STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Ipilimumab for previously treated unresectable malignant melanoma

This report was commissioned by the NIHR HTA Programme as project number 08/209/01

Completed August 19, 2011

CONTAINS CIC and AIC

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Title: Ipilimumab for previously treated unresectable malignant melanoma

Produced by: Liverpool Reviews and Implementation Group (LR*i*G)

Authors: Adrian Bagust, Professor of Modelling in Health, LR*i*G, University of Liverpool

Angela Boland, Associate Director, LR*i*G, University of Liverpool

Helen Davis, Assistant Director, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

Rumona Dickson, Director, LR*i*G, University of Liverpool

Yenal Dundar, Research Fellow, LR*i*G, University of Liverpool

Ernie Marshall, Consultant in Medical Oncology, Clatterbridge Centre for Oncology

Gerlinde Massey, Research Assistant (Clinical Effectiveness), *LRiG*, University of Liverpool

Michaela Blundell, Research Associate (Medical Statistician), LR*i*G, University of Liverpool

Correspondence to: Ms Rumona Dickson, Director, LR*i*G , University of Liverpool, Room 2.12, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 19 August, 2011

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 08/209/01

Declared competing interests of the authors:

Dr. Ernie Marshall has received consultancy fees for advisory board work for Bristol-Myers Squibb.

Acknowledgements:

We would like to thank Professor Ruth Plummer for her comments on the content of this report. Professor Plummer has received consultancy fees and reimbursement for attending a symposium from Bristol-Myers Squibb. Also thanks to Dr. James Larkin for his response to our clinical questions. Dr. Larkin has received honoraria from Bristol-Myers Squibb.

Rider on responsibility for report:

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Dickson R, Boland A, Bagust A, Blundell M, Massey G, Dundar Y, Davis H and Marshall E. Ipilimumab for previously treated unresectable malignant melanoma: A Single Technology Appraisal. LRiG, The University of Liverpool, 2011.

Contributions of authors:

Rumona Dickson	Project lead, drafted clinical results section and supervised production of the final report
Adrian Bagust	Critical appraisal of the economic model
Angela Boland	Critical appraisal of the economic evidence
Michaela Blundell	Critical appraisal of the clinical statistical approach
Helen Davis	Critical appraisal of the manufacturer's submission
Gerlinde Massey	Critical appraisal of the clinical evidence
Yenal Dundar	Cross checking of manufacturer's search strategies
Ernie Marshall	Critical appraisal of the clinical sections of the manufacturer's submission

All authors read and commented on draft versions of the ERG report.

Table of contents

Abbreviations

AE(s)	adverse event(s)
AUC	area under curve
BORR	best overall response rate
BSA	body surface area
CI	confidence interval
CR	complete response
ECOG	Eastern Co-operative Oncology Group
EMA	European Medicines Agency
ERG	Evidence Review Group
FDA	Food and Drug Administration
gp100	gp100 peptide vaccine
HR	hazard ratio
HRQoL	health related quality of life
ICER	incremental cost-effectiveness ratio
IL-2	Interluken-2
Ipi	Ipilimumab
irAE	Immune-related adverse event
ITT	intention to treat
LYG	life years gained
MS	manufacturer's submission
NICE	National Institute for Health and Clinical Excellence
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PR	partial response
PSA	probabilistic sensitivity analysis
QALY	quality adjusted life years
RCT	randomised controlled trial
RECIST	Response Evaluation Criteria in Solid Tumours
SAE	Serious adverse event
SmPC	Summary of Product Characteristics
STA	Single Technology Appraisal

1 SUMMARY

1.1 *Scope of the submission*

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost-effectiveness evidence submitted to the National Institute for Health and Clinical Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence have been submitted to NICE from Bristol-Myers Squibb Pharmaceuticals Ltd (BMS) in support of the use of ipilimumab (Yervoy™) as a treatment for advanced (unresectable or metastatic) melanoma in adults who have received prior therapy.

The European Medicines Agency (EMA) has recently issued a marketing authorisation for ipilimumab. The recommendations state that patients should receive the full induction regimen of four doses regardless of the appearance of new lesions or the growth of existing lesions; tumour assessment is scheduled to be carried out only after completion of the 12 week induction regimen. The EMA accepted the pharmacovigilance risk management plan proposed by the manufacturer.

1.2 *Summary of submitted clinical-effectiveness evidence*

The clinical evidence provided in the manufacturer's submission (MS) is derived from three studies. The MDX010-20 trial provided the primary clinical evidence; CA184-007 was included to provide safety evidence; it is unclear why CA184-022 was included. None of the submitted trials compared ipilimumab with any of the comparators listed in the NICE scope/decision problem.

The MDX010-20 trial was an international, multicentre double blind, three armed randomised controlled trial (RCT) that compared ipilimumab+gp100, ipilimumab+placebo and gp100+placebo in a 3:1:1 randomisation schedule in patients with unresectable stage III/IV melanoma. The trial included 676 patients and the primary outcome was overall survival (OS).

The reported median OS and 95% confidence interval (CI) in the three groups (ipilimumab+gp100, ipilimumab+placebo and gp100+placebo) are 10.0 months (8.5 to 11.5), 10.1 months (8.0 to 13.8) and 6.4 months (5.5 to 8.7) respectively. When comparing OS benefit in the ipilimumab+gp100 arm vs gp100+placebo, the MS reports a statistically significant hazard ratio (HR) of 0.68 (95% CI 0.55 to 0.85) that favours ipilimumab+gp100. A similar statistically significant difference is reported in the comparison of ipilimumab+placebo vs gp100+placebo with a reported HR of 0.66 (95% CI 0.51 to 0.87). Benefit from treatment varies from patient to patient and only a small subgroup of patients appears to exhibit long-term survival. To date no patient characteristics or biomarkers have been identified which can prospectively target treatments to the minority of patients most likely to benefit from treatment with ipilimumab.

The most important adverse events (AE) are noted to be immune-related AEs (irAE). The MS states that these events are manageable and reversible in most cases. The EMA has accept a pharmacovigilance programme proposed by Bristol-Myers Squibb to monitor and treat these events. Clinical opinion and recent trial data indicate that as clinicians become more familiar with the use of immunotherapy they are able to identify and treat these AEs in a timely and proactive manner.

1.3 *Summary of the submitted cost-effectiveness evidence*

In the absence of any relevant UK economic evaluations, the manufacturer submitted a *de novo* economic evaluation comparing ipilimumab vs best supportive care (BSC) in patients with previously treated unresectable malignant melanoma (base case); the manufacturer also considers the use of other drug treatments in the sensitivity analysis.

The manufacturer constructed an EXCEL-based cost-utility model. The model is a cohort model with one cohort receiving ipilimumab and the other cohort receiving BSC. The approach used in the evaluation is a "partitioned-survival" model and is similar to a Markov cohort model. However, unlike a Markov model in which the transitions are modelled explicitly using transition probabilities, the "partitioned-survival" model calculates the proportion of patients in each treatment cohort that are expected to be in each health state at any time after treatment initiation. There are four mutually exclusive states in the model: baseline disease, non-progressive disease, progressive disease and death. The economic model adopts a 30 year time horizon. The perspective adopted in the economic evaluation was that of the NHS and Personal Social Services (PSS) and costs and benefits were discounted at 3.5% per annum.

Clinical effectiveness data from the MDX010-020 trial were used to populate the base case analysis in the submitted economic model. Health related quality of life (HRQoL) data were collected in the MDX010-20 trial using the EORTC QLQ-C30 and the short form 36 (SF-36); the manufacturer therefore had to map HRQoL data using validated mapping algorithms in line with the NICE reference case. The manufacturer's base case incremental cost-effectiveness ratio (ICER) for ipilimumab vs BSC is £60,737 per QALY gained; when ipilimumab is compared to all other (chemotherapy) comparators in the sensitivity analysis the ICERs are reduced. The manufacturer showed the ICERs to be robust when subjected to extensive deterministic, scenario and probabilistic sensitivity analysis (PSA). The cost effectiveness acceptability curve from the manufacturer's PSA shows that there is approximately a 14% chance of ipilimumab being cost effective vs BSC at a threshold of £50,000 per QALY gained. The manufacturer makes the case that ipilimumab should be considered under NICE's 'End of Life' criteria.

1.4 *Commentary on the robustness of submitted evidence*

1.4.1 Strengths

The use of ipilimumab in the treatment of patients with unresectable melanoma yields an OS benefit in a patient population who previously had access to only limited and unproven treatment regimens.

Using immunotherapy to treat patients with unresectable melanoma has raised the non-standard AE profile of ipilimumab. Experience shows that, as clinical experience with ipilimumab is gained, health care professionals are more able to identify and treat any AEs and irAEs that may occur.

Other than the manufacturer's approach to estimating OS, the minor corrections/amendments to the economic model suggested by the ERG do not affect the size of the base-case ICER.

1.4.2 Weaknesses

The clinical evidence presented by the manufacturer is primarily derived from a single RCT that does not compare ipilimumab to any of the comparators listed in the final scope issued by NICE. The MS assumes that the comparator in the key trial (gp100) is clinically equivalent to best BSC. The ERG notes that OS in the gp100 arm appears to be slightly lower than standard estimates of OS in this patient population.

All of the patients in the key RCT are HLA-A2*0201-positive. It is estimated that HLA positive patients make up approximately one quarter to a third of the total patient population. The manufacturer makes a convincing case that the clinical effectiveness of ipilimumab is not influenced by the HLA status of patients.

Although manageable, the AE and irAE profiles of ipilimumab require that a pharmacovigilance programme be put in place by the manufacturer as a condition of the marketing authorisation.

The ERG is of the opinion that the manufacturer has overestimated the OS gains associated with ipilimumab in the economic model.

1.4.3 Areas of uncertainty

The most clinically effective dose of ipilimumab is yet to be determined. The manufacturer has planned a randomised study to compare 3mg/kg vs 10mg/kg of ipilimumab with a focus on efficacy and safety in patients with advanced melanoma.

There appear to be subgroup of patients who may benefit more from treatment than others, but as yet it is not possible to predict who these patients will be prior to treatment.

Immunotherapy appears to yield different patient responses to treatment and, in the case of advanced melanoma, patients may move from progressive disease into stable disease or partial response. Existing criteria for measuring cancer tumour progression (e.g. the modified WHO or RECIST criteria) do not appear to be sufficiently adequate to assess the treatment response to ipilimumab.

The health outcomes of melanoma patients treated in the UK appear to differ from the health outcomes of similar patients treated in other parts of Europe and this divergence may indicate differences in patient management styles across European countries. It is therefore uncertain whether the clinical effectiveness data used in the economic model are generalisable to patients treated in the UK.

1.5 *Key issues*

None of the trials described in the MS contain a comparator listed in the final scope issued by NICE. In the clinical section of the MS the manufacturer instead argues that gp100 is clinically equivalent to BSC. However, patient outcomes in the gp100 arm of the trial are less favourable than might be expected in untreated patients.

Clinical data suggest that ipilimumab provides survival benefit to a small subgroup of patients. At present it is not possible to predict who these patients will be prior to treatment.

Recent marketing authorisation recommends that all four doses of ipilimumab be given to patients, as tolerated, regardless of disease progression. It is not known how this might affect the costs or OS benefit if ipilimumab was to be approved for use in clinical practice.

It appears to be clear from the clinical data provided by the manufacturer that ipilimumab yields an OS benefit over gp100. However, approaches to estimating survival for this group of patients are complex and open to debate. The ERG considers that the manufacturer has overestimated the survival benefit associated with ipilimumab and in doing so underestimates the size of the estimated base-case ICER. From the clinical data available, the ERG was unable to estimate the true size of the OS benefit related to ipilimumab and offers a commentary on why this is not possible.

The MS makes a case for ipilimumab to be considered under NICE's 'End of Life' criteria.

1.6 *What are the key factors influencing the size of the ICER?*

The key factor influencing the size of the ICER is the size of the OS benefit gained by ipilimumab that is used in the economic model. The ERG considers that the manufacturer substantially overestimates the OS benefit associated with ipilimumab. The ERG is of the opinion that it is currently impossible to estimate the OS benefit gained from use of ipilimumab from the key RCT as data are currently unsuitable for projective modelling. Based on the results of an exploratory analysis,

the ERG calculated a revised ICER using an OS estimate that is half (16.3 months) the size of the manufacturer's OS gain (33.8 months); the ERG's revised ICER is £96,717 per QALY gained; the ERG believes that use of 16.3 months is also likely to overestimate the size of the OS benefit gained from use of ipilimumab.

2 BACKGROUND

2.1 *Critique of manufacturer's description of underlying health problems*

In the context section of the manufacturer's submission (MS) (Section 2), the manufacturer describes the key issues related to the underlying health problem, including epidemiology, prognosis and treatment objectives. A summary of this section of the MS is presented in Box 1.

Box 1: Epidemiology, prognosis and treatment

Epidemiology and aetiology

Melanoma is the most aggressive form of skin cancer and is generally fatal if undetected and untreated. The mean age of diagnosis is 50 years, which is earlier than for most other cancers, but approximately 20% of cases occur in young adults aged between 15 and 39 years. Its incidence is slightly higher in males than in females.[1] Its global incidence is increasing faster than all other types of cancer.[2] Some estimates suggest that the incidence of melanoma is doubling every 10-20 years,[3] and the mortality rate continues to increase faster than that of most other cancers.[2]

Prognosis

In its early stages, melanoma is normally asymptomatic and, if detected before it has spread, it can be cured. However, 10% of cutaneous melanomas will have already metastasised by the time they are diagnosed. For stage III disease (i.e. regional lymph nodes involved) 5-year survival rates range from 40% to 50%, while stage IV disease (i.e. the melanoma has spread to distant sites) has an extremely poor prognosis (5-year survival rate is approximately 5-15%; median survival is 6-9 months).[4]

Treatment objectives

A small minority of people with advanced disease can still have their entire tumour removed. People with unresectable stage III or IV (metastatic) disease are usually managed by a specialist oncologist and first-line standard of care normally involves the administration of dacarbazine. Radiotherapy, immunotherapy and combination chemotherapy have been studied in randomised clinical trials. Limited treatment options are currently available for second or subsequent line therapy[5] and there is no standard of care in this setting. Up to now, there have been no approved therapies for previously treated advanced disease. Dacarbazine, vindesine, interferon and carboplatin are amongst the treatments used but these offer limited benefit. None of these agents have demonstrated a significant survival benefit in randomised phase III clinical studies.[4] Participation in clinical trials is the main treatment option for these patients.

It is worth noting that malignant melanoma is the least common but also the most serious type of skin cancer. In 2008 there were 11,767 new cases diagnosed in the UK, and 2067 deaths.[6] UK incidence rates have increased more rapidly over the past 25 years than any of the top ten cancers in males and females, and malignant melanoma was the sixth most common cancer diagnosed in females in 2008.[6] The mortality rate in people aged 65 years and older has almost tripled since 1979 from four deaths per 100,000 to 11.4 deaths per 100,000 in 2008.[6]

2.2 *Critique of manufacturer's overview of current service provision*

The manufacturer provides a detailed description of current service provision for the treatment of malignant melanoma (MS, pg 27). A summary of the clinical pathway of care for patients with malignant melanoma as described in the MS is presented in Box 2. The Evidence Review Group (ERG) considers this to be an accurate description of the current pathway of care in England and Wales.

Box 2: Summary of treatment approaches

There is no standard treatment for malignant melanoma. The definitive treatment of primary cutaneous melanoma is a wide local excision of the tumour. If complete removal of the tumour is not possible by surgical excision, then it is termed unresectable. The treatment options for an unresectable melanoma include:[7]

Radiotherapy - this has a very limited role in the management of patients with malignant melanoma, as it is generally regarded as a radio-resistant tumour. It is occasionally used for localised metastases and in the palliative setting.
Chemotherapy - is often used for patients with malignant melanoma. There is no evidence to support the use of adjuvant chemotherapy following surgery.
Vaccines - use is still experimental.
No active treatment - occasionally, for some patients, no active treatment with best supportive care (BSC) may be the most appropriate course of action.

When a patient has unresectable stage III or stage IV disease, the mainstay of treatment is systemic therapy. Systemic therapies include immunotherapy (e.g. interferon alfa), chemotherapy (e.g. dacarbazine, carboplatin, paclitaxel), immuno/biochemotherapy, and experimental vaccine immunotherapy.[8, 9]

In aggressive metastatic disease multi-agent chemotherapy containing paclitaxel and carboplatin or cisplatin, vindesine and dacarbazine produce partial responses and stabilisations in a meaningful number of patients.[8] Patients are often referred to clinical trials, due to questionable survival rates with the current treatment options. However, not all patients are eligible for available clinical trials (i.e. do not meet the inclusion criteria - for example being fit enough to receive a treatment).

When a treatment of an unresectable malignant melanoma has not been successful, there are no approved therapies or agreed standards of care for 'previously treated' advanced melanoma (i.e. no standard second-line therapy).

The MS (pg 26) also refers to relevant published guidelines on the treatment of malignant melanoma. A summary list of the guidelines is presented in Box 3.

Box 3: Summary of guidelines

In the absence of a validated standard of care, the melanoma treatment guidelines of the British Association of Dermatologists (2010),[10] the Royal College of Physicians (2007),[11] the European Dermatology Forum (2009),[12] the European Association of Dermato-Oncology and the European Organisation of Research and Treatment of Cancer,[13] the European Society for Medical Oncology[8] and the National Comprehensive Cancer Network in the US[14] all recommend that healthcare providers utilise clinical trials to manage advanced melanoma. However, as mentioned above, it is important to realise that not all patients are eligible for such clinical trials, and so have limited further treatment options.

As previously mentioned, there is currently no accepted standard of care for first-line treatment of patients with unresectable malignant melanoma and no approved second-line treatment. Expert clinical opinion sought by the ERG suggests that for patients with unresectable stage III and IV malignant melanoma, up to 30% of patients are not fit for chemotherapy and therefore receive no active treatment in the first-line setting. The remaining 70% receive dacarbazine and a small proportion of patients are entered into clinical trials. Second-line treatment is given to approximately 20% of patients who relapse, and treatment is predominantly carboplatin based chemotherapy or treatment within a clinical trial. The manufacturer states that ipilimumab "...offers advanced melanoma patients, and their clinicians, a legitimate treatment option where previously none existed" (MS, pg 28).

The comparison of treatment for patients with malignant melanoma in England to other parts of Europe was highlighted in the MS and in the manufacturer's clarification response describing data from the Melody study.[15, 16] In this study, patient record data were retrospectively collected from 31 centres in Europe [France (n=10), Italy (n=11) and the UK (n=10)]; data were collected from presenting patients from July 2005 to June 2006 with follow-up to May 2008. The ERG requested data in relation to UK and non-UK patients receiving second-line or subsequent treatments excluding those in clinical trials and those receiving immunotherapy. As can be seen in [redacted] it appears that survival in the UK is considerably worse for patients beginning conventional second-line systemic therapy than for patients treated in France and Italy. It is important to note that the data are somewhat limited in that there were data for only 23 UK patients while there were 203 patients in the non-UK group.

Reasons for a survival difference can only be speculated upon and, when asked, the ERG clinical reviewers suggested that some aspects of care may vary between UK and non-UK clinical practice e.g. differences in the frequency of follow-up assessment of patients who were previously successfully treated, the length of time in first-line treatment, differences in the selection of patients who are offered second-line treatment and/or the more conservative approach to treatment in the UK.

[1]

3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

The MS (pg 32) presents the decision problem issued by NICE and the manufacturer's rationale for a single deviation from this decision problem (Table 1).

Table 1 Decision problem and manufacturer's deviations from scope

Key parameter	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
Population	People with previously treated unresectable stage III or IV malignant melanoma	People with previously treated unresectable stage III or IV malignant melanoma	No difference
Intervention	Ipilimumab	Ipilimumab	No difference
Comparator(s)	Best supportive care, carboplatin-based chemotherapy, dacarbazine	Best supportive care, carboplatin-based chemotherapy, dacarbazine	No difference*
Outcomes	Overall survival, progression-free survival, response rate, adverse effects of treatment, health-related quality of life	Overall survival, progression-free survival, response rate, adverse effects of treatment, health-related quality of life	No difference
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	The cost effectiveness of treatments will be expressed in terms of cost per quality-adjusted life year The time horizon will be 30 years Costs will be considered from an NHS and Personal Social Services perspective	30 years is sufficiently long to reflect any differences in costs or outcomes between the technologies being compared
Subgroups to be considered	None	None	No difference
Special considerations, including issues related to equity or equality	None	None	No difference

*The ERG does not agree that there is no difference as the comparator used in the key trial is gp100 and this is also used as a proxy for BSC in the economic evaluation

3.1 *Population*

The population under consideration is made up of patients with unresectable stage III or IV malignant melanoma who have been previously treated and this matches the patient population described in the NICE scoping document[5] and the EU marketing authorisation for ipilimumab.[17, 18] The key clinical evidence submitted is derived from a trial[19] that only recruited HLA-A2*0201-positive patients; this was unavoidable due to the choice of gp100 as a comparator. However, the ERG considers that the manufacturer makes a convincing case that the clinical effectiveness of ipilimumab is unaffected by HLA status.

In addition, the evidence submitted provides details of patients excluded from the trial and these are outlined in Section 4, the most important of these are the exclusion of patients with ocular or untreated brain metastases and those patients with a variety of previously treated cancers. These exclusion criteria are considered appropriate.

3.2 *Intervention*

Ipilimumab has been co-developed by Medarex and BMS. It is a fully human monoclonal antibody that blocks the cytotoxic T-lymphocyte antigen 4 (CTLA-4) to promote anti-tumour immunity. It has received a marketing authorisation for the treatment of patients who have been previously treated for advanced melanoma.[17, 18] In terms of this appraisal ipilimumab is delivered intravenously with an induction dose of 3mg/kg every 3 weeks for a total of four doses. It is worth noting that in the Summary of Product Characteristics (SmPC)[20] it is recommended that:

> The recommended induction regimen of YERVOY® is 3 mg/kg administered intravenously over a 90-minute period every 3 weeks for a total of 4 doses. Patients should receive the entire induction regimen (4 doses) as tolerated, regardless of the appearance of new lesions or growth of existing lesions. Assessments of tumour response should be conducted only after completion of induction therapy (pg1).[20]

Omission of scheduled doses is recommended for a range of grade 2/3 adverse events (AEs) with continuation when the symptoms have resolved to grade 1. Permanent discontinuation of the intervention is recommended for a range of grade 3/4 AEs, especially immune-related AEs (irAEs).

Of note is that based on additional trial data, the FDA[21] has recently (March, 2011) approved the use of ipilimumab in both the first and second-line treatment of malignant melanoma. Both the European Medicines Agency (EMA) and the Food and Drug Administration (FDA) acknowledge the many facets of the benefit-risk balance of ipilimumab and therefore have accepted a pharmacovigilance risk management plan be implemented by the manufacturer.

3.3 *Comparators*

Given that there are no current recommendations for the second-line treatment of patients with malignant melanoma, the comparators in the scope represent a range of current treatment options used in UK clinical practice. These treatments include carboplatin-based chemotherapy and dacarbazine which are used in patients who are fit enough to receive treatment and/or BSC for those who are not. The majority of the clinical discussion in the MS is focussed on the clinical evidence derived from the manufacturer's pivotal trial[19] which compares ipilimumab vs ipilimumab+gp100 vs gp100 and does not include any of the comparators stated in the final scope issued by NICE.

The manufacturer makes a case in their submission that gp100 is equivalent to BSC. This seems to be somewhat of a contradiction. In their clarification response the manufacturer states that gp100 was used because of its single-agent activity and its previously demonstrated ability to enhance the efficacy of IL-2.[22, 23] In their clarification response they goes on to say that gp100 vaccine is a 'well-studied and clinically active investigational agent' (pg 16). Clinical debate in this area continues with some who advocate that such vaccines may in fact be detrimental to the health status of patients.[24]

Clinical advisors to the ERG indicated that they did not consider that gp100 vaccine was the ideal comparator but that its use in the trial was reasonable; they also stated that the direction of the effect that gp100 vaccine may potentially have on outcomes is not known. The EMA states[17] that gp100 is an "experimental anti-cancer agent and the effect of gp100 monotherapy on the overall survival of melanoma patients is not exactly known. The observed median OS of gp100 of 6.4 months [in the trial] was somewhat lower than those seen in recent phase III trials curve which are typically around 7-8 months" (pg 48/49). However, the EMA did ask the manufacturer to provide the results of an additional analysis that added an overall 1.5 months benefit to the gp100 group on the assumption that patients in this group could indeed have had negative outcomes as a result of the treatment; the EMA was satisfied with the results of this additional analysis.

3.4 *Outcomes*

The MS provides evidence related to overall survival (OS), progression-free survival (PFS), AEs and health related quality of life (HRQoL). However, the nature of the disease and the action of the intervention are such that the published literature questions the ability of standard outcome measures (with the exception of OS) to reflect accurately the effectiveness of immunotherapy treatments in general and ipilimumab in particular. These issues are reflected in the manufacturer's product information that recommends continuation of treatment even if there is evidence of the identification of new lesions or the growth of existing lesions.[17]

In a recent publication Saenger and Wolchok[25] reported on the variety of patient responses to ipilimumab noting that the kinetics of the responses of patients are heterogeneous and different from those of chemotherapy and even other immunotherapy agents. They go on to report that patients may even present with mixed responses to treatment (e.g. patients may present with stable disease that was preceded by disease progression). They attribute these differences to the time needed to reach and maintain the anti-tumour response.

A paper[26] co-authored by a subset of authors of the published trial reports submitted for evidence in this appraisal, outlines the outcome of a set of workshops held during 2004 and 2005 where 200 oncologists, immune-therapists and regulatory experts were convened to examine the differences in patient response and measurement of outcomes in patients with melanoma. This paper[26] argues that the standard WHO[27] or revised RECIST criteria[28] used to measure solid tumour response to chemotherapy are not appropriate for measuring response to immunotherapy. It is interesting to note that the authors of the paper[26] state that they use outcome data from the manufacturer's phase II trial programme; they list three different trials,[29-31] yet data from only two are included.[29, 31]

The results of the workshops outline four response scenarios reported by the participants. These include:

- Response in baseline lesions
- Stable disease with a slow steady decline in total tumour volume
- Response after initial increase in total tumour volume
- Reduction in total tumour burden after the appearance of new lesions

The conclusion drawn is that treatment should be continued even in light of what might appear to be disease progression and this, as noted earlier, is included in the SmPC for use of ipilimumab.[20]

In a commentary on these proposed new outcome measures Ribas et al[32] note some of the difficulties of adopting this new system pointing out that the proposed new criteria have not been validated and that they would not be applicable in trials where immunotherapy is compared to standard chemotherapy. This is less of an immediate issue for this appraisal since the key outcome of importance is OS. However, this is an important issue that will need to be addressed as the use of immunotherapy treatments increases in clinical practice.

Adverse events are a key outcome in this appraisal for two reasons. The first is that the intervention is immunotherapy and not chemotherapy and therefore the AE profile is significantly different to the clinical practice experience of many clinicians. That is, the most important AEs are related to immune-related adverse events (irAEs) which require rapid intervention with newly established treatment protocols. The second is that the serious adverse event rates described in the trial evidence

provided are greater than would normally be expected and include relatively high rates of treatment related deaths.

Comparative data are difficult to identify. However, in a recently published[33] phase III international multi-centre randomised controlled trial (RCT) of 502 patients, ipilimumab (10 mg/kg) in combination with dacarbazine was compared with dacarbazine plus placebo in patients with previously untreated metastatic melanoma. The types of AEs in this trial were similar to those seen in earlier studies of ipilimumab, but the rates of elevated liver-function values were higher and there were fewer cases of gastrointestinal events than expected. The trial authors consider that the noticeable shift in the rates of AEs associated with ipilimumab may be due to its combination with dacarbazine, which is known to cause hepato-toxicity when used as a monotherapy. In addition, grade 3/4 AEs occurred in 56.3% of patients treated with ipilimumab and dacarbazine, which was significantly higher ($p<0.001$) than in those treated with dacarbazine and placebo (27.5%).

The MS indicates that AEs were generally medically manageable and usually reversible. Clinical advice to the ERG indicated that, although initial AE rates were high and of concern, clinical experience with the drug and the introduction of robust AE treatment algorithms mean that such events are now identified quickly and managed more proactively leading to improved patient outcomes. It is also worth noting that in the trial discussed above there were no drug-related deaths reported. [33]

3.5 *Economic analysis*

The MS argues appropriately that given the age of onset of the disease and its poor prognosis that the time horizon should be limited to 30 years; this is the only identified deviation from the final scope issued by NICE.

4 CLINICAL EFFECTIVENESS

Table 2 provides an outline of the key background/clinical information and its location within the MS. Its purpose is to signpost the reader to the main areas of background/clinical information within the MS.

Table 2 Key clinical information in the MS

Key information	Section in the MS
Description of the technology	1
Context	2
Equity and equality	3
Statement of decision problem	4
Literature search	5
Search strategies	5.1; Appendix 2 and 6
Study selection	5.2
Summary of methodology and study characteristics of relevant RCTs	5.3; Appendix 3
Summary of methodology and study characteristics for indirect comparisons	N/A
Critical appraisal of relevant RCTs	5.4; Appendix 3
Results: efficacy and tolerability from direct evidence	5.5
Results: efficacy and tolerability from indirect evidence	N/A
Adverse events	5.9
Interpretation of clinical evidence	5.10

4.1 *Critique of manufacturer's search strategy*

The MS reports the conduct of two systematic reviews of the literature to identify relevant clinical evidence related to ipilimumab. The first was conducted early in the development of the submission and spanned 1970 to April 8, 2010 and the second is a review update to cover the period of January 2010 to May 2011. Three electronic databases were searched (Medline, EMBASE and the Cochrane Library). Search terms appropriately included a combination of free-text and index terms combined with drug name used as free-text. The search strategies were reviewed and considered to be appropriate.

In addition, three sets of conference abstracts (first review) including American Society of Clinical Oncology (ASCO, annual meetings 2008-2010), European Society of Clinical Oncology (ESMO, annual meetings 2008-2010) and European Association of Dermato-Oncology (EADO, annual meetings 2008-2010) were searched. Additional conference sites that were also searched included:

Perspectives in Melanoma (2008-2011) and Annual International Congress of the Society for Melanoma Research (2008-2011).

Given that it is likely that for all clinical studies in this area the manufacturer will have sponsored the trial or at least supplied the drugs, it is somewhat surprising that the manufacturer did not include a search of their own internal database of studies.

4.2 *Critique of manufacturer's inclusion criteria and study selection*

The inclusion criteria for the review of the evidence were provided by the manufacturer and were appropriate. However, there were inconsistencies in the reporting of the decisions taken during the application of the review inclusion criteria. Although the review inclusion criteria indicated that trials with comparator treatments would be included in fact only trials that provided evidence related to ipilimumab were ultimately considered.

Three studies that examined comparator interventions were identified in the manufacturer's first review but are never mentioned again. Only one[34] of these is listed in the references list in the MS and is a comparison of sorafenib in combination with carboplatin and paclitaxel in the second-line treatment of malignant melanoma. The second is a paper cited as Zimpfer (2003) and it is likely that it is a trial[35] of weekly paclitaxel vs paclitaxel and carboplatin as second-line therapy in disseminated melanoma. The third trial is listed as Eisen (2010) and it is likely that it is a trial[36] assessing the effectiveness of lenalidomide vs placebo. In addition, a trial by Maio et al[37] is mentioned in the text of the MS as having been included in the first review but it had not been previously mentioned. It is unlikely that further discussion of the results of these trials[34-37] or inclusion of the data (e.g. in an indirect comparison) in the MS would have provided any added value. However, the inconsistencies in reporting reflect poorly on the quality of the review of the literature conducted by the manufacturer.

The MS concludes that two trials should be included for consideration of clinical effectiveness; MDX010-20[19] referred to as Hodi 2010 and CA184-022[31] referred to as Wolchok 2010. The first is a randomised double blind phase III trial while the second is a phase II dose ranging trial. It is worth noting that there is an internal inconsistency in the document whereby the inclusion of this second trial is listed as presenting either clinical or only safety data (MS, pg 84 and 87).

The inclusion criteria for the literature search were limited to RCTs of second-line treatments (explicitly excluding studies with mixed first and second-line treatment). However, the MS indicates that study CA184-007[30] which is referred to as Weber 2009 was selected to provide safety and tolerability data. This trial includes a mixed population of treatment naive and previously treated patients and used a dose of ipilimumab that is three times larger than the dose being considered in this

appraisal; the trial also included use of a steroid in one arm to test whether the steroid decreased the number of AEs experienced by patients. No other rationale for the use of this trial is provided.

It is not unusual in circumstances where trial data are limited to broaden the inclusion criteria to allow for the identification of such safety data. However, it is advisable to be very specific as to the inclusion criteria that are used in the selection and inclusion of such studies so that any possible bias may be identified. Given that all of the studies of ipilimumab appear to have been sponsored by the manufacturer it is likely that they have a database of all study results and could therefore have interrogated a much broader sample of data related to patients who had received the treatment drug to develop a more comprehensive safety profile. Such an analysis was provided to the EMA for their consideration.[17]

The ERG identified at least one other RCT that included a mixed treatment population where patients in one arm of the trial received the dose of ipilimumab being considered in this appraisal.[38] The results of this study were found on-line through the Bristol-Myers Squibb website but no publication of the results was identified. In addition there is a single arm trial[29] of patients receiving second-line treatment with a 10mg/kg dose of the drug that might have provided comparative safety data.

A recent independent review[39] identified the same published studies as the searches run by the manufacturer and the ERG. The ERG is therefore confident that no other relevant published studies were missed in the search. As noted above the ERG identified one unpublished study[38] that is not mentioned in the submission, the results of which could have expanded the data available for analysis but it is unlikely that its inclusion would have changed the outcomes.

4.3 *Critique of included studies*

A critique of the quality of the trials submitted in the MS is provided below. The MS provided a quality assessment of the included trials and these are presented in Appendix 1.

4.3.1 MDX010-20

The MDX010-20[19] trial is a large, three armed, double-blind, international, multicentre phase III trial. The ERG considers that the design of the trial was appropriate however the numerous protocol amendments that occurred mean that the robustness of the results may be brought into question. For example, there was a change in the primary outcome of the trial from that of best overall response rate (BORR) to that of OS.

The MDX010-20[19] trial recruited and centrally randomised 676 patients from 125 centres in 13 countries in North America, South America, Europe and Africa. As there is no current standard of care for these patients the researchers selected the gp100 vaccine for its additive effect to ipilimumab. Such an additive effect has previously been demonstrated in the literature.[23, 40, 41] Patients were

randomly assigned in a 3:1:1 ratio to one of three treatment options: ipilimumab+gp100, ipilimumab+gp100 placebo, or gp100+ipilimumab placebo.

Patients were stratified according to baseline metastasis stage (M0, M1a, or M1b vs M1c) and previous interleukein 2 (IL-2) therapy, or no previous IL-2 therapy. The ERG considers the method of randomisation used in the MDX010-20[19] trial to be adequate. The manufacturer presents details of the baseline characteristics of the three arms of the MDX010-20[19] trial (MS, pg 50). These show that the patients in the three arms were generally well balanced for key baseline characteristics.

Concealment of treatment in the MDX010-20[19] trial was achieved through use of placebos for both ipilimumab and gp100. Patients and site personnel were blinded to treatment allocation. It is noted that the manufacturer was responsible for the collection of trial data and that the dispensing pharmacists and the manufacturer were aware of the patient assignments in the trial. It is unclear what effect this may have had on the reported outcomes of the trial.

Of the 676 patients recruited from 13 countries (Argentina, Belgium, Brazil, Canada, Chile, France, Germany, Hungary, The Netherlands, South Africa, Switzerland, UK and US), only 8% were from the UK, however, a significant proportion (38%) were from European Union countries. Due to the use of the gp100 vaccine all patients in the trial were tested and were recruited if HLA-A*0201-positive. A recently established database estimates that the proportion of such patients in Western Europe would be between 20 and 30%.[42] This figure is consistent with the fact that the manufacturer reports screening 1783 patients to identify the 676 (38%) patients included in the trial. Under the assumption that the CTLA-4 blockade action of ipilimumab is independent of HLA status, the MS therefore concludes that the results of the study can be generalised to HLA-A*201 negative patients.

The ERG is satisfied that the patients in the MDX010-20[19] trial are representative of patients in UK clinical practice.

4.3.2 CA184-022

CA184-022[31] is a double-blind, multicentre, dose-ranging phase II RCT. The ERG considers that the trial is of reasonable quality.

The trial recruited 217 patients from 66 centres in 12 countries, and randomised patients on a 1:1:1 ratio to receive either ipilimumab 0.3 mg/kg, 3mg/kg or 10mg/kg every 3 weeks for four cycles (induction) followed by maintenance therapy every 3 months.

Randomisation was performed using an interactive voice response system and patients were assigned a unique identification number. Patients were stratified according to previous treatment (IL-2,

dacarbazine, fotemustine, or temozolomide vs other treatments). The ERG is satisfied that the method of randomisation used was appropriate.

Details of the baseline characteristics of the patients in the CA184-022 trial[31] are presented in the MS (pg 51) and generally the treatment arms are well balanced. However, although the number of subjects with M1c-stage disease was similar between the 3mg/kg and the 10mg/kg groups (50% and 51% respectively) it was higher in the 0.3 mg/kg group (62%). Also, there were more males in the 0.3mg/kg group (71%) compared to 67% and 61% in the 3mg/kg and 10mg/kg groups, respectively.

The CA184-022 trial[31] did not report the methods used for blinding of patients, clinicians and staff to treatment allocations. However, pharmacists and a biostatistician were not blinded to treatment allocation.

Of the 217 patients recruited from 12 countries (Canada, USA, Brazil, Argentina, Greece, Belgium, Germany, Czech Republic, Hungary, South Aftrica and Australia) a total of 85 were from five non-UK European countries. The manufacturer points out that of the 72 patients assigned to the licensed 3mg/kg arm, 28 were from non-UK European centres. The ERG is satisfied that the patients in the CA184-022 trial[31] are representative of patients in the UK.

4.3.3 CA184-007

CA184-007[30] is a double-blind, multicentre, phase II RCT. The ERG considers the trial to be of reasonable quality.

The trial randomised 115 patients on a 1:1 ratio to receive open-label ipilimumab (10mg/kg at weeks 1, 4, 7 and 10) with either concomitant oral budesonide or placebo. Patients were recruited from 11 centres in six countries (Canada, Israel, Italy, Peru, UK and USA). Randomisation was performed using an interactive voice response system, and randomisation was stratified according to previous treatment received. Blinding was achieved through use of placebos, however patients developing grade >2 diarrhoea or other irAEs discontinued budesonide/placebo and commenced open-label budesonide/other steroids. The ERG is satisfied that the method of randomisation used was appropriate. Details of the baseline characteristics of the patients in the CA184-007 trial are presented in the MS (pg 52) and generally the treatment arms are well balanced. However, more patients received previous systemic therapy in the budesonide arm compared to the placebo arm.

The ERG notes that the patients in these three trials are slightly younger and have a better performance status (PS) than patients typically seen in UK clinical practice. This is consistent with the patient profiles in clinical trials.

Table 3 Included study details

	MDX010-20[19]	CA 184-022[31]	CA 184-007[30]
Location	125 centres 13 countries	66 centres 12 countries	11 centres 6 countries
Design	RCT – Phase III Double blind	RCT – Phase II – dose ranging trial Double blind	RCT – Phase II Double blind
Patient inclusion criteria	Inclusion: Unresectable stage III or IV melanoma that had been previously treated ECOG status 0 or 1 HLA-A2*0201 Exclusion: Any other cancer from which disease free for <5 years (except treated /cured basal cell or squamous cell skin cancer, superficial bladder cancer or treated carcinoma in situ of cervix, breast or bladder) Primary ocular melanoma Previous anti-CTLA-4 antibody or cancer vaccine Autoimmune disease Active untreated metastases in CNS Pregnancy or lactation Concomitant non-study anticancer therapy or immunosuppressant Long-term use of corticosteroids	Inclusion: Unresectable stage III or IV melanoma that had been previously treated Exclusion: Concomitant therapy with any anti-cancer agent; immunosuppressive agents; any non-oncology vaccine therapy; surgery or radiotherapy, other investigational anti-cancer therapies; or chronic use of systemic corticosteroids Previous treatment with other investigational products, including cancer immunotherapy, within 30 days Previous treatment in another ipilimumab clinical trial or prior treatment with a CD137 agonist, CTLA-4 inhibitor or agonist Autoimmune disease: a documented history of inflammatory bowel disease, including ulcerative colitis and Crohn's disease, symptomatic autoimmune disease (e.g. rheumatoid arthritis, systemic progressive sclerosis, systemic lupus erythematosus, autoimmune vasculitis Evidence of brain metastases Any other malignancy from which disease-free for < 5 years, (except adequately treated and cured basal or squamous cell skin cancer, superficial bladder cancer or carcinoma *in situ* of the cervix) Primary ocular or mucosal melanoma Pregnancy	Inclusion: Unresectable stage III or IV melanoma that were treatment naive or who had been previously treated ECOG status 0 or 1 Exclusion: Any other cancer from which disease free for <5 years (except treated /cured basal cell or squamous cell skin cancer, superficial bladder cancer or treated carcinoma in situ of cervix, breast or bladder) Primary ocular melanoma Previous anti-CTLA-4 antibody Autoimmune disease Active untreated metastases in CNS Long-term use of corticosteroids Received investigational drugs within 4 weeks of starting protocol therapy

	MDX010-20[19]	CA 184-022[31]	CA 184-007[30]
Interventions and comparators	Ipi 3 mg/kg + gp100 (n=403) Ipi 3 mg/kg + gp100 placebo (n=137) Ipi placebo + gp100 (n=136)	Ipi10 mg/kg (n=72) Ipi 3 mg/kg (n=72) Ipi 0.3 mg/kg (n=73)	Prophylactic budesonide 9 mg + ipi 10 mg/kg (n=58) Prophylactic placebo + ipi 10 mg/kg (n=57)
Primary and secondary outcomes	OS BORR Modified WHO criteria[27] PFS at 24 weeks Duration of response	BORR Disease control, PFS at week 24 Modified WHO criteria[27] Median OS Duration of response 1yr survival	AE – specifically diarrhoea BORR, OR, duration of response, time to response Modified WHO criteria[27] OS 1 yr survival Disease control rate
Timeframe	2004-2008	2006-2007	2005-2007
Duration of follow-up	55 months	Median follow-up for OS 8.3-10.7 months	38 months
Gender male	54-61%	61-71%	67-74%
% Melanoma stage M1c	71-73%	50-62%	48-51%
% Received 4 study doses	57-64%	51-72%	55-61%
Mean age range	55.6-57.4	56-59	Median 59-61
Performance status 1 or 2	99%	99%	97%

4.3.4 Description and critique of the statistical approach used

MDX010-20

A total of 676 patients were recruited from 125 centres in 13 countries in North America, South America, Europe and Africa. For the results of a trial with so many centres to be meaningfully interpreted, the manner in which the protocol is implemented should be clear and similar across all centres. This is because with so many investigators in different countries, general clinical practice will always be an issue and the results of a trial can only be generalisable if it is executed efficiently. According to the clarification response, there were 21 major protocol deviations (3.1% of patients) and they were comparable across the trial arms (3% vs 2.9% vs 3.6% for ipilimumab+gp100, ipilimumab+placebo and gp100+placebo, respectively). However, one site in particular was found to have multiple protocol deviations and inadequate recordkeeping so a sensitivity analysis was performed to assess the impact of the removal of the results from this site on the OS result. The impact was found to be minor (HR=0.70; 95% CI 0.56 to 0.86 for ipilimumab+gp100 vs gp100+placebo).

It should be noted that it was originally planned that an independent review committee would review and evaluate all response data but when the primary outcome was changed from BORR to OS an independent review committee was no longer considered necessary so the final response analyses were based on investigator reported responses.

The ERG is concerned about the timing of the change and the degree of blinding that was in place when the change was made to the primary outcome. Recruitment took place from September 2004 until August 2008 with the final data cut-off being on 19th June 2009. The change to the primary outcome was made in the sixth protocol amendment (dated 15th January 2009), five months after the end of recruitment. The study is described as being double blind, the sponsor and pharmacists were aware of treatment allocations but patients and all other study personnel were blinded. When the decision was made to change the primary outcome, the sample size was revised using the new outcome based on a simulation using the collected blinded survival data from the study along with historical literature. The ERG is concerned about the preservation of the blinding as it is likely that the ipilimumab+gp100 group would be easily identifiable due to the 3:1:1 randomisation ratio.

There were six amendments to the original protocol. The first was put in place before the beginning of the study and it included changing the primary outcome from objective response rate (ORR) to BORR, adapting the study design to a one stage subject enrolment process rather than two stages and the inclusion of an additional stratification factor. The randomisation ratio was also changed from 4:1:1 to 3:1:1. The second amendment had no impact on the statistical analysis. The third changed the metastatic status stratification to TNM (tumour, lymph and metastasis) status and added an extra level

to it (M0). Clarification was provided on the major durable response rate and the duration of serious adverse event (SAE) reporting was increased from 4 weeks to 10 weeks. The fourth and fifth amendments had no impact on the statistical analysis. The sixth amendment changed the primary endpoint from BORR to OS and added secondary comparisons between the other treatment arms.

The original sample size calculation was based on the initial primary endpoint of BORR. It was calculated to provide 90% power to detect a 10% improvement in BORR (assuming 5% response rate in gp100 arm) assuming a two-sided 0.05 alpha level. Allowing for a 10% drop-out rate, 750 patients in total (450: 150: 150 for ipilimumab+gp100, ipilimumab+placebo and gp100+placebo respectively) were required to be enrolled. When the decision was made to change the primary outcome, the sample size calculation was revised using the new endpoint of OS based on the comparison of ipilimumab+gp100 vs gp100+placebo. Based on a simulation which used the collected blinded data from the study along with historical literature, it was identified that a total of 385 deaths in the ipilimumab+gp100 arm and the gp100+placebo arm would give 90% power. To observe 385 events in these groups it was estimated that approximately 481 deaths would be required across the three groups (assuming events are distributed evenly among the three treatment arms according to the randomisation ratio). As mentioned previously the ERG is concerned with the level of blinding that existed when the data were used to perform the simulation. In actual fact, 676 patients were enrolled onto the study and 525 experienced an event (died), exceeding the number required in the sample size calculation.

A *post hoc* sample size calculation showed that the 219 events in 273 patients in the ipilimumab+placebo and gp100+placebo arms provided 80% power to detect a difference in OS between the groups, at a 2-sided 0.05 alpha level, assuming ipilimumab+placebo had the same effect as ipilimumab+gp100.

As noted earlier the manufacturer and the pharmacists were not blinded to the patient allocation. The ERG is also concerned that investigators may have been able to guess which treatment patients were on because of the notable differences in AEs rates across the groups – especially lower AE rates in the gp100 arm of the trial. It was also highlighted in the clarification response from the manufacturer that patients receiving gp100 had higher rates of injection site reactions which could have been noted by patients and other study personnel. However, OS is an objective outcome and could not have been influenced by the investigators knowledge of the treatments. This may however have influenced the more objective secondary objectives.

All efficacy analyses were performed on the intention-to-treat population as well as the evaluable patient population (all intention-to-treat patients in the study who received any amount of ipilimumab or gp100 and had per protocol tumour evaluations). The statistical methods that were used to analyse

the outcomes in the trial are presented in Table 4. The ERG is generally satisfied with the statistical methods that have been used. However, there was no hierarchical testing rule specified for the testing of secondary endpoints so p-values cannot be interpreted.

Table 4 Efficacy analysis

Outcomes	**Method of Analysis**
OS (comparisons of ipilimumab+gp100 vs gp100+placebo, ipilimumab+placebo vs gp100+placebo and ipilimumab+gp100 vs ipilimumab+placebo), time to response, PFS	Plotted using Kaplan-Meier survival curves, median survival time estimates reported along with95% confidence intervals (CIs) computed using Brookmeyer and Crowley method, treatment differences compared using stratified log rank test (stratified by baseline TNM status and prior or no prior treatment with IL-2) and effect of prognostic factors and baseline disease status evaluated using Cox proportional hazards model.
BORR, disease control rate, major durable response rate	Summarised by treatment group using descriptive statistics and 95% CIs, treatment differences compared using stratified Cochran-Mantel-Haenszel test/Fisher's exact test.
Duration of Response	Summarised by treatment group using descriptive statistics and 95% CIs for responder patients only.
Progression-free survival rate (at week 12), survival rate (at 12 months, 18 months and 24 months)	Estimated using the Kaplan-Meier product-limit method by treatment groups, along with corresponding 95% CIs.
Quality of life	Treatment differences compared using an Analysis of Variance (ANOVA) model with treatment group as the main factor or using non-parametric methods when categorical data are observed.

No subgroup or adjusted analyses were presented in the MS. However, the manufacturer provided subgroup analyses in response to a clarification request from the ERG regarding prior treatments. The three treatment arms were relatively comparable in terms of number/type of prior treatments received by patients. [redacted] (reproduced from the clarification response see pg 14) shows that the treatment benefit for ipilimumab+gp100 is consistent across the different prior treatment subgroups.

[redacted]

[redacted]2

Treatment discontinuation rates were high in the trial (36% of patients who were treated in the ipilimumab+gp100 arm discontinued, 33% in the ipilimumab+placebo arm and 41% in the gp100+placebo arm). [redacted] below shows that the rate at which patients discontinued was comparable across the three treatment arms.

Figure 3

Further to this, the manufacturer provided information on treatments that patients received after participating in the trial. This information was collected by telephone contact and was not available for all patients (approximately 75% of all patients provided this information). The proportions of patients who went on to receive further treatment were comparable across the three treatment arms (70.2%, 63.4% and 71.3% of patients in the ipilimumab+gp100, ipilimumab+placebo and gp100+placebo arms respectively, who had follow-up contact received post-progression therapy). On the whole the numbers were comparable for the different types of therapy but there were some slight fluctuations. The effect of these additional treatments on OS is not measureable but the ERG clinical advisors indicate that, given the lack of available effective third-line therapies, these treatments are unlikely to have made a difference to OS estimates across the treatment arms of the trial.

As stated previously, the MS identified three relevant studies[19, 30, 31] describing ipilimumab for previously treated unresectable malignant melanoma. However, as one of these[31] was a dose-ranging study for ipilimumab monotherapy and another[30] was a comparison of ipilimumab 10mg/kg (not the licensed dose) either with prophylactic budesonide or with prophylactic placebo, no meta-analyses were performed. The ERG agrees that it would have been inappropriate to perform any meta-analyses. Furthermore the ERG agrees that there were insufficient data to perform mixed treatment comparisons.

4.4 *Clinical-effectiveness results*

Clinical results from the three included trials[19, 30, 31] submitted as a part of the MS are presented in Table 5.

Although two trials[19, 31] are included in the MS as part of the clinical effectiveness evidence, the second of these (CA 184-022)[31] is a dose ranging trial in which all participants received ipilimumab. Only clinical effectiveness data derived from MDX010-20[19] are discussed in this section.

As noted earlier, the MDX010-020 trial[19] was based on an assumption of the additive effect of gp100 vaccine and therefore in the trial the ipilimumab+gp100 group was three times larger than either the ipilimumab+placebo or the gp100+placebo groups. The primary outcome of the trial was OS. The reported median OS and 95% CI in these three groups are 9.95 months (8.5 to 11.5), 10.12 months (8.0 to 13.8) and 6.44 months (5.5 to 8.7).

In their clarification response, the manufacturer provided an updated survival analysis including **only patients who received all four treatments.** The results of this updated analysis are considerably more favourable than those in the original analysis (in all treatment arms). Median survival times are [redacted]

When comparing OS benefit in the ipilimumab+gp100 arm to gp100+placebo the MS reports a hazard ratio (HR) of 0.68 (95% CI 0.55 to 0.85) demonstrating a statistically significant difference. A similar statistically significant difference is reported in the comparison of ipilimumab+placebo with gp100+placebo with a reported HR of 0.66 (95% CI 0.51 to 0.87). As can be noted through examination of the Kaplan-Meier curves for OS there is a high risk of death in the first 18 months of the trial, after which death rates level off with a small but steady decrease. Reasons for the lack of response in a large proportion of patients and a more complete and long term response in a small proportion of patients are not explained in the MS. Research in this area is continuing to attempt to identify specific patient characteristics or biomarkers that might allow for the prediction of patients most likely to benefit from treatment.[43, 44]

Table 5 Outcomes

Outcomes[a]	MDX010-20[19]		CA 184-022[31]		CA 184-007[30]	
Overall survival - median months (95% CI)	Ipi + gp100 (n=403) Ipi + placebo (n=137) gp100 + placebo (n=136)	10.0 (8.5;11.5) 10.1 (8.0;13.8) 6.4 (5.5;8.7)	10mg/kg (n=72) 3mg/kg (n=72) 0.3 mg/kg (n=73)	11.4 (6.9;16.1) 8.7 (6.9;12.1) 8.6 (7.7;12.7)	Ipi + budesonide (n=58) Ipi + placebo (n=57)	17.7 (6.8;not reached) 19.3(12.0;not reached)
[redacted]	[redacted]	[redacted]				
BORR (CR/PR) % (95% CI)	Ipi + gp100 (n=403) Ipi + placebo (n=137) gp100 + placebo (n=136)	6% (4;8) 11% (6;17) 2% (<1;5)	10mg/kg (n=72) 3mg/kg (n=72) 0.3 mg/kg (n=73)	11% (5;21) 4% (<1;12) 0% (0;5)	Ipi + budesonide (n=58) Ipi + placebo (n=57)	12%(5;23) 16% (8;28)
Stable disease	Ipi + gp100 (n=403) Ipi + placebo (n=137) gp100 + placebo (n=136)	14% 18% 10%	10mg/kg (n=72) 3mg/kg (n=72) 0.3 mg/kg (n=73)	18% 22% 14%	Ipi + budesonide (n=58) Ipi + placebo (n=57)	19% 19%
Progressive disease	Ipi + gp100 (n=403) Ipi + placebo (n=137) gp100 + placebo (n=136)	60% 51% 65%	10mg/kg (n=72) 3mg/kg (n=72) 0.3 mg/kg (n=73)	50% 60% 59%	Ipi + budesonide (n=58) Ipi + placebo (n=57)	59% 51%
Not evaluable, missing or unknown	Ipi + gp100 (n=403) Ipi + placebo (n=137) gp100 + placebo (n=136)	21% 20% 24%	10mg/kg (n=72) 3mg/kg (n=72) 0.3 mg/kg (n=73)	31% 17% 27%	Ipi + budesonide (n=58) Ipi + placebo (n=57)	10% 14%
Survival at 1 year % (95% CI)	Ipi + gp100 (n=403) Ipi + placebo (n=137) gp100 + placebo (n=136)	44%(39;49) 46% (37;54) 25% (18;33)	10mg/kg (n=72) 3mg/kg (n=72) 0.3 mg/kg (n=73)	49% (37;60) 40% (28;51) 40% (28;51)	Ipi + budesonide (n=58) Ipi + placebo (n=57)	56% (43;69) 62% (49;75)
Survival at 2 years % (95% CI)	Ipi + gp100 (n=403) Ipi + placebo (n=137) gp100 + placebo (n=136)	22% (17;26) 24% (16;32) 14% (4;20)	10mg/kg (n=72) 3mg/kg (n=72) 0.3 mg/kg (n=73)	30% (19;41) 24% (14;35) 18% (10;28)	Ipi + budesonide (n=58) Ipi + placebo (n=57)	45% (27;54) 42% (28;55)

[a]Note – data from published papers and MS; ipi=ipilimumab; SD=standard deviation; CI=confidence interval; BOOR=best overall response rate; CR=complete response; PR=partial response; NR=not reported

Continuation and response rates

As can be seen in Table 3 the proportion of patients in MDX010-20[19] who received a full regimen of four doses of ipilimumab ranges from 57 to 64%. The most common reason for discontinuation was reported as disease progression. Given the wording of the EMA marketing authorisation[17] for ipilimumab which states that patients should complete the entire course of four treatments regardless of disease progression, it may be expected that a higher proportion of patients will receive four doses of the drug in clinical practice than was received by patients in the trial.[19]

Measuring response rates, as previously discussed, is problematic. However, reported response rates are low. The MDX010-20[19] trial reported that the highest percentage of patients with BORR was 10.9% in the ipilimumab+placebo arm of the trial. They go on to report that 28.5% of patients in the ipilimumab+placebo arm were evaluated as having a CR, PR or SD.

As noted earlier a small proportion of patients respond to treatment and as yet there are no predictive markers to identify who those patients are prior to treatment. The EMA[45] reports a number of ancillary analyses that were carried out in an attempt to identify possible sub-groups of patients who might benefit (or failed to receive any benefit) from the treatment. However, the subgroups were small and they determined that no conclusions could be drawn from this *post hoc* analysis.

Adverse events

Other outcomes of critical importance are the rates of AEs and irAEs. As can be seen from Table 6 the rate of AEs is high in all groups with over 95% of all patients experiencing an AE. Of note is that there were 14 (2.2%) deaths in the MDX010-20[19] trial related to the study drugs and seven of these were associated with irAEs. As noted earlier the manufacturer will implement an additional pharmacovigilance programme as part of the requirement of the marketing authorisation for ipilimumab.[17]

The most important AEs are noted to be irAEs. The MS states that these events are manageable and reversible in most cases. Treatment protocols for AEs for these events have been evaluated[46] and are included in the pharmacovigilance plan designed by the manufacturer and accepted by the EMA.[17] Comments from clinicians who have experience in using this therapy concur with the manufacturer's view that these irAEs are manageable and that, as experience with ipilimumab grows, AEs are being identified quickly and treated more proactively now than in the past. Evidence for this was derived from the recent first-line trial[33] of ipilimumab and dacarbazine in which there were no reported deaths from AEs.

Table 6: Adverse events

Adverse events[a]	MDX010-20[19]		CA 184-022[31]		CA 184-007[30]	
Any AE	Ipi + gp100 (n=380) Ipi + placebo (n=131) gp100 + placebo (n=132)	98% 98% 97%	10mg/kg (n=71) 3mg/kg (n=71) 0.3 mg/kg (n=72)	100% 97% 94%	Ipi + budesonide (n=58) Ipi + placebo (n=57)	NR
Any AE (grade 3 or 4)	Ipi + gp100 (n=380) Ipi + placebo (n=131) gp100 + placebo (n=132)	51% 56% 52%	10mg/kg (n=71) 3mg/kg (n=71) 0.3 mg/kg (n=72)	41% 30% 29%	Ipi + budesonide (n=58) Ipi + placebo (n=57)	90% 95%
Any irAE (all grades)	Ipi + gp100 (n=380) Ipi + placebo (n=131) gp100 + placebo (n=132)	58% 62% 32%	10mg/kg (n=71) 3mg/kg (n=71) 0.3 mg/kg (n=72)	70% 65% 26%	Ipi + budesonide (n=58) Ipi + placebo (n=57)	81% 84%
irAE (grades 3 or 4)	Ipi + gp100 (n=380) Ipi + placebo (n=131) gp100 + placebo (n=132)	10% 15% 3%	10mg/kg (n=71) 3mg/kg (n=71) 0.3 mg/kg (n=72)	25% 7% 0%	Ipi + budesonide (n=58) Ipi + placebo (n=57)	41% 38%
AE leading to discontinuation	Ipi + gp100 (n=380) Ipi + placebo (n=131) gp100 + placebo (n=132)	7% 10% 3%				
Treatment related deaths	Ipi + gp100 (n=380) Ipi + placebo (n=131) gp100 + placebo (n=132)	2.1% 3.1% 1.5%				
AE with outcome of death	Ipi + gp100 (n=380) Ipi + placebo (n=131) gp100 + placebo (n=132)	6% 10% 6%	10mg/kg (n=71) 3mg/kg (n=71) 0.3 mg/kg (n=72)	unclear	Ipi + budesonide (n=58) Ipi + placebo (n=57)	NR

[a]Note – data from published papers and MS; ipi=ipilimumab; AE=adverse event; irAE=immune-related adverse event; NR=not reported

The manufacturer provided the ERG with additional information during the clarification process and additional AE data from the MDX010-20 trial[19] are presented in [redacted]7. In their response the manufacturer highlighted that the majority of AEs in the gp100+placebo group were not serious and a large proportion were related to injection site reactions. This raises a concern about the adequacy of the blinding in the trial if almost 20% of patients had a specific reaction to gp100. Of note also is that in the arm of the trial with the least benefit there were two (1.5%) treatment related deaths.

[redacted]7

	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

4.5 *Summary of submitted evidence*

The MS includes three studies in which it presents the clinical[19, 31] and safety[30] data related to the use of ipilimumab in patients who have previously been treated for malignant melanoma.

The submitted data from the key trial MDX010-20[19] demonstrate an OS benefit for patients receiving ipilimumab alone or with gp100 compared to those who received gp100 alone. For both comparisons, the difference in OS benefit is reported as being statistically significant.

The cost of this benefit is the high proportion (16 to 21%) of SAEs associated with ipilimumab which include drug related treatment deaths of 2.5 to 3.1%. The most serious adverse events are irAEs which the manufacturer states can be medically managed.

4.5.1 Critique of submitted evidence syntheses

This is a relatively new drug which has been investigated in manufacturer sponsored studies. The MS provides the results of a search for studies that is appropriate except for the fact that a search of their own internal databases was not reported. Given that the manufacturer holds the data related to the studies conducted to assess the effectiveness of ipilimumab it might have made more extensive use of its own database to identify unpublished studies or to identify data related to drug safety. The rationale for the selection of CA184-07[30] to provide safety data is not immediately apparent.

The results of the search therefore identified only published studies. The inclusion criteria were clearly presented but were not consistently applied and it is unclear how the decisions regarding study inclusion were made, especially in relation to the safety data that are provided.

A number of issues related to trial protocol amendments and analysis were identified. However, given that the (modified) primary outcome was OS, none were considered serious enough to throw doubt on the conduct of the key trial or its results.

Data are descriptively presented with no attempt to combine the data, which is appropriate given the heterogeneity between studies (e.g. differences in drug doses and comparators). The MS is limited by the fact that no studies were identified that included ipilimumab compared to any of the comparators listed in the final scope issued by NICE although the MS makes a case that the gp100 vaccine is clinically equivalent to BSC. Data from the pivotal study[19] demonstrated an OS benefit in favour of ipilimumab+gp100 over gp100+placebo and ipilimumab+placebo over gp100+placebo.

The fact that the MS contains a dose ranging trial highlights the fact that the most beneficial dose of ipilimumab has not been determined and both the FDA[21] and the EMA[17, 45] have added requirements for further investigations that will allow for comparison of the effectiveness of 3mg/kg vs 10mg/kg regimens.

4.5.2 Clinical-effectiveness summary

Clinical results

- Clinical evidence from the key trial demonstrates that ipilimumab provides a statistically significant OS benefit to patients compared to gp100
- Increased rates of AEs and irAEs are seen in patients receiving ipilimumab compared with those observed in other malignant melanoma trials.

Clinical issues

- The comparator used in the key trial does not match any of the comparators listed in the final scope issued by NICE
- Given the negative outcomes in the gp100+placebo arm of the trial, consideration needs to be given to the manufacturer's claim that gp100 is clinically equivalent to BSC
- The overall clinical benefit of ipilimumab is not realised by all patients and there is currently no available markers to predict the subgroups of patients who may benefit the most from treatment
- The SAE and irAE rates in the trial are higher than have been observed in other similar trials of patients with malignant melanoma
- Reported AE and irAE require the implementation of different treatment algorithms for identification and treatment of patients
- Current marketing authorisation recommends continuation of treatment for the full four dose schedule regardless of disease progression with discontinuation only in the event of unresolved serious AE
- Current standard assessment criteria for disease progression may not be adequate to measure patient response to ipilimumab immunotherapy. The timing of assessment is important as, if assessment is performed too early (e.g. before 12 weeks), changes in the patient's health may be missed.

5 ECONOMIC EVALUATION

5.1 *Introduction*

This section provides a structured critique of the economic evidence submitted by the manufacturer of ipilimumab. The two key components of the economic evidence presented in the MS are (i) a systematic review of the relevant economic literature and (ii) a report of the manufacturer's *de novo* economic evaluation. See Table 8 for a summary of the key information points. The manufacturer also provided an electronic version of the EXCEL based economic model.

Table 8 Key information in the MS

Key information	Section (MS)
Details of the systematic review of the economic literature	6.0
Model structure	6.2
Technology	6.2
Clinical parameters and variables	6.3
Measurement and valuation of health effects and adverse events	6.4
Resource identification, valuation and measurement	6.5
Sensitivity analysis	6.6
Results	6.7
Validation	6.8
Subgroup analysis	6.9
Strengths and weaknesses of economic evaluation	6.10
Assessment of factors relevant to other parties	7.0

5.2 *Overview of manufacturer's cost-effectiveness review*

The MS provides a brief description of the review of published cost-effectiveness evidence undertaken by the manufacturer. The databases searched and the search terms used appear to be reasonable and both inclusion and exclusion criteria are explicitly stated. The search by the manufacturer did not identify any relevant studies for inclusion in the review. Although there is no mention of searching within in-house databases for relevant studies, the ERG is confident that no relevant published studies are available for inclusion in the review.

In summary, the manufacturer did not identify any papers that had evaluated the cost effectiveness of ipilimumab in patients with previously treated unresectable malignant melanoma. However, during the review process, 16 full publications were reviewed by the manufacturer to determine if they matched the inclusion criteria for the review. As a result, six studies were excluded from the review. The remaining ten studies were identified as being potentially relevant; however, **none** of the studies identified were both economic models and UK based. The manufacturer summarises the ten potentially relevant studies in the MS (MS, Table 18) and provides a quality assessment of the paper

by Dixon et al[47] published in 2006 as this was the only study conducted in a UK setting. The study by Dixon et al[47] was funded by Roche Pharmaceuticals and the manufacturer concludes that summary values such as life years gained (LYG) and cost per QALY estimates are not comparable to the analysis performed in this MS, as the patient population is different (first-line only).

5.3 *Overview of manufacturer's economic evaluation*

The manufacturer undertook a *de novo* economic evaluation of ipilimumab as a treatment for patients with previously treated unresectable (Stage III or IV) malignant melanoma. In the base-case analysis, ipilimumab is compared with BSC. In the sensitivity analysis, ipilimumab is compared with other comparators such as carboplatin-based chemotherapy and dacarbazine.

5.3.1 Description of manufacturer's economic model

The manufacturer constructed an EXCEL-based cost-utility model. The model is a cohort model with one cohort receiving ipilimumab and the other cohort receiving BSC. The approach used in the evaluation is a "partitioned-survival" model and is similar to a Markov cohort model. However, unlike a Markov model in which the transitions are modelled explicitly using transition probabilities, the "partitioned-survival" model calculates the proportion of patients in each treatment cohort that are expected to be in each health state at any time after treatment initiation.

There are four mutually exclusive states in the model: baseline disease, non-progressive disease, progressive disease and death. Figure 4 shows the model health states.

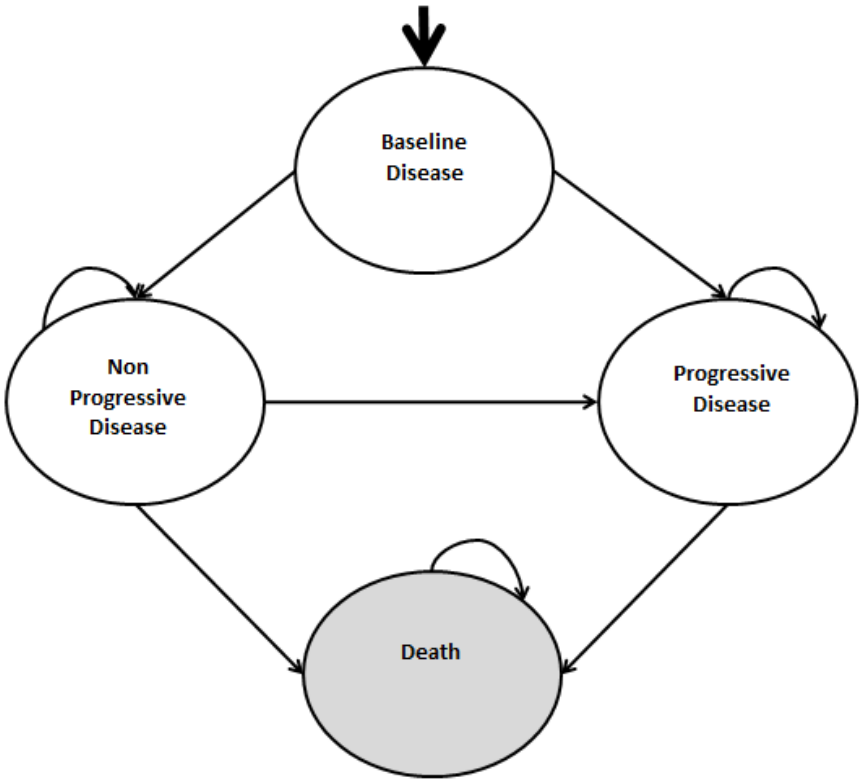


Figure 4 Model health states

5.3.2 Parameters and values

The manufacturer provides a full list of the variables applied in the economic model (MS, Section 6.3). Table 9 lists the key parameters and values identified by the ERG.

Table 9 Summary of variables applied in the model

Parameter	Value used	Distribution	Source
Patient/treatment characteristics			
Average patient body weight	81.7 kg	Normal SD=18.1	UK patients in MDX010-20; compassionate use programme
Patient starting age	56 years	Normal SD=13.4	MDX010-20
% Male	59%		MDX010-20
Ipilimumab dose (mg/kg)	3		MDX010-20
Average number of 200mg vials	0.99		Weights from UK patients in MDX010-20 and compassionate use programme
Average number of 50mg vials	1.24		
Ipilimumab: days between administrations	21		MDX010-20
Survival			
Ipilimumab PFS parameter alpha - second curve	-7.2476	Normal SE=0.354	MDX010-20 (Section 6.3)
Ipilimumab OS parameter alpha - second curve	-6.081	Multivariate normal covariance matrix	MDX010-20 (Section 6.3)
Ipilimumab OS parameter beta - second curve	-0.0032		MDX010-20 (Section 6.3)
BSC PFS parameter alpha - second curve	-6.4148	Normal SE=0.302	MDX010-20 (Section 6.3)
BSC OS parameter alpha - second curve	-6.4148	Normal SE=0.302	MDX010-20 (Section 6.3)
Unit costs			
Ipilimumab administration 1st attendance	£271	Normal SD=124*	NHS Reference costs 09/10
Ipilimumab administration other attendances	£284	Normal SD=61*	NHS Reference costs 09/10
Ipilimumab unit cost	£3,750		BMS, 50mg vial cost
Utilities			
Utility of stable disease	0.81	Beta SD=0.140	EORTC mapped values from MDX010-20
Utility of progressive disease	0.77	Beta SD=0.162	

* Standard deviations calculated using upper and lower quartile values

5.3.3 Treatment effectiveness within the MS

The manufacturer uses the clinical effectiveness data from the MDX010-20 trial[19] in the economic evaluation. The manufacturer combines the two ipilimumab arms from the trial (ipilimumab and ipilimumab+gp100) in order to give the best available estimate of the survival of ipilimumab-treated patients. This is viewed by the manufacturer as a conservative approach as it assumes that gp100 does not have a significant impact on efficacy when the trial data show that the use of gp100 marginally reduces efficacy (PFS and OS). Data on PFS and OS for patients receiving BSC were unavailable. The manufacturer uses the gp100 data from the MDX010-20 as a proxy for BSC data; gp100 and BSC are assumed to be clinically equivalent.

In order to estimate PFS and OS accurately for the two groups of patients, the manufacturer presents the results of two different approaches to parametric curve fitting as part of the survival modelling exercise undertaken. Strategy 1 involves a single curve fit approach and appears to demonstrate that none of the curves (Weibull, Exponential, Lognormal, Log-logistic, Gompertz) fits the Kaplan-Meier data from the MDX010-010 trial.[19] Strategy 2 involves a 2-part curve fit. This approach uses the Kaplan-Meier estimates of PFS and OS for the first 18 months and then uses the 'best-fit' parametric curves beyond 18 months; the 18 month point was chosen to represent the point at which data have started to flatten in both of the ipilimumab arms. The manufacturer concludes that the 'best-fit' curves are as follows:

- Exponential for PFS in ipilimumab
- Gompertz for OS in ipilimumab
- Exponential for OS in BSC
- PFS in BSC arm represented by OS arm

The manufacturer then limits the curve fitting to 60 months and includes only background mortality from this point on; the data appear to show that patients surviving beyond 60 months exhibit long-term survival and die due to natural causes.

5.3.4 Population

The population represented in the economic evaluation consists of patients with advanced (unresectable or metastatic) melanoma in adults who have received prior therapy i.e. people with previously treated unresectable stage III or IV malignant melanoma. The patients in the economic evaluation appear to be matched closely with the patients in the MDX010-20 trial.[19]

5.3.5 Comparator technology

As there are currently no EU licensed second- or third-line treatments, the comparator to ipilimumab may be BSC (as in the base case) or other active treatments (as in the sensitivity analysis). The manufacturer states that use of BSC as the comparator in the base case economic evaluation is a conservative assumption as most of the active therapies available (e.g. dacarbazine, paclitaxel, paclitaxel and carboplatin, carboplatin, cisplatin) are more costly than BSC in terms of drugs administration or AE costs. The manufacturer does not explicitly define what comprises BSC for this group of patients; however it is clear from the MS that no drugs costs are included in the BSC cost.

5.3.6 Health related quality of life

Health related quality of life data were collected in the MDX010-20 trial[19] using the EORTC QLQ-C30 and the short form 36 (SF-36); the manufacturer therefore had to map data using validated mapping algorithms in line with the NICE reference case.

The EORTC QLQ-C30 values were mapped from the 971 trial observations using "mapping algorithm 3" as described in the paper by Rowan et al.[48] Algorithm 3 was chosen by the manufacturer as it had the following properties: a low mean absolute error (0.046); no inconsistencies within the model; lowest number of large errors (>5 or 10% out). SF-36 observations (n=963) were mapped using the SF-6D algorithm; the mean root square of the model is 0.088 and the model was validated using three datasets.

Following mapping of individual patient observations, the average for each patient for the "progressed" and "not progressed" states was calculated and the averages for the overall health states were calculated based upon the individual patient averages in order to avoid bias towards patients with more observations.

The manufacturer also conducted a systematic review in order to identify studies which included HRQoL data for patients with metastatic melanoma. One study by Beusterien et al[49] was identified by the manufacturer which included 63 patients from the UK and 77 patients from Australia. A summary of the HRQoL values from the three different sources is presented in Table 10.

Table 10 Summary of HRQoL values available

Health state	Utility value: EORTC QLQ-C30	Utility value: SF-36	Utility value: Beusterien et al
Progression free disease	0.80 (95% CI [0.53;0.97])	0.64	0.77
Progressive disease	0.76 (95% CI [0.46;0.97])	0.62	0.59
Difference	0.04	0.02	0.18

The manufacturer states that EORTC QLQ-C30 trial data are utilised in the economic model as these data most closely meet the NICE reference case and are relatively consistent with the values published in the Beusterien et al[49] paper. The manufacturer explores the impact of changing the source of the utilities in the model on the size of the ICER in the sensitivity analysis.

Adverse events

Adverse event rates for patients receiving ipilimumab in the model were estimated using results from the MDX010-20 trial,[19] AE rates for the BSC population were estimated from the gp100 arm of the trial.

The model includes the impact of AEs on costs. In the ipilimumab arm grade 3+ AEs with an incidence of at least 3% were included as well as grade 2 diarrhoea and colitis. The costs utilised for AEs have been taken from research conducted by Oxford Outcomes.[50] Table 40 in the MS provides a summary of the per patient per episode AE costs used in the model. The costs are based upon a microcosting approach which estimates the proportion of patients treated as inpatients versus outpatients for each condition and assigns a cost to each type of treatment. The costs for endocrine disorders are assumed to be incurred every 6 months; all other AE costs are assumed to be incurred once at the start of the model.

The impact of AEs on utilities is included in the health state utility weights derived from the HRQoL data collected in the MDX010-20 trial.[19]

5.3.7 Resources and costs

The manufacturer conducted a literature search to identify papers relating to the costs of resource use associated with ipilimumab. No papers were identified by the manufacturer.

NHS costs

Ipilimumab is administered as a 90 minute intravenous infusion and therefore is costed as a day case attendance. NHS Reference Costs (HRG SB 13Z [first attendance] and HRG SB 15Z [subsequent attendance]) were used in the economic evaluation.

Intervention and comparators' costs

As ipilimumab has been recently licensed, there are no British National Formulary (BNF) costs available. The manufacturer has recently confirmed the list price of ipilimumab as £3750 per 50mg vial and £15,000 per 200mg vial. The manufacturer estimates that the average number of vials used per patient dose is as follows: 1.24 X 50mg + 0.99 X 200mg. In the base case economic evaluation, there is wastage and the model rounds the usage per patient to the nearest vial.

The manufacturer does not explicitly mention the costs of BSC under the heading of 'intervention and comparators' costs' in the MS.

Health state costs

The manufacturer presents detailed disease management micro-costing information in the MS (MS, Table 37) and discusses six different cost categories:

- On initiation of treatment (one-off)
- On treatment pre-progression (monthly)
- BSC cost (monthly)
- On progression cost (one-off)
- Palliative care off treatment (monthly)
- Terminal care (one-off)

In addition, the MS presents detailed unit costs associated with ipilimumab and also provides a list of health states and associated costs as used in the economic model. These tables are replicated in this report (see Table 11 and Table 12).

Table 11: Unit costs associated with the technology in the economic model

Items	Intervention (confidence interval)	Reference to section in MS	Comparator 1 (confidence interval)	Reference to section in MS
Drug costs	£3750 for 50mg vial £15,000 for 200mg vial	Section 1.1 and 1.1	£0 - BSC assumed no drug costs	Section 6.5
Treatment initiation cost	£365 (£19 to £538)	Section 6.5.	£0 - BSC assumed no treatment initiation	
Drug administration costs	£271 (£67 to £474) 1st administration £284 (£184 to £384) subsequent administrations	Section 6.5.	£0 - BSC assumed no drug administration	
Monthly cost of routine treatment pre progression	£162 (£92 to £231)	Section 6.5.	£378 (£244 to £511)	
Cost on progression	£648 (£338 to £958)	Section 6.5.	£648 (£33 to £958)	
Monthly cost of treatment post progression prior to palliative care	£378 (£244 to £511)	Section 6.5.	£378 (£244 to £511)	
Monthly cost of palliative care (4 months)	£838 (£295 to £1642)	Section 6.5.	£838 (£295 to £1642)	
Terminal care cost	£5401 (£0 to £13,752)	Section 6.5.	£5401 (£0 to £13,752)	

Data source: Oxford Outcomes[50](2011), Improving choice at end of life, King's Fund[51] (2008)

Table 12: List of health states and associated costs in the economic model

Health states	Items	Value	Reference to section in MS
Progression-free disease	Drug costs	Ipilimumab= £19,565 per dose; BSC=£0	Section 6.5
	One off treatment initiation cost	Ipilimumab=£365; BSC=£0	
	Drug administration	Ipilimumab= £271(1st administration), £284 per administration thereafter BSC= £0	
	Routine treatment per month	Ipilimumab=£162; BSC=£378	
Progressive disease	One off cost on progression	£648	
	Routine treatment per month	£378	
	Palliative care per month (4 months before death)	£838	
Death	One off terminal care cost	£5401	

5.3.8 Perspective, time horizon and discounting

The economic evaluation was undertaken from the perspective of the NHS and Personal Social Services (PSS). The time horizon set was 30 years (lifetime) and this was considered by the manufacturer to be adequate to capture any complete differences between comparators (as per the NICE reference case). Both costs and benefits were discounted at 3.5% per annum. The manufacturer presents the cost-effectiveness results without discounting in the sensitivity analysis.

5.3.9 Model validation

The manufacturer details a number of steps that were taken to validate the model (MS, Section 6.8) including:

- Estimates of PFS and OS from the final model were checked against values calculated in a separate spreadsheet – results were the same
- Extensive one-way sensitivity analyses were conducted on all model inputs and results were reviewed to ensure that changes in cost and effectiveness measures were consistent with expectations, given model specifications
- Random checks were made on model inputs compared with source data
- In terms of internal validity, as noted above the survival functions used to generate estimates of PFS and OS for ipilimumab are very close to those obtained based on the empirical (Kaplan-Meier) survival distributions (see MS, Section 6.3).

In addition, the model was presented (face-to-face) to four practicing UK clinicians to check the face validity of the model. The model was also presented to an advisory board of six UK health economists. Finally, the model was checked by a senior health economist. The manufacturer considered the feedback from all of those involved in the peer-review process and changes were made to the model and documentation where appropriate.

5.3.10 Results included in the MS

In the base case economic evaluation, the manufacturer presents cost-effectiveness results for the comparison of ipilimumab vs BSC as shown in Table 13. Table 14 and Table 15 show disaggregated results for the incremental costs and benefits of ipilimumab vs BSC.

Table 13 Base-case results

Technology	Total			Incremental			ICER (incremental cost per QALY)
	Costs	Life years gained	QALYs	Costs	Life years gained	QALYs	
BSC	£12,837	1.33	1.01				
Ipilimumab	£96,188	3.19	2.38	£83,351	1.86	1.37	£60,737

Table 14 Summary of QALY gain by health state (discounted)

Health state	QALY Intervention (Ipilimumab)	QALY Comparator (BSC)	Increment	Absolute increment	% Absolute increment
Progression free	0.48	0.25	0.23	0.23	17%
Post progression	1.90	0.76	1.14	1.14	83%
Total	2.38	1.01	1.37	1.37	100%

Table 15 Summary of costs by heath state (discounted)

Health state	Cost Intervention (Ipilimumab)	Cost Comparator (BSC)	Increment	Absolute increment	% Absolute increment
Progression free	£78,739	£1,908	£76,831	£76,831	92%
Post progression	£17,449	£10,929	£6,520	£6,520	8%
Total	£96,188	£12,837	£83,351	£83,351	100%

5.3.11 Sensitivity analyses

Methods

The manufacturer carried out a wide range of sensitivity analysis: one-way sensitivity analysis, scenario analysis and probabilistic sensitivity analysis (PSA).

One-way sensitivity analysis: the manufacturer conducted deterministic analysis on key variables (n=16) using the 5% and 95% confidence intervals for the variables considered. The only variable not included in the sensitivity analysis was the dose (mg/kg) of ipilimumab as this dose is fixed.

Scenario analysis: eight different scenario analyses are discussed in the MS and are described in the MS (MS, Table 41).

Probabilistic sensitivity analysis was conducted by repeated sampling; 1000 Monte-Carlo simulations were performed to provide sufficient runs to allow PSA results to stabilise.

Results

The manufacturer presents a summary tornado diagram (MS, Figure 27) showing all variables that cause a change of >= +/- £200 on the base case ICER. The variable which most affects the size of the ICER is the utility assumed for PD. Other variables which significantly affect the size of the ICER include: the curve fit parameters for OS for ipilimumab; the price of ipilimumab; the curve fit parameters assumed for OS for BSC; the cost of palliative care; and the patient's starting age.

The manufacturer shows detailed sensitivity analysis results (MS, Tables 48 to 50) for the following variables: vial sharing allowed (ICER falls to £55,824 per QALY gained); minimum dose of ipilimumab (ICER falls to £38,387 per QALY gained); maximum dose of ipilimumab (ICER rises to £88,788 per QALY gained). When patients receive all four doses of ipilimumab the ICER is £70,163 per QALY gained. The impact of changing the utility value associated with PD is most influential and the manufacturer shows that utility values in the range of 0.6 to 0.8 yield ICERs in the range of £73,854 per QALY gained to £58,411 per QALY gained.

The manufacturer also shows detailed results of the scenario analyses undertaken (MS, Table 51 to Table 58). These results are summarised by the ERG in Table 16.

Table 16 ERG summary of manufacturer's scenario analyses results

Scenario description	ERG summary
(1) No discounting	ICER **reduces** to £42,871
(2) Alternative comparators to BSC	ICER **reduces** in all situations
(3) Alternative utility estimates	Use of SF-36/SF-6D utility values from MDX010-20 trial and Beuresterien et al paper **increases** the size of the ICER; using drug specific utility values **reduces** the size of the ICER; adjusting utilities for age only slightly affects the ICER
(4) Maximum dosing assumptions	ICER **increases** when: patients receive all 4 doses (£70,163 per QALY gained); 50% more patients receive induction; ICER **decreases** when 50% fewer patients receive induction
(5) Alternative curve fits	ICER **reduces** when: one curve fit/BSC arm/ best AIC /without background mortality; one curve fit/BSC arm/ Weibull /without background mortality; two part curve fit/best AIC/without background mortality; two part curve fit/best AIC/with background mortality; two part curve fit/IPI only/best AIC/ with background mortality only after 5 years; two part curve fit/IPI only/Weibull/with background mortality only after 5 years ICER **increases** when: one curve fit/both arms/best AIC/without background mortality; one curve fit/both arms/Weibull/without background mortality; two part curve fit/Weibull/ without background mortality; two part curve fit/both arms/Weibull/with background mortality only after 5 years
(6) Use of alternative data for ipilimumab	Use of ipilimumab only data **reduces** the size of the ICER; use of ipilimumab+gp100 data **increases** the size of the ICER
(7) Use of alternative time horizons	As expected reducing the time horizon **increases** the size of the ICER; when a lifetime horizon is used, the ICER decreases
(8) Use of alternative weight data	Using UK patient weights from MDX010-20 trial slightly **increases** the size of the ICER; using weights from the compassionate use programme very slightly **decreases** the size of the ICER

A scatterplot of the manufacturer's PSA results is presented in Figure 5. The cost effectiveness acceptability curve from the manufacturer's PSA shows that there is approximately a 14% chance of ipilimumab being cost effective vs BSC at a threshold of £50,000 per QALY gained.

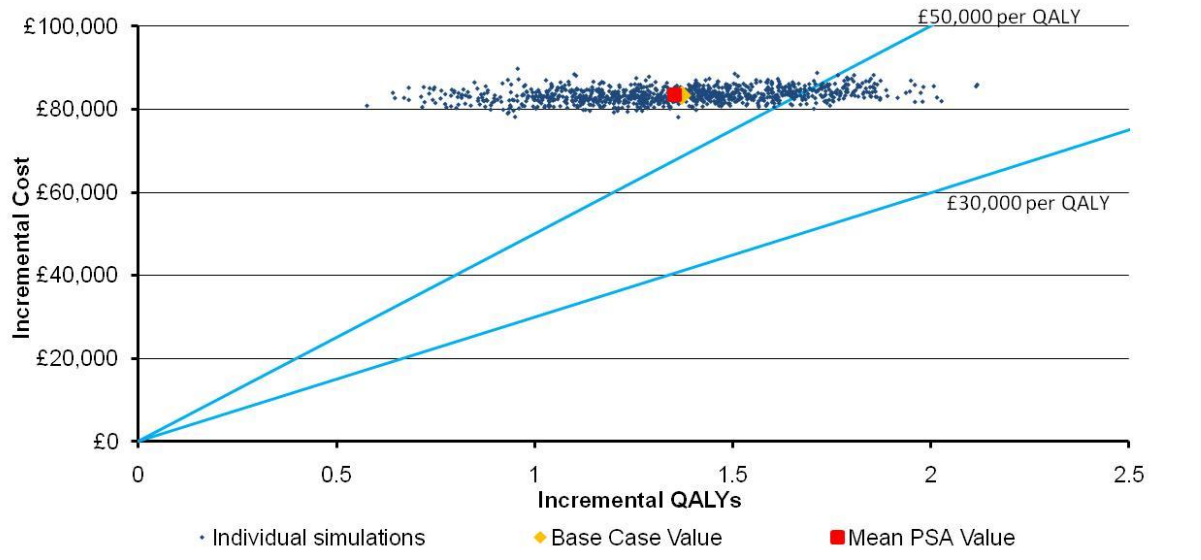


Figure 5 Scatterplot of probabilistic sensitivity results

5.4 *Assessment of manufacturer's economic model*

This section summarises the ERG's assessment of the manufacturer's economic model against (i) NICE reference case checklist and (ii) Drummond 10-point checklist.

Table 17 shows how closely the manufacturer's submitted economic evaluation accords with the requirements for a base-case analysis set out in the NICE reference case checklist. The single difference between the manufacturer's approach and the NICE reference case checklist is that the manufacturer's economic evaluation considers a 30 year time horizon rather than a lifetime horizon.

Table 18 summarises the ERG's appraisal of the economic evaluation conducted by the manufacturer using the Drummond 10-point checklist. In addition to concerns about the intervention and the comparator, the ERG has several important criticisms of the submitted economic evaluation. The manufacturer compares ipilimumab vs BSC. The intervention described in the model is ipilimumab – clinical data are derived from combining two arms of the MDX010-020 trial[19] where there are three times as many patients in the ipilimumab+gp100 arm as in the ipilimumab only arm; no details of the combination process are presented. The comparator described in the model is BSC; whether or not BSC is truly reflected by the use of gp100 vaccine is not known. The ERG proposed several corrections/modifications to the submitted economic model including re-estimation of costs (drug acquisition costs and AE costs) and re-calculation of utilities for the two health state parameters using an improved age-adjustment. The ERG also suggested that the continuity correction method used in the model was not needed and that the background mortality logic described was incorrectly applied. The ERG considers that the main weakness of the model is related to the manufacturer's overestimation of survival gains and that the base case ICER is substantially underestimated. The ERG agrees with the manufacturer that survival estimation for this group of patients is complex and is of the opinion that further work in this area is required in order to generate meaningful ICERs.

Table 17 NICE reference case

Attribute	Reference case	Does the *de novo* economic evaluation match the reference case?
Decision problem	The scope developed by the Institute	Yes
Comparator(s)	Alternative therapies routinely used in the NHS	Yes - the manufacturer compared ipilimumab with BSC (using gp100 data as a proxy for BSC). There is no gold standard comparator for 2nd line treatment of patients with multiple melanoma. Clinical expert advice agrees that BSC is a relevant comparator
Perspective costs	NHS and Personal Social Services	Yes
Perspective benefits	All health effects on individuals	Yes
Form of economic evaluation	Cost-effectiveness analysis	Yes
Time horizon	Sufficient to capture differences in costs and outcomes	Yes (30 year time horizon)
Synthesis of evidence on outcomes	Systematic review	Yes - treatment effectiveness was derived from the pivotal ipilimumab vs gp100 RCT for patients who had been previously treated
Outcome measure	Quality adjusted life years	Yes – QALYs and incremental cost per QALYs were estimated by the manufacturer
Health states for QALY	Described using a standardised and validated instrument	HRQoL data were collected in the key trial and EORTC QLQ-C30 data were mapped to EQ-5D scores using algorithm 3 in the Rowan et al publication EQ-5D scores are based on TTO benefit valuation methods and the source of preference data for valuation of changes in HRQoL is a representative sample of the general public
Benefit valuation	Time-trade off or standard gamble	
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	
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
Sensitivity analysis	Probabilistic sensitivity analysis	Yes - deterministic, scenario and probabilistic analyses were undertaken by the manufacturer

PSS= Personal Social Services; TTO= time trade off

Table 18 Critical appraisal checklist

Item	Critical appraisal	ERG comment
Was a well-defined question posed in answerable form?	Yes	However, question was difficult to answer primarily because there are no gold standard treatments for this group of patients and also because many of the treatments offered to these patients are being used off-licence
Was a comprehensive description of the competing alternatives given?	Yes	To a certain extent. The manufacturer assumes that the efficacy of BSC=gp100; however, there is some evidence to suggest that gp100 might be less efficacious than BSC and the ERG considered that this view was not fully discussed in the MS. The MS does not include a description of BSC
Was the effectiveness of the programme or services established?	Yes	Whether or not BSC=gp100 is not known; gp100 is an experimental anti-cancer agent whose effectiveness is not well established. The manufacturer combines the two ipilimumab arms of the pivotal RCT in order to make best use of the trial data – in doing so, efficacy of ipilimumab is slightly reduced but a larger data set becomes available. The manufacturer does not elaborate on the methods used to combine the data from the two arms of the trial – there are three times as many patients in the ipilimumab+gp100 arm as in the ipilimumab only arm
Were all the important and relevant costs and consequences for each alternative identified?	Yes	ERG notes that several of the NHS Reference costs were mislabelled (e.g. text stated out-patient price but used day-case price)
Were costs and consequences measured accurately in appropriate physical units?	Partial	ERG modified drug acquisition costs to reflect different weight distribution of males and females; ERG is of the opinion that all drug administration costs should be estimated as day-case costs; ERG identified a background mortality logic error
Were the cost and consequences valued credibly?		ERG considers that the manufacturer's estimates of OS are substantially flawed (over-estimated); two health state utility parameters used in the model require an age-adjustment; ERG corrected an error in the AE costs which reduced the size of the ICER by approximately £1000 per QALY gained
Were costs and consequences adjusted for differential timing?	Yes	Discounting was applied appropriately. ERG does not consider that the continuity correction method adopted was appropriate or correctly applied
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	
Was allowance made for uncertainty in the estimates of costs and consequences?	Yes	Deterministic, scenario and probabilistic sensitivity analyses were performed
Did the presentation and discussion of study results include all issues of concern to users?	Partial	The effect of an increased/decreased price of ipilimumab was not considered as part of the sensitivity analysis exercise described in the MS

5.5 *Detailed critique of manufacturer's economic model*

5.5.1 Model design and implementation

The submitted EXCEL economic model is generally well constructed with adequate references to data sources. The ERG initially found the model logic quite difficult to follow due to the extensive use of range name labels in formulae. It would have been helpful to include a list of range names used in the model with a simple description of each parameter's meaning, and its location by worksheet/cell. In addition, a simple table of worksheet names with a description of the function of each sheet in relation to other sheets would be helpful.

5.5.2 Survival projection

This appraisal of a treatment for advanced or metastatic cancer is unusual in two respects:

- on the basis of results from a single RCT the manufacturer of ipilimumab claims a substantial improvement in mean survival, for some patients amounting to several years;

- the natural history and prognosis for this condition are not well understood, and there is currently no recognised standard treatment which has been shown to give real benefits over BSC for many patients.

A study published in 1999[52] (involving reanalysis of eight trials of IL-2 in treating patients with metastatic melanoma) illustrates a common interpretive problem facing researchers in this field. Figure 6 reproduces Figure 2 of Atkins et al review[52] which shows a very high initial mortality rate, such that about 80% of patients had died within 2 years. However, those patients surviving after 2 years of follow-up then appear to have suffered little or no mortality for a further 9 years. This remarkable pattern of response is replicated in other trials including the MDX010-20 trial[19] and observational studies.[52]

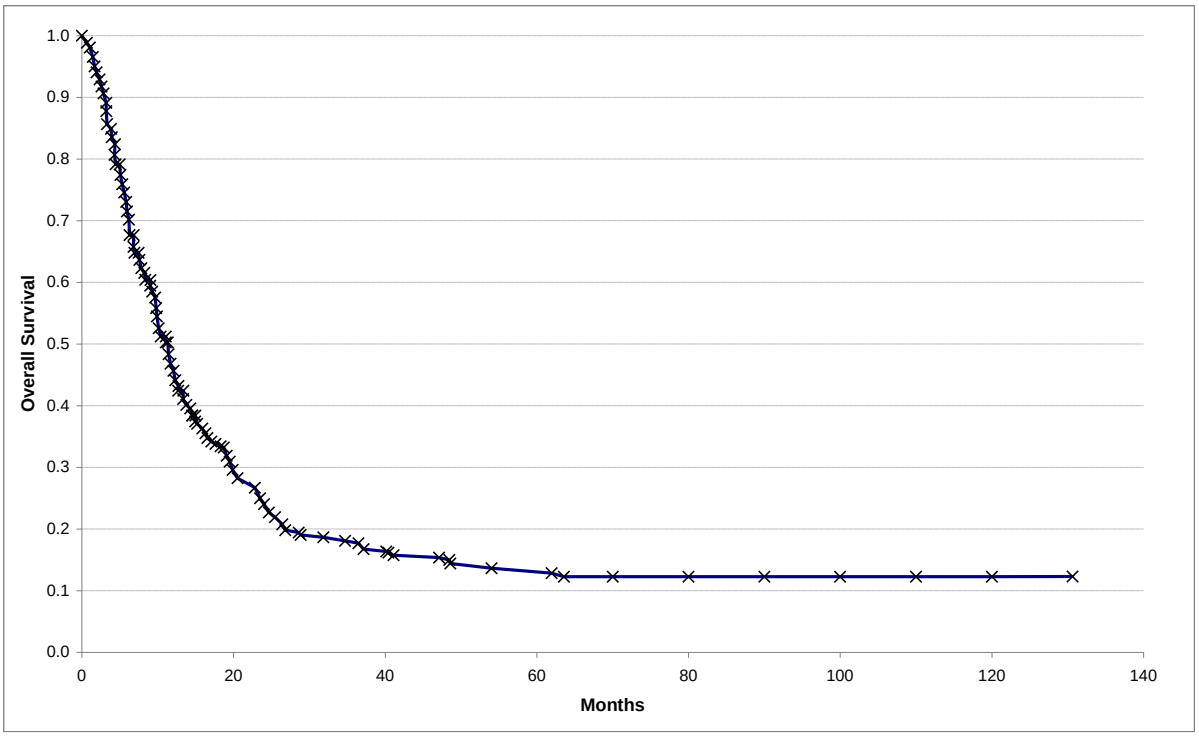


Figure 6 Kaplan-Meier survival for Atkins interleukin-2 study

An initial high attrition rate is common in the later stages of advanced and metastatic disease, but the extended survival tail is problematic without suggesting that a minority of patients are effectively cured of their condition by a treatment normally expected to yield only modest benefits. The only other possible explanation is that this population is severely heterogeneous exhibiting different prospects of survival for distinct subgroups. The particular practical difficulty in using such data to populate a decision model is that the information available for analysis is at its weakest exactly at the point where enhanced survival is likely to generate the most added life years from the novel therapy. The natural conclusion is therefore that although the MDX010-20 trial[19] may have been adequately powered to demonstrate that a survival advantage exists for ipilimumab, it was probably underpowered to provide reliable quantification of that benefit, requiring substantially more patients (especially in the non-ipilimumab arm) surviving after 2 years follow-up. Moreover, in order to furnish sufficient information to characterise the long-term pattern of survival, a trial would also need to extend the follow-up period by several years.

In section 6.3 of the MS (pgs 103-114) a detailed description is provided of the steps taken by the manufacturer in their attempt to carry out parametric modelling of the MDX010-20 trial[19] data as a basis for projecting outcomes beyond the follow-up period. The first approach was to fit standard survival functions to data from the full duration of the trial, but this was found to be unsuccessful as the extreme inflexion in the survival curve could not be credibly replicated with a single parametric function. The second strategy was to fit 2-part models to time periods before and after 18 months

follow-up. This achieved some improvement in representing the trial data, but some problems remained, which required additional adjustments to be applied to the 2-part survival curves in the final decision model:

- in the absence of sufficient progression-free patients alive in the gp100 arm of the trial beyond 18 months, the model authors employed the same Log-logistic survival model to represent both PFS and OS for the comparator in the decision model. This has the effect of ensuring that there can be no patients in a post-progression condition for the comparator arm at any time after 18 months, thus introducing a notable bias into one arm of the model;

- it was observed that the fitted OS functions beyond 18 months follow-up generate mortality risks lower than those in the general population at a comparable age, and as a consequence the model predicted significant numbers of patients surviving to unreasonably advanced ages (beyond 100). To counter this anomaly, the modellers replaced calculated model mortality risks by the mortality experienced in the general population beyond 5 years follow-up.

Taken together these adjustments indicate that the modellers failed to achieve a coherent and credible interpretation of the MDX010-20 trial[19] data on which to predict future outcomes, and to allow reliable estimates of patient benefit to be made. In particular, reversion to general mortality rates implies that any patient surviving beyond 5 years of second-line systemic treatment is effectively completely cured of their metastatic disease. No evidence has been submitted to support such a strong claim.

It is also noteworthy that the Gompertz function employed as the basis for projecting OS in the ipilimumab arms is frequently associated with especially 'long tails' so might be expected to overstate future benefits for the intervention therapy.

5.5.3 Costs

Drug acquisition costs

Ipilimumab doses are calculated as 3mg/kg of body weight. In the manufacturer's base case analysis, doses are calculated individually for UK patients in the MDX010-20 trial[19] combined with patients in the compassionate use programme assuming unused part-used vials are not shared (full wastage). This approach is preferable to using simple averages of body weight, but is still likely to include some bias due to not recognising that males and females exhibit different weight distributions, and that the relative proportions of males and females in the trial and the compassionate use programme are very different, with larger numbers of patients in the compassionate use programme than in the trial.

The ERG has applied a more structured approach to costing ipilimumab doses. Examination of the weight distributions of UK patients in the trial confirms the general pattern in human populations that body weight in both sexes is skewed and is better described by a Log-normal than a normal distribution. Log-normal parameters were estimated separately from the male and female UK trial patients using the method of moments, and the proportion of patients requiring each whole number of 50mg vials to receive a dose of 3mg/kg was calculated. The overall mean number of vials per patient dose was then weighted by the relative proportions of males and females, which was estimated from the number of deaths in 2008 in England and Wales from malignant melanoma (54.5% males, 45.5% females).[53] The ERG calculation led to a mean of 5.161 50mg vials (or equivalent in 200mg vials) per patient-dose, compared to 5.217 in the base case analysis, indicating that the model ipilimumab acquisition costs should be reduced by 1.07%. This has the effect of reducing the base case ICER by £563 per QALY gained.

The manufacturer's base case uses BSC as the comparator so that there are no drug costs for systemic treatment. However, sensitivity analyses are undertaken for a range of different drugs commonly used for second-line therapy. The ERG has estimated the acquisition cost of the most common of these options – dacarbazine. The standard dose is calculated by body surface area (BSA) at 850mg per m^2. Unlike body weight, BSA conforms to a normal distribution and parameters provided separately for males and females in the UK trial population by the manufacturer were used to estimate a mean cost per patient-dose. The ERG's estimate is £55.13 compared to £53.35 used in the manufacturer's model – an increase in 3.3%. This also slightly improves the ICER in favour of ipilimumab for sensitivity analyses involving dacarbazine.

Drug administration costs

The submitted model employs the latest NHS Reference Costs[54] for the cost of administration of systemic treatments. In most cases (ipilimumab and most comparators) the values used are for day-case sessions (although mislabelled as inpatient sessions); the only exception is for two oral medications (temozolomide and imatinib) where outpatient costs are used instead. The ERG considers this to be inappropriate as in most cases it is likely that all patients will be seen in a specialist day-case unit. This change has the effect of reducing the cost of the first attendance from £171 to £152, but to increase the cost of subsequent attendances from £171 to £284. However, since the base case analysis does not involve any systemic treatments as comparators, these alterations have no impact on the size of the base case ICER.

Health state costs: disease monitoring, supportive care and terminal care

The submitted model includes two one-off costs (at the start of treatment, and on progression), three state-related costs (BSC, pre-progression/on treatment, and palliative care/off treatment) and an estimate for terminal care. All of these costs appear to be well sourced, and to be given realistic values. Model results do not appear to be very sensitive to these parameters.

Adverse event costs

[redacted]

Comparison of the model logic with the account presented in the MS indicates that an implementation error has been made which inflates the costs of treating AEs. This results from including the duration of an AE event in the calculation of costs, when the unit costs in the model already take account of the average duration of each event. Thus where fatigue is treated for 15 days, the expected cost of £11.21 for ipilimumab patients (6.1% of £162.51) is increased to £168.20. Overall, the correct value of AE costs for initial treatment with ipilimumab is £139.50 in place of £1612.35 used in the manufacturer's base case. This error is present in the ipilimumab model arm but has been corrected in the comparator arm, with the result that the base case ICER for ipilimumab is reduced by about £1073 per QALY when the error is corrected.

5.5.4 Utility values

The submitted model offers five options for the scheme of utility values to apply to health state utility calculations, selecting mapped values from EORTC-QLQ-C30 data from the pivotal trial for the base case analysis. In the absence of any time-trade-off utility measurements for this patient population being available in the published literature, and the omission of such a utility measure in the trial, the use of a mapping exercise to convert HRQoL data to a NICE preferred utility scale was inevitable. Conversion of SF-36 data using the SF-6D algorithm has been documented to produce reduced utility measures and reduced incremental effects compared to EQ-5D and therefore leads to inconsistency with EQ-5D based evaluations (the NICE norm). The only other cited source (Beusterian et al[49]) which suggests much greater decrements for progressive disease is not based on patient data, and

therefore may not capture the unusual characteristics of advanced melanoma disease in which simple categories based on a single determination of disease progression may be less meaningful for those patients benefitting from extended survival. The ERG considers that the manufacturer's case for using the EORTC-QLQ-C30 data mapped to EQ-5D utility values is the best approach available, albeit clearly less satisfactory than direct measurement.

The utility values used in the model are adjusted over time using a complex procedure in which the proportions of the chosen population (trial, compassionate use or combined) at each age in years at entry are projected unchanged over time, and the corresponding age-related standard utility values as reported by Kind et al[55] are applied to re-estimate a mean utility value at each succeeding year. The ratio of this resulting utility to the base line utility is then used each succeeding year as a multiplier to adjust the utilities calculated by the model.

This approach involves at least three important problems:

- it is inconsistent with the overall structure of the model which treats all patients as beginning treatment with a single fixed age, set to a predetermined average value

- the projection of patients over time makes no allowance for cumulative mortality (i.e. assumes all patients are immune from death)

- it assumes that the age-adjustment to utility should be proportional to the unadjusted utility.

The ERG has analysed the results by Kind et al[55] and concludes that the relationship between age and mean utility in the general population is linear and not proportional, amounting to -0.004114 per year of survival beyond baseline. The effect of applying this simpler alternative method of adjustment is to increase the base case estimated ICER by £1054 per QALY gained.

5.5.5 Minor amendments and corrections

Background mortality logic error

In order to prevent projected mortality rates falling below those of the general population the submitted model includes an option to switch to background population mortality rates after 5 years follow-up. A logic error has been detected which has the effect of implementing this change 1 year earlier than expected. When corrected this decreases the base case ICER by £2188 per QALY gained.

Projective survival logic error

Errors have been detected in the implementation of Weibull functions (not the 'broken curve' option) for PFS and OS in the 'ipilimumab only' population and the combined ipilimumab population. This is due to referencing elapsed time in years rather than days as required by the model parameters. This error only affects supporting results displays, but it does not affect the main logic of the model so does not alter the size of the estimated ICER.

Continuity correction

The submitted model updates all estimates daily for the first 5 years and then weekly thereafter. With this frequency of updating no continuity correction is appropriate. However, the manufacturer has implemented a 'half-cycle' correction which is only valid for simple 'smooth' patterns of outcome and resource use. If a continuity correction were deemed necessary, a 'mid-cycle' correction would be more suitable for this model design. The impact of this issue is in fact minimal.

5.6 *Summary of ERG's critique of the manufacturer's submitted economic model*

The ERG proposes a number of minor corrections/modifications to the manufacturer's submitted economic model. None of the suggested changes leads to a substantial increase or decrease in the size of manufacturer's base case ICER. For example, correcting drug acquisition costs to reflect the different weight distributions of males and females leads to a 1% decrease in the size of the ICER and addressing an error in AE costs leads to a reduced ICER by approximately £1000 per QALY gained. The main weakness of the manufacturer's model is the estimation of mean OS – the ERG considers that the manufacturer's approach to estimating survival gains is inconsistent and is flawed. In Section 6 the ERG discusses two different approaches to estimating OS from the clinical data available.

6 ADDITIONAL WORK CARRIED OUT BY THE ERG

6.1 *Alternative approaches to estimating survival gains*

The ERG has carried out further statistical modelling aiming to develop methods which correspond more closely to the MDX010-20 trial[19] data, and thus avoid the inconsistencies described previously.

Initially, the ERG considered whether any of the available baseline patient characteristics were able to distinguish between patients dying early and those surviving into the 'long tail' of the distribution. Information was requested from the manufacturer comparing the proportion of patients exhibiting each factor who were alive at 215 days, with the proportion of patients who died or were censored for survival prior to 215 days. The manufacturer provided the information requested in Tables B3.1-4 of the clarification response (see Appendix 2). This showed modest but significant differences in all of the clinical variables already known to be indicators of extended survival, but no significant difference for gender or the age of the patient. However, a consistent additional strong indicator of extended survival was the achievement of an objective response to treatment (complete or partial response). Therefore the ERG undertook a two-way analysis, using both BORR at 24 weeks and treatment arm as strata within a Kaplan-Meier analysis for each of the following outcomes: OS, PFS and post-progression survival (PPS). The results indicated clearly that analysis by response category (response, stable disease, progressive disease and 'not recorded') was sufficient to explain much of the variation in outcomes apparent between the trial arms. This accorded with the observation by Atkins et al[52] that the patients in the 'long tail' of the OS distribution were predominantly responders to treatment. The ERG therefore attempted detailed modelling of the trial data by response category, combining patients in all three arms, and then estimated the expected survival in each trial arm by assigning the values obtained for each response group weighted by the relative proportions of patients in each response group for the trial treatment arms.

The ERG considered two approaches to estimating OS from the trial data:

- by direct modelling of OS trial outcome data
- by separate modelling of PFS and PPS, combining estimates to yield an estimate of OS.

6.1.1 Modelling of overall survival: ERG approach

Direct modelling of overall survival

For direct modelling of OS, four response category models were developed and calibrated from the MDX010-20 trial OS data, combining all three treatment arms. All models employed an initial event-free period and 2-part exponential models where warranted by the cumulative hazard data. The Kaplan-Meier data and fitted models are illustrated in Figure 7. It should be noted that the number of cases and events available for analysis for the response (R) group is very limited (40 patients and only four deaths) and therefore there is considerable uncertainty attached to the choice of model for these patients.

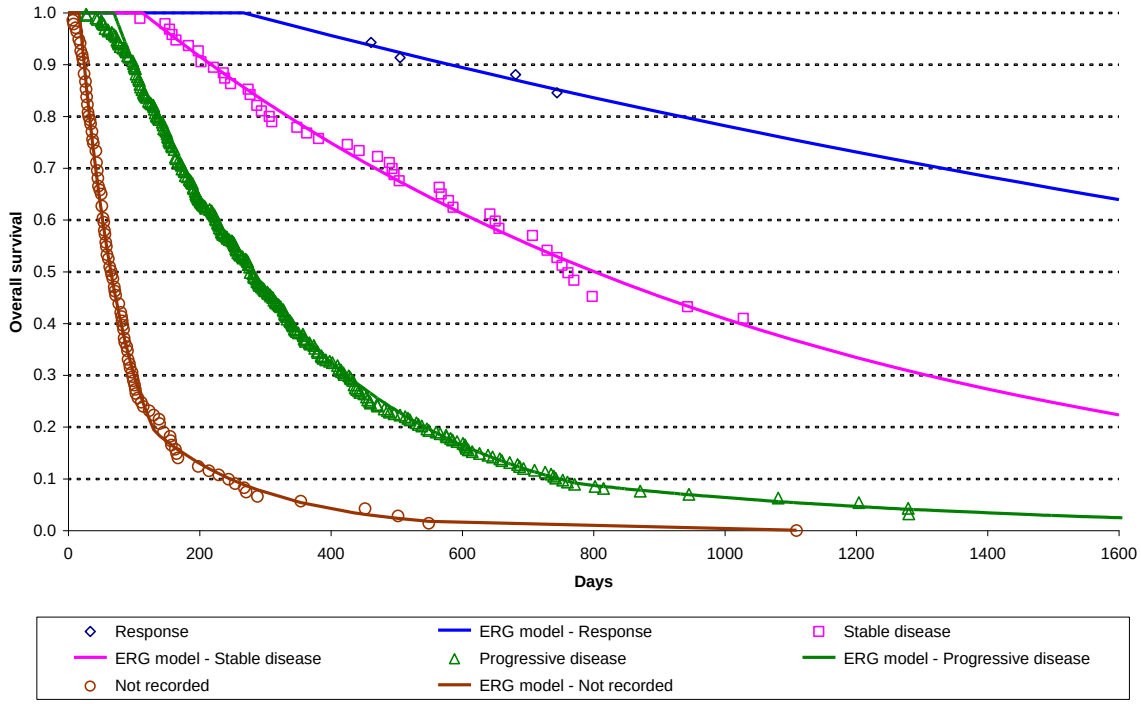


Figure 7 Overall survival modelled by response category

Indirect modelling of overall survival: ERG approach

For indirect modelling of OS, PFS and PPS data were modelled separately in a similar manner.

For PFS, the fitted models were more complex requiring 2- or 3-part models to be developed.

For PPS, the trial data were more limited; for all response groups, simple exponential functions were adopted for modelling.

The correspondence between Kaplan-Meier results and fitted PFS and PPS models are shown in Figure 8 and Figure 9.

Both approaches yielded much smaller estimates of OS than obtained by the manufacturer's model, with life extension of between 5 and 17 months. However, when the results were compared with the trial Kaplan-Meier results by treatment, the fit obtained with both models was found to be unacceptable, presumably due to additional variation not accounted for by response category alone.

An attempt was made to reduce uncertainty and improve accuracy by using only two response categories (combining responding and stable disease patients, and combining the progressive and 'not recorded' categories). However, estimates of survival gain obtained by this approach failed even to reach the level demonstrated by the area under the curve (AUC) measurements directly from the trial data, and were therefore considered unreliable.

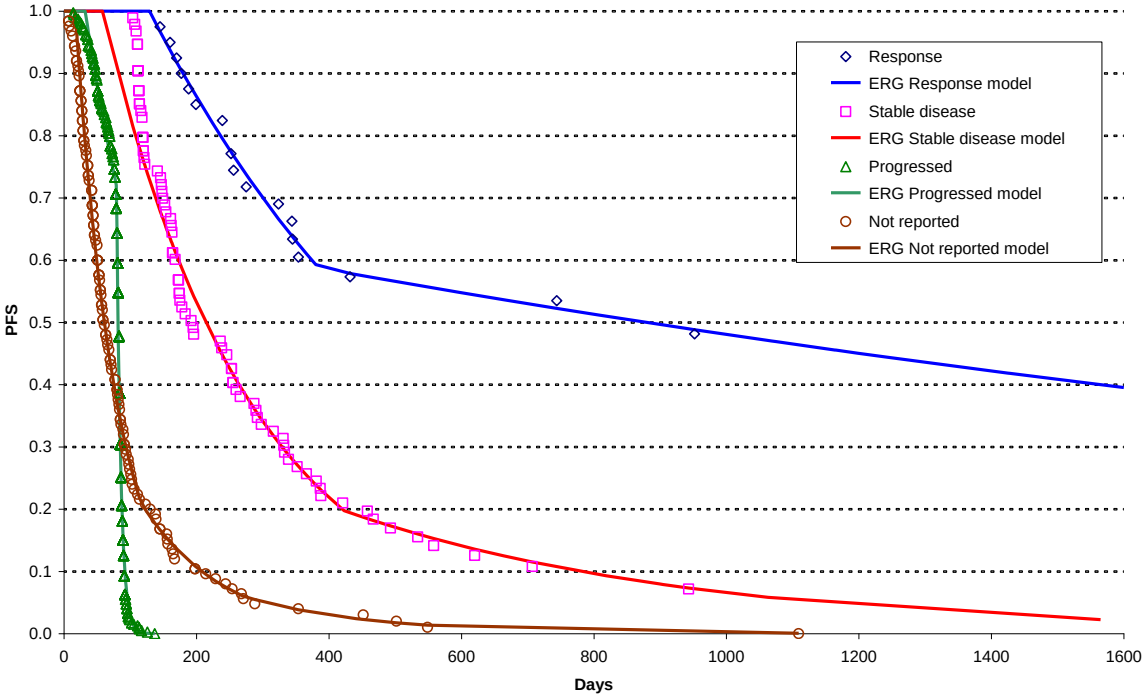


Figure 8 Progression-free survival modelled by response category

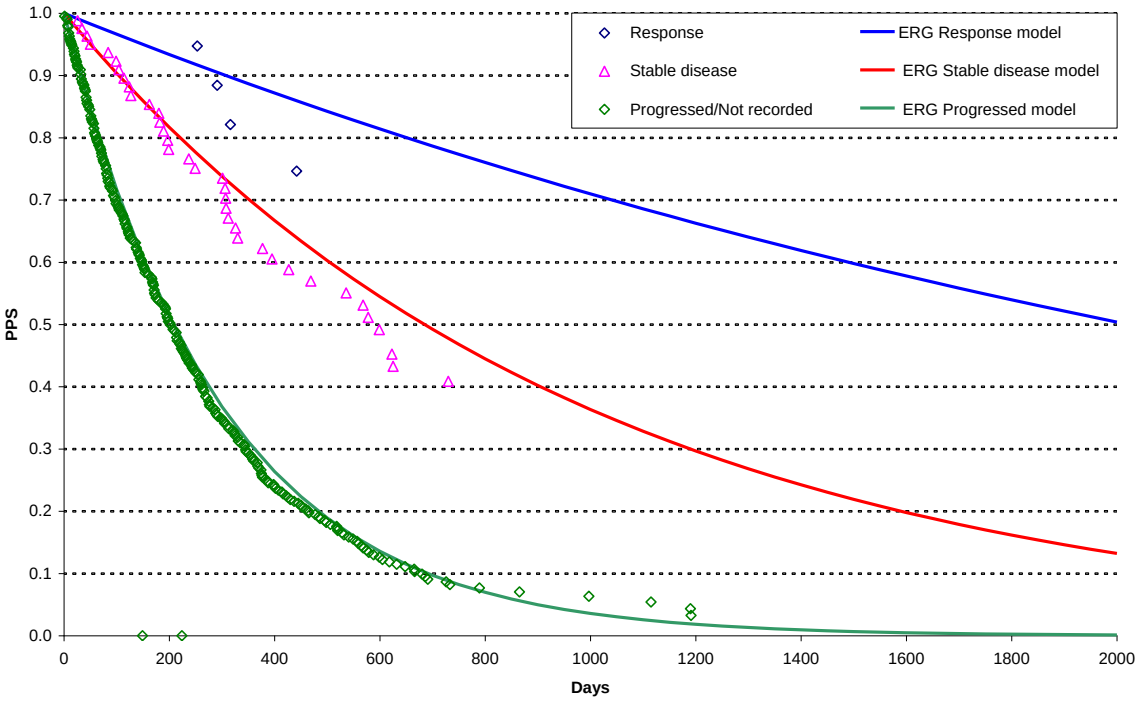


Figure 9 Post-progression survival modelled by response category

Pragmatic exploration of survival differences

In the absence of an obvious parametric method for projecting OS, the ERG undertook an exploration of survival data from the MDX010-20 trial[19] and its representation in the submitted model, with a view to deriving at least a simple method of projecting outcomes to illustrate the direction and approximate magnitude of future gains, even if particular estimates might only be considered approximate.

The approach taken was to consider whether a mixture of survival gain derived directly from the trial results together with simple long-term hazard trends might yield reasonable estimates, minimising uncertainty due to projection assumptions. This is achieved by:

- calculating the area under the Kaplan-Meier curve (AUC) to a common late time point beyond which both the intervention and the comparator arms could be seen to be following long-term trendlines;

- projecting further life expectancy based on calibrating an appropriate parametric function.

This approach was applied by the ERG to the analysis of both OS and PFS data from the MDX010-20 trial.[19] For OS, the time point at which AUC was augmented by a long-term time trend was 770 days after randomisation, and for PFS this occurred at 365 days. In all cases cumulative hazard plots showed that a simple linear trend was appropriate for long-term projection, equivalent to a simple exponential survival function (Figures 10 and 11).

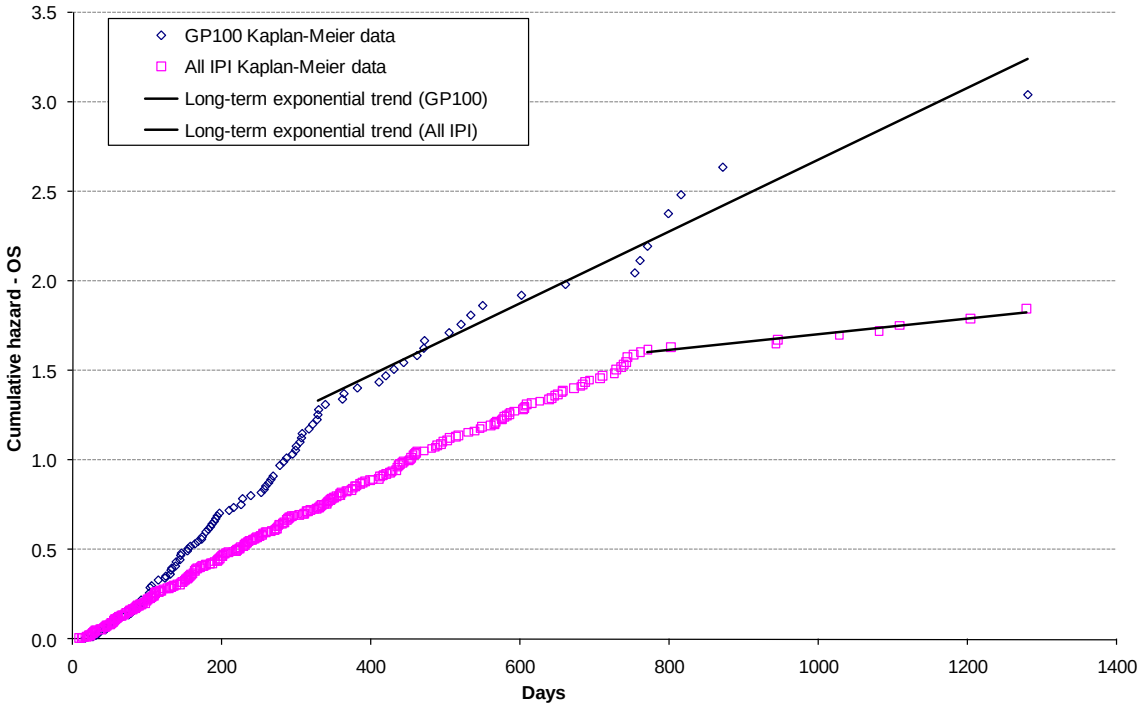


Figure 10 Long-term OS projection trends used in ERG exploratory analysis

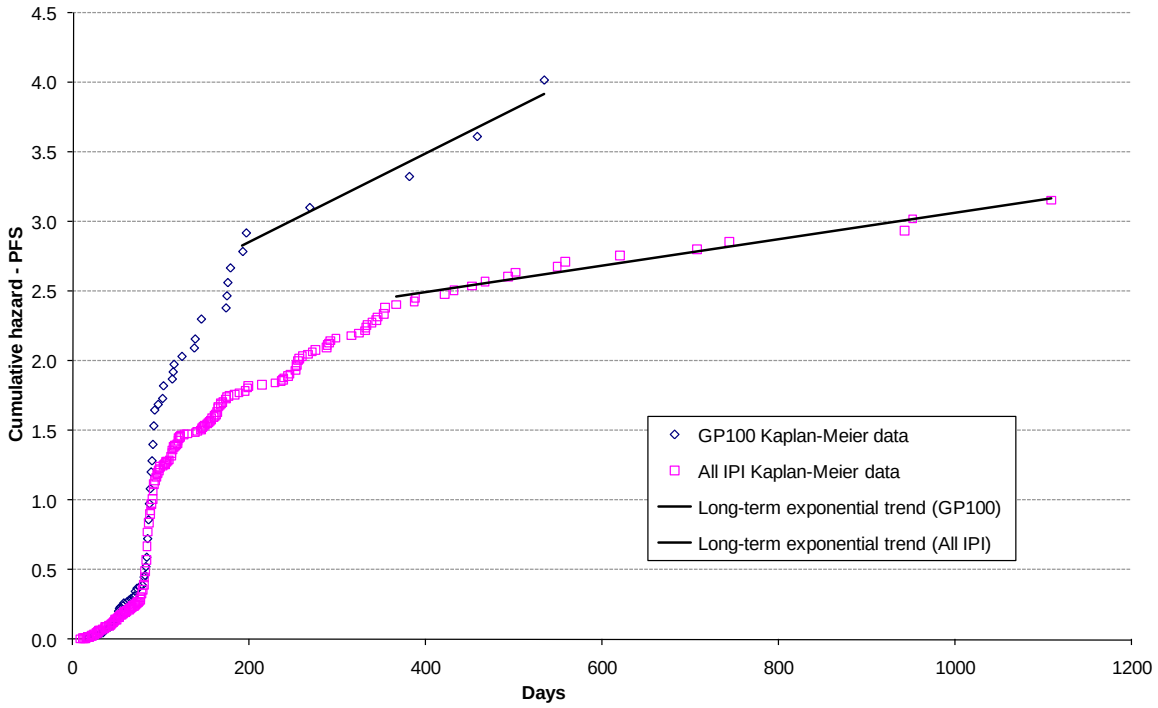


Figure 11 Long-term PFS projection trends used in ERG exploratory analysis

The combined results obtained by this method are shown in Table 19, suggesting a mean lifetime survival gain (undiscounted) of about 16 months, which is less than half the value calculated by the manufacturer's model in the base case analysis (33.77 months).

Table 19 ERG's exploratory hybrid estimation of mean OS using AUC+long-term projection

	GP100		Ipilimumab		Difference	
	Days	Months	Days	Months	Days	Months
Overall survival						
AUC from 0 to 770 days	286.7	9.4	373.3	12.3	86.7	2.8
Projection >770 days	55.3	1.8	463.8	15.2	408.5	13.4
Total estimated OS	342.0	11.2	837.2	27.5	495.2	16.3
Progression-free survival						
AUC from 0 to 365 days	95.9	3.1	122.3	4.0	26.4	0.9
Projection >365 days	14.1	0.5	96.6	3.2	82.5	2.7
Total estimated PFS	109.9	3.6	218.9	7.2	108.9	3.6

6.1.2 Summary of ERG survival estimation

The ERG analyses cannot be considered definitive, since the evidence available from the key clinical trial is inadequate to furnish a firm basis for discriminating between alternative long-term projection models. It is evident that the population of advanced and metastatic patients exhibits a high degree of heterogeneity in terms of clinical outcomes. However, to date no patient characteristics or biomarkers have been identified which can prospectively target treatments to the small minority of patients most likely to benefit from the very substantial life extension which may be possible as a result of treatment with ipilimumab. The practical problem of reliably estimating the magnitude of such benefits from the results of clinical trials with limited follow-up also remains unresolved. However, the ERG is of the opinion that using an approach based on response subgroups is a promising line of development which offers the prospect of more robust estimation limiting the need to introduce arbitrary adjustments to model logic to overcome anomalies. Unfortunately, the volume and duration of patient data available from the MDX010-20 trial[19] data proved to be inadequate to achieve survival projections which the ERG could confidently commend as a basis of decision making. Instead the ERG has provided a simple exploratory estimation scheme which is consistent with the trial data, and suggests that the manufacturer's model is likely to have substantially overestimated the extent of survival benefit that is likely to accrue from ipilimumab treatment.

The manufacturer's base case results show undiscounted mean life years (OS) of 19.5 months for the BSC comparator (based on gp100 data), and 53.3 months for ipilimumab therapy (based on combined ipilimumab trial arms), indicating an incremental gain in OS of 33.8 months. This may be compared with the ERG's exploratory estimates of 11.2 months for gp100 and 27.5 months for the combined ipilimumab arms (a gain in OS of 16.3 months). This suggests a two-fold difference in incremental survival estimates, which can be expected to generate ICERs substantially greater than those estimated for the manufacturer's base case analysis. Clearly this is a major issue in the determination of the cost effectiveness of ipilimumab.

6.2 *Cost-effectiveness results using ERG model revisions*

6.2.1 ERG revisions to base case analysis

The ERG has made minor amendments to the manufacturer's model in order to test the individual and combined effects of the errors and problems previously discussed. The derivation of an adjusted base case analysis are summarised in

Table 20 Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG

	BSC		Ipilimumab		Incremental		ICER
	Cost per patient	QALYs per patient	Cost per patient	QALYs per patient	Cost per patient	QALYs per patient	Cost per QALY gained
Manufacturer's base case analysis	**£12,837**	**1.0077**	**£96,188**	**2.3800**	**£83,351**	**1.3723**	**£60,737**
Correct background mortality logic error	£12,079	0.8790	£96,382	2.4127	£84,303	1.5337	£54,966
Correct AE costs logic error	£12,837	1.0077	£94,715	2.3800	£81,878	1.3723	£59,664
Amend age-adjustment to utility values	£12,837	1.0002	£96,188	2.3491	£83,351	1.3489	£61,791
ERG estimate of ipilimumab costs	£12,837	1.0077	£95,416	2.3800	£82,579	1.3723	£60,175
ERG revised base case analysis (all four changes)	£12,079	0.8743	£94,137	2.3811	£82,059	1.5067	£54,462
ERG revised base case analysis + exploratory survival modelling	**£11,027**	**0.7043**	**£88,618**	**1.5066**	**£77,591**	**0.8022**	**£96,717**

.

Three of the four ERG changes reduce the ICER slightly and the revised base case ICER is reduced by £2760 to £57,977 per QALY gained by use of ipilimumab.

In addition, the results of the ERG's exploratory survival estimates have been applied to the ERG revised base case results to indicate the likely impact of employing this alternative approach to projection.

The minor difference in ICERs between the manufacturer's base case analysis and the ERG's revised base case implies that the sensitivity analyses reported by the manufacturer remain an appropriate indication of the impact of model uncertainty, with the exception of the survival estimation where the ERG's exploratory survival projection constitutes the key univariate sensitivity analysis.

6.3 *Summary*

The corrections and amendments implemented by the ERG to the manufacturer's model have only a limited effect on the base case cost-effectiveness results, slightly improving the case for use of ipilimumab. However, there remains a serious issue concerning the manner of projecting survival benefit beyond the available trial evidence. The ERG was unable to arrive at a reliable estimate of the gain in mean OS using the evidence currently available. However, the ERG has prepared a simple and plausible exploratory analysis which yields OS gains approximately half the size of those claimed by the manufacturer. Using these projections substantially increases the calculated ICER to over £96,000 per QALY gained, suggesting that the true ICER is likely to be greater than that obtained with the revised base case analysis (

Table 20 Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG

	BSC		**Ipilimumab**		**Incremental**		**ICER**
	Cost per patient	QALYs per patient	Cost per patient	QALYs per patient	Cost per patient	QALYs per patient	Cost per QALY gained
Manufacturer's base case analysis	**£12,837**	**1.0077**	**£96,188**	**2.3800**	**£83,351**	**1.3723**	**£60,737**
Correct background mortality logic error	£12,079	0.8790	£96,382	2.4127	£84,303	1.5337	£54,966
Correct AE costs logic error	£12,837	1.0077	£94,715	2.3800	£81,878	1.3723	£59,664
Amend age-adjustment to utility values	£12,837	1.0002	£96,188	2.3491	£83,351	1.3489	£61,791
ERG estimate of ipilimumab costs	£12,837	1.0077	£95,416	2.3800	£82,579	1.3723	£60,175
ERG revised base case analysis (all four changes)	£12,079	0.8743	£94,137	2.3811	£82,059	1.5067	£54,462
ERG revised base case analysis + exploratory survival modelling	**£11,027**	**0.7043**	**£88,618**	**1.5066**	**£77,591**	**0.8022**	**£96,717**

).

Table 20 Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG

	BSC		Ipilimumab		Incremental		ICER
	Cost per patient	QALYs per patient	Cost per patient	QALYs per patient	Cost per patient	QALYs per patient	Cost per QALY gained
Manufacturer's base case analysis	**£12,837**	**1.0077**	**£96,188**	**2.3800**	**£83,351**	**1.3723**	**£60,737**
Correct background mortality logic error	£12,079	0.8790	£96,382	2.4127	£84,303	1.5337	£54,966
Correct AE costs logic error	£12,837	1.0077	£94,715	2.3800	£81,878	1.3723	£59,664
Amend age-adjustment to utility values	£12,837	1.0002	£96,188	2.3491	£83,351	1.3489	£61,791
ERG estimate of ipilimumab costs	£12,837	1.0077	£95,416	2.3800	£82,579	1.3723	£60,175
ERG revised base case analysis (all four changes)	£12,079	0.8743	£94,137	2.3811	£82,059	1.5067	£54,462
ERG revised base case analysis + exploratory survival modelling	**£11,027**	**0.7043**	**£88,618**	**1.5066**	**£77,591**	**0.8022**	**£96,717**

7 END OF LIFE

7.1 *Introduction*

This section provides an overview of the manufacturer's case for ipilimumab as an 'End of Life' treatment for patients with previously treated malignant melanoma. The NICE 'End of Life' treatment criteria[56] has three key points: (i) treatment is indicated for patients with a short life expectancy, normally less than 24 months and (ii) there is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with NHS treatment and (iii) the treatment is licensed or otherwise indicated for small patient populations.

7.2 *NICE End of Life treatment criteria*

7.2.1 Patient life expectancy of less than 24 months

The manufacturer makes the case that the prognosis of untreated patients with previously treated malignant melanoma is very poor. The manufacturer states (MS, pg 14) that for patients with stage III disease (i.e. regional lymph nodes involved), 5-year survival rates range from 40% to 50%, while stage IV disease (i.e. the melanoma has spread to distant sites) has an extremely poor prognosis (5-year survival rate is approximately 5-15%; median survival is 6 to 9 months). The ERG agrees with the manufacturer's estimate of life expectancy.

7.2.2 Life extension of at least 3 months

In the clinical section of the MS, the manufacturer reports data (MS, pg 112) showing that compared to BSC, ipilimumab increases both median OS (10.0 months [303 days] vs 6.4 months [196 days]) and **estimated** mean OS (23.1 months [703 days] vs 12.5 months [381 days]). The manufacturer considers that such a level of survival benefit is in excess of what is considered significant in terms of NICE 'End of Life' guidance. The ERG does not believe the manufacturer's estimate of life extension to be robust.

The ERG considers that true value of the extension in mean OS is considerably less than the manufacturer's base case (33.8 months) derived from the economic model. The ERG's exploratory analysis gives 16.3 months, which it considers to be more credible, but this cannot be considered robust since both patient numbers and follow-up time were inadequate to reliably project the experience of the small number of patients with extended survival. The evidence of the trial strongly indicates that there is a genuine life extension, and that it is highly likely that the mean life extension exceeds 3 months, but the true size of the benefit remains unclear.

7.2.3 Licensed for a small population

The manufacturer states the following (MS, pg 26);

[redacted]

The ERG is of the opinion that the number of patients eligible for treatment with ipilimumab falls within NICE's description of a small population (<7000 patients).

8 DISCUSSION

8.1 *Summary of clinical-effectiveness issues*

The clinical evidence presented in the MS is primarily derived from RCT[19] evidence which demonstrates an OS benefit for patients with unresectable melanoma who are treated with ipilimumab compared with gp100 vaccine. Unfortunately the comparator (gp100) in the key trial[19] is not one that was listed as a comparator in the final scope issued by NICE.

Over 60% of patients tolerated the full four courses of treatment. However, only a small proportion of patients responded to treatment and to date there are limited data available that would assist in predicting the patients most likely to benefit from treatment with ipilimumab. Current EMA market authorisation[17] states that all patients should receive all four courses of treatment and that treatment should be delayed or discontinued only in the case of AEs that cannot be resolved. This is in part linked to the fact that a patient's response to treatment may change during the first few months.

Standard measures of response to cancer treatments may not apply to this intervention and patient population. There is therefore a need to re-examine definitions of disease progression and, as noted above, due to the time delay between receiving treatments and definitive response, assessment should not take place until after all four treatment doses have been administered (e.g. 12 weeks).

Both the FDA[21] and the EMA[17, 45] acknowledged the increased and differing AE profile of ipilimumab in comparison to standard chemotherapy treatments. In response, a pharmacovigilance programme has been put in place by Bristol-Myers Squibb to assist clinicians to identify and treat AEs and irAEs. The MS claims that these AEs are manageable when health care professionals use, what are now becoming standard, specific treatment protocols; however, it will take time for such protocols to become routinely used in clinical practice in the UK NHS.

The most clinically effective dose of ipilimumab is still unknown and the manufacturer has been directed to conduct further research to compare the current 3mg/kg dose vs a 10mg/kg dose.

8.2 *Summary of cost-effectiveness issues*

The ERG offers a detailed critique of the manufacturer's model and suggested minor corrections/amendments related to background mortality logic, AE costs, utility age-adjustment and costs of ipilimumab. Taken together, the ERG's ICER is slightly lower than the manufacturer's base case ICER. The ERG also identified a major weakness in the economic model and considers that the manufacturer has substantially overestimated the size of the OS benefit associated with ipilimumab. The ERG agrees with the manufacturer that the size of the OS benefit is likely to be more than 3

months. However, the ERG is of the opinion that, based on the currently available data, it is impossible to estimate a reliable and robust OS gain for ipilimumab. The ERG prepared a simple and plausible exploratory analysis and estimated OS gains for ipilimumab approximately half the size of those claimed by the manufacturer (16.3 months vs 33.8 months). Using these projections, the ERG substantially increased the calculated ICER to over £96,000 per QALY gained.

8.3 *Implications for research*

Research is required to determine which patients are most likely to benefit from this treatment and also to determine the most clinically effective dose of ipilimumab.

9 REFERENCES

1. International Agency on Cancer Research. GLOBOCAN - Cancer Incidence and Mortality Worldwide in 2008. 2008; Available from: http://globocan.iarc.fr/.
2. Lens MB. Global perspectives of contemporary epidemiological trends of cutaneous malignant melanoma. Br J Dermatol. 2004; 150:179-85.
3. Garbe C, McLeod G, Buettner P. Time trends of cutaneous melanoma in Queensland, Australia and Central Europe. Cancer 2000. Cancer. 2000; 89:1269-78.
4. Agarwala SS. Current systemic therapy for metastatic melanoma. Expert Review of Anticancer Therapy. 2009; 9(5):587-95.
5. National Institute for Health and Clinical Excellence (NICE). Final scope for the appraisal of ipilimumab for previously treated unresectable malignant melanoma London 2011 [cited 2011 May]; Available from: http://guidance.nice.org.uk/TA/WaveCRS2/48/Scope/pdf/English.
6. Cancer Research UK. Skin Cancer statistics - UK. 2011 [cited 2011 July]; Available from: http://info.cancerresearchuk.org/cancerstats/types/skin/.
7. National Institute for Health and Clinical Excellence (NICE). Improving outcomes for people with skin tumours including melanoma: the manual. 2006 [cited 2011 July]; Available from: http://www.nice.org.uk/nicemedia/live/10901/28906/28906.pdf.
8. Dummer R. Melanoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. 2010 [cited 2011 July]; Available from: http://annonc.oxfordjournals.org/content/21/suppl_5/v194.full.pdf+html.
9. Lui P, Cashin R, Machado M, Hemels M, Corey-Lisle P, Einarson T. Treatments for metastatic melanoma: synthesis of evidence from randomized trials. Cancer Treat Rev. 2007; 33:665-80.
10. British Association of Dematologists. Revised UK guidelines for the management of cutaneous melanoma 2010. 2010 [cited 2011 July]; Available from: http://www.bad.org.uk/Portals/_Bad/Guidelines/Clinical%20Guidelines/Melanoma%20guidelines%202010.pdf.
11. Royal College of Physicians and British Association of Dermatologists. The prevention, diagnosis, referral and management of melanoma of the skin: concise guidelines. No 7. 2007; Available from: http://www.bad.org.uk/Portals/_Bad/Guidelines/Clinical%20Guidelines/RCP%20Melanoma%20Guidelines%202007.pdf.
12. European Dermatology Forum. Diagnosis and treatment of melanoma. European consensus-based interdisciplinary guideline. 2009 [cited 2011 July]; Available from: http://www.euroderm.org/edf/index.php.
13. Garbe C, Peris K, Hauschild A, O S, Middleton M, Spatz A, *et al.* Diagnosis and treatment of melanoma: European consensus-based interdisciplinary guideline. Eur J Cancer. 2010; 46(2):270-83.
14. American Cancer Society and National Comprehensive Cancer Network. Melanoma: Version 2 2010.
15. Middleton M, Marples M, Harries M, Wagstaff J, Dalgleish A, Osborne R, *et al.* Treatment Patterns and Outcomes in Advanced Melanoma in UK: a Retrospective Longitudinal Survey (MELODY Study). Poster LB175. National Cancer Research Institute Cancer Conference; Liverpool, UK 2010.
16. Oxford Outcomes. MELODY study resource use data analysis. Vancouver BC, Canada: Bristol-Myers Squibb 2010.
17. European Medicines Agency. Assessment report for Yervoy (ipilimumab). London 2011.

18. European Commission. Commission implementing decision for Yervoy - ipilimumab. Brussels 2011.
19. Hodi F, O'Day S, McDermott D, Weber R, Sosman J, Haanen J, *et al.* Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010; 363(8):711-23.
20. European Medicines Agency. Summary of Product Charateristics -Yervoy 2011; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/002213/WC500109299.pdf.
21. Food and Drug Administration (FDA). Yerboy (ipilimumab): Risk evaluation and mitigation strategy (REMS) - severe immune-mediated adverse reactions. 2011; Available from: http://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm249770.htm.
22. Bristol-Myers Squibb. Ipilimumab - clarification letter response 26 July 2011.
23. Parkhurst M, Salgaller m, Southwood S, Robbins P, Sette A, Rosenberg S, *et al.* Improved induction of melanoma-reactive CTL withpeptides from the melanoma antigen gp100 modified at HLA-A*0201. J Immunol. 1996; 157(6):2539-48.
24. Eggermont AMM. Therapeutic vaccines in solid tumours: Can they be harmful? Eur J Cancer. 2009; 45(12):2087-90.
25. Saenger Y, Wolchok J. The heterogeneity of the kinetics of response to ipilimumab in metastatic melanoma: patient cases. Cancer Immunity. 2008; 8:1-7.
26. Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, Lebbe C, *et al.* Guidelines for the evaluation of immune therapy activity in solid tumors: Immune-related response criteria. Clin Cancer Res. 2009; 15(23):7412-20.
27. World Health Organization. Oncology Modified WHO Response Criteria. [cited 2011 July]; Available from: https://www.ctnbestpractices.org/resources/study-patient-management/subjectassessments/Oncology%20response%20defintion%20by%20Modified%20WHO%20criteria.doc/view.
28. Eisenhauer E, Therasse P, Bogarts J, Schwartz L, Sargent D, Ford R, *et al.* New gidelines to evaluate the response to treatment in solid tumors. Eur J Cancer. 2009; 45:228-47.
29. O'Day SJ, Maio M, Chiarion-Sileni V, Gajewski TF, Pehamberger H, Bondarenko IN, *et al.* Efficacy and safety of ipilimumab monotherapy in patients with pretreated advanced melanoma: A multicenter single-arm phase II study. Ann Oncol. 2010; 21(8):1712-7.
30. Weber J, Thompson J, Hamid O, Minor D, Amin A, Ron I, *et al.* A randomized, double-blind, placebo-controlled, phase II study comparing the tolerability and efficacy of ipilimumab administered with or without prophylactic budesonide in patients with unresectable stage III or IV melanoma. Clin Cancer Res. 2009; 15(17):5591-8.
31. Wolchok J, Neyns B, Linette G, Negrier S, Lutzky J, Thoma sL, *et al.* Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study. Lancet Oncology. 2010; 11(2):155-64.
32. Ribas A, Chmielowski B, Glaspy J. Do we need a different set of response assessment criteria for tumor immunotherapy. Clin Cancer Res. 2009; 15(23):7116-8.
33. Robert C, Thomas L, Bondarenko I, O'Day S, Weber J, Garbe C, *et al.* Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. N Engl J Med. 2011; 364(26):2517-26.
34. Hauschild A, Agarwala S, Trefzer J, Hogg D, Rober C, Hersey P, *et al.* Results of a Phase III, randomized, placebo-controlled study of sorafenib in combination with carboplatin and paclitaxel as second-line treatment in patients with unresectable stage iii or stage iv melanoma. J Clin Oncol. 2009; 27(17):2823-30.
35. Zimpfer-Rechner C, Hofmann U, Figl R, Becker J, Trefzer U, Keller I, *et al.* Randomized phase II study of weekly paclitaxel versus paclitaxel and carboplatin as second-line therapy in disseminated melanoma: a multicentre trial of the Dermatologic Co-operative Oncology Group (DeCOG). Melanoma Res. 2003; 13(5):531-6.
36. Eisen T, Trefzer J, Hamilton A, Hersey P, Millward M, Knight R, *et al.* Results of a multicenter, randomized, double-blind phase 2/3 study of lenalidomide in the treatment of pretreated relapsed or refractory metastatic malignant melanoma. Cancer. 2010; 116(1):146-54.

37. Maio M, Lebbe C, Neyns B, O'Day S, Wolchok J, Weber J, *et al.* Three-year survival rates for patients with advanced melanoma who received ipilimumab at 10 mg/kg in Phase II Trials. Perspectives in Melanoma XIV; September 17-18; Amsterdam, The Netherlands2010.
38. Bristol-Myers Squibb. Final Clinical Study Report for Study CA184004. 2010; Available from: http://ctr.bms.com/pdf//CA184-004ST.pdf.
39. Culver ME, Gatesman ML, Mancl EE, Lowe DK. Ipilimumab: a novel treatment for metastatic melanoma. Ann Pharmacother. 2011; 45(4):510-9.
40. Atria P, Phan GQ, Maker AV, Robinson MR, Quezado MM, Yang JC, *et al.* Autoimmunity correlates with tumor regression in patients with metastatic melanoma treated with anti-cytotoxic T-lymphocyte antigen-4. J Clin Oncol. 2005; 23(25):6043-53.
41. Slingluff C, Yamshchikov G, Neese P, Galavottie H, Eastham S, Engelhard V, *et al.* Phase I Trial of a melanoma vaccine with gp100280–288 peptide and tetanus helper peptide in adjuvant:immunologic and clinical outcomes. Clin Cancer Res. 2001; 7:3012-24.
42. Gonzalez-Galarza F, Christmas S, Middleton D, Jones A. Allele frequency net: a database and online repository for immune gene frequences in worldwide populations. Nucleic Acids Res. 2011; 39:D913-D9.
43. Allison JP. Checkpoint blockade in tumor immunotherapy: New in-sights and opportunities. J Immunother. 2009; 32(9):986.
44. Bouwhuis MG, Ten Hagen TLM, Suciu S, Eggermont AMM. Autoimmunity and treatment outcome in melanoma. Curr Opin Oncol. 2011; 23(2):170-6.
45. European Medicines Agency. Summary of opinion Yervoy (Ipilibumab). London 2011; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/002213/WC500106522.pdf.
46. Chin K, Ibrahim R, Berman D, Yellin M, Lowy I, Lin R, *et al.* Treatment guidelines for the management of immune-related adverse events in patients treated with ipilimumab, an anti-CTLA4 therapy. Ann Oncol. 2008; 19 (S8):viii244-viii5.
47. Dixon S, Walters S, Turner L, Hancock B. Quality of life and cost-effectiveness of interferon-alpha in malignant melanoma: results from randomised trial. Br J Cancer. 2006; 94(4):492-8.
48. Rowan D, Brazier J, Young T, Gaugris S, Craig B, King M, *et al.* Deriving a preference-based measure for cancer using the EORTC QLQ-C30. Value in Health. 2011 (in press).
49. Beusterien KM, Szabo SM, Kotapati S, Mukherjee J, Hoos A, Hersey P, *et al.* Societal preference values for advanced melanoma health states in the United Kingdom and Australia. Br J Cancer. 2009; 101(3):387-9.
50. Oxford Outcomes. Advanced melanoma resource use and costs in Europe. Vancouver, BC Canada: Bristol-Myers Squibb 2011.
51. Addicott R, Dewar S. Improving choice at end of life; a descriptive analysis of the impact and costs of the Marie Curie Delivery Choice Programe in Lincolnshire. London: Kings Fund 2008.
52. Atkins M, Lotze M, Dutcher J, Fisher R, Weiss G, Margolin K, *et al.* High-dose recombinant Interleuken 2 therapy for patients with metastatic melanoma: analysi of 270 patients treated between 1985 and 1993. J Clin Oncol. 1999; 17(7):2105-16.
53. Office for National Statistics. Mortality statistics: All cause England and Wales 2008 London: TSO; 2010 [cited 2011 July]; Available from: http://www.statistics.gov.uk/cci/nscl.asp?ID=5014.
54. Department of Health. National Schedule of Reference Costs 2007-08 published 8 May 2009. 2009 [cited 2010 January]; Available from: <http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_098945>.
55. Kind P, Hardman G, Macran S. UK Population Norms for EQ-5D. Discussion paper 172. York, UK: University of York Centre for Health Economics 1999.
56. National Institute for Health and Clinical Excellence. Supplementary Advice to the Appraisal Committees - End of Life Treatments. 2009; Available from: http://www.nice.org.uk/aboutnice/howwework/devnicetech/endoflifetreatments.jsp?domedia=1&mid=88ACDAE5-19B9-E0B5-D422589714A8EC6D.

10 APPENDICES

Appendix 1 Quality assessment of submitted trials

Trial MDX010-20[19]

Study question	How is the question addressed in the study?	Grade (yes/no/not clear/NA)	ERG comment
Was randomisation carried out appropriately?	Adequate: Patients were randomly assigned to one of three study groups using centralised scheme with stratification according to baseline metastases stage (M0, M1a or M1b) [Publication and study protocol]	Yes	
Was the concealment of treatment allocation adequate?	Adequate: Placebo for both ipilimumab and vaccine were used; The pharmacist at each study site was unblinded to study medication; other study site personnel and patients were blinded to patient assignment.	Yes	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Adequate: Metastases stages among the three arms were comparable; Previous systemic therapy (including IL-2 therapy) was similar	Yes	Agree
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)?	Adequate: All site personnel including clinicians, data management, statisticians and patients were blinded. (except pharmacists were unblinded to the study medication)	Yes	Medarex were aware of the patient assignment
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Adequate: The drop rates among the groups were similar; Patients who didn't start treatment after randomisation were 22/403, 6/137 and 5/136 respectively; Patients who discontinued were 135/403, 43/137 and 54/136 (Supplementary Appendix)	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Adequate: Results for all the mentioned outcomes were presented in the publication	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Adequate: Except for safety analysis, all efficacy outcomes were analysed using randomised (ITT) population	Yes	Agree

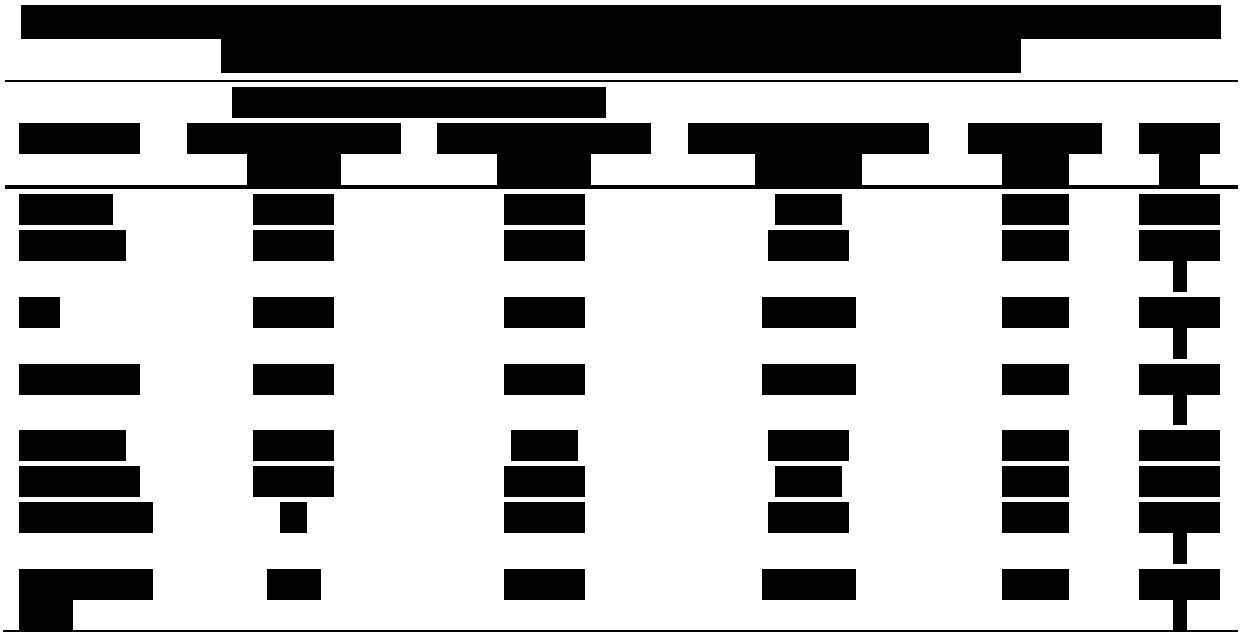
Trial CA184-22[31]

Study question	How is the question addressed in the study?	Grade (yes/no/not clear/N/A)	ERG comment
Was randomisation carried out appropriately?	Adequate: Patients were randomised in a 1:1:1 ratio by using interactive voice response system (IVRS) and were assigned unique identification number	Yes	Agree
Was the concealment of treatment allocation adequate?	Adequate: The IVRS assigned a unique identification number and the system assigned the patient to one of the three treatment groups on the basis of a randomisation schedule generated.	Yes	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Adequate: Demographics and baseline characteristics were consistent across treatment groups; Patients were similar across the groups with respect to M-stage and ECOG status.	Yes	Agree?
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)?	Adequate: Patients, treating doctors, and doctors' staff were blinded to the treatment allocation; whereas pharmacists, data monitoring committee were aware of the allocations. Unblinded personnel involved in the study have minimal or no effect on study bias	Yes	Pharmacists and manufacturer not blinded
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Adequate: Across the three treatment groups - 0.3mg/kg, 3.0mg/kg, and 10mg/kg - treatment discontinuations (received fewer than 4 doses) were similar; 23/73, 21/72 and 35/72, respectively. One patient from each of the three treatment groups, having been randomised, did not receive any ipilimumab dose due to not meeting a study criterion (n=2) and progressive disease (n=1)	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Adequate: Outcomes specified were reported in results section	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Adequate: Efficacy analyses were done by ITT, whereas safety analyses included only patients who received at least one dose of ipilimumab	Yes, for efficacy analyses; No for safety analyses	Agree

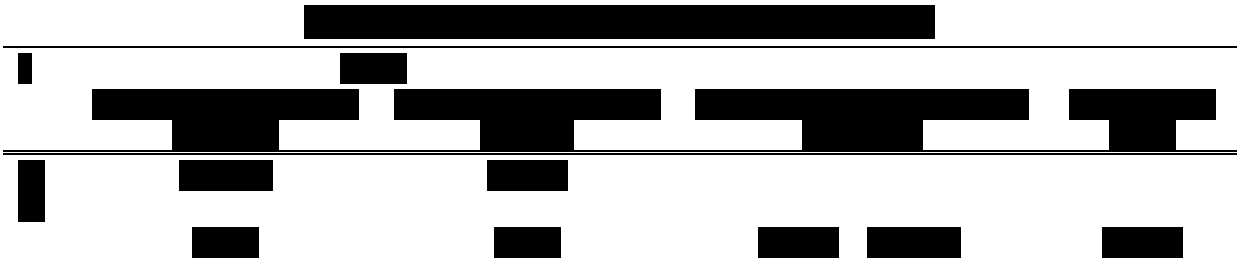
Trial CA184-007[30]

Study question	How is the question addressed in the study?	Grade (yes/no/not clear/N/A)	ERG comment
Was randomisation carried out appropriately?	Adequate: Patients were randomised 1:1 to treatment arms; Randomisation was determined using a telephone interactive voice response system (IVRS) using a permuted block procedure.	Yes	Agree
Was the concealment of treatment allocation adequate?	Adequate: The IVRS randomisation suggests the treatment allocation was concealed adequately.	Yes	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Adequate: More patients received previous systemic therapy in the budesonide arm compared to the placebo arm. However, other baseline demographic and characteristics were generally similar.	No	Agree
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)?	Adequate: Blinded oral medication was self-administered by patients; Radiologic tumour assessments in patients were done by an independent review committee.	Yes	Agree
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Adequate: Treatment discontinuations in each arm were similar, i.e., discontinuation due to disease progression was 25/58 in budesonide arm and 30/57 in placebo arm; discontinuation due to AEs was 17/58 and 17/57 respectively.	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Adequate: Outcomes specified were reported in results section.	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Adequate: All randomly allocated patients (n=115) were included in the efficacy as well as tolerability analyses.	Yes	Agree

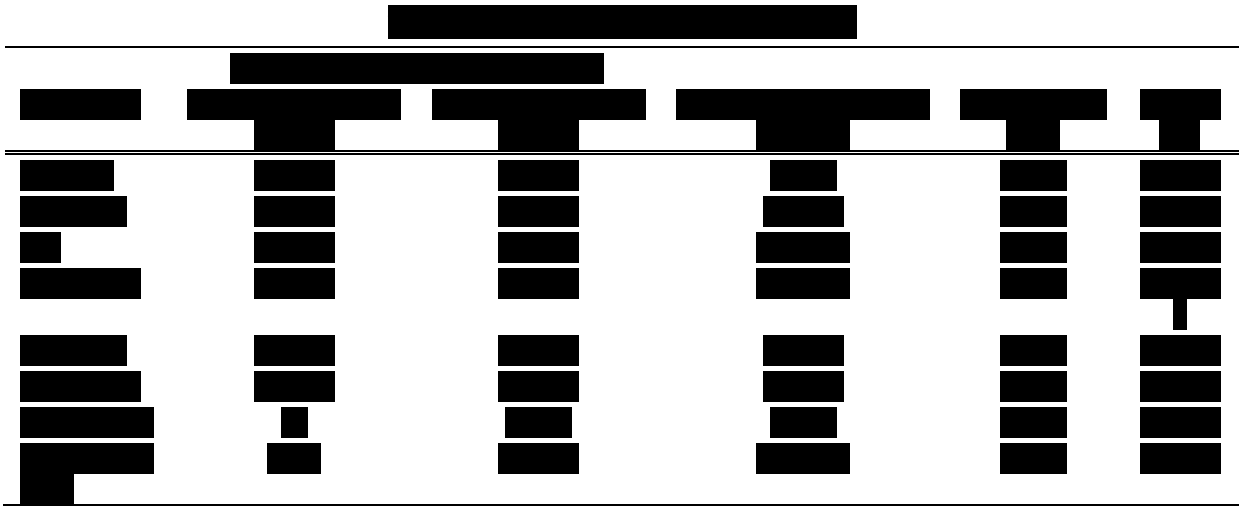
Appendix 2 Overall survival – sub-group analysis from clarification response



* Test based upon Exact confidence intervals (inverting a two-sided test). Agresti A & Min Y (2001). On small-sample confidence intervals for parameters in discrete distributions. Biometrics 57: 963-971.



** T-test assuming unequal variances.



* Test based upon Exact confidence intervals (inverting a two-sided test). Agresti A & Min Y (2001). On small-sample confidence intervals for parameters in discrete distributions. Biometrics 57: 963-971.

[redacted]

** T-test assuming unequal variances.