Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE

Imatinib for the adjuvant treatment of gastrointestinal stromal tumours

Produced by Southampton Health Technology Assessments Centre (SHTAC)

Authors Jeremy Jones
Petra Harris
Jonathan Shepherd
Keith Cooper

Correspondence to Jonathan Shepherd
Southampton Health Technology Assessments Centre
University of Southampton
[redacted], [redacted]
[redacted], [redacted]
[redacted]
[redacted]

Date completed 7th April 2014

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 13/47/01.

Declared competing interests of the authors

None

Acknowledgements

We are very grateful to Professor Peter Simmonds, Consultant Medical Oncologist at [redacted], for providing expert clinical advice during the production of this report.

We also thank Dr Jill Colquitt, Senior Research Fellow, SHTAC, for acting as internal editor for the ERG report, and Karen Welch, Information Specialist, SHTAC, for appraising the literature search strategies, and Jackie Bryant, Principal Research Fellow, SHTAC, for assistance with appraising the manufacturer's submission.

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:

Jones J, Harris P, Shepherd J, Cooper K. Imatinib for the adjuvant treatment of gastrointestinal stromal tumours. A Single Technology Appraisal. Southampton Health Technology Assessments Centre (SHTAC), 2014.

Contributions of authors

J Jones (Principal Research Fellow) critically appraised the health economic systematic review and the economic evaluation and drafted the report; P Harris (Research Fellow) critically appraised the clinical effectiveness systematic review and drafted the report; J Shepherd (Principal Research Fellow) critically appraised the clinical effectiveness systematic review, drafted the report and project managed the review; K Cooper (Senior Research Fellow) critically appraised the health economic systematic review and the economic evaluation and drafted the report.

Word count:

34,209

TABLE OF CONTENTS

1 Introduction to ERG Report 15
2 BACKGROUND 15
2.1 Critique of manufacturer's description of underlying health problem 15
2.2 Critique of manufacturer's overview of current service provision 15
2.3 Critique of manufacturer's definition of decision problem 18
3 CLINICAL EFFECTIVENESS 20
3.1 Critique of manufacturer's approach to systematic review 20
3.2 Summary statement of manufacturer's approach 38
3.3 Summary of submitted evidence 39
3.4 Summary 50
4 ECONOMIC EVALUATION 51
4.1 Overview of manufacturer's economic evaluation 51
4.2 Critical appraisal of the manufacturer's submitted economic evaluation 53
4.3 Additional work undertaken by the ERG 95
4.4 Summary of uncertainties and issues 100
5 End of life 101
6 Innovation 101
7 DISCUSSION 101
7.1 Summary of clinical effectiveness issues 101
7.2 Summary of cost effectiveness issues 102
8 REFERENCES 103

LIST OF TABLES

Table 1: Miettinen risk stratification of primary GIST by mitotic count, size and site (Table A-2 taken from page 22 of MS) 17
Table 2: Additional RCT identified by ERG searches 21
Table 3: Selected main differences in baseline demographic and pathological characteristics of patients as reported in the primary analysis (full population) and for primary analysis patients retrospectively identified as the Miettinen high risk sub-population 26
Table 4: Manufacturer and ERG assessment of trial quality 30
Table 5: Quality assessment (CRD criteria) of MS review 39
Table 6: RFS probabilities based on 5-year follow-up analysis of the ACOSOG Z9001 trial 40
Table 7: RFS probabilities based on 5-year follow-up analysis of the SSGXVIII/AIO trial 42
Table 8: Interim RFS results for the EORTC 62024 trial 43
Table 9: Interim 5-year IFS results for the EORTC 62024 trial 43
Table 10: Summary of OS in the ACOSOG Z9001 trial 44
Table 11: Summary of OS in the SSGXVIII/AIO trial 44
Table 12: Results of comparative non-RCTs in high risk patients 46
Table 13: Results of non-comparative non-RCTs in high risk patients 47
Table 14: Statistically significant differences in most frequently recorded AEs between 1 and 3-years of adjuvant imatinib therapy 48

Table 15: Base case cost effectiveness results 53
Table 16: Critical appraisal checklist of economic evaluation 54
Table 17: NICE reference case requirements 55
Table 18: Analysis points for ACOSOG Z9001 trial and populations included at each analysis 67
Table 19: Hazard ratios for recurrence-free survival with adjuvant imatinib in base case analysis & five year update (response to clarification request C1) 68
Table 20: discontinuation of non-adjuvant imatinib, due to disease progression and adverse events in the manufacturer's model 70
Table 21: Discontinuation of sunitinib, due to disease progression and adverse events in the manufacturer's model 71
Table 22: Overall survival. Data sources and parameter estimates for manufacturer's model 73
Table 23: Health state utility values used in the economic model 79
Table 24: Drug acquisition costs 85
Table 25: Unit costs associated with monitoring and follow up visits in the MS 86
Table 26: Unit costs for adverse events in the MS 87
Table 27: Comparison of MS model results with ACOSOG Z9001 and SSGXVII/AIO clinical trials 90
Table 28: No adjuvant treatment vs 1 year (reproduced from MS B-51 – B-53) 91
Table 29: Base case cost effectiveness results 95
Table 30: Scenario A: Off treatment HR reduced to 75% after 5 years 96
Table 31: Scenario B: Off treatment HR halved after 5 years: 96
Table 32: Scenario C: Off treatment HR reduced to 25% after 5 years: 96
Table 33: Scenario analysis increasing post recurrence death rate in treatment arms . 97
Table 34: Probability of recurrence at selected time points for MS base case and ERG analysis with no extended off treatment benefit 98
Table 35: ERG additional analyses 99

LIST OF FIGURES

Figure 1: RFS rates for Miettinen high risk sub-population receiving imatinib or placebo in the ACOSOG Z9001 study based on the 5-year follow-up analysis 41
Figure 2: Kaplan–Meier estimate of recurrence-free survival in the Miettinen high risk sub-population in the SSGXVIII/AIO trial 42
Figure 3: Kaplan–Meier estimate of OS in the Miettinen high risk population in the SSGXVIII/AIO trial 45
Figure 4: MS Model structure (MS Figure B-10 page 129) 56
Figure 5: Survival functions applied in the economic model 65
Figure 6: Observed recurrence-free survival (3-years adjuvant imatinib) versus alternative parametric survival extrapolations 66
Figure 7: Resource use assumptions included in the model 83
Figure 8: Cost-effectiveness acceptability curves (MS Figure B-24) 94

LIST OF ABBREVIATIONS

AE	Adverse event
ACOSOG	American College of Surgeons Oncology Group
ASCO	American Society of Clinical Oncology
AT	Adjuvant treatment
BSC	Best supportive care
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CRD	Centre for Reviews and Dissemination
CSR	Clinical study report
CT	Computerised tomography
EORTC	European Organisation for Research and Treatment of Cancer
ERG	Evidence review group
GIST	Gastrointestinal stromal tumour
GP	General practitioner
HR	Hazard ratio
HRG	Health resource group
HRQoL	Health related quality of life
ICER	Incremental cost-effectiveness ratio
IFS	Imatinib-failure-free survival
IPCW	Inverse Probability of Censoring Weights
IPE	Iterative Parameter Estimation Algorithm
ITT	Intention-to-treat analysis
KM	Kaplan-Meier
LY	Life year
mITT	Modified intention-to-treat analysis
MS	Manufacturer's submission
NIH	National Institutes of Health
NR	Not reported
OP	Out-patient
OS	Overall survival
PAS	Patient access scheme
PSA	Probabilistic sensitivity analysis
PSS	Personal social services
QALY	Quality adjusted life year
QoL	Quality of life
RPSFTM	Rank preserving structural failure time model
RFS	Recurrence free survival
RCT	Randomised controlled trial
SA	Sensitivity analysis
SmPC	Summary of product characteristics
SSGXVIII/AIO	Scandinavian Sarcoma Group and the Sarcoma Group of the Arbeitsgemeinschaft Internistische Onkologie
STA	Single technology assessment
TKI	Tyrosine kinase inhibitor
WTP	Willingness To Pay

SUMMARY

Scope of the manufacturer submission

This appraisal topic is a review of TA196 'imatinib for the adjuvant treatment of gastrointestinal stromal tumours'. The scope of the submission was in line with the NICE scope, that is to assess the clinical and cost-effectiveness of imatinib as an adjuvant treatment for adults who are at significant risk of relapse following resection of c-KIT (CD117)-positive gastrointestinal stromal tumour (GIST) and in line with the significant risk population in the marketing authorisation. The comparator is observation after surgery (no adjuvant therapy). The TA196 guidance was based on mainly one trial comparing 1-year of adjuvant imatinib with placebo (the ACOSOG Z9001 trial). At that time the NICE Appraisal Committee were aware of on-going clinical trials and this review of TA196 includes the longer-term evidence in the appraisal of the clinical and cost-effectiveness of imatinib.

Summary of submitted clinical effectiveness evidence

The manufacturer's submission (MS) for the clinical effectiveness evidence to NICE included:

i) a systematic literature review to identify all studies reporting on the clinical effectiveness and safety of imatinib in the adjuvant setting of GIST.
ii) three randomised controlled trials (RCTs) (the ACOSOG Z9001 trial, the SSGXVIII/AIO trial, and the EORTC 62064 trial). One RCTs is only available as an interim analysis reported in a conference abstract (EORTC 62064).
iii) twelve non-RCTs of varying relevance to the decision problem, with some failing to report the risk category of the included patients and some failing to report a classification system for the reported risk categories.

Meta-analysis was not performed and would not have been feasible due to methodological differences between the included RCTs (and between the RCTs and non-randomised studies). An indirect comparison of two of the RCTs was conducted to inform the economic analysis.

The population and the comparisons of the three RCTs varied. The ACOSOG Z9001 trial compared 1-year of adjuvant imatinib therapy following surgical resection against placebo, based on patients at any level of risk of recurrence (the trial was conducted prior to the introduction of risk categorisation). The SSGXVIII/AIO trial compared 1-year of adjuvant imatinib therapy following surgical resection with 3-years therapy based on patients with a high

Superseded – see erratum

risk of recurrence of GIST (based on modified US National Institutes of Health NIH Consensus Criteria). The EORTC 62064 trial compared 2-years of imatinib therapy following surgical resection with observation only (no treatment) based on patients with intermediate or high risk GIST (based on NIH Consensus Criteria). The SSGXVIII/AIO trial and EORTC 62064 trials were open-label.

The submission provides treatment effect estimates for the full trial populations (intention to treat, ITT) and retrospectively analysed high risk sub-populations. The Miettinen risk classification scheme was used in preference to other classification schemes as this is recommended by UK clinical guidelines.

All three RCTs reported longer recurrence free survival (RFS) associated with adjuvant imatinib treatment. In the ACOSOG Z9001 trial 1-year imatinib compared to no adjuvant treatment was associated with longer RFS at 5-year follow-up (full population HR 0.718 (95% CI 0.531 to 0.971); p =0.0305; Miettinen high risk sub-population: HR 0.608 (0.417 to 0.886; p = 0.009), while in the SSGXVIII/AIO trial 3-year imatinib treatment was significantly associated with longer RFS compared to 1-year treatment at 5-year follow-up (full population: HR 0.46; 95% CI 0.32 to 0.65; p < 0.0001; Miettinen high risk sub-population: HR 0.43; 95% CI 0.30 to 0.62; p < 0.001). The EORTC 62064 trial showed a difference between imatinib (84%) and no adjuvant treatment at 3-years (66%) in RFS, but similar results at 5-years (69% vs 63%, respectively) (based on interim data and caution is advised in the interpretation of these results).

The results for overall survival (OS) across the two trials which reported this outcome were mixed. The ACOSOG Z9001 trial had few deaths overall and there was no statistically significant difference between 1-year treatment and no adjuvant treatment (full population only: 2-years HR 0.66; 95% CI 0.22 to 2.03, p = 0.47; 5-years HR 0.816; 95% CI 0.488 to 1.365; p = 0.4385). The SSGXVIII/AIO trial (relatively smaller, but evaluating a longer treatment period), reported comparatively more deaths and at 5-years follow-up there was a statistically significantly longer OS associated for 3-years imatinib treatment compared to 1-year treatment (full population: HR 0.45; 95% CI 0.22 to 0.89; p = 0.019; Miettinen high risk sub-population: HR 0.39; 95% CI 0.19 to 0.79; p = 0.007). Differences between the two trials in terms of patient characteristics or other variables may explain the differences in the overall death rates seen. However, neither of the trials was statistically powered for OS and caution is necessary in the interpretation of the results. The EORTC 62064 trial reported imatinib-failure-free survival (IFS;

Superseded – see erratum

the trial's primary end-point) and 5-year IFS was similar between the 2-year imatinib group and the no adjuvant therapy group (full population HR 0.80; 98.5% CI, 0.51 to 1.26; $p = 0.23$); high risk GIST population: $p = 0.11$). However, once again this is based on interim data and caution is advised in the interpretation of these results.

Evidence from the comparative non-RCTs (as well as the non-comparative) reporting a high risk patient group is supportive of some of the findings of the three included RCTs, in that imatinib if taken for 3 or more years is associated with better OS and RFS than taken for shorter periods.

Adverse events (AE) were reported by two of the RCTs for the full trial populations (rather than the high risk sub-populations). There was a greater incidence of combined grade 3/4 AEs in the imatinib group in the ACOSOG Z9001 trial (30.0% vs 18.3% placebo), with the most common grade 3 or 4 events being neutropenia, abdominal pain, dermatitis, nausea and elevated alanine aminotransferase levels. At the 5-year analysis, there were a higher number of withdrawals due to AEs in the imatinib group (1.7%) compared to placebo (0.3%) and a slightly higher percentage of deaths in the placebo group than the imatinib group (9.3% vs 7.2%). In the SSGXVIII/AIO trial, the incidence of any AE was similar for 1-year or 3-year adjuvant imatinib treatment (99% vs 100%). Incidences of any grade 3 or 4 event were statistically significantly higher in the adjuvant imatinib 3-year group compared to the 1-year group (32.8% vs 20.1% respectively; $p = 0.006$), with the most common reported grade 3 or 4 events being leukopenia and diarrhoea. Discontinuations were double that for the 3-year imatinib group compared to the 1-year group (25.8% vs 12.9%), reflected in higher discontinuations due to AEs in patients treated for 3-years (13.6% vs 7.7%, respectively).

Health-related quality of life was not reported in any of the three RCTs.

The MS does not report sub-group analyses (e.g. by tumour genetic mutation site) though these are available in the journal publication for the SSGXVIII/AIO trial.

Summary of submitted cost effectiveness evidence

The manufacturer's submission to NICE includes:

i) a review of published economic evaluations of adjuvant imatinib with surgical resection compared with surgical resection alone for adult patients with GIST.

ii) a report of an economic evaluation undertaken for the NICE STA process. The cost effectiveness of adjuvant imatinib for 1 and 3-years is compared with no treatment for adult patients with surgical resection for GIST.

A systematic search of the literature was conducted by the manufacturer to identify economic evaluations of treatments for GIST. The review identified two studies evaluating the cost effectiveness of adjuvant imatinib for the treatment of GIST.

The manufacturer's own cost effectiveness analysis uses a multi-state Markov model to estimate the cost-effectiveness of adjuvant imatinib for the treatment of GIST compared with no treatment. The model adopted a lifetime horizon, with a monthly cycle length. Discount rates of 3.5% were applied to both benefits and costs. The model consists of nine health states. Patients can remain recurrence-free, have a recurrent GIST (first or second recurrence), and have progressive disease or die (from GIST or other causes). The model was based upon the one submitted for the previous NICE appraisal for adjuvant imatinib (TA196). Clinical data for early transitions (primary recurrence after surgery with and without adjuvant imatinib, discontinuation of adjuvant imatinib) in the model are based upon the results from the adjuvant imatinib trials (ACOSOG Z9001 and SSGXVIII/AIO). Later transitions in the model, for treatment outside the adjuvant setting are based on data from other clinical trials in patients with advanced GIST.

Results from the economic model are presented as incremental cost per quality adjusted life years (QALY) gained for 3-year adjuvant imatinib treatment compared to 1-year adjuvant treatment and no treatment. For the base case analysis, the incremental cost per QALYs gained were £3509 for 1-year adjuvant imatinib versus no adjuvant treatment, £8390 for 3-years adjuvant treatment versus no adjuvant treatment, and £16,006 for 3-years adjuvant imatinib versus 1-year adjuvant imatinib.

The model explores structural and parameter uncertainty in one-way deterministic sensitivity analyses and probabilistic sensitivity analyses (PSA). The deterministic sensitivity analysis indicated that the model was most sensitive to the time horizon and the treatment HR for the on and off treatment phase. Scenario analyses were also conducted for alternative parametric distributions, dose escalation to 800 mg imatinib following recurrence, change to the proportion of patients moving to BSC (progressive disease) following recurrence and extended survival after recurrence. The results of the PSA found that the likelihood of 1-year and 3-years of

imatinib treatment being cost effective at a threshold of £20,000 per QALY is 41.7% and 58.3% respectively, and at a threshold of £30,000 per QALY is 30.0% and 69.1% respectively.

Commentary on the robustness of submitted evidence

Strengths

The assessment of clinical effectiveness is based on a systematic review generally conducted to a reasonable standard and reported in adequate detail. However, the search strategy for clinical effectiveness was not fully up to date, necessitating the ERG to update the search. One additional potentially relevant phase II RCT was identified by the ERG, though it evaluated treatment duration of less than a year and therefore is relatively less informative than the other RCTs included which evaluate treatment effects up to 3-years. The ERG is not aware of any other relevant studies that have not been included in the MS.

The three RCTs identified in the clinical effectiveness evidence were generally well designed and executed, although two of the RCTs changed their primary outcome measure after randomisation (with agreement), and two were open-label. Five year follow-up data are available for adjuvant treatment lasting up to 3-years, thus providing evidence for effectiveness of longer-term treatment than one year treatment data considered in the previous appraisal (NICE TA196).

The approach taken in the submission to model GIST is reasonable and consistent with the clinical pathway for GIST.

The model results have been validated against the outcomes from the clinical trials and show a reasonable fit for recurrence free survival for the 1-year and 3-year adjuvant imatinib treatment arms.

Weaknesses and Areas of uncertainty

In order to meet the patient population stipulated in the scope (patients at significant risk of disease recurrence – used synonymously with the term ‘high risk’ in the MS), the manufacturer has performed retrospective sub-population analyses of the RCTs to identify high risk patients. These analyses vary in size as a proportion of the randomised population, with the lowest being 28% in the ACOSOG Z9001 trial. Differences between the treatment arms of the trials at baseline in patient characteristics were more pronounced in the Miettinen sub-populations than

the full populations, indicating selection bias. These sub-populations are most likely underpowered for RFS and OS, though results were not significantly different between the full trial population and the high risk sub-populations (confidence intervals did not cross 1).

