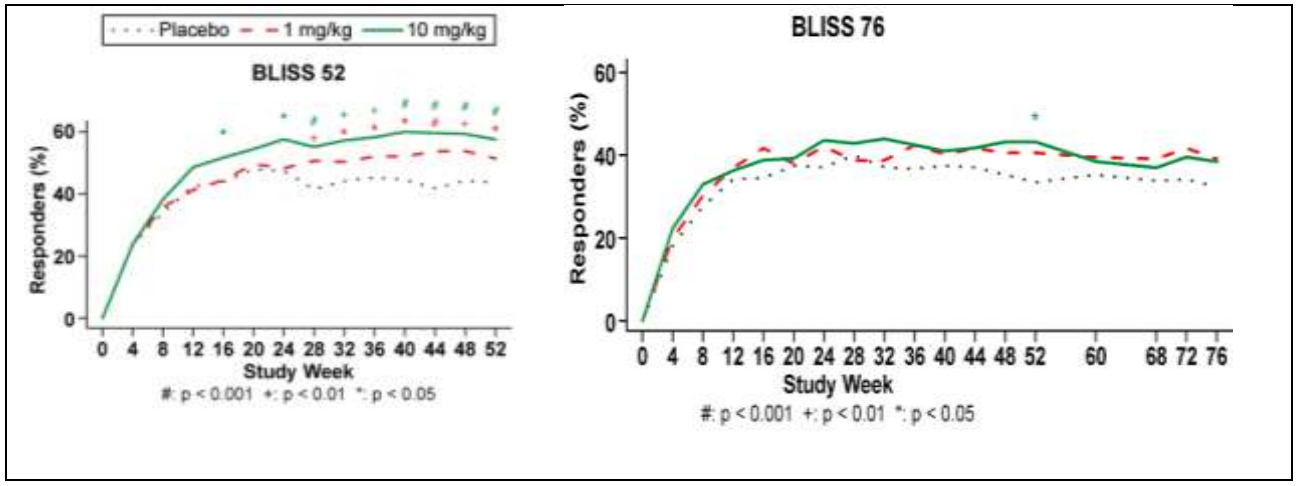


Figure 3: SRI percent responders over follow up (from HGS Briefing Document to FDA)

The ERG note that the percentage SRI responders observed at various follow up times is a group response and does not reflect sustained SRI response at the individual level. The graph line showing percentage of responders across the duration of the trials rose and fell at various follow up times, thus an individual non-responder could later improve sufficiently to be classified as a responder.

Modified SRI

To be classified as an SRI responder a patient is required to have a SELENA-SLEDAI score that is reduced by ≥ 4 points relative to baseline. A 4 point reduction in SELENA-SLEDAI can be achieved by normalisation of serological manifestations only (e.g. anti-dsDNA antibodies and complement). The MS presented an analysis of a modified SRI response in which the increased DNA binding and low complement items were removed from the SELENA-SLEDAI component of the SRI; the analysis was performed in patients who still had a SELENA SLEDAI score ≥ 4 at baseline after points for low complement and increased DNA binding were removed from the scale. During the clarification process the manufacturer provided modified SRI results for the Target or high disease activity population; these plus the information from the MS Page 111 are summarised in Table 13

Table 14: Modified SRI response at week 52

	BLISS-523		BLISS-763		Pooled Total Population2		High Disease Activity Subgroup Pooled Total		High Disease Activity Subgroup BLISS-52		High Disease Activity Subgroup BLISS-76	
Change from Baseline at Week 522	**Placebo N = 264**	**10 mg/kg N = 259**	**Placebo N = 255**	**10 mg/kg N = 245**	**Placebo N = 519**	**10 mg/kg N = 504**	**Placebo N = 203**	**10 mg/kg N = 193**	**Placebo N = 107**	**10 mg/kg N = 112**	**Placebo N = 96**	**10 mg/kg N = 81**
n(%) responders	127 (48.1%)	158 (61.0%)	92 (36.1%)	109 (44.5%)	219 (42.2%)	267 (53.0%)	42 (39.3%)	73 (65.2%)	29 (30.2%)	43 (53.1%)	71 (35.0%)	116 (60.1%)
OR (95% CI)[1]	-	-	-	-	-	-	-	3.0 (1.7, 5.2)	-	2.5 (1.4, 4.8)	-	2.8 (1.8, 4.2)
P-value1	-	-	-	-	-	-	-	0.0001	-	0.0036	-	<0.0001
10 mg/kg vs placebo difference	-	(11.9%)	-	(8.4%)	-	(10.8%)	-	(25.9%)	-	(25.9%)	-	(25.1%)
P-value	-	0.0038	-	0.0604	-	0.0006	-	-	-	-	-	-

[1] ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for baseline SELENA SLEDAI score (≤ 9 vs. ≥ 10), baseline proteinuria level (< 2 g/24 hour vs. ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs. other). For pooled data analysis, study was also included as an additional covariate

[2] Defined as SRI response with serology components (increased DNA binding and low complement items) removed

[3] Information extracted from HGS Briefing Document to FDA Figure 9.51

Table 14: Modified SRI response at week 52

	BLISS-52[3]		BLISS-76[3]		Pooled Total Population[2]		High Disease Activity Subgroup Pooled Total		High Disease Activity Subgroup BLISS-52		High Disease Activity Subgroup BLISS-76	
Change from Baseline at Week 52[2]	**Placebo N = 264**	**10 mg/kg N = 259**	**Placebo N = 255**	**10 mg/kg N = 245**	**Placebo N = 519**	**10 mg/kg N = 504**	**Placebo N = 203**	**10 mg/kg N = 193**	**Placebo N = 107**	**10 mg/kg N = 112**	**Placebo N = 96**	**10 mg/kg N = 81**
n(%) responders	127 (48.1%)	158 (61.0%)	92 (36.1%)	109 (44.5%)	219 (42.2%)	267 (53.0%)	42 (39.3%)	73 (65.2%)	29 (30.2%)	43 (53.1%)	71 (35.0%)	116 (60.1%)
OR (95% CI)[1]	-	-	-	-	-	-	-	3.0 (1.7, 5.2)	-	2.5 (1.4, 4.8)	-	2.8 (1.8, 4.2)
P-value[1]	-	-	-	-	-	-	-	0.0001	-	0.0036	-	<0.0001
10 mg/kg vs placebo difference	-	(11.9%)	-	(8.4%)	-	(10.8%)	-	(25.9%)	-	(25.9%)	-	(25.1%)
P-value	-	0.0038	-	0.0604	-	0.0006	-	-	-	-	-	-

[1] ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for baseline SELENA SLEDAI score (≤ 9 vs. ≥ 10), baseline proteinuria level (< 2 g/24 hour vs. ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs. other). For pooled data analysis, study was also included as an additional covariate

[2] Defined as SRI response with serology components (increased DNA binding and low complement items) removed

[3] Information extracted from HGS Briefing Document to FDA Figure 9.51

The MS did not specify patient numbers for this analysis and so data from the HGS Briefing Document to the FDA.[5] Figure 4 shows the percentage of modified SR responders (from HGS Briefing Document to FDA).

In the HGS/FDA[5] analysis there is little difference in response between 1mg/kg and 10mg/kg groups for BLISS-52.

The number of patients at risk was not specified. A stronger response was observed for the Target populations than for the total populations and statistical significance was reached in both trials.

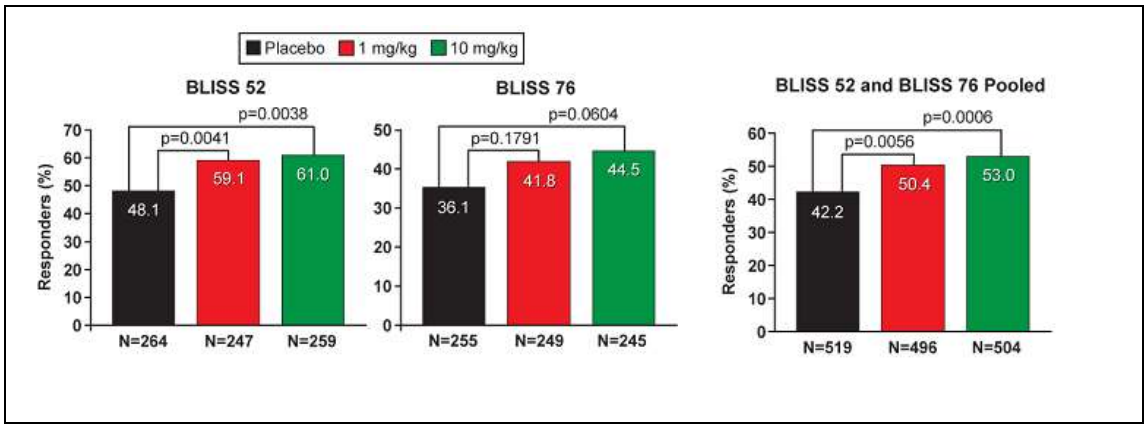


Figure 4: Modified SR percentage of responders (from HGS Briefing Document to FDA)

Subcomponents of the SRI response

Table 15 summarises the week 52 results for the three subcomponents of the composite SRI response (based on MS Table 5.16 and clarification Table A6.1).

Table 15: Results for subcomponents of SRI at week 52 (adjusted)

	BLISS-52		BLISS-76		Pooled Total Population[4]		High Disease Activity Subgroup Pooled Total		High Disease Activity Subgroup BLISS-52		High Disease Activity Subgroup BLISS-76	
	Placebo N = 287	10mg/kg N = 290	Placebo N = 275	10mg/kg N = 273	Placebo N = 562	10mg/kg N = 563	Placebo N = 203	10mg/kg N = 193	Placebo N = 107	10mg/kg N = 112	Placebo N = 96	10mg/kg N = 81
4-point reduction in SELENA-SLEDAI	132 (46.0%)	169 (58.3%)	98 (35.6%)	128 (46.9%)	230 (40.9%)	297 (52.8%)	84 (41.4%)	125 (64.8%)	47 (43.9%)	76 (67.9%)	37 (38.5%)	49 (60.5%)
Observed difference vs placebo (%)	-	12.3	-	11.3	-	11.9	-	23.4	-	24.0		22.0
OR (95% CI)[1] vs placebo	-	1.71 (1.21,2.41)	-	1.63 (1.15,2.32)	-	1.68 (1.3,2.2)	-	2.6 (1.7,3.9)	-	2.8 (1.6,4.8)	-	2.4 (1.3,4.4)
P-value[1]		0.0024		0.0062	-	< 0.0001	-	< 0.0001		0.0004	-	0.0063
No New 1A/2B BILAG domain scores	210 (73.2%)	236 (81.4%)	179 (65.1%)	189 (69.2%)	389 (69.2%)	425 (75.5%)	125 (61.6%)	145 (75.1%)	68 (63.6%)	88 (78.6%)	57 (59.4%)	57 (70.4%)
Observed difference vs placebo (%)	-	8.2	-	4.1	-	6.3	-	13.6	-	15.0	-	11.0
OR (95% CI)[1,2] vs placebo	-	1.62 (1.09,2.42)	-	1.20 (0.84,1.73)	-	1.4 (1.1,1.8)	-	1.9 (1.2,3.0)	-	2.3 (1.2,4.2)	-	1.6 (0.9, 3.1)
P-value[1,2]		0.0181		0.3193	-	0.0190	-	0.0034	-	0.0099	-	0.1297
No worsening in PGA	199 (69.3%)	231 (79.7%)	173 (62.9%)	189 (69.2%)	372 (66.2%)	420 (74.6%)	119 (58.6%)	142 (73.6%)	64 (59.8%)	86 (76.8%)	55 (57.3%)	56 (69.1%)
Observed difference vs placebo (%)	-	10.4	-	6.3	-	8.4	-	15.0	-	17.0	-	11.8
OR (95% CI)[1,3] vs placebo	-	1.74 (1.18,2.55)	-	1.32 (0.92,1.90)	-	1.5 (1.2,2.0)	-	2.0 (1.3,3.1)	-	2.3 (1.3,4.2)	-	1.6 (0.9,3.0)
P-value[1,3]	-	0.0048	-	0.1258	-	0.0017	-	0.0015	-	0.0063	-	0.1312

[1] Odds Ratio (95% confidence interval) and p-value were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, covariates include baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate

[2] Additional covariate: baseline BILAG domain involvement (at least 1A/2B)

[3] Additional covariate: baseline PGA score

[4] No significant treatment-by-study interactions were observed (all p > 0.287)

The three subcomponents of the composite SRI outcome were: [i] an improved SELENA SLE DAI score by ≥ 4 points; [ii] a BILAG index showing no new grade A organ involvement or no two grade B organ involvements (i.e. no worsening by one new A or two new B BILAG indices); [iii] a PGA score that has not increased by more than 0.3 points (i.e. no worsening in PGA by ≥ 0.3).

The percentage of patients at week 52 that achieved a SLEDAI score reduction of ≥ 4 points was defined as a major secondary outcome. For the whole population, both trials delivered more responders in the belimumab group than the placebo group (P = 0.0024 and P = 0.0062 for BLISS-52 and BLISS-76, respectively).

Results at week 52 for the other two SRI subcomponents (i.e. no worsening in BILAG index and no worsening in PGA score) were defined as non-major secondary outcomes. The percentage of patients in the whole population that satisfied the BILAG and PGA criteria in BLISS-52 was greater for belimumab relative than placebo (significant at P = 0.0181 and P = 0.0048 for BILAG and PGA, respectively); however, for BLISS-76 the differences between belimumab and placebo were considerably smaller and neither component reached statistical significance in favour of belimumab (P = 0.319 and P = 0.1258 for BILAG and PGA, respectively). According to results reported in the HGS Briefing Document to the FDA (Table 9.20, Page 102) the 1mg/kg belimumab dose regimen in BLISS-76 performed slightly better than 10mg/kg for both the PGA and BILAG subcomponents at week 52.

The corresponding results for the target population supplied during the clarification process are also summarised in Table 15 Pooled across trials, all three SRI components at week 52 were supportive of belimumab relative to placebo and delivered significant effects. However, for BLISS-76 the PGA and BILAG results at week 52 for the target population were considerably weaker (P = 0.1312 and P = 0.1297, respectively) than for BLISS-52 or the pooled target population.

Major secondary outcomes

The MS identified five pre-specified major secondary outcomes. These included the SRI response at week 76 and the percentage of patients with a ≥ 4 point SLEDAI improvement at week 52, each of which have been discussed in the preceding sections. The other three major secondary outcomes were: mean change in PGA score at week 24, percentage of patients with prednisone reductions ≥ 25% from baseline to ≤ 7.5 mg/day during weeks 40 to 52 (in subjects whose baseline dose was > 7.5 mg/day); mean change in SF36 PCS at week 24. These are discussed in this section.

Change in PGA score at week 24 was presented in MS Table 5.18 and the relevant results from this are shown in Table 16 below. For the whole population in BLISS-52 the change in PGA score (week 24 relative to baseline) for both groups indicated disease improvement and was greater in the belimumab group (-0.54) than placebo group (-0.39; P = 0.0003 in support of belimumab). For BLISS-76 the difference between groups was very small and in favour of placebo (-0.49 placebo and -0.48 belimumab) and did not reach statistical significance (P = 0.7987). For the Target HDAP pooled across trials belimumab delivered a greater reduction in PGA score than placebo (P = 0.028 with mean changes of -0.42 and -0.52 for placebo and belimumab, respectively). Target population results by trial are not available.

Table 16: Mean change in PGA score at week 24 (taken from MS Table 5.18)

Major secondary endpoint at Week 24	BLISS-52		BLISS-76		Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 287	10mg/kg N = 290	Placebo N = 275	10mg/kg N = 273	Placebo N = 562	10mg/kg N = 563	Placebo N = 203	10mg/kg N = 193
Mean ± SE	-0.39 ± 0.03	-0.54 ± 0.03	-0.49 ± 0.04	-0.44 ± 0.03	-0.44 ± 0.02	-0.49 ± 0.02	-0.42 ± 0.04	-0.52 ± 0.04
LS Mean ± SE[1]	-0.35 ± 0.04	-0.50 ± 0.04	-0.49 ± 0.05	-0.48 ± 0.05	-0.40 ± 0.03	-0.48 ± 0.03	-0.41 ± 0.05	-0.53 ± 0.05
P-value[1]	-	0.0003	-	0.7987	-	0.0167	-	0.0268

[1] All statistics, including the difference in LSM (least square means), were from ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for the baseline PGA score, baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.

The mean change in PGA at week 52 was submitted as an additional secondary outcome. The results are shown in Table 17.

Table 17: Mean change in PGA score at week 52 (taken from MS Table 5.18)

Other secondary endpoints Week 52	BLISS-52		BLISS-76		Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 287	10mg/kg N = 290	Placebo N = 275	10mg/kg N = 273	Placebo N = 562	10mg/kg N = 563	Placebo N = 203	10mg/kg N = 193
Mean ± SE	-0.48 ± 0.04	-0.67 ± 0.04	-0.46 ± 0.04	-0.49 ± 0.04	-0.47 ± 0.03	-0.58 ± 0.03	-0.41 ± 0.05	-0.62 ± 0.05
LS Mean ± SE[1]	-0.38 ± 0.05	-0.57 ± 0.05	-0.47 ± 0.06	-0.55 ± 0.06	-0.40 ± 0.04	-0.54 ± 0.04	-0.36 ± 0.06	-0.59 ± 0.06
P-value[1]	-	0.0001	-	0.1159	-	< 0.0001	-	0.0003

[1] All statistics, including the difference in LSM (least square means), were from ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for the baseline PGA score, baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.

In BLISS-52 a larger improvement in PGA score was observed for the 10mg/kg group than for placebo (P = 0.0001) whereas in BLISS-76 the difference between treatments was trivial (P = 0.115). For the pooled populations 10mg/kg was superior to placebo (P = 0.0003).

The HGS Briefing Document to the FDA[5] provided graphed results for mean change in PGA through successive clinic visits for all three randomised groups. These are shown below in Figure 5 for BLISS-76.

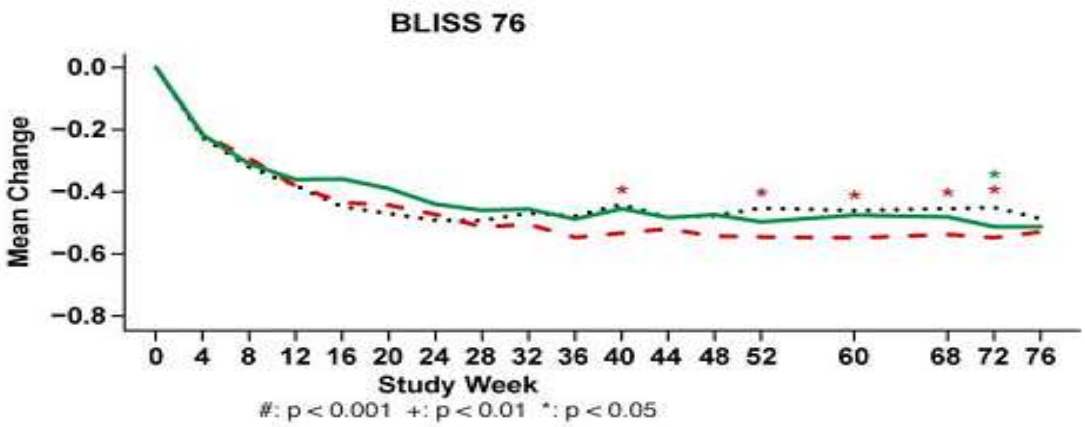


Figure 5: Mean change in PGA score in BLISS-76

Across 76 weeks of follow up in BLISS-76 the 1mg/kg dose regimen appeared to outperform the 10mg/kg regimen. Baseline differences (MS Table 5.9) were similar between treatment groups.

The mean change in SF-36 PCS scores at week 24 relative to baseline, a major secondary outcome, showed little difference between belimumab and placebo groups in BLISS-52 (P = 0.8870), or in BLISS-76 (P = 0.6601), or in the Target population pooled across trials (P = 0.4276).

Change in SF-36 PCS scores at week 52 was specified as a non-major secondary outcome. No significant improvement was observed for BLISS-76 or Target populations (P = 0.5134 and P = 0.1124, respectively) however in BLISS-52 the difference between belimumab and placebo arms (4.18 vs. 2.96) was sufficient to reach statistical significance (P = 0.0247).

Reduction in steroid use between weeks 40 and 52 for those patients receiving ≥ 7.5 mg/day prednisone at baseline was specified as a major secondary outcome. The results submitted summarised in

Table 18.

Table 18: Prednisone reduction Weeks 40 through 52 – Phase 3 trials

	BLISS-52		BLISS-76		Pooled Total Population[4]		High Disease Activity Subgroup Pooled Total		High Disease Activity Subgroup BLISS-52		High Disease Activity Subgroup BLISS-76	
	Placebo N = 192	10 mg/kg N = 204	Placebo N = 126	10 mg/kg N = 120	Placebo N = 318	10 mg/kg N = 324	Placebo N = 126	10 mg/kg N = 126	Placebo N = 76	10 mg/kg N = 81	Placebo N = 50	10 mg/kg N = 45
No. %[1] Response[2]	23 12.0%	38 18.6%	16 12.7%	20 16.7%	39 12.3%	58 17.9%	9 7.1%	20 15.9%	4 5.3%	15 18.5%	5 10.0%	5 11.1%
Observed difference vs Placebo	-	6.65	-	3.97	-	5.64	-	8.73	-	13.5	-	1.1
OR (95% CI)[3] vs placebo	-	1.75 (0.99, 3.08)	-	1.26 (0.61, 2.60)	-	1.57 (1.01, 2.45)	-	2.43 (1.05, 5.65)	-	4.11 (1.29, 13.2)	-	0.88 (0.21, 3.60)
P-value[3]	-	0.0526	-	0.5323	-	0.0451	-	0.0389	-	0.0171	-	0.8586

[1] Includes only subjects with baseline prednisone > 7.5 mg/day

[2] Any subject who withdrew from the study prior to the Day364 (Week 52) visit, missed the Day 364 (Week 52) visit (± 28 day window allowed), and/or received a protocol-prohibited medication or a dose of allowable (but protocol-restricted) medication that resulted in treatment failure designation prior to the Day 364 (Week 52) visit was considered a treatment failure for prednisone reduction

[3] Odds Ratio (95% confidence interval) and p-value were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, the covariates include baseline prednisone level, baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate

[4] Obtained from a logistic regression by adding study and the treatment-by-study interaction to the above model

In BLISS-52 and BLISS-76 at baseline 68.6% and 44.9% of patients respectively were receiving ≥ 7.5 mg/day prednisone. The percentage that reduced steroid use in weeks 40 to 52 by the pre-specified amount was greater in the belimumab arm than the placebo arm in both trials, however the difference (belimumab vs. placebo) failed to reach statistical significance in either trial: 18.6% vs. 12.0% in BLISS-52 (P = 0.0526 from logistic regression including baseline covariates) and 16.7% vs. 12.7% in BLISS-76 (P = 0.5323).

For the Target or HDAP pooled across trials 15.9% and 7.1% reduced steroid use in the 10mg/kg belimumab and placebo groups respectively (P = 0.0389 from logistic regression). The results from BLISS-52 supported belimumab (P = 0.171) whereas in BLISS-76 differences between treatments were trivial (P = 0.8586). The HGS Briefing Document to the FDA[5] provided results for reduction in steroid use for all three treatment arms. Table 9-16 from the HGS Briefing Document is shown in Figure 6.

Table 9-16 Reduction in steroid use - Phase 3 trials[1]

	BLISS 52			BLISS 76			Both Studies		
	Placebo N = 192	1 mg/kg N = 204	10 mg/kg N = 204	Placebo N = 126	1 mg/kg N = 130	10 mg/kg N = 120	Placebo N = 318	1 mg/kg N = 334	10 mg/kg N = 324
No. (%) Response[2]	23 (12.0%)	42 (20.6%)	38 (18.6%)	16 (12.7%)	25 (19.2%)	20 (16.7%)	39 (12.3%)	67 (20.1%)	58 (17.9%)
OR (95% CI) vs placebo		1.89 (1.08, 3.31)	1.75 (0.99, 3.08)		1.57 (0.78, 3.14)	1.26 (0.61, 2.60)		1.77 (1.15, 2.73)	1.57 (1.01, 2.45)
P-value[3]		0.0252	0.0526		0.2034	0.5323		0.0097	0.0451
Treatment by study interaction p-value[4]		NA	NA		NA	NA		0.7020	0.5177

[1] Average prednisone (equivalent) dose reduced by ≥ 25% from baseline to ≤ 7.5 mg/day during Weeks 40-52 in the subgroup of patients who were receiving > 7.5 mg/day of prednisone at baseline

[2] Any patient who withdrew from the study prior to the Day 364 (Week 52) visit, missed the Day 364 (Week 52) visit (± 28 day window allowed), and/or received a protocol-prohibited medication or a dose of allowable (but protocol-restricted) medication that resulted in treatment failure designation prior to the Day 364 (Week 52) visit was considered a treatment failure for prednisone reduction.

[3] Odds ratio (95% confidence interval) and p-values were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, the covariates include baseline prednisone level, baseline SELENA SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate. P-values nominal.

[4] Obtained from a logistic regression by adding study and the treatment-by-study interaction to the above model.

Figure 6: Reduction in steroid use Phase II trials (Taken from HGS Table 9-16)

It is noticeable that again there was little difference in effectiveness between the 1mg/kg and 10mg/kg dose regimens, in BLISS-76 a better outcome was recorded with 1mg/g than with 10mg/kg, and that the results from BLISS-52 were more strongly supportive of belimumab than those from BLISS-76.

Further secondary outcomes submitted

Flares

Time to first flare and to first severe flare was reported in MS Figures 5.9 to 5.13.

In BLISS-52 the time to first flare was delayed by 10 mg/kg belimumab relative to placebo (HR 0.76, 95% CI: 0.63 – 0.91, P = 0.0036) (Figure 7).

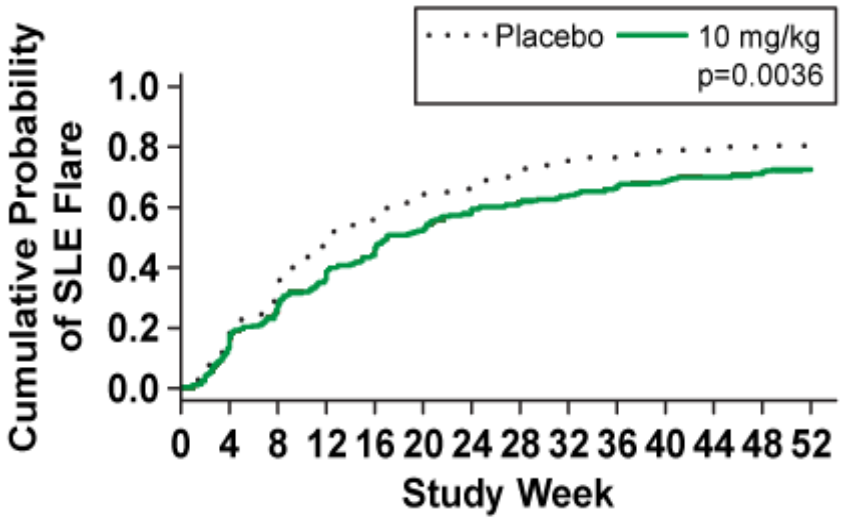


Figure 7: Time to first flare; BLISS-52 (Taken from MS Figure 5.10)

In BLISS-76 there was no difference between groups in time to first flare (P = 0.4796; Figure 8).

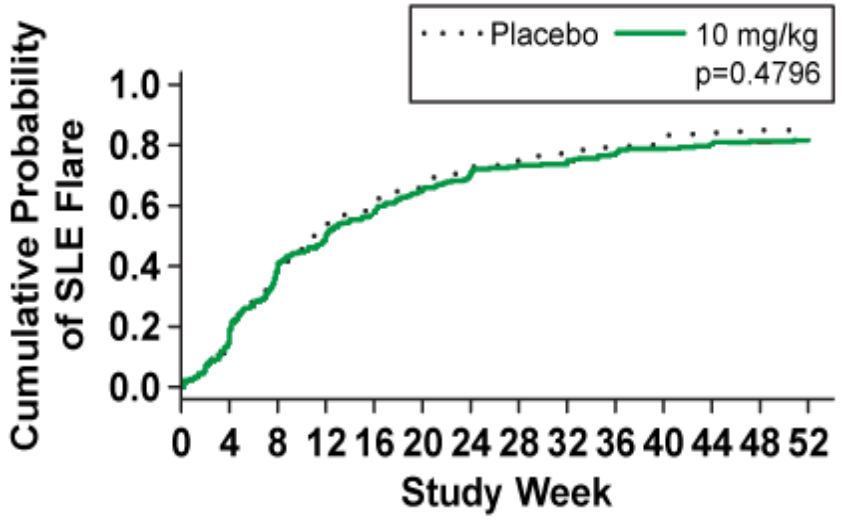


Figure 8: Time to first flare; BLISS-76 (Taken from MS Figure 5.11)

When the whole populations from the BLISS trials were pooled the difference between treatments reached statistical significance (P = 0.0120; Figure 9).

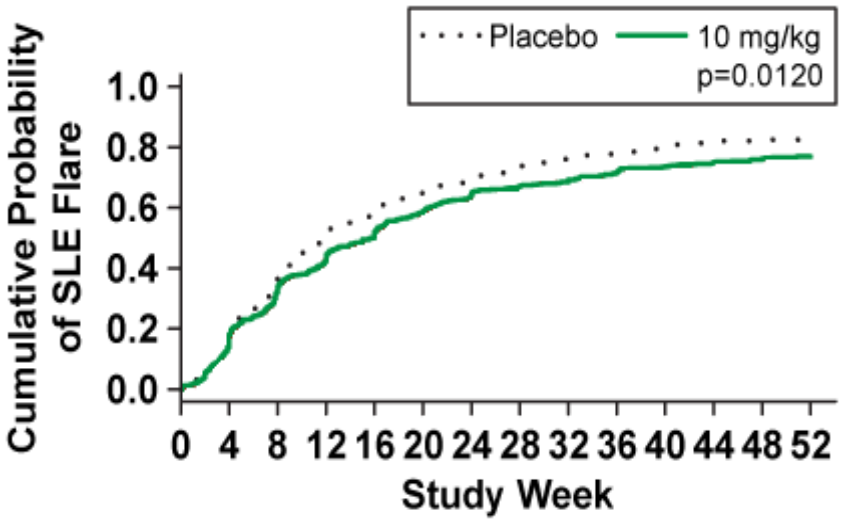


Figure 9: Time to first flare; pooled whole populations (Taken from MS Figure 5.12)

For the high disease activity Target population pooled across trials, belimumab significantly delayed time to first flare relative to placebo (P = 0.007; Figure 10).

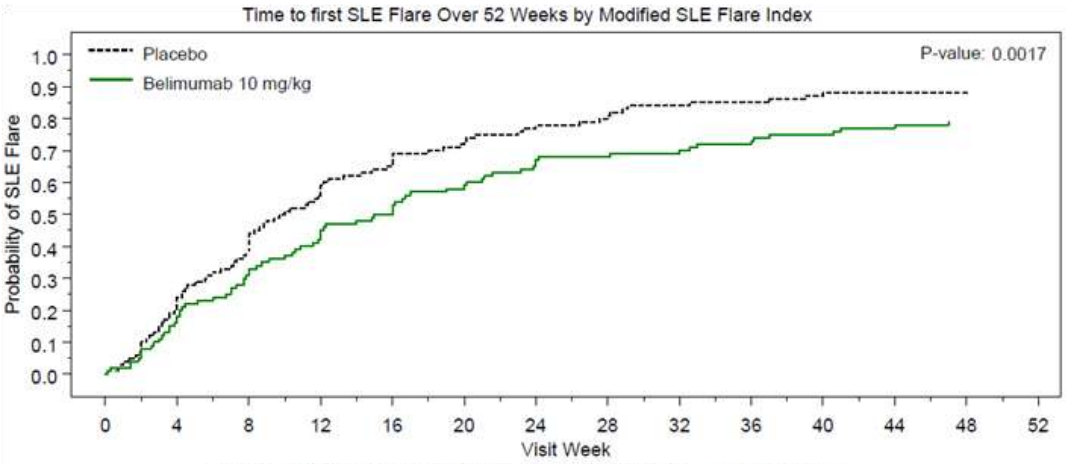


Figure 10: Time to first flare; pooled Target populations (Taken from MS Figure 5.13)

In BLISS-52 the time to first severe flare was delayed by 10 mg/kg belimumab relative to placebo P = 0.0055; Figure 11).

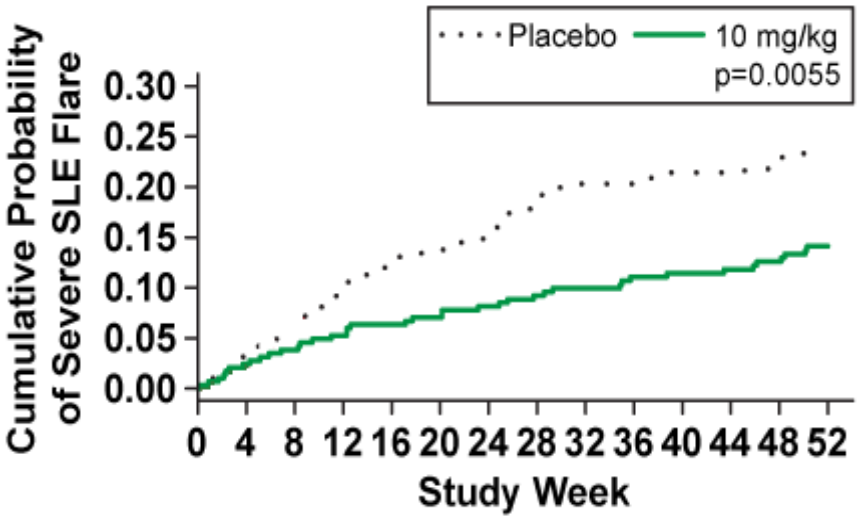


Figure 11: Time to first severe flare; BLISS-52 (Taken from MS Figure 5.10)

In BLISS-76 belimumab somewhat delayed time to first severe flare in (HR 0.72, 95% CI 0.50–1.05, P = 0.0867; Figure 12).

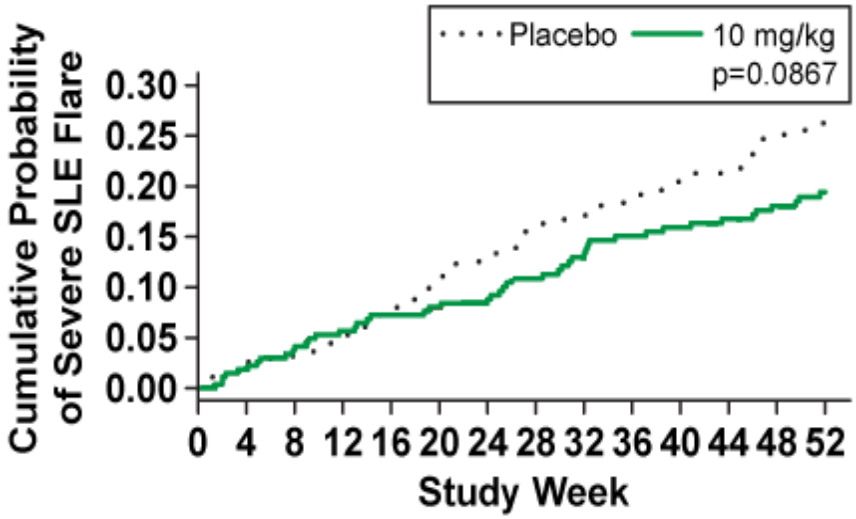


Figure 12: Time to first severe flare; BLISS-76 (Taken from MS Figure 5.11)

When the whole populations from the BLISS trials were pooled the difference between treatments for time to first severe reached statistical significance (P = 0.0011; Figure 13).

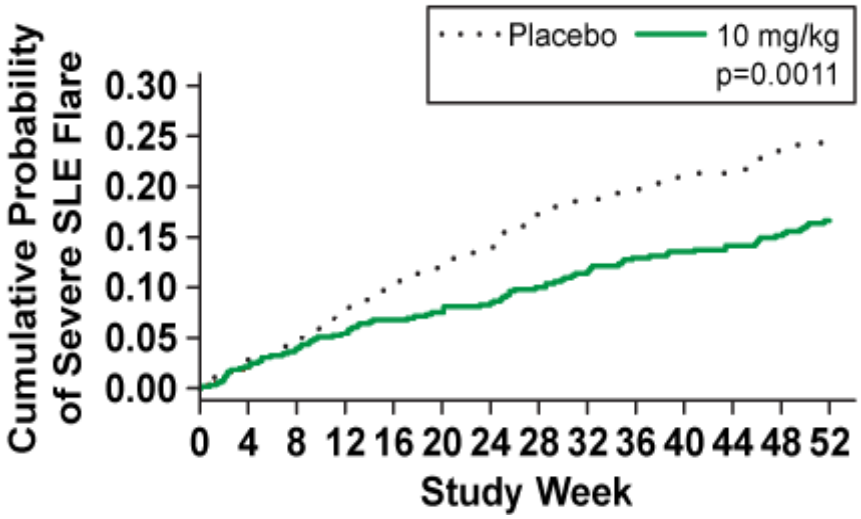


Figure 13: Time to first severe flare; pooled whole population (Taken from MS Figure 5.12)

For the high disease activity Target population pooled across trials, belimumab significantly delayed time to first severe flare relative to placebo (P = 0.0028; Figure 14).

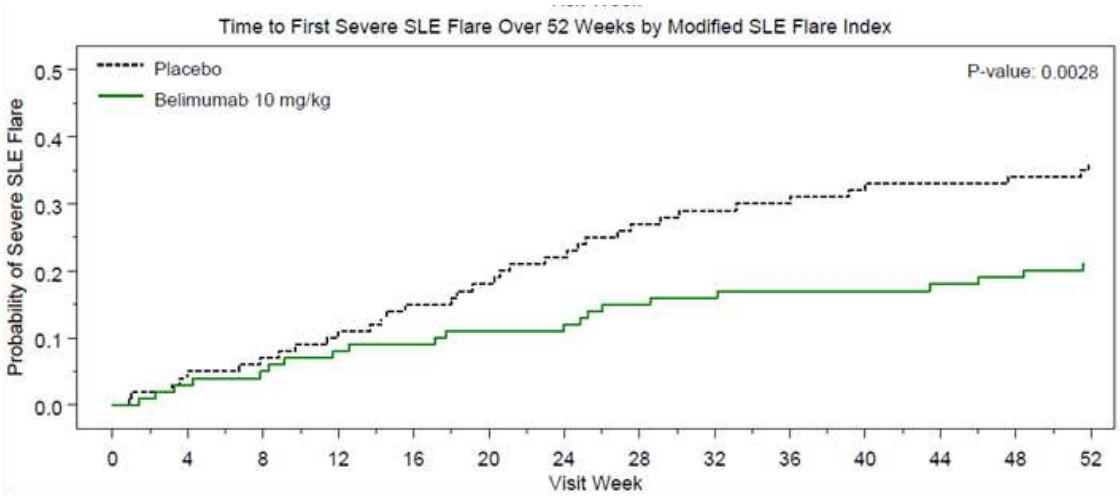


Figure 14: Time to first severe flare; pooled Target population (Taken from MS Figure 5.13)

The HGS Briefing Document to the FDA[5] provided the graphs shown in Figure 15 depicting results for all three treatment arms.

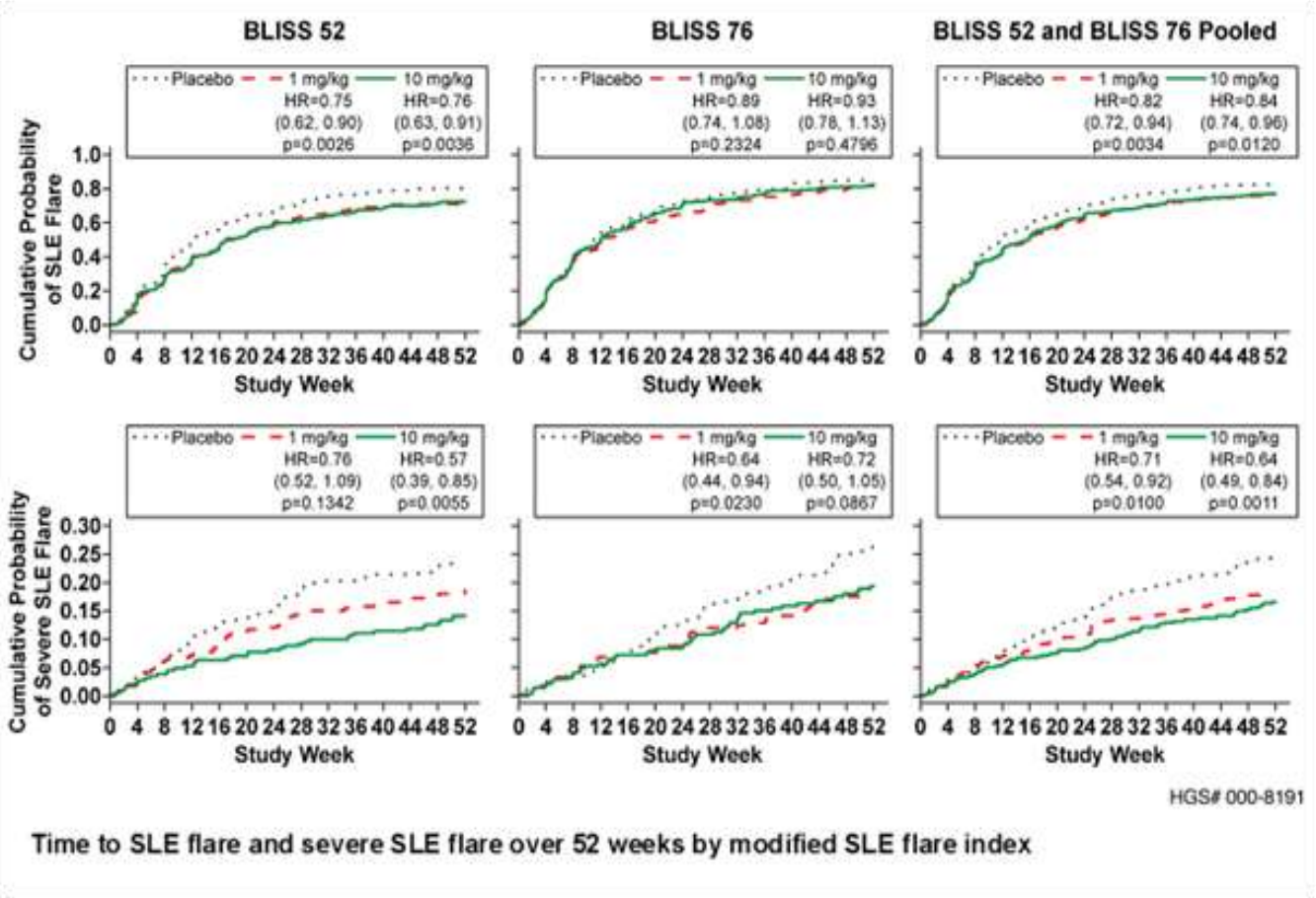


Figure 15: Time to first flare (taken from HGS Briefing Document to the FDA)

It is noticeable that in both trials the 1 mg/kg belimumab dose regimen was as effective as the 10 mg/kg dose regimen in extending time to first flare, and that for BLISS-76 this also applies for severe flares. For both flares and severe flares the results from BLISS-52 were more supportive of belimumab than those from BLISS-76.

SLICC/ACR Damage Index

There was no meaningful difference between the belimumab and placebo groups in the change in SLICC/ACR Damage Index at Week 52 compared with baseline.

FACIT-fatigue index

The mean change FACIT fatigue score from baseline was reported in MS Figures 5.14 to 5.17 shown in Figure 16.

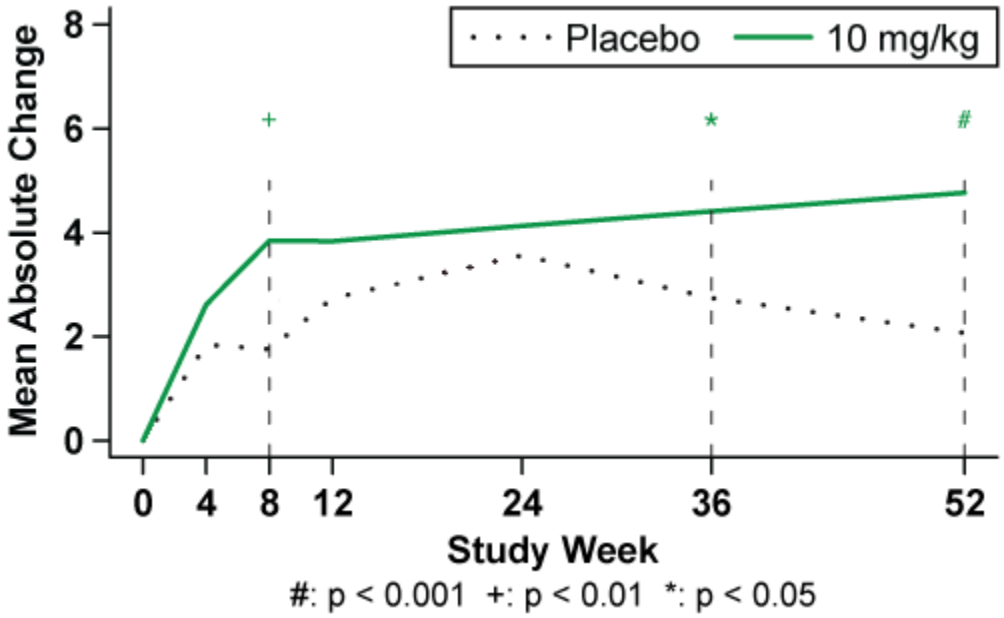


Figure 16: Mean change in FACIT-Fatigue score – BLISS-52 (Taken from MS Figure 5.14)

At week 52 relative to baseline the belimumab group had greater improvement in FACIT-Fatigue score than the placebo group (4.8 belimumab and 2.1 placebo in BLISS 52; 4.6 and 2.9 in BLISS-76). The difference was significant for BLISS-52 ($P < 0.001$) but not for BLISS-76 ($P \geq 0.05$) (Figure 17 and Figure 18).

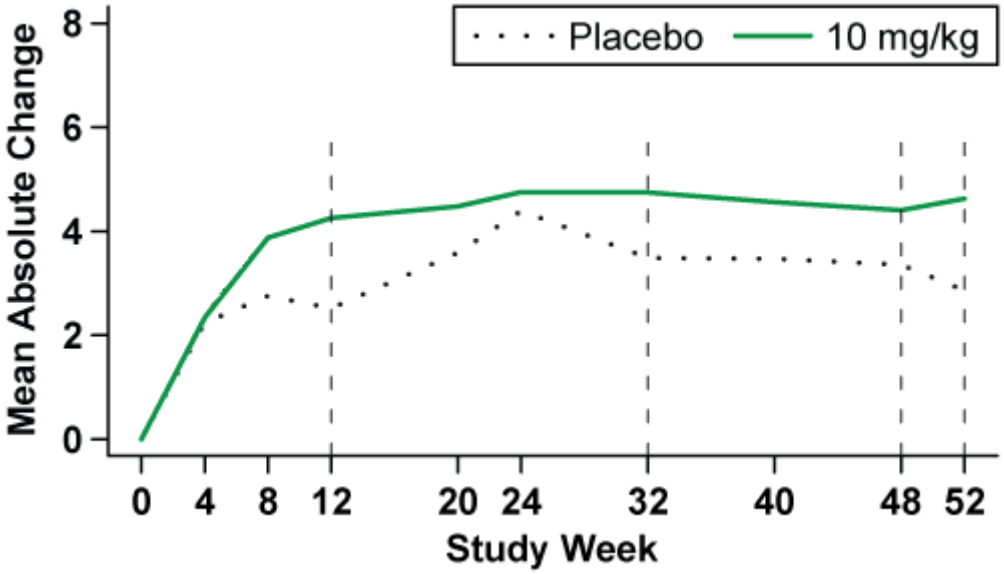


Figure 17: Mean change in FACIT-Fatigue score – BLISS-76 (Taken from MS Figure 5.15)

For the whole population pooled across trials the difference was statistically in favour of belimumab at week 52 (Figure 18).

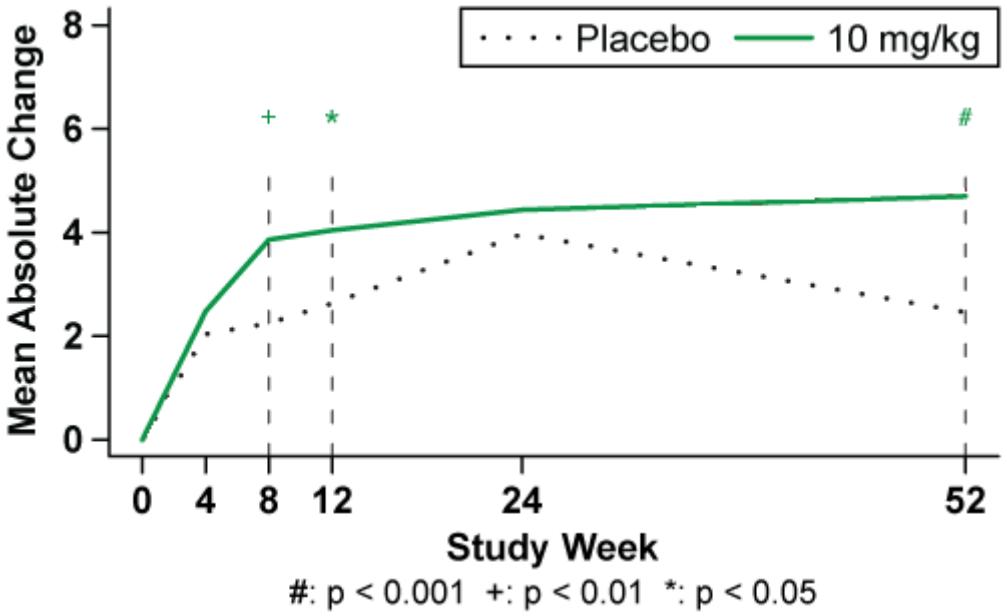


Figure 18: Mean change in FACIT-Fatigue – Pooled Total Population (Taken from MS Figure 5.16)

While for the target population pooled across trials at weeks 8 and 12 the difference between groups was statistically in support of belimumab ($P < 0.05$) however the difference between groups then diminished; at week 52 there was no longer a significant difference (see Figure 19).

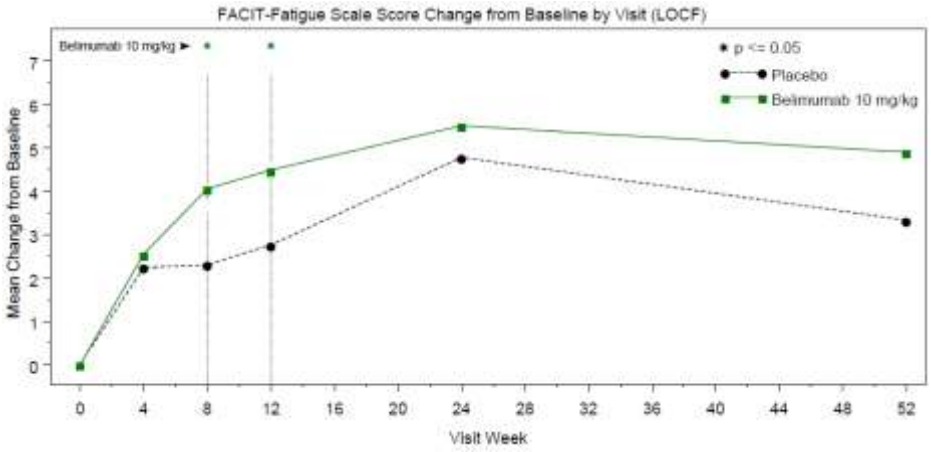


Figure 19: Mean change in FACIT-Fatigue – pooled Target population (Taken from MS Figure 5.17)

The HGS Briefing Document to the FDA provided results for all three treatment arms at week 52. In BLISS-52 the 10mg/kg dose was more effective than the 1 mg/kg but for BLISS-76

the reverse was the case. The BLISS-76 result is shown in Figure 20 together the mean change in SF-36 Vitality domain score.

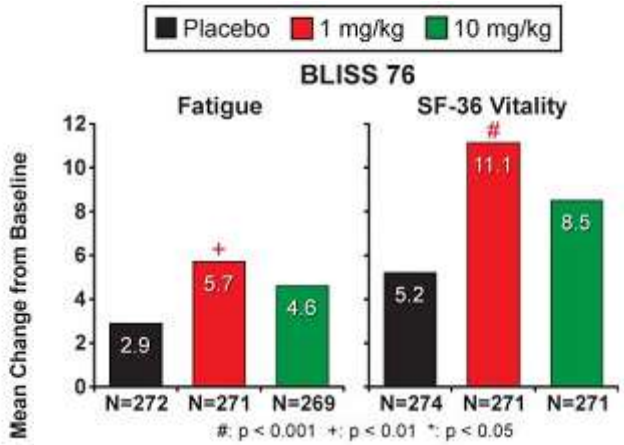


Figure 20: Mean change in FACIT and SF-36 vitality score by week 52 (Taken from HGS Briefing Document to FDA see Figure 9-35)

EQ-5D

There was no significant difference between belimumab and placebo in the absolute change of EQ-5D score from baseline in either trial or pooled total populations during clinic visits. The results for the 10 mg/kg belimumab and placebo groups in BLISS-76 were indistinguishable. For the pooled target population the difference between 10 mg/kg and placebo groups reached statistical significance in favour of belimumab at week 24 ($P \leq 0.01$), but the difference had almost completely faded by week 52 MS Figure 5.21 Page 135).

Results for the mean change in SELENA SLEDAI score from baseline at week 52 were submitted in MS Table 5.17 (Page 113) and clarification response Table A6.1 and are summarised in Table 19. There was no significant difference between belimumab and placebo in the absolute change of EQ-5D score from baseline in either trial or pooled total populations during clinic visits. The results for the 10mg/kg belimumab and placebo groups in BLISS-76 were indistinguishable. For the pooled Target population the difference between 10mg/kg and placebo groups reached statistical significance in favour of belimumab at week 24 ($P \leq 0.01$), but the difference had almost completely faded by week 52 MS Figure 5.21 Page 135).

Results for the mean change in SELENA SLEDAI score from baseline at week 52 were submitted in MS Table 5.17 (Page 113) and clarification response Table A6.1 and are summarised in Table 19.

Table 19: Mean change and mean percent change in SLEDAI score week 52

	BLISS-52		BLISS-76		Pooled Total Population[2]		High Disease Activity Subgroup Pooled Total		High Disease Activity Subgroup BLISS-52		High Disease Activity Subgroup BLISS-76	
Change from Baseline at Week 52	**Placebo N = 287**	**10mg/kg N = 290**	**Placebo N = 275**	**10mg/kg N = 273**	**Placebo N = 562**	**10mg/kg N = 563**	**Placebo N = 203**	**10mg/kg N = 193**	**Placebo N = 107**	**10 mg/kg N = 112**	**Placebo N = 96**	**10 mg/kg N = 81**
Mean change from baseline (± SE)	-3.57 ± 0.24	4.97 ± 0.27	-2.77 ± 0.25	-3.70 ± 0.27	-3.18 ± 0.18	-4.36 ± 0.19	-4.1 ± 0.3	-5.8 ± 0.3	-4.1 ± 0.4	-6.3 ± 0.5	-4.0 ± 0.5	-5.2 ± 0.5
P-value[1]	-	< 0.0001	-	0.0063	-	< 0.0001	-	0.0005	-	0.0008	-	0.1705
Mean % change (± SE)	-34.76 ± 2.50	-45.60 ± 2.45	-25.97 ± 2.72	-35.94 ± 2.80	-30.47 ± 1.85	-40.93 ± 1.86	-30.5 (2.3)	-45.5 (2.4)	-	-	-	-
P-value[1]	-	0.0018	-	0.0073	-	< 0.0001	-	< 0.0001	-	-	-	-

[1] ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for baseline SELENA SLEDAI score (≤ 9 vs. ≥ 10), baseline proteinuria level (< 2 g/24 hour vs. ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs. other). For pooled data analysis, study was also included as an additional covariate

[2] No treatment-by-study interactions observed (all p-values > 0.367)

Both absolute SLEDAI score reduction from baseline, and percent reduction relative to baseline score, were greater for the 10mg/kg group than for the placebo group; this was consistent and significant for the whole BLISS population (separately by trial and for pooled populations) and for the pooled Target or high disease activity population. For the whole population, results favoured belimumab more strongly in BLISS-52 than BLISS-76. The by-trial results for the Target population are shown below. They indicate stronger support for belimumab in BLISS-52 in which the difference between groups in absolute reduction in SLEDAI score was about double that in BLISS-76 in which the difference between groups was not significant (P = 0.1705).

The HGS Briefing Document to the FDA[5] (see Figure 21) showed the percentage change in SLEDAI score (relative to baseline) throughout the two trials; this is reproduced in Figure 21 for the mean change in FACIT and SF-36 vitality score by week 52.

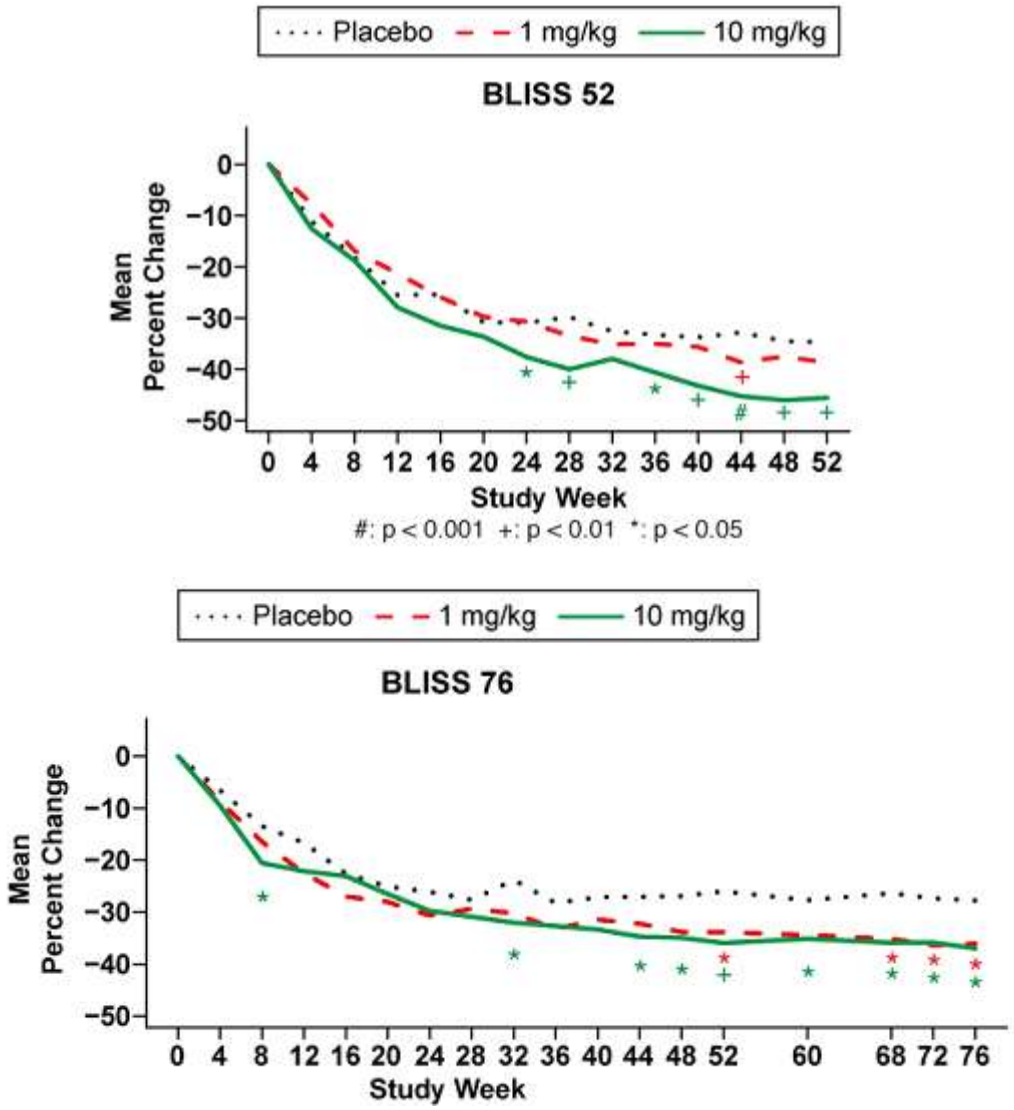


Figure 21: Percentage change in SLEDAI score according to treatment arm (Taken from HGS Briefing Document Figure 9-29)

These results support a dose response relationship in BLISS-52, but the difference between 1mg/kg and 10mg/kg dose regimens in the BLISS-76 trial is relatively trivial.

Safety

The submission pooled results from three RCTs: BLISS-52, BLISS-76 and LBSL02. LBSL02 lasted 52 weeks, preceded the BLISS trials, and was conducted in North America (98% patients from the USA), and did not employ the SRI composite outcome measure. The LBSL02 trial randomised 449 patients to one of four treatments: SoC + placebo, SoC + 1mg/kg belimumab, SoC + 4mg/kg belimumab, and SoC + 10mg/kg belimumab. Although

all patients had a history of auto-immunity, at recruitment 30% currently lacked anti-nuclear antibodies.

There were 15 deaths during the controlled phase of the three trials; 3 in the placebo group (n=675), and 12 in the belimumab groups (n=1458) with 6 each in the 10mg/kg and 1mg/kg groups respectively. One death in the 1mg/kg belimumab group followed 15 weeks after the patient stopped belimumab treatment. The causes of death were various and are listed in Table 20 (based on FDA Briefing Package, Table 34).[3] When deaths are rated according to exposure these results translate to: 0.43/100 patient years for placebo (95% CI: 0.09, 1.27) and 0.79/100 patient years for belimumab (95% CI: 0.41, 1.38).[3] There were two completed suicides in the belimumab groups (none in placebo); a further suicide was observed during the LBSL99 extension study. These were not judged to be associated with belimumab since the patients concerned had a history of depression and SLE is associated with an increased risk of depression and suicide.

Table 20: Deaths occurring during controlled phase of belimumab RCTs

Study group	Age/Sex	Cause	Days from 1'st infusion	Days from last infusion	Pertinent History
Placebo	45yo/F	Myocardial Infarction	328	19	Presented to ER with new onset chest and epigastric pain and had a cardiopulmonary arrest.
Placebo	25yo/F	Cardiac Arrest, secondary to sepsis	70	11	Concomitant Meds: Prednisolone, methotrexate, diclofenac and ibuprofen. Developed bacterial gastroenteritis and dehydration complicated by vasculitis and became septic (blood culture positive for Staph. Saprophyticus) despite antibiotics and supportive medical care.
Placebo	18yo/F	Unknown	225	84	Hospitalized 2 months prior to death for acute abdominal pain secondary to portal/mesenteric/renal vein and vena cava thrombosis and acute pancreatitis.
Belimumab 1mg/kg	43yo/F	Suicide	32	20	H/O Depression on antidepressant (citalopram). Reported to have worsening depression prior to committing suicide.
Belimumab 1mg/kg	46yo/F	Unknown	56	28	H/O Asthma, clostridial gastroenteritis, eosinophilia and QT prolongation on EKG. Concomitant Meds: ibuprofen, hydroxychloroquine, mycophenolate, prednisone and lisinopril. Pt. developed nausea, vomiting and weakness while camping and was found to be dehydrated due to unspecified gastrointestinal illness at local ER where she died despite resuscitative measures.
Belimumab 1mg/kg	52yo/F	Ovarian cancer	21	7	Positive family H/O ovarian cancer. H/O Vaginal bleeding prior to study entry that evolved to include left lower abdominal pain, vaginal pain, pelvic cramping and diarrhea by the 9th study medication that was followed by a diagnosis of advanced ovarian cancer on laporotomy.
Belimumab 1mg/kg	32yo/F	Sepsis, secondary to cellulitis	13	13	Concomitant Meds: Methylprednisolone, mycophenolate, thalidomide, and ibuprofen. Developed cellulitis and died as a result of sepsis despite antibiotics and supportive medical care.
Belimumab 1mg/kg	58yo/F	Ischemic stroke	345	34	H/O hypertension. Anti-cardiolipin antibody negative at screening. Concomitant meds: Prednisolone, hydroxychloroquine, bioprolol.
Belimumab 1mg/kg	25yo/F	Respiratory failure /SLE flare	216	104	Patient died due to respiratory arrest more than 15 weeks after the patient discontinued the trial due to acute renal failure. Post study withdrawal, the patient was hospitalized and experienced oliguria, uremic syndrome, sepsis, polyserositis, ascites, intestinal edema, anemia, and alveolar hemorrhage.
Belimumab 10mg/kg	40yo/F	Respiratory failure secondary to sepsis	257	33	Pt. developed aspiration pneumonia status post seizure, became septic and died due to respiratory failure despite antibiotics and aggressive supportive medical care (respirator).
Belimumab 10mg/kg	47yo/F	Cardiac arrest (SLE flare)	77	21	H/O Diabetes mellitus, pericardial excision, serositis, antiphospholipid syndrome, pulmonary hypertension, and heart failure. Concomitant Meds: Azathioprine, methotrexate and prednisone. Hospitalized after c/o severe headache with vomiting associated with fever, chills and productive cough with bilateral pleural effusions and lymphopenia attributed to SLE flare with CNS involvement. She was treated with corticosteroids and NSAIDs but died due to cardiac arrest.
Belimumab 10mg/kg	53yo/F	Bacterial sepsis	331	25	H/O Obesity, pulmonary fibrosis. Developed septic shock (blood cultures positive for MRSA) and multi-organ failure secondary to infected herpes zoster lesions despite antibiotics. Concomitant meds: Methyplprednisone, azathioprine, chloroquine, salbutamol, acenocoumarol, sertraline, and omeprazole.
Belimumab 10mg/kg	20yo/F	Infectious diarrhea	336	28	Had SLE flare with cutaneous vasculitis and hypochromic anemia. Started on antibiotics and increased corticosteroids but developed infectious diarrhea and died en route to hospital. Concomitant meds: Prednisolone, azathioprine, hydroxychloroquine, levofloxacin, iron, ciprofloxin/tinidazole, and fluconazole.
Belimumab 10mg/kg	23yo/F	Suicide	272	13	H/O Depressed mood and psychotic disorder; autoimmune thyroiditis, and drug-induced hepatitis. Committed suicide following conflict with parent. Concomitant meds: methylprednisone, azathioprine, hydroxychloroquine, meloxicam, levothyroxine, and rebamipide.
Belimumab 10mg/kg	33yo/F	Respiratory Failure From Presumed Pulmonary Embolus	128	8	H/O chronic cholecystitis. Pt. developed dyspnea eight days after her last study infusion and died en route to the hospital. (No autopsy.) Concomitant meds: Prednisone, levothyroxine, and ceftriaxone.

Adverse events

In all treatment groups > 90% of patients experienced at least 1 AE. The most commonly occurring AEs were headache, upper respiratory tract infection, arthralgia, nausea, UTI, diarrhoea and fatigue.

The percentage of patients experiencing at least one serious AE and at least one serious AE was very similar between placebo and belimumab groups ranging 13.5% to 18.6%, there was a very slight numerical excess with belimumab. The most frequent serious AEs (≥ 1% in any treatment group) were pneumonia, pyrexia, UTI, cholelithiasis, and cellulitis. The percentage of patients experiencing at least one severe AE was 15.4% for the placebo group and 16% across the belimumab groups; the most common severe adverse events were not identified.

Infections

Infections occurred slightly more frequently in patients treated with belimumab compared to placebo. The most frequent infections were URTI, UTI, nasopharyngitis, sinusitis, and bronchitis.

Infusion / hypersensitivity reactions

Occurrence of infusion plus hypersensitivity reactions was similar between belimumab and placebo-treated patients (17% and 14.7%, respectively). Of 1458 belimumab treated patients, 15 experienced hypersensitivity reactions on the day of infusion compared to one of 675 placebo-treated patients.[3] Five discontinuations resulted from hypersensitivity reactions amongst 1458 belimumab patients and none among 675 patients receiving placebo.

The most frequent infections were URTI, UTI, nasopharyngitis, sinusitis, and bronchitis. Of these, nasopharyngitis and bronchitis occurred more commonly with belimumab treatment compared to placebo. Two opportunistic infections occurred, both in the belimumab 10mg/kg group: disseminated CMV infection on day 62; and an Acinetobacter bacteremia on day 15. Four infections were related to deaths: sepsis (placebo group); infectious diarrhea (belimumab 10mg/kg group); cutaneous infection leading to sepsis (belimumab 10mg/kg group); and cellulitis leading to sepsis (belimumab 1mg/kg group).[3]

3.2.6 Pooling of trial data

NICE requests that: "*For each outcome for each included RCT, the following information should be provided*" …."*The size of the effect*" … "*The unit of measurement*".

The submission pooled results from two trials, both for the whole BLISS populations and for the Target populations. The manufacturer considered that the pooled trial results were most

appropriate for the decision problem, and importantly it was pooled results for both populations that were entered into the economic model. The MS stated as shown in Box 11.

Box 11: Taken from Page 21 and 97 of the MS

"For the purpose of this submission pooled efficacy from the two pivotal Phase III studies is considered most relevant to the decision problem." Refer to MS Page 21

AND

"Pooling is appropriate given that the trials were essentially identical in design and in the analysis of the primary endpoint, the p-values for the treatment-by-study interaction were not significant (interaction p-values > 0.5)." Refer to MS Page 97

Comment

The trials were conducted according to very similar protocols and used the same primary end point so the lack of significant treatment-by-study interaction was not surprising; see section 4.2.4 of this report.

The submission initially supplied only pooled results for the Target population, therefore the ERG requested "by-trial" results and further justification for pooling. The response to this request is shown in Box 12 (for the full response see Appendix 5).

Box 12: From the manufacturer's clarification response

"….one must then determine whether the relative treatment effect is different in one study compared with the other study when evaluating whether two studies are similar enough to pool. Each of the Phase 3 studies achieved statistical significance for belimumab 10mg/kg on the pre-specified primary endpoint of SRI response at Week 52; therefore, these nearly identical, studies provide independent replication of results."

Comment

There remain doubts as to whether the pooled trial results are relevant for patients in England and Wales. The BLISS trials were run globally, recruited 1684 patients in 226 centres across 32 countries and involved a large number of different investigators. Although both trials achieved statistical significance on the primary end point, they were dissimilar in underlying patient groups e.g. by ethnicity and in effect size for almost all outcomes, with BLISS-52

providing stronger results than BLISS-76; furthermore there were inconsistencies with regard to an expected dose response relationship.

The ERG is concerned that the pooled results are mainly driven by those from the BLISS-52 trial (conducted in Pacific-Asia and South America) while results in the BLISS-76 trial, conducted in North America and Europe, were only marginally in favour of belimumab relative to placebo and reached statistical significance only for two overlapping outcomes (SRI responders at week 52, and percentage of patients with ≥ 4point reduction in SLEDAI score at week 52 which itself is a component of the SRI). The extent to which these concerns extend to the target population was not possible to gauge from the initial submission because only pooled results were presented. The ERG therefore requested clarification on trial specific target populations and the manufacturer's justification for pooling across trials.

For the target population there was again a greater contribution from BLISS-52 to the pooled results both in terms of number of patients (BLISS-52 contributed 55% of whole target population and 58% of those that received belimumab) and in effectiveness (BLISS-52 provided greater effect sizes compared to BLISS-76 for SRI week 52, modified SRI week 52, percentage with SLEDAI reduction by ≥ 4points, SLEDAI mean change by week 52, no new BILAG 1A/2B, no worsening in PGA and reduction in steroid usage weeks 40 to 52). The ERG therefore remain concerned that the pooled trial results dilute the rather less positive findings most relevant to the decision problem by including additional data from a less relevant population, and that target population results by trial should have been included in sensitivity analysis in the economic model.

Trial baseline SLEDAI scores used in economic model

The manufacturer's economic analysis (section 5) made use of data from an SLE cohort studied at JHU so as to model cost effectiveness of belimumab for the whole and Target populations from the BLISS trials. The JHU cohort experienced relatively mild disease compared to patients in BLISS and particularly compared to the BLISS the high disease activity Target population. During the clarification process the ERG requested the distribution of SLEDAI scores at year one and last year of observation for patients in the JHU cohort. The SLEDAI scores shown in Figure 22: SLEDAI scores for Target and JHU populations (from clarification document)illustrates the differences between Target and JHU populations, (year one and last year scores are shown for JHU cohort, Figure B17.2 of the clarification document, the clarification response did not make clear which was year one and which last year).

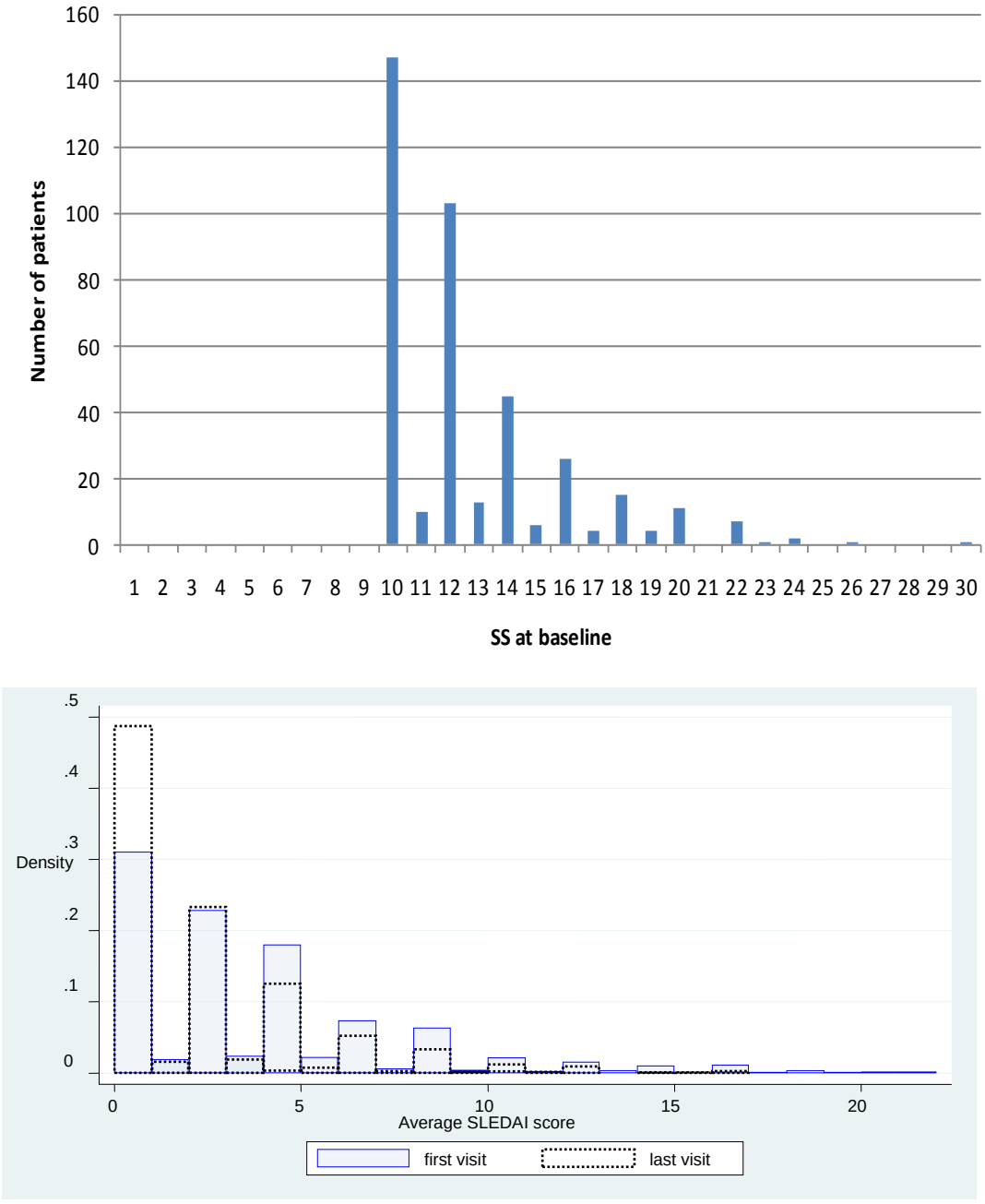


Figure 22: SLEDAI scores for Target and JHU populations (from clarification document)

The median follow up for the JHU cohort was 6.9 years. During this time organ damage progression was monitored and for economic analysis this was modelled (using parametric time event analyses) so as to be able to predict organ damage progression for BLISS patients according to their observed SLEDAI changes. Yet for the JHU cohort the difference between year one and last year in average SLEDAI scores is small. The ERG considers that this indicates some inadequacy in using the short term measure of disease activity (i.e. SLEDAI) to model how a group of patients will progress to organ damage.

3.2.7 Conclusions

Efficacy evidence came from two multicentre international industry sponsored RCTs (BLISS-52 and BLISS-76) comparing SoC plus belimumab with SoC plus placebo; each trial had three arms: placebo, 1mg/kg belimumab and 10mg/kg belimumab dose regimens. Data for the 1mg/kg arms was excluded from the submission, but results available in the public domain were considered in the ERG's assessment. Outcomes for six populations were presented: whole populations from BLISS trials, whole populations pooled across BLISS trials, Target population from BLISS trials and Target populations pooled across BLISS trials.

The Target population was a high disease activity subgroup identified from post hoc exploration of effectiveness of the primary outcome. There were more noticeable within-trial baseline imbalances (10mg/kg vs. placebo) for the Target population than for the whole population. The Target population results are not necessarily equivalent to those that would be obtained from a randomised trial in this population.

The primary outcome was specified as the percentage of responders at week 52 according to a novel composite disease activity measure (SRI). This outcome was statistically in favour of belimumab (10mg/kg vs. placebo) in both trials. For both whole and Target populations, results from BLISS-52 were more favourable for belimumab than results from BLISS-52.

For all secondary outcomes in BLISS-76 effect sizes were insufficient to be confident that effects could not be accounted for by chance. For several outcomes, including percentage responders and time to first flare in BLISS-76, although formal statistical tests were not performed, it appeared that the 1mg/kg dose regimen was as effective, or more effective, than the 10mg/kg dosage.

Geographical and racial differences between BLISS trials indicate that efficacy results from BLISS-76, rather than from BLISS-52 or pooled BLISS populations, are more generalisable to the UK.

On most safety outcomes placebo and belimumab performed equally. There were more deaths under belimumab than placebo; on a "per patient year of exposure" basis the rate for belimumab was about double that for placebo although this finding could have occurred by chance. Causes of death were various and most were those associated with the condition of SLE. There was insufficient evidence to determine if there was any association between belimumab and mortality.

4 ECONOMIC EVALUATION

4.1 Introduction

Including a one page summary of structure, assumptions and sources, with signposting to Tables.

Patient population and subgroups under consideration

The submission outlines that there are three groups under consideration:

- The patient population as observed from pooling the All BLISS patient data;
- The anticipated license patient population of Anti-dsDNA+ve and low (C3 or C4);
- The Target population which restricts the patient population to the licensed patient population with an SS score at baseline of at least 10.

With the exception of Table 6.49 of MS, the analysis presented within the main body of the submission relates to the All BLISS patient population. Little detail is presented for the anticipated license population, though the base case results for this group are presented within Table 6.49 of the MS. The inputs and results for the Target population are presented in section 6.9 of the MS.

Given the anticipated license as stated within the submission, the ERG review of the economics does not focus upon the All BLISS inputs, though the base case results for this group are presented. The brief summary of the base case results for the anticipated license patient population is also presented. But unless otherwise stated the inputs to and results of the modelling within the ERG review of the economics relate to the Target population.

Implementation of the electronic model

The manufacturer model is embedded within Excel, but apart from some very basic pre-modelling data adjustment the Excel element of this is confined to being a database of input values and a store of the model results. The modelling is implemented using Visual Basic (VB) programming.

Prior to running the model the user is allowed to change various pre-specified settings within the model, such as the subgroup to be analysed. The model uploads the relevant set of input parameters into memory, calculates the model using the VB code and outputs the results to an Excel worksheet.

The VB programming is well organised and compact with no obviously superfluous code. But it is relatively complex and lengthy with little to no explanatory comment, running to 53 Pages when printed out in Arial 8pt. This has made it difficult for the ERG to confidently explore structural scenarios other than those pre-specified within the model within the STA time constraints.

Stopping rule and clinical effectiveness estimates

Note that the economics applies a stopping rule within the belimumab arm at week 24: those not experiencing a change in their SELENA-SLEDAI (SS) score of at least 4 by week 24 are assumed to stop belimumab treatment. Conceptually, this gives rise to two groups within each arm:

- Belimumab week 24 responders;
- Belimumab week 24 non-responders;
- SoC week 24 responders;
- SoC week 24 non-responders.

When reviewing the economics of the submission, it should be borne in mind that the actual experience of the belimumab week 24 non-responders at week 52 as reported within the trials is not used within the model for these patients.

4.2 Manufacturer's submission

4.2.1 Economic literature search

The searches undertaken by the manufacturer to identify cost-effectiveness evidence were conducted on 18th March 2011. Seven databases were searched (Medline (Pubmed), Medline In-Process (Pubmed), Embase, EconLit, CRD databases (HTA database, Database of Abstracts of Reviews of Effectiveness (DARE), NHS Economic Evaluation Database (EED). In addition, searches were conducted in Research Papers in Economics (RePEc), a clinical trials database (ClinicalTrials), the websites of the American College of Rheumatology, the USA FDA and EMA.

The search utilised terms to identify patient group (lupus) and the intervention (belimumab). Terms to identify comparators were not included. In the Pubmed search, the restriction to title and abstract in the belimumab section of the search strategy has resulted in 7 fewer hits compared to the same line in the clinical effectiveness search strategy. In line 1 of the Pubmed search, lupus would automatically have been mapped by Pubmed, resulting in the inclusion of the wrong MeSH heading Lupus Vulgaris. The MeSH heading Lupus

Erythematosus, Systemic was not included. However, because Pubmed also searched for lupus in all fields, papers with this MeSH heading should still have been picked up.

In the Embase search several lines include major mistakes resulting in it being unclear as to how the database would have performed the search. For example, line 1 starts with "exp AND" and includes two database index terms that do not exist: 'lupus'/exp, 'sle'/exp. Testing this line of the search by entering it exactly as reported results in 23221 (06/05/11), which is far fewer than the 58059 reported in the search strategy. Many of the lines in Embase that should have been searching the EMTREE headings were entered very differently and resulted in far fewer hits (e.g. the EMTREE heading in the filter at line 11: exp Economic Evaluation/ (which when tested brought back 166263 hits on 06/05/11) was entered as "exp AND economic AND ('evaluation'/exp OR evaluation)" and resulted in only 977 hits, which may have led to important studies being missed.

No date or language limits were applied. A search filter was applied to the Embase and Medline searches to limit them to a particular type of evidence (economic studies), which was not appropriate in Pubmed in light of the small number of studies retrieved before the filter was applied (41). The manufacturer states that the search filter used was the CRD sensitive economics filter. However, the version the manufacturer uses in Pubmed does not match that given in CRD's NHS Economics Evaluation Database Handbook 2007[16] and there are also some mistakes in the translation of some elements (e.g. the MeSH heading in the filter exp "Costs and Cost analysis"/ was entered as: costs AND "cost analysis". Fortunately, when this was translated by Pubmed the correct MeSH heading was searched, but several unusual combinations of free text terms were also searched (e.g. costs"[All Fields] AND "cost"[All Fields] AND "analysis"[All Fields]).

Overall, the Pubmed and Embase searches appear to be of a poor quality and may have resulted in important studies being missed (see Appendix 2 for further details).

4.2.2 Manufacturer's direct drug cost and administration

Belimumab dose and direct drug cost

To calculate the direct drug cost for belimumab the manufacturer has assumed whole vial use that minimises wastage. Belimumab is available in 400mg and 120mg vials, at unit costs of £381 and £114 respectively excluding the PAS, these yielding the same cost per mg. [redacted] Since 120mg is not a factor of 400mg, the

minimum cost dose can in theory use anything up to nine 120mg vials. Cost minimisation suggests that where appropriate multiple vials of 120mg should be used. For instance, six 120mg vials provides a total available dose of 720mg whereas one 400mg vial and three 120mg vials provide a total available dose of 760mg: a 72kg patient is most cheaply dosed with six 120mg vials while a 73kg patient is most cheaply dosed with one 400mg vial and three 120mg vials.

The manufacturer applies the above dosing calculations to the distribution of patient weights of the pooled trial data to arrive at a weighted average belimumab drug cost. These are differentiated by patient subgroups to yield average drug costs per administration of £671 for the belimumab arm pooled for All BLISS, £650 when this is restricted to the anti-dsDNA+ve, low (C3 or C4) group and £654 when this is restricted to the Target population. [redacted]

Belimumab administration cost

Belimumab administration is assumed to require 2 hours of dedicated nursing time which is costed using 2010 Patient Social Service Research Unit (PSSRU) rates for a senior hospital staff nurse at £126 per administration. There is no specific allowance for any consumables within this administration costing

4.2.3 Model Structure

The model is implemented as a patient level simulation due to the complexity of SLE and the large number of health states that this implies. This inevitably makes it and its electronic implementation relatively complex, but an outline of the model and the data sources is reasonably simple to present. Within this summary it is simplest to separate the model elements by the source of the data feeding into them:

1. Trial data:
    a. The baseline characteristics for each patient being simulated including SS score at baseline and whether there is involvement for each of the 12 SLICC organs modelled
    b. The likelihood of response at week 24 in the belimumab arm, defined as a change of at least 4 in the SS score from baseline
    c. The change in SS score between baseline and week 52 for belimumab week 24 responders

d. The change in SS score between baseline and week 52 for SoC, this also being applied to belimumab week 24 non-responders
e. The "natural" discontinuation rates for belimumab week 24 responders after week 24
f. The direct effect of the SS score upon quality of life
g. The direct effect of the SS score upon treatment cost

2. JHU cohort data:
    a. The evolution of the SS score in the SoC arm after week 52, with belimumab week 24 responders being assumed to retain the absolute advantage in SS score over the SoC arm while they remain on treatment
    b. Implicit in the above the Adjusted Mean SLEDAI (AMS): the adjusted mean SS score since baseline
    c. The relationship between the SS score and steroid use
    d. The main survival model
    e. The likelihood of developing involvement for each of the 12 SLICC organs modelled if the organ concerned is not involved at baseline
3. Other data drawn from the broader literature:
    a. The standardised mortality rate for a given SS score
    b. The quality of life impact of organ involvement
    c. The additional cost for each organ involvement

The baseline characteristics, likelihood of response, week 52 SS scores and discontinuation rates are differentiated by subgroup within the model (1.a. - 1.e.). All other relationships are not. A number of Tables from the manufacturer's submission are replicated within what follows for ease of reference.

Trial data element 1.a.Baseline patient characteristics

For each patient level simulation the patient characteristics are randomly sampled from the underlying distribution; e.g. for the Target population the likelihood of the patient being female is based upon the 94.2% of the pooled trial data and a drawing on a Bernoulli distribution (Tables 6.37, 6.38 and 6.39 of the MS for the Target population). The patient is then cloned within the model for running through the SoC arm of the model and the belimumab arm of the model.

Trial data element 1.b.Likelihood of response in the belimumab arm

This is differentiated by the patient baseline SS score and drawn on a Bernoulli distribution (Table 6.42 of the submission for the Target population).

Trial data element 1.c and 1.d.Change in SS score from baseline to week 52

Unlike the likelihood of response, this is not drawn from lookup Tables based upon SS score at baseline and treatment arm. The manufacturer pools the trial data and uses regression analysis to derive coefficients for the percentage reduction in a patient's baseline SS score dependent upon whether the patient was in the SoC arm, the belimumab arm and if in the belimumab arm whether they were a week 24 responder (See Table 21; Adapted from Table 6.41 of MS for the Target population).

Table 21: Linear regression of coefficients for $SS_{52}=(1+\beta)SS_0$: Target population

	β	s.e.	P value
SoC	-34.9%	2.2%	< 0.01
Belimumab	-34.3%	4.6%	< 0.01
Belimumab week 24 responders	-28.0%	5.2%	< 0.01

As the model assumes that week 24 non-responders within the belimumab arm cease treatment and experience the SoC SS scores at week 52 the central estimates of this are more transparently rearranged as shown in Table 22.

Table 22: Rearranged linear regression of coefficients for $SS_{52}=(1+\beta)SS_0$: Target population

	β	1+β
SoC	-34.9%	65.1%
Belimumab week 24 non-responders	-34.9%	65.1%
Belimumab week 24 responders	-62.3%	37.7%

For a belimumab week 24 responder, the absolute difference at week 52 in SS scores between the belimumab week 24 responder and her SoC arm clone is assumed to be maintained while she remains on belimumab treatment; i.e. 27.4% of her baseline SS score up to discontinuation.

Trial data element 1.e. natural discontinuation rates for belimumab week 24 responders

The written submission is not entirely explicit as to the modelled natural discontinuation rates but it appears that for belimumab week 24 responders a daily hazard of discontinuation is calculated based upon the overall rate of discontinuations between day 168 and day 532. A six monthly natural discontinuation rate is calculated for the first year; i.e. presumably

subsequent to the assessment of response at week 24, with an annual rate being calculated thereafter. The day 168 to day 532 proportion remaining on treatment among belimumab week 24 responders is given as 0.891 in the electronic model for the All BLISS data set, and as 0.920 for the Target population.

A key aspect of this data is that it must relate to discontinuations between day 168 and day 532 and not to discontinuations between baseline and day 532, as the latter would cause the model to overestimate discontinuation rates among belimumab week 24 responders (Table 6.41 of the MS for the Target population). There is some lack of detail within the submission around this variable: its source, the period it relates to, any pooling of data between BLISS-52 and BLISS-76 given their differing duration after day 164; any evidence of difference between BLISS-52 and BLISS-76 for this variable; and, the reasons for the observed discontinuations. The latter may be particularly important given that this variable is used to extrapolate over the time horizon of the model with the cost effectiveness estimate being quite sensitive to it.

Within the model the impact of these continuation/discontinuation rates are graphed below, as submitted by the manufacturer in response to ERG clarification question B24. The upper curve is the Target population. Note that Figure 23 shows the effect of both discontinuations and mortality on those remaining on belimumab treatment.

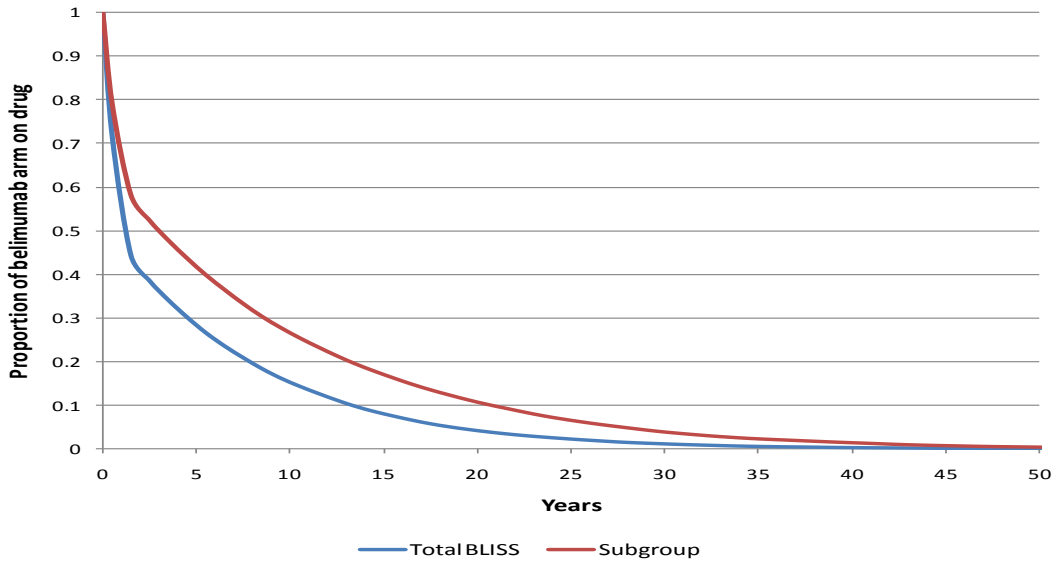


Figure 23: Continuation rates among belimumab week 24 responders

As explored later, a high discontinuation rate improves the estimated cost effectiveness of belimumab; i.e. it is more cost effective if belimumab has an initial effect and high response

rate but that patients experiencing a response rapidly discontinue belimumab treatment thereafter.

Trial data element 1.f. SS score direct impact upon QoL

Within the BLISS trials EQ-5D data was collected and transformed to HRQoL values using the Dolan algorithm[17]. Regression analysis then related these values to patients' SS scores, age, sex, ethnicity and organ damage. The statistically significant organ damage parameters were retained in order to control for their impact within the regression analysis, but organ damage was then set to be zero to generate the "clean" utility function for a patient with no organ damage:

"Clean" QoL = 1.297 – 0.145 * ln(Age) – 0.054 * ethnicity – 0.009 * SS score

where ethnicity is 1 if black and 0 if not. Note that the above corresponds with the clean utility function as in Table 6.20 of the submission and the electronic model, not with the function given in the text of the submission which appears to be incorrect.

Trial data element 1.g. SS score direct impact upon costs:

Limited detail is provided within the submission on the resource use questionnaire administered during the LBSL02 phase II trial. This was apparently a retrospective data collection administered at baseline, day 168 and day 365. 2006 PSSRU and NHS reference costs[18] were applied to this data, and the aggregate six monthly costs from baseline to day 168 and day 168 to day 365 were regressed on patients' SS score severity class during this period. The SS score severity class took a value of 0 to 3, this being determined by a patient's maximum observed SS score during the relevant 6 month period: 0 for a maximum SS score of 0; 1 for a maximum SS score of between 1 and 4; 2 for a maximum SS score of between 5 and 12; and, 3 for a maximum SS score of over 12. This regression analysis based on SS score severity class was then mapped back onto SS scores as outlined in Figure 6.16 of the submission, with the CPI being used to inflate the figures to 2010 prices and the six monthly costs being doubled to yield an annual cost. This yields the final direct cost function of Table 6.25 of the MS (Table 23). Within this, it should be noted that while both constant and derived coefficient were estimated as being significant, the regression had an R^2 of only 0.01.

Table 23: Manufacturer estimated SS direct annual cost function table

SS score	Cost	SS score	Cost	SS score	Cost
0	£1,152	5	£1,625	10	£1,931
1	£1,286	6	£1,681	11	£2,005

2	£1,419	7	£1,736	12	£2,079
3	£1,514	8	£1,792	13+	£2,153
4	£1,569	9	£1,857		

JHU cohort data elements 2.a.and 2.b. Evolution of SS and AMS scores after week 52

For the evolution of the SS score beyond week 52 the manufacturer originally estimated the regression equation:

But the manufacturer views this as providing a relatively poor fit to the data from the belimumab phase II trial, as shown in Figure 24. For the modelling the manufacturer adjusts the constant and applies the equation:

on the grounds of it better reflecting the phase II trial data.

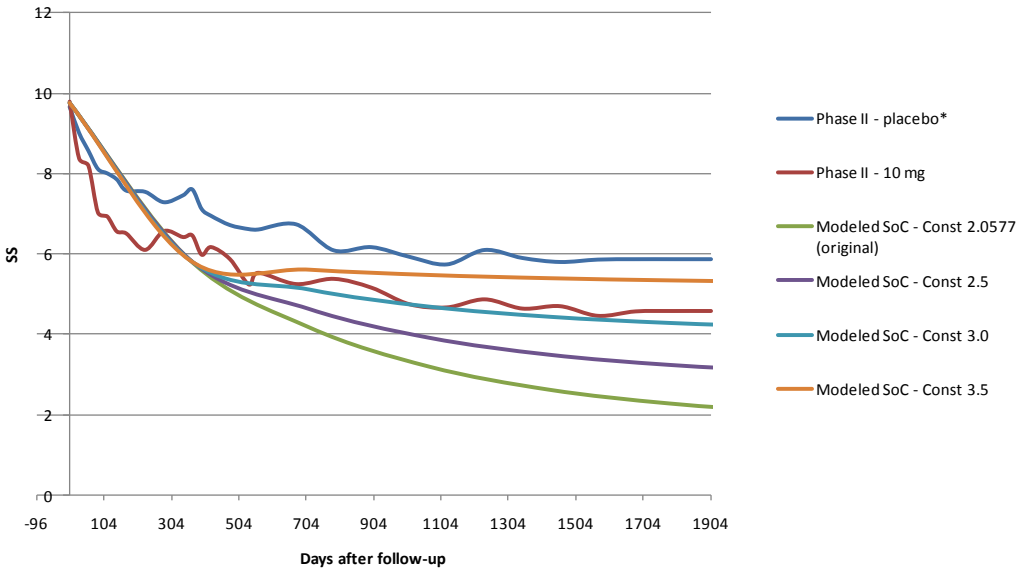


Figure 24: Medium term SS natural history model

The AMS score was developed to measure disease severity over time as opposed to the SS score which only reflects disease activity over the preceding 10 days. AMS is calculated as the area under the curve of disease activity measurements between two time-points. The area under the curve is then divided by time of follow-up to provide an average score over the period of interest. In Figure 24: Medium term SS natural history modelabove the ERG assume that the area under the curves shown can be used to represent AMS. However, the MS references to AMS score may refer to either the "AMS over lifetime" or the "average mean SLEDAI up to current time" which are presumably calculated in the same manner: the area under the SS score curve divided by time elapsed.

The AMS is calculated as the area under the SS score curve between two time points divided by the time of follow-up to provide an average score over the period of interest.

The AMS score refers to either the "AMS over lifetime" or the "average mean SLEDAI up to current time" which are presumably the same[a]: the area under the SS score curve divided by time elapsed.

The SS curve determines the AMS, with it being the AMS that is used within the survival model and the organ damage natural history model.

JHU cohort data element 2.c. Steroid use:

[a] Note that within the written there seems to be some occasional looseness of wording around SS and AMS, with there being some instances of the AMS referring to an annual AMS score; i.e. the average SS score over one year rather than from baseline.

The manufacturer argues that the steroid doses and changes to these as seen in the BLISS trials are not representative of the likely steroid dose reductions that would be possible with belimumab. Given this, the manufacturer fits a random effects model to the JHU data which estimates the steroid dose as a function of the average SS score within the year being simulated.

Table 24: Steroid use as a function of SS score Table 6.11

	Coef (95% CI)	P-value
Constant	3.410 (0.617-0.823)	< 0.001
SS score within year average	0.720 (3.073-3.747)	< 0.001

Note that despite the arguments around the representativeness of the JHU cohort for the All BLISS population and modelling as outlined in the previous section, no alternative forms for the steroid use equation were explored. In response to ERG clarification question C2 the manufacturer justifies this on the basis of there being little difference in the baseline steroid doses between the JHU cohort, 9.95mg/day, and the pooled All BLISS 10.78mg/day.

JHU cohort data element 2.d. Survival model

Initial univariate survival analysis within an exponential regression model framework found a range of variables within the JHU cohort data to be predictors of mortality, including age and duration of disease which were each statistically significant. Through a process of multivariate stepwise covariate selection a sensitivity analysis around the proper distribution (Exponential and Weibull) the range of variables included in the survival model were reduced. Within this process age was not included due to concerns around it having a high positive correlation with disease duration. Age at diagnosis and disease duration was chosen instead for this selection process. Within the multivariate stepwise covariate selection disease duration was further eliminated, leaving only age at diagnosis [Table 6.12 and Table 5 of Appendix 21[b]].

Note that both the AMS and the Cumulative Average Prednisone Dose (CAPD) up to current time are explanatory variables within the JHU cohort survival model. As outlined above, these are determined by a patient's SS curve.

[b] Table 5 of appendix 21 outlines that the model adopted, model 5, has the second highest AIC of the four models this is reported for.

Since age and disease duration have been eliminated from the JHU cohort survival model, the manufacturer reintroduces age to the calculation of mortality risk using an SLE Standard Mortality Ratio (SMR) differentiated by age group drawn from the Bernatsky 2006[19] reference coupled with general population mortality rates as outlined below.

JHU cohort data element 2.e.Risk of developing organ involvement

Through a similar analysis to the survival model, the manufacturer estimates individual risk equations for the development of individual SLICC item organ involvement [Table 6.14 and for more detail Appendix 21 of the MS]. These individual models are described within the model as "JHU – AMS forced in, involvement removed".

The AMS is an explanatory variable within the models of the risk of: CVD, gastrointestinal, musculoskeletal, neuropsychiatric, ocular, peripheral vascular, pulmonary, renal and skin involvement. But is not an explanatory variable within the models of the risk of: diabetes, malignancy or gonadal failure involvement.

The CAPD is an explanatory variable within the models of the risk of: CVD, diabetes, gastrointestinal, musculoskeletal, ocular, gonadal failure and skin involvement. But is not an explanatory variable within the models of the risk of: malignancy, neuropsychiatric, peripheral vascular, pulmonary or renal involvement.

Any new involvement of an organ is assumed to occur at the average SLICC score for that organ observed as observed across the JHU cohort. As the manufacturer notes, this will tend to overestimate the SLICC score for that organ when involvement occurs, but this bias is likely to wane as time and the model progresses. The net overall impact of the assumption of a constant SLICC score at the average of that observed in the JHU cohort for newly incident organ involvement is consequently ambiguous.

Literature element 3.a.SLE SMR by age group

Due to age not being within the JHU cohort derived survival model, the manufacturer uses a set of age dependent SMRs for SLE patients relative to the general population as derived from the Bernatsky 2006[19] reference: 19.2 age 16-24, 8.0 age 25-39, 3.7 age 40-59 and 1.4 age 60+.

To calculate the likelihood of a patient dying during a cycle the model first derives the probability of death for this patient from the JHU survival model. The probability of death from the JHU survival model for a patient at the average value of the covariates observed

within the JHU cohort is then calculated. Dividing the first by the second yields the patient's hazard of death compared to the "average" JHU cohort patient.

This hazard is then multiplied by the age dependent SLE SMR as drawn from the Bernatsky 2006[19] reference and the age dependent general population risk of mortality, with the derived mortality rate then being adjusted back to being a probability [see MS Table 6.13 and untitled Table immediately after for a worked example].

Literature element 3.b HRQoL impact of organ involvement

Utility values for each SLICC element were drawn where possible by the manufacturer from HTAs available on the NICE website.

Paralleling the assumption that the average SLICC score for new organ involvement would be the average observed across the JHU cohort, the weights attached to each SLICC element utility value are the proportion of those elements observed within the JHU cohort. The resulting weighted average is raised to the power of the average SLICC score for those with that organ involvement within the JHU cohort as given in Table 6.16 of the MS.

For instance, for the calculation of the pulmonary involvement HRQoL based upon the text of the submission is 0.70 in Table 16.19 of the MS (refer to Table 25).

Table 25: HRQoL calculation pulmonary involvement from Table 16.19

SLICC Element	HRQoL	JHU %	Weighted	JHU SLICC	Final
Pulmonary hypertension	0.61	33%	0.20		
Pulmonary fibrosis	0.73	42%	0.31		
Shrinking lung (Chest XRay)	1.00	2%	0.02		
Pleural fibrosis (Chest XRay)	1.00	20%	0.20		
Pulmonary infarction/resection	0.94	4%	0.04		
Average across pulmonary			0.77	1.31	0.70

These organ involvement HRQoL values are applied multiplicatively. For a patient having developed more than one SLICC organ involvement, only the lowest HRQoL multiplier is applied to the "clean" utility.

Literature element 3.c Cost impact of organ involvement

A similar approach is undertaken for the cost impacts of organ involvement as for the QoL impacts, only with the number of patients in the JHU cohort experiencing the individual elements among those having had an event within the organ class giving rise to the weight to apply. These weights can sum to more than one due to a patient being able to experience more than one event. As with the calculation of the quality of life impacts this will tend to overestimate costs in the incident year and early years after incidence.

As these cost elements are less well documented in the submission than the HRQoL elements the full set is outlined below, with more detail being available in Appendix 28 of the submission. There are some minor discrepancies between the figures in Table 26 and those given in Table 6.26 of the MS for reasons that are unclear, but these will not affect results.

Table 26: Average costs for organ involvement

	Unit Costs			Average total	
	Year 1	Year 2+	Weight	Year 1	Year 2+
Ocular				**£1,518**	**£17**
Cataract	£1,553		96%		
Retinal damage / optic atrophy	£103	£64	27%		
Neuropsychiatric				**£3,659**	**£1,131**
Cognitive impairment			24%		
OR major psychosis	£1,122	£1,122	8%		
Seizures requiring therapy for 6 months	£826		19%		
Cerebral vascular accident ever or resection excl mal.	£8,066	£2,266	38%		
Cerebral vascular accident ever or resection >1	£8,066	£2,266	2%		
Cranial or peripheral neuropathy			43%		
Transverse myelitis	£4,772	£2,386	4%		
Renal				**£1,746**	**£2453**
				To max	**£6479**
Pulmonary				**£9,678**	**£9,603**
Pulmonary hypertension	£22,488	£22,488	43%		
Pulmonary fibrosis			55%		
Shrinking lung (on chest radiograph)			3%		
Pleural fibrosis (on chest radiograph)			26%		
Pulmonary infarction or resection	£1,539		5%		
Cardiovascular				**£3,402**	**£500**
Angina or Coronary Artery Bypass Graft	£4,196	£368	31%		
Myocardial infarction	£4,322	£368	36%		
Cardiomyopathy	£724	£724	35%		
Valvular disease (dias/sys murmur)			25%		
Pericarditis x 6 mth or pericardiectomy	£2,079		14%		
Peripheral vascular				**£2,955**	**£591**
Significant tissue loss ever	£10,375	£368	21%		
Significant tissue loss > 1 site			0%		
Venous thrombosis with swelling	£1,501	£936	55%		
Gastrointestinal				**£2678**	**£0**
Infarction or resection of bowel	£2,848		93%		
Resection > 1 site	£2,848		1%		
Pancreatic insufficiency enzyme replacement			3%		
Musculoskeletal				**£5,372**	**£1,903**
				To min	**£1,514**
Muscle atrophy / weakness			11%		
Deforming or erosive arthritis	£3,112	£3,112	26%		
Osteoporosis with fracture or vert. collapse	£8,118	£1,148	49%		
Avascular necrosis	£1,359		37%		
Avascular necrosis 2	£1,359		3%		
Osteomyelitis			2%		
Ruptured tendon			12%		
Diabetes				**£2,313**	**£2,313**
Diabetes mellitus sufficient for some intervention	£2,313	£2,313	100%		
Malignancy				**£6,056**	**£0**
Malignant tumours	£6,056		100%		

4.2.4 Base case deterministic results

The base case deterministic results are presented below in Table 27, Table 28, Table 29, Table 30 and Table 31.

Both Table 5 and 6 of the MS and also the default belimumab costs in the electronic model, suggest that the [redacted]

Table 27: Base case deterministic results: All BLISS

		Without PAS		With PAS	
	SoC	Belimumab	Net	Belimumab	Net
Undiscounted survival Life Years	30.47	31.97	1.50	31.97	1.50
Discounted quantities					
Belimumab direct drug cost	..	£31,687	£31,687	[redacted]	[redacted]
Total cost	£97,583	£133,167	£35,584	[redacted]	[redacted]
QALYs	9.55	9.98	0.43	9.98	0.43
Base Case ICER			£82,909		[redacted]

Table 28: Base case deterministic results: Anticipated license population

		Without PAS		With PAS	
	SoC	Belimumab	Net	Belimumab	Net
Undiscounted survival Life Years	32.82	34.96	2.13	34.96	2.13
Discounted quantities					
Belimumab direct drug cost		£36,796	£36,796	[redacted]	[redacted]
Total cost	£103,591	£143,895	£40,303	[redacted]	[redacted]
QALYs	10.11	10.72	0.61	10.72	0.61
Base Case ICER			£66,170		[redacted]

Table 29: Base case deterministic results: Target population

		Without PAS		With PAS	
	SoC	Belimumab	Net	Belimumab	Net
Undiscounted survival Life Years	31.93	34.87	2.93	34.87	2.93
Discounted quantities					
Belimumab direct drug cost	0	£47,008	£47,008	[redacted]	[redacted]
Total cost	£105,366	£157,291	£51,925	[redacted]	[redacted]
QALYs	9.81	10.61	0.81	10.61	0.81
Base Case ICER			£64,410		[redacted]

From the above the difference between the estimates of cost effectiveness for the anticipated license population and the Target population are relatively minor. Given this, from an economic point of view it is unclear why the manufacturer niches belimumab to only those with an SS score of at least 10 at baseline: around 67% (n=396) of the anticipated license population (n=592) within the trials.

Table 30: Base case organ involvement to death MS Table 6.43: Target population

	SoC	Belimumab	Net
Cardiovascular	23.9%	21.3%	-2.6%
Diabetes	17.9%	19.2%	1.3%
Gastrointestinal	22.1%	25.0%	3.0%
Malignancy	32.0%	34.1%	2.2%
Musculoskeletal	48.5%	48.9%	0.4%
Neuropsychiatric	44.7%	45.8%	1.1%
Ocular	35.1%	36.0%	0.8%
Peripheral vascular	21.5%	20.8%	-0.7%
Premature gonadal failure	7.2%	7.2%	0.0%
Pulmonary	39.9%	36.8%	-3.1%
Renal	24.3%	19.2%	-5.1%
Skin	7.9%	7.9%	0.0%

Table 31: Base case discounted costs: Target population

	SoC	Belimumab	Net
Belimumab therapy	..	£47,008	£47,008
Belimumab administration	..	£9,059	£9,059
Belimumab total costs	..	£56,067	£56,067
Other Costs			
SS score related costs	£27,882	£28,130	£248
Organ damage costs			
Cardiovascular	£1,838	£1,633	-£205
Diabetes	£2,493	£2,731	£238
Gastrointestinal	£359	£399	£40
Malignancy	£998	£1,031	£33
Musculoskeletal	£9,758	£10,114	£356
Neuropsychiatric	£6,434	£6,719	£286
Ocular	£392	£391	-£1
Peripheral vascular	£1,380	£1,339	-£41
Premature gonadal failure	£0	£0	£0
Pulmonary	£42,692	£39,652	-£3,040
Renal	£11,139	£9,083	-£2,056
Skin	£0	£0	£0
Total other costs	£105,366	£101,224	-£4,142
Total costs	£105,366	£157,291	£51,925

As outlined in Table 31 the direct drug costs of belimumab account for around 90% of the total net costs with the administration costs of belimumab accounting for another 17% of the total net costs: taken together roughly 108% of the total net cost. Administration costs are a significant proportion of the total direct cost of belimumab: 16% without the PAS and approximately [redacted] with the PAS.

The main anticipated cost savings arise from reduced rates of pulmonary disease and renal involvement, which generate cost offsets of around -6% and -4% of the total net costs respectively. Some additional costs are associated with belimumab due to the anticipated undiscounted survival gain of 2.93 life years causing some complications to occur in a higher proportion of patients and to be experienced for longer.

4.2.5 Base case probabilistic results

The base case probabilistic results do not appear to be presented within the written submission, with only the CEACs being presented. Refer to Figure 25, Figure 26 and

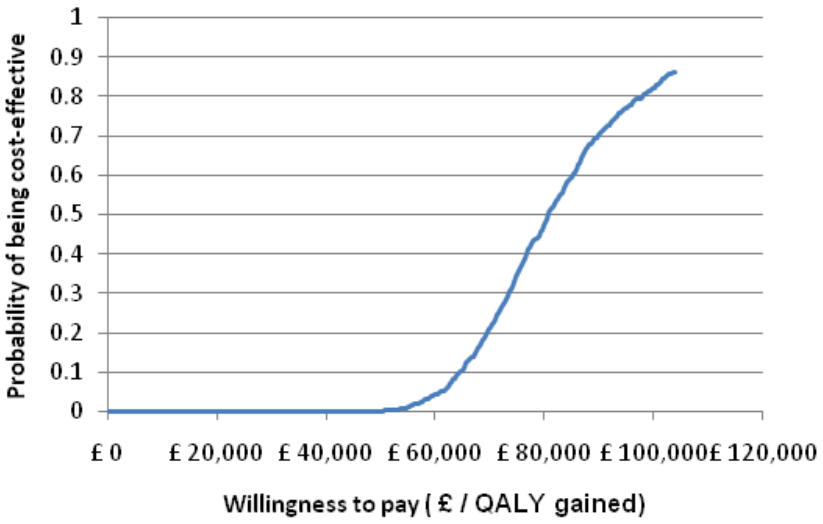


Figure 27.

Figure 25: CEAC without PAS (MS Fig 6.27) All BLISS

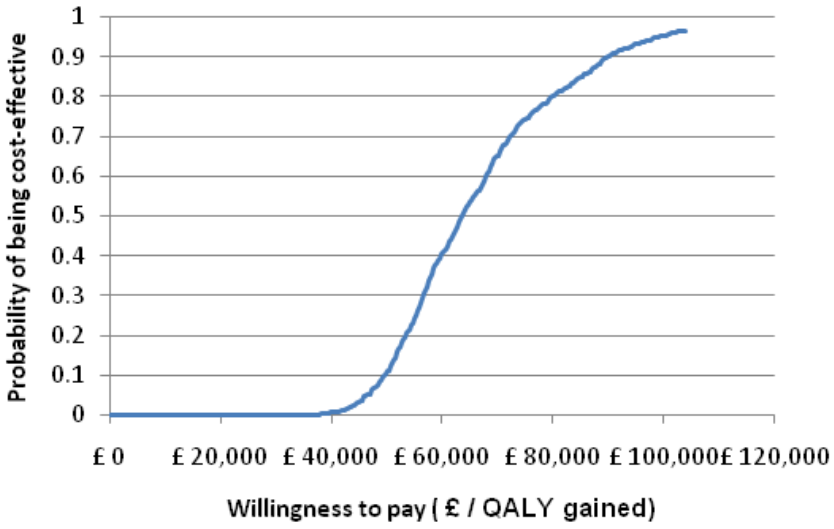


Figure 26: CEAC without PAS (MS Fig 6.41) Target population

[redacted]

Figure 27: CEAC with PAS, corrected in response to clarification question: Target population

[redacted]

Manufacturer's sensitivity and scenario analysis

The presentation of the sensitivity analyses within the submission and the responses to ERG clarification questions are mainly limited to those for the Target population. The results of

sensitivity analyses presented by the manufacturer for other groups are only presented here when that for the Target population is not supplied by the manufacturer.

As outlined in the summary of the model structure for the evolution of SS scores the manufacturer estimated a regression from the JHU cohort that related the change in the SS score to the average SS score in the previous period, and to gender, ethnicity and age. The constant estimated from this regression was 2.0577. The modelling applied a value of 3.0 on the basis of 2.0577 providing a poor fit to the evolution of SS scores within the All BLISS data. The effect of this upon the cost effectiveness for the All BLISS group is as in Table 32.

Table 32: Varying the constant in the SS change regression: All BLISS

Regression constant	ICER
2.0577	£93,654
2.5	£85,394
3.0	£82,909
3.5	£80,988

The univariate sensitivity analyses undertaken by the manufacturer are presented graphically as a tornado diagram in Figure 6.37 of the MS, with the values underlying this shown in Table 33. These values relate to the without PAS scenario for which the base case Incremental Cost-Effectiveness Ratio (ICER) is estimated as £64,410 per QALY.

Table 33: Manufacturer univariate sensitivity analysis – Target population

	Base	Low		High	
Variable	**Value**	**Value**	**ICER**	**Value**	**ICER**
Belimumab wk 24 responders % SS change baseline to 52 weeks	-0.28	-0.383	£49,393	-0.173	£103,840
All belimumab patients % SS change baseline to 52 weeks	-0.343	-0.437	£50,335	-0.251	£96,031
AMS on mortality	0.213	0.085	£85,677	0.333	£50,962
Annual % belimumab week 24 responders on treatment	0.92	0.863	£54,518	0.981	£85,893
Ln(age) of the "clean utility" regression	-0.145	-0.18	£78,448	-0.103	£53,263
Constant coefficient in "clean utility" regression	1.297	1.146	£79,243	1.426	£55,493
SoC patients % SS change baseline to 52 weeks	-0.349	-0.394	£77,351	-0.307	£55,581
AMS of the natural history Pulmonary model	0.139	0.06	£73,044	0.216	£55,216
Constant of the natural history Neuropsychiatric model	-7.396	-9.934	£61,333	-5.117	£76,231
Ln(age) of natural history Neuropsychiatric model	0.607	0.026	£61,514	1.226	£76,261
AMS coefficient of the natural history Renal model	0.323	0.228	£69,696	0.412	£56,744
Constant of the adjusted natural history SS model	3	2.202	£73,226	3.934	£61,871
Constant coefficient from the natural history PV model	-11.69	-16.47	£65,935	-6.81	£55,396
Ln(age) of natural history Pulmonary model	1.232	0.593	£70,841	1.916	£79,571
Constant of the natural history Renal model	-8.29	-9.01	£67,867	-7.56	£60,057

Many of the key variables within the sensitivity analyses tornado diagram are as would be expected: including the changes in SS scores, and the impact of SS scores through the AMS upon mortality,

As for the All BLISS sensitivity analysis the adjustment to the constant for the adjusted natural history model of the evolution of the SS score, demonstrates a similar impact as in the model for the Target population. The value of 2.202 (close to the original estimate of 2.0577) gives a cost effectiveness estimate of £73,226 per QALY. The alternative estimate of 3.9 results in a cost effectiveness estimate of £61,871 per QALY. One point to note is that this is

non-linear: an increase of 0.934 improves the cost effectiveness estimate by £2,539 per QALY while a smaller reduction of 0.798 worsens the cost effectiveness estimate by £8,816 per QALY.

Given the cost outputs of the model as previously summarised, the specifications of the natural history models for pulmonary and renal disease also have an impact. Perhaps more surprising is the influence of the neuropsychiatric natural history model.

Note that the manufacturer has presented additional information in Tables B8.2 and B8.3 in the response to ERG clarification questions. These outline the sensitivity of net QALYs and net costs to the univariate sensitivity analyses. The neuropsychiatric natural history model mainly impacts upon QALYs. Within net costs, the renal natural history model also has an impact.

A key variable that has not been particularly explored or explained within the submission is the assumed rate of continuation and discontinuations among belimumab week 24 responders. The ICERs reported in MS provide insufficient detail. Table 34 shows that both the net QALYs and the net cost are increasing in the annual continuation rate for belimumab week 24 responders, as would be anticipated. Note that the longer belimumab week 24 responders are estimated to remain on belimumab treatment the worse the estimated cost effectiveness is for belimumab.

Table 34: Sensitivity to continuation rate for belimumab week 24 responders: Target population

	Low value	Central value	High value
Annual continuation rate	0.863	0.920	0.981
Net QALY	0.649	0.806	1.165
Net Cost	£35,386	£51,925	£100,094
ICER	£54,518	£64,410	£85,893

The manufacturer also presented a range of scenario analyses for the Target population for the without PAS scenario. Compared to the base case estimate of £64,410 per QALY:

- Excluding the continuation rule at week 24 worsens the ICER to £72,207 per QALY
- Tightening the continuation rule to SS change ≥ 6 improves the ICER to £50,114 per QALY
- A 12% higher belimumab price worsens the cost effectiveness to £71,297 per QALY
- Vial sharing improves the cost effectiveness to £61,671 per QALY

- Using the original natural history model worsens the cost effectiveness to £77,707 per QALY
- An administration cost of £159 worsens the cost effectiveness to £67,353 per QALY

4.2.6 Base case deterministic results

Base case results

Simulating 50,000 patients within the deterministic model results in mean estimates that cross check with those presented within the submission.

4.2.7 Base case probabilistic results

Due to the patient level simulations for the base case deterministic results being run across 50,000 simulations, running the model for a sufficiently large number of iterations for the probabilistic modelling to reliably generate a central estimate and associated CEAC takes around one week. Time constraints have precluded the ERG from running the model probabilistically for all the subgroups and for any of the sensitivity analyses.

The results of an ERG probabilistic run of the model for the Target patient population, retaining 50,000 patients over 1,000 iterations, cross check with that of the manufacturer, the likelihood of belimumab being cost effective for a given willingness to pay (WTP):

- 0% at £30,000 per QALY
- 1% at £40,000 per QALY
- 11% at £50,000 per QALY
- 40% at £60,000 per QALY

In addition, the central estimates from the ERG probabilistic run of the model are an additional 0.84 QALYs at an additional cost of £55,166 to yield a central estimate of cost effectiveness of £65,530 per QALY. Both net QALYs and net costs are slightly higher than the deterministic run of the model, 0.81 and £51,925 respectively, but both rise in roughly equal proportion and the central estimate from the probabilistic modelling of £65,530 per QALY is little different from the deterministic estimate of £64,410 per QALY.

4.2.8 Data inputs

For correspondence between written submission and electronic model related to the Target population please refer to Table 35.

Table 35: Correspondence between MS and electronic model: Target population

Written submission		Electronic model		Correspondence
Model parameters specific to Target population				
Data	Table	Worksheet	Cells	Correspondence
Baseline demographics	6.37	Baseline Patient Characteristics Subgroup BLISS data	D9 Q7:Q62, P64, P78, P69, P240	Yes
Baseline disease activity	6.38	Baseline Patient Characteristics Subgroup BLISS data	D21:D32 Q185:Q215 P258:Q258 P260:Q260 P262:Q262 P265:Q265 P267:Q267 P269:Q269 P271:Q271 P273:Q273	Yes
Baseline SDI items	6.39	Subgroup BLISS data	P245:T256	Yes
Discontinuation rates	6.40	Treatment Effect Baseline Patient Characteristics Subgroup BLISS data	L33:L34 AD38 Q185:Q215 AR9:AR39 P276:Q276	
Week 52 SS regression	6.41	Treatment Effect	D12:E14 P221:Q223	Yes
Week 24 response by SS	6.42	Baseline Patient Characteristics Subgroup BLISS data	AD7:AD38AR9:AR39	Yes
Model parameters common to all subgroups				
Data	Table	Worksheet	Cells	Correspondence
Long term SS regression	6.9	Natural History Short	D11:D16	As per the text of the submission, the constant within Table 6.9 is increased from 2.0577 to 3.0000. The electronic model corresponds with the constant being 3.0000
Steroid use	6.11	PSA Inputs	HY7:HZ7	Yes
Mortality weibull	6.12	Natural History Model data	AG8:AG52	Yes
SLE SMR by age	6.13	PSA Inputs	IB7:IE7	Yes
Organ damage tte	6.14	Natural History Model data	AI8:AT52	Yes
JHU cohort characteristics	6.15	Hopkins Patient Characteristics	D8:D38	Possibly not. The electronic model highlights a larger range of cells in white than the five rows within Table 6.15: 19 cells in total. These cells are described within the electronic model as "*Average characteristics imputed to simulate non-trial*

Written submission		Electronic model		Correspondence
				characteristics that are used to determine long term outcome risks". Some of these elements of the electronic model may relate to those given in Table 6.16 of the submission Baseline hypertension within the electronic model is given as 53.1% compared to the 15.8% annual risk within Table 6.15. Note that the electronic model given the annual infection probability as 15.8%
JHU SLICC scores	6.16	Hopkins Characteristics	D8:D38?	Possibly not. There is no ready read across between D8:D38 of Hopkins Characteristics and Table 6.16. For instance, the electronic model give a value of 9.7% renal damage (mean) while Table 6.16 gives a renal score of 1.83. Even if the renal damage among the 9.7% had been at the maximum SLICC damage level of 4 it is difficult to see how this can result in a renal score of 1.83
Organ HRQoL impact	6.19	QoL Inputs	X9:AI10	Yes
"Clean" utility equation	6.19	QoL Inputs	D10:D13	No. The electronic model applies the values given in Table 6.20 and not those given in Table 6.19 and in the body of the text of the submission
"Clean" utility equation	6.20	QoL Inputs	D10:D13	Yes
Cost for a given SS score	6.25	Other cost inputs	D9:D29	Yes
Organ damage cost	6.26	Other cost inputs	W8:AH9	Partial. The year 1 and year 2 costs in the electronic model are as per Table 6.26 But note that the musculoskeletal annual cost declines between year 2 and year 17 from £1,903 to £1514 Also note that the renal annual cost increases

Written submission		Electronic model		Correspondence
				indefinitely from year 2 onwards from £2,453 to £6,749 by year 50
Belimumab annual costs	6.27	Belimumab cost	I23:I38	Yes

4.2.8.1 Model structure

Belimumab dose and direct drug cost

There appear to be minor errors in the calculation of the average belimumab drug cost per administration. For instance, for a patient weight of 50kg the manufacturer calculates that this is most cheaply administered using five 120mg vials at a cost of £571 per administration. This results in a total available dose of 600mg and wastage of 100mg, when a combination of one 400mg and one 120mg dose results in a total available dose of 520mg and wastage of only 20mg.

Note that if the simpler approach of using 400mg vials being used for the dose up to a multiple of 400mg with anything in addition to this being topped up through use of 120mg vials, wastage and the average drug cost would increase.

The drug cost for a patient of the mean patient weight can be calculated on the same basis as the manufacturer uses for individual patient drug cost calculations. Within an individual patient simulation model the approach of the manufacturer is correct. But much of the modelling submitted to NICE employs Markov modelling of a cohort of representative patients. This may use a weighted average drug cost, but it is also not unknown for the drug cost for the representative patient to be calculated. See Table 36.

Table 36: Belimumab average direct cost per administration

	Manufacturer	ERG	400mg	Mean		Median	
All BLISS	£671	£664	£694	67kg	£686	63kg	£610
Anti-dsDNA+ve, low (C3/C4)	£650	£642	£671	65kg	£686	62kg	£610
Target population	£654	£646	£674	65kg	£686	63kg	£610

Any errors within the manufacturer calculations are slight and will not materially affected the estimates of cost effectiveness. For the most severe subgroup of patients adopting the approach of using 400mg and topping up with 120mg rather than minimising waste has a similar effect to costing at the mean patient weight, with both increasing the average drug cost by a little over 3%.

As the belimumab drug costs account for around 90% of the estimated total net cost for both the All BLISS population and the most severe subgroup with an SS ≥ 10 at baseline, any change in the belimumab drug costs will lead to a roughly proportionate change in the cost effectiveness estimate.

Modelling the evolution of SS scores during the 1st year

The likelihood of the patient dying or developing organ involvement is not directly determined by the SS score, but this flow through to the AMS score which does determine mortality and organ involvement. The SS score has direct impacts upon patient utilities and the costs of treatment. The SS score also determines the patient steroid use, which in turn further affects the likelihood of mortality and organ involvement. The SS score is the key variable within the modelling.

The manufacturer response to ERG clarification question B3 outlines that the change in the SS score from baseline to week 52, SS_{52}-SS_0, for belimumab week 24 non-responders is calculated using the week 52 linear regression for the SoC arm. The reason for this is that belimumab week 24 non-responders cease belimumab treatment at week 24 and as a consequence at week 52 are receiving only SoC.

But calculating SS_{52} for belimumab week 24 non-responders using the week 52 linear regression for the SoC arm is likely to be incorrect. Both the SoC arm and the belimumab arm of the trials had week 24 responders and week 24 non-responders. For the Target population these are shown in Table 37 below.[c]

Table 37: Week 24 response rates – Target population

	BLISS-52		BLISS-76		Pooled	
	Resp	NResp	Resp	NResp	Resp	NResp
SoC	59%	41%	44%	56%	52%	48%
Belimumab	71%	29%	62%	38%	67%	33%
OR	1.75		2.07		1.93	

By definition the 52% of patients within the pooled SoC arm with response at week 24 had a change of at least 4 in their SS score at week 24, while the 48% without response at week 24 had a change of less than 4 in their SS score at week 24. Similarly by definition, the 33% within the pooled belimumab arm without response at week 24 had a change of less than 4 in their SS score at week 24.

[c] Based upon the patient numbers reported in table A7.1 of the manufacturer response to ERG clarification question A7.

It seems likely to be more appropriate to model SS_{52} for the belimumab week 24 non-responders based upon the changes in SS scores among the SoC week 24 non-responders than upon changes in SS scores across all SoC patients, the latter being an average across week 24 responders and week 24 non-responders. This is underlined by the response of the manufacturer to ERG clarification question A7 and associated Table A7.1 outlining the mean changes in SS scores at week 24 and week 52 by arm and week 24 responder status. See Table 38.

Table 38: SS changes at week 24 and by week 52 by week 24 status – Target population

	BLISS-52			BLISS-76			Pooled		
Week 24 status	Resp	NResp	All	Resp	NResp	All	Resp	NResp	All
Mean change from baseline at week 24: SS_{24}-SS_0									
SoC	-6.7	-0.7	-4.2	-7.1	-1.3	-3.9	-6.9	-1.1	-4.1
Belimumab	-7.4	-0.8	-5.5	-7.0	-1.0	-4.7	-7.3	-0.9	-5.2
Mean change from baseline at week 52: SS_{52}-SS_0									
SoC	-5.4	-2.3	-4.1	-6.1	-2.3	-4.0	-5.7	-2.3	-4.1
Belimumab	-7.5	-3.3	-6.3	-6.8	-2.5	-5.2	-7.2	-2.9	-5.8

Among week 24 non-responders the average changes in SS score at week 24 were very similar between the SoC arms and the belimumab arms: for instance these were -1.1 and -0.9 respectively within the All BLISS data.

It could be argued that among week 24 non-responders the average changes in SS score have started to slightly divergence between the SoC arms and the belimumab arms by week 52: -2.3 and -2.9 respectively within the All BLISS data. But for modelling purposes these figures are not particularly relevant. The stopping rule of the modelling is applied at 24 weeks. As a consequence, the evolution of SS scores among week 24 non-responders in the belimumab arm between week 24 and week 52 within the trials is of less interest. It reflects continued use of belimumab between week 24 and week 52 during the trials, when the modelling assumes that these patients will no longer receive belimumab between week 24 and week 52. It is presumably for this reason that SS_{52} has to be modelled for belimumab week 24 non-responders rather than drawn directly from trial data.

In the light of the above (see Table 38), for the belimumab week 24 non-responders it would seem to have been more appropriate to base SS_{52}-SS_0 upon the parallel change for SoC week 24 non-responders than that for the SoC arm as a whole: -2.3 rather than -4.1 within the All BLISS data. This is quite a large difference of -1.8 for belimumab week 24 non-responders given the overall changes in average SS scores within the trials. Belimumab week 24 non-responders make up 33% of the belimumab Target population within the trials.

The trial data reported in Table 39 is illustrative of the assumption underlying the manufacturer model, with the figures relating to trial data. The model makes a parallel assumption, but estimating SS_{52}-SS_0 uses the linear regressions as outlined in Table 6.41 of the MS and the manufacturer response to question C6 of the ERG clarification questions (see Table 39).

Table 39: Linear regression of SS_{52}-SS_0 central parameter estimates – Target population

	BLISS-52	BLISS-76	Pooled
SS_{52}-SS_0 SoC	-0.3629	-0.3341	-0.3493
SS_{52}-SS_0 belimumab	-0.3746	-0.3153	-0.3435
SS_{52}-SS_0 belimumab week 24 responders	-0.2626	-0.2827	-0.2800

Where SS_{52}-SS_0 SoC is the average SS change within the SoC arm, SS_{52}-SS_0 belimumab is the average SS change within the belimumab arm and SS_{52}-SS_0 belimumab week 24 is the additional average SS change within the belimumab arm among those showing a response at week 24. Note that these are regression coefficients and multiplicative: e.g. from the All BLISS data a patient within the SoC arm with, for example, $SS_0 = 10.00$ has a central estimate of SS_{52}-SS_0 = -0.349*10 = -3.49 hence $SS_{52} = 6.51$.

Note that the data in Table 39 will include any additional treatment effect from belimumab between week 24 and week 52 among belimumab week 24 non-responders. To the extent that this effect exists it will tend to lead to an overestimate of the effectiveness of for SS_{52}-SS_0 belimumab within the regression model, but this will tend to net out through a reduction in the estimate of SS_{52}-SS_0 belimumab week 24 responders. Within the modelling it is only really the sum of these two coefficients that it applied.

As outlined within the manufacturer response to the ERG clarification question B3, SoC patients are assumed to have the SoC coefficient applied, belimumab week 24 non-responders are assumed to have the SoC coefficient applied and belimumab week 24 responders have the sum of the two belimumab coefficients applied. Given this the model application of the results of the regression can be more transparently presented as below. These are then applied to the trial mean of $SS_0 \approx 12.7$ among the Target population to result in a modelled estimate for the trial mean which can then be compared with the actual trial mean for some simple triangulation and model validation.

Table 40: SS_{52}-SS_0 model versus trial – Target population

All BLISS	Coeff.	Modelled SS_0=12.7	Trial mean
SS_{52}-SS_0 SoC week 24 non-responders	-0.349	-4.4	-2.3
SS_{52}-SS_0 SoC week 24 responders	-0.349	-4.4	-5.7
SS_{52}-SS_0 SoC All (weighted average)		-4.4	-4.1
SS_{52}-SS_0 belimumab week 24 non-responders	-0.349	-4.4	-2.9
SS_{52}-SS_0 belimumab week 24 responders	-0.623	-7.9	-7.2
SS_{52}-SS_0 belimumab All (weighted average)		-6.8	-5.8

As anticipated, data shown in Table 40 suggests that assuming that belimumab week 24 non-responders have the same change in SS scores as the average for the SoC arm systematically overestimates the average change in the SS score within belimumab arm. Note that this also does not correct for any impact of belimumab week 24 non-responders being assumed to cease treatment at week 24, which may suggest that the -2.9 trial mean will be an overestimate of the likely effect in clinical practise and the value that should be applied within the modelling.

Note also that the electronic copy of the model includes the coefficients for the parallel regression of the change in SS scores at week 24: SS_{24}-SS_0

Table 41 mirrors the week 24 trial data, but appears unduly pessimistic for the belimumab week 24 non-responders when coupled with the average baseline of $SS_0 \approx 12.7$. This regression is not used within the current model implementation and is not presented within the written submission.

Table 41: Linear regression of SS_{24}-SS_0 central parameter estimates – Target population

	BLISS-52	BLISS-76	Pooled
SS_{24}-SS_0 SoC	n.a.	n.a.	-0.3525
SS_{24}-SS_0 belimumab	n.a.	n.a.	-0.0003
SS_{24}-SS_0 belimumab week 24 responders	n.a.	n.a.	0.5755

The manufacturer in response to ERG clarification question C8 confirms that the central estimates of the modelled evolution of SS scores for SoC, belimumab week 24 non-responders and belimumab week 24 responders for the first 10 years[d] of the modelling are shown in Table 42.

[d] For all the SS and AMS elements only the first 10 years modelling is presented for illustrative purposes, while the modelling extends to the patient lifetime.

Table 42: Modeled evolution of SS scores – Target population

Year	SoC	Bel. week 24 non-responders	Bel. week 24 responders
0	12.7	12.7	12.7
1	8.2	8.1	4.9
2	6.5	6.4	3.4
3	5.5	5.5	2.6
4	4.9	4.9	2.3
5	4.6	4.6	2.2
6	4.4	4.4	2.2
7	4.3	4.3	2.2
8	4.2	4.2	2.3
9	4.1	4.1	2.4
10	4.1	4.1	2.5

On the basis of a 67:33 split between belimumab 24 week responders and non-responders, the weighted average SS score within the belimumab arm is modelled as being 5.95 at the end of year 1: a net gain over SoC of around 2.2. This compares with a net observed gain from belimumab over SoC in the trials of 1.7: the modelled net improvement is around 30% greater than that observed in the trials. If the SoC week 24 non-responder change of -2.3 at week 52 is assumed to apply to the belimumab week 24 non-responder at week 52 on the grounds that these patients are assumed to cease belimumab treatment from week 24 the net "observed" gain from belimumab over SoC in the trials falls to 1.5: the modelled net improvement consequently increases to around 45% over that "observed" in the trials.

In short, given the trial data for the average changes in SS scores for SoC and belimumab week 24 responders and non-responders, it seems unwarranted to assume that belimumab week 24 non-responders will have the same change in SS score at week 52 as the average observed across the SoC arm. This overestimates the impact of belimumab upon SS scores.

Modelling of AMS scores

It is not the SS score but the AMS score that directly contributes to the likelihood of a patient dying and a patient developing cardiovascular, gastrointestinal, musculoskeletal, neuropsychiatric, ocular, peripheral vascular, pulmonary, renal and/or skin involvement. Renal involvement further determines the likelihood of cardiovascular involvement.

As explained previously the AMS is calculated as the area under the SS score curve between two time points divided by the time of follow-up to provide an average score over the period of interest. Within the response to ERG clarification question C8 the manufacturer clarified the central estimates of the SS scores and AMS as below in Table 43.

Table 43: Modeled evolution of AMS scores: manufacturer clarification – Target population

Year	SoC		Bel. week 24 non-responders		Bel. week 24 responders	
	SS	AMS	SS	AMS	SS	AMS
0	12.7	12.7	12.7	12.7	12.7	12.7
1	8.2	10.4	8.1	10.4	4.9	9.9
2	6.5	9.1	6.4	9.1	3.4	8.4
3	5.5	8.2	5.5	8.2	2.6	7.4
4	4.9	7.6	4.9	7.5	2.3	6.8
5	4.6	7.1	4.6	7.0	2.2	6.3
6	4.4	6.7	4.4	6.7	2.2	5.9
7	4.3	6.4	4.3	6.4	2.2	5.6
8	4.2	6.1	4.2	6.1	2.3	5.4
9	4.1	5.9	4.1	5.9	2.4	5.2
10	4.1	5.8	4.1	5.8	2.5	5.0

Running the model with a 100% probability of response for belimumab and a 0% probability of response[e] for belimumab results in the same SS scores as above for SoC, Belimumab week 24 non-responders and belimumab week 24 responders.

For both the SoC arm and the belimumab week 24 non-responders, the AMS at time T as reported in the response to clarification question C8 cross checks with being the average of the SS values t = 0…..T. But for the belimumab week 24 responders the AMS at time T as reported in the response to clarification question C8 bears no relation to the ERG cross check (see Table 44).

[e] Implemented by setting the *Baseline Patient Chars* worksheet AD7:AD38 = 1 for 100% response and = 0 for 0% response and taking the values from the *Results* worksheet columns BR:BS.

Table 44: Modelled evolution of AMS scores: Belimumab week 24 responders – Target population

Year	Belimumab week 24 responders		
	SS	Manufacturer AMS	ERG AMS
0	12.7	12.7	12.7
1	4.9	9.9	8.8
2	3.4	8.4	7.0
3	2.6	7.4	5.9
4	2.3	6.8	5.2
5	2.2	6.3	4.7
6	2.2	5.9	4.3
7	2.2	5.6	4.1
8	2.3	5.4	3.9
9	2.4	5.2	3.7
10	2.5	5.0	3.6

The values submitted in response to the ERG clarification question are averages across many individual patient iterations. Some non-linearity or rounding approximation may have crept into the figures. But it is difficult to reconcile the ERG cross check of the AMS for SoC and belimumab week 24 non-responders with the discrepancies between the ERG cross check and the manufacturer reported values of the AMS for belimumab week 24 responders.

It is unclear whether the above discrepancy is due to an error in the manufacturer response to the ERG clarification question, an error in the VB coding of the model or an error in interpretation by the ERG. The manufacturer figures for the AMS may incorporate discontinuations within the belimumab week 24 responder figures.

Ignoring this discrepancy for the moment, by definition the AMS introduces memory of previous SS scores to the modelling. As a consequence of this, even when there is no modelled contemporaneous difference in SS scores between the arms at a particular point, the AMS scores will retain a memory of previous differences in SS scores between the arms. Any errors in the calculation of SS scores within the first year will, even if largely washed out over a relatively short period due to a high discontinuation rate within the belimumab arm, continue to be carried forward by the AMS.

Patients within the JHU cohort were recruited at somewhat lower values in their SS scores compared to the BLISS trials, Figure 6.8 of the MS suggesting an average SS score in their first year of around 2.8 with the subsequent AMS being reasonably level or declining slightly over the 17 years of data presented within Figure 6.8. At a minimum it seems possible that

JHU cohort type patients would have a history of lower SS scores prior to being eligible for recruitment to the BLISS trials. Any history of SS scores prior to baseline will tend to dampen the impact that changes in SS scores at baseline have upon the AMS.

As noted by the manufacturer, the AMS tends to smooth out changes in the SS scores. But by not taking into account a patient's SS score prior to the model baseline, the model effectively ignores this smoothing out effect and exaggerates the impact that the changes in SS scores at baseline have upon the AMS and upon the net difference in the AMS between the arms. For the AMS at time T from model baseline for belimumab week 24 responders from the written submission it seems that the intention of the model is to calculate it as:

With the net difference in the AMS at time T between belimumab week 24 responders and SoC as:

But this ignores the previous history of SS scores since diagnosis D years prior to the model baseline. Taking this into account results in an AMS at time T for belimumab week 24 responders of:

With the net difference in the AMS at time T between belimumab week 24 responders and SoC as:

Which implies that:

Manufacturer clarifications on SS scores suggests that SS scores among belimumab week 24 responders are never worse than the contemporaneous SS score for SoC, and typically appear

to be modelled as being superior[f]. Since (T+D)/T is greater than one this implies that modelled superiority in AMS for belimumab week 24 responders over SoC systematically overstates the actual superiority in AMS for belimumab week 24 responders over SoC due to not having taken into account the previous patient history.

Any overstatement of effect upon the AMS arising because of this will be larger during the early years of the modelling, and for those patients with a long prior history of SLE. The average duration of disease at recruitment was 6 -7 years within the Target population. The average duration of disease at recruitment to the JHU cohort was a little over 5 years.

Calculation of steroid dose

The ERG expert opinion is that steroid use is variable, but that the tapering allowed within the trials was not unrepresentative of UK practice. In response to ERG clarification question A8 the manufacturer has clarified that within the BLISS trials the evolution of steroid use in the Target population was as shown in Table 45.

Table 45: Average steroid use (mg): BLISS Target population

	BLISS-52		BLISS-76		All BLISS	
	SoC	Belim.	SoC	Belim.	SoC	Belim.
Week 24 responders:	n=63	n=80	n=42	n=50	n=105	n=130
Baseline	12.4 ± 8.6	13.7 ± 10.8	8.9 ± 8.2	10.9 ± 7.6	11.0 ± 8.6	12.6 ± 9.8
Week 24	14.4 ± 11.7	11.5 ± 8.0	9.1 ± 6.7	11.4 ± 8.4	12.3 ± 10.3	11.5 ± 8.1
Week 52	10.5 ± 6.3	8.6 ± 5.9	7.5 ± 6.2	11.9 ± 21.9	9.3 ± 6.4	9.8 ± 14.1
Week 24 non-responders:	n=44	n=32	n=54	n=31	n=98	n=63
Baseline	13.4 ± 8.1	13.6 ± 9.3	11.3 ± 9.2	9.5 ± 8.9	12.2 ± 8.7	11.6 ± 9.3
Week 24	17.8 ± 30.2	14.1 ± 10.1	20.5 ± 54.8	12.5 ± 8.8	19.2 ± 44.8	13.3 ± 9.4
Week 52	13.3 ± 7.0	12.1 ± 9.3	9.9 ± 9.2	9.9 ± 8.3	11.4 ± 8.3	11.0 ± 8.8
Overall:	n=107	n=112	n=96	n=81	n=203	n=193
Baseline	12.8 ± 8.4	13.7 ± 10.4	10.3 ± 8.8	10.4 ± 8.1	11.6 ± 8.6	12.3 ± 9.6
Week 24	15.7 ± 20.6	12.1 ± 8.6	14.9 ± 39.5	11.8 ± 8.5	15.3 ± 30.6	12.0 ± 8.5
Week 52	11.4 ± 6.6	9.4 ± 6.9	8.6 ± 7.8	11.2 ± 18.6	10.2 ± 7.3	10.1 ± 12.9

Note that the standard deviations associated with these estimates suggests quite strongly skewed data. The manufacturer further notes in response to ERG clarification question A8 that "*corticosteroid taper during the study was determined strictly at the investigators'*

[f] Note that these are average figures across the 50,000 patient simulated. Within the patient level modelling clinical effectiveness estimates are treated deterministically, and as a consequence this seems likely to apply to each individual belimumab patient and its clone that is modelled within the 50,000 simulations.

discretion. There were no protocol mandates regarding dose reduction. The total dose of corticosteroids could be adjusted as clinically required during the first 24 weeks of the study; corticosteroid use beyond pre-specified dose limits resulted in the patient being designated as a non-responder."

A crude assessment of the above might suggest that steroid use at baseline is typically lower within BLISS-76 than within BLISS-52, which may reflect the different geographic recruitment for the two trials with BLISS-76 locations being more relevant to the UK.

The modelling ignores the above steroid use data on the grounds of it being unrepresentative of UK practice, choosing instead to use the relationship derived from the JHU cohort. The steroid dose is modelled as a linear function of the AMS score: steroid dose (mg/day) = 3.41 + 0.72 * AMS_T [Table 6.11].

This links with the trial based linear regression of changes in SS scores $SS_{52} = (1+\beta) * SS_0$ which implies that $AMS_{52} \equiv (SS_{52} + SS_0)/2 = (2+\beta)/2 * SS_0$ where for SoC $\beta = -0.349$ and for belimumab week 24 responders $\beta = -0.623$. Given a central baseline of $SS_0 = 12.7$ for the Target population this implies $AMS_{52} = 10.5$ for SoC and $AMS_{52} = 8.7$ for belimumab. For the Target population this in turn implies estimated daily steroid doses at the central baseline SS scores of:

- 12.6mg at baseline
- 11.0mg at week 52 for SoC
- 9.7mg at week 52 for belimumab week 24 responders

While the trial steroid use data appears to be quite skewed, the average steroid doses for the Target population are:

- For BLISS-52
    - 12.8mg at baseline and 11.4mg at week 52 for SoC
    - 13.7mg at baseline and 8.6mg at week 52 for belimumab week 24 responders
- For BLISS-76
    - 10.3mg at baseline and 8.6mg at week 52 for SoC
    - 10.9mg at baseline and 11.9mg at week 52 for belimumab week 24 responders
- For All BLISS
    - 11.6mg at baseline and 10.2mg at week 52 for SoC
    - 12.6mg at baseline and 9.8mg at week 52 for belimumab week 24 responders

It can be argued that the trial data as presented, and in particular the trial data from BLISS-76, does not triangulate particularly well with the steroid doses estimated within the model[g].

Note also that while the absolute difference in SS scores between a belimumab 24 week responder and her SoC clone is maintained while the patient remains on belimumab, the SoC SS score is modelled as falling over time as shown in Table 6.9 and Figure 6.7 of the MS. As a consequence of this, the absolute difference in steroid dose between a belimumab 24 week responder and her SoC clone will be modelled as falling over time.

Calculation of the CAPD

There is limited detail within the submission on the calculation of the CAPD. It seems possible that this may be subject to the same source of bias as the calculation of the AMS, if a patient's prior history and CAPD to date at baseline is not taken into account.

Calculation of the mortality probability

The JHU model survival excludes age and duration of disease on statistical grounds. The SLE SMRs from the Bernatsky reference[19] are then applied to the ratio of the JHU modelled patient specific probability of death with the JHU modelled SLE average probability of death. This appears quite convoluted. A more natural approach might have been to have reconsidered the treatment of age within the JHU survival model, and given the age ranges within the JHU cohort to assess whether the derived model would be applicable outside a certain age range.

The manufacturer justifies the application of the SLE SMRs as drawn from the Bernatsky reference[19] on the grounds of the JHU survival model being unrepresentative of older patients and so unsuitable for extrapolating into old age. This may be the case, but the SLE SMRs drawn from the Bernatsky reference[19] are larger for younger patients than older patients. Admittedly in the younger age group these will be being applied to lower general population mortality risks.

There is the concern that the multiplication by the Bernatsky SMRs may tend to exaggerate the impact of the covariates within the JHU cohort survival model, and of any differences in the values of the covariates as modelled between the arms of the model.

[g] Within this it should be borne in mind that the week 52 steroid dose for belimumab week 24 responders within the BLISS-76 trial has a particularly high standard deviation.

It can also be noted that the Bernatsky 2006 reference[19] provides a number of cuts of the data for the SMR estimates unadjusted and from a "*multivariate hierarchical regression to determine the independent effects of the factors examined (sex, age group, SLE duration, calendar year period of SLE diagnosis, country) on the relative SMR estimates among SLE patients*". Refer to Table 46.

Table 46: Bernatsky SLE SMRs

Unadjusted SMRs by	Unadjusted SMR	95% CI	Revised SMR	95% CI
Gender				
Female	2.5	2.3 – 2.7	1.2	1.0 – 1.4
Male	1.9	1.7 – 2.2	1.0	(ref)
Age				
16-24	19.2	14.7 – 24.7	..	..
25-40	8.0	7.0 – 9.1	..	..
< 40 (above pooled)	10.7	9.5 – 11.9	6.4	5.5 – 7.5
40-59	3.7	3.3 – 4.0	2.6	2.3 – 3.0
60+	1.4	1.3 – 1.5	1.0	(ref)
Duration SLE years				
< 1	5.4	4.7 – 6.3	7.7	5.9 – 10.2
1 – 4	2.5	2.2 – 2.8	3.2	2.5 – 4.1
5 – 9	2.1	1.9 – 2.4	2.4	1.8 – 3.0
10 – 19	2.0	1.8 – 2.3	1.8	1.4 – 2.2
20+	2.0	1.7 – 2.0	1.0	(ref)

The manufacturer argument for the need to apply the SLE SMRs to the JHU cohort survival model centres on speculation that the JHU cohort survival model does not correctly estimate survival probabilities for older SLE patients who were insufficiently represented within the JHU cohort. A possible approach would be to validate the JHU cohort survival model by examining to what extent the estimates of survival probabilities conform to the SMR estimates given within Bernatsky: both relative to the general population and relative to other SLE patients. If the model results triangulate well with the Bernatsky SMRs it is a good fit, if not it is not a good fit and needs revision. But it would seem sensible to check this first, prior to any ad hoc revisions.

The argument as to why the cohort survival model requires adjustment by the Bernatsky SMRs when modelling patients who are of a similar age to those within the JHU cohort is unclear.

There is also some concern around the SMR values applied from Bernatsky. A recent UK based study found somewhat lower SMRs for SLE as shown in Table 47.

Table 47: SMRs for cohort of UK SLE patients: Caroline Gordon (22 June 2011, personal communication)

Age	SMR	CI
20 - 24	[redacted]	[redacted]
25 – 34	[redacted]	[redacted]
35 – 44	[redacted]	[redacted]
45 – 54	[redacted]	[redacted]
55 – 64	[redacted]	[redacted]
65 – 74	[redacted]	[redacted]
75 – 84	[redacted]	[redacted]
≥ 85	[redacted]	[redacted]

[redacted]

Calculation of the SS score direct effect upon treatment costs

The one year observational cost study divided the data into two six month periods and examined the relationship between the SS score severity class 0, 1, 2 or 3 and the patient's 6 month cost. The SS score severity class was determined by the maximum SS score observed during the relevant 6 month period. To arrive at an annual cost related to the SS score severity class during the six months, the six monthly costs are simply doubled within the submission.

Given that annual costs are required for the model, there is the obvious question of why the cost data is not analysed on an annual basis. SS scores will have varied over the one year observational study, and the maxima are likely to have differed between the two 6 month periods for some if not all patients. It would be anticipated that costs will be highest in the period immediately around any peak in SS score, and will tend to fall away either side of this. Doubling the six monthly cost will tend to have projected a patient's high costs during one period into what was actually a lower cost period.

To labour the point, suppose that a flare leads to an SS score of 12 and a hospital admission of two weeks duration for a particular patient. If the observational cost data had been analysed on a monthly basis and the patient's peak in the SS score due to the flare was observed in the

data, the SS score of 12 or SS score severity class of 3 would in this instance be associated with a hospital admission of two weeks duration. The corollary of the manufacturer approach to annualisation would be to multiply this by 12, leading to the conclusion that in this instance an SS severity index of 3 sees the patient spend half the year admitted to hospital.

While the bias caused by the manufacturer approach will be less than that outlined in the hypothetical example above, it seems likely that it will have tended to exaggerate the association between the SS score average over the year and annual treatment costs. A simpler approach that averages patients' SS scores over the year and relates these to their annual cost would seem to be more in line with the natural history model and probably less likely to lead to bias.

Calculation of HRQoL and Cost impacts of newly incident organ involvement

It can be noted that the assumption that newly incident organ damage will be at the average SLICC score for that SLICC element within the JHU cohort has a possible double impact. As noted by the manufacturer, incident cases will by definition be 1 when incident rather than the average SLICC score for that element within the JHU cohort. But the element being involved at incidence may also tend to be a less serious element. The weighting given to the utility values within each SLICC score is also the average prevalence of the elements within the organ SLICC score within the JHU cohort. If the more serious elements tend to occur later, then not only will the number of elements being involved at incidence be overestimated, their seriousness might be as well.

The above applies with equal force to the costs associated with newly incident organ involvement.

Double counting of treatment costs

Within the manufacturer model there is a direct causal link between the SS score and the incidence of new organ involvement. This is perfectly reasonable. But the costs associated with the SS score are estimated entirely separately from the costs associated with individual organ involvement. Due to the positive association between SS scores and organ involvement, adding the cost associated with SS scores and the cost associated with organ involvement is likely to have double counted these costs to some degree.

The extent of this bias may be limited if rates of organ involvement within the observational cost study conducted during the phase II trial were low. In some sense, there is a need for the corollary of the "clean" utility function on the cost side of the modelling: a "clean" SS score cost function stripped of the impact of organ involvement upon costs.

4.2.9 ERG reconciliation of durations of organ involvement and undiscounted organ costs

The modelling for the target population results in quite large estimates of the net discounted cost savings from reduced pulmonary involvement with belimumab: £3,040 which is around 6% of the total net discounted cost estimate of £51,925. The model output also outlines that the average duration of pulmonary involvement is modelled as 9.87 years within the SoC arm and 9.50 years within the belimumab arm, these appearing to be undiscounted figures (see Table 48).

Table 48: ERG cross check of modelled pulmonary costs – Target population

	SoC	Belimumab	Net
Average survival undiscounted	31.93	34.87	2.93
Pulmonary involvement			
Baseline	2.5%	2.5%	0.0%
Final	39.9%	36.8%	-3.1%
Pulmonary costs cross check			
Pulmonary average duration (D)	9.87	9.50	-0.37
Pulmonary cost year 1 (£Y1)	£9,678	£9,678	
Pulmonary cost year 2+ (£Y2)	£9,603	£9,603	
Total costs (£Y1 + £Y2*(D-1))	£94,896	£91,308	-£3,587
Modelled costs			
Undiscounted costs from model	£94,852	£91,262	-£3,590
Discounted costs from model	£42,692	£39,652	-£3,040
Pulmonary involvement duration			
Pulmonary average duration (affected)	24.75	25.82	1.07

Given this, the modelled average duration of pulmonary involvement can be coupled with the average costs in year 1 and year 2+ for pulmonary involvement to arrive at the average undiscounted costs for pulmonary involvement. This cross check appears to tally very closely with the summary of the model output: within the SoC arm average undiscounted costs for pulmonary involvement of 94,896 compared to the £94,852 reported in the model output, and £91,308 compared to £91,262 for the belimumab arm. This is in part due to year 1 and year 2 costs being very similar for pulmonary involvement which means that the mean organ

duration is sufficient to characterise the undiscounted mean costs; the distribution of organ duration does not have to be taken into account.

But this cross check is based upon the average duration of pulmonary involvement reported within the model output being that applicable across the patient cohort; i.e. including those who are modelled as not developing pulmonary involvement. This implies an average duration of pulmonary involvement among those with pulmonary involvement at baseline or developing pulmonary involvement over the period of the model of 24.75 years for SoC and 25.82 years for belimumab. The has been confirmed as correct by the company that developed the model for the manufacturer, which in turn imples the following average undiscounted durations of organ involvement and average undiscounted organ cost among those having the relevant organ involved at some point during the modelling. Note that organ involvement at baseline was low. These can then be conditioned by the percentages having the relevant organ involved at some point during the modelling to arrive at the avergage organ cost across the cohort as a whole; i.e. including those not having the relevant organ involved at some point during the modelling (see Table 49).

Table 49: Mean undiscounted organ durations and costs – Target population

	Among those with the organ involved				Across the whole patient group			
	Duration		Undiscounted cost		Involvement		Undiscounted cost	
	SoC	Belim	SoC	Belim	SoC	Belim	SoC	Belim
Cardiovascular	23.48	24.53	£14,787	£15,313	24%	21%	£3,527	£3,260
Diabetes	14.72	15.68	£34,408	£36,656	18%	19%	£6,173	£7,035
Gastrointestinal	20.92	22.55	£2,696	£2,697	22%	25%	£595	£675
Malignancy	13.73	14.86	£6,119	£6,120	32%	34%	£1,955	£2,089
Musculoskeletal	23.16	24.83	£40,285	£42,833	49%	49%	£19,552	£20,952
Neuropsychiatric	24.98	26.36	£30,782	£32,349	45%	46%	£13,761	£14,826
Ocular	22.42	23.57	£1,897	£1,917	35%	36%	£666	£690
PV	17.01	18.02	£12,532	£13,130	22%	21%	£2,698	£2,729
Gon. Failure	24.53	25.78	£0	£0	7%	7%	£0	£0
Pulmonary	24.75	25.82	£237,795	£248,049	40%	37%	£94,852	£91,262
Renal	22.16	23.22	£103,220	£108,974	24%	19%	£25,060	£20,947
Skin	31.47	34.11	£0	£0	8%	8%	£0	£0

4.2.10 Comparison with NICE reference case

Table 50 provides a comparison between the MS basecase submission and the NICE reference case.

Table 50: Comparison with NICE reference case

Attribute	Reference case and TA Methods guidance	Does the *de novo* economic evaluation match the reference case
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	The main comparison is between belimumab adjunctive to standard therapy and standard therapy alone The NICE scope also includes rituximab as a comparators
Patient group	As per NICE scope	The manufacturer niches belimumab to those within the anticipated license with an SS score of at least 10
Perspective costs	NHS and Personal Social Services (PSS)	Yes
Perspective benefits	All health effects on individuals	Yes
Form of economic evaluation	Cost-effectiveness analysis	Cost utility analysis
Time horizon	Sufficient to capture differences in costs and outcomes	Yes
Synthesis of evidence on outcomes	Systematic review	As there is no consideration of rituximab or cyclophosamide there is no requirement for a synthesis of the evidence as the comparator is the standard care arm of the trials. The only synthesis of the trial data is the pooling of BLISS-52 with BLISS-76
Outcome measure	Quality adjusted life years (QALYs)	Yes
Health states for QALY	Described using a standardised and validated instrument	Yes. The "clean" utility linked to SS scores is derived from EQ-5D trial data The HRQoL impacts from further organ involvement are drawn from a range of studies within the literature
Benefit valuation	Time-trade off or standard gamble	Yes. The "clean" utility linked to SS scores using EQ-5D trial data applies the standard social tariffs from Dolan REFto arrive at utility values
Source of preference data for valuation of changes in HRQL	Representative sample of the public	Yes. The "clean" utility applies the standard social tariffs from Dolan to arrive at utility values
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes
Probabilistic modelling	Probabilistic modelling	Probabilistic modelling is presented for the base case results
Sensitivity analysis		A wide range of univariate sensitivity analyses and scenario analyses are included

4.3 ERG additional scenario and sensitivity analysis

The model runs 50,000 patient simulations for reliable convergence which takes some time.

[redacted]

[redacted]

Belimumab administration cost

The £126 per administration for two hours of senior nursing time cross checks with the 2009-10 PSSRU health care costs, though including qualification costs increases this slightly to £140 per administration.

ERG expert opinion is in line with the manufacturer in noting that the administration cost for belimumab would be similar to that for tocilizumab which is also a one hour IV infusion. Tocilizumab was recently reviewed by NICE for rheumatoid arthritis in TA 198 within which there was clearly some debate throughout over the assessment up to the FAD about the appropriate administration cost to apply. The ERG noted the availability of tariffs and reference costs for Health Research Group (HRG) codes HD23A to HD23C for Inflammatory Spine, Joint or Connective Tissue Disorders. Refer to Table 51 and Table 52.

Table 51: 2011 NHS Tariffs

Combined Daycase / Elective tariff		
HD23A	with Major CC	£1,730
HD23B	with CC	£595
HD23C	without CC	£471

Table 52: 2009 - 10 reference costs

Daycase		
HD23A	with Major CC	£ 455
HD23B	with CC	£ 412
HD23C	without CC	£ 432

Within the manufacturer's assessment of these costs, the £432 day case reference cost is adjusted pro rata by the number of hours required to yield an administration cost of £115 (£432 * 7.5/2.0). The manufacturer argues that adopting £126 per administration is as a result conservative. Adjusting the reference cost in this manner may be questionable.

Within TA198 the ERG noted that "*The administration cost of each infusion of tocilizumab was assumed to be £142. This originated from the first version of the Birmingham Rheumatoid Arthritis Model (BRAM) model calculations using 0.5 day case administration cost from the 2001 version of the PSSRU Unit Costs of Health and Social Care. That*

administration cost was calculated to be £124… This has then been inflated from 2004 to 2008 to get £142 which, according to the submission, has since been used in a couple of STAs including the Abatacept appraisal (TA141)." Further correcting this for inflation led to a final administration cost of £154 with the Final Appraisal Determination (FAD) indicating that while there was uncertainty around this cost £154 was acceptable, with this flowing through to the costing template for tocilizumab.

Since the day case cost taken from the PSSRU is relatively dated, there is an argument for adopting a similar approach but taking half of the current reference cost of £432 to yield £216 rather than the 7.5/2 suggested by the manufacturer. The most stringent approach would be to apply the full day case cost of £432. Applying these costs[h] results in the following cost effectiveness estimates. Refer to Table 53.

Table 53: Belimumab administration cost sensitivity analyses – Target population

		Without PAS		With PAS	
	SoC	Belimumab	Net	Belimumab	Net
QALYs	9.81	10.61	0.81	10.61	0.81
Admin cost @ £126	..	£9,059	£9,059	£9,059	£9,059
Total cost					
Admin @ £115	£105,366	£156,500	£51,134	[redacted]	[redacted]
Admin @ £126	£105,366	£157,291	£51,925	[redacted]	[redacted]
Admin @ £154	£105,366	£159,304	£53,938	[redacted]	[redacted]
Admin @ £216	£105,366	£163,761	£58,395	[redacted]	[redacted]
Admin @ £432	£105,366	£179,290	£73,924	[redacted]	[redacted]
ICERs					
Admin @ £115			£63,429		[redacted]
Admin @ £126			£64,410		[redacted]
Admin @ £154			£66,907		[redacted]
Admin @ £216			£72,436		[redacted]
Admin @ £432			£91,699		[redacted]

Patient age at baseline

Holding all other variables constant and setting the patient age at baseline to be 30, 40 and 50 results in cost effectiveness estimates of £65,498 per QALY, £62,695 per QALY and £55,439 per QALY respectively; i.e. for otherwise identical patients the cost effectiveness of belimumab improves as the age at first administration increases.

The manufacturer response to ERG clarification question B2 confirmed that these estimates for age 30 and age 50 are correct. The rationale provided by the manufacturer for this initially counter intuitive result lies in the calculated AMS at baseline for a 30 year old and a 50 year old being the same within the hypothetical example. The AMS has the same additive effect within the $\lambda_p = X'\beta$ for the 30 year old as for the 50 year old, since by assumption all patient

[h] Due to time constraints these estimates are derived by applying a multiplier to the total administration costs estimated within the base case run for the Target population. For instance, for the £432 day case cost the multiplier applied is £432/£126 = 3.43.

variables within X'β are the same for the 30 year old and the 50 year old other than the AMS. This rolls through to the patient hazard of death relative to the JHU cohort "average". It seems that it is at this point that the modelling of the 30 year old and the 50 year old diverge, with the age specific Bernatsky SLE SMRs being applied in conjunction with the age specific general population mortality rate.

In the light of what still appears to be counterintuitive result around baseline age, this may argue for a review of the modelling of mortality and the application of the Bernatsky SLE SMRs to the JHU cohort survival model.

Linear regression of SS_{52}-SS_0 central parameter estimates

The base case uses the pooled trial data to estimate the linear regression of SS_{52}-SS_0. As already outlined, the parameter estimates for this regression differ quite considerably between the two trials. All parameter estimates have p values of less than 0.1%.

From Table 54, for BLISS-52 the overall effect for belimumab week 24 responders is a central parameter estimate of -0.6372, as compared to -0.5980 for BLISS-76. But this has to be read in conjunction with the estimates for SoC which are also larger in BLISS-52 than for BLISS-76. The net difference between the overall effect of belimumab week 24 responders over that of SoC is -0.2743 for BLISS-52 and -0.2639 for BLISS-76 which given their similarity as naturally close to that estimated from the pooled data: -0.2742.

Table 54: Linear regression of SS_{52}-SS_0 central parameter estimates – Target population

	BLISS-52	BLISS-76	Pooled
SS_{52}-SS_0 SoC	-0.3629	-0.3341	-0.3493
SS_{52}-SS_0 belimumab	-0.3746	-0.3153	-0.3435
SS_{52}-SS_0 belimumab week 24 responders	-0.2626	-0.2827	-0.2800

For the estimated net SS change among belimumab week 24 responders over that of SoC to be so similar between the trials is surprising given the absolute mean changes as previously reported and repeated in Table 55. There would appear to be a larger absolute advantage within BLISS-52 compared to that within BLISS-76: 3.4 (-7.5 vs. -4.1) compared to 2.8 (-6.8 vs. -4.0).

Table 55: SS changes at week 52 by week 24 responder status and by trial – Target population

	BLISS-52			BLISS-76			Pooled		
Week 24 responder	Resp	NResp	All	Resp	NResp	All	Resp	NResp	All
Mean change from baseline at week 52: SS_{52}-SS_0									
SoC	-5.4	-2.3	**-4.1**	-6.1	-2.3	**-4.0**	-5.7	-2.3	-4.1
Belimumab	**-7.5**	-3.3	-6.3	**-6.8**	-2.5	-5.2	-7.2	-2.9	-5.8

Applying the estimates from the BLISS-52 and BLISS-76 results in economic estimates (see Table 56).

Table 56: Effect upon economic estimates of SS_{52}-SS_0 source: Target population

		Without PAS		With PAS	
	SoC	Belimumab	Net	Belimumab	Net
BLISS-52 as source for SS_{52}-SS_0 regression					
Total cost	£105,195	£157,102	£51,907	[redacted]	[redacted]
QALYs	9.84	10.64	0.80	10.64	0.80
ICER			£64,950		[redacted]
BLISS-76 as source for SS_{52}-SS_0 regression					
Total cost	£105,518	£157,469	£51,951	[redacted]	[redacted]
QALYs	9.77	10.55	0.78	10.55	0.78
ICER			£66,318		[redacted]

As anticipated, the cost effectiveness estimates are not particularly different between the application of the pooled regression coefficients, the BLISS-52 regression coefficients and the BLISS-76 regression coefficients.

Steroid dose use equation

The trial evidence for a steroid dose effect between the arms may be open to question, particularly within BLISS-76. This effect can be removed through a sensitivity analysis that slightly arbitrarily applies constant steroid dose of 10mg/day for all patients[i], together with

[i] Implemented within the *PSA Inputs* worksheet by setting HY7=0 and HZ7=10 for 10mg and 8 for 8mg, and HY7=0. 6799 and HZ7=3.197 for the post 1 Jan 2000 JH data

another that reduces it to 8mg/day. But note that this not only equalises the steroid dose between the arms but also equalises it between patients of differing SS score severity at baseline which is likely to be unrealistic. Unfortunately the electronic model is not easily amended to permit different steroid dosing based upon the individual patient baseline SS score undifferentiated by arm.

Retaining the differentiation of steroid use by arm, in response to ERG clarification question B21 the manufacturer re-estimates the steroid dose equation with a dummy for data that was pre-2000. The dummy was statistically significant and of the anticipated sign at 1.433, with the regression constant and coefficient for the SS score falling to 0.6799. These values can be used for a third sensitivity analysis. Refer to Table 57.

Table 57: Steroid doses: Target population

		Without PAS		With PAS	
	SoC	Belimumab	Net	Belimumab	Net
Constant steroid dose of 10mg/day					
Total cost	£103,261	£154,453	£51,192	[redacted]	[redacted]
QALYs	9.64	10.38	0.74	10.38	0.74
ICER			£68,766		[redacted]
Constant steroid dose of 8mg/day					
Total cost	£104,816	£156,561	£51,745	[redacted]	[redacted]
QALYs	9.79	10.55	0.76	10.55	0.76
ICER			£68,278		[redacted]
Post 1 January 2000 JHU data regression					
Total cost	£105,692	£157,877	£52,186	[redacted]	[redacted]
QALYs	9.84	10.65	0.81	10.65	0.81
ICER			£64,369		[redacted]

Arbitrarily equalising the steroid dose between the arms of the model to a constant 10mg/day or 8/mg per day does affect the overall patient experience and cost, but the net effect of the 10mg/day and the 8/mg is similar. For both, the net costs show limited change from the base case but the net benefits fall away slightly faster, resulting in reasonably similar cost effectiveness estimates of around £68,500 per QALY without the PAS [redacted]. The revised post 1 January 2000 regression has no practical

impact upon the results of the model. This may suggest that the net outcomes of the model are not particularly driven by the level of steroid dose, but differentiated by arm it has some impact with this mainly affecting the QALY side of the cost effectiveness equation.

Modelling of mortality and application of SMR

The requirement to apply the SMRs drawn from the Bernatsky reference within the modelling remains unclear to the ERG, particularly for when the patient being modelled is within the age range of the JHU cohort. But if it is reasonable to apply SLE SMRs within the mortality modelling, there is an additional concern over whether the SMRs from the Bernatsky reference are representative. [redacted]

Table 58: SMRs for sensitivity analysis

Age	Base case	Sens. analysis
16 – 24	19.2	5.3
25 – 40	8.0	3.7
40 – 59	3.7	2.6
60+	1.4	1.4

The SMRs reported in Table 58 result in the following model outputs (Table 59).

Table 59: Sensitivity analysis around SMRs – Target population

		Without PAS		With PAS	
	SoC	Belimumab	Net	Belimumab	Net
Survival LY - undiscounted	35.00	37.58	2.58	37.58	2.58
Total cost - discounted	£116,657	£168,095	£51,438	[redacted]	[redacted]
QALYs - discounted	10.53	11.26	0.726	11.26	0.73
ICER			£70,860		[redacted]

[redacted]

Given the apparently quite large undiscounted additional survival of 2.93 years within the Target population, the contribution of the various coefficients within the JHU cohort survival function of Table 6.12 can be explored. This is most simply achieved by sequentially setting each of the individual coefficients within Table 6.12 to equal zero with the remainder taking the values within Table 6.12 and running the model[j]. It is recognised that , as with the stepwise elimination of coefficients that led to the JHU cohort survival model, further eliminating coefficients within the JHU cohort survival model would if correctly undertaken change the values of the remaining coefficients.

Ordering results by their impacts upon net undiscounted survival results in the following estimates compared to the 2.93 additional life years of the Target population base case: AMS 0.90 life years; renal 2.59 life years; CAPD 2.68 life years; PVD 2.70 life years; age at diagnosis 3.14 life years; diabetes 3.06 life years; gastrointestinal 2.99 life years; malignancy 2.98 life years; and, infection 2.90 life years. Musculoskeletal and cholesterol have no impact upon the net undiscounted survival. Refer to Table 60.

[j] Implemented within the Nat Hist Model Data worksheet by setting the relevant cell within cells AG11:AG52 equal to zero

Table 60: Removing the AMS coefficient from JHU cohort survival function - Target population

		Without PAS		With PAS	
	SoC	Belimumab	Net	Belimumab	Net
Survival LY - undiscounted	20.20	21.10	0.90	21.10	0.90
Total cost - discounted	£127,598	£174,022	£46,424	[redacted]	[redacted]
QALYs - discounted	11.12	11.55	0.43	11.55	0.43
ICER			£106,912		[redacted]

The data reported in Table 60 underlines the importance of the AMS coefficient within the JHU cohort survival function to the anticipated additional 2.93 life years from belimumab use within the Target population. This should be read in conjunction with the concerns around the calculation of the SS score and the resultant calculation of the AMS score. It also highlights the possible significance of applying the Bernatsky SMRs to the patient mortality hazard as drawn from the JHU cohort survival function.

SLICC organ involvement at baseline

The model through random drawings simulates a range of organ involvements at baseline within the 50,000 patient simulated. The central estimate of cost effectiveness average across these. Given this it is illustrative to explore the scenarios of: no organ involvement at baseline; all organs having a SLICC score of 1 at baseline; and, individual organs having a SLICC score of 1 at baseline with no other organ involvement[k]. The net effects reported below relate to the addition of belimumab to SoC. Refer to Table 61.

[k] Implemented within the *Subgroup BLISS data* worksheet by setting cells P245:P256 equal to cell Q64 and cells Q245:T256 equal to 0, and for any organ involvement at SLICC score 1 setting the relevant cell(s) within cells Q245:Q256 equal to cell Q64 with the corresponding cells within P245:P256 equal to 0

Table 61: SLICC involvement at baseline – Target population

SLICC = 1 involvement	None	All	CV	Diabetes*	GI*	Malign*	MSK*
SoC undiscounted LYs	33.45	1.40	33.00	26.05	28.53	22.04	29.84
Net undiscounted LYs	2.82	0.32	3.05	3.13	3.19	3.35	3.13
Net disc. QALYs	0.84	0.13	0.67	0.88	0.86	0.95	0.75
Net disc. Costs ex PAS	£51,018	£16,067	£51,846	£53,135	£51,130	£49,048	£53,094
ICER ex PAS	£60,486	£122,796	£77,635	£60,240	£59,583	£51,759	£71,048
ICER with PAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SLICC = 1 involvement	NP	Ocular	PV*	GF	Pulm	Renal*	Skin
SoC undiscounted LYs	33.05	33.02	23.11	32,87	33.06	26.97	32.86
Net undiscounted LYs	3.08	2.91	3.05	3.00	2.97	2.89	3.11
Net disc. QALYs	0.63	0.83	0.88	0.84	0.61	0.81	0.82
Net disc. Costs ex PAS	£53,303	£51,624	£50,449	£52,530	£65,233	£58,222	£51,033
ICER ex PAS	£84,963	£62,420	£57,486	£62,206	£107,729	£71,932	£61,875
ICER with PAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
* Within the JHU cohort survival function of Table 6.12 of the submission							
CV – cardiovascular; GI – gastrointestinal; Malign – malignancy; MSK – musculoskeletal; NP – neuropsychatric; PV – peripheral vascular; GF – gonodal failure; Pulm – pulmonary							

Table 61 illustrates that of the organs not entering the JHU cohort survival function, assuming their individual involvement at baseline with no other organ involvement at baseline has a similar effect upon the anticipated patient survival as there being no organ involvement at all at baseline: an average survival in the SoC arm of a little over 33 years. Belimumab is anticipated to provide around an additional 3 life years.

The impact upon net QALYs is more marked. But it must be borne in mind that within the model multiple organ involvement only sees the HRQoL multiplier for the worst organ involved being applied. Neuropsychiatric involvement or pulmonary involvement are the worst, with HRQoL multipliers of 0.71 and 0.69 respectively, and their involvement from baseline effectively limits QALY gains to those arising from additional survival.

Cardiovascular and musculoskeletal disease with HRQoL multipliers of 0.76 and 0.79 respectively also have this effect but to a lesser extent, as their involvement at baseline leaves open the possibility of the subsequent development of neuropsychiatric involvement and/or pulmonary involvement.

Of those organs entering the JHU cohort survival function, assuming the individual involvement at baseline of malignany or peripheral vascular has the largest impact upon anticipated survival in the SoC arm, the anticipated additional survival from belimumab remains fairly constant at around 3 life years.

Given the individual impacts of organ involvement and the JHU cohort survival function, it may be slightly surprising for the scenario of all organ systems being involved at baseline to result in an average survival within the SoC arm of only 1.40 years. It is only in this admittedly extreme scenario that the anticipated additional survival from belimumab drops noticeably below 3 life years.

Patients may differ at baseline in terms of their organ involvement. For organs within the JHU cohort survival function this is modelled as affecting their anticipated survival under SoC. But almost regardless of their anticipated survival under SoC, adding belimumab to SoC appears to be modelled as yielding a fairly constant additional 3 years survival. This may again highlight the centrality of the modelling of the impact of belimumab on the SS score, and by implication the AMS score, upon model outcomes.

Pulmonary involvement costs and HRQoL

The costs of pulmonary involvement are based upon 90% of patients requiring average direct drug costs of £1571 per month plus 100% of patients requiring £316 other resource use to give a total monthly cost of £1,730. Within this the direct drug cost if only sildenafil was used would be somewhat less at only £348. In response to an ERG clarification question the manufacturer has run two additional sensitivity analyses: one applying the costs from sildenafil and the other excluding all pulmonary costs. These result in cost effectiveness estimates for the Target population of £66,807 per QALY and £68,182 per QALY respectively.

The HRQoL impacts are mainly sourced from the same HTA monograph that examines pulmonary arterial hypertension, these relating to the pulmonary arterial hypertension functional classes II, III and IV. An additional HRQoL value for functional class I is drawn from the Zisman pulmonary fibrosis paper, but as this is assumed to only apply to 1% of

pulmonary arterial hypertension patients it has no impact upon the calculations. This leads to the modelling applying an HRQoL multiplier for pulmonary arterial hypertension involvement of 0.61 as shown in Table 62.

Table 62: Pulmonary arterial hypertension average HRQoL

PAH functional class	HRQoL	% patients
I	0.73	1%
II	0.67	24%
III	0.60	63%
IV	0.52	12%
Weighted Average	0.61	

As already outlined this can be applied within the overall pulmonary HRQoL calculation as shown in Table 63.

Table 63: HRQoL calculation pulmonary involvement from Table 16.19

	HRQoL	JHU %	Weighted	JHU SLICC	Final
Pulmonary hypertension	0.61	33%	0.20		
Pulmonary fibrosis	0.73	42%	0.31		
Shrinking lung (Chest XRay)	1.00	2%	0.02		
Pleural fibrosis (Chest XRay)	1.00	20%	0.20		
Pulmonary infarction/resection	0.94	4%	0.04		
Average across pulmonary			0.77	1.31	0.70

Sensitivity analyses around this parameter do not appear to have been conducted. To explore its impact upon model outputs it can in effect be removed from the modelling, due to the values being treated as multiplicative by setting it equal to 1.00[l]. This results in a central estimate of £65,812 per QALY suggesting that results are not particularly sensitive to this variable.

[l] Implemented within the *QoL Inputs* worksheet by setting cells AG9:AG58=1

5 DISCUSSION

5.1.1 Clinical Effectiveness

Across many outcomes whilst the pooled data appear promising, the effect size for patients in BLISS-52 favoured belimumab to a much greater extent than those in BLISS-76, this applied for both the whole and the Target (high disease activity) populations. The effect sizes in favour of belimumab for the whole population in BLISS-76 were modest and for the most part showed no significant difference between belimumab and placebo groups. BLISS-76 is likely to be more representative of the proposed patient population in England and Wales.

Drawing on FDA data there appeared to be little difference in effectiveness between 1mg/kg and 10mg/kg dose regimens. The reasons for, and implications of, the differences between trials and a lack of dose response between doses for BLISS-76, are worthy of discussion but on available evidence cannot be resolved. The reason BLISS-76 patients were relatively unresponsive to belimumab is unlikely to be attributable to recruitment of patients with inactive disease because all were auto-immune positive at entry and had a SLEDAI score ≥ 6 points.

BLISS-76 patients had longer established disease, had more developed organ damage, were older, and were receiving less steroid dosage than those in BLISS-52, and these differences may have contributed to differing responses to therapy. The most obvious differences between trials were in the geographical distribution of study centres and in the racial make-up of the populations. These might be reflected in differences in response to therapy and in the nature of standard of care practices. Ninety two percent of BLISS-76 study centres, but none of the BLISS-52 centres, were located in North America + West Europe. The LBSL02 phase II RCT (100% of the trial centres located in USA + Canada) preceded the BLISS trials and failed to demonstrate effectiveness of belimumab (primary outcomes: percent mean change in SLEDAI score at week 24, and median time to first flare). However, post hoc analysis of LBSL02 data did identify a subgroup of patients (~ 70% of the total) who responded better and who exhibited auto-immune positive disease at trial entry. This population became the focus of the subsequent Phase III BLISS trials. The failure of the LBSL02 trial to show an effect was attributed to the inclusion of inappropriate patients lacking auto-antibodies at recruitment.

Belimumab is an expensive drug and the proposal is that it should be administered at monthly intervals at a dose of 10mg/kg. Because of some doubt regarding the relative effectiveness of differing dose regimens it is possible that in practice belimumab may be used at lower than

10mg/kg; it seems important that if this should happen that good data on effectiveness of reduced dose regimens should be collected.

Target population and proposed licence population

The focus of the MS was the high disease activity "Target population" which represents a subgroup of the proposed "licence population" (in turn a subpopulation of the pooled BLISS population). The primary end point, which was the percentage of responders at week 52 according to the novel composite SRI outcome measure, was very similar for pooled Target population and pooled "licence population" with respectively 19.8% and 24.8% extra responders for belimumab compared to placebo (belimumab vs. placebo odds ratio = 2.7 for both populations). Furthermore the cost-effectiveness of belimumab in each population was essentially the same (base case ICER £64,410 and £66,170 / QALY respectively). Given these results, there appear small grounds on which to distinguish patients in the Target population from those in the proposed licence population on the basis of either clinical or cost-effectiveness and a SLEDAI score cut-off of 10 points, appears to be an arbitrary criterion that would be difficult to implement in practice. One effect of selecting the Target population in preference to the "licence population" is to considerably reduce the manufacturer's calculation of total budget impact of introducing belimumab across the country. (MS section 7).

Belimumab vs. rituximab

No head-to-head trial comparing belimumab with rituximab has been conducted. The ERG and the manufacturer disagree about the commonality of outcome measures available from belimumab and rituximab trials, but concur that a credible indirect comparison is not feasible on the grounds of large difference between trial populations. The ERG note that the primary outcome measure in the relevant Rituximab trial may be a more stringent test of therapeutic effect than that used in the BLISS trials, and therefore are not convinced by the manufacturer's implication that belimumab is necessarily a more effective drug.

Efficacy of belimumab for different SLE manifestations

In the BLISS trials the most commonly involved SLE manifestations were musculoskeletal (60%), mucocutaneous (59%), hematologic (16%), general (11%), renal (11%) and vasculitis (7%). Direct evidence for a beneficial effect of belimumab on other manifestations, such as pulmonary, renal or central nervous system manifestations, is not available.

5.1.2 Cost Effectiveness

The manufacturer presents a complex and impressive natural history model of the evolution of SLE. The visual basic modelling as far as has been assessed to date by the ERG is sophisticated and appears correct. There appear to be some data input discrepancies between the written submission and the electronic model.

It is also unclear from an economic point of view why the manufacturer niches belimumab to those with a baseline SLEDAI score of at least 10. The cost effectiveness estimates for the anticipated license population and the Target population are very similar. Over time it also seems possible that those within the anticipated license population may fall within the Target population.

The base case estimates for a patient falling within the Target population are that belimumab will:

- Increase undiscounted survival by 2.93 years
- Increase discounted patient benefits by 0.81 QALYs
- Increase discounted costs by £51,925 [redacted]
- Cost £64,410 per QALY [redacted]

There are a number of ERG concerns with the modelling of the submission. If these concerns are justified, addressing them appears more likely to worsen the estimate of the cost effectiveness of belimumab than improve it.

- Assuming that belimumab week 24 non-responders will experience the average SS score within the SoC arm seems likely to have over-estimated the average impact upon SS scores within the belimumab arm. The SS scores drive the analysis and any error in their calculation is likely to have a major impact on results
- Not taking into account a patient's history before entry into the trial may further exaggerate the impact upon the AMS of belimumab compared to SoC
- The steroid use data within the trials has been passed over within the modelling.
- The calculation of the cumulative average steroid dose may be subject to a bias similar to that of the calculation of the AMS
- Maintaining the net gain in SS score for a belimumab week 24 responder compared to the parallel patient in the SoC arm while the belimumab week 24 responder remains on treatment may be optimistic
- There is some lack of clarity around the reasons for patients' discontinuation and the derivation of the 8% annual discontinuation rate among belimumab week 24

responders, and of the reasonableness of extrapolating using this value. A low discontinuation rate worsens the cost effectiveness of belimumab

- The requirement to adjust the JHU cohort survival model by SMRs from the literature is unclear and may have tended to exaggerate the impact of the individual covariates within the JHU cohort survival model
- The analysis of the observational cost data on a six monthly basis in order to relate it to the maximum SS score during that period then doubling it to arrive at the annual relationship appears peculiar given that the observational cost data was collected over a year. It may also lead to bias
- The separate estimation of a cost per organ involved may have double counted costs estimated within the SS score cost function to some degree
- There appear to be some discrepancies in the reported model outputs for the average durations of organ involvement, the annual costs of these and the discounted total costs of these organ involvements. There are as a consequence concerns around the calculation of the cost offsets from reduced organ involvement arising from belimumab

5.2 Implications for research

It is unlikely that an industry sponsored trial will be conducted to compare belimumab with rituximab or other new biological interventions for SLE. The cost of a sufficiently powered study to discriminate between such treatments is likely to be too great for such studies to be undertaken independently of industry sponsorship. In view of the relative expense of belimumab and the lack of clear demonstration of a dose response relationship it is possible that in the real world belimumab may be employed at doses less than 10mg/kg. Useful research could be undertaken to monitor such usage and the 24 week response rates elicited.

Due to the paucity of long-term evidence for the continued benefit of belimumab and its safety, monitoring and surveillance of patients who have been treated with belimumab are therefore necessary. Further investigation is needed in patients excluded in the current BLISS-52 and BLISS-76 trials who had severe lupus nephritis or central nervous system manifestations of the disease. The two trials were limited in the inclusions of black patients, who for example account for approximately 25% of lupus patients in the USA. These patients also tend to have more severe disease than the general lupus population. In an earlier Phase II study of belimumab, black patients did significantly better than non-black patients. In contrast the reported Phase III trials found black patients treated with belimumab performed worse than those given placebo. These discrepancies needed to be considered further.

Although BLyS (B-Lymphocyte stimulator) is raised in SLE, reducing its activity with belimumab in SLE patients appears to have only a very modest effect. In RCTs a large proportion of patients in the belimumab group responded, but the placebo group response indicated that many would have responded irrespective of receiving belimumab. In a Targeted population with higher response rates the effect of belimumab remained relatively modest. On this basis, research should be directed at identifying additional factors that independently, or together with BLyS play a role in the pathology of SLE. Until such factors are identified it is probable that the traditional armamentarium of interventions will remain core for the treatment of most SLE patients.

6 REFERENCE LIST

1. ACR Ad Hoc Committee on Systemic Lupus Erythematosus Response Criteria. The American College of Rheumatology response criteria for systemic lupus erythematosus clinical trials: Measures of overall disease activity. *Arthritis and Rheumatism* 2004;50:3418-26.

2. Isenberg DA, Gordon C. From BILAG to BLIPS--disease activity assessment in lupus past, present and future. *Lupus* 2000;9:651-4.

3. U.S.Food and Drugs Administration. *FDA briefing package.* 2010. URL: http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/arthritisdrugsadvisorycommittee/ucm233579.pdf (accessed 23 June 2011).

4. Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. *J Rheumatol* 2005;32:811-9.

5. U.S.Food and Drugs Administration. *Arthritis Advisory Committee Meeting Briefing Document for the 16 November 2010 Meeting: Belimumab: BENLYSTA® (belimumab): Treatment of Systemic Lupus Erythematosus: BLA 125370.* 13/10/2010. URL: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisDrugsAdvisoryCommittee/UCM233581.pdf (accessed 23 June 2011).

6. Wallace DJ, Stohl W, Furie RA, Lisse JR, McKay JD, Merrill JT *et al*. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. *Arthritis & Rheumatism* 2009;61:1168-78.

7. Navarra SV, Guzman RM, Gallacher AE, Hall S, Levy RA, Jimenez RE *et al*. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. *Lancet* 2011;377:721-31.

8. van Dis I, Kromhout D, Geleijnse JM, Boer JM, Verschuren WM. Evaluation of cardiovascular risk predicted by different SCORE equations: the Netherlands as an example. *Eur J Cardiovasc Prev Rehabil* 2010;17:244-9.

9. Swaak AJ, van den Brink HG, Smeenk RJ, Manger K, Kalden JR, Tosi S *et al*. Systemic lupus erythematosus: clinical features in patients with a disease duration of over 10 years, first evaluation. *Rheumatology (Oxford)* 1999;38:953-8.

10. Merrill, J. T., Neuwelt, C. M., Wallace, D. J., Shanahan, J. C., Latinis, K. M., Oates, J. C., Utset, T. O., Gordon, C., Isenberg, D. A., Hsieh, H. J., Zhang, D, and Brunetta, P. G. Efficacy and safety of rituximab in patients with moderately to severely active systemic lupus erythematosus: results from the randomized, double-blind phase II/III study EXPLORER. Conference: 72nd Annual Scientific Meeting of the American College of Rheumatology. Abstract L12. 2008. 2008 Oct 24;

11. Gordon C, Merrill JT, Wallace DJ, Hsieh H, Brunetta P. BILAG versus SELENA-SLEDAI scoring in systemic lupus erythematosus (SLE) patients in EXPLORER. *Arthritis and Rheumatism* 2009;60:281.

12. Scottish Intercollegiate Guideline Network. *Search Filters.* 21/04/2011. URL: http://www.sign.ac.uk/methodology/filters.html (accessed 23 June 2011).

13. Wallace DJ, Stohl W, Furie RA, Lisse JR, McKay JD, Merrill JT *et al*. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. *Arthritis & Rheumatism* 2009;61:1168-78.

14. Navarra SV, Guzman RM, Gallacher AE, Hall S, Levy RA, Jimenez RE *et al*. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. *Lancet* 2011;377:721-31.

15. Gladman DD, Urowitz MB, Kagal A, Hallett D. Accurately describing changes in disease activity in Systemic Lupus Erythematosus. *J Rheumatol* 2000;27:377-9.

16. Centre for Reviews and Dissemination. *NHS Economics Evaluation Database Handbook.* 2007. URL: www.york.ac.uk/inst/crd/pdf/nhseed-handb07.pdf (accessed 27 June 2011).

17. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35:1095-108.

18. Department of Health. *NHS reference costs 2005-06.* 07/12/2006. URL: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_062884 (accessed 9 April 2011).

19. Bernatsky S, Boivin JF, Joseph L, Manzi S, Ginzler E, Gladman DD *et al*. Mortality in systemic lupus erythematosus. *Arthritis and Rheumatism* 2006;54:2550-7.

20. Buyon JP, Petri MA, Kim MY, Kalunian KC, Grossman J, Hahn BH *et al*. The effect of combined estrogen and progesterone hormone replacement therapy on disease activity in systemic lupus erythematosus: a randomized trial. *Ann Intern Med* 2005;142:953-62.

21. Petri M, Buyon J, Kim M. Classification and definition of major flares in SLE clinical trials. *Lupus* 1999;8:685-91.

22. Petri M, Kim MY, Kalunian KC, Grossman J, Hahn BH, Sammaritano LR *et al*. Combined oral contraceptives in women with systemic lupus erythematosus. *N Engl J Med* 2005;353:2550-8.

7 APPENDICES

7.1 Appendix 1 SLE Flare index

The SLE Flare Index categorizes SLE flare as "mild or moderate" or "severe" based on 6 variables[20,21,22] (check that this is the correct Petri et al, 2005):

- Change in SELENA SLEDAI score from the most recent assessment to current.
- Change in signs or symptoms of disease activity.
- Change in prednisone dosage.
- Use of new medications for disease activity or hospitalization.
- Change in PGA score.
- Hospitalization for SLE activity (severe flare only).

Applied as follows (*Taken from HGS Briefing Document to the FDA Oct 2010*[5]):

SLE Flare Index

Mild to Moderate Flare	Severe Flare
Change (increase) in SELENA SLEDAI score of 3 points or more (but not more than 12)	Change (increase) in SELENA SLEDAI score to > 12
New/worse: Discoid, photosensitive, profundus, cutaneous vasculitis, bullous lupus Nasopharyngeal ulcers Pericarditis Arthritis Fever (SLE)	New/Worse: CNS-SLE Vasculitis Nephritis Myositis Plts < 60,000 Hemolytic anemia: Hb <70 g/L or decrease in Hb > 30 g/L **Requiring**: double prednisone, or prednisone increase to > 0.5 mg/kg/day, or hospitalization
Increase in prednisone, but not to > 0.5 mg/kg/day	Increase in prednisone to > 0.5 mg/kg/day
Added NSAID or hydroxychloroquine for SLE activity	New cyclophosphamide, azathioprine, methotrexate, or mycophenolate for SLE activity
≥ 1.0 increase in PGA score, but not to more than 2.5	Hospitalization for SLE activity Increase in PGA score to > 2.5.

7.2 Appendix 2 Assessment of manufacturer's search strategies

Appraised using PRESS CHECKLIST

Checklist developed by: Sampson M, McGowan J, Lefebvre C, Moher D, Grimshaw J. PRESS: Peer Review of Electronic Search Strategies. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2008. Available from: http://www.cadth.ca/media/pdf/477_PRESS-Peer-Review-Electronic-Search-Strategies_tr_Appendices.pdf

Search for non-randomised studies

1. Translation: Is the search question translated well into search concepts?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

2. Operators: Are there any mistakes in the use of Boolean or proximity operators?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

3. Subject headings: Are any important subject headings missing or have any irrelevant ones been included?

☑ Adequate

Needs revision Provide an explanation or example

_____

_____

4. Natural language: Are any natural language terms or spelling variants missing, or have any irrelevant ones been included? Is truncation used optimally?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

5. Spelling & syntax: Does the search strategy have any spelling mistakes, system syntax errors, or wrong line numbers?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

6. Limits: Do any of the limits used seem unwarranted or are any potentially helpful limits missing?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

Uses SIGN's Observational study filters for Embase and Medline)

7. Adapted for db: Has the search strategy been adapted for each database to be searched?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

Other notes:

Initial number in report is 14. Number in combined total for databases is 14.

Search doesn't include comparators, but section 5.1.1 of report implies that their plan was only to look for non-RCTs for belimumab.

Search for RCTs

1. Translation: Is the search question translated well into search concepts?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

2. Operators: Are there any mistakes in the use of Boolean or proximity operators?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

3. Subject headings: Are any important subject headings missing or have any irrelevant ones been included?

☑ Adequate

Needs revision Provide an explanation or example

_____

_____

4. Natural language: Are any natural language terms or spelling variants missing, or have any irrelevant ones been included? Is truncation used optimally?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

5. Spelling & syntax: Does the search strategy have any spelling mistakes, system

syntax errors, or wrong line numbers?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

6. Limits: Do any of the limits used seem unwarranted or are

any potentially helpful limits missing?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

Uses SIGN RCT filter for Embase, Sections of the search for Medline appear exactly the same as the SIGN RCT filter for Medline, but several lines are missing covering relevant publication types and other small differences are noted. SIGN filter may have been updated and the version used here is older? – new publication types)

7. Adapted for db: Has the search strategy been adapted for each database to be

searched?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

Other notes:

Initial number in flow diagram in section 5.1 of report is 3774. Number in combined total for databases in Appendix is 3776

In section 5.1 of the report it is stated that 39 full publications and 4 conference proceedings were included, but no details have been given for most of these (only 11 are mentioned in Tables 5.2, 5.3 and 5.4). Ideally we should see a list of all 43 publications in tabular form with clear reasons for exclusion.

Search for economic studies

1. Translation: Is the search question translated well into search

concepts?

Adequate

☑ Needs revision Provide an explanation or example

_____

The searches are not as well done as the searches for the clinical effectiveness section. In Pubmed, the use of title/abstract in the belimumab section of the search strategy has resulted in 7 fewer hits compared to the same sections in the clinical effectiveness search strategy. In Embase several lines include major mistakes resulting in it being unclear how the database would have performed the search. For example, line #1 starts with "exp AND" and includes two EMTREE headings that do not exist: 'lupus'/exp, 'sle'/exp. Testing this line of the search by entering it exactly as reported results in 23221 (06/05/11), which is far fewer than the 58059 reported in the search strategy. There are similar problems in the economic filter section of the search strategy.

The search strategies do not include comparators, but they state in report that they do not intend to search for these (is this reasonable?).

_____

2. Operators: Are there any mistakes in the use of Boolean or

proximity operators?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

3. Subject headings: Are any important subject headings missing or have any

irrelevant ones been included?

☐ Adequate

☑ Needs revision Provide an explanation or example

_____

The basic search in Pubmed for lupus was automatically mapped by Pubmed, resulting in the inclusion of the MeSH heading Lupus Vulgaris. The MeSH heading Lupus Erythematosus, Systemic was not included. However, because Pubmed also searched for lupus in all fields papers with this MeSH heading would have been picked up.

---

4. Natural language: Are any natural language terms or spelling variants missing, or

have any irrelevant ones been included? Is

truncation used optimally?

☑ Adequate

☐ Needs revision Provide an explanation or example

---

5. Spelling & syntax: Does the search strategy have any spelling mistakes, system

syntax errors, or wrong line numbers?

☑ Adequate

☐ Needs revision Provide an explanation or example

---

6. Limits: Do any of the limits used seem unwarranted or are

any potentially helpful limits missing?

Adequate

☑ Needs revision Provide an explanation or example

As the numbers found in the subject part of the search were so small for Medline (44 in Pubmed), the use of a filter was inappropriate. It is stated in section 6.1 of report that the CRD sensitive economics filters for Pubmed and Embase were used, but the version used in Pubmed does not match that given in CRD's NHS Economics Evaluation Database Handbook 2007. The versions of the filter that used to be on CRD's website are no longer there after the website restructure so it is possible that they have been updated. However, there are some clear discrepancies in the translation of some elements (i.e.the MeSH heading in the filter exp "Costs and Cost analysis"/ was entered as costs AND "cost analysis". Fortunately, this was translated by Pubmed correctly, but as well as including the correct MeSH heading several odd combinations of free text terms were searched, such as (costs"[All Fields] AND "cost"[All Fields] AND "analysis"[All Fields]). Many of the lines in Embase that should have been searching the EMTREE headings were entered

very differently and resulted in massively fewer hits (e.g. the EMTREE heading in the filter exp Economic Evaluation/ (which when tested brought back 166263 hits on 06/05/11) was entered as "exp AND economic AND ('evaluation'/exp OR evaluation)" and resulted in only 977 hits.

7. Adapted for db: Has the search strategy been adapted for each database to be

searched?

☑ Adequate

☐ Needs revision Provide an explanation or example

_____

_____

Other notes:

The initial number in the flow diagram tallies with those in the search strategies.

In section 5.1 of the report it is stated that the 14 excluded studies are listed in section 9.1, Appendix 10, but only 2 are actually listed in this Appendix.

7.3 Appendix 3 List of 43 publications from manufacturer's clinical study search

Inclusion / exclusion table from manufacturer's clarification document

Table A16.1 Summary of publications of RCTs reviewed and their reasons for exclusion

Publication	Reason for exclusion
1. Wallace DJ, Stohl W, Furie RA, Lisse JR, McKay JD, Merrill JT, *et al*. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. Arthritis Care and Research. 2009 15;61 (9):1168-78.	Included (LBSL02).
2. Furie RA, Petri MA, Wallace DJ, Ginzler EM, Merrill JT, Stohl W, *et al*. Novel evidence-based systemic lupus erythematosus responder index. Arthritis & Rheumatism. [Research Support, N.I.H., Extramural]. 2009 Sep 15;61(9):1143-51.	Included. Linked to LBSL02.
3. Furie R, Stohl W, Ginzler EM, Becker M, Mishra N, Chatham W, *et al*. Biologic activity and safety of belimumab, a neutralizing anti-B-lymphocyte stimulator (BLyS) monoclonal antibody: a phase I trial in patients with systemic lupus erythematosus. Arthritis research & therapy. 2008;10 (5):R109.	Included (LBSL01).
4. Navarra S, Ilianova E, Bae SC, Guzman R, *et al*. Belimumab, a BLYS-specific inhibitor, reduced disease activity, flares, and steroid use in patients with seropositive systemic lupus erythematosus (SLE): BLISS-52 study. EULAR. 2010:Abstract SAT0204.	Included (BLISS-52).
5. Tanasescu C, Gallacher A, Garcia M, Littlejohn G, Saaibi D, *et al*. Belimumab, a BLYS-specific inhibitor, significantly improved physical functioning, fatigue, and other health-related quality of life (HRQOL) measures in patients with seropositive systemic lupus erythematosus (SLE): BLISS-52 study. EULAR. 2010:abstract SAT0206.	Included (BLISS-52).
6. D'Cruz D, Tanasescu C, Navarra S, Guzman R, *et al*. Belimumab, a BLYS-specific inhibitor, reduced disease activity, flares and prednisone use in patients with active seropositive SLE: Phase 3 BLISS-52 study. BSR. 2010: abstract OP3.	Included (BLISS-52).
7. Furie R, Zamani O, Wallace D, Tegzova D, *et al*. Belimumab, a BLyS-Specific Inhibitor, Reduced Disease Activity and Severe Flares in Seropositive SLE Patients: BLISS-76 Study Results through Wk 76 ACR. 2010:Abstract 1454.	Included (BLISS-76).
8. Petri M, Van Vollenhoven RF, Zamani O, Furie RA, *et al*. Belimumab, a BLyS-Specific Inhibitor, Reduces Disease Activity and Severe Flares in Seropositive Systemic Lupus Erythematosus (SLE) Patients: BLISS-76 Study. International Journal of Rheumatic Diseases; Asia Pacific League of Associations of Rheumatology 2010;13:suppl. 1: 110-5, abstract 0281.	Included (BLISS-76).
9. Wallace DJ, Kalunian KC, Petri MA, Strand CV, *et al*. Epratuzumab Demonstrates Clinically Meaningful Improvements in Patients with Moderate to severe Systemic Lupus Erythematosus (SLE): Results from EMBLEM™, at Phase IIb Study ACR. 2010:Abstract 1452.	Investigational drug. Not yet available in the UK.
10. Carneiro JRM, Sato EI. Randomized double-blind clinical study with methotrexate in systemic lupus erythematosus. (Portuguese]. Revista Brasilliana de Reumamologia. 1999;39 (4):203-10.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
11. Islam N, Hossain M, Atiqul Haq S, Noor Alam M, *et al*. Efficacy and safety of methotrexate (MTX) in articular and cutaneous manifestations of systemic lupus erythematosus. EULAR. 2006:Abstract THU0273.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Focus on articular and cutaneous manifestations only.
12. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, *et al*. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: The randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis and Rheumatism. 2010 January;62 (1):222-33.	Included.
13. Andrade-Ortega L, Irazoque-Palazuelos F, Lopez-Villanueva R, Barragan-Navarro Y, Bourget-Pietrasanta F, Diaz-Ceballos MDLT, *et al*. Efficacy of rituximab versus cyclophosphamide in lupus patients with severe manifestations. A randomized and multicentre study. [Spanish]. Reumatologia Clinica. 2010 September;6 (5):250-5.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Excluded patients on other immunosuppressants (except antimalarials). Cyclophosphamide is not a relevant comparator.
14. Fortin PR, Abrahamowicz M, Ferland D, Lacaille D, Smith CD, Zummer M. Steroid-sparing effects of methotrexate in systemic lupus erythematosus: A double-blind, randomized, placebo-controlled trial. Arthritis Care and Research. 2008 15;59 (12):1796-804.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Excluded patients taking azathioprine.
15. Carneiro JRM, Sato EI. Double blind, randomized, placebo controlled clinical trial of methotrexate in systemic lupus erythematosus. Journal of Rheumatology. 1999;26 (6):1275-9.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
16. Barile-Fabris L, Ariza-Andraca R, Olguin-Ortega L, Jara LJ, Fraga-Mouret A, Miranda-Limon JM, *et al*. Controlled clinical trial of IV cyclophosphamide versus IV methylprednisolone in severe neurological manifestations in systemic lupus erythematosus. Annals of the Rheumatic Diseases. 2005 Apr;64 (4):620-5.	Included patients with severe neurological involvement. Cyclophosphamide is not a relevant comparator.
17. Fries JF, Sharp GC, McDevitt HO, Holman HR. Cyclophosphamide therapy in systemic lupus erythematosus and polymyositis. Arthritis and Rheumatism. 1973 1973;16 (2):154-62.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Included patients with polymyositis. Cyclophosphamide is not a relevant comparator.

18. Dussán KB, Magder L, Brodsky RA, Jones RJ, Petri M. High dose cyclophosphamide performs better than monthly dose cyclophosphamide in quality of life measures. Lupus. 2008(12):1079-85.	Cyclophosphamide is not a relevant comparator.
19. Gonzalez-Lopez L, Cardona-Munoz EG, Celis A, Garcia-De La Torre I, Orozco-Barocio G, Salazar-Paramo M, *et al*. Therapy with intermittent pulse cyclophosphamide for pulmonary hypertension associated with systemic lupus erythematosus. Lupus. 2004;13 (2):105-12.	Cyclophosphamide is not a relevant comparator. Included patients with CNS lupus and lupus nephritis.
20. Petri M, Brodsky RA, Jones RJ, Gladstone D, Fillius M, Magder LS. High-dose cyclophosphamide versus monthly intravenous cyclophosphamide for systemic lupus erythematosus a prospective randomized trial. Arthritis and Rheumatism. 2010 May;62 (5):1487-93.	Cyclophosphamide is not a relevant comparator. Included patients with CNS lupus and lupus nephritis.
21. Bykerk V, Sampalis J, Esdaile JM, Choquette D, Senecal JL, Danoff D, *et al*. A randomized study of the effect of withdrawing hydroxychloroquine sulfate in systemic lupus erythematosus. New England Journal of Medicine. 1991;324 (3):150-4.	Withdrawal study in patients with stable SLE.
22. Tsakonas E, Joseph L, Esdaile JM, Choquette D, Senecal JL, Cividino A, *et al*. A long-term study of hydroxychloroquine withdrawal on exacerbations in systemic lupus erythematosus. Lupus. 1998;7 (2):80-5.	Withdrawal study in patients with stable SLE.
23. Levy RA, Vilela VS, Cataldo MJ, Ramos RC, Duarte JLMB, Tura BR, *et al*. Hydroxychloroquine (HCQ) in lupus pregnancy: Double-blind and placebo-controlled study. Lupus. 2001;10 (6):401-4.	Study in pregnant patients.
24. Bezerra EL, Vilar MJ, da Trindade Neto PB, Sato EI. Double-blind, randomized, controlled clinical trial of clofazimine compared with chloroquine in patients with systemic lupus erythematosus. Arthritis and rheumatism. 2005(10):3073-8.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Clofazimine not available in the UK. Focus on cutaneous manifestations only.
25. Meinão IM, Sato EI, Andrade LE, Ferraz MB, Atra E. Controlled trial with chloroquine diphosphate in systemic lupus erythematosus. Lupus. 1996(3):237-41.	Chloroquine not available in the UK.
26. Danowski A, Magder L, Petri M. Flares in Lupus: Outcome Assessment Trial (FLOAT), a comparison between oral methylprednisolone and intramuscular triamcinolone. Journal of Rheumatology. 2006 January;33 (1):57-60.	Study in patients presenting with mild or moderate flare.
27. Dougados M, Job-Deslandre C, Amor B, Menkes CJ. Danazol therapy in systemic lupus erythematosus. A one-year prospective controlled trial on 40 female patients. Clinical Trials Journal. 1987;24 (2):191-200.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus. Danazol is not considered standard of care.
28. Bootsma H, Spronk P, Derksen R, De Boer G, Wolters-Dicke H, Hermans J, *et al*. Prevention of relapses in systemic lupus erythematosus. Lancet. 1995;345 (8965):1595-9.	Study designed to look at prevention of relapses in patients presenting with a rise in anti-dsDNA.
29. Edwards JC, Snaith ML, Isenberg DA. A double blind controlled trial of methylprednisolone infusions in systemic lupus erythematosus using individualised outcome assessment. Annals of the rheumatic diseases. 1987(10):773-6.	Study in patients with severe SLE presenting with an acute exacerbation.
30. Dammacco F, Della Casa Alberighi O, Ferraccioli G, Racanelli V, Casatta L, Bartoli E. Cyclosporine-A plus steroids versus steroids alone in the 12-month treatment of systemic lupus erythematosus. International Journal of Clinical and Laboratory Research. 2000;30 (2):67-73.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
31. Denburg SD, Carbotte RM, Denburg JA. Corticosteroids and neuropsychological functioning in patients with systemic lupus erythematosus. Arthritis and Rheumatism. 1994 Sep;37 (9):1311-20.	Study was designed to assess the effects of corticosteroids on nervous system functioning as well as disease-related symptoms in patients with mild SLE and mild neuropsychiatric symptoms.
32. Hahn BH, Kantor OS, Osterland CK. Azathioprine plus prednisone compared with prednisone alone in the treatment of systemic lupus erythematosus. Report of a prospective controlled trial in 24 patients. Annals of Internal Medicine. 1975 Nov;83(5):597-605.	Study in severe, life-threatening systemic lupus erythematosus.
33. Mackworth-Young CG, David J, Morgan SH, Hughes GR. A double blind, placebo controlled trial of intravenous methylprednisolone in systemic lupus erythematosus. Annals of the rheumatic diseases. 1988(6):496-502.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
34. Tseng CE, Buyon JP, Kim M, Belmont HM, Mackay M, Diamond B, *et al*. The effect of moderate-dose corticosteroids in preventing severe flares in patients with serologically active, but clinically stable, systemic lupus erythematosus: Findings of a prospective, randomized, double-blind, placebo-controlled trial. Arthritis and Rheumatism. 2006 Nov;54 (11):3623-32.	Included patients with inactive disease defined as a S EDAI score ≤ 4.
35. Mease PJ, Ginzler EM, Gluck OS, Schiff M, Goldman A, Greenwald M, *et al*. Effects of prasterone on bone mineral density in women with systemic lupus erythematosus receiving chronic glucocorticoid therapy. Journal of Rheumatology. 2005 Apr;32 (4):616-21.	Study designed to examine the effects of prasterone on bone mineral density in female patients with mild to moderate systemic lupus erythematosus.
36. Petri MA, Mease PJ, Merrill JT, Lahita RG, Iannini MJ, Yocum DE, *et al*. Effects of prasterone on disease activity and symptoms in women with active systemic lupus erythematosus: Results of a multicentre randomized, double-blind, placebo-controlled trial. Arthritis and Rheumatism. 2004 Sep;50 (9):2858-68.	Included patients with SLEDAI > 2.
37. Petri MA, Lahita RG, Van Vollenhoven RF, Merrill JT, Schiff M, Ginzler EM, *et al*. Effects of prasterone on corticosteroid requirements of women with systemic lupus erythematosus: A double-blind, randomized, placebo-controlled trial. Arthritis and Rheumatism. 2002;46 (7):1820-9.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
38. Sanchez-Guerrero J, Fragoso-Loyo HE, Neuwelt CM, Wallace DJ, Ginzler EM, Sherrer YRS, *et al*. Effects of prasterone on bone mineral density in women with active systemic lupus erythematosus receiving chronic glucocorticoid therapy. Journal of Rheumatology. 2008 August;35 (8):1567-75.	Study designed to examine the effects of prasterone on bone mineral density in female SLE patients.
39. Chang DM, *et al*. Dehydroepiandrosterone treatment of women with mild-to-moderate systemic lupus erythematosus. Arthritis & Rheumatism. 2002;46(11):2924-27.	Included patients with SLEDAI > 2.
40. Hartkamp A, Geenen R, Godaert GLR, Bijl M, Bijlsma JWJ, Derksen RHWM. Effects of dehydroepiandrosterone on fatigue and well-being in women with quiescent systemic lupus	No requirement for patients to have active autoantibody-positive systemic

erythematosus: A randomised controlled trial. Annals of the Rheumatic Diseases. 2010 June;69 (6):1144-7.	lupus erythematosus.
41. Nordmark G, Bengtsson C, Larsson A, Karlsson FA, Sturfelt G, Ronnblom L. Effects of dehydroepiandrosterone supplement on health-related quality of life in glucocorticoid treated female patients with systemic lupus erythematosus. Autoimmunity. 2005 Nov;38 (7):531-40.	No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
42. Van Vollenhoven RF, Engleman EG, McGuire JL. Dehydroepiandrosterone in systemic lupus erythematosus: Results of a double-blind, placebo-controlled, randomized clinical trial. Arthritis and Rheumatism. 1995 Dec;38 (12):1826-31.	Study in mild to moderate SLE. No requirement for patients to have active autoantibody-positive systemic lupus erythematosus.
43. Gordon C, Wallace DJ, Shinada S, Kalunian KC, Forbess L, Braunstein GD, *et al*. Testosterone patches in the management of patients with mild/moderate systemic lupus erythematosus. Rheumatology. 2008 Mar;47 (3):334-8.	Included patients with mild to moderate SLE defined by SELENA-SLEDAI ≥ 2.

7.4 Appendix 4 Demographic details for BLISS total and Target populations (MS Table 5.9)

	BLISS-52				BLISS-76				Pooled Total Population			
	Placebo N = 287	**1mg/kg N = 288**	**10mg/kg N = 290**	**All N = 865**	**Placebo N = 275**	**1mg/kg N = 271**	**10mg/kg N = 273**	**All N = 819**	**Placebo N = 562**	**1mg/kg N = 559**	**10mg/kg N = 563**	**All N = 1684**
SLE Disease duration (yr)[1]												
Mean ± SD	5.93 ± 6.17	4.96 ± 4.58	5.03 ± 5.07	5.31 ± 5.32	7.42 ± 6.72	7.93 ± 7.13	7.20 ± 7.45	7.52 ± 7.10	6.66 ± 6.48	6.40 ± 6.13	6.08 ± 6.42	6.38 ± 6.35
SELENA SLEDAI score												
≥ 10	158 (55.1%)	139 (48.3%)	160 (55.2%)	457 (52.8%)	141 (51.3%)	144 (53.1%)	136 (49.8%)	421 (51.4%)	299 (53.2%)	283 (50.6%)	296 (52.6%)	878 (52.1%)
Mean ± SD	9.70 ± 3.62	9.56 ± 3.78	9.97 ± 3.88	9.75 ± 3.76	9.80 ± 3.97	9.70 ± 3.65	9.52 ± 3.64	9.67 ± 3. 5	9.75 ± 3.79	9.63 ± 3.71	9.75 ± 3.77	9.71 ± 3.76
PGA score												
< 1	43 (15.0%)	38 (13.2%)	32 (11.0%)	113 (13.1%)	33 (12.0%)	39 (14.4%)	51 (18.7%)	123 (15.0%)	76 (13.5%)	77 (13.8%)	83 (14.7%)	236 (14.0%)
1 - < 2	195 (67.9%)	207 (71.9%)	212 (73.1%)	614 (71.0%)	196 (71.3%)	189 (69.7%)	175 (64.1%)	560 (68.4%)	391 (69.6%)	396 (70.8%)	387 (68.7%)	1174 (69.7%)
≥ 2	49 (17.1%)	43 (14.9%)	46 (15.9%)	138 (16.0%)	46 (16.7%)	43 (15.9%)	47 (17.2%)	136 (16.6%)	95 (16.9%)	86 (15.4%)	93 (16.5%)	274 (16.3%)
BILAG organ domain involvement												
at least 1A or 2B	166 (57.8%)	166 (57.6%)	172 (59.3%)	504 (58.3%)	187 8.0%)	173 (63.8%)	160 (58.6%)	520 (63.5%)	353 (62.8%)	339 (60.6%)	332 (59.0%)	1024 (60.8%)
at least 1A	52 (18.1%)	58 (20.1%)	54 (18.6%)	164 (19.0%)	37 (13.5%)	38 (14.0%)	24 (8.8%)	99 (12.1%)	89 (15.8%)	96 (17.2%)	78 (13.9%)	263 (15.6%)
SLICC Damage Index score (Mean ± SD)	0.55 ± 0.93	0.60 ± 1.06	0.55 ± 1.00	0.57 ± 1.00	0.99 ± 1.45	1.04 ± 1.39	0.94 ± 1.38	0.99 ± 1.41	0.77 ± 1.23	0.81 ± 1.25	0.74 ± 1.21	0.77 ± 1.23
SLICC Damage Index score = 0	182 (63.4%)	190 (66.0%)	193 (66.6%)	565 (65.3%)	145 (52.7%)	125 (46.1%)	145 (53.1%)	415 (50.7%)	327 (58.2%)	315 (56.4%)	338 (60.0%)	980 (58.2%)
SLICC Damage Index score = 1	70 (24.4%)	56 (19.4%)	60 (20.7%)	186 (21.5%)	66 (24.0%)	76 (28.0%)	62 (22.7%)	204 (24.9%)	136 (24.2%)	132 (23.6%)	122 (21.7%)	390 (23.2%)
SLICC Damage Index score ≥ 2	35 (12.2%)	42 (14.6%)	37 (12.8%)	114 (13.2%)	64 (23.3%)	69 (25.5%)	66 (24.2%)	199 (24.3%)	99 (17.6%)	111 (19.9%)	103 (18.3%)	313 (18.6%)
Proteinuria (g/24 hour)												
≥ 2	21 (7.3%)	26 (9.0%)	19 (6.6%)	66 (7.6%)	11 (4.0)	7 (2.6%)	15 (5.5%)	33 (4.0%)	32 (5.7%)	33 (5.9%)	34 (6.0%)	99 (5.9%)
Mean ± SD	0.62 ± 1.15	0.63 ± 1.13	0.54 ± 0.91	0.60 ± 1.07	0.39 ± 0.81	0.33 ± 0.65	0.40 ± 0.73	0.37 ± 0.74	0.50 ± 1.00	0.48 ± 0.94	0.48 ± 0.83	0.49 ± 0.93

[1] Time elapsed between date of SLE diagnosis and the date of informed consent.

BLISS whole population serological (MS Tables 5.10)

	BLISS-52				BLISS-76				Pooled Total Population			
	Placebo N = 287	**1mg/kg N = 288**	**10mg/kg N = 290**	**All N = 865**	**Placebo N = 275**	**1mg/kg N = 271**	**10mg/kg N = 273**	**All N = 819**	**Placebo N = 562**	**1mg/kg N = 559**	**10mg/kg N = 563**	**All N = 1684**
Anti-dsDNA positive (≥ 30 IU/mL)	205 (71.4)	221 (76.7%)	218 (75 2%)	644 (74.5%)	174 (63.3%)	171 (63.1%)	179 (65.6%)	524 (64.0%)	379 (67.4%)	392 (70.1%)	397 (70.5%)	1168 (69.4%)
Anti-Smith positive (≥ 15 U/mL)	101/287 (35.2%)	102/288 (35.4%)	105/287 (36.6%)	308/862 (35.7%)	72/269 (26.8%)	69/269 (25.7%)	75/265 (28.3%)	216/803 26.9%)	173/556 (31.1%)	171/557 (30.7%)	180/552 (32.6%)	524/1665 (31.5%)
IgG >ULN (16.18 g/L)	146 (50.9%	140 (48.6%)	151 (52.1%)	437 (50.5%)	108 (39.3%)	105 (38.7%)	94 (34.4%)	307 (37.5%)	254 (45.2%)	245 (43.8%)	245 (43.5%)	744 (44.2%)
Complement												
Normal/high C3 and C4	102 (35.5%)	100 (34.7%)	89 (30.7%)	291 (33.6%)	113 (41.1%)	122 (45.0%)	112 (41.0%)	347 (42.4%)	215 (38.3%)	222 (39.7%)	201 (35.7%)	638 (37.9%)
Low C3 or C4, but not both	78 (27.2%)	55 (19.1%)	75 (25.9%)	208 (24.0%)	65 (23.6%)	57 (21.0%)	60 (22.0%)	182 (22.2%)	143 (25.4%)	112 (20.0%)	135 (24.0%)	390 (23.2%)
Low C3 (< 900 mg/L)	132 (46.0%)	148 (51.4%)	147 (50.7%)	427 (49.4%)	116 (42.2%)	100 (36.9%)	115 (42.1%)	331 (40.4%)	248 (44.1%)	248 (44.4%)	262 (46.5%)	758 (45.0%)
Low C4 (< 16 mg/dL)	160 (55.7%)	173 (60.1%)	180 (62.1%)	513 (59.3%)	143 (52.0%)	141 (52.0%)	147 (53.8%)	431 (52.6%)	303 (53.9%)	314 (56.2%)	327 (58.1%)	944 (56.1%)
Low C3 and C4	107 (37.3%)	133 (46.2%)	126 (43.4%)	366 (42.3%)	97 (35.3%)	92 (33.9%)	101 (37.0%)	290 (35.4%)	204 (36.3%)	225 (40.3%)	227 (40.3%)	656 (39.0%)
BLyS (above LOQ, ≥ 0.5 ng/mL)	273/283 (96.5%)	273/285 (95.8%)	281/285 (98.6%)	827/853 (97.0%)	268/271 (98.9%)	267/270 (98.9%)	263/268 (98.1%)	798/809 (98.6%)	541/554 (97.7%)	540/555 (97 3%)	544/553 (98.4%)	1625/1662 (97.8%)

BLISS whole population concomitant medications (MS Tables 5.11)

	BLISS-52				BLISS-76				Pooled Total Population			
	Placebo N = 287	**1mg/kg N = 288**	**10mg/kg N = 290**	**All N = 865**	**Placebo N = 275**	**1mg/kg N = 271**	**10mg/kg N = 273**	**All N = 819**	**Placebo N = 562**	**1mg/kg N = 559**	**10mg/kg N = 563**	**All N = 1684**
Total corticosteroid use	276 (96.2%)	276 (95.8%)	278 (95.9%)	830 (96.0%)	212 (77.1%)	211 (77.9%)	200 (73.3%)	623 (76.1%)	488 (86.8%)	487 (87.1%)	478 (84.9%)	1453 (86.3%)
Prednisone or equivalent												
> 0 to ≤ 7.5 g/day	84 (29.3%)	72 (25.0%)	74 (25.5%)	230 (26.6%)	86 (31.3%)	81 (29.9%)	80 (29.3%)	247 (30.2%)	170 (30.2%)	153 (27.4%)	154 (27.4%)	477 (28.3%)
> 7.5 to < 2 mg/day	136 (47.4%)	133 (46.2%)	131 (45.2%)	400 (46.2%)	76 (27.6%)	96 (35.4%)	81 (9.7%)	253 (30.9%)	212 (37.7%)	229 (41.0%)	212 (37.7%)	653 (38.8%)
≥ 20 mg/day	56 (19.5%)	71 (24.7%)	73 (25.2%)	200 (23.1%)	50 (18.2%)	34 (12.5%)	39 (14.3%)	123 (15.0%)	106 (18.9%)	105 (18.8%)	112 (19.9%)	323 (19.2%)
Antimalarials	201 (70.0%)	195 (67.7%	185 (63.8%)	581 (67.2%)	180 (65.5%)	171 (63.1%)	168 (61.5%)	519 (63.4%)	381 (67.8%)	366 (65.5%)	353 (62.7%)	1100 (65.3%)
Other immunosuppressants	122 (42.5%)	120 (41.7%)	123 (42.4%)	365 (42.2%)	154 (56.0%)	153 (56.5%)	148 (54.2%)	455 (55.6%)	276 (49.1%)	273 (48.8%)	271 (48.1%)	820 (48.7%)
1 immunosupressant	111 (38.7%)	116 (40.3%)	118 (40.7%)	345 (39.8%)	140 (50.9%)	143 (52.8%)	140 (51.3%)	423 (51.6%)	251 (44.7%)	259 (46.3%)	258 (45.8%)	768 (45.6%)
2 immunosupressants	11 (3.8%)	4 (1.4%)	5 (1.7%)	20 (2.3%)	13 (4.7%)	10 (3.7%)	8 (2.9%)	31 (3.8%)	24 (4.3%)	14 (2.5%)	13 (2.3%)	51 (3.0%)
Azathioprine	67 (23.3%)	71 (24.7%)	84 (29.0%)	222 (25.7%)	57 (20.7%)	52 (19.2%)	58 (21.2%)	167 (20.4%)	124 (22.1%)	123 (22.0%)	142 (25.2%)	389 (23.1%)
Methotrexate	35 (12.2%)	24 (8.3%)	20 (6.9%)	79 (9.1%)	60 (21.8%)	53 (19.6%)	39 (14.3%)	152 (18.6%)	95 (16.9%)	77 (13.8%)	59 (10.5%)	231 (13.7%)
Mycophenolate	19 (6.6%)	16 (5.6%)	17 (5.9%)	52 (6.0%)	42 (15.3%)	45 (16.6%)	50 (18.3%)	137 (16.7%)	61 (10.9%)	61 (10.9%)	67 (11.9%)	189 (11.2%)
Cyclosporin	6 (2.1%)	5 (1.7%)	2 (0.7%)	13 (1.5%)	5 (1.8%)	4 (1.5%)	5 (1.8%)	14 (1.7%)	11 (2.0%)	9 (1.6%)	7 (1.2%)	27 (1.6%)
Leflunomide	2 (0.7%)	-	3 (1.0%)	5 (0.6%)	3 (1.1%)	7 (2.6%)	1 (0.4%)	11 (1.3%)	5 (0.9%)	7 (1.3%)	4 (0.7%)	16 (1.0%)
Cyclophosphamide	2 (0.7%)	3 (1.0%)	1 (0.3%)	6 (0.7%)	2 (0.7%)	2 (0.7%)	2 (0.7%)	6 (0.7%)	4 (0.7%)	5 (0.9%)	3 (0.5%)	12 (0.7%)
NSAIDs	59 (20.6%)	56 (19.4%)	58 (20.9%)	173 (20.0%)	119 (43.3%)	114 (42.1%)	101 (37.0%)	334 (40.8%)	178 (31.7%)	170 (30.4%)	159 (28.2%)	507 (30.1%)

Demographic characteristics of Target population (Table A3.1 clarification document)

		BLISS-52		BLISS-76		Combined BLISS	
		SoC	**10mg/kg**	**SoC**	**10mg/kg**	**SoC**	**10mg/kg**
	Total enrolled	**n=107**	**n=112**	**n=96**	**n=81**	**n=203**	**n=193**
Gender	Male	10 (9.3%)	3 (2.7%)	6 (6.3%)	4 (4.9%)	16 (7.9%)	7 (3.6%)
	Female	97 (90.7%)	109 (97.3%)	90 (93.8%)	77 (95.1%)	187 (92.1%)	186 (96.4%)
Race	Caucasian	29 (27.1%)	23 (20.5%)	61 (63.5%)	54 (66.7%)	90 (44.3%)	77 (39.9%)
	Asian	40 (37.4%)	53 (47.3%)	5 (5.2%)	4 (4.9%)	45 (22.2%)	57 (29.5%)
	Black/African American	1 (0.9%)	6 (5.4%)	13 (13.5%)	7 (8.6%)	14 (6.9%)	13 (6.7%)
	Alaskan/Native American	37 (34.6%)	30 (26.8%)	17 (17.7%)	16 (19.8%)	54 (26.6%)	46 (23.8%)
	Hawaiian/Pacific Islander	0	0	0	0	0	0
	Multiracial	1 (0.9%)	0	0	0	1 (0.5%)	0
	Hispanic origin	55 (51.4%)	46 (41.1%)	28 (29.2%)	25 (30.9%)	83 (40.9%)	71 (36.8%)
Age	Years						
	n ≤ 45 yrs	93 (86.9%)	100 (89.3%)	78 (81.3%)	65 (80.2%)	171 (84.2%)	165 (85.5%)
	n > 45 to < 65	14 (13.1%)	12 (10.7%)	17 (17.7%)	16 (19.8%)	31 (15.3%)	28 (14.5%)
	n ≥ 65 to < 75	0	0	1 (10%)	0	1 (0.5%)	0
Weight (kg)	Mean (SD)	62.1 ± 13.9	61.4 ± 14.1	688 ± 13.7	70.0 ± 16.7	65.2 ± 14.2	65.0 ± 15.8
	Range	34.7-127.6	39.5-128.5	45.4-108.6	47.0-131.7	34.7-127.6	39.5-131.7
Region & country	USA/Canada	0	0	45 (46.9%)	24 (29.6%)	45 (22.2%)	24 (12.4%)
	W Europe/Israel	0	0	24 (25.0%)	30 (37.0%)	24 (11.8%)	30 (15.5%)
	E Europe	10 (9.3%)	11 (9.8%)	12 (12.5%)	12 (14.8%)	22 (10.8%)	23 (11.9%)
	America excluding USA/Canada	56 (52.3%)	48 (42.9%)	15 (15.6%)	15 (18.5%)	71 (35.0%)	63 (32.6%)
	Asia	39 (36.4%)	53 (47.3%)	0	0	39 (19.2%)	53 (27.5%)
	Australia	2 (1.9%)	0	0	0	2 (1.0%)	0

BLISS Target population disease characteristics (Table A3.1 Clarification document)

		BLISS-52		BLISS-76		Combined BLISS	
		SoC	**10mg/kg**	**SoC**	**10mg/kg**	**SoC**	**10mg/kg**
	Total enrolled	**n=107**	**n=112**	**n=96**	**n=81**	**n=203**	**n=193**
SLE duration yrs; mean (SD)		6.70 ± 6.96	5.26 ± 4.99	7.42 ± 6.40	7.94 ± 7.47	7.04 ± 6.69	6.38 ± 6.28
BILAG organ involvement	At least 1A or 2B	65 (60.7%)	78 (69.6%)	78 (81.3%)	58 (71.6%)	143 (70.4%)	136 (70.5%)
	At least 1A	18 (16.8%)	25 (22.3%)	21 (21.9%)	7 (8.6%)	39 (19.2%)	32 (16.6%)
	At least 1A or 1B	99 (92.5%)	103 (92.0%)	94 (97.9%)	78 (96.3%)	193 (95.1%)	181 (93.8%)
	No A or B	8 (7.5%)	9 (8.0%)	2 (2.1%)	3 (3.7%)	10 (4.9%)	12 (6.2%)
SELENA-SLEDAI mean (SD)		12.6 ± 3.0	12.8 ± 3.6	13.0 ± 3.5	12.4 ± 2.9	12.8 ± 3.3	12.6 ± 3.3
PGA category	0 to 1	15 (14.0%)	13 (11.6%)	8 (8.3%)	8 (9.9%)	23 (11.3%)	21 (10.9%)
	>1 to 2.5	91 (85.0%)	97 (86.6%)	86 (89.6%)	71 (87.7%)	177 (87.2%)	168 (87.0%)
	>2.5 to 3	1 (0.9%)	2 (1.8%)	2 (2.1%)	2 (2.5%)	3 (1.5%)	4 (2.1%)
SLICC Damage index; mean (SD)		0.6 ± 1.0	0.5 ± 0.9	0.8 ± 1.4	0.8 ± 1.2	0.7 ± 1.2	0.6 ± 1.0
Prednisone or equivalent dose	0 mg/day	5 (4.7%)	4 (3.6%)	15 (15.6%)	12 (14.8%)	20 (9.9%)	16 (8.3%)
	>0 - ≤7.5 mg/day	26 (24.3%)	27 (24)	31 (32.3%)	24 (29.6%)	57 (28.1%)	51 (26.4%)
	> 7.5 mg/day	76 (71.0%)	81 (72.3%)	50 (52.1%)	45 (55.6%)	126 (62.1%)	126 (65.3%)
Average prednisone or equivalent dose; mean (SD) mg/day		12.8 ± 8.4	13.7 ± 10.4	0 3 8.8	10.4 ± 8.1	11.6 ± 8.6	12.3 ± 9.6

7.5 Appendix 5 Justification for pooling results across trials

The ERG points for clarification requested "Please provide further justification for pooling results for Target Population". The manufacturer's response is shown in full below:

Pooling studies under nearly identical protocols but with subjects of varying demographic and baseline characteristics can be justified by extension of the same principles outlined in ICH E9 (Statistical Principles of Clinical Trials) for which different centres in a single multicentre trial are pooled together, i.e. adherence to a common protocol that has been implemented in the same way at all centres using the same standardised procedures and evaluation criteria (as has been done in these studies), and a homogeneous treatment effect across studies (as is the case with these studies). In particular, when pooling the data across these studies, we considered study design, inclusion and exclusion criteria relative to disease severity, and whether the studies were run contemporarily such that the SoC treatment options were similar. These studies followed very similar protocols, were of nearly identical design, had identical inclusion and exclusion criteria, and were conducted over the same time period. Nevertheless, given the heterogeneous presentation of SLE disease and the fact that the Phase 3 program was run globally, one should expect to have variation in the patient population, both within the studies (e.g. between different centres) and between the studies (analogous to differences between centres within the same study).

Since it has been established that the conduct of the studies was effectively the same, one must then determine whether the relative treatment effect is different in one study compared with the other study when evaluating whether two studies are similar enough to pool. Each of the Phase 3 studies achieved statistical significance for belimumab 10mg/kg on the pre-specified primary endpoint of SRI response at Week 52; therefore, these nearly identical, studies provide independent replication of results. While pooling is not necessary to establish the effectiveness of belimumab, it was considered appropriate in order to evaluate treatment effects in high disease activity subgroups of interest, given that the individual studies were not designed to provide sufficient power to demonstrate effectiveness within subgroups. When the two Phase 3 studies were pooled a test for a treatment-by-study interaction was undertaken for the SRI analysis and the treatment-by-study interaction was >0.5. Likewise, for the Target Population of high disease activity, the treatment-by-study interaction was >0.7.

Additionally, a multivariate logistic regression model was developed in order to determine predicators of SRI response. Of the characteristics highlighted as being different between the two studies neither age, race, proteinuria, nor raised IgG were predictors of response. SLICC Damage Score and

complement levels (and their interaction terms with treatment) were included in the final model and neither study (p=0.54) nor the treatment-by-study interaction (p=0.95) was a predictor of SRI response. This result further substantiates that the study is not a predictor of SRI response, thus we believe that is reasonable and valid to pool the two studies.

National Institute for Health and Clinical Excellence

Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus

Please find enclosed the ERG report prepared for this appraisal.

You are asked to check the ERG report from Warwick Evidence to ensure there are no factual inaccuracies contained within it. If you do identify any factual inaccuracies you must inform NICE by **5pm, 6 July 2011** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The attached proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 Section 1.2.1 Primary outcome, Page 13, Para 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The corresponding adjusted odds ratios for a response in BLISS-52 and in BLISS-76 were respectively 1.83 (95% CI: 1.30, 2.59; P = 0.0006) and 1.52 (95% CI: 1.07, 2.15; P = 0.027).	The corresponding adjusted odds ratios for a response in BLISS-52 and in BLISS-76 were respectively 1.83 (95% CI: 1.30, 2.59; P = 0.0006) and 1.52 (95% CI: 1.07, 2.15; P = 0.0207).	P-value incorrect.	Minor amendment; see errata sheet.

Issue 2 Section 1.2.1 Primary outcome, Page 13, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
For the high disease activity subgroup (Target population) pooled across trials the difference in percentage of responders between the belimumab group and placebo group was 24.8% and the adjusted odds ratio was 2.7 (95% CI: 1.8, 4.1; P < 0.0001).	For the high disease activity subgroup (Target population) pooled across trials the difference in percentage of responders between the belimumab group and placebo group was 24.8% and the adjusted odds ratio was 2.7 (95% CI: 1.8, 4.1; P = 0.0001).	Sign in P-value incorrect.	P sign correct according to Table 5.15 of MS which states: P < 0.0001

Issue 3 Section 1.2.1 Primary outcome, Page 13, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
For the Target population in BLISS-76 the difference in percentage of responders between the belimumab group and placebo group was 22.4% and the adjusted odds ratio was 2.5 (95% CI: 1.3, 4.6; P < 0.0045).	For the Target population in BLISS-76 the difference in percentage of responders between the belimumab group and placebo group was 22.4% and the adjusted odds ratio was 2.5 (95% CI: 1.3, 4.6; P = 0.0045).	Sign in P-value incorrect.	Minor amendment; see errata sheet.

Issue 4 Section 1.2.2 Secondary outcomes, Page 14, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The values reported in the quoted text below are the least square mean differences for PGA change and not the raw mean changes. This is inconsistent with how all other results from this type of analysis are reported in this section.	For BLISS-76 the difference between groups was very small and in favour of placebo (-0.49 placebo and -0.44 belimumab) and did not reach statistical significance (P = 0.7987).	Maintains consistency with how all other results are reported using raw means not the estimated least square means. Least square mean presented. Changed to raw mean value.	Minor amendment; see errata sheet.

Issue 5 Section 1.2.2 Secondary outcomes, Page 14, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
For the Target population pooled across trials belimumab delivered a greater reduction in PGA score than placebo (P = 0.028 with mean changes of -0.42 and -0.52 for placebo and belimumab, respectively).	For the Target population pooled across trials belimumab delivered a greater reduction in PGA score than placebo (P = 0.0268 with mean changes of -0.42 and -0.52 for placebo and belimumab, respectively).	P-value incorrect.	Minor amendment; see errata sheet.

Issue 6 Section 1.2.2 Secondary outcomes, Page 15, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Reduction in steroid use was specified as a major secondary outcome. In BLISS-52 at baseline 68.6% of patients were receiving ≥ 7.5 mg/day prednisone.	Reduction in steroid use was specified as a major secondary outcome. In BLISS-52 at baseline 68.6% of patients were receiving > 7.5 mg/day prednisone.	Wrong sign used.	Minor amendment; see errata sheet.

Issue 7 Section 1.2.2 Secondary outcomes, Page 15, Para 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The mean change in SF 36 PCS scores was specified as, a major secondary outcome.	The mean change in SF 36 PCS scores at week 24 was specified as, a major secondary outcome.	Specificity around timing of major secondary endpoint is important for interpretation.	Not accepted. The first words of the following sentence make it clear that text refers to week 24.

Issue 8 Section 1.2.2 Secondary outcomes, Page 16, Para 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
For the Target population pooled across trials, relative to placebo, belimumab significantly delayed time to both first flare (P = 0.007) and to first severe flare (P = 0.0028).	For the Target population pooled across trials, relative to placebo, belimumab significantly delayed time to both first flare (P = 0.0017) and to first severe flare (P = 0.0028).	P-value incorrect.	Minor amendment; see errata sheet.

Issue 9 Section 1.2.2 Secondary outcomes, Page 16, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Data reported elsewhere[5] were: BLISS-52 score change 0.06 and 0.04 for placebo and belimumab groups respectively, P for difference 0.4222; BLISS-76 score change 0.05 and 0.03 for placebo and belimumab groups respectively, P for difference 0.3415.	Data reported elsewhere[5] were: BLISS-52 score change 0.05 and 0.03 for placebo and belimumab groups respectively, P for difference 0.4222; BLISS-76 score change 0.06 and 0.04 for placebo and belimumab groups respectively, P for difference 0.3415.	Results presented against incorrect trials.	Minor amendment; see errata sheet.

Issue 10 Section 1.2.2 Secondary outcomes, Page 17, Para 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
In BLISS-52 belimumab delivered a greater percentage of responders than did placebo (P = 0.0038); in BLISS-76 the difference in favour of belimumab failed to reach the conventional level of statistical significance (P = 0.064).	In BLISS-52 belimumab delivered a greater percentage of responders than did placebo (P = 0.0038); in BLISS-76 the difference in favour of belimumab failed to reach the conventional level of statistical significance (P = 0.0604).	P-value incorrect.	Minor amendment; see errata sheet.

Issue 11 Section 1.2.3 Safety, Page 17, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Although all patients had a history of auto-immunity, at recruitment 30% currently lacked anti-nuclear antibodies.	Although all patients had a history of auto-immunity, at recruitment approximately 30% lacked anti-nuclear antibodies.	Actual figure was 28.7%.	Minor amendment; see errata sheet.

Issue 12 Section 1.2.3 Safety, Page 17, Para 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
There were 15 deaths during the controlled phase of the three trials; 3 in the placebo group (n=675), and 12 in the belimumab groups (n=1458) with 6 each in the 10mg/kg and 1mg/kg groups respectively.	There were 14 deaths during the controlled phase of the three trials; 3 in the placebo group (n=675), and 11 in the belimumab groups (n=1458) with 6 in the 10mg/kg and 5 in the 1mg/kg groups respectively.	Only 14 deaths occurred during the controlled phase of the three trials. An additional death occurred in the belimumab 1 mg/kg group 15 weeks after the patient stopped belimumab treatment.	See errata sheet.

Issue 13 Section 1.2.3 Safety, Page 17, Para 5

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The percentage of patients experiencing at least one serious AE and at least one serious AE was very similar between placebo and belimumab groups ranging from13.5% to 18.6%, with a very slight numerical excess in the belimumab group.	The percentage of patients experiencing at least one serious AE ~~and at least one serious AE~~ was very similar between placebo and belimumab groups ranging from13.5% to 18.6%, with a very slight numerical excess in the belimumab group.	Duplicate text.	Minor amendment; see errata sheet

Issue 14 Section 1.2.3 Safety, Page 17, Para 5

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The percentage of patients experiencing at least one severe AE was 15.4% for the placebo group and 16% across the belimumab groups;	The percentage of patients experiencing at least one severe AE was 15.4% for the placebo group and 16% across the belimumab 1 mg/kg and 10mg/kg groups;	Clarity around treatment groups being compared aids interpretation.	The MS states that "*The 4 mg/kg dose was only studied in Study LBSL02 and had a safety and tolerability profile comparable to the placebo group and other belimumab dose groups*". This being the case ERG cannot identify any advantage in changing the wording in the way suggested.

Issue 15 Section 1.2.3 Safety, Page 17, Para 6

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Occurrence of infusion plus hypersensitivity reactions was similar between belimumab and placebo-treated patients (17% and 14.7%, respectively).	Occurrence of infusion plus hypersensitivity reactions was similar between belimumab 1 mg/kg and 10mg/kg and placebo-treated patients (17% and 14.7%, respectively).	Clarity around treatment groups being compared aids interpretation.	The MS states that "*The 4 mg/kg dose was only studied in Study LBSL02 and had a safety and tolerability profile comparable to the placebo group and other belimumab dose groups*". This being the case ERG cannot identify any advantage in changing the wording in the way suggested.

Issue 16 Section 1.2.3 Safety, Page 18, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
There were four non-melanoma skin cancers: two basal cell carcinomas, and two squamous cell carcinomas (1 in the placebo group, 3 in the belimumab 1mg/kg group).	There were four non-melanoma skin cancers: two basal cell carcinomas, and two squamous cell carcinomas (1 in the placebo group, 3 in the belimumab 10mg/kg group).	Incorrect belimumab group identified.	See errata sheet.

Issue 17 Section 1.6.2 Weakness and areas of uncertainty, Page 25, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Similarly subgroup analysis (Manufacturer's clarification document) of the primary outcome according to geographical region (USA/Canada, Western Europe, Eastern Europe, America-excluding USA/Canada, Asia) indicated that response was strongest in America-excluding-USA/Canada and weakest in USA/Canada and Western Europe.	~~Similarly~~ In the subgroup analysis (Manufacturer's clarification document) of the primary outcome according to geographical region (USA/Canada, Western Europe, Eastern Europe, America-excluding USA/Canada, Asia) indicated that response was strongest in Eastern Europe and Asia and weakest in USA/Canada and America excluding USA/Canada.	Results of subgroup analysis have been reported incorrectly. Treatment difference in the primary outcome is outlined below (in order of response): Eastern Europe - 37.5% (p=0.137) Asia - 26.9% (p=0.0112) Western Europe/Australia/Israel – 25.9% (p=0.0550) USA/Canada - 21.1% (p=0.0848) America excluding USA/Canada - 16.8% (p=0.0525)	ERG based their statement on Table A9.1 of the clarification response. The statement refers to the whole population from BLISS trials and is correct according to Table A9.1. The figures listed by the manufacturer are new and do not appear in the MS or in the clarification document.

Issue 18 Section 3.1 Population, Page 35, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The reported clinical benefit observed in BLISS-76 in the following sentence is inaccurate: "This would be of little consequence if the clinical results were consistent between trials; however this was not so for some outcomes and in general very little clinical benefit was observed in BLISS-76 compared to some benefits in BLISS-52."	This would be of little consequence if the clinical results were consistent between trials; however this was not so for some outcomes ~~and in general very little clinical benefit was observed in BLISS-76 compared to some benefits in BLISS-52.~~	The results of the primary endpoint and some secondary endpoints in BLISS-76 were clinically and statistically significant.	See errata sheet ("relatively" substituted for "very").

Issue 19 Section 2.5 Comparators, Page 37, Bullet 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
[redacted]	[redacted]	The details of the PAS are CIC	ERG has brought this to the attention of NICE.

Issue 20 Section 3.6 Other relevant factors, Page 40, Para 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The statement concerning the marginal relevance of women of childbearing age is inaccurate.	Special considerations and issues raised in the manufacturer's scope include: 1) the innovative nature of belimumab for SLE; 2) the inability of the utility method to capture the QoL of SLE patients sufficiently sensitively; and 3) the impact of SLE on particular ethnic groups and on women of childbearing age. The proposed SPC specifies that belimumab should not be administered to pregnant women or to those planning pregnancy ~~and therefore the special consideration relating to women of childbearing age appears to be of marginal relevance.~~	Women of childbearing age are a significant proportion of patients with SLE who may require treatment and who may not be pregnant or planning pregnancy.	Accepted; see errata sheet.

Issue 21 Section 4.1.3.1 Important included studies, Page 45, Table 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The percentage of patients enrolled from USA and Canada in study LBSL02 are listed incorrectly.	USA (99%) Canada (1%)	Incorrect percentages presented. 446 patients from the USA and 3 patients from Canada.	Not accepted. ERG took %s from FDA briefing package Table 2 which states: 98% and 2%. Wallace et al do not report on this. ERG has no other way of checking the correct value. The difference is too trivial to warrant changing. Therefore, the supplied percentage from FDA briefing package Table 2 is retained.

Issue 22 Section 4.1.3.1 Important included studies, Page 45, Table 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Reported infusion days for rituximab and placebo in the EXPLORER study have been reported incorrectly.	*** Placebo by IV infusion on days 1, 15, 168 and182 # Rit = Rituximab on days 1, 15, 168 and182	Incorrect reporting of infusion days in the EXPLORER study.	See errata sheet.

Issue 23 Section 4.2.1 Scope and synopsis of the studies providing clinical evidence, Page 47, Table 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The percentage of patients enrolled from USA and Canada in study LBSL02 are listed incorrectly.	USA (99%) Canada (1%)	Incorrect percentages presented. 446 patients from the USA and 3 patients from Canada.	Not accepted. ERG took %s from FDA briefing package Table 2 which states: 98% and 2%. Wallace et al do not report on this. ERG has no other way of checking the correct value. The difference is too trivial to warrant changing. See issue 21 above.

Issue 24 Section 4.2.1 Scope and synopsis of the studies providing clinical evidence, Page 48, Para 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The following statement is incorrect: Results for Target population patients who received the 1mg/kg regimen are not in the public domain and the manufacturer stated that they were unable to supply these results within the time constraints of the clarification process because of the large amount of other information that was requested.	Results for Target population patients who received the 1mg/kg regimen are not in the public domain ~~and the manufacturer stated that they were unable to supply these results within the time constraints of the clarification process because of the large amount of other information that was requested.~~	We responded to all of the clarification questions provided by the ERG. There were no questions relating to the 1 mg/kg regimen, and hence we did not provide these.	Not accepted. The ERG considers the statement a fair reflection of events. The manufacturer responded to the first clarification request forwarded by NICE saying they had statistical resource issues; a compromise was reached about the information requested which resulted in the omission of 1 mg / kg results for the target population. The manufacturer did supply all information asked for in the reduced clarification request.

Issue 25 Section 4.2.3 Description and critique of manufacturers outcome selection, Page 52, Table 6

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Two of the outcomes listed as 'other outcome reported' were specified secondary outcomes.	Amend the Outcome specification to 'Specified secondary outcome' for the following outcomes: No worsening in PGA score by ≥ 0.3 No new BILAG 1A/2B domain scores	Provides clarity around which outcomes were specified secondary outcomes.	Not accepted. The ERG statement is based on MS Table 5.12 page 84 to 90 titled "*Primary and secondary outcomes of the RCTs*". This does NOT list these outcomes as pre-specified.

Issue 26 Section 4.2.5.2 BLISS trials: demography of patients, Page 59, Table 10

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Number of patients with Rash in the Placebo arm of BLISS-76 is incorrect.	Please amend to 187 (68%).	Incorrect value.	See errata sheet.

Issue 27 Section 4.2.5.2 BLISS trials: demography of patients, Page 59, Table 11

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The number of patients in the belimumab 10 mg/kg arm of BLISS-76 is incorrect.	Please amend to n=81.	Incorrect value.	See errata sheet.

Issue 28 Section 4.2.5.3 BLISS trial results by outcome, Page 60, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The corresponding adjusted odds ratios for a response in BLISS-52 and in BLISS-76 were respectively 1.83 (95% CI: 1.30, 2.59; P = 0.0006) and 1.52 (95% CI: 1.07, 2.15; P = 0.027).	The corresponding adjusted odds ratios for a response in BLISS-52 and in BLISS-76 were respectively 1.83 (95% CI: 1.30, 2.59; P = 0.0006) and 1.52 (95% CI: 1.07, 2.15; P = 0.0207).	P-value incorrect.	Minor amendment; see errata sheet.

Issue 29 Section 4.2.5.3 BLISS trial results by outcome, Page 61, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
For the Target population pooled across trials and in BLISS-52, at many times, a significantly greater response was observed for the belimumab group relative to placebo group (significance tests uncorrected for multiple testing), however, for BLISS-76 the only time a significantly (P < 0.05) greater response was observed for the belimumab group was at week 52.	For the whole population and Target population pooled across trials and in BLISS-52…	More accurately describes the response observed.	Not accepted. Section refers to target population. Grammar nuance not error.

Issue 30 Section 4.2.5.3 BLISS trial results by outcome, Page 64, Table 14

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG response</th></tr>
<tr>
<td>This table has been repeated.

In addition, results for the High Disease Activity Subgroup have been mixed up.</td>
<td>Please delete one instance.

Please amend table as follows:
<table>
<tr><th colspan="2">High Disease Activity Subgroup Pooled Total</th><th colspan="2">High Disease Activity Subgroup BLISS-52</th><th colspan="2">High Disease Activity Subgroup BLISS-76</th></tr>
<tr><th>Placebo N = 203</th><th>10 mg/kg N = 193</th><th>Placebo N = 107</th><th>10 mg/kg N = 112</th><th>Placebo N = 96</th><th>10 mg/kg N = 81</th></tr>
<tr><td>71 (35.0%)</td><td>116 (60.1%)</td><td>42 (39.3%)</td><td>73 (65.2%)</td><td>29 (30.2%)</td><td>43 (53.1%)</td></tr>
<tr><td>-</td><td>2.8
(1.8, 4.2)</td><td>-</td><td>3.0
(1.7, 5.2)</td><td>-</td><td>2.5
(1.4, 4.8)</td></tr>
<tr><td>-</td><td><0.0001</td><td>-</td><td>0.0001</td><td>-</td><td>0.0036</td></tr>
<tr><td>-</td><td>(25.1%)</td><td>-</td><td>(25.9%)</td><td>-</td><td>(22.9%)</td></tr>
<tr><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
</table>
</td>
<td>Table has been repeated. Results for the High Disease Activity Subgroup have been presented incorrectly.</td>
<td>See errata sheet.</td>
</tr>
</table>

Issue 31 Section 4.2.5.3 BLISS trial results by outcome, Page 65, Table 15

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The observed difference in the High Disease Activity Subgroup Pooled Total population for the % of patients with no new 1A/2B BILAG domain scores is incorrect.	Please amend from 13.6 to 13.5.	Incorrect value.	Minor amendment; see errata sheet.

Issue 32 Section 4.2.5.3 BLISS trial results by outcome, Page 66, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
For BLISS-76 the difference between groups was very small and in favour of placebo (-0.49 placebo and -0.48 belimumab) and did not reach statistical significance (P = 0.7987).	For BLISS-76 the difference between groups was very small and in favour of placebo (-0.49 placebo and -0.44 belimumab) and did not reach statistical significance (P = 0.7987).	Maintains consistency with how all other results are reported using raw means not the estimated least square means. Least square mean presented. Changed to raw mean value.	See errata sheet.

Issue 33 Section 4.2.5.3 BLISS trial results by outcome, Page 66, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
For the Target HDAP pooled across trials belimumab delivered a greater reduction in PGA score than placebo (P = 0.028 with mean changes of -0.42 and -0.52 for placebo and belimumab, respectively).	For the Target HDAP pooled across trials belimumab delivered a greater reduction in PGA score than placebo (P = 0.0268 with mean changes of -0.42 and -0.52 for placebo and belimumab, respectively).	P-value incorrect.	Minor amendment; see errata sheet.

Issue 34 Section 4.2.5.3 BLISS trial results by outcome, Page 68, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
In BLISS-52 a larger improvement in PGA score was observed for the 10mg/kg group than for placebo (P = 0.0001) whereas in BLISS-76 the difference between treatments was trivial (P = 0.115).	In BLISS-52 a larger improvement in PGA score was observed for the 10mg/kg group than for placebo (P = 0.0001) whereas in BLISS-76 the difference between treatments was trivial (P = 0.1159).	P-value incorrect.	Minor amendment; see errata sheet.

Issue 35 Section 4.2.5.3 BLISS trial results by outcome, Page 69, Table 18

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The observed difference in the High Disease Activity Subgroup in BLISS-52 is incorrect.	Please amend from 13.5 to 13.2.	Incorrect value.	Minor amendment; see errata sheet.

Issue 36 Section 4.2.5.3 BLISS trial results by outcome, Page 69, Para 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
In BLISS-52 and BLISS-76 at baseline 68.6% and 44.9% of patients respectively were receiving ≥ 7.5 mg/day prednisone.	In BLISS-52 and BLISS-76 at baseline 68.6% and 44.9% of patients respectively were receiving > 7.5 mg/day prednisone.	Wrong sign used.	Minor amendment; see errata sheet.

Issue 37 Section 4.2.5.3 BLISS trial results by outcome, Page 72, Para 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
For the high disease activity Target population pooled across trials, belimumab significantly delayed time to first flare relative to placebo (P = 0.007; Figure 10).	For the high disease activity Target population pooled across trials, belimumab significantly delayed time to first flare relative to placebo (P = 0.0017; Figure 10).	P-value incorrect.	Minor amendment; see errata sheet.

Issue 38 Section 4.2.5.3 BLISS trial results by outcome, Page 73, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The word ‘flare’ has been left out of the sentence below: When the whole populations from the BLISS trials were pooled the difference between treatments for time to first severe reached statistical significance (P = 0.0011; Figure 13).	When the whole populations from the BLISS trials were pooled the difference between treatments for time to first severe flare reached statistical significance (P = 0.0011; Figure 13).	The word ‘flare’ left out in the description of the this outcome.	Minor amendment; see errata sheet.

Issue 39 Section 4.2.5.3 BLISS trial results by outcome, Page 78, Para 2-4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Text has been repeated.	There was no significant difference between belimumab and placebo in the absolute change of EQ-5D score from baseline in either trial or pooled total populations during clinic visits. The results for the 10 mg/kg belimumab and placebo groups in BLISS-76 were indistinguishable. For the pooled target population the difference between 10 mg/kg and placebo groups reached statistical significance in favour of belimumab at week 24 (P ≤ 0.01), but the difference had almost completely faded by week 52 MS Figure 5.21 Page 135). Results for the mean change in SELENA SLEDAI score from baseline at week 52 were submitted in MS Table 5.17 (Page 113) and clarification response Table A6.1 and are summarised in Table 19. ~~There was no significant difference between belimumab and placebo in the absolute change of EQ-5D score from baseline in either trial or pooled total populations during clinic visits. The results for the 10mg/kg belimumab and placebo groups in BLISS-76 were indistinguishable. For the pooled Target population the difference between 10mg/kg and placebo groups reached statistical significance in favour of belimumab at week 24 (P ≤ 0.01), but the difference had almost completely faded by week 52 MS Figure 5.21 Page 135).~~ ~~Results for the mean change in SELENA SLEDAI score from baseline at week 52 were submitted in MS Table 5.17 (Page 113) and clarification response Table A6.1 and are summarised in Table 19.~~	Duplication of text.	Minor amendment; see errata sheet.

Issue 40 Section 4.2.5.3 BLISS trial results by outcome, Page 79, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Sentence on FACIT and SF-36 included incorrectly following a sentence about SLEDAI score.	The HGS Briefing Document to the FDA[5] (see Figure 21) showed the percentage change in SLEDAI score (relative to baseline) throughout the two trials; this is reproduced in Figure 21 ~~for the mean change in FACIT and SF-36 vitality score by week 52.~~	Incorrect statement.	See errata sheet.

Issue 41 Section 4.2.5.3 BLISS trial results by outcome, Page 80, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The submission pooled results from three RCTs: BLISS-52, BLISS-76 and LBSL02. LBSL02 lasted 52 weeks, preceded the BLISS trials, and was conducted in North America (98% patients from the USA), and did not employ the SRI composite outcome measure.	The submission pooled results from three RCTs: BLISS-52, BLISS-76 and LBSL02. LBSL02 lasted 52 weeks, preceded the BLISS trials, and was conducted in North America (99% patients from the USA), and did not employ the SRI composite outcome measure.	Incorrect percentage used. 446 of 449 patients from the USA.	ERG took %s from FDA briefing package table 2 which states: 98% and 2%. Wallace et al do not report on this. ERG have no other way of checking the correct value. The difference is too trivial to warrant changing. See issue 21 above.

Issue 42 Section 4.2.5.3 BLISS trial results by outcome, Page 80-81

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Percentage of patients at recruitment lacking anti-nuclear antibodies was 28.7%.	Although all patients had a history of auto-immunity, at recruitment approximately 30% ~~currently~~ lacked anti-nuclear antibodies.	Incorrect value.	See errata sheet.

Issue 43 Section 4.2.5.3 BLISS trial results by outcome, Page 83, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The percentage of patients experiencing at least one severe AE was 15.4% for the placebo group and 16% across the belimumab 1 mg/kg and 10 mg/kg groups	The percentage of patients experiencing at least one severe AE was 15.4% for the placebo group and 16% across the 1 mg/kg and 10 mg/kg belimumab groups	Clarity around treatment groups being compared aids interpretation.	The MS states that "*The 4 mg/kg dose was only studied in Study LBSL02 and had a safety and tolerability profile comparable to the placebo group and other belimumab dose groups*". This being the case ERG cannot identify any advantage in changing the wording in the way suggested.

Issue 44 Section 4.2.5.3 BLISS trial results by outcome, Page 83, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Occurrence of infusion plus hypersensitivity reactions was similar between belimumab and placebo-treated patients (17% and 14.7%, respectively).	Occurrence of infusion plus hypersensitivity reactions was similar between belimumab 1 mg/kg and 10 mg/kg and placebo-treated patients (17% and 14.7%, respectively).	Clarity around treatment groups being compared aids interpretation.	The MS states that "*The 4 mg/kg dose was only studied in Study LBSL02 and had a safety and tolerability profile comparable to the placebo group and other belimumab dose groups*". This being the case ERG cannot identify any advantage in changing the wording in the way suggested.

Issue 45 Section 4.2.5.3 BLISS trial results by outcome, Page 83, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Of 1458 belimumab treated patients, 15 experienced hypersensitivity reactions on the day of infusion compared to one of 675 placebo-treated patients.	Of 1458 belimumab treated patients, 14 experienced hypersensitivity reactions on the day of infusion compared to one of 675 placebo-treated patients.	Incorrect value.	See errata sheet.

Issue 46 Section 4.2.6 Pooling of trial data, Page 85, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 85 of the ERG report states: "The SLEDAI scores shown in Figure 22 illustrates the differences between Target and JHU populations, (year one and last year scores are shown for JHU cohort, Figure B17.2 of the clarification document, the clarification response did not make clear which was year one and which last year)." GSK believes that the last sentence is inaccurate and should be deleted.	Text should read "The SLEDAI scores shown in Figure 22 illustrates the differences between Target and JHU populations, (year one and last year scores are shown for JHU cohort, Figure B17.2 of the clarification document)."	The clarification response to Question B17 did make it clear which were the first and last year values on the chart. The following footnote was presented under the graph "Note: the Blue bars represent a histogram of the SLEDAI score at first visit. The Red outline bars represent a histogram of at the patient's last visit SLEDAI score in the cohort."Corrected for accuracy but has minimal impact on the results.	Accepted. See errata sheet.

Issue 47 Section 4.2.7 Conclusion, Page 87, Para 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
For both whole and Target populations, results from BLISS-52 were more favourable for belimumab than results from BLISS-52.	For both whole and Target populations, results from BLISS-52 were more favourable for belimumab than results from BLISS-76.	Typographical error.	See errata sheet.

Issue 48 Section 4.2.7 Conclusion, Page 87, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
This statement is incorrect: "For all secondary outcomes in BLISS-76 effect sizes were insufficient to be confident that effects could not be accounted for by chance."	For ~~all~~ some secondary outcomes in BLISS-76 effect sizes were insufficient to be confident that effects could not be accounted for by chance.	There were some secondary outcomes in BLISS-76 where the p-value reached its predefined level for statistical significance.	The word "non-major" was omitted, see errata sheet.

Issue 49 Section 6.1.1 Clinical Effectiveness, Page 142, Para 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The primary end point, which was the percentage of responders at week 52 according to the novel composite SRI outcome measure, was very similar for pooled Target population and pooled "licence population" with respectively 19.8% and 24.8% extra responders for belimumab compared to placebo (belimumab vs. placebo odds ratio = 2.7 for both populations).	The primary end point, which was the percentage of responders at week 52 according to the novel composite SRI outcome measure, was very similar for pooled Target population and pooled "licence population" with respectively 24.8% and 19.8% extra responders for belimumab compared to placebo (belimumab vs. placebo odds ratio = 2.7 for both populations).	Results presented against incorrect trials.	See errata sheet.

Issue 50 Section 5.2.3 Model Structure, Page 96, Para 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The sign of the coefficient quoted for the previous year's SS score (SS_{t-1}) in the regression equations for the evolution of the SS score beyond Week 52 is incorrect. The ERG report states the	The two regression equations should be as follows:	The corrected equation enables estimation of all subsequent year SS scores which is key to estimating level of disease activity during each model cycle.	This is an error in the ERG report and should be corrected as suggested by the manufacturer; see errata sheet..

following two equations:			

Issue 51 Section 5.2.3 Model Structure, Page 100, Para 2

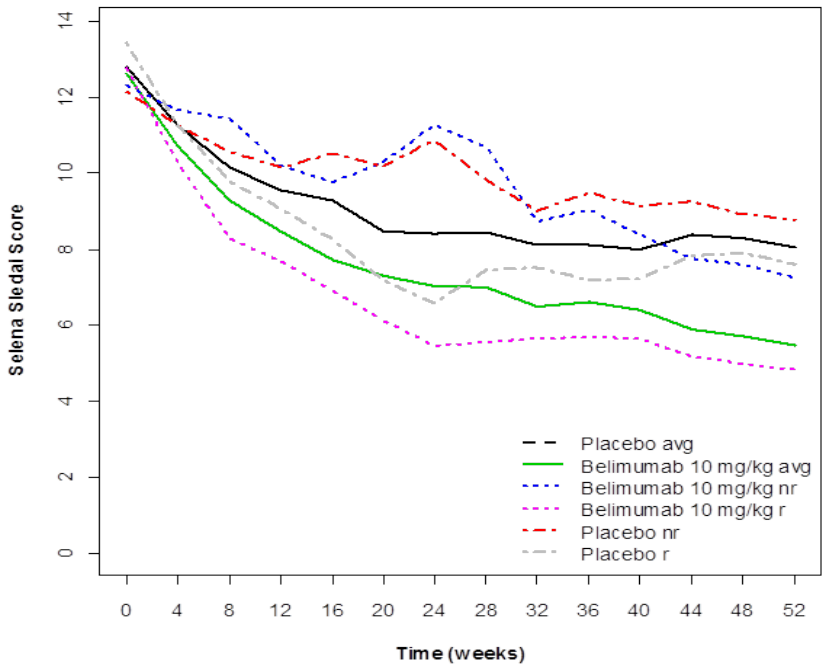
Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
The terminology is incorrect in this section. It should state organ damage, which is irreversible whereas organ involvement relates to the presence of symptoms at a point in time (ie items in the SELENA-SLEDAI instrument).	We propose the corrected text: *"Literature element 3.c Cost impact of **organ damage**"* *"A similar approach is undertaken for the cost impacts of **organ damage** as for the QoL impacts, only with the number of patients in the JHU cohort experiencing the individual elements among those having had an event within the organ class giving rise to the weight to apply."*	The model simulates organ damage over time. Organ involvement is not considered in the base case. This is an important distinction in appraising the economic modeling.	The revision suggested by the manufacturer is accepted by the ERG. See errata sheet.

Issue 52 Section 5.2.5 Base case probabilistic results, Page 108, Para 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG report states in this section: '*A key variable that has not been particularly explored or explained within the submission is the assumed rate of continuation and discontinuations among belimumab week 24 responders.*' In fact the estimated discontinuation rate was explored in both the univariate and probabilistic sensitivity analyses for both populations detailed in the GSK submission document. For the Total Population results see Pages 269-270, Figures 6.23, 6.24, 6.25 for the variable 'Natural Discontinuation Responders' in the tornado diagrams and the following text on page 273 *'Discontinuation probabilities for patients satisfying the six-month treatment continuation rule affect both incremental benefits and costs and thereby the ICER. For example, lower probabilities for natural discontinuation lead to higher incremental QALYs with belimumab compared with the base case value but significantly increased drug costs resulting in higher ICERs".* Similarly for the Target Population pages 291-292, Figures 6.37, 6.38, 6.39 for the variable 'Natural Discontinuation Responders' in the tornado diagrams and Tables B8.1, B8.2 and B8.3 in the response to clarification questions, and the following text on page 297 'Discontinuation probability affects both incremental QALYs and costs and thereby the ICER. For example, lower probabilities for natural discontinuation lead to higher incremental QALYs with belimumab compared with the base case value but significantly increased drug costs resulting in higher ICERs.'	GSK suggest that the sentence should be ignored as it is inaccurate.	Deleting this sentence will provide a fairer picture of all the variables explored in the sensitivity analyses as part of the decision problem.	The sensitivity analyses around the continuation rates are outlined in detail in the last table of section 4.2.6 of the ERG report. The explanation/ exploration intended by the ERG was: what is it that underlies this discontinuation rate and in the light of this is it reasonable to extrapolate this into the long term? No revision required.

Issue 53 Section 5.2.9.1 Model structure, Page 113-118

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
GSK does not agree with the ERG's suggestion that the belimumab non-responders should take the SS score at week 52 of the SoC non-responders and not the SS score of all the SoC patients. The data for change from baseline to week 52 in SS score for the Total BLISS population provided in Section 5.3 of our submission document and for the Target population in Table 7.1 in our response to the clarification questions was based on a last observation carried forward (LOCF) analysis and not on data that was available at week 52. We regret that this was not made clear in the Table titles in our submission document and response to clarification questions. The LOCF analysis had been pre-specified in the data analysis plan for SS score data. In the economic model however only the actual values recorded at week 52 were used to estimate change from baseline in SS score. This is believed to be more appropriate for patients on belimumab as it is more reflective of what happened to patients in the BLISS trial ie patients who withdrew were excluded. When only the raw data at week 52 are used then the mean change in the belimumab non-responders (-4.1) is close to the mean change seen for all SoC patients (-4.4) at week 52 and not to the mean change for just the SoC non-responders. Please see graph below for pattern of change for each arm of the model over the one year trial period based on observed cases. The belimumab non-responders have a relatively high mean disease score at 6 months. It is likely that these patients have a higher chance of discontinuation in the trial due to higher disease activity. As such, using LOCF, the week 52 data would contain a large number of 'old' values, giving an estimate higher than that expected in reality at week 52, if you assume that in clinical pratice these non-responders will no longer be receiving belimumab.	The text and tables provided on pages 113 through to 118 of the ERG report suggesting that the belimumab non-responders should take the SS score values for the SoC non-responders at week 52 should not be accepted as a more appropriate methodology for the economic modelling for this decision problem. We stand by our assumption that the belimumab non-responders should take the SS score values observed for all the SoC patients at week 52 as is currently simulated in the model. We believe this is a fair assumption based on the observed data in the BLISS trials.	We believe the assumption we used for belimumab non-responders taking the average SoC SS score at week 52 is appropriate in the economic model. It is important that this assumption is accepted as valid as it does have an important impact on the cost-effectiveness results.	In assessing this it is important to remember that the economic model assumes belimumab non-responders come off belimumab at week 24. The trial data reflects belimumab non-responders remaining on belimumab to week 52 and as such is not a good guide to what might be expected among belimumab non-responders coming off belimumab at week 24. The graph included by the manufacturer illustrates belimumab non-responders being similar to SoC non-responders at week 24. No revision required.

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
			

Issue 54 Section 5.2.9.1 Model structure, Page 119, Para 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG report states: "The values submitted in response to the ERG clarification question are averages across many individual patient iterations. Some non-linearity or rounding approximation may have crept into the figures. But it is difficult to reconcile the ERG cross check of the AMS for SoC and belimumab week 24 non-responders with the discrepancies between the ERG cross check and the manufacturer reported values of the AMS for belimumab week 24 responders. It is unclear whether the above discrepancy is due to an error in the manufacturer response to the ERG clarification question, an error in the VB coding of the model or an error in interpretation by the ERG. The manufacturer figures for the AMS may incorporate discontinuations within the belimumab week 24 responder figures."	We can confirm that there was an error in the response to the clarification questions provided by the modelling agency and that the values for AMS calculated by the ERG are the correct ones in Table 44 on Page 33. The modelling agency has confirmed that the VB code in the model is correct and so the model should correctly calculate AMS scores over time for belimumab week 24 responders.	To clarify the cause of the discrepancy. However the estimated cost-effectiveness will not be affected as the code in the model is correct.	No revision required.

Issue 55 Appendix 4, Page 159

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Table 5.8 from our submission has been left out of Appendix 4.	Include Table 5.8 from our submission.	Page 57 of the ERG report refers to Tables 5.8 to 5.11 in Appendix 4, however, only Tables 5.9 to 5.11 are included in Appendix 4. Readers will not be able to locate Table 5.8.	See errata sheet (text change to page 57 that refers to Appendix 4).

Belimumab for the treatment of autoantibody-positive systemic lupus erythematosus.

ERG REPORT: ERRATA SHEET

Page / location	Original	Change / Replacement
Pg 13 para 3, last line	P = 0.027	P = 0.0207
Pg 13 para 4, last line	P < 0.0045	P = 0.0045
Pg 14 para 2, line 5	And -0.48 belimumab	and -0.44 belimumab
Pg 14 para 2, line 7	P = 0.028	P = 0.0268
Pg 15 para 2, line 2	≥ 7.5 mg/day	> 7.5 mg/day
Pg 16 para 3, last line	P = 0.007	P = 0.0017
Pg 16 para 4, line 3 (SLICC/ACR)	BLISS-52	BLISS-76
Pg 16 para 4, line 4 (SLICC/ACR)	BLISS-76	BLISS-52
Pg 17 line 3	P = 0.064	P = 0.0604
Pg 17 para 2, line 7	30%	~30%
Pg 17 para 3, line 1	15 deaths	14 deaths
Pg 17 para 3, line 2	12 deaths	11
Pg 17 para 3, line 2	6 each in the 10mg/kg and 1mg/kg	6 in the 10mg/kg and 5 in the 1mg/kg
Pg 17 para 5	and at least one serious AE	Deleted duplicate text
Pg 18 para 2, last line	3 in the belimumab 1mg/kg	3 in the belimumab 10 mg/kg
Pg 35 para 2, penultimate line	very little clinical benefit	relatively little clinical benefit
Pg 37 Sect 3.3 bullet 4	Identified as CIC	Has been highlighted as CIC
Pg 40 para 3, last 2 lines	and therefore the special consideration relating to women of childbearing age appears to be of marginal relevance.	and therefore the special consideration relating to women of childbearing age is of relevance.
Page 45, Table 3 footer	*** Placebo .. on days 0, 14, 167 and181 # Rit = Rituximab on days 0 and 14	*** Placebo .. on days 1, 15, 167 and181 # Rit = Rituximab on days 1, 15, 167 and 181
Page 57, 4.2.5.2 line 2	Tables 5.8 to 5.11	Tables 5.9 to 5.11
Pg 59 Table 10 Rash placebo	87 (68%)	187 (68%)
Pg 59 Table 11 BLISS-76 row 2	Belimumab 10mg/kg (N=96)	Belimumab 10mg/kg (N=81)
Pg 60 last para, end of line 5	1.52 (95% CI: 1.07, 2.15; P = 0.027).	1.52 (95% CI: 1.07, 2.15; P = 0.0207).
Pg 63 Table 13	Duplicate tables (incorrect)	The correct table has been added
Pg 64 Table 14	Duplicate tables: 13 and 14 (incorrect)	This table has been **deleted**
Pg 65 Table 15, row 9, col 9	13.6%	13.5%
Pg 66 last para, line 6	and -0.48 belimumab	and -0.44 belimumab
Pg 66 last para, penultimate line	placebo (P = 0.028….	placebo (P = 0.0268….
Pg 67 para 2, line 2 - 3	difference between treatments was trivial (P = 0.115).	difference between treatments was trivial (P = 0.1159).
Pg 69 Table 18, row 4 column 11	13.5%	13.2%
Pg 69 para 1, line 2	≥ 7.5 mg/day prednisone	> 7.5 mg/day prednisone
Pg 72 para 1, line 2	P = 0.007	P = 0.0017
Pg 73 para 2, line 2	first severe reached	first severe flare reached
Pg 78 para 2, last 2 sentences are repeated text	The results for ….Figure 5.21 Page 135).	Deleted text: The results for ….Figure 5.21 Page 135).
Pg 79 last line	for the mean change in FACIT and SF-36 vitality score by week 52.	This sentence has been **deleted**
Pg 81 line 1	30%	~ 30%
Pg 83 Para 4, line 3	15 experienced hypersensitivity reactions	14 experienced hypersensitivity reactions
Pg 85 para 4, last sentence	the clarification response did not make clear which was year one and which last year	This sentence has been **deleted**
Pg 87, last word of para 3	BLISS-52	BLISS-76
Pg 87, para 4 line 1	For all secondary outcomes in BLISS-76	For all non-major secondary outcomes in BLISS-76
Pg 96 equation one	$2.0577 – 0.5837*SS_{t-1}$	$2.0577 + 0.5837*SS_{t-1}$
Pg 96 equation two	$3.0 – 0.5837*SS_{t-1}$	$3.0 + 0.5837*SS_{t-1}$
Pg 100 3c. Heading & Line 1	...organ involvement	...organ damage
Pg 142 para 2, line 5	with respectively 19.8% and 24.8%	with respectively 24.8% and 19.8%
Pg 144	last bullet point in section 6.1.2	Deleted last bullet point

were: an improved SLEDAI score by ≥ 4 points; a BILAG index showing no new grade A organ involvement or no two grade B organ involvements; a PGA score that has not increased by more than 0.3 points. The primary end point in both trials was the proportion of responders at week 52 relative to baseline according to the SRI.

In summary, the primary efficacy endpoint was the response rate at week 52, assessed with SLE Responder Index (SRI). A responder was defined as having a reduction of at least 4 points in the SELENA-SLEDAI score, no new BILAG A organ domain score, no more than 1 new BILAG B organ domain score, and no worsening in PGA score (increase < 0.3) at week 52 compared with baseline.

In both trials at 52 weeks SoC + 10mg/kg belimumab delivered a greater percentage of responders than did SoC + placebo. The difference in percentage of responders in the belimumab group relative to placebo group was 14% in BLISS-52and 9.4% in BLISS-76 and 11.8% for the whole population pooled across trials. The corresponding adjusted odds ratios for a response in BLISS-52 and in BLISS-76 were respectively 1.83 (95% CI: 1.30, 2.59; P = 0.0006) and 1.52 (95% CI: 1.07, 2.15; P = 0.0207).

For the high disease activity subgroup (Target population) pooled across trials the difference in percentage of responders between the belimumab group and placebo group was 24.8% and the adjusted odds ratio was 2.7 (95% CI: 1.8, 4.1; P < 0.0001). For the Target population in BLISS-52 the difference in percentage of responders between the belimumab group and placebo group was 25.9% and the adjusted odds ratio was 3.0 (95% CI: 1.7, 5.2; P < 0.0001). For the Target population in BLISS-76 the difference in percentage of responders between the belimumab group and placebo group was 22.4% and the adjusted odds ratio was 2.5 (95% CI: 1.3, 4.6; P = 0.0045).

The percentage of responders was also reported at multiple follow up times. For the Target population pooled across trials and in BLISS-52, at many times, a significantly greater response was observed for the belimumab group relative to placebo group (significance tests uncorrected for multiple testing), however for BLISS-76 the only time a significantly (P < 0.05) greater response was observed for the belimumab group was at week 52.

1.2.2 Secondary Outcomes

The pre-specified major secondary outcomes were: the SRI response at week 76; the percentage of patients with a ≥ 4 point SLEDAI improvement at week 52; mean change in PGA score at week 24, percentage of patients with prednisone reductions ≥ 25% from baseline to ≤7.5 mg/day during weeks 40 to 52 (in subjects whose baseline dose was > 7.5 mg/day); mean change in SF-36 PCS at week 24.

The major secondary outcome of percentage of SRI responders at week 76 failed to reach statistical significance (odds ratio and P value not submitted; odds ratio 1.31, 95% CI: 0.92 – 1.87, P = 0.1323 by logistic regression, taken from the FDA briefing package).[3]

Mean change in PGA score at week 24 was defined as a major secondary outcome. For the whole population in BLISS-52 the change in PGA score (week 24 relative to baseline) for both groups indicated disease improvement and was greater in the belimumab group (-0.54) than placebo group (-0.39; P = 0.0003 in support of belimumab). For BLISS-76 the difference between groups was very small and in favour of placebo (-0.49 placebo and -0.44 belimumab) and did not reach statistical significance (P = 0.7987). For the Target population pooled across trials belimumab delivered a greater reduction in PGA score than placebo (P = 0.0268 with mean changes of -0.42 and -0.52 for placebo and belimumab, respectively). Target population data was not been provided for the change in PGA score separately for the BLISS-52 and BLISS-76.

Components of the SRI at Week 52

A further major secondary outcome was the percentage of patients at week 52 that achieved a SLEDAI score reduction of ≥ 4 points. Both trials delivered a significantly greater percentage for belimumab than for placebo (P = 0.0024 and P = 0.0062 for BLISS-52 and BLISS-76, respectively). Similarly, the Target population data delivered a significantly greater percentage for belimumab (P = 0.0004 and P = 0.0063 for BLISS-52 and BLISS-76, respectively).

Results at week 52 for the other two SRI components (i.e. the BILAG index and PGA score) were submitted (non-major secondary outcomes). The percentage of patients in the whole population that satisfied BILAG and PGA criteria in BLISS-52 was greater for belimumab than placebo (significant at P = 0.0181 and P = 0.0048 for BILAG and PGA, respectively); however, for BLISS-76 the differences between belimumab and placebo were smaller and neither component reached statistical significance in favour of belimumab (P = 0.319 and P = 0.1258 for BILAG and PGA, respectively). Similarly, the percentage of patients in the Target population which satisfied BILAG and PGA criteria in BLISS-52 was greater for belimumab than placebo (P = 0.0099 and P = 0.0063 for BILAG and PGA, respectively); however, for BLISS-76 the differences between belimumab and placebo were far more modest and did not reach conventional statistical significance (P = 0.1297 and P = 0.1312 for BILAG and PGA, respectively).

In summary, in BLISS-52 for the total population and for the high disease activity subgroup, belimumab at 10mg/kg delivered significantly more responders at 52 weeks than placebo for

SLEDAI score reduction of ≥ 4 points, no worsening in PGA, and no worsening in BILAG. However, for BLISS-76 at 52 weeks total population and high disease activity subgroup, a significant response with belimumab 10mg/kg compared to placebo was only seen with the 4-point reduction in SELENA-SLEDAI component (difference between belimumab and placebo = 22%, odds ratio = 2.4 [95% CI: 1.3, 4.4; P < 0.0063]).

Reduction in steroid usage

Reduction in steroid use was specified as a major secondary outcome. In BLISS-52 at baseline 68.6% of patients were receiving > 7.5 mg/day prednisone. The corresponding percentage for BLISS-76 was 44.9%. The percentage of these patients whose steroid use was reduced in weeks 40 to 52 by the pre-specified amount was greater in the belimumab arm than the placebo arm in both trials. The difference (belimumab vs. placebo) failed to reach statistical significance in either trial: 18.6% vs. 12.0% in BLISS-52 (P = 0.0526 from logistic regression including baseline covariates) and 16.7% vs. 12.7% in BLISS-76 (P = 0.5323). For the Target population pooled across trials 15.9% and 7.1% reduced steroid use in belimumab and placebo groups, respectively (P = 0.0389 from logistic regression). For the Target population in the BLISS-52 trial there was a large difference in reduced steroid use between belimumab and placebo groups (18.5% and 5.3% respectively; odds ratio = 4.11, 95% CI: 1.29, 13.2; P = 0.0171). For the Target population in the BLISS-76 trial there was no difference between groups (11.1% and 10% reduced steroid use in belimumab and placebo groups respectively; odds ratio = 0.88, 95% CI: 0.21, 3.60; P = 0.8586).

Quality of life

The mean change in SF 36 PCS scores was specified as, a major secondary outcome. At week 24 relative to baseline it showed little difference between belimumab and placebo groups in BLISS-52 (P = 0.8870), or in BLISS-76 (P = 0.6601), or in the Target population pooled across trials (P = 0.4276).

Change in SF 36 PCS scores at week 52 was also specified as a non-major secondary outcome. No significant improvement was observed for BLISS-76 or Target populations (P = 0.5134 and P = 0.1124, respectively) however in BLISS-52 there was a difference between belimumab and placebo arms (4.18 vs. 2.96) (P = 0.0247).

In BLISS-52 over the course of the study there was a statistically non-significant difference (P value not submitted) in favour of belimumab relative to placebo in the absolute change of EuroQoL 5 Dimensions (EQ-5D) score from baseline, however the results for belimumab and placebo groups in BLISS-76 were indistinguishable. For the pooled Target population the

difference between 10mg/kg and placebo groups reached statistical significance in favour of belimumab at week 24 ($P \leq 0.01$), but the difference almost completely faded by week 52.

SLEDAI flare index

Other specified non-major secondary efficacy outcomes for which results were submitted included: time to first SLE flare (assessed using the SLEDAI Flare Index which categorizes flares as "mild or moderate" or "severe" based on 6 variables (see Appendix 1); disease progression at week 52 relative to baseline assessed using the SLICC/ACR index; fatigue over the course of the study estimated using the FACIT-Fatigue scale4 which ranges from 0 to 52 (0 is the worst possible score and 52 is the best).

In BLISS-52 the time to first flare and time to first severe flare were delayed by belimumab relative to placebo (HR 0.76, 95% CI: 0.63 – 0.91, P = 0.0036; HR 0.57, 95% CI: 0.39 – 0.85, P = 0.0055, respectively). In BLISS-76 there was no difference between groups in time to first flare (P = 0.4796) but relative to placebo belimumab somewhat delayed time to first severe flare (HR 0.72, 95% CI: 0.50 – 1.05, P = 0.0867). For the Target population pooled across trials, relative to placebo, belimumab significantly delayed time to both first flare (P = 0.0017) and to first severe flare (P = 0.0028).

SLICC/ACR organ damage

There was little difference between placebo and belimumab groups in terms of change in SLICC/ACR score at week 52; precise values by trial were not submitted. Data reported elsewhere[5] were: BLISS-76 score change 0.06 and 0.04 for placebo and belimumab groups respectively, P for difference 0.4222; BLISS-52 score change 0.05 and 0.03 for placebo and belimumab groups respectively, P for difference 0.3415.

Fatigue

At week 52 relative to baseline the belimumab group had greater improvement in FACIT-Fatigue score than the placebo group (4.8 belimumab and 2.1 placebo in BLISS-52, $P < 0.001$; 4.6 and 2.9 in BLISS-76, P = 0.05). For the Target population at weeks 8 and 12 the difference between groups was statistically in favour of belimumab ($P < 0.05$) however the difference between groups then diminished and at week 52 the difference no longer reached conventional statistical significance.

Modified SRI response

The results for a non-pre-specified secondary outcome, the "modified SRI" at week 52, were submitted. In the "modified SRI" serological improvements (2 points each for anti-dsDNA antibodies and for complement) were not allowed to contribute toward a ≥4 points reduction

in SLEDAI score. In BLISS-52 belimumab delivered a greater percentage of responders than did placebo (P = 0.0038); in BLISS-76 the difference in favour of belimumab failed to reach the conventional level of statistical significance (P = 0.0604).

1.2.3 Safety

The submission pooled results from three randomised controlled trials (RCTs): BLISS-52, BLISS-76 and LBSL02, providing information on 675 patients who received placebo and 1458 who were exposed to belimumab. LBSL02 lasted 52 weeks, preceded the BLISS trials, was conducted in North America (98% patients from the USA) randomised 449 patients to one of four treatments: SoC + placebo, SoC + 1mg/kg belimumab, SoC + 4mg/kg belimumab, and SoC + 10mg/kg belimumab. Although all patients had a history of auto-immunity, at recruitment ~30% currently lacked anti-nuclear antibodies. This trial did not employ the SRI composite outcome.

Deaths

There were 14 deaths during the controlled phase of the three trials; 3 in the placebo group (n=675), and 11 in the belimumab groups (n=1458) with 6 in the 10mg/kg and 5 in the 1mg/kg groups respectively. One further death in the 1mg/kg belimumab group followed 15 weeks after the patient stopped belimumab treatment. The causes of death were various.

Adverse events

In all treatment groups > 90% of patients experienced at least one adverse event (AE). The most commonly occurring AEs were headache, upper respiratory tract infection, arthralgia, nausea, urinary tract infection (UTI), diarrhoea and fatigue.

The percentage of patients experiencing at least one serious AE was very similar between placebo and belimumab groups ranging from13.5% to 18.6%, with a very slight numerical excess in the belimumab group. The most frequent serious AEs (≥ 1% in any treatment group) were pneumonia, pyrexia, UTI, cholelithiasis, and cellulitis. The percentage of patients experiencing at least one severe AE was 15.4% for the placebo group and 16% across the belimumab groups; the most common severe adverse events were not identified.

Occurrence of infusion plus hypersensitivity reactions was similar between belimumab and placebo-treated patients (17% and 14.7%, respectively).

Infections

Infections occurred slightly more often in patients treated with belimumab compared to placebo. The most frequent infections were upper respiratory tract infection (URTI), UTI, nasopharyngitis, sinusitis, and bronchitis.

Malignancy

Five solid organ malignancies were reported across the trials: a stomach carcinoid (placebo group, day 202); a breast cancer (belimumab 1mg/kg, day 102); a cervical cancer (belimumab 1mg/kg, day 439); an ovarian cancer (belimumab 1mg/kg, day 21, patient died); and a thyroid cancer (belimumab 1mg/kg, day 378). There were four non-melanoma skin cancers: two basal cell carcinomas, and two squamous cell carcinomas (1 in the placebo group, 3 in the belimumab 10mg/kg group).

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The submission omitted results for the 1mg/kg groups from the two pivotal trials. Therefore from information in the manufactures submission (MS) alone, the consistency of results across the whole data set could not be fully assessed and it was not possible to gauge the evidence for a dose response relationship. However, data for the 1mg/kg groups is available in the public domain (FDA documents pertaining to the Human Genome Sciences (HGS) Briefing Document to the FDA[3,5]) and the ERG have considered this information in critiquing the submitted evidence.

Even without the 1mg/kg group results the MS provided clinical evidence for a large number of outcomes reported for six separate populations (whole population from BLISS-52, whole population from BLISS-76, pooled whole populations from BLISS-52 plus BLISS-76, pooled Target populations from BLISS-52 plus BLISS-76, and after the clarification process Target population from BLISS-52 and Target population from BLISS-76. Additionally, AEs for LBSL02 were included.[6]

The most noticeable aspect of the submitted results was the relative lack of evidence for clinical effectiveness of belimumab seen in the BLISS-76 trial. Although at week 52 for the pre-specified primary outcome measure the percent responders (SRI) reached statistical significance in favour of belimumab (P = 0.027), at no other time point did this outcome reach significance. Furthermore, all major and non-major secondary outcome results submitted, except for a $\geq$ 4 point SLEDAI improvement at week 52 which is a subcomponent of the SRI response, likewise failed to reach statistical significance including: PGA change at week 24 and 52, SRI responders at week 76, reduction in use of steroids week 40 to 52, SF-36 change

3.1 Population

The manufacturer's scope specified two populations: the Phase III trial population (adults with active autoantibody-positive SLE), and a High Disease Activity Subgroup (HDAS).

The submitted clinical effectiveness evidence came from two multicentre international Phase III RCTs (BLISS-52 and BLISS-76). The geographical location of study centres differed considerably between trials. In BLISS-52 there were 90 centres: in 13 countries in Latin America there were 38 centres (Argentina, Brazil, Chile, Colombia and Peru), in Asia-Pacific there were 41 centres (Australia, Hong Kong, India, Korea, Philippines and Taiwan) and in Eastern Europe there were eleven centres (Romania and Russia). In BLISS-76 there were 136 centres in 19 countries in North America (Canada, Costa Rica, Mexico, Puerto Rico and USA) and Europe (Austria, Belgium, Czech Republic, France, Germany, Israel, Italy, The Netherlands, Poland, Romania, Slovakia, Spain, Sweden and UK); North America (65 centres) and Europe (62 centres) contributed 93% of the centres in BLISS-76. These geographical differences were reflected in racial differences between the populations in the two trials. Although both trials included adults with auto-antibody positive active SLE it appears clear that the population in BLISS-76 is more likely to be similar to that in England and Wales than that from BLISS-52. It is reasonable to assume that the results from BLISS-76 will be more generalisable to the UK. This would be of little consequence if the clinical results were consistent between trials; however this was not so for some outcomes and in general relatively little clinical benefit was observed in BLISS-76 compared to some benefits in BLISS-52.

The manufacturer's scope also specified a HDAS termed the "Target" population and described as the focus of the submission. The identification of the Target population, and the evidence for clinical effectiveness of belimumab in the Target population, arose from post hoc analyses of the two BLISS trials. The rationale for this deviation from NICE scope was largely on economic grounds in that cost effectiveness was more favourable. Because the BLISS-76 trial subpopulation is more likely to match high disease activity patients in the UK than the BLISS-52 subpopulation, it is arguable that the BLISS-76 Target population is the most appropriate.

The Target population was defined as: *"Adults with active autoantibody-positive systemic lupus erythematosus with evidence for serological disease activity (low complement, positive anti-dsDNA) and SELENA-SLEDAI ≥ 10"* [MS Page 53]. There are undoubtedly patients in the UK who correspond to the Target population; however, according to expert clinical opinion, the SELENA SLEDAI is not commonly used to define high disease activity and it

3.3 Comparators

Three comparators were identified in the NICE final scope: standard of care (SoC), rituximab, and cyclophosphamide. The clinical effectiveness and economic sections of the submission did not quantitatively consider rituximab or cyclophosphamide as comparators, only SoC was formally assessed.

The MS justifies the exclusion of rituximab as a comparator on the following grounds:

- There has been no head to head trial of rituximab versus belimumab;
- Outcome measures used in rituximab and belimumab trials have differed to the extent that there is little possibility of undertaking meaningful indirect comparison meta-analysis;
- Rituximab has not been shown to be effective versus SoC whereas belimumab has, therefore by implication belimumab is unlikely to be less effective than rituximab;

[redacted]

Regarding effectiveness, although the primary end point was not reached in the Phase II/III rituximab RCT (EXPLORER)[10] the ERG's clinical expert indicated that the EXPLORER end point was more stringent than that in the BLISS trials because it registered a sustained response (once a patient was classified as a non-responder they remained so classified for the remainder of the trial), whereas the primary end point in BLISS was a group response in which a non-responder could later become classified as a responder for the primary end point at week 52.

A literature search undertaken by ERG revealed published information on SLEDAI and SF 36 changes in the EXPLORER trial which might have been used for comparison with the BLISS trials. Furthermore, RCTs for both drugs recorded BILAG changes thus offering the potential for an indirect comparison to be undertaken[10,11]. For these reasons the ERG requested clarification regarding the manufacturer's justification for not considering rituximab as a comparator.

The manufacturer responded with further justification as shown in **Error! Reference source not found.**.

items)[2]. A score is calculated for each system depending on the SLE clinical manifestations (or signs and symptoms) present and whether they are new, worse, the same, improving, or not present in the last 4 weeks compared with the previous 4 weeks. A BILAG A score is given for a disease manifestation considered sufficiently severe to normally require high-dose steroids (prednisolone > 20 mg/day or equivalent) and/or immunosuppressive / cytotoxic agents under normal circumstances. A more moderate manifestation, which it would be considered appropriate to treat with lower dose steroids, antimalarial drugs or NSAIDs, constitutes a BILAG B score. A mild symptomatic manifestation that would require only symptomatic therapy (e.g. analgesics and NSAIDs) constitutes a BILAG C score. If there are no current symptoms, but the system has previously been involved, then a BILAG D score is recorded, while if the system has never been involved, a BILAG E score is assigned.

3.5 Economic analysis

The manufacturer's economic analysis is in line with that stipulated in the NICE scope. The MS presented its economic assessment in terms of incremental cost per QALY and has modelled outcomes using a lifetime horizon. Costs are considered from an NHS and PSS perspective.

3.6 Other relevant factors

Special considerations and issues raised in the manufacturer's scope include: 1) the innovative nature of belimumab for SLE; 2) the inability of the utility method to capture the QoL of SLE patients sufficiently sensitively; and 3) the impact of SLE on particular ethnic groups and on women of childbearing age. The proposed SPC specifies that belimumab should not be administered to pregnant women or to those planning pregnancy and therefore the special consideration relating to women of childbearing age is of relevance.

There were no issues identified in the NICE scope.

A publication describing the industry phase I study LBSL01 was listed as "included" but this study was not discussed in the MS. The manufacturer's stated reason for this is reproduced in Box 1.

Box 1: Reasons for not including study LBS01

As this was a small (n=70) exploratory study of limited duration, designed primarily to demonstrate safety and tolerability in humans, it does not reflect the proposed clinical use of belimumab and therefore will be excluded from further discussion.

4.1.3.1 Important included studies

No RCTs were found that compared belimumab with an alternative active intervention. The most important belimumab studies identified were three industry-sponsored RCTs conducted in adults comparing belimumab plus standard care with placebo plus standard care (trials: LBSL02[13], BLISS-52[7], and BLISS-76) together with an uncontrolled extension (LBSL99) of LBSL02. One rituximab RCT (EXPLORER trial[10]) was included in narrative discussion of potential comparators. Brief details of these studies are shown in Table 1.

Table 1: Important studies included in manufacturer's submission

ID Year ψ	Study type	Study duration	Patient Age, yr	Treatment Groups ¥	N (ITT)	Countries (% enrolled)
LBSL02 2006	Phase 2 Efficacy and Safety	52 wk	20 - 75	Bel 1mg/kg IV** Bel 4mg/kg IV* Bel 10mg/kg IV* Placebo**	114 111 111 113	USA (98%) Canada (2%)
BLISS-76 2009	Phase 3 Efficacy and Safety	76 wk	18 - 73	Bel 1mg/kg IV* Bel 10mg/kg IV* Placebo**	271 273 275	USA and Canada (53%) West Europe (25%) East Europe (11%) Latin America (11%)
BLISS-52 2009	Phase 3 Efficacy and Safety	52 wk	18 - 71	Bel 1mg/kg IV* Bel 10mg/kg IV* Placebo**	288 290 287	Latin America (50%) Asia (38%) East Europe and Australia (13%)
LBSL99 2006	Safety extension of L02			Bel 10mg/kg IV*	296	USA and Canada (100%)
EXPLORER	Phase 2/3 Efficacy and Safety	52 wk	16 - 75	Rit 1000mg# Placebo***	169 88	North America (100%)

ψ Year study subject enrolment ended
¥ All treatments were additional to standard care
* Bel = Belimumab 1, 4, or 10mg/kg administration by IV infusion on days 0, 14, 28, and every 28 days thereafter
** Placebo by IV infusion on days 0, 14, 28, and every 28 days thereafter
*** Placebo by IV infusion on days 1, 15, 167 and 181
# Rit = Rituximab on days 1, 15, 167 and 181

Table 2: Patient eligibility for BLISS-52 and BLISS-76 (From MS Table 5.7)

Trial no. (acronym)	Inclusion criteria	Exclusion criteria
C1057 **(BLISS-52)**	Adult patients (aged ≥ 18 years) who met the American College of Rheumatology criteria for systemic lupus erythematosus and had active disease (score ≥ 6 at screening on SELENA-SLEDAI) were eligible for enrolment. Other inclusion criteria were unequivocally positive ANA (titre ≥ 1:80) or anti-dsDNA antibody (≥ 30IU/mL), and a stable treatment regimen with fixed doses of prednisone (0–40mg/day), or non-steroidal anti-inflammatory, antimalarial, or immunosuppressant drugs for at least 30 days before the first study dose	The main exclusion criteria were severe active lupus nephritis or CNS lupus; pregnancy; and previous treatment with any B-lymphocyte-targeted drug (including rituximab), intravenous cyclophosphamide within 6 months of enrolment, and intravenous Ig or prednisone (>100 mg/day) within 3 months
C1056 **(BLISS-76)**	As per BLISS-52	As per BLISS-52
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee		

Relative to the whole trial population imbalance between treatment arms was more pronounced for the Target population in both trials, especially in BLISS-76 (see Appendix 4 of this report).

4.2.5.2 BLISS trials: demography of patients

Demographic characteristics of patients in BLISS-52 and BLISS-76 were presented in MS Tables 5.9 to 5.11 (see **Error! Reference source not found.** of this report). Patients were mostly young females (74% ≤ 45 years of age; 94% female), a population which is representative of patients with SLE.

Selected characteristics for placebo and 10mg/kg groups taken from MS Table 5.8 are shown below in **Error! Reference source not found.**. Amongst all treatment arms pooled across the two studies 47% of patients were white, 23% American Indian, 21% Asian, and 8.8% black, however there were large differences in the racial makeup between the two studies reflecting the racial distributions in the geographic regions in which the trial centres were located. The substantial differences between trials in geographical and in racial distributions seen for the whole population were mirrored in the Target population **Error! Reference source not found.** (**Error! Reference source not found.**).

Table 3: Baseline SELENA SLEDAI involvement: whole population in BLISS trials

	BLISS-52				BLISS-76			
Condition (weight)	**Placebo (N=287)**	**Belimumab 1mg/kg (N=288)**	**Belimumab 10mg/kg (N=290)**	**Total (N=865)**	**Placebo (N=275)**	**Belimumab 1mg/kg (N=271)**	**Belimumab 10mg/kg (N=273)**	**Total (N=819)**
Organic Brain Syndrome (8)	0	2 (1%)	0	2 (1%)	1 (0%)	2 (1%)	3 (1%)	6 (1%)
Lupus HA (8)	4 (1%)	2 (1%)	4 (1%)	10(1%)	1 (0%)	4 (2%)	9 (3%)	14 (2%)
Vasculitis (8)	20 (7%)	16 (6%)	28 (10%)	64 (7%)	17 (6%)	20 (7%)	10 (4%)	47 (6%)
Arthritis (4)	165 (58%)	169 (59%)	173 (60%)	507 (59%)	206 (75%	193 (71%)	191 (70%	590 (72%)
Hematuria (4)	15 (5%)	16 (6%)	16 (6%)	47(5%)	5 (2%)	7 (3%)	8 (3%)	20 (2%)
Proteinuria (4)	50 (19%)	54(19%)	41 (14%)	145 (17%)	29 (11%)	23 (9%)	26 (10%)	78 (10%)
Rash (2)	176 (61%)	176 (61%)	182 (63%)	534 (62%)	187 (68%)	180 (66%)	154 (56%)	521 (64%)
Alopecia (2)	150 (52%)	138 (48%)	158 (55%)	446 (52%)	30 (47%	137 (51%)	116 (43%	383 (47%)
Mucosal Ulcers (2)	71 (25%)	52 (18%)	58 (20%)	181 (21%)	74 (27%)	57 (21%)	78 (29%)	209 (26%)
Low Complement (2)	183 (64%)	186 (65%)	198 (68%)	567 (66%)	160 (58%	149 (55%)	159 (58%	468 (57%)
Inc. DNA Binding (2)	205 (71%)	220 (76%)	218 (75%)	643 (74%)	175 (64%	168 (62%)	176 (65%	519 (63%)
Leukopenia (1)	18 (6%)	12 (4%)	9 (3%)	39 (5%)	16 (6%)	22 (8%)	23 (8%)	61 (7%)

Table 4: Baseline SELENA SLEDAI involvement: in the Target population in BLISS Trials

	BLISS-52		BLISS-76		Combined BLISS	
Condition (weight)	**Placebo (N=107)**	**Belimumab 10mg/kg (N=112)**	**Placebo (N=96)**	**Belimumab 10mg/kg (N=81)**	**Placebo (N=203)**	**Belimumab 10mg/kg (N=193)**
Organic Brain Syndrome (8)	0	0	1 (1.0%)	0	1 (0.5%)	0
Lupus HA (8)	1 (0.9%)	3 (2.7%)	0	2 (2.5%)	1 (0.5%)	5 (2.6%)
Vasculitis (8)	15 (14.0%)	19 (17.0%)	10 (10.4%)	5 (6.2%)	25 (12.3%)	24 (12.4%)
Arthritis (4)	65 (60.7%)	76 (67.9%)	83 (86.5%)	63 (77.8%)	148 (72.9%)	139 (72.0%)
Hematuria (4)	9 (8.4%)	7 (6.3%)	3 (3.1%)	6 (7.4%)	12 (5.9%)	13 (6.7%)
Proteinuria (4)	31 (29.0%)	28 (25.0%)	17 (17.7%)	21 (25.9%)	48 (23.6%)	49 (25.4%)
Rash (2)	74 (69.2%)	75 (67.0%)	72 (75.0%)	52 (64.2%)	146 (71.9%)	127 (65.8%)
Alopecia (2)	66 (61.7%)	69 (61.6%)	50 (52.1%)	38 (46.9%)	116 (57.1%)	107 (55.4%)
Mucosal Ulcers (2)	28 (26.2%)	20 (17.9%)	30 (31.3%)	22 (27.2%)	58 (28.6%)	42 (21.8%)
Low Complement (2)	107 (100.0%)	112 (100.0%)	96 (100.0%)	80 (98.8%)*	203 (100.0%)	192 (99.5%)
Inc. DNA Binding (2)	107 (100.0%)	112 (100.0%)	96 (100.0%)	81 (100.0%)	203 (100.0%)	193 (100.0%)
Leukopenia (1)	6 (5.6%)	4 (3.6%)	7 (7.3%)	10 (12.3%)	13 (6.4%)	14 (7.3%)

A specified major secondary outcome was the percentage of SRI responders at week 76. There was only a small difference between placebo and 10mg/kg belimumab (odds ratio and P value not submitted; odds ratio 1.31, 95% CI: 0.92 – 1.87, P = 0.1323 by logistic regression, taken from the FDA HGS briefing document.[5]

Relative to the whole trial population imbalance between treatment arms was more pronounced for the Target population in both trials, especially in BLISS-76 (Appendix 4).

Patients from BLISS-52 contributed more patients to the pooled Target population than did patients from BLISS-76 (55% and 45% respectively, and contributed a greater proportion of the patients receiving 10mg/kg belimumab (58% and 42% from each trial respectively); therefore effectiveness results pooled across trials will tend to reflect BLISS-52 outcomes more than BLISS-76.

4.2.5.3 BLISS trial results by outcome

Primary outcome: SRI at week 52

The pre-specified primary outcome in the BLISS trials was the proportion of responders at week 52 defined according to the composite SRI outcome measure. The results were provided in MS Table 5.15 and clarification Table A6.1 and summarised below in Table: 5.

Table: 5 Primary efficacy endpoint (SRI) at Week 52 (dropout-failure)

	BLISS-52		**BLISS-76**		**Pooled Total Population[4]**		**High Disease Activity Subgroup Pooled Total**		**High Disease Activity Subgroup BLISS-52**		**High Disease Activity Subgroup BLISS-76**	
SRI at Week 52	**Placebo N = 287**	**10mg/kg N = 290**	**Placebo N = 275**	**10mg/kg N = 273**	**Placebo N = 562**	**10mg/kg N = 563**	**Placebo N = 203**	**10mg/kg N = 193**	**Placebo N = 107**	**10mg/kg N = 112**	**Placebo N = 96**	**10 mg/kg N = 81**
No. (%) Response	125 (43.6%)	167 (57.6%)	93 (33.8%)	118 (43.2%)	218 (38.8%)	285 (50.6%)	77 (37.9%)	121 (62.7%)	44 (41.1%)	75 (67.0%)	33 (34.4%)	46 (56.8%)
Observed difference vs placebo (%)	-	14.03	-	9.41	-	11.8	-	24.8	-	25.9	-	22.4
OR (95% CI)[1] vs placebo	-	1.83 (1.30, 2.59)	-	1.52 (1.07, 2.15)	-	1.68 (1.3, 2.2)	-	2.7 (1.8, 4.1)	-	3.0 (1.7, 5.2)	-	2.5 (1.3, 4.6)
P-value[1]	-	0.0006	-	0.0207	-	< 0.0001	-	< 0.0001		0.0001	-	0.0045

[1] Odds Ratio (95% confidence interval) and p-values were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, covariates include baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.

In both trials SoC + 10mg/kg belimumab delivered a greater percentage of responders than SoC + placebo. The difference in percentage of responders in the belimumab group relative to placebo group for the whole population was 14% in BLISS-52 and 9.4% in BLISS-76. The corresponding adjusted odds ratios for a response in BLISS-52 and in BLISS-76 were respectively 1.83 (95% CI: 1.30, 2.59; P = 0.0006) and 1.52 (95% CI: 1.07, 2.15; P = 0.0207). For the Target population pooled across trials the difference in percentage of responders in the belimumab group relative to placebo group was 24.8% and the adjusted odds ratio was 2.7 (95% CI: 1.8, 4.1; P < 0.0001). In BLISS-52 and BLISS-76 Target populations the difference

Modified SRI

To be classified as an SRI responder a patient is required to have a SELENA-SLEDAI score that is reduced by ≥ 4 points relative to baseline. A 4 point reduction in SELENA-SLEDAI can be achieved by normalisation of serological manifestations only (e.g. anti-dsDNA antibodies and complement). The MS presented an analysis of a modified SRI response in which the increased DNA binding and low complement items were removed from the SELENA-SLEDAI component of the SRI; the analysis was performed in patients who still had a SELENA SLEDAI score ≥ 4 at baseline after points for low complement and increased DNA binding were removed from the scale. During the clarification process the manufacturer provided modified SRI results for the Target or high disease activity population; these plus the information from the MS Page 111 are summarised in Table 13 Error! Reference source not found.

Table 6 Duplicate table removed

The MS did not specify patient numbers for this analysis and so data from the HGS Briefing Document to the FDA.[5] Figure 1 shows the percentage of modified SR responders (from HGS Briefing Document to FDA).

In the HGS/FDA[5] analysis there is little difference in response between 1mg/kg and 10mg/kg groups for BLISS-52.

The number of patients at risk was not specified. A stronger response was observed for the Target populations than for the total populations and statistical significance was reached in both trials.

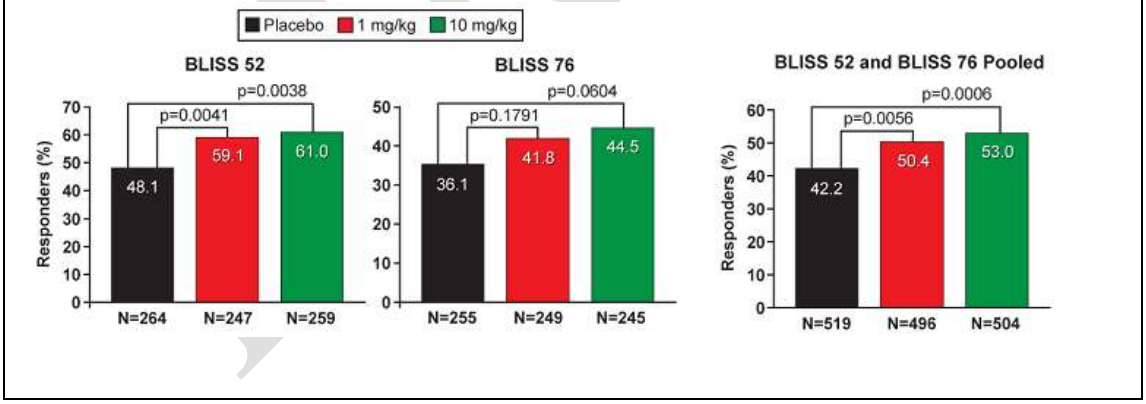


Figure 1: Modified SR percentage of responders (from HGS Briefing Document to FDA)

Subcomponents of the SRI response

Table 7 summarises the week 52 results for the three subcomponents of the composite SRI response (based on MS Table 5.16 and clarification Table A6.1).

Table 7: Results for subcomponents of SRI at week 52 (adjusted)

	BLISS-52		BLISS-76		Pooled Total Population[4]		High Disease Activity Subgroup Pooled Total		High Disease Activity Subgroup BLISS-52		High Disease Activity Subgroup BLISS-76	
	Placebo N = 287	10mg/kg N = 290	Placebo N = 275	10mg/kg N = 273	Placebo N = 562	10mg/kg N = 563	Placebo N = 203	10mg/kg N = 193	Placebo N = 107	10mg/kg N = 112	Placebo N = 96	10mg/kg N = 81
4-point reduction in SELENA-SLEDAI	132 (46.0%)	169 (58.3%)	98 (35.6%)	128 (46.9%)	230 (40.9%)	297 (52.8%)	84 (41.4%)	125 (64.8%)	47 (43.9%)	76 (67.9%)	37 (38.5%)	49 (60.5%)
Observed difference vs placebo (%)	-	12.3	-	11.3	-	11.9	-	23.4	-	24.0		22.0
OR (95% CI)[1] vs placebo	-	1.71 (1.21,2.41)	-	1.63 (1.15,2.32)	-	1.68 (1.3,2.2)	-	2.6 (1.7,3.9)	-	2.8 (1.6,4.8)	-	2.4 (1.3,4.4)
P-value[1]		0.0024		0.0062	-	< 0.0001	-	< 0.0001		0.0004	-	0.0063
No New 1A/2B BILAG domain scores	210 (73.2%)	236 (81.4%)	179 (65.1%)	189 (69.2%)	389 (69.2%)	425 (75.5%)	125 (61.6%)	145 (75.1%)	68 (63.6%)	88 (78.6%)	57 (59.4%)	57 (70.4%)
Observed difference vs placebo (%)	-	8.2	-	4.1	-	6.3	-	13.5	-	15.0	-	11.0
OR (95% CI)[1,2] vs placebo	-	1.62 (1.09,2.42)	-	1.20 (0.84,1.73)	-	1.4 (1.1,1.8)	-	1.9 (1.2,3.0)	-	2.3 (1.2,4.2)	-	1.6 (0.9, 3.1)
P-value[1,2]		0.0181		0.3193	-	0.0190	-	0.0034	-	0.0099	-	0.1297
No worsening in PGA	199 (69.3%)	231 (79.7%)	173 (62.9%)	189 (69.2%)	372 (66.2%)	420 (74.6%)	119 (58.6%)	142 (73.6%)	64 (59.8%)	86 (76.8%)	55 (57.3%)	56 (69.1%)
Observed difference vs placebo (%)	-	10.4	-	6.3	-	8.4	-	15.0	-	17.0	-	11.8
OR (95% CI)[1,3] vs placebo	-	1.74 (1.18,2.55)	-	1.32 (0.92,1.90)	-	1.5 (1.2,2.0)	-	2.0 (1.3,3.1)	-	2.3 (1.3,4.2)	-	1.6 (0.9,3.0)
P-value[1,3]	-	0.0048	-	0.1258	-	0.0017	-	0.0015	-	0.0063	-	0.1312

[1] Odds Ratio (95% confidence interval) and p-value were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, covariates include baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate

[2] Additional covariate: baseline BILAG domain involvement (at least 1A/2B)

[3] Additional covariate: baseline PGA score

[4] No significant treatment-by-study interactions were observed (all p > 0.287)

The three subcomponents of the composite SRI outcome were: [i] an improved SELENA SLE DAI score by ≥ 4 points; [ii] a BILAG index showing no new grade A organ involvement or no two grade B organ involvements (i.e. no worsening by one new A or two new B BILAG indices); [iii] a PGA score that has not increased by more than 0.3 points (i.e. no worsening in PGA by ≥ 0.3).

The percentage of patients at week 52 that achieved a SLEDAI score reduction of ≥ 4 points was defined as a major secondary outcome. For the whole population, both trials delivered more responders in the belimumab group than the placebo group (P = 0.0024 and P = 0.0062 for BLISS-52 and BLISS-76, respectively).

Results at week 52 for the other two SRI subcomponents (i.e. no worsening in BILAG index and no worsening in PGA score) were defined as non-major secondary outcomes. The percentage of patients in the whole population that satisfied the BILAG and PGA criteria in BLISS-52 was greater for belimumab relative than placebo (significant at P = 0.0181 and P = 0.0048 for BILAG and PGA, respectively); however, for BLISS-76 the differences between belimumab and placebo were considerably smaller and neither component reached statistical significance in favour of belimumab (P = 0.319 and P = 0.1258 for BILAG and PGA, respectively). According to results reported in the HGS Briefing Document to the FDA (Table 9.20, Page 102) the 1mg/kg belimumab dose regimen in BLISS-76 performed slightly better than 10mg/kg for both the PGA and BILAG subcomponents at week 52.

The corresponding results for the target population supplied during the clarification process are also summarised in Table 7 Pooled across trials, all three SRI components at week 52 were supportive of belimumab relative to placebo and delivered significant effects. However, for BLISS-76 the PGA and BILAG results at week 52 for the target population were considerably weaker (P = 0.1312 and P = 0.1297, respectively) than for BLISS-52 or the pooled target population.

Major secondary outcomes

The MS identified five pre-specified major secondary outcomes. These included the SRI response at week 76 and the percentage of patients with a ≥ 4 point SLEDAI improvement at week 52, each of which have been discussed in the preceding sections. The other three major secondary outcomes were: mean change in PGA score at week 24, percentage of patients with prednisone reductions ≥ 25% from baseline to ≤ 7.5 mg/day during weeks 40 to 52 (in subjects whose baseline dose was > 7.5 mg/day); mean change in SF36 PCS at week 24. These are discussed in this section.

Change in PGA score at week 24 was presented in MS Table 5.18 and the relevant results from this are shown in Table 8 below. For the whole population in BLISS-52 the change in PGA score (week 24 relative to baseline) for both groups indicated disease improvement and was greater in the belimumab group (-0.54) than placebo group (-0.39; P = 0.0003 in support of belimumab). For BLISS-76 the difference between groups was very small and in favour of placebo (-0.49 placebo and -0.44 belimumab) and did not reach statistical significance (P = 0.7987). For the Target HDAP pooled across trials belimumab delivered a greater reduction in PGA score than placebo (P = 0.0268 with mean changes of -0.42 and -0.52 for placebo and belimumab, respectively). Target population results by trial are not available.

Table 8: Mean change in PGA score at week 24 (taken from MS Table 5.18)

Major secondary endpoint at Week 24	BLISS-52		BLISS-76		Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 287	10mg/kg N = 290	Placebo N = 275	10mg/kg N = 273	Placebo N = 562	10mg/kg N = 563	Placebo N = 203	10mg/kg N = 193
Mean ± SE	-0.39 ± 0.03	-0.54 ± 0.03	-0.49 ± 0.04	-0.44 ± 0.03	-0.44 ± 0.02	-0.49 ± 0.02	-0.42 ± 0.04	-0.52 ± 0.04
LS Mean ± SE[1]	-0.35 ± 0.04	-0.50 ± 0.04	-0.49 ± 0.05	-0.48 ± 0.05	-0.40 ± 0.03	-0.48 ± 0.03	-0.41 ± 0.05	-0.53 ± 0.05
P-value[1]	-	0.0003	-	0.7987	-	0.0167	-	0.0268

[1] All statistics, including the difference in LSM (least square means), were from ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for the baseline PGA score, baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.

The mean change in PGA at week 52 was submitted as an additional secondary outcome. The results are shown in Table 9.

Table 9: Mean change in PGA score at week 52 (taken from MS Table 5.18)

Other secondary endpoints Week 52	BLISS-52		BLISS-76		Pooled Total Population		High Disease Activity Subgroup	
	Placebo N = 287	10mg/kg N = 290	Placebo N = 275	10mg/kg N = 273	Placebo N = 562	10mg/kg N = 563	Placebo N = 203	10mg/kg N = 193
Mean ± SE	-0.48 ± 0.04	-0.67 ± 0.04	-0.46 ± 0.04	-0.49 ± 0.04	-0.47 ± 0.03	-0.58 ± 0.03	-0.41 ± 0.05	-0.62 ± 0.05
LS Mean ± SE[1]	-0.38 ± 0.05	-0.57 ± 0.05	-0.47 ± 0.06	-0.55 ± 0.06	-0.40 ± 0.04	-0.54 ± 0.04	-0.36 ± 0.06	-0.59 ± 0.06
P-value[1]	-	0.0001	-	0.1159	-	< 0.0001	-	0.0003

[1] All statistics, including the difference in LSM (least square means), were from ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for the baseline PGA score, baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate.

In BLISS-52 a larger improvement in PGA score was observed for the 10mg/kg group than for placebo (P = 0.0001) whereas in BLISS-76 the difference between treatments was trivial (P = 0.1159). For the pooled populations 10mg/kg was superior to placebo (P = 0.0003).

The HGS Briefing Document to the FDA[5] provided graphed results for mean change in PGA through successive clinic visits for all three randomised groups. These are shown below in **Error! Reference source not found.** for BLISS-76.

Table 10: Prednisone reduction Weeks 40 through 52 – Phase 3 trials

	BLISS-52		BLISS-76		Pooled Total Population[4]		High Disease Activity Subgroup Pooled Total		High Disease Activity Subgroup BLISS-52		High Disease Activity Subgroup BLISS-76	
	Placebo N = 192	10 mg/kg N = 204	Placebo N = 126	10 mg/kg N = 120	Placebo N = 318	10 mg/kg N = 324	Placebo N = 126	10 mg/kg N = 126	Placebo N = 76	10 mg/kg N = 81	Placebo N = 50	10 mg/kg N = 45
No. %[1] Response[2]	23 12.0%	38 18.6%	16 12.7%	20 16.7%	39 12.3%	58 17.9%	9 7.1%	20 15.9%	4 5.3%	15 18.5%	5 10.0%	5 11.1%
Observed difference vs Placebo	-	6.65	-	3.97	-	5.64	-	8.73	-	13.2	-	1.1
OR (95% CI)[3] vs placebo	-	1.75 (0.99, 3.08)	-	1.26 (0.61, 2.60)	-	1.57 (1.01, 2.45)	-	2.43 (1.05, 5.65)	-	4.11 (1.29, 13.2)	-	0.88 (0.21, 3.60)
P-value[3]	-	0.0526	-	0.5323	-	0.0451	-	0.0389	-	0.0171	-	0.8586

[1] Includes only subjects with baseline prednisone > 7.5 mg/day

[2] Any subject who withdrew from the study prior to the Day364 (Week 52) visit, missed the Day 364 (Week 52) visit (± 28 day window allowed), and/or received a protocol-prohibited medication or a dose of allowable (but protocol-restricted) medication that resulted in treatment failure designation prior to the Day 364 (Week 52) visit was considered a treatment failure for prednisone reduction

[3] Odds Ratio (95% confidence interval) and p-value were from logistic regression for the comparison between each belimumab dose and placebo with covariates. For individual studies, the covariates include baseline prednisone level, baseline SELENA-SLEDAI score (≤ 9 vs ≥ 10), baseline proteinuria level (< 2 g/24 hour vs ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). For pooled data analysis, study was also included as an additional covariate

[4] Obtained from a logistic regression by adding study and the treatment-by-study interaction to the above model

In BLISS-52 and BLISS-76 at baseline 68.6% and 44.9% of patients respectively were receiving > 7.5 mg/day prednisone. The percentage that reduced steroid use in weeks 40 to 52 by the pre-specified amount was greater in the belimumab arm than the placebo arm in both trials, however the difference (belimumab vs. placebo) failed to reach statistical significance in either trial: 18.6% vs. 12.0% in BLISS-52 (P = 0.0526 from logistic regression including baseline covariates) and 16.7% vs. 12.7% in BLISS-76 (P = 0.5323).

For the Target or HDAP pooled across trials 15.9% and 7.1% reduced steroid use in the 10mg/kg belimumab and placebo groups respectively (P = 0.0389 from logistic regression). The results from BLISS-52 supported belimumab (P = 0.171) whereas in BLISS-76 differences between treatments were trivial (P = 0.8586). The HGS Briefing Document to the FDA[5] provided results for reduction in steroid use for all three treatment arms. Table 9-16 from the HGS Briefing Document is shown in **Error! Reference source not found.**.

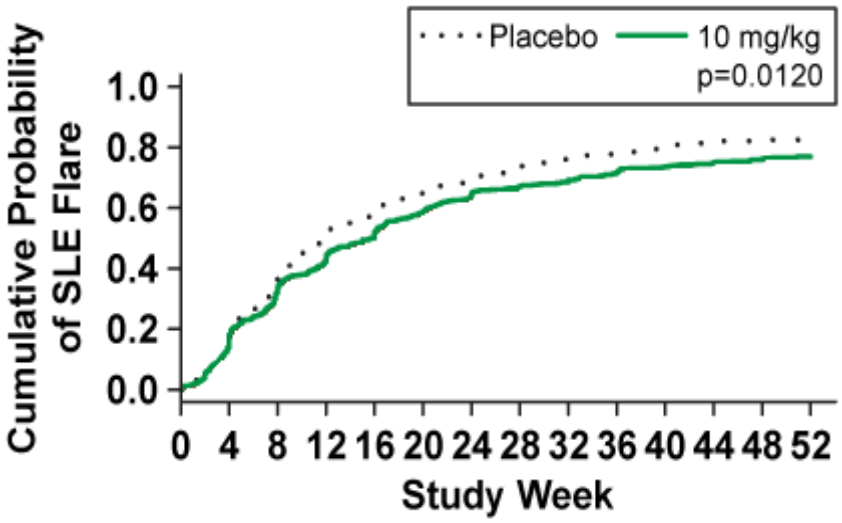


Figure 2: Time to first flare; pooled whole populations (Taken from MS Figure 5.12)

For the high disease activity Target population pooled across trials, belimumab significantly delayed time to first flare relative to placebo (P = 0.0017; Figure 3).

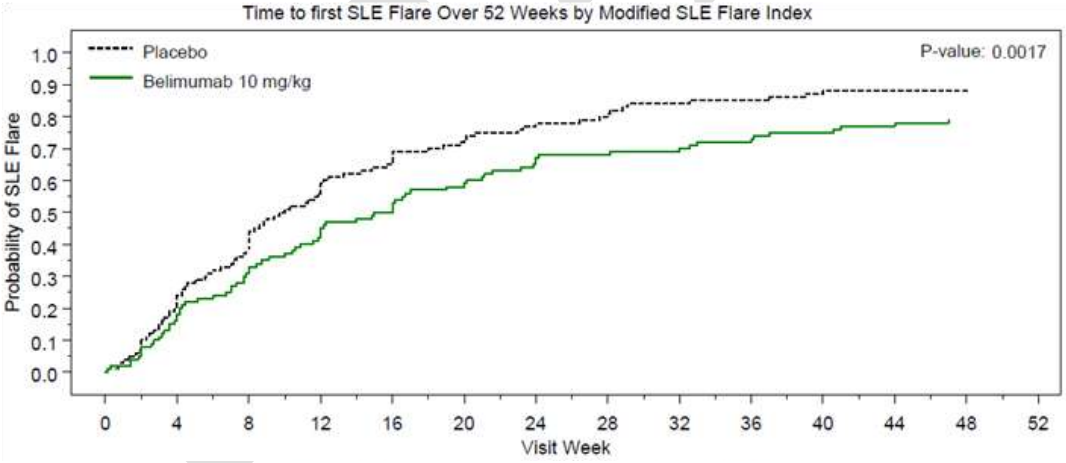


Figure 3: Time to first flare; pooled Target populations (Taken from MS Figure 5.13)

In BLISS-52 the time to first severe flare was delayed by 10 mg/kg belimumab relative to placebo P = 0.0055; Figure 4).

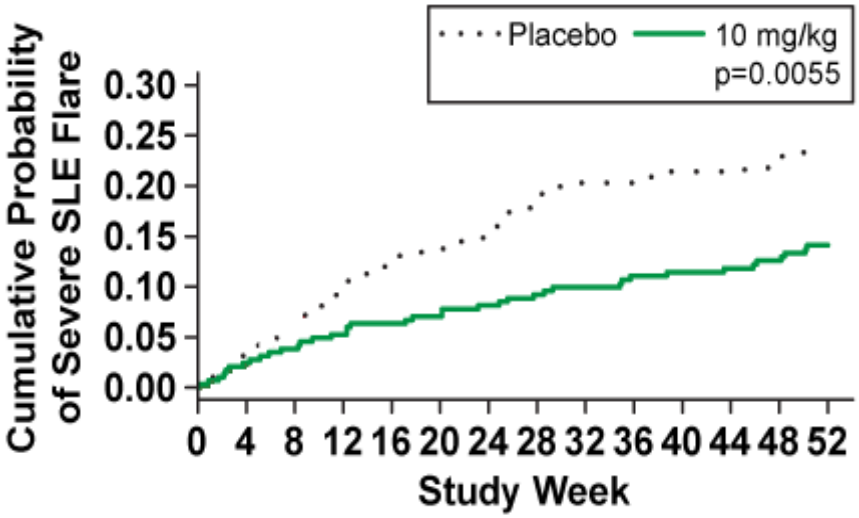


Figure 4: Time to first severe flare; BLISS-52 (Taken from MS Figure 5.10)

In BLISS-76 belimumab somewhat delayed time to first severe flare in (HR 0.72, 95% CI 0.50–1.05, P = 0.0867; Figure 5).

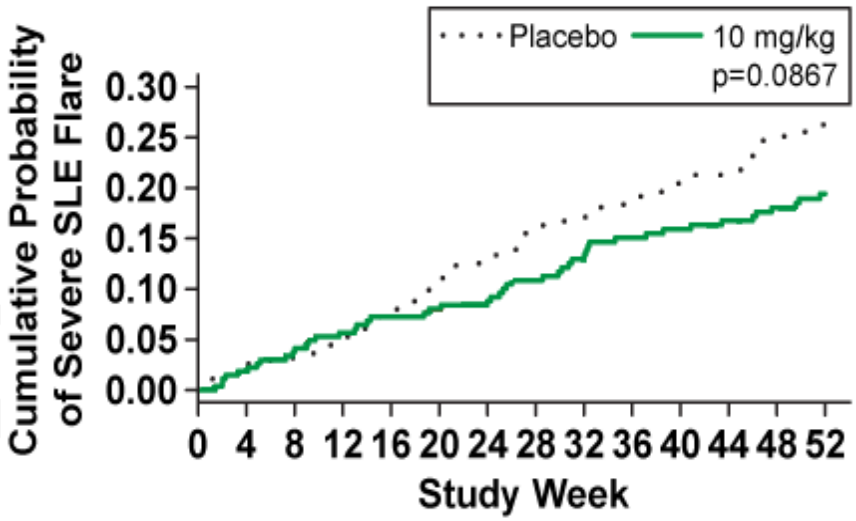


Figure 5: Time to first severe flare; BLISS-76 (Taken from MS Figure 5.11)

When the whole populations from the BLISS trials were pooled the difference between treatments for time to first severe flare reached statistical significance (P = 0.0011; **Error! Reference source not found.**).

the reverse was the case. The BLISS-76 result is shown in Figure 6 together the mean change in SF-36 Vitality domain score.

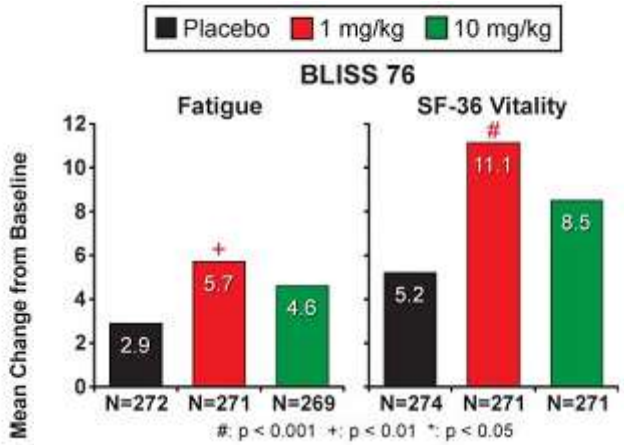


Figure 6: Mean change in FACIT and SF-36 vitality score by week 52 (Taken from HGS Briefing Document to FDA see Figure 9-35)

EQ-5D

There was no significant difference between belimumab and placebo in the absolute change of EQ-5D score from baseline in either trial or pooled total populations during clinic visits. The results for the 10 mg/kg belimumab and placebo groups in BLISS-76 were indistinguishable. For the pooled target population the difference between 10 mg/kg and placebo groups reached statistical significance in favour of belimumab at week 24 ($P \leq 0.01$), but the difference had almost completely faded by week 52 MS Figure 5.21 Page 135).

Results for the mean change in SELENA SLEDAI score from baseline at week 52 were submitted in MS Table 5.17 (Page 113) and clarification response Table A6.1 and are summarised in Table 11. There was no significant difference between belimumab and placebo in the absolute change of EQ-5D score from baseline in either trial or pooled total populations during clinic visits. The results for the 10mg/kg belimumab and placebo groups in BLISS-76 were indistinguishable.

Results for the mean change in SELENA SLEDAI score from baseline at week 52 were submitted in MS Table 5.17 (Page 113) and clarification response Table A6.1 and are summarised in Table 11.

Table 11: Mean change and mean percent change in SLEDAI score week 52

	BLISS-52		BLISS-76		Pooled Total Population[2]		High Disease Activity Subgroup Pooled Total		High Disease Activity Subgroup BLISS-52		High Disease Activity Subgroup BLISS-76	
Change from Baseline at Week 52	**Placebo N = 287**	**10mg/kg N = 290**	**Placebo N = 275**	**10mg/kg N = 273**	**Placebo N = 562**	**10mg/kg N = 563**	**Placebo N = 203**	**10mg/kg N = 193**	**Placebo N = 107**	**10 mg/kg N = 112**	**Placebo N = 96**	**10 mg/kg N = 81**
Mean change from baseline (± SE)	-3.57 ± 0.24	4.97 ± 0.27	-2.77 ± 0.25	-3.70 ± 0.27	-3.18 ± 0.18	-4.36 ± 0.19	-4.1 ± 0.3	-5.8 ± 0.3	-4.1 ± 0.4	-6.3 ± 0.5	-4.0 ± 0.5	-5.2 ± 0.5
P-value[1]	-	< 0.0001	-	0.0063	-	< 0.0001	-	0.0005	-	0.0008	-	0.1705
Mean % change (± SE)	-34.76 ± 2.50	-45.60 ± 2.45	-25.97 ± 2.72	-35.94 ± 2.80	-30.47 ± 1.85	-40.93 ± 1.86	-30.5 (2.3)	-45.5 (2.4)	-	-	-	-
P-value[1]	-	0.0018	-	0.0073	-	< 0.0001	-	< 0.0001	-	-	-	-

[1] ANCOVA model for the comparison between each belimumab dose and placebo, adjusted for baseline SELENA SLEDAI score (≤ 9 vs. ≥ 10), baseline proteinuria level (< 2 g/24 hour vs. ≥ 2 g/24 hour equivalent) and race (African descent or indigenous-American descent vs. other). For pooled data analysis, study was also included as an additional covariate

[2] No treatment-by-study interactions observed (all p-values > 0.367)

Both absolute SLEDAI score reduction from baseline, and percent reduction relative to baseline score, were greater for the 10mg/kg group than for the placebo group; this was consistent and significant for the whole BLISS population (separately by trial and for pooled populations) and for the pooled Target or high disease activity population. For the whole population, results favoured belimumab more strongly in BLISS-52 than BLISS-76. The by-trial results for the Target population are shown below. They indicate stronger support for belimumab in BLISS-52 in which the difference between groups in absolute reduction in SLEDAI score was about double that in BLISS-76 in which the difference between groups was not significant (P = 0.1705).

The HGS Briefing Document to the FDA[5] (see **Error! Reference source not found.**) showed the percentage change in SLEDAI score (relative to baseline) throughout the two trials; this is reproduced in Figure 21.

all patients had a history of auto-immunity, at recruitment ~30% currently lacked anti-nuclear antibodies.

There were 15 deaths during the controlled phase of the three trials; 3 in the placebo group (n=675), and 12 in the belimumab groups (n=1458) with 6 each in the 10mg/kg and 1mg/kg groups respectively. One death in the 1mg/kg belimumab group followed 15 weeks after the patient stopped belimumab treatment. The causes of death were various and are listed in **Error! Reference source not found.** (based on FDA Briefing Package, Table 34).[3] When deaths are rated according to exposure these results translate to: 0.43/100 patient years for placebo (95% CI: 0.09, 1.27) and 0.79/100 patient years for belimumab (95% CI: 0.41, 1.38).[3] There were two completed suicides in the belimumab groups (none in placebo); a further suicide was observed during the LBSL99 extension study. These were not judged to be associated with belimumab since the patients concerned had a history of depression and SLE is associated with an increased risk of depression and suicide.

Adverse events

In all treatment groups > 90% of patients experienced at least 1 AE. The most commonly occurring AEs were headache, upper respiratory tract infection, arthralgia, nausea, UTI, diarrhoea and fatigue.

The percentage of patients experiencing at least one serious AE and at least one serious AE was very similar between placebo and belimumab groups ranging 13.5% to 18.6%, there was a very slight numerical excess with belimumab. The most frequent serious AEs (≥ 1% in any treatment group) were pneumonia, pyrexia, UTI, cholelithiasis, and cellulitis. The percentage of patients experiencing at least one severe AE was 15.4% for the placebo group and 16% across the belimumab groups; the most common severe adverse events were not identified.

Infections

Infections occurred slightly more frequently in patients treated with belimumab compared to placebo. The most frequent infections were URTI, UTI, nasopharyngitis, sinusitis, and bronchitis.

Infusion / hypersensitivity reactions

Occurrence of infusion plus hypersensitivity reactions was similar between belimumab and placebo-treated patients (17% and 14.7%, respectively). Of 1458 belimumab treated patients, 14 experienced hypersensitivity reactions on the day of infusion compared to one of 675 placebo-treated patients.[3] Five discontinuations resulted from hypersensitivity reactions amongst 1458 belimumab patients and none among 675 patients receiving placebo.

The most frequent infections were URTI, UTI, nasopharyngitis, sinusitis, and bronchitis. Of these, nasopharyngitis and bronchitis occurred more commonly with belimumab treatment compared to placebo. Two opportunistic infections occurred, both in the belimumab 10mg/kg group: disseminated CMV infection on day 62; and an Acinetobacter bacteremia on day 15. Four infections were related to deaths: sepsis (placebo group); infectious diarrhea (belimumab 10mg/kg group); cutaneous infection leading to sepsis (belimumab 10mg/kg group); and cellulitis leading to sepsis (belimumab 1mg/kg group).[3]

4.2.6 Pooling of trial data

NICE requests that: "*For each outcome for each included RCT, the following information should be provided*" …."*The size of the effect*" … "*The unit of measurement*". The submission pooled results from two trials, both for the whole BLISS populations and for the Target populations. The manufacturer considered that the pooled trial results were most

providing stronger results than BLISS-76; furthermore there were inconsistencies with regard to an expected dose response relationship.

The ERG is concerned that the pooled results are mainly driven by those from the BLISS-52 trial (conducted in Pacific-Asia and South America) while results in the BLISS-76 trial, conducted in North America and Europe, were only marginally in favour of belimumab relative to placebo and reached statistical significance only for two overlapping outcomes (SRI responders at week 52, and percentage of patients with ≥ 4point reduction in SLEDAI score at week 52 which itself is a component of the SRI). The extent to which these concerns extend to the target population was not possible to gauge from the initial submission because only pooled results were presented. The ERG therefore requested clarification on trial specific target populations and the manufacturer's justification for pooling across trials.

For the target population there was again a greater contribution from BLISS-52 to the pooled results both in terms of number of patients (BLISS-52 contributed 55% of whole target population and 58% of those that received belimumab) and in effectiveness (BLISS-52 provided greater effect sizes compared to BLISS-76 for SRI week 52, modified SRI week 52, percentage with SLEDAI reduction by ≥ 4points, SLEDAI mean change by week 52, no new BILAG 1A/2B, no worsening in PGA and reduction in steroid usage weeks 40 to 52). The ERG therefore remain concerned that the pooled trial results dilute the rather less positive findings most relevant to the decision problem by including additional data from a less relevant population, and that target population results by trial should have been included in sensitivity analysis in the economic model.

Trial baseline SLEDAI scores used in economic model

The manufacturer's economic analysis (section 5) made use of data from an SLE cohort studied at JHU so as to model cost effectiveness of belimumab for the whole and Target populations from the BLISS trials. The JHU cohort experienced relatively mild disease compared to patients in BLISS and particularly compared to the BLISS the high disease activity Target population. During the clarification process the ERG requested the distribution of SLEDAI scores at year one and last year of observation for patients in the JHU cohort. The SLEDAI scores shown in **Error! Reference source not found.**illustrates the differences between Target and JHU populations (year one and last year scores are shown for JHU cohort, Figure B17.2 of the clarification document).

4.2.7 Conclusions

Efficacy evidence came from two multicentre international industry sponsored RCTs (BLISS-52 and BLISS-76) comparing SoC plus belimumab with SoC plus placebo; each trial had three arms: placebo, 1mg/kg belimumab and 10mg/kg belimumab dose regimens. Data for the 1mg/kg arms was excluded from the submission, but results available in the public domain were considered in the ERG's assessment. Outcomes for six populations were presented: whole populations from BLISS trials, whole populations pooled across BLISS trials, Target population from BLISS trials and Target populations pooled across BLISS trials.

The Target population was a high disease activity subgroup identified from post hoc exploration of effectiveness of the primary outcome. There were more noticeable within-trial baseline imbalances (10mg/kg vs. placebo) for the Target population than for the whole population. The Target population results are not necessarily equivalent to those that would be obtained from a randomised trial in this population.

The primary outcome was specified as the percentage of responders at week 52 according to a novel composite disease activity measure (SRI). This outcome was statistically in favour of belimumab (10mg/kg vs. placebo) in both trials. For both whole and Target populations, results from BLISS-52 were more favourable for belimumab than results from BLISS-76.

For all non-major secondary outcomes in BLISS-76 effect sizes were insufficient to be confident that effects could not be accounted for by chance. For several outcomes, including percentage responders and time to first flare in BLISS-76, although formal statistical tests were not performed, it appeared that the 1mg/kg dose regimen was as effective, or more effective, than the 10mg/kg dosage.

Geographical and racial differences between BLISS trials indicate that efficacy results from BLISS-76, rather than from BLISS-52 or pooled BLISS populations, are more generalisable to the UK.

On most safety outcomes placebo and belimumab performed equally. There were more deaths under belimumab than placebo; on a "per patient year of exposure" basis the rate for belimumab was about double that for placebo although this finding could have occurred by chance. Causes of death were various and most were those associated with the condition of SLE. There was insufficient evidence to determine if there was any association between belimumab and mortality.

JHU cohort data elements 2.a.and 2.b. Evolution of SS and AMS scores after week 52

For the evolution of the SS score beyond week 52 the manufacturer originally estimated the regression equation:

But the manufacturer views this as providing a relatively poor fit to the data from the belimumab phase II trial, as shown in Figure 24. For the modelling the manufacturer adjusts the constant and applies the equation:

on the grounds of it better reflecting the phase II trial data.

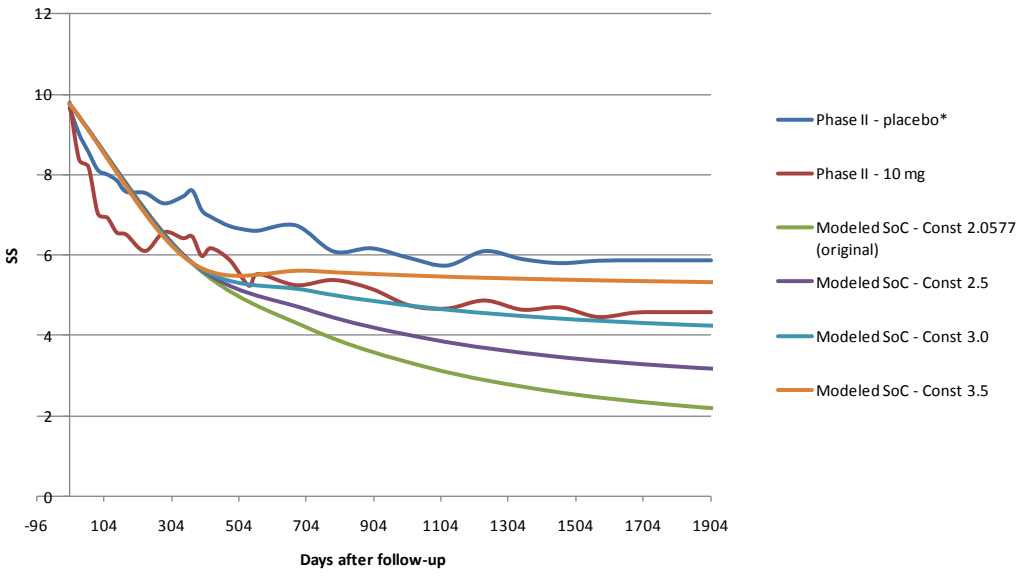


Figure 7: Medium term SS natural history model

The AMS score was developed to measure disease severity over time as opposed to the SS score which only reflects disease activity over the preceding 10 days. AMS is calculated as the area under the curve of disease activity measurements between two time-points. The area under the curve is then divided by time of follow-up to provide an average score over the period of interest. In Figure 24: Medium term SS natural history modelabove the ERG assume that the area under the curves shown can be used to represent AMS. However, the MS references to AMS score may refer to either the "AMS over lifetime" or the "average mean SLEDAI up to current time" which are presumably calculated in the same manner: the area under the SS score curve divided by time elapsed.

Table 12: HRQoL calculation pulmonary involvement from Table 16.19

SLICC Element	HRQoL	JHU %	Weighted	JHU SLICC	Final
Pulmonary hypertension	0.61	33%	0.20		
Pulmonary fibrosis	0.73	42%	0.31		
Shrinking lung (Chest XRay)	1.00	2%	0.02		
Pleural fibrosis (Chest XRay)	1.00	20%	0.20		
Pulmonary infarction/resection	0.94	4%	0.04		
Average across pulmonary			0.77	1.31	0.70

These organ involvement HRQoL values are applied multiplicatively. For a patient having developed more than one SLICC organ involvement, only the lowest HRQoL multiplier is applied to the "clean" utility.

Literature element 3.c Cost impact of organ damage

A similar approach is undertaken for the cost impacts of organ damage as for the QoL impacts, only with the number of patients in the JHU cohort experiencing the individual elements among those having had an event within the organ class giving rise to the weight to apply. These weights can sum to more than one due to a patient being able to experience more than one event. As with the calculation of the quality of life impacts this will tend to overestimate costs in the incident year and early years after incidence.

As these cost elements are less well documented in the submission than the HRQoL elements the full set is outlined below, with more detail being available in Appendix 28 of the submission. There are some minor discrepancies between the figures in **Error! Reference source not found.** and those given in Table 6.26 of the MS for reasons that are unclear, but these will not affect results.

10mg/kg; it seems important that if this should happen that good data on effectiveness of reduced dose regimens should be collected.

Target population and proposed licence population

The focus of the MS was the high disease activity "Target population" which represents a subgroup of the proposed "licence population" (in turn a subpopulation of the pooled BLISS population). The primary end point, which was the percentage of responders at week 52 according to the novel composite SRI outcome measure, was very similar for pooled Target population and pooled "licence population" with respectively 24.8% and 19.8% extra responders for belimumab compared to placebo (belimumab vs. placebo odds ratio = 2.7 for both populations). Furthermore the cost-effectiveness of belimumab in each population was essentially the same (base case ICER £64,410 and £66,170 / QALY respectively). Given these results, there appear small grounds on which to distinguish patients in the Target population from those in the proposed licence population on the basis of either clinical or cost-effectiveness and a SLEDAI score cut-off of 10 points, appears to be an arbitrary criterion that would be difficult to implement in practice. One effect of selecting the Target population in preference to the "licence population" is to considerably reduce the manufacturer's calculation of total budget impact of introducing belimumab across the country. (MS section 7).

Belimumab vs. rituximab

No head-to-head trial comparing belimumab with rituximab has been conducted. The ERG and the manufacturer disagree about the commonality of outcome measures available from belimumab and rituximab trials, but concur that a credible indirect comparison is not feasible on the grounds of large difference between trial populations. The ERG note that the primary outcome measure in the relevant Rituximab trial may be a more stringent test of therapeutic effect than that used in the BLISS trials, and therefore are not convinced by the manufacturer's implication that belimumab is necessarily a more effective drug.

Efficacy of belimumab for different SLE manifestations

In the BLISS trials the most commonly involved SLE manifestations were musculoskeletal (60%), mucocutaneous (59%), hematologic (16%), general (11%), renal (11%) and vasculitis (7%). Direct evidence for a beneficial effect of belimumab on other manifestations, such as pulmonary, renal or central nervous system manifestations, is not available.

- responders, and of the reasonableness of extrapolating using this value. A low discontinuation rate worsens the cost effectiveness of belimumab
- The requirement to adjust the JHU cohort survival model by SMRs from the literature is unclear and may have tended to exaggerate the impact of the individual covariates within the JHU cohort survival model
- The analysis of the observational cost data on a six monthly basis in order to relate it to the maximum SS score during that period then doubling it to arrive at the annual relationship appears peculiar given that the observational cost data was collected over a year. It may also lead to bias
- The separate estimation of a cost per organ involved may have double counted costs estimated within the SS score cost function to some degree

6.2 Implications for research

It is unlikely that an industry sponsored trial will be conducted to compare belimumab with rituximab or other new biological interventions for SLE. The cost of a sufficiently powered study to discriminate between such treatments is likely to be too great for such studies to be undertaken independently of industry sponsorship. In view of the relative expense of belimumab and the lack of clear demonstration of a dose response relationship it is possible that in the real world belimumab may be employed at doses less than 10mg/kg. Useful research could be undertaken to monitor such usage and the 24 week response rates elicited.

Due to the paucity of long-term evidence for the continued benefit of belimumab and its safety, monitoring and surveillance of patients who have been treated with belimumab are therefore necessary. Further investigation is needed in patients excluded in the current BLISS-52 and BLISS-76 trials who had severe lupus nephritis or central nervous system manifestations of the disease. The two trials were limited in the inclusions of black patients, who for example account for approximately 25% of lupus patients in the USA. These patients also tend to have more severe disease than the general lupus population. In an earlier Phase II study of belimumab, black patients did significantly better than non-black patients. In contrast the reported Phase III trials found black patients treated with belimumab performed worse than those given placebo. These discrepancies needed to be considered further.

Although BLyS (B-Lymphocyte stimulator) is raised in SLE, reducing its activity with belimumab in SLE patients appears to have only a very modest effect. In RCTs a large proportion of patients in the belimumab group responded, but the placebo group response indicated that many would have responded irrespective of receiving belimumab. In a