The manufacturer states that patients classified as at moderate risk of recurrence are not included in the submission because evidence for this sub-population is less developed, and because there is uncertainty in the prognosis for patients at 'intermediate' (moderate) risk. The ERG notes that there is an unspecified proportion of patients in the ACOSOG Z9001 that would be classified as at moderate risk, but there would be only a small proportion of patients in the SSGXVIII/AIO trial that could be similarly classed as moderate.

The results of the ACOSOG Z9001 trial are confounded by the high degree of cross-over to imatinib by recurrence-free placebo patients when the study became unblinded. The results of the placebo arm of this trial are used as a baseline for comparison to adjuvant imatinib in the manufacturer's economic model. Following the main submission document the manufacturer submitted to NICE and the ERG a supplemental report using various statistical methods to adjust for patient cross-over in the trial. These methods have advantages as well as limitations in terms of assumptions made and their applicability to the trial, and all produced RFS and OS estimates that were lower (to varying degrees) than the ITT analysis and therefore more favourable to imatinib. The manufacturer's favoured method produced HRs that are similar to a per protocol analysis that simply censors switchers at the time of cross-over, and that both of these approaches give HRs that were only slightly lower than the ITT analysis. These results are not formally incorporated into the manufacturer's assessment of cost-effectiveness. It is likely that adjustment of the treatment effects for cross-over would lower the ICERs.

The EORTC 62064 trial currently only provides limited interim results. In common with the ACOSOG Z9001 trial, this study provides a direct comparison with no adjuvant treatment but over a longer-time period (2-years). Data from this study, if fully available, would obviate the need for an indirect comparison with no adjuvant treatment (though it would only be for a 2-year and not a 3-year treatment period), and would potentially not be subject to the limitations of patient cross-over seen in the ACOSOG Z9001 trial.

None of the RCTs identified in the clinical effectiveness systematic review were conducted solely in the UK and the applicability of the evidence to NHS practice and to the UK GIST

Superseded - see erratum

population could be questioned. However, expert clinical opinion suggests that there are no important differences.

The manufacturer suggests that improvements in RFS associated with adjuvant imatinib therapy could be expected to translate into better HRQoL, but no HRQoL data were collected in the RCTs. The submission uses HRQoL data from a trial of patients with advanced GIST treated with sunitinib, but there is a lack methodological detail on procedures for valuation and an absence of information on the characteristics of patients in the study which limits the ability to critically appraise the valuations. The ERG is not aware of any other relevant HRQoL data in patients with treated for GIST, and suggest caution in the interpretation of the evidence.

The extrapolation of disease recurrence after the trial end-points is uncertain. There is no reliable data available to inform the choice of parametric distribution. The manufacturer's choice of the Gompertz distribution produces the most favourable results for adjuvant imatinib. However other parametric distributions such as the log-logistic, exponential or Weibull distributions may be more plausible.

The MS has assumed that patients continue to benefit from adjuvant imatinib after treatment has finished. This assumption appears optimistic and is likely to produce results favourable to adjuvant imatinib. Upon request by the ERG, the manufacturer supplied additional analyses that investigated the effect of a reduced treatment effect beyond 5-years, producing ICERs ranging from £4569 to £34,683 for different assumptions and across the different comparisons.

The MS does not report sub-group analyses (e.g. by tumour genetic mutation site) as requested in the NICE scope (where evidence is available). Only the SSGXVIII/AIO trial reported treatment effectiveness (in terms of RFS) for sub-groups (available only in the trial journal publication). Generally there were similar effects for the sub-groups as the main trial population (i.e. favouring 3-year treatment), though there was uncertainty for tumour genetic mutation site where numbers of patients for some mutation groupings were small.

There are some minor coding errors in the model for the calculation of health state medical costs and utility which the ERG has corrected.

Summary of additional work undertaken by the ERG

The ERG conducted the following additional analyses:

- A corrected base case, with correction of coding errors in the manufacturers model;
- A series of sensitivity analyses for this corrected base case, including alternative assumptions regarding the off treatment effect of adjuvant imatinib, the parametric distribution used for modelling recurrence-free survival, resistance to imatinib and the mortality estimates used for the recurrence health states.

In the sensitivity analyses the ERG found the results vary considerably with changes to assumptions of the parametric distribution used to model recurrence-free survival. The cost effectiveness of adjuvant imatinib remained below £30,000 per QALY for all ERG analyses.

1 Introduction to ERG Report

This report is a critique of the manufacturer's submission (MS) to NICE from Novartis Pharmaceuticals on the clinical effectiveness and cost effectiveness of imatinib for the adjuvant treatment of gastrointestinal stromal tumours (GIST). It identifies the strengths and weakness of the MS. Clinical experts were consulted to advise the ERG and to help inform this review.

Clarification on some aspects of the MS was requested from the manufacturer by the ERG via NICE on 27th February 2014. A response from the manufacturer via NICE was received by the ERG on 14th March 2014 and this can be seen in the NICE evaluation report for this appraisal.

2 BACKGROUND

The manufacturer's submission (MS) is an update of a previous submission made for NICE Technology Appraisal (TA) 196 in 2010.[1] The original MS included the ACOSOG Z9001 trial, which compared 1-year of adjuvant imatinib with placebo. At that time the NICE Appraisal Committee were aware of on-going clinical trials and scheduled an update of the appraisal to incorporate longer-term evidence for the clinical and cost-effectiveness of imatinib.

2.1 Critique of manufacturer's description of underlying health problem

The MS provides an appropriate description of GIST including its incidence, common locations within the gastrointestinal tract, and methods of diagnosis. The MS cites incidence of GIST estimates to be between 9 and 14.5 per million population.[2] A study reported an annual incidence of 13.2 per million based on a retrospective analysis of data from UK patients (January 1987 to December 2003).[3] However, guidelines published by the Royal College of Pathologists (updated in 2012) suggest that some epidemiological studies have shown higher estimated incidence of GIST of around 15 per million of population per annum (approximately 900 new cases per year in the UK).[4]

2.2 Critique of manufacturer's overview of current service provision

The MS overview of current service provision provides a clinical care pathway shown in the MS in Figure A-1 (page 26). While the current service provision appears to be accurate, the ERG's clinical advisor suggests that the pathway does not consider potential neo-adjuvant imatinib (i.e.

given before surgical resection, with the goal being a reduction in tumour size that may facilitate complete surgical resection and/or increase the likelihood of organ preservation)(NB. However, neo-adjuvant treatment is not specifically mentioned in the NICE scope). Sunitinib (Sutent®) is included in the care pathway, but no background information is given on this drug. According to the UK GIST guidelines, patients with unresectable and/or metastatic GISTs showing progression or intolerance on imatinib should be considered for switching to sunitinib (50 mg/day for 4 weeks followed by 2 weeks rest - 6 week cycle).[5] For patients with tolerability issues, lower daily doses of sunitinib given continuously may be considered (e.g. 37.5 mg/day). Acquisition costs are £3138.80 for 50mg (28-cap pack), with the first treatment cycle free to the NHS.[6]

The MS provides an informative discussion of the development of risk stratification schemes for GIST recurrence. The first widely accepted risk classification was the US NIH Consensus Criteria developed in 2002 (also referred to as the Fletcher 2002 criteria[7]), based on tumour size and tumour mitotic count. More recently a classification scheme by Miettinen was introduced which included tumour size, mitotic count and tumour location.[8] Current UK clinical guidelines[5] recommend that at diagnosis, all patients with c-KIT (CD117)-positive GIST are stratified as being at very low, low, moderate or high risk of recurrence according to Miettinen 2006 criteria[8] (see Table 1, taken from MS page 22). The ERG clinical advisor agreed that this is an appropriate and widely used risk stratification instrument, which is at present the best risk stratification instrument available as other prognostic and/or predictive factors still need validation. The MS does not appear to have omitted any other important risk classification schemes.

The MS discusses additional factors that are under investigation as predictors of recurrence (e.g. tumour rupture; age at diagnosis; gene mutations) (MS section 2.1). It is noted that current evidence does not suggest that these are independent risk factors and therefore that they should be incorporated into existing risk classification schemes. The choice to focus on the Miettinen risk classification scheme in the MS therefore seems appropriate, however, as the evidence base evolves in the future it is likely that risk classification schemes will undergo revision to incorporate additional factors, and future clinical trials of adjuvant treatments will adopt such schemes, with the aim of better patient selection and improved treatment outcomes.

Table 1: Miettinen risk stratification of primary GIST by mitotic count, size and site (Table A-2 taken from page 22 of MS)

Tumour parameters		Risk of developing progressive disease or metastases during long-term follow-up (%)			
Tumour size	**Mitotic count**	**Tumour location**			
		Gastric	**Jejunal/ileal**	**Duodenal**	**Rectal**
≤ 2 cm	≤ 5/50 HPF	None (0)	None (0)	None (0)	None (0)
> 2 to ≤ 5 cm		Very low (1.9)	Low (4.3)	Low (8.3)	Low (8.5)
> 5 to ≤ 10 cm		Low (3.6)	Moderate (24)	NA[b]	NA[b]
> 10 cm		Moderate (10)	**High (52)**	**High (34)**	**High (57)[a]**
≤ 2 cm	> 5/50 HPF	None (0)[a]	**High (50)[a]**	NA[b]	**High (54)**
> 2 to ≤ 5 cm		Moderate (16)	**High (73)**	**High (50)**	**High (52)**
> 5 to ≤ 10 cm		**High (55)**	**High (85)**	**NA[b]**	**NA[b]**
> 10 cm		**High (86)**	**High (90)**	**High (86)**	**High (71)**

HPF, high-power fields. [a] Very small case numbers. [b] Insufficient data.

The MS notes that imatinib may also be indicated for adult patients with KIT (CD117)-positive, unresectable and/or metastatic malignant GIST (MS section 2.2). This is of significance as two of the health states in the manufacturer's model are for patients with advanced/metastatic GIST (though treated with sunitinib - see section 4.2.1).

The MS lacks details such as type of patients and number of patients already receiving adjuvant imatinib. Guidelines suggest that that the majority of GIST patients are adults with a median age of 50 to 60 years, with perhaps a slight male predominance.[4] A UK study spanning 17 years reported a mean age of 64.4 years, with a slight female pre-dominance.[3] The ERG's clinical advisor commented that in local practice the standard length of treatment would be around 3-years (in common with the SSGXVIII/AIO trial). Around 90% of high risk patients offered imatinib would accept it as the drug is fairly well tolerated (see below).

The MS states that up to 90% of patients with a high risk of recurrence undergoing resection have an adverse outcome such as recurrence, metastasis or GIST-related death. Median time to recurrence may be < 2-years following complete resection, with a 5-year recurrence-free survival (RFS) of 20% (MS section 2.3 page 24). In contrast, in a UK study of patients who underwent surgical resection and were categorised at high risk of recurrence, 25% developed recurrence with mortality at 37% (mean follow up of 6.7 years).[4]

No discussion of primary resistance (no response to therapy) or secondary resistance (resistance that develops whilst taking imatinib after an initial response) to imatinib in the adjuvant or the advanced disease setting is provided. The ERG asked the manufacturer for clarification regarding definitions of resistance and how this is considered in the MS (see section 4.2.4 – sub section 'Resistance').

2.3 Critique of manufacturer's definition of decision problem

Population

The NICE scope states that the population for assessment should be adults who are at significant risk of relapse following resection of KIT (CD117)-positive GIST, based on the licenced indication. The licence specifies that patients with low or very low risk of recurrence should not receive adjuvant treatment. The Summary of Product Characteristics (SmPC) does not clearly define the criteria for significant risk, but it is suggested in the MS that it encompasses patients at high risk of recurrence and intermediate/moderate risk of recurrence (MS page 127).

The manufacturer notes that the UK GIST guidelines recommend adjuvant treatment in patients at high risk of recurrence. The guidelines also recommend the Miettinen 2006 criteria,[8] for risk stratification, and the highest Miettinen risk strata is high risk (Table 1). The manufacturer therefore considers significant risk as analogous to Miettinen high risk in their submission, though in NICE TA196[1] they considered both moderate risk and high risk Miettinen categories as analogous to significant risk and presented incremental cost-effectiveness ratios (ICERs) respectively. The current MS does not include patients at moderate risk of recurrence in the assessment of clinical and cost-effectiveness. Their justification is that evidence for patients at moderate risk is less developed and therefore the uncertainty is greater for those patients (MS section 7.2.1). For example, it is noted that the pivotal SSGXVII/AIO trial only included patients deemed to be at high risk of recurrence. Furthermore, European guidelines suggest uncertainty in the prognosis for patients at 'intermediate' (moderate) risk. The ERG considers that inclusion of Miettinen high risk patients is appropriate for the MS given that it is recommended by UK guidelines and is an appropriate and widely used risk stratification instrument in practice. The ERG considers that omission of patients with moderate risk from the MS is a limitation, but notes that there is limited clinical RCT evidence available (generally only the ACOSOG Z9001 trial could provide clinical effectiveness evidence for patients at moderate as well high risk of

recurrence, though the MS does not report the proportion of patients classified as Miettinen moderate risk – see section 3.1.3 and section 3.1.6).

Intervention

The description of the intervention in the decision problem reflects its use in the UK and is appropriate for the NHS, including licensed indication and relevant dose. Imatinib (Glivec) was approved in Europe in April 2009 for 1-year of adjuvant treatment of adult patients with a significant risk of relapse following resection of KIT (CD117)-positive GIST, with a recommended dose of 400 mg once daily. This was amended from 1 to 3-years in 2012. Acquisition costs were reported as £1724.39 for 30 x 400 mg or £862.19 for 60 x 100 mg (MS page 18).

Comparators

As stated in the MS (page 28), there is no currently accepted alternative to imatinib as adjuvant therapy for patients at significant risk of relapse following resection of primary GIST. The comparator stipulated in the NICE scope is observation after surgery (no adjuvant therapy).

Outcomes

The outcomes appear to be appropriate to the decision problem. These are overall survival (OS), recurrence-free survival (RFS), adverse events and Health related quality of life (HRQoL). There were no outcomes from the scope omitted from the decision problem, and some of the studies included in the manufacturer's systematic review of clinical effectiveness reported additional outcomes (see section 3.1.5).

Economic analysis

The MS states that the economic analysis has been conducted from the perspective of the NHS in England and Wales (MS page 10) and the analysis is therefore appropriate for the NHS. Cost categories were based on the NHS and Personal Social Services (PSS) perspective including treatment acquisition costs, management costs, monitoring costs (recurrence-free GIST) and adverse event costs.

Other relevant factors

The decision problem in the NICE scope specifies that if evidence allows, sub-group analysis by baseline risk of relapse and tumour genetic mutational status should be considered. Only the SSGXVIII/AIO trial reported pre-defined sub-groups (age, tumour site and size, mitotic count/50HPF local and central, tumour rupture, completeness of surgery and tumour mutation site). These appear to be relevant sub-groups.

There are no issues with regard to equity or equality.

3 CLINICAL EFFECTIVENESS

3.1 Critique of manufacturer's approach to systematic review

3.1.1 Description of manufacturer's search strategy

The MS reports separate searches for studies of clinical effectiveness, cost effectiveness, health related quality of life (utility values) and resource use data. The MS search strategies are considered overall to be of a reasonable quality, employing the correct use of Boolean operators and set combinations, adapted per database. The databases chosen match the minimum criteria set by NICE (i.e. Medline, Medline In-Process, Embase, The Cochrane Library). There were some minor indexing and truncation issues in the searches and it was noted that some papers were indexed on Medline as 'Postoperative Period', which was not in the search strategy. However, on checking the relevant papers, these were included in the MS reference list.

In addition, there were a few minor inconsistencies between the clinical and cost/quality of life searches. For example, Science Citation Index and Conference Proceedings Index were used in the cost but not in the clinical searches. The approach in the clinical searches was to search specific conferences. Medline and Ovid are not specified as host databases in the text for in the clinical effectiveness searches, but Ovid is recorded for the cost searches. The clinical effectiveness searches show the return number of hits per line, which are absent in the cost effectiveness and quality of life searches. Cost effectiveness and quality of life filters have been applied within the one search linked to the disease terms. It would represent best practice to run these as separate searches for greater transparency, especially in absence of number of hits per line being documented, although the more pragmatic approach can be time effective. The search strategy appeared to be of reasonable quality.

Searches of electronic databases for clinical effectiveness studies were conducted until April 2013, and searches for conference proceedings were conducted up to until June 2013. However, cost-effectiveness searches were conducted up to December 2013 (details in MS Appendix 2). The MS provided sufficient detail for a reproduction of their search methods (i.e. specified databases, dates of searches and search strategies). Given that the clinical effectiveness searches were not up to date, the ERG has therefore updated them to 18th February 2014 for electronic bibliographic databases and to 19th February 2014 for on-going trial searches (see below for details).

The MS does not report a separate search to identify adverse drug reactions. This appears a reasonable approach as the ERG considers that adverse event search filters are of questionable value and that side effects are not always reported in abstracts on bibliographic databases. No search of grey literature or hand searching was reported.

The ERG conducted the clinical effectiveness update searches using a slightly adjusted strategy (on all years in all the databases) using an RCT filter (the original MS search was for RCTs and non-RCTs). Searches identified one additional phase II RCT reported in a conference abstract and a poster (see Table 2). However, the data are likely to be of limited value to this appraisal as the trial compared 6 months with 12 months adjuvant imatinib for intermediate or high risk GIST patients (as will be discussed in section 3.1.3 of this report, RCT evidence from longer treatment periods is available).

Table 2: Additional RCT identified by ERG searches

Authors	Date	Title
Muguruma *et al.*[9]	2013	Randomized phase II study of 6 versus 12 months of adjuvant imatinib for patients with intermediate- or high-risk GIST
Yamamoto *et al.*[10]	2013	Multicentre randomized phase II trial of adjuvant imatinib for 6 versus 12 months in patients with intermediate or high risk GIST: Interim analysis results.

The interim analysis (median follow-up time of 33 months) showed that 6 months of adjuvant imatinib was inferior in efficacy to 12 months treatment in terms of RFS.[9] It was concluded that shortening of the adjuvant imatinib duration is not recommended for intermediate or high risk GIST patients.

While no systematic search of trial databases was undertaken, the MS reported searching for relevant conferences, supplemented by an electronic review of ASCO (the American Society of Clinical Oncology) abstracts. The ERG elected to search the following: UKCRN, clinicaltrials.gov, controlled-trials.com, WHO ICTRP, Cancer.gov/clinicaltrials and http://www.cancerresearchuk.org/cancer-help/trials/. Searches conducted by the ERG did not identify any additional relevant new conference abstracts.

3.1.2 Statement of the inclusion/exclusion criteria used in the study selection.

The MS states that the inclusion/exclusion criteria are detailed in the flow chart (MS Figure B-2). While no criteria are specified in the flow chart, the information is provided in the appendices (10.2.6 page 253-354). The inclusion criteria are clearly stated and are based on patients with GIST (any risk) being treated with adjuvant imatinib, reporting recurrence-free (or equivalent) and overall survival in any prospective and retrospective study including case series. Excluded were sub-groups of GIST (e.g. rectal GIST), neoadjuvant imatinib and other TKIs, studies not specifically reporting data for adjuvant imatinib (e.g. reporting for neoadjuvant and/or adjuvant imatinib, or for with and without imatinib), studies reporting data for <20 patients receiving adjuvant imatinib, and studies reported in non-English language (MS section 10.2.6 page 253-4). No definition of risk was applied and risk was not limited to 'significant risk' as per the scope.

No limits as to the quality of the RCTs were placed in the inclusion criteria. Setting was not used as inclusion criteria and does not appear to be a relevant factor.

A flow chart with the numbers of references included and excluded at each stage was presented, and appears to be correct. It is unclear why the electronic title and abstract screening was conducted twice (see illustration Figure B-2 page 38). Clarification requested from the manufacturer (see clarification request A3) states that the first round of screening focussed on the inclusion criteria and the second on the exclusion criteria. Findings from the first round of screening influenced the second round in that two extra items were added to the exclusion criteria: exclusion based on the number of patients and exclusion of GIST sub-groups. The MS did not provide a list of excluded studies and the ERG was unable to check whether any studies were excluded inappropriately. Following a clarification request (see clarification request A4), a reference list of the six studies was provided with reasons for their exclusion and

the ERG concluded that the exclusions were appropriate. The ERG is not aware of any other potential bias in the selection of studies.

It should be noted that the screening of references was carried out by one researcher, with a random quality check of 30% of all articles selected (MS section 6.2.2 page 37, repeated in appendix 10.2.7 page 254) by a second researcher, with a third researcher resolving any disputes. No justification for this approach was provided. Guidance for undertaking systematic reviews in health care recommends that all papers are independently assessed by more than one researcher, as this increases the reliability of the decision process.[11]

3.1.3 Identified studies

The MS included three RCTs led by separate clinical groups:

- American College of Surgeons Oncology Group ACOSOG Z9001 is a randomised, double-blind, placebo-controlled, phase II, multi-centre trial conducted in the United States. The trial compares 400 mg/day of adjuvant imatinib with placebo for 1-year after surgical resection in patients at any level of risk of GIST recurrence. The information is based on a published paper (DeMatteo and colleagues[12]), an abstract (Corless and colleagues[13]), unpublished data based on a 5-year follow-up for the full study population[14] and the retrospective Miettinen high risk sub-population.[15]
- Scandinavian Sarcoma Group and the Sarcoma Group of the Arbeitsgemeinschaft Internistische Onkologie SSGXVIII/AIO is a randomised, open-label, phase III, multi-centre trial conducted in Nordic countries and Germany. The trial compares 1-year with 3-years of 400 mg/day of adjuvant imatinib therapy after surgical resection in patients considered to have a high risk of GIST recurrence (based on modified NIH Consensus Criteria). The information is based on a published paper (Joensuu and colleagues[16]) and unpublished data for the retrospective Miettinen high risk sub-population.[17]
- European Organisation for Research and Treatment of Cancer EORTC 62024 is a randomised controlled, open-label, phase III, multi-centre trial launched in 2004 and conducted in Europe including the UK. The trial compares 2-years of 400 mg/day imatinib therapy following surgical resection with observation only post-surgery in intermediate- or high risk GIST patients. The information is based on a conference abstract (Casali and colleagues[18]) reporting interim results.

The MS provided summary details for all three RCTs. MS Table B-5 (page 42) summarises the methodology of the RCTs, including location, design, duration, intervention and comparator. MS Table B-6 (page 46) summarises the inclusion and exclusion criteria of the three RCTs, while baseline characteristics per treatment group are summarised for each RCT in MS Table B-7 (page 47-49). The primary and secondary outcomes for the three RCTs are summarised in MS Table B-8 (page 50). A summary of statistical analyses including sample size, power calculations and details of data management for patient withdrawals is presented in MS Table B-9 (page 53). CONSORT flow-charts with patient numbers (marked AIC[a] for details of the 5-year follow-up for the ACOSOG Z9001 trial) are reported for only the ACOSOG Z9001 and the SSGXVIII/AIO trials (page 56 and 57), as the abstract for the EORTC 62024 trial does not provide this information. Apart from the Miettinen high risk sub-population of interest, no additional sub-group data are reported.

Electronic copies of the included trials, clinical study reports (CSR) and unpublished data were provided by the manufacturer.

Both the ACOSOG Z9001 and the SSGXVIII/AIO trials received some funding from the manufacturer of imatinib (Novartis) amongst others. There was insufficient information in the MS to establish if the EORTC 62024 trial received any funding from the manufacturer. Clarification requested from the manufacturer (see clarification request A6) established that Novartis also provided some funding to this trial.

Non-randomised trials

The MS identified 12 non-RCTs (summarised in MS Table B-16 page 82). Three of the studies were retrospective[19-21] and one study was a review of two studies (ACOSOG Z9000 and Z9001).[22] Of these, two studies were set in the USA[19;20;22] and one in China.[21]

Of the eight remaining non-RCTs, two had historical controls (one set in Sweden[23] and one in South Korea[24]) and six were prospective studies.[25-27;27-29] However, three of the prospective studies had no control arm.[28-30] Of the prospective studies, three studies were set in China,[25-27] two in Japan[28;30] and the remaining study in Africa, Asia, Europe, the Middle East, the Pacific, and Russia.[29]

[a] The AIC status has been removed since this report was written

None of these studies used the Miettinen scheme[8] for categorising patients' risk or classified patients at 'significant risk'. Four studies reported no details on the patient's risk of GIST recurrence,[19-22] four studies were based on patients with low/intermediate and high risk of GIST recurrence,[23;25;26;29] and four on patients with high risk of GIST recurrence.[24;27;28;30] Seven of the non-RCTs did not report a risk stratification scheme.[19-23;26] Of the remaining five non-RCTs,[24;25;27;28;30] all based their risk stratification on the NIH Consensus Criteria, however one used a modified NIH version[27] and one added c-KIT exon 11 mutations as a criteria.[24]

While the MS summarised study details and results data in MS Table B-16 (page 82 - 96), not all of the studies were discussed in the text. Of the five non-RCTs reporting safety data, only two studies were used in the section on adverse events (MS section 6.9.3 page 104) to illustrate how well tolerated adjuvant imatinib therapy was in patients treated following GIST resection and at high risk of recurrence. No reasons for the non-inclusion of the three remaining studies or references are provided. Due to the limited value of the non-randomised trials, based on their design and restricted information available in five of the studies based on conference abstracts only,[19;23;26;28;29] only a summary is provided by the ERG.

Baseline characteristics of the included RCTs

Generally, baseline characteristics between the treatment arms in the three RCTs (of the full populations) appear to be balanced. There were some minor differences between the treatment groups in the ACOSOG Z9001 trial, with slight differences in male gender (47% imatinib vs 54% placebo), ECOG status and tumour origin (see Table 3) (MS Table B-12 page 66), and R1 resection margins (10% imatinib vs 7% placebo). Similarly, there were some slight differences in the SSGXVIII/AIO trial, including male gender (52% imatinib 1yr vs 49% imatinib 3yrs), R0/R1 margins and modified consensus classification risk group (MS Table B-15 page 72), and tumour origin (stomach: 49% imatinib 1yrs vs 53% imatinib 3yrs; small intestine 37% imatinib 1yr vs 31% imatinib 3yrs) (see Table 3).

It was unclear where the baseline data for the EORTC 62024 trial[18] were from, as they were not in the conference abstract. Clarification provided by the manufacturer (see clarification request A5) stated that the information was based on an ASCO 2013 slide presentation, which the ERG was unable to access. There were some slight differences between trial arms in ECOG status (see Table 3) (MS Table B-7 page 48) and the risk category review diagnosis (Low risk: 4%

imatinib vs 3% no treatment/observation only; Intermediate: 36% imatinib vs 33% observation only). It should be noted that between 23% - 27% of patients received no risk category review.

Differences between treatment arms in baseline characteristics of the Miettinen high risk sub-population are more pronounced than in the full populations. These data are AIC[b] for both trials, but not available for the EORTC 62024 trial.

Table 3: Selected main differences in baseline demographic and pathological characteristics of patients as reported in the primary analysis (full population) and for primary analysis patients retrospectively identified as the Miettinen high risk sub-population

ACOSOG Z9001	**Full population**		**Miettinen high risk sub-population**	
Characteristic	**Placebo n = 354**	**Imatinib n = 359**	**Placebo n = 81**	**Imatinib n = 84**
ECOG performance status, n (%)				
0	265 (74.9)	281 (78.3)	52 (64.2)	64 (76.2)
1	81 (22.9)	74 (20.6)	24 (29.6)	19 (22.6)
2	8 (2.3)	4 (1.1)	5 (6.2)	1 (1.2)
Primary tumour site, n (%)				
Stomach	235 (66.4)	209 (58.2)	43 (53.1)	30 (35.7)
Small intestine	102 (28.8)	125 (34.8)	1 (1.2)	1 (1.2)
Rectum	5 (1.4)	5 (1.4)	1 (1.2)	3 (3.6)
Other	12 (3.4)	18 (5.0)	36 (44.4)	50 (59.5)
Not available	0	2 (0.6)	0	0
Completeness of surgery, n (%) Complete resection (R0)	330 (93.2)	325 (90.5)	72 (88.9)	77 (91.7)
Tumour size range, cm				
≥ 3 to < 6, n (%)	149 (42.1)	143 (39.8)	7 (8.6)	13 (15.5)
≥ 6 to < 10.0, n (%)	119 (33.6)	123 (34.3)	22 (27.2)	30 (35.7)
≥ 10.0, n (%)	86 (24.3)	93 (25.9)	52 (64.2)	41 (48.8)
Data above reproduced from MS Table B-12 page 66 of the MS.				
SSGXVIII/AIO	**Full population**		**Miettinen high risk sub-population**	

[b] The AIC status of these ACOSOG Z9001 trial data has been removed since this report was written

Characteristic	1 year n = 199	3 years n = 198	1 year n = 142	3 years n = 139
ECOG performance status 0, n (%)	169 (85)	170 (86)	118 (83.1)	121 (87.1)
Primary tumour site, n (%) Stomach	97 (49)	105 (53)	54 (38.0)	58 (41.7)
Small intestine	74 (37)	62 (31)	62 (43.7)	56 (40.3)
Colon or rectum	16 (8)	19 (10)	15 (10.6)	13 (9.4)
Other	11 (6)	11 (6)	10 (7.0)	11 (7.9)
Not available	1 (1)	1 (1)	1 (0.7)	1 (0.7)
Completeness of surgery, n (%) Complete resection (R0)	169 (85)	160 (81)	116 (81.7)	107 (77.0)
Microscopic residual tumour suspected (R1)	29 (15)	37 (19)	26 (18.3)	32 (23.0)
Tumour rupture present, n (%)	35 (17.6)	44 (22.2)	18 (12.7)	25 (18.0)
Median tumour size range, cm	9 (2 to 35)	10 (2 to 40)	NA	NA
< 5.1, n (%)	29 (15)	18 (9)	17 (12.0)	7 (5.0)
> 5.1 to 10.0, n (%)	91 (46)	81 (41)	60 (42.2)	57 (41.0)
> 10.0, n (%)	78 (39)	98 (50)	64 (45.1)	74 (53.2)
Not available, n (%)	1 (1)	1 (1)	1 (0.7)	1 (0.7)
* Error in SSGXVIII/AIO baseline characteristic for Median mitotic count - not available for 3-year imatinib: shown as 3 92.1), presumed to be 3 (2.1). **Data above reproduced from MS Table B-15 page 72 of the MS.**				

EORTC 62024 **Characteristic**	**Full population**		**Miettinen high risk sub-population not available**
	Imatinib 2 years n=454	**Observation only n=454**	
ECOG status			
0	399 (87.9)	380 (83.7)	
1	54 (11.9)	74 (16.3)	
2	1 (0.2)	0	
Margins			
R0	381 (83.9)	381 (83.9)	
R1	70 (15.4)	72 (15.9)	
R2	1 (0.2)	1 (0.2)	
Unknown	2 (0.4)	0	
Tumour origin			

Gastric	250 (55.1)	253 (55.7)	
Other	204 (44.9)	201 (44.3)	
Risk category (NIH Consensus Criteria) review diagnosis			
Very low	0	2 (0.4)	
Low	19 (4.2)	13 (2.9)	
Intermediate	162 (35.7)	150 (33.0)	
High	168 (37.0)	168 (37.0)	
Not reviewed	105 (23.1)	121 (26.7)	
Data above reproduced from MS Table B-7 page 48 - 49 of the MS.			

There are other minor differences not shown in this table

For the Miettinen high risk sub-population in the in the ACOSOG Z9001 trial, the percentage of patients in the imatinib arm was higher for baseline characteristics such as ECOG performance status 0, primary tumour sites in the rectum and other, complete resection (R0) and tumour size range ≥ 3 to < 6 and ≥ 6 to < 10.0. The percentage of patients in the placebo arm was higher for ECOG performance status 1 and 2, primary tumour sites in the stomach, tumour rapture present and tumour size range ≥ 10.0. For the Miettinen high risk sub-population in the in the SSGXVIII/AIO trial, the percentage of patients in the 1-year imatinib arm was higher for primary tumour sites in the small intestine, colon or rectum, complete resection (R0) and median tumour size range< 5.1 and > 5.1 to 10.0. The percentage of patients in the 3-year imatinib arm was higher for primary tumour sites in the stomach, microscopic residual tumour suspected (R1) and median tumour size range > 10.0. There were other minor differences not shown in Table 3 or discussed by the ERG. It is unclear if the imbalances in baseline characteristics between the treatment arms in the Miettinen high risk sub-population in either ACOSOG Z9001 or SSGXVIII/AIO are statistically significant. The EORTC 62024 abstract did not report baseline characteristics for the high risk group.

There are differences in patient characteristics between the trials, mainly due to varying inclusion criteria. To be included in the ACOSOG Z9001 trial, adult patients had to have a complete resection of c-KIT (CD117)-positive GIST ≥ 3 cm in size, have an ECOG status of > 2, be tumour free within 28 days of trial entry and could be at any risk of recurrence. To be included in the SSGXVIII/AIO trial, adult patients had to have a c-KIT (CD117)-positive GIST removed at open surgery 1 to 12 weeks prior to randomisation and be at a high risk of recurrence based on modified US NIH Consensus Criteria. To be included in the EORTC 62024

trial, adult patients had to have localised c-KIT (CD117)-positive GIST of intermediate or high risk of recurrence based on NIH Consensus Criteria and no previous medical therapy.

All included RCTs appear to meet the inclusion criteria of the submission. However, as stated earlier in section 3.1.1, the ERG's clinical effectiveness update searches identified one additional phase II RCT published in 2013 in a conference abstract and in a poster.[9;10] The trial compared six months with 12 months adjuvant imatinib for intermediate or high risk GIST patients and, combined with the restricted information available in the abstracts, may therefore of limited value to this appraisal.

The MS identified two on-going clinical phase II, open label non-RCTs from clinicaltrials.gov, both sponsored by Novartis. Brief details of the trials were reported (MS section 1.6). One is a 5-year study of adjuvant imatinib therapy in 91 patients at significant risk of recurrence following complete resection of primary GIST – the Post-resection Evaluation of Recurrence-free Survival for gastrointestinal Stromal Tumours (PERSIST-5) (NCT00867113). At the completion of treatment patients will be followed-up for 2-years to assess survival, status of response and quality of life. The trial is scheduled to complete in 2018.

The other study is an open-label study (NCT 01172548) assessing the safety and efficacy of 2-years of adjuvant imatinib therapy compared with historical data, though risk status of patients was not reported. The estimated completion date is March 2014.

The MS states that it is not clear whether any results from these studies will be available in the next 12 months. No additional potentially relevant ongoing trials were identified by the ERG.

3.1.4 Description and critique of the approach to validity assessment

The MS quality assessed all studies including the non-randomised trials following Centre for Reviews and Dissemination (CRD) criteria. A summary of the quality assessment for the included RCTs is presented in MS Table-10 (page 59), with a more in-depth table presented in the appendices (MS Appendix 3, section 10.3.1 page 254; non-randomised trials MS Appendix 7, section 10.7.1 page 259).

The ERG repeated the quality assessment of the RCTs. The manufacturer's quality assessment was based on criteria specified by NICE (see Table 4). There were some differences between the quality assessment judgements of the MS and the ERG.

Table 4: Manufacturer and ERG assessment of trial quality

		ACOSOG Z9001	SSGXVIII/AIO	EORTC 62024
1. Was randomisation carried out appropriately?	**MS:**	Yes	Yes	Not clear
	ERG:	Yes	Yes	Not clear
Comment: while randomisation was possible for the whole population, for the Miettinen high risk sub-population however, randomisation of patient characteristics between the treatment and control group may no longer have been maintained.				
2. Was concealment of treatment allocation adequate?	**MS:**	**Not clear**	**N/A**	**N/A**
	ERG:	**Not clear**	**Not clear**	**Not clear**
Comment: In the SSGXVIII/AIO trial no clear statement was made, however, page 1266 of the trial journal paper[16] states that the results of randomisation were communicated to study centres by fax, therefore lessening the risk that clinicians and investigators recruiting patients would be aware of the allocation sequence.				
3. Were groups similar at outset in terms of prognostic factors?	**MS:**	**Yes**	**Yes**	**Yes**
	ERG:	**Not clear**	**Not clear**	**Not clear**
Comment: treatment groups were generally similar for the full population for all three RCTs, but differences between the treatment arms in the Miettinen high risk sub-population were more pronounced in both the ACOSOG Z9001 and the SSGXVIII/AIO trial (described as well-balanced MS page 71). The MS did not report any statistical testing for baseline characteristics between the treatment arms of the Miettinen high risk sub-population. While baseline characteristics in the EORTC 62024 trial were reported in the MS, it is unclear where the information is from. Clarification provided by the manufacturer (see clarification request A5) stated that the information was based on an ASCO 2013 slide presentation, which the ERG was unable to access. Characteristics appear to be mostly similar, but it is unclear if the data are based on the full population or only the high risk group.				
4. Were care providers, participants and outcome assessors blind to treatment allocation?	**MS:**	**Yes**	**No**	**No**
	ERG:	**Partly**	**No**	**No**
Comment: the ACOSOG Z9001 trial was a double-blind, placebo controlled trial, but it is not clear if outcome assessors were blinded (publication only states that patients and investigators were blinded). The SSGXVIII/AIO trial and the EORTC 62024 trial were described as open-label. It is assumed that as not otherwise stated, the outcome assessors were not blind to treatment allocation.				
5. Were there any unexpected imbalances in drop-outs between groups?	**MS:**	**No**	**No**	**No**
	ERG:	**No**	**Yes**	**Not clear**
Comment: there were no unexpected reported imbalances in drop-outs for the full population in the ACOSOG Z9001 trial. There were more early discontinuations in the 3-year group than the 1-year group (51 vs 25) in the SSGXVIII/AIO trial, although the ITT analysis and censoring should account for this. The EORTC 62024 abstract reports 17% of imatinib treated patients discontinued early, but there is no information for the control arm and no breakdown for the high risk patient group.[18]				
6. Is there any evidence that authors measured more outcomes than reported?	**MS:**	**No**	**No**	**No**
	ERG:	**No**	**No**	**No**
Comment: no evidence of selective reporting				
7. Did the analysis include an	**MS:**	**Yes**	**Yes**	**Not clear**

ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?		**Not clear**	**Yes**	**Not clear**
	ERG:	**Yes** **Not clear**	**Yes** **Yes**	**Not clear** **Not clear**
Comment: the ITT analysis refers to the full population, not the Miettinen high risk sub-population.				

3.1.5 Description and critique of manufacturer's outcome selection

The MS indicated that all outcomes stated in the scope (OS, RFS, AEs of treatment and HRQoL) are covered, however no data for HRQoL were collected by the three included RCTs (MS section 6.10.2 page 110).

The MS appears to report all relevant trial outcomes. In the ACOSOG Z9001 trial, RFS was defined as 'the time from patient registration to the development of tumour recurrence or death from any cause' and OS as 'the time from patient registration to death from any cause'. RFS in the SSGXVIII//AIO trial was defined as 'the time period from the date of randomisation to the earliest date of recurrence (first date at which the physician suspected GIST recurrence leading to cytological or histological confirmation or radiological evidence of recurred GIST) or death from any cause' (MS page 52). OS was defined as 'the time period from the randomisation date to death from any cause plus 1 day, was a secondary endpoint' (MS page 52). The EORTC 62024 trial reports OS, RFS and the outcome 'imatinib failure-free survival' (IFS), where failure was defined as the time at which patients had to be changed to treatment with a different tyrosine kinase inhibitor owing to disease relapse or recurrence. The trial investigators describe this as a new end-point for the adjuvant setting and it was designed to incorporate secondary resistance. The manufacturer notes that this is not a generally recognised end-point and has not been included in other studies of adjuvant GIST (MS Table B-8 page 50). The ERG notes that, while this was not used in the other RCTs of adjuvant treatment in the submission, a similar endpoint has been used in an RCT of patients with controlled advanced GIST to assess the effects of interrupted or continuous imatinib treatment (the BFR14 trial, Blay and colleagues (2007)[31]). In that trial the (secondary) outcome was 'time to imatinib resistance', calculated from the date of random assignment to the date of progression under imatinib 400 mg/d or date of last follow-up. The ERG clinical advisor suggested that an outcome such as IFS is more relevant than RFS in the adjuvant treatment setting as imatinib is more likely to suppress rather than eradicate residual disease in patients with GIST - therefore it is more likely that it will delay rather than prevent recurrence. The advisor also noted that there are concerns about

accelerated development of secondary resistance in patients receiving adjuvant imatinib. Therefore outcomes that specifically take into account resistance in the adjuvant setting are relevant.

3.1.6 Description and critique of the manufacturer's approach to trial statistics

All three RCTs in the MS aimed to assess the clinical effectiveness of imatinib as adjuvant treatment for GIST, but had slightly differing hypotheses. The investigators of the ACOSOG Z9001 trial proposed that adjuvant treatment with imatinib would improve RFS compared with placebo in patients who underwent resection of localised, primary GIST. In contrast, the hypothesis of the SSGXVIII/AIO RCT was that longer than 1-year of adjuvant imatinib treatment might be beneficial, and compared 3-years of imatinib administration to 1-year of administration as adjuvant treatments for patients who were considered to have a high risk of GIST recurrence following surgery. The hypothesis of the EORTC 62024 trial is not explicit in the conference abstract,[18] but the MS reports the hypothesis objective to be 'IFS in patients with localised GIST treated with adjuvant imatinib (IM)' (MS Table B-9 page 54).

The MS presents results for RFS, OS and adverse events (MS Table B-8 page 50) for the ACOSOG Z9001 and the SSGXVIII/AIO RCTs. RFS was the primary outcome in both of these RCTs (though in the ACOSOG Z9001 trial the primary outcome was originally OS but this was changed 6 months before the first planned interim analysis – the rationale is explained on MS page 51). For the EORTC 62024 trial the MS presents IFS and RFS. The original primary outcome in this trial was OS, but this was changed to IFS by the trial's independent data monitoring committee (the full rationale is explained on MS page 77).

All three trials were powered statistically for their primary outcomes. In the ACOSOG Z9001 trial a total of 803 patients were required for a median RFS of 4.9 years in the imatinib group with an HR of 0.71 (a 40% improvement in RFS) (MS Table B-9 page 53). The SSGXVIII/AIO RCT assumed an HR of 0.44 in favour of the 3-year imatinib group. At least 110 events were required in the efficacy population, with 160 patients needed in each group, increased to 200 patients per group to account for an assumed drop-out rate of 20% (MS Table B-9 page 53 - 54). In the EORTC 62024 trial, a planned accrual of 400 patients was escalated to 900 patients, presumably when the primary outcome was changed from OS to IFS. No further details are given in the conference abstract on the statistical power calculation.

Both the ACOSOG Z9001 and SSGXVIII/AIO RCTs used Cox proportional hazards regression models (MS Table B-9 page 53 - 54). The ACOSOG Z9001 model was stratified by tumour size for RFS, but was not stratified for OS because of the few recorded deaths. In the SSGXVIII/AIO trial survival between groups was compared using the Kaplan-Meier life-table method and unstratified log-rank test (*p* values) or an unstratified Cox proportional hazards model (HRs). HRs with 95% CIs and Kaplan-Meier plots are given for RFS and OS in both trials. In the SSGXVIII/AIO trial all *p*-values were two-sided and not adjusted for multiple testing. In both trials the Cox proportionality assumption was tested. The ACOSOG Z9001 trial used Schoenfeld residuals, and the assumption was reported to be valid for all the analyses (no further information given). Schoenfeld residuals[32] are one of a number of accepted methods of testing the proportionality assumption.[33] In the SSGXVIII/AIO trial the assumption was tested by evaluating the time varying interaction of the log transformation of time to the event by the treatment arm variable. The assumption of proportional hazards would be supported by "parallel" lines in a log-log plot for each treatment group. There was no statistically significant interaction, indicating that the assumption of proportionality was supported (page 1788 of the CSR[17]). In the log-log plot there was some degree of non-proportionality in the lines for the treatment groups, but the curves did not cross (page 1803 of the CSR[17]). To further examine the nature of any non-proportionality over time, a smoothed plot of Schoenfeld residuals against log(time) was presented (page 1788 of the CS[17]).

In the ACOSOG Z9001 trial censoring time was defined as the last date that a patient was known to be alive and recurrence free (MS Table B-13 page 67). No further information is given on censoring of patients who withdrew or were lost to follow-up. In the SSGXVIII/AIO trial patients who were alive without recurrence were censored on the date of last follow-up. Patients lost to follow-up were censored on the date of the last follow-up visit. It is not clear how censoring was performed in the EORTC 62024 trial.

Both the ACOSOG Z9001 and SSGXVIII/AIO trials conducted intention-to-treat (ITT) analyses (MS Table B-9 page 53 – 54). In the ACOSOG Z9001 trial patients were analysed by randomised group, for RFS and OS (all randomised patients, n=713). (NB. Placebo patients who did not have a recurrence at study unblinding were permitted to cross-over to receive imatinib (MS section 6.3.8 page 54), thus confounding results – see below). The safety population comprised all patients receiving at least one dose of their assigned treatment (n= 682). In the SSGXVIII/AIO trial there were three analysis populations:

- modified full population (also referred to as the modified ITT population (mITT) in the trial journal publication[16]) – all randomised patients who signed informed consent (n=397);
- the efficacy population – patients who signed consent and had centrally confirmed GIST and did not have metastases resected prior to study entry (n=358);
- safety population – patients who took at least one dose of the study medication (n=392).

The trial journal publication reports RFS and OS results for both the mITT and the efficacy population, whereas the MS just reports the mITT. The population analysis in the EORTC 62024 trial was not explicitly stated, other than of 908 patients randomised, 835 were eligible.

Only the SSGXVIII//AIO trial reported predefined exploratory patient sub-group analyses, giving RFS estimates for the mITT population for the following variables: age, tumour site, tumour size, local and central mitotic count, tumour rupture, completeness of surgery and tumour mutation site. As results for these sub-groups are given for the mITT population and not the Miettinen high risk sub-population they are only reported in the trial journal publication (in Figure 3)[16], not the MS. It is unlikely that the trial is sufficiently powered for the sub-group analysis. Furthermore, the manufacturer commented in their response to the ERG clarification questions (see clarification request A7) that the number of patients in the sub-groups would be smaller if the analyses were restricted to the Miettinen high risk population. The manufacturer's economic evaluation therefore does not estimate cost-effectiveness for these patient sub-groups (though mitotic count, tumour size and tumour site are criteria for classifying risk recurrence status, and are therefore taken into account in cost-effectiveness estimates for the Miettinen high risk patients). As described later in this report (section 3.3) results of the sub-group analyses were generally similar to those of the full trial population, with statistically significant effects for 3 - years compared to 1-year of treatment for most sub-groups.

The ACOSOG Z9001 trial was not designed to investigate sub-group analyses, but did assess RFS by tumour size ((≥3–<6 cm, ≥6–<10 cm, or ≥10 cm, see Figure 3 in the trial journal publication[12]).

Of the three RCTs, the EORTC trial is currently only available as a planned interim analysis, carried out after 115 IFS events with a median follow-up of 4.7 years. The other two RCTs report long-term results at 5-years.

Classification of recurrence risk status

To meet the scope for the appraisal, the manufacturer conducted retrospective sub-population analyses of the ACOSOG Z9001 and SSGXVIII/AIO RCTs, specifically to identify patients classed as at significant (high) risk of recurrence using the Miettinen risk classification 2006 criteria.[8] As discussed in section 2.3 of this report, the manufacturer considers Miettinen high risk to be analogous to 'significant risk' (included, but not explicitly defined, in the SmPC for adjuvant imatinib).

The ACOSOG Z9001 RCT was designed before the introduction of risk stratification schemes for GIST, and patients were stratified only by tumour size, which was the main known risk factor for recurrence at that time. In this trial there were more patients in the smallest tumour size category (41%), compared to the medium size category (around 34%) and to the largest size category (25%), indicating a patient population predominantly at low or medium risk of recurrence. The SSGXVIII/AIO RCT was specifically conducted to assess adjuvant treatment in patients who were considered to have a high risk of GIST recurrence, and stratification was performed using the US NIH Consensus Criteria.[7] The two RCTs therefore varied in the proportion of enrolled patients classified at high risk of recurrence.

Of the 713 randomised patients in the ACOSOG Z9001 trial, data from a total of 627 (88%) patients were available for risk classification (at the time of the primary analysis data from 556 (78%) patients were available, as reported in NICE TA196[1], with data from a further 71 patients available at the five year follow-up). A total of 165 (23%) patients were retrospectively classified as high risk of recurrence on the Miettinen criteria and these were evenly distributed between the trial arms (n=84 in the imatinib group and n=81 in the placebo arm). At 5-year follow-up, when data from the further 71 patients were available 36 were classified as at high risk, bringing the total number of high risk patients to 201 (28%) (103 in the imatinib group and 98 in the placebo group). Note that the MS does not report the proportion of patients who could be classified as Miettinen moderate risk (in relation to the earlier discussion of the omission of moderate risk patients from the MS - section 2.3 of this report).

Of the 397 patients analysed in the SSGXVIII/AIO RCT, a total of 281 (71%) patients (142 in the 1-year group and 139 in the 3-year group) were retrospectively classified as Miettinen high risk. It is of note that under the NIH Consensus Criteria 90% of patients were classified as at high risk of recurrence (despite the eligibility criteria stating patients had to be high risk), with the remaining 10% at intermediate risk, low risk, or undetermined risk; MS Table B-7 page 48).

Therefore the proportion of randomised patients available for the Miettinen high risk sub-population analysis is lower than the proportion in the mITT analysis (NIH Consensus Criteria), potentially further reducing the statistical power of this sub-population. Of the 908 patients randomised in the EORTC 62024 trial a total of 336 (37%) were classified as at high risk of recurrence by the NIH Consensus Criteria (MS Table B-7 page 49). The conference abstract reports that the 336 were high risk GIST by local pathology, but that there were 528 (58%) patients at high risk by centrally reviewed pathology (NB. IFS is reported for both sets of patients in the conference abstract – see section 3.3 of this report). The manufacturer has not been able to conduct sub-population analysis based on the Miettinen criteria for this trial (they have no access to the data; MS section 6.2.6 page 41).

Patient cross-over

The design of the ACOSOG Z9001 trial permitted patients in the placebo group to cross-over to receive imatinib in the event of a recurrence, or if without a recurrence at the point of study unblinding (primary efficacy analysis on 12th April 2007; MS page 65). MS Figure B-11 (page 140) provides a flowchart of the number of placebo patients that did not experience a recurrence and who crossed over to imatinib, for both the ITT population and the Miettinen high risk sub-population. Placebo patient cross-over in the event of a recurrence is accounted for by censoring in the Kaplan-Meier survival analysis. However, as the MS acknowledges, the 5-year follow-up analysis is confounded by the majority of the placebo patients who were recurrence-free at the time of study unblinding opting to cross-over to active treatment for 1-year (MS page 69).

In the MS the manufacturer reports 'supportive analyses' for RFS and OS which removed recurrence free patients who crossed over from placebo to 1-year of imatinib treatment after 12 April 2007 (MS page 69 - 70) (see section 3.3 of this report – the ERG presumes that this analysis relates to the full trial population rather than the Miettinen high risk population). Since the MS was written the manufacturer submitted to NICE and the ERG a report which attempts to adjust for the confounding effect of cross-over in the Miettinen high risk population.[34] The report critiques and applies three methods for accounting for patient cross-over in survival analyses: rank preserving structural failure time model (RPSFTM)[35]; the Iterative Parameter Estimation Algorithm (IPEA)[36], and Inverse Probability of Censoring Weights (IPCW)[37]. "Exploratory" adjusted RFS and OS results are given using these three methods as well as "naïve" per protocol analyses which censors crossovers at time of switch or which excludes them

altogether. The report gives a detailed appraisal of the strengths and weaknesses of each approach, noting that all have advantages and disadvantages and no single approach has strengths that would make it more appropriate overall for this trial. However, on balance, the report proposes that the most reliable method in this instance is the IPCW. The ERG notes that the application of all the methods produces HRs for RFS and OS that are lower (to varying degrees) than the ITT analysis and therefore more favourable to imatinib (Table 1 and Table 2 in the supplemental report)[34]. The ERG agrees that all the methods have advantages as well as limitations and that the IPCW method appears to be appropriate. It is also noteworthy that the IPCW method produces HRs that are similar to a per protocol analysis that simply censors switchers at the time of cross-over, and that both of these approaches give HRs that are only slightly lower than the ITT analysis (HRs approximately 0.1 to 0.2 lower, compared to bigger differences for some of the other methods, so a more conservative estimation). As these estimates were only made available after the MS had been submitted to NICE the cost-effectiveness estimates in the submission do not account for the confounding effect of cross-over. The ERG considers that inclusion of HRs that adjust for patient cross-over will likely lower the ICERs for adjuvant imatinib.

Summary

All three trials were powered statistically for their primary outcomes. In two of the trials the original primary outcome was OS, but in both cases this was changed (subject to approval from authorities) to outcomes that reflected time to recurrence during the trials due to prognostic improvement in survival in GIST patients noted from other studies. The trials used Cox proportional hazards regression models to estimate treatment effects, and satisfactorily tested the proportionality assumption. ITT analysis, using appropriate methods, was performed in the ACOSOG Z9001 and SSGXVIII/AIO trials. The latter trial also reported an efficacy analysis. The EORTC trial is currently only available as a planned interim analysis at a median follow-up of 4.7 years, whereas the other two trials have fully published primary analyses and long-term follow-up results in the MS.

Treatment effect estimates (RFS and OS) for high risk patients in the MS are based on retrospectively classified sub-population analyses, varying in size, and are most likely underpowered. However, as reported in section 3.3 of this report, results for RFS and OS were not significantly different between the full trial population and the high risk sub-populations (confidence intervals did not cross 1).

The 5-year follow-up analysis of the ACOSOG Z9001 trial is confounded by the majority of the placebo patients who were recurrence-free at the time of study unblinding opting to cross-over to active treatment for 1-year. A number of statistical methods to account for patient cross-over in survival analyses are proposed in a supplemental report. All have advantages and disadvantages and no single approach has overall strengths, though the IPCW method is favoured by the manufacturer, with caveats. These estimates (which are slightly more favourable to imatinib) are not currently reflected in the manufacturer's cost-effectiveness analyses. The ERG considers that inclusion of HRs that adjust for patient cross-over will likely lower the ICERs for adjuvant imatinib.

3.1.7 Description and critique of the manufacturer's approach to the evidence synthesis

A narrative synthesis is provided with results reported in tables, text and Kaplan-Meier plots (MS section 6.5, page 59).

Meta-analysis was not performed though an indirect comparison of the ACOSGO Z9001 and SSGXVIII/AIO trials was conducted, to inform the economic analysis (MS section 6.7, page 79 and MS section 7.3, page 151). It should be noted that according to the manufacturer, the indirect comparison does not follow standard statistical methods as its only purpose was to populate the economic model (MS section 6.7.1 page 79). The ERG comments on the methodology used in section 0 of this report.

3.2 Summary statement of manufacturer's approach

Table 5 provides an assessment of the quality of the manufacturer's systematic review. The MS states that screening titles and abstracts for inclusion was conducted by one researcher and a second performed a random quality check of 30% of all articles selected. As previously stated, guides on conducting systematic reviews recommend that tiles and abstracts are screened independently by an additional person.[11] It is not clear whether full reports were screened by only one researcher or whether a second person performed a check. It is stated that all data extraction was fully validated by a second reviewer, but it is not clear whether this also includes quality assessment.

Table 5: Quality assessment (CRD criteria) of MS review

CRD Quality Item: score Yes/ No/ Uncertain	
1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question?	YES – see MS Appendix 2 (MS section 10.2.6).
2. Is there evidence of a substantial effort to search for all relevant research? i.e. all studies identified	Uncertain - The databases chosen match the minimum criteria set by NICE. However, the clinical effectiveness searches were updated only as far as April 2013 (databases) and June 2013 (abstracts), and the ERG identified an additional relevant abstract, presented at a conference in January 2013.
3. Is the validity of included studies adequately assessed?	YES – the criteria suggested by NICE have been used (MS Table B-10 page 59, and appendix 10.3 page 254, and appendix 10.7 page 259 for non-RCT evidence). The MS does not provide a narrative summary or discussion of the methodological quality of the evidence base as a whole though. Also, it is not explicit whether quality assessment judgements were performed independently by more than one researcher, or were checked by a second researcher.
4. Is sufficient detail of the individual studies presented?	YES – Details are tabulated in MS Section 6.3 (page 41), with CONSORT flow charts given for the two published RCTs (page 56 – 57), marked AIC[c] for details of the 5-year follow-up for the ACOSOG Z9001 trial.
5. Are the primary studies summarised appropriately?	YES – a narrative synthesis was appropriate given that meta-analysis would not have been feasible due to methodological differences between the included RCTs (and between the RCTs and non-randomised studies).

The evidence submitted generally reflects the decision problem in the MS, though, as noted above, to assess the clinical effectiveness of imatinib in high risk patients the MS has had to perform retrospective sub-population analysis of the RCTs, which is subject to methodological weaknesses (as detailed above).

Overall there is low chance of systematic error in the systematic review of the MS based on the methods employed. However, there are limitations in the search strategy, and in inclusion screening, as detailed above.

3.3 Summary of submitted evidence

Summary of results recurrence free survival (RFS)

The MS provides two sets of RFS analyses for the ACOSOG Z9001 trial:

[c] The AIC status has been removed since this report was written

(i) the primary analysis (cut-off date of April 2007, with a median follow-up of 19.7 months) (MS Table B-13 and Figure B-5, page 67 - 68), including 165 patients classified as Miettinen high risk; and

(ii) the 5-year analysis (cut-off date of 15th March 2011, median follow-up of 46.3 months) including a total of 201 patients classified as Miettinen high risk (an additional 36 patients above the primary analysis).

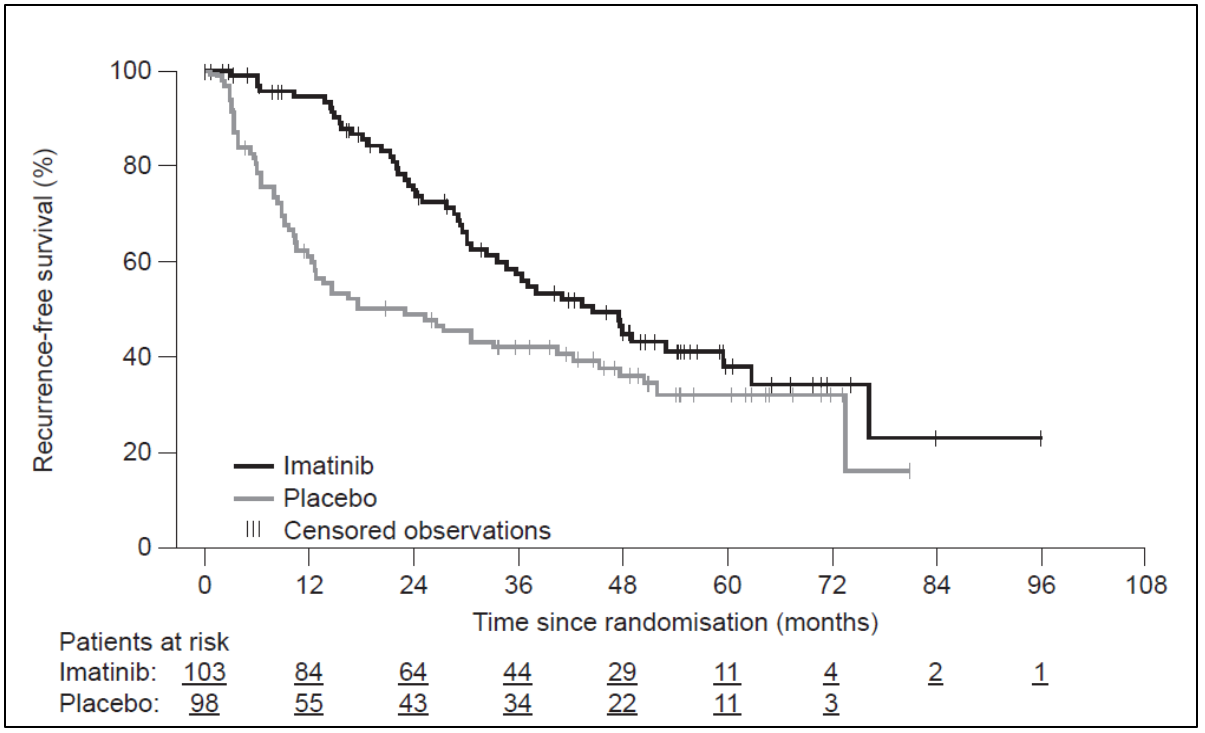
In both analyses RFS probabilities are given for the full ITT population and for the Miettinen high risk sub-population. In the ERG report we report both the ITT and Miettinen high risk sub-populations, but only for the 5-year analysis as this provides an assessment of longer-term follow-up.

Table 6 and Figure 1 show the RFS probabilities for 1-year imatinib treatment and for placebo, based on the 5-year follow-up analysis.

Table 6: RFS probabilities based on 5-year follow-up analysis of the ACOSOG Z9001 trial

	Full population		**Miettinen high risk sub-population**	
Time period	**Imatinib (n = 359)**	**Placebo (n = 354)**	**Imatinib (n = 103)**	**Placebo (n = 98)**
1 year	98.1 (96.5 to 99.6)	85.7 (82.0 to 89.5)	94.6 (90.0 to 99.2)	61.0 (51.1 to 71.0)
5 years	72.8 (67.1 to 78.4)	68.4 (63.0 to 73.8)	37.9 (25.9 to 49.9)	32.1 (21.6 to 42.6)
HR	0.718 (0.531 to 0.971)		0.608 (0.417 to 0.886)	
p-value	0.0305		0.009	

Adapted from MS Table B-14 page 69. Full data by up to 8 years is provided in the MS table. 95% CIs in parentheses.



(Reproduced from MS Figure B-6, page 70)

Figure 1: RFS rates for Miettinen high risk sub-population receiving imatinib or placebo in the ACOSOG Z9001 study based on the 5-year follow-up analysis

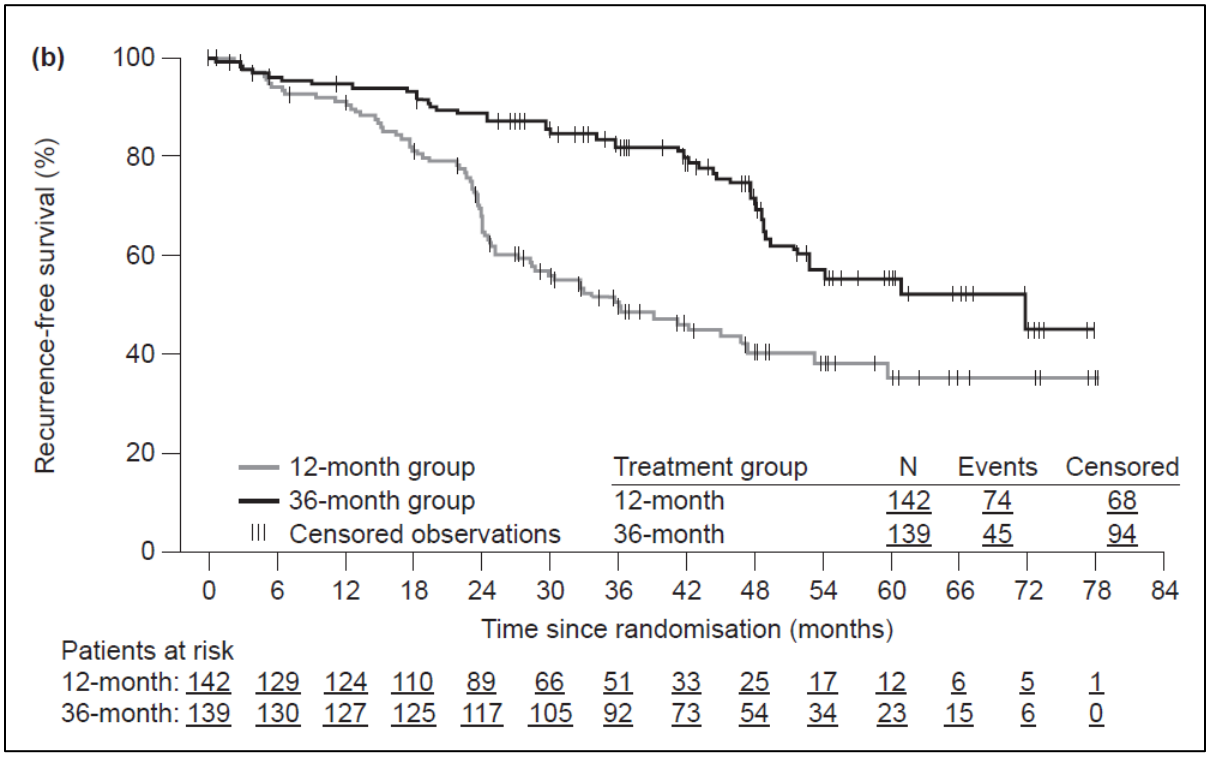
There was a statistically significant treatment effect in both the full ITT population (HR 0.718 (0.531 to 0.971; *p* = 0.0305) and the high risk sub-population (HR 0.608 (0.417 to 0.886; *p* = 0.009), with a slightly lower HR in the latter indicating increased benefit for these patients (an HR of less than 1 indicates a treatment effect in favour of imatinib). As can be seen from Figure 1, the difference in RFS between the trial arms increasingly narrowed over the follow-up period, indicating the attenuation of the treatment effect over time, potentially caused by cross-over of placebo patients to imatinib. The manufacturer reported a 'supportive analysis' which removed patients who crossed-over to placebo when the study became unblinded, with an HR of 0.671 (95% CI 0.491 to 0.919, *p* = 0.0123) which is slightly lower than the HR for the full ITT population (to which, it is presumed, this analysis relates to).

Table 7 and Figure 2 show the KM curve for the Miettinen high risk sub-population in the SSGXVIII/AIO trial (the MS also provides the KM curve for the mITT population – see MS Figure B-7, page 74). Results are reported for a median duration of follow-up of 54 months for the mITT population. In the Miettinen high risk sub-population there was a statistically significant

increase in RFS for patients treated for 3-years compared with 1-year, HR 0.43; 95% CI 0.30 to 0.62; $p < 0.001$). This was similar to the mITT population (0.46; 95% CI 0.32 to 0.65; $p < 0.0001$). Differences between trial arms were apparent from 18 months.

Table 7: RFS probabilities based on 5-year follow-up analysis of the SSGXVIII/AIO trial

	mITT population		**Miettinen high risk sub-population**	
Time period	**Imatinib 1 year**	**Imatinib 3 years**	**Imatinib 1 year**	**Imatinib 3 years**
18 months	86.8%	94.3%	81.6%	93.4%
4 years			40.2%	72.0%
5 years	47.9%	65.6%		
5-year HR	0.46; 95% CI 0.32 to 0.65		0.43; 95% CI 0.30 to 0.62;	
p-value	$p < 0.0001$		$p < 0.001$	



(Reproduced from MS Figure B-7(b) page 74)

Figure 2: Kaplan–Meier estimate of recurrence-free survival in the Miettinen high risk sub-population in the SSGXVIII/AIO trial

In the mITT population the median time to recurrence was 53.2 months for the 1-year group, but it was not reached for the 3-year group. In the Miettinen high risk sub-population the median

time to recurrence was 35.9 months in the 1-year adjuvant imatinib group and 71.8 months in the 3-year adjuvant imatinib group. For the mITT population, there was no significant difference in the hazard of GIST recurrence or death between the two trial arms during the first year or 3 - years after randomisation (HR 0.64; 95% CI, 0.26 to 1.57, and HR 1.31; 95% CI, 0.65 to 2.62, respectively), but a significant difference emerged during 1 to 2-years and 2 to 3-years after randomisation (HR 0.26; 95% CI 0.13 to 0.53; and HR 0.17; 95% CI 0.07 to 0.39, respectively). It is not reported whether this was also the case for the Miettinen high risk sub-population. The MS only reports results for the mITT population, however the trial journal publication[16] reports that RFS results were similar between the efficacy population (patients who signed informed consent, had centrally confirmed GIST, and did not have metastases resected prior to study entry) and the mITT population (randomised patients who signed informed consent).

Finally, interim results are available from the conference abstract of the EORTC 62024 trial. The median follow-up was 4.7 years, with 835 of 908 (92%) randomised patients available for assessment. Table 8 reports RFS, showing a difference between imatinib and no adjuvant treatment at 3-years, but an attenuation of the difference by 5-years.

Table 8: Interim RFS results for the EORTC 62024 trial

Time period	Imatinib 2 years	No adjuvant therapy
3 years	84%	66%
5 years	69%	63%

The conference abstract reports p<0.001 but it is not explicit whether this applies to the 3-years RFS or the 5-years RFS, or the whole period.

The 5-year IFS (imatinib-failure-free survival, the trial's primary end-point) was similar between the 2-year imatinib group and the no adjuvant therapy group (Table 9). The HR was 0.80 (98.5% CI, 0.51 to 1.26; p = 0.23). Likewise, the 5-year IFS was similar between the trial arms in the sub-population of patients with NIH Consensus Criteria classified high risk GIST (p=0.11).

Differences between the treatment arms in the high risk population were not statistically significant (p = 0.44). However, as this is only an interim analysis, caution is advised in the interpretation of these results.

Table 9: Interim 5-year IFS results for the EORTC 62024 trial

Time period	Imatinib 2 years	No adjuvant therapy
Full population, n=835	87%	84%
High risk population, n=528	79%[a]	73%

[a] Reported as 77% in the MS (page 77) and 79% in the conference abstract.[18]

Summary of results for overall survival

There were few deaths in the ACOSOG Z9001 trial with non-statistically significant HRs at 2 and 5-years (Table 10, results given in the MS for the full ITT population only). A sensitivity analysis that censored for placebo patients eligible to cross-over to receive imatinib gave a slightly lower HR of 0.746; 95% CI 0.441 to 1.262; *p* = 0.2725.

Table 10: Summary of OS in the ACOSOG Z9001 trial

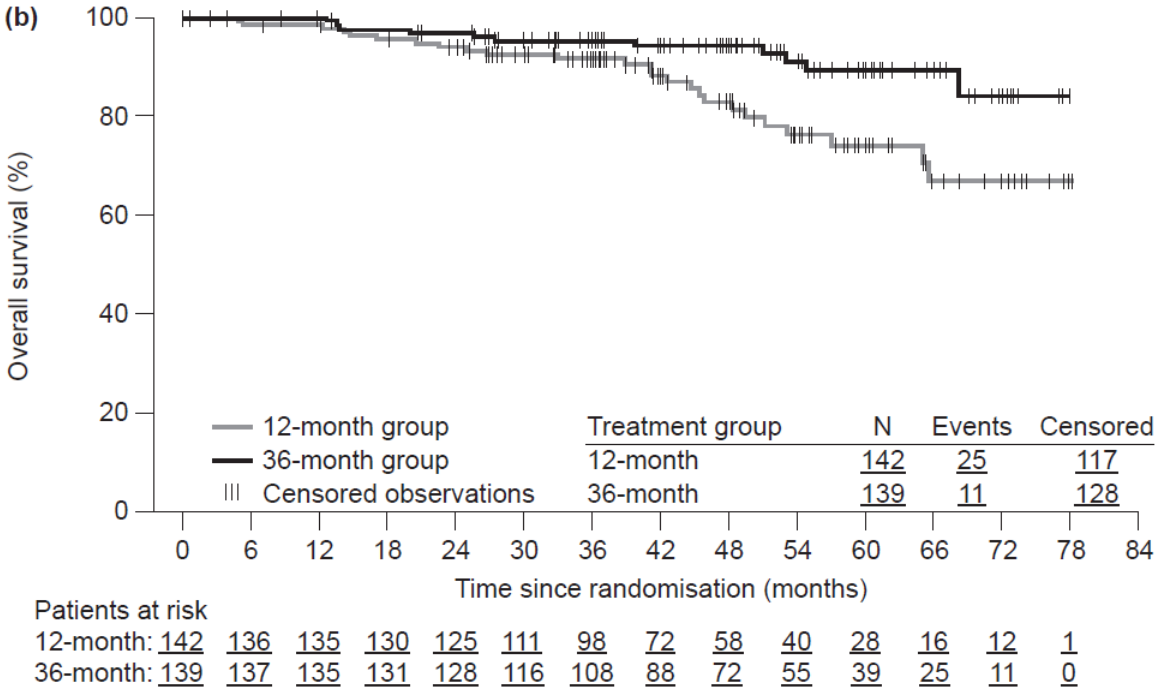
	Full population	
Time period	**Imatinib (n = 359)**	**Placebo (n = 354)**
2 years	98.8%	97.6%
2 year HR	0.66; 95% CI 0.22 to 2.03, *p* = 0.47[a]	
5 years[b]	91.3%	91.1%
5 year HR	0.816; 95% CI 0.488 to 1.365; *p* = 0.4385	

[a] *p* value from CSR report.[15] [b] Median follow-up of 60.2 months.

In the SSGXVIII/AIO trial a total of 25 patients in the 1-year imatinib group died during the study compared to 12 in the 3-year imatinib group. As Table 11 and Figure 3 show, 5-year OS was statistically significantly longer for patients treated with imatinib for 3-years, compared to 1-year (HR 0.45; 95% CI 0.22 to 0.89; *p* = 0.019). This was also the case for the Miettinen high risk sub-population (0.39; 95% CI 0.19 to 0.79; p = 0.007).

Table 11: Summary of OS in the SSGXVIII/AIO trial

	mITT population		Miettinen high risk sub-population	
Time period	**Imatinib 1 year**	**Imatinib 3 years**	**Imatinib 1 year**	**Imatinib 3 years**
5 years	81.7%	92.0%	74.2%	89.5%
HR	0.45; 95% CI 0.22 to 0.89; *p* = 0.019		0.39; 95% CI 0.19 to 0.79; *p* = 0.007	



(Reproduced from MS Figure B 8(b), page 76)

Figure 3: Kaplan–Meier estimate of OS in the Miettinen high risk population in the SSGXVIII/AIO trial

The conference abstract available for the EORTC 62024 trial reported 5-year OS as 100% in the 2-year imatinib group and 99% in the no adjuvant therapy group, but did not report OS for the high risk GIST population.

Summary of results for RFS and OS for non-RCTs in a high risk population

As can be seen from the comparative non-RCTs reporting a high risk patient group in Table 12 and non-comparative trials in Table 13, results are supportive of the three included RCTs in the submission for RFS and OS, where reported. Generally, imatinib if taken for 3 or more years is associated with better RFS and OS than taken for shorter periods.

Table 12: Results of comparative non-RCTs in high risk patients

Study, Risk category (system)	Treatment	Results, HR (95% CI)
Conely et al. 2012[19] High risk (NR)	Intervention: short-term imatinib treated (6 months to 1 year) Follow-up: 884 days Control: long-term imatinib treated (≥ 2 years) Follow-up: 963 days	*Disease recurrence rates:* Imatinib 6 months to 1 year: 7.3% Imatinib ≥ 2 years: 1.8%; $p < 0.01$ Adjusted risk of recurrence (short- versus long-term) 4.77 (1.98 to 11.48); $p < 0.01$ *Mortality:* Imatinib 6 months to 1 year: 6.9% Imatinib ≥ 2 years: 2.3%; $p < 0.01$ Adjusted risk of mortality (short- versus long-term) 3.44 (1.53 to 7.75); $p < 0.01$
Li et al. 2011[25] Intermediate or high risk (NIH Consensus Criteria)	Intervention: 3-years Imatinib Control: no treatment Follow-up: 45 months (median)	*RFS (high risk versus control):* 1 year: 100% versus 82% 2 years: 97% versus 43% 3 years: 85% versus 31% HR 0.159 (0.066 to 0.381); $p = 0.000$
Nilsson et al. 2010[23] low/intermediate risk and high risk (NR)	Intervention: Imatinib (duration NR) Control: historical Follow-up: NR	*5-year RFS (imatinib versus historic controls):* 85% versus 35%; $p < 0.001$ *5-year OS (imatinib versus historic controls):* Palliative: 55% versus 5%; $p < 0.001$
Li al. 2009[26] Intermediate or high risk of recurrence (NR)	Intervention: 3-years Imatinib (20 months median) Control: no treatment Follow-up: 30 months (median)	*2-year RFS in high risk patients (imatinib versus controls):* 91.5% versus 46.2%; $p < 0.001$ HR 0.107 (0.031 to 0.370); $p < 0.001$
Jiang et al. 2011[27] High risk (NIH modified classification)	Intervention: Imatinib Follow-up: 33.8 months (median) Control: no treatment Follow-up: 44 months (median)	*RFS (imatinib versus surgery only):* Year 1: 100% versus 70.9% Year 2: 88.0% versus 37.8% Year 3: 88.0% versus 27.5% HR 0.122 (0.041 to 0.363); $p = 0.000$

NR, not reported. NB. Table does not include non-RCTs that did not specify the risk group of the patients, as it is unclear if the population in these trials meets the criteria specified in the NICE scope.

Table 13: Results of non-comparative non-RCTs in high risk patients

Study, Risk category (system)	Treatment	Results (95% CI)
Kanda et al. 2013[30] High risk (NIH Consensus Criteria)	Imatinib 48 weeks or confirmation of tumour recurrence Follow-up: 3-years	*RFS*: 1-year: 94.7% (88.9 to 100) 2-year: 71.1% (58.5 to 83.7) 3-year: 57.3% (43.7 to 70.8)
Kang et al. 2013[24] High risk (NIH Consensus Criteria plus c-KIT exon 11 mutations)	Imatinib 2 years unless evidence of disease recurrence or unacceptable toxicity Follow-up: 56.7 months (median)	*RFS (58.9 months median):* 1 year: 97.9% 2 years: 93.6% 3 years: 78.7% 4 years: 62.1% 5 years: 46.0% Median overall survival: Not reached
Nishida et al. 2009[28] High risk (NIH Consensus Criteria)	Imatinib 1-year Follow-up: 109 weeks (median)	3-year RFS: 59% 3-year OS: 87%
Yalcin et al. 2012[29] Intermediate/high risk (NR)	Imatinib 2-years Analysis at follow-up of 1-year	1 year RFS: 0.95 (mean) (0.907 to 0.993)

Summary of results for adverse events

As previously stated, the MS does not report a separate search to identify adverse drug reactions. A brief overview of safety data from all three included RCTs assessing imatinib therapy in the adjuvant setting of GIST is presented (MS Table B-21 page 105). The safety data are based on the full trial populations and not limited to the Miettinen high risk sub-populations. No safety data were reported in the EORTC 62024 conference abstract. NB. A table in the MS presenting adverse events (AEs) across randomised groups (MS Table B-17 page 101) contains no data. This is a generic table in the NICE STA submission template for manufacturers. It is not clear whether the manufacturers intended to add data to the table or to remove the table.

For the ACOSOG Z9001 trial, incidence of AEs by the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) grade 1 - 5 are reported (MS Table B-18 page 102), with a greater incidence of combined grade 3 and 4 AEs in the imatinib group (30.0% (reported as 31% in the MS) vs 18.3% placebo). At the time of the primary outcome analysis, approximately 50% of patients in both groups had completed 1-year of study treatment (MS page 65). Approximately 25% of patients in each trial arm had discontinued therapy (see MS Figure B-3). Details taken from the trial journal publication[12] show that premature

discontinuations due to AEs were statistically significantly higher in the imatinib group compared to the placebo group (16% vs 3%, respectively, $p < 0.0001$).

The most common grade 3/4 events were neutropenia, abdominal pain, dermatitis, nausea and elevated alanine aminotransferase levels. At the 5-year analysis, there were a higher number of withdrawals due to AEs in the imatinib group (1.7%) compared to placebo (0.3%) (MS Figure B-3). There was a slightly higher percentage of deaths in the placebo group than the imatinib group (9.3% vs 7.2%) (MS Figure B-3).

In the SGVXIII/AIO trial, the incidence of any AE was similar for 1-year or 3-year adjuvant imatinib treatment (99% vs 100%). The 3-year adjuvant imatinib therapy profile was described as similar to that of 1-year. However, there were some statistically significantly higher AEs in the 3-year group (see Table 14). Cardiac AEs and diagnosis of secondary cancer were higher in the 1-year group, but no statistical comparisons were reported.

Table 14: Statistically significant differences in most frequently recorded AEs between 1 and 3-years of adjuvant imatinib therapy

SGVXIII/AIO trial **Events**	**Imatinib 1yr** **(n = 194)**	**Imatinib 3yrs** **(n = 198)**	***p*-value**
Haematological			
Leukopenia	67 (34.5)	93 (47.0)	0.01
Non-haematological			
Periorbital oedema	115 (59.3)	147 (74.2)	0.002
Diarrhoea	85 (43.8)	107 (54.0)	0.04
Muscle cramps	60 (30.9)	97 (49.0)	< 0.001
Biochemical			
Elevated blood lactate dehydrogenase	84 (43.3)	119 (60.1)	0.001
Elevated serum creatinine	59 (30.4)	88 (44.4)	0.005
Cardiac AEs	8 (4.1%)	4 (2.0%)	NR
Diagnosis of secondary cancer	14 (7.2%)	13 (6.6%)	NR

NR, Not reported. Data partly copied from MS Table B-19, page 103.

Incidences of any grade 3 or 4 event were statistically significantly higher in the adjuvant imatinib 3-year group compared to the 1-year group (32.8% vs 20.1% respectively; $p = 0.006$),

with the most common reported grade 3 or 4 AEs leukopenia (3.0 vs 2.1%, respectively) and diarrhoea (2.0% vs 0.5%, respectively). Discontinuations were double that for the 3-year imatinib group compared to the 1-year group (25.8% vs 12.9%). This was reflected in higher discontinuations due to AEs in patients treated for 3-years compared to 1-year (13.6% vs 7.7% year 1). No statistical comparisons of discontinuations between the two time periods were reported in either the publication or the MS.

The manufacturer states that five of the non-RCTs provided safety data on the extended use of adjuvant therapy for periods of 3-years. However, only two of these studies are used to illustrate AEs, one in patients at intermediate or high risk of GIST recurrence[27] and one in patients at high risk.[24] No explanations about the reasons for not including the remaining three trials are given and the MS fails to report the appropriate references. After identification of the relevant non-RCTs by the ERG, inspection of the trials suggests that one possible explanation for the exclusion of these trials may be the lack of a comparator arm.[25;28;30]

Summary of Health related quality of life

As stated earlier, none of the three RCTs included in the MS reported HRQoL and neither does this appear to have been reported in the non-RCTs (MS summary table B-16 page 82).

Sub-group analyses results

The NICE scope for the appraisal specifies that, if evidence allows, sub-group analyses by baseline risk of relapse and tumour genetic mutational status should be considered. However, the MS does not report sub-group analyses, either for the full mITT population or the Miettinen high risk sub-population. The SSGXVIII/AIO trial journal publication[16] reports RFS for pre-planned exploratory sub-groups of patients, for the mITT population. HRs are presented according to variables which may be predictive of tumour recurrence such as tumour site, tumour size and tumour mutation site. The results of the sub-group analyses were similar to those of the mITT population, with statistically significant effects for 3-years compared to 1-year of treatment. In the genetic mutational status sub-group there was a statistically significant treatment effect favouring 3-years treatment for patients with the KIT exon 11 mutation ($p < 0.001$), but not for the other mutations (KIT exon 9, wild type or other) or for patients with no mutation. The authors note that the number of patients with these other mutations/ no mutation was smaller than number with the KIT exon 11 mutation. The ERG advises caution in the

interpretation of the sub-group analyses in general due to their exploratory nature, and because some of the groups are likely to be under-powered.

Mixed Treatment Comparison results

A mixed treatment comparison was not reported in the MS. However, an indirect comparison was performed and is discussed in section 4.2.4.

3.4 Summary

The ERG considers that the MS contains a generally unbiased estimate of the treatment effect on the basis of the RCT evidence presented. The RCTs have been well conducted though sanctioned changes in trial end-points did occur during the course of two trials. The main limitation is that treatment effects for high risk patients in the MS are based on retrospective sub-population analyses, varying in the proportion of randomised patients (the lowest being 28%), and are most likely underpowered. Differences between the treatment arms of the trials at baseline in patient characteristics were more pronounced in the Miettinen sub-populations than the full populations, indicating selection bias. Although results for the full population and high risk sub-populations were in agreement (in terms of showing significant differences between trial arms for RFS and OS), caution is necessary in the interpretation of the results for the reasons stated.

All three RCTs included in the MS reported longer RFS associated with adjuvant imatinib treatment, evident for patients irrespective of their risk of recurrence and also for the sub-population of patients classified as at high risk. One year adjuvant imatinib compared to no adjuvant treatment was associated with longer RFS at 5-year follow-up, though the difference between the two lessened over time. Three-year adjuvant imatinib treatment was significantly associated with longer RFS compared to 1-year treatment at 5-year follow-up. Again, the difference between the two lessened during follow-up. Notably there was no significant difference in RFS between 1 and 3-years imatinib treatment during the first year after randomisation (which would be expected) or 3-years after randomisation, but a significant difference emerged during 1 to 2-years and 2 to 3-years after randomisation. The manufacturer suggests this indicates the benefit of prolonged adjuvant treatment. It is not reported whether this was also the case for the Miettinen high risk sub-population, however, as stated earlier, the majority of patients enrolled in the trial were classified as high risk according to the modified US

NIH Consensus. The manufacturer suggests that improvements in RFS associated with adjuvant imatinib therapy could be expected to translate into better HRQoL (MS page 108), but there is no data to support this supposition.

The two RCTs which measured impact of treatment on OS reported differing results. (NB. neither of these two RCTs were statistically powered for OS). In the ACOSOG Z9001 trial there were few deaths overall, and at 2 and 5-years there was no statistically significant difference in OS. The manufacturer suggests that the 5-year data are confounded by the high degree of cross-over to imatinib by recurrence-free placebo patients when the study became unblinded (MS page 108). However, in additional analyses to adjust for patient cross-over in a supplemental report, the difference between trial arms generally remained non-statistically significant. In the SSGXVIII/AIO trial, which was relatively smaller than the ACOSOG Z9001 trial, there were comparatively more deaths and at 5-years there was statistically significantly longer OS associated with 3-year imatinib treatment compared to 1-year. The manufacturer suggests that extending the time that patients remain recurrence-free has a beneficial impact on survival extending to periods of 5-years and longer. Although there may be differences between the two trials in terms of patient characteristics or other variables which may explain the differences in the overall death rates seen, the available evidence suggests that extending imatinib treatment for 3-years is associated with longer overall survival.

4 ECONOMIC EVALUATION

4.1 Overview of manufacturer's economic evaluation

The manufacturer's submission to NICE includes:

iii) a review of published economic evaluations of adjuvant imatinib with surgical resection compared with surgical resection alone for adult patients with GIST.

iv) a report of an economic evaluation undertaken for the NICE STA process. The cost effectiveness of adjuvant imatinib for 1 and 3-years is compared with no treatment for adult patients with surgical resection for GIST.

Manufacturer's review of published economic evaluations

A systematic search of the literature was conducted by the manufacturer to identify economic evaluations of treatments for GIST. The inclusion and exclusion criteria for the systematic

review are listed in Table B-22 of the MS (page 114). The inclusion criteria state that economic evaluations of adult patients with a GIST for adjuvant imatinib with surgical resection compared to surgical resection alone or adjuvant imatinib with surgical resection for a different time period would be included. Abstracts and non-English language studies were excluded.

Nine studies were identified from screening 642 titles and abstracts. Of these seven studies were excluded, mainly as they were not a full paper or were not an economic evaluation. Two studies were included for full review.[38;39]

CEA Methods

The manufacturer's cost effectiveness analysis (CEA) uses a Markov model to estimate the cost-effectiveness of adjuvant treatment with imatinib compared with no treatment in adult patients with GIST treated with surgical resection. The model adopted a lifetime horizon, with a monthly cycle length. Discount rates of 3.5% were applied to both benefits and costs. The model consists of nine health states. Patients can remain recurrence-free, have a recurrent GIST (first or second recurrence), and have progressive disease or die (from GIST or other causes).

The probability of disease recurrence was estimated from clinical-effectiveness data from the published pivotal phase III trials of adjuvant imatinib (ACOSOG Z9001 and SSGXVII/AIO). The treatment effect was estimated for two distinct periods: the period patients received adjuvant imatinib ("on treatment" period) and the period immediately after cessation of adjuvant imatinib ("off-treatment" period).

Quality of life is captured by utility values, which are assigned for patients in different phases of disease according to health state. Utility weights used in the economic model were identified through a systematic review of the literature, as no quality of life data were collected from the clinical trials.

The current UK guidelines[5] were used to determine the frequency of visits and tests associated with the different disease states. Costs associated with outpatient attendance, CT scans, blood counts, liver function tests, surgery for recurrence and adverse events (AEs) requiring hospitalisation were estimated from the NHS reference costs. The frequency and types of AEs included were based on the most frequently reported grade 3/4 AEs in the SSGXVII/AIO trial.

CEA Results

The results from the economic evaluation are presented for 1-year and 3-year adjuvant treatment (Table B-54, MS section 7.7.6, page 220) as incremental cost per Quality Adjusted Life Year (QALY) gained for adjuvant imatinib for 1-year and 3-years compared with no treatment.

For the base case an incremental cost per QALY gained of £3509 is reported for adjuvant imatinib 1-year treatment compared with no treatment (see Table 15). The ICER for 3-year adjuvant imatinib compared with 1-year adjuvant imatinib was £16,006 per QALY. The results were most sensitive to changes to the hazard ratio for treatment effect and the time horizon of the model.

Table 15: Base case cost effectiveness results

	Total Per Patient:		Incremental:		incremental analysis	Compared with no treatment
	Costs	QALYs	Costs	QALYs	ICER (Cost/QALY Gained)	
No treatment	£47,292	3.83				
Adjuvant imatinib 1yr treatment	£55,136	6.07	£7,844	2.24	£3,509	£3,509
Adjuvant imatinib 3yrs treatment	£78,068	7.50	£22,931	1.43	£16,006	£8,390

The MS summarises the results of the probabilistic sensitivity analyses (PSA) stating that the likelihood of 1-year and 3-years of imatinib treatment being cost effective at a threshold of £20,000 per QALY is 41.7% and 58.3% respectively, and at a threshold of £30,000 per QALY is 30.0% and 69.1% respectively.

The MS states that adjuvant imatinib given for 1-year or 3-years in patients with GIST at high risk of recurrence therefore represents an efficient use of NHS resources.

4.2 Critical appraisal of the manufacturer's submitted economic evaluation

Manufacturer's review of published economic evaluations

The manufacturer completed a review of economic evaluations of treatments for GIST. The inclusion and exclusion criteria for the systematic review are listed in Table B-22 of the MS (page 114). Abstracts and non-English language studies were excluded. Nine studies were

identified from screening 642 titles and abstracts. Of these seven studies were excluded, mainly as they were not a full paper or were not an economic evaluation. Two studies were included for full review.[38;39] The MS provides a tabulated summary of these studies and a quality assessment checklist.[40]

The two identified studies were both for patients in the SSGXVIII/AIO trial and were from the perspective of the Netherlands[38] and USA.[39] Both used Markov state-transition models. The MS notes that these studies differed in the method used for the extrapolation of RFS beyond the observed period, post recurrence mortality, resource costs, and the sources used for utility weights. The ICER for 3-years of adjuvant imatinib compared with 1-year of adjuvant imatinib was €29,872 per QALY in Majer and colleagues (2013)[38] and $62,600 per QALY in Sanon and colleagues (2013).[39] Both economic evaluations received funding from Novartis.

Critical appraisal of manufacturer's submitted economic evaluation

The ERG have considered the methods applied in the economic evaluation in the context of the critical appraisal questions listed in Table 16 below, drawn from common checklists for economic evaluation methods (e.g. Drummond and colleagues[40]).

Table 16: Critical appraisal checklist of economic evaluation

Item	Critical Appraisal	Reviewer Comment
Is there a well defined question?	Y	
Is there a clear description of alternatives?	Y	Surgery alone or surgery plus adjuvant imatinib for two different treatment durations (one year or three years)
Has the correct patient group / population of interest been clearly stated?	Y	Submission is based on patients at high risk of recurrence (using Miettinen criteria) while the license refers to significant risk (scope refers to "imatinib within its licensed indication for the adjuvant treatment of gastrointestinal stromal tumours").
Is the correct comparator used?	Y	
Is the study type reasonable?	Y	
Is the perspective of the analysis clearly stated?	Y	
Is the perspective employed appropriate?	Y	
Is effectiveness of the intervention established?	Y	Trial-based evidence for 1-year adjuvant treatment compared with placebo (surgery with no adjuvant therapy), based on sub-group analysis. Indirect comparison required for 3-years vs no adjuvant therapy (also based on sub-group analysis of trial 3-year adjuvant imatinib compared with 1-year adjuvant imatinib)

Has a lifetime horizon been used for analysis (has a shorter horizon been justified)?	Y	Base case model runs for 600 monthly cycles (50 years) from a mean starting age of 61. It would be more appropriate to terminate at 100 years.
Are the costs and consequences consistent with the perspective employed?	Y	
Is differential timing considered?	Y	
Is incremental analysis performed?	Y	
Is sensitivity analysis undertaken and presented clearly?	Y	

NICE reference case

The NICE reference case requirements have also been considered for critical appraisal of the submitted economic evaluation in Table 17.

Table 17: NICE reference case requirements

NICE reference case requirements:	Included in submission	Comment
Decision problem: As per the scope developed by NICE	Y	
Comparator: Alternative therapies routinely used in the UK NHS	Y	
Perspective on costs: NHS and PSS	Y	
Perspective on outcomes: All health effects on individuals	Y	
Type of economic evaluation: Cost effectiveness analysis	Y	
Synthesis of evidence on outcomes: Based on a systematic review	Y	No systematic searches reported for additional clinical effectiveness parameters in the model
Measure of health benefits: QALYs	Y	Health/ treatment state utility values derived from studies unrelated to clinical trials providing clinical effectiveness evidence
Description of health states for QALY calculations: Use of a standardised and validated generic instrument	Y	Use EQ-5D based valuations.
Method of preference elicitation for health state values: Choice based method (e.g. TTO, SG, not rating scale)	?	Method of valuation for health states in one study is not clear
Source of preference data: Representative sample of the public	?	Uncertainty over valuation set for EQ-5D from one study.
Discount rate: 3.5% p.a. for costs and health effects	Y	

N/A=not applicable otherwise use yes or no. P.a., per annum. Notes: ? = uncertain;. Only no, ? or N/A need qualification in the comments column.

4.2.1 Modelling approach / Model Structure

The MS economic model consists of a multi-state Markov model with nine health states: no recurrence and no treatment; no recurrence and imatinib adjuvant therapy; post recurrence and 400mg imatinib; no recurrence and completed imatinib adjuvant therapy; post recurrence and sunitinib; sunitinib second-line therapy; best supportive care (BSC); death from GIST; death from other cause (Figure 4). The model's cycle length is 1 month. Costs and QALYs were calculated over a lifetime horizon (50 years) and discounted at 3.5% per annum. The analyses are conducted from the perspective of the NHS in England and Wales.

The model structure and the possible transitions between health states are shown in **Error! eference source not found.** (MS Figure B-10 page 129). Patients on adjuvant imatinib start in the no recurrence and on adjuvant imatinib therapy health state (state B), and those on no treatment start in the no recurrence and no treatment health state (state A). Patients on adjuvant imatinib may discontinue treatment due to AEs or through disease recurrence. Patients who complete treatment move to the no recurrence and completed adjuvant imatinib therapy health state and remain in this health state until disease progression (state C). These patients remain in this health state until death or second disease recurrence, when they will either be treated with sunitinib (state G) or BSC (state H).

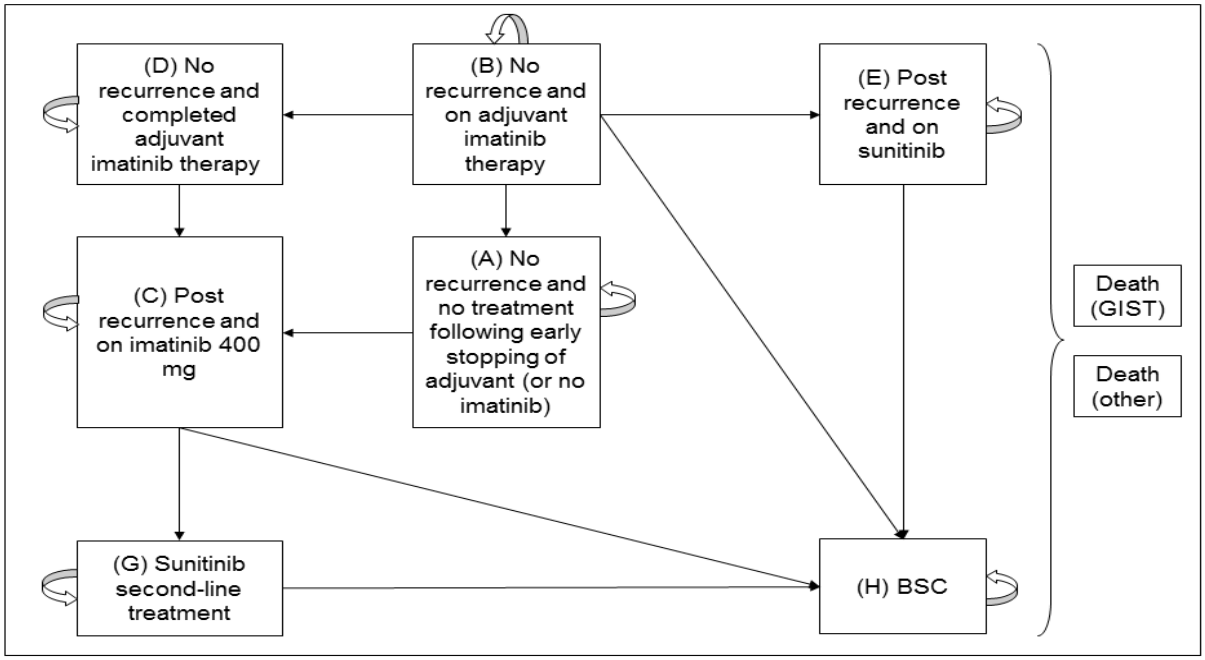


Figure 4: MS Model structure (MS Figure B-10 page 129)

The two terminating states in the model are death from GIST and death from any other cause. Death from GIST is from any post-recurrence health state and patients may die from non-GIST related mortality from any health state.

The model structure is based upon the structure used in the previous submission to NICE for TA196.[1] The MS states that the model structure and treatment pathways were informed by discussion with five UK clinicians. It states that a Markov modelling approach is appropriate "in order to allow modelling of disease progression over time" (MS Table B-32). The MS does not discuss alternative model structures that may have been used, for example survival models. The current model structure differs from the previous submission in that there is no longer a post recurrence health state for dose escalation to imatinib 800 mg (as this treatment was not approved in NICE guidance TA86 for patients with unresectable and/or metastatic GIST[41]). In addition, there is the possibility to move from the post recurrence and imatinib therapy health state to best supportive care. This was added in response to comments in the previous appraisal. The ERG considers that the MS clinical pathway and model structure to be appropriate and relevant to UK clinical practice.

In the MS economic model, following a first recurrence, patients move to the health state C and receive imatinib first line treatment. The MS states that about 15% of these patients would receive further surgery. This is not modelled explicitly in the model but the cost is included. The ERG is unclear whether this group has been fully captured and whether they would have additional health benefits that have not been included. The ERG asked the manufacturer for clarification on this issue (see clarification request B2). The manufacturer clarified that these patients incurred the costs of surgery and the costs of receiving imatinib and experience outcomes (in terms of AEs, probability of recurrence and quality of life) estimated for imatinib-treated patients.

The manufacturer assumes that the treatment effect for the 'off treatment period' remains constant. The ERG asked for justification for this assumption. The manufacturer stated that this was based on the clinical data from the pivotal phase III trials where results for the SSGXVII/AIO trial are reported for a median duration of follow-up of 54 months and therefore offer robust data for 5 to 6 years. The ERG asked the manufacturer to provide scenarios to demonstrate the effect of varying this assumption (see section 4.3).

The manufacturer assumes a long-term survival benefit for imatinib and that the post-recurrence GIST mortality is equal across all treatment arms. The ERG asked for justification of this assumption. The manufacturer stated that this assumption was adopted as there is no evidence to suggest this is not the case. The ERG asked the manufacturer to provide a scenario varying this assumption (see section 4.3).

4.2.2 Patient Group

The patient group in the economic model are described and discussed in section 7.2.1 (pages 127 - 128) of the MS and are defined as adult patients at high risk of relapse following resection of KIT (CD117)-positive GIST. Risk of recurrence in this patient group is defined on the basis of the Miettinen criteria (discussed in section 2.2 of this report). This definition of the patient group does not exactly match the imatinib license, which refers to “patients at significant risk”[42] (acknowledged by the manufacturer, see page 127 of MS), nor does it exactly match the entry criteria for the clinical trials used to derive the clinical effectiveness evidence used in the model.

As stated earlier in this report, the eligibility criteria for ACOSOG Z9001 trial specified that patients could be at any risk of recurrence. The MS reports that sub-group analyses were conducted for patients retrospectively classified as at high risk of recurrence using the Miettinen criteria – using additional (unspecified) data. The ERG requested clarification regarding the additional data (see clarification request B8). The clarification response indicated that risk stratification in the trial was based on tumour size. However, since tumour specimens were collected in the trial, a retrospective analysis of mitotic count was conducted (in response to questions from European regulatory authorities regarding risk classification) and combined with available data on tumour size and location to identify the sub-group at high risk of recurrence using the Miettinen criteria. An additional report submitted by the manufacturer in response to the ERG clarification request indicates that mitotic index data were available for 78% (556/713) of cases in the ACOSOG Z9001 ITT population.

The eligibility criteria for SSGXVIII/AIO trial were based on the modified US NIH Consensus Criteria[43;44] (see Table B-6, page 46 of MS) used to identify patients at high risk of recurrence. These criteria are not the same as those originally used for risk stratification in the ACOSOG Z9001 trial or the Miettinen criteria. Section 6.3 of the MS reports that sufficient data were collected in the trial to classify patients at high risk of recurrence by the Miettinen criteria (see page 45 of the MS) – baseline data, recurrence-free survival and overall survival data are

reported for both the full trial population and the Miettinen high-risk sub-group in section 6.5 of the MS (pages 71 – 76). The discrepancy between the SSGXVIII/AIO trial inclusion criteria and the Miettinen high-risk classification is not discussed in the Cost-effectiveness / modelling sections of the MS describing the derivation of data to populate the model (section 7.3 - clinical parameters and variables). As a result it is not entirely clear whether the clinical effectiveness parameters in the model, derived from the SSGXVIII/AIO trial, are based on the full trial population or only the sub-group classified as at high risk of recurrence by the Miettinen criteria.

No further sub-groups were considered. In particular the MS presents no evidence presented for patients who might be considered at “significant risk”, but not at high risk.

4.2.3 Interventions and comparators

The comparator in the economic model is specified as “No treatment” and is modelled (for primary recurrence following resection) using data from the placebo arm of the ACOSOG Z9001 (note all patients in the ACOSOG Z9001 underwent surgical resection, with patients randomised to the control arm receiving placebo adjuvant therapy). This comparator corresponds well to the scoped definition of “observation after surgery (no adjuvant therapy)” – although patients in both arms of the ACOSOG Z9001 had more intensive follow-up than would be indicated as normal practice (11 evaluation visits in the first twelve months after surgery (nine in first six months) including liver function, creatinine tests and full blood count, and three-monthly evaluation visits the following year).

The model regards 3-years of adjuvant imatinib as the intervention, modelled (for primary recurrence following resection) using data from the 3-year treatment arm of the SSGXVIII/AIO trial and an indirect comparison with the ACOSOG Z9001 trial. However the analysis also includes 1-year of adjuvant imatinib (modelled using data from the ACOSOG Z9001) – the base case results present a fully incremental analysis across all three treatment strategies. This is consistent with the scope for this appraisal, which does not specify a duration of adjuvant imatinib treatment.

4.2.4 Clinical Effectiveness

The main clinical effectiveness parameters included in the model are:

- **Recurrence-free survival/ risk of recurrence with or without adjuvant imatinib** based on data from ACOSOG Z9001 and SSGXVIII/AIO trials (described and discussed on pages 139 - 152 of MS) – which involved estimation of three main sets of effectiveness parameters
    - baseline (MS page 139 - 142)
    - on-treatment (MS page 145 - 149)
    - off-treatment (MS page 149 - 152).
- **Probability of discontinuation in patients receiving adjuvant imatinib** based on data from SSGXVIII/AIO trial (described and discussed on MS page 155)
- **Recurrence and probability of discontinuation in patients receiving non-adjuvant imatinib** based on a published trial report[45] (described and discussed on MS page 155)
- **Recurrence and probability of discontinuation in patients receiving sunitinib second line** based on a published trial report[46] (described and discussed on MS page 154)
- **Mortality in patients receiving imatinib** (first line – i.e. non-adjuvant) based on a published trial report[45] (described and discussed on MS page 154)
- **Mortality in patients receiving sunitinib** (second line) based on a published trial report[46] (described and discussed on MS page 155)
- **Mortality in patients receiving BSC** based on a trial report [47], an epidemiological study [48] and a post-hoc sub-group analysis used in a decision model [49] (described and discussed on MS pages 153-154)
- **Mortality from other causes** (MS page 155)

The key clinical effectiveness parameter included in the model is the risk of primary recurrence following surgery (with or without adjuvant imatinib therapy) which was based on data from the ACOSOG Z9001 and SSGXVIII/AIO trials. The clinical trial data used to estimate the baseline risk of recurrence and the effectiveness of adjuvant therapy are presented and discussed in section 6 (Clinical evidence) of the MS and are critically appraised in section 3.1 of this report. However, the methods for deriving the baseline risk of recurrence (in patients at high risk of recurrence) and hazard ratios for 1-year and 3-year adjuvant imatinib to be applied to the estimated baseline hazard function are only presented in section 7 (Cost effectiveness) of the MS and are appraised below.

The later transitions in the model (i.e. for patient treatment outside of the adjuvant setting) model inputs are based on data from trials other than those reviewed in section 6 (clinical

evidence) of the MS. No searches are reported for data to populate the model and no critical appraisal of studies used to populate the model is presented in the MS. These are discussed in turn below.

Additional assumptions are included in the model regarding the proportion of patients likely to undergo additional surgery following recurrence and proportion able to receive sunitinib following recurrence. These assumptions are discussed later in this section.

Recurrence-free survival/ risk of recurrence for patients undergoing surgery without adjuvant imatinib - baseline

The probability of recurrence in patients treated with surgical resection only was based on data from the placebo arm of the ACOSOG Z9001. Additional analyses, not reported or discussed in the clinical effectiveness section of the MS (MS section 6.5.3 pages 65 - 70 of the MS for presentation of results from the ACOSOG Z9001 trial) were undertaken to populate the economic model. The MS states that data for patients in the placebo arm classified as at high risk of recurrence, by the Miettinen criteria, in the 5-year follow up analysis (n=98) were used to derive the baseline risk of progression in this patient group. However, only data for these patients prior to unblinding (i.e. before they were eligible to cross-over) was used in the analysis. The MS does not state clearly the maximum follow-up for placebo patients prior to cross-over (censoring) so the ERG cannot judge the duration over which the baseline survival function was modelled.

Five different parametric survival models (exponential, Weibull, gamma, log-logistic and Gompertz) were fitted to the recurrence-free Kaplan-Meier data, using maximum-likelihood methods in R. The MS reports that goodness of fit was assessed visually and using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) statistics, with the face validity of extrapolations beyond the trial data assessed using published survival data and clinical judgement. There was very little difference in the goodness of fit statistics for the different parametric survival models (MS Table B-28 page 141) with loglogistic distribution fitting best.

The long-term survival extrapolations varied substantially for different parametric functions (demonstrated clearly in Figure B-12 on page 143 of the MS) and judgements between functions was primarily made on the basis of the validity of these extrapolations at 5-years (not

far beyond the maximum follow-up in the dataset which appears to be around 48 months for the placebo arm). The MS reports that 5-year survival based on the exponential, Weibull and gamma models was around 5-6% whereas 5-year survival in the log-logistic and Gompertz models was between 10% and 15%. The MS uses data from a single study reporting long term outcomes for patients classified as at high risk of recurrence, using the Miettinen criteria, who did not receive adjuvant therapy[50] to suggest that (depending on tumour size) 10% to 26% of high risk patients may remain disease-free at periods of follow up greater than 10 years. These data are used to justify the rejection of the exponential, Weibull and gamma models – this seems reasonable given the limited supporting data reported. However the MS then states that the Gompertz model will be used for the base case analysis without any further discussion or comparison with the loglogistic model.

Recurrence-free survival/ risk of recurrence for patients undergoing surgery with adjuvant imatinib

Having derived the probability of recurrence in patients treated with surgical resection only, from the placebo arm of the ACOSOG Z9001 trial, the MS discusses possible approaches to modelling recurrence in patients treated with adjuvant imatinib in addition to surgical resection. The MS rejects what they term the traditional approach – deriving a single treatment effect for 1-year of adjuvant imatinib using the active treatment arm in the ACOSOG Z9001 trial and a single treatment effect for 3-years of adjuvant imatinib based on an adjusted indirect comparison using data from both the ACOSOG Z9001 trial and the SSGXVIII/AIO – in favour of an approach involving the estimation of two treatment effects – an on-treatment effect and a post-treatment effect. The adoption of this approach is supported by reference to a plot showing the Kaplan-Meier curves for each arm in both trials (Figure B-14, page 146 of the MS). This is difficult to interpret (due to showing four Kaplan-Meier curves with 95% confidence intervals for each curve super-imposed). However it does appear to show changes in the shape of the survival curves for the adjuvant treatment arms in the trials shortly after end of adjuvant treatment. These trends may have been more apparent, and the cut–points more easily identified and justified, by plotting the hazard function rather than the Kaplan-Meier curves.

Recurrence-free survival/ risk of recurrence for patients undergoing surgery with adjuvant imatinib - on-treatment hazard ratios

The on-treatment recurrence-free hazard ratio was estimated using a direct comparison between the placebo and 1-year adjuvant imatinib arms in the ACOSOG Z9001 trial, using the

Cox proportional hazards model with the data truncated at 12 months, giving a value of 0.111 (95% CI, 0.043 to 0.281). The proportional hazards assumption was tested by visual inspection of the complementary log-log plot of the Kaplan-Meier curves (MS Figure B-15 page 148), which the MS argues were “roughly parallel”. The ERG notes that there do not appear to be gross deviations. However the curves are closer together on the left side of the chart. The interpretation of this figure is confused by the super-imposition of 95% confidence intervals for each curve. The ERG requested clarification regarding this hazard ratio (see clarification request B8), which is lower than the hazard ratio of 0.265 (95% CI 0.148 TO 0.477) reported in Table B-13 (page 67 of the MS) for the Miettinen high risk population. The ERG requested an indication of whether the difference between the two results was primarily related to truncating the data at 12 months or due to the retrospective re-classification of additional high risk patients. The manufacturer’s response does not appear to address this question.

In the absence of a direct comparison (between 3-year of adjuvant imatinib and placebo) the MS assumed that the hazard ratio observed for 1-year adjuvant imatinib compared with placebo can also be applied for each year of the 3-year of adjuvant imatinib strategy, justifying this by reference to the complementary log-log plot of the Kaplan-Meier curves for the four trial arms (placebo and 1-year adjuvant imatinib arms in the ACOSOG Z9001 trial, 1-year and 3-year adjuvant imatinib arms from the SSGXVIII/AIO trial) arguing that this does not indicate any time dependence in treatment effect.

The ERG were concerned that hazard ratios derived using Cox proportional hazards models were to be applied in the model to a range of parametric survival functions and requested clarification from the manufacturer and a rationale for not using hazard ratios derived using the same parametric survival functions as those used in the model (see clarification request B12). The manufacturer responded by stating that “using curves from different trials is likely to introduce bias”. This doesn’t appear to answer the ERG’s concern over the appropriateness of combining hazard ratios derived using a semi-parametric model (Cox) with fully parametric survival functions. The ERG’s concern was not that the model should use curves from different trials, but that the treatment effects should be derived using methods that are consistent with the estimation of the baseline recurrence-free survival function.

Recurrence-free survival/ risk of recurrence for patients undergoing surgery with adjuvant imatinib - off-treatment hazard ratios

Estimation of the off-treatment hazard ratios for 1-year and 3-year adjuvant imatinib required further manipulation of the trial datasets. The first stage (described in the MS as "estimating the Kaplan-Meier curves for the post-treatment phase") involved removing all patients who experienced recurrence or were censored during their planned duration of adjuvant imatinib in both trials. This provided a dataset for, post-treatment, recurrence-free survival for patients who were recurrence-free at the end of their planned duration of adjuvant treatment. The data were not truncated for patients who underwent surgical resection with no adjuvant treatment.

Off-treatment hazard ratios were estimated using these new, derived datasets and are defined in the MS as HR1 (1-year adjuvant imatinib vs placebo) and HR2 (3-year adjuvant imatinib vs 1-year adjuvant imatinib). The values estimated for HR1 and HR2 were 0.519 (95% CI 0.297 to 0.906) and 0.633 (95% CI 0.392 to 1.123).

The off-treatment hazard ratio for 3-year adjuvant imatinib vs placebo (HR3) was calculated using a standard adjusted-indirect comparison method (see below), although the MS contained a documentation error regarding the calculation of the adjusted-indirect comparison. The ERG requested clarification regarding this and were informed that, due to the need to change the reference category for HR2 (to be 1-year adjuvant imatinib vs 3-year adjuvant imatinib) the inverse (1/0.519) was used in the calculation (see clarification request B5).

$$\ln(HR3) = \ln(HR2) - \ln(HR1)$$

$$SEln(HR3) = SEln(HR1)^2 + SEln(HR2)^2$$

where ln() natural logarithm and SEln() indicates the standard error of the natural log.

Using this corrected calculation the hazard ratio for 3-year adjuvant imatinib vs placebo (HR3) gives a value of 0.344 (95% CI 0.160 to 0.741).

Recurrence-free survival/ risk of recurrence for patients undergoing surgery with & without adjuvant imatinib – summary and estimated survival curves

Figure 5 shows the Kaplan-Meier curves and fitted survival functions (using the Gompertz function) for surgery with no adjuvant therapy (ACOSOG Z9001 placebo arm), surgery with 1-year adjuvant imatinib (ACOSOG Z9001 imatinib 12 months arm) and surgery with 3-years of adjuvant imatinib (SSGXVIII/AIO imatinib 36 months arm) derived using the methods described

Superseded - see erratum

above. The implication of the survival extrapolations – and the choice of Gompertz survival function - are shown clearly in this figure with recurrence-free survival probabilities at 9 years of approximately 40% for patients receiving 3-years of adjuvant imatinib, approximately 30% for patients receiving 1-year of adjuvant imatinib and a little less than 10% for patients having surgery only.

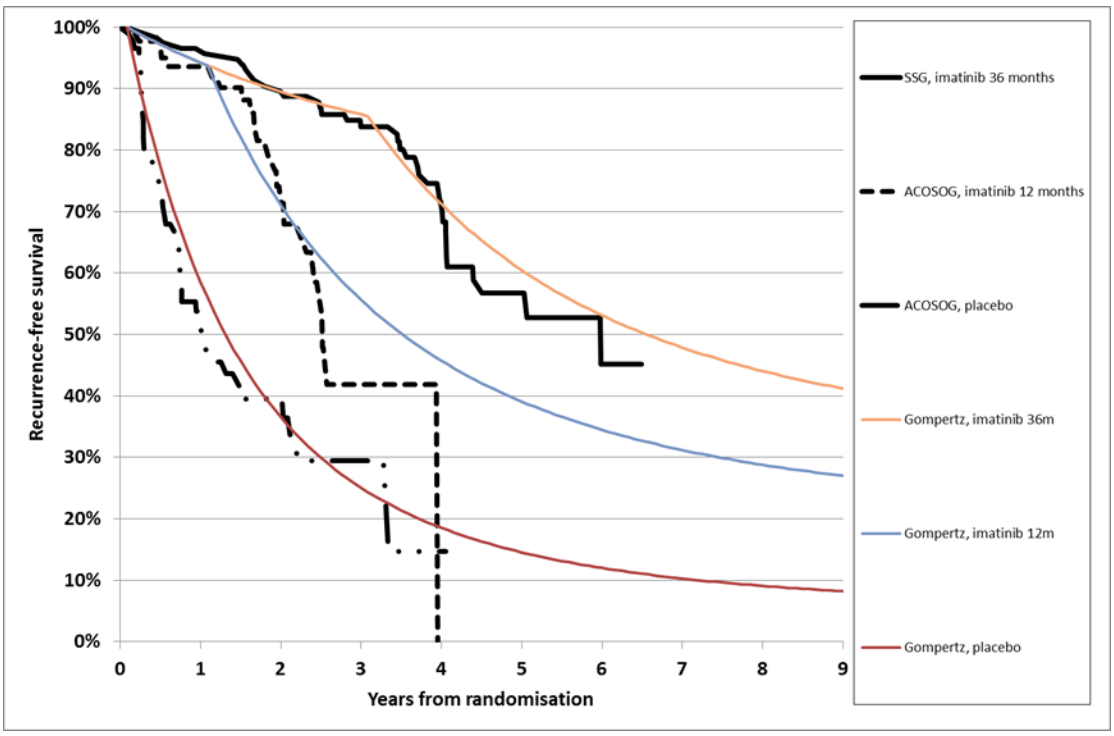


Figure 5: Survival functions applied in the economic model

The ERG are concerned about the face validity of these survival extrapolations based on the Gompertz function. Figure 6, derived by the ERG, shows the Kaplan-Meier plot for "observed" recurrence-free survival in the 3-year adjuvant imatinib arm of the SSGXVIII/AIO trial, with the five candidate survival extrapolations over a duration of 20 years (approximately half the duration extrapolated in the economic model). The curve derived using the Gompertz function is levelling off suggesting a long term maintenance of RFS in around 30% of patients undergoing surgery for GIST and receiving 3-years of adjuvant imatinib. The ERG is concerned this may not be appropriate in a population initially identified as being at high risk of recurrence. This compares with approximately 20% recurrence-free survival at 20 years, using the loglogistic model or approximately 5% using the other candidate functions.

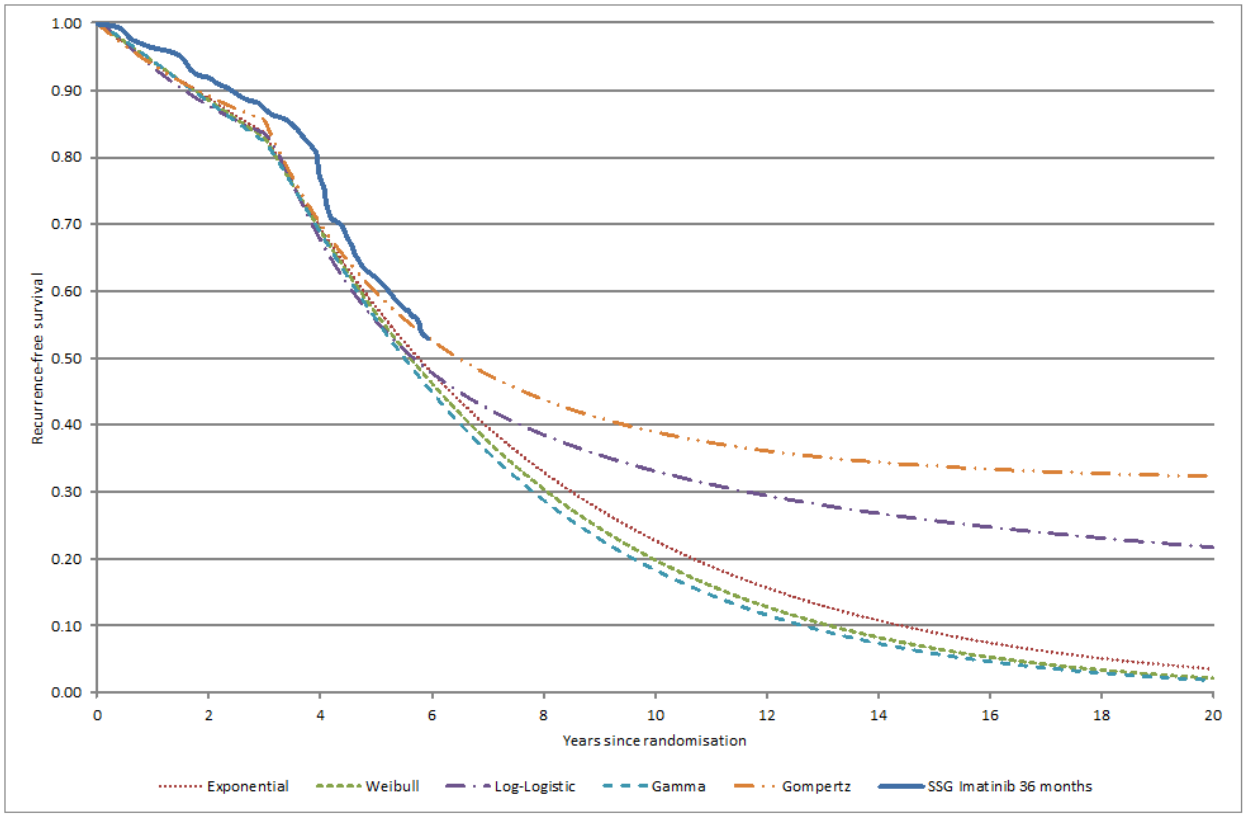


Figure 6: Observed recurrence-free survival (3-years adjuvant imatinib) versus alternative parametric survival extrapolations

The MS stated that a crossover analysis for the ACOSOG Z9001 five year update analysis was being undertaken at the same time as the submission, but did not include any findings since final results were not expected until February 2014. The ERG requested clarification regarding the scope of the analysis and whether this would be used to update the economic model accompanying the submission. This clarification request was amended by NICE to indicate their procedures regarding the submission of further information (see clarification request C1). The manufacturer confirmed that the crossover analysis was complete – this has been forwarded to the ERG (as discussed earlier in section 3.1.6 of this report). The manufacturer also included an additional analysis in their response to this request, presenting the HRs applied in the base case in the MS and an equivalent set of HRs derived from the five year update data.

The ERG are concerned that there is potential for confusion over terms used in the MS and responses to clarification regarding the data used in analyses undertaken to populate the economic model. The MS uses the term “primary analysis” to refer to analyses undertaken in the ACOSOG Z9001 trial using data prior to unblinding (12th April 2007). However this term is

used to refer to analysis of three populations – the full trial population and two overlapping groups of Miettinen high-risk patients (which are a sub-group of the full trial population) – see Table 18.

Table 18: Analysis points for ACOSOG Z9001 trial and populations included at each analysis

Term in MS/ clarification	Population	Source
"Primary analysis"	Full trial population (Imatinib n=359; Placebo n = 354) Miettinen high-risk (Imatinib n=84; Placebo n = 81) Miettinen high-risk (Imatinib n=103; Placebo n = 98)	DeMatteo and colleagues[12] MS section 6.5.3 Table B-13 [a] MS section 7.3.2 [b]
Five year update	Full trial population (Imatinib n=359; Placebo n = 354) Miettinen high-risk (Imatinib n=103; Placebo n = 98)	MS section 6.5.3 Table B-14 MS section 6.5.3 Table B-14 [c]

[a] this analysis is conducted at the same time point as the original primary analysis (i.e. prior to unblinding) and should therefore be unaffected by cross-over in the placebo arm

[b] this analysis is conducted at the same time point as the original primary analysis (i.e. prior to unblinding) and the retrospective sub-group analysis of Miettinen high-risk patients (a) but includes additional cases which were identified at the five update.

[c] this analysis is conducted after unblinding and therefore would be affected by cross-over in the placebo arm

The base case analysis in the economic model uses clinical parameters derived using the dataset identified by superscript b in Table 18. The manufacturer's response to clarification request C1 includes HRs estimated using the dataset identified by superscript c in Table 18, which includes placebo patients who crossed-over to adjuvant imatinib at study unblinding. This would be expected to result in a less favourable off-treatment recurrence-free survival HR for adjuvant imatinib (see Table 19 reproduced from the manufacturer's clarification response). As a result the cost effectiveness results for adjuvant imatinib reported in the clarification request using the five year update data are less favourable than in the base case.

Table 19: Hazard ratios for recurrence-free survival with adjuvant imatinib in base case analysis & five year update (response to clarification request C1)

	Current base case	5 year update (unadjusted)
HR on treatment	0.111	0.112
HR off treatment (1-year adjuvant imatinib)	0.519	0.727
HR on treatment (3-year adjuvant imatinib)	0.344	0.482

The manufacturer has not presented any analyses at the 5-year time point that account for cross-over, although the overall HR for recurrence (combining on- and off-treatment effects) is more favourable [the HR adjusted for cross-over is 0.5 (95% CI 0.32 to 0.78) compared with the unadjusted estimate of 0.61 (95% CI 0.42 to 0.89)]. The ERG is unable to provide additional cost effectiveness analyses based on the cross-over adjusted estimates as only the combined (on- and off-treatment effects) HR is reported in the cross-over analysis.

Probability of discontinuation in patients receiving adjuvant imatinib

The probability of discontinuation for patients receiving adjuvant imatinib was based on data from the SSGXVIII/AIO trial, adopting a different rate for the first six months compared with the remaining planned duration of treatment. The ERG requested clarification on why only data from the SSGXVIII/AIO trial were used (see clarification request B11) – the manufacturer's clarification response states that this was a pragmatic decision made for the sake of simplicity. The ERG requested clarification on the basis for assuming different discontinuation rates in the first six months of treatment (see clarification request B13). The MS states that this decision was entirely data driven. The ERG also requested clarification on why different discontinuation rates, due to adverse events, were applied for patients receiving 1-year adjuvant imatinib and patients receiving 3-year adjuvant imatinib, as this seemed inconsistent with other assumptions in the model and with statements made elsewhere in the MS (see clarification request B10). The manufacturer accepted that the approach taken in the submission was inconsistent and provided additional analyses applying consistent assumptions for 1-year and 3-year adjuvant imatinib – these had limited impact on the cost effectiveness results.

The model assumes that all patients who discontinue adjuvant imatinib due to adverse events will be eligible for, and accept, non-adjuvant imatinib following recurrence. Expert clinical advice to the ERG stated that the logic of this model structure is reasonable - adverse events that

patients deem unacceptable in the adjuvant setting maybe more acceptable when offered the same treatment for disease recurrence. However it is not certain that all patients will accept the same treatment.

Recurrence and probability of discontinuation in patients receiving non-adjuvant imatinib

The probability of discontinuation due to recurrence in patients receiving imatinib for primary recurrence following surgery is based on the proportion of treatment failures due to disease progression reported by Verweij and colleagues.[45] This trial was conducted between February 2001 and February 2002 in 13 countries with eligibility criteria reported in Table 20. Patients in the trial were randomised (473 in each arm) to receive the standard dose of imatinib (400 mg) either once or twice daily. Verweij and colleagues[45] report that, at 2-years of follow up, 56% of patients randomised to receive 400 mg imatinib once daily experienced treatment failure (consisting of 53% progression and 3% being deaths by any other cause). The equivalent figures for patients randomised to receive 400 mg imatinib twice daily were 48%, with 44% due to progression and 4% deaths from other causes.

The MS states that the monthly probability of discontinuation due to progression is based on the value for imatinib once daily. The calculation is based on a reported value of 44% (which does not agree with the value for imatinib once daily reported by Verweij and colleagues) and yields an estimated value of 0.034 (monthly rate (mr) = -1/2*ln(0.44), converted to a monthly transition probability as 1-exp(-mr)). This calculation over-estimates the progression probability and would lead to a modelled discontinuation proportion (in the absence of other transitions) of 56%. Calculating the monthly rate correctly for this proportion (-1/2*ln(1-0.44) = 0.024) and converting this to monthly transition probability gives an estimate of 0.024. However, as noted above the progression proportion (44%) used in the MS appears to relate to imatinib twice daily. Using the proportion reported Verweij and colleagues[45] for imatinib once daily gives a monthly transition probability of 0.031 (1-exp(-(1/2*ln(1-0.53))). Given the small difference between this final value and that applied in the original submission this is likely to have minimal impact on the cost effectiveness results.

Superseded - see erratum

Table 20: discontinuation of non-adjuvant imatinib, due to disease progression and adverse events in the manufacturer's model

Source	Study population	Model Input Parameter	Extracted data	Transition probability
Verweij et al[45]	Adults with histologically proven advanced or metastatic GIST characterised by cKIT expression, WHO PS ≤ 4	Discontinuation due to progression (Imatinib 400 mg once daily)	Proportion with progression = 44% at 2 years	0.034[a] (0.024)[b]
			Proportion with progression = 53% at 2 years	0.031
		Discontinuation due to AE	Proportion discontinue due to toxic effects = 7%	0.0029[c]

[a] Transition probability reported in the MS and used in the model. [b] Transition probability estimated by ERG. [c] Transition probability reported in MS and used in model (based on 760 days median follow-up)

The monthly probability of discontinuing non-adjuvant imatinib was based on the proportion discontinuing due to toxic effects in the trial reported by Verweij and colleagues.[45] The trial report did not provide any information on when the discontinuations occurred – as a result the MS estimated the probability based on the median duration of follow-up for the trial.

All patients discontinuing non-adjuvant imatinib are treated in the model as experiencing progression and are therefore eligible for second-line treatment with sunitinib (although a proportion (10%) are considered unsuitable for sunitinib and progress to BSC). This assumption is inappropriate for those patients discontinuing non-adjuvant imatinib due to adverse effects (as they are not modelled as having experienced progression), but arises from the model being structured using treatment states rather than health states. Given the relatively low transition probability this is unlikely to have a substantial impact on the model results.

Recurrence and probability of discontinuation in patients receiving sunitinib second line

The probability of discontinuation due to recurrence in patients treated with sunitinib is based on the median progression-free survival reported by Demetri and colleagues (2006).[46] This trial was conducted between December 2003 and January 2005 in multiple centres (56) in 11 countries using eligibility criteria reported in Table 21. Demetri and colleagues report a median progression-free survival of 24.1 weeks (95% CI 11.1 – 28.3) for sunitinib. For use in the model the progression-free survival duration is converted to a monthly rate, by first calculating the survival duration in months ((24.1/52)*12 = 5.56), assuming an exponential survival function to

derive a monthly rate (monthly rate = -1/5.56*ln(0.5)) and converting this to a transition probability as above.

Table 21: Discontinuation of sunitinib, due to disease progression and adverse events in the manufacturer's model

Source	Study population	Model Input Parameter	Extracted data	Transition probability
Demetri et al.[46]	Adults with histologically proven malignant GIST not amenable to surgery/radiation/combination therapy with curative intent, confirmed failure of prior imatinib therapy, ECOG PS ≤ 1 and adequate hepatic/ renal/ cardiac function	Discontinuation due to progression	Median progression-free survival = 24.1 weeks	0.117
		Discontinuation due to adverse events	Discontinuation due to adverse events = 9% Median number of days on treatment = 56	0.050[a] (0.034)[b] (0.017)[c]

[a] Transition probability reported in the MS and used in the model. [b] Transition probability estimated by ERG assuming 84 days of sunitinib treatment (median 2 cycles). [c] Transition probability estimated by ERG averaging discontinuations across 21.4 weeks of progression-free survival.

Table 21 also reports a value applied in the model for discontinuation of sunitinib due to adverse events. The ERG has concerns regarding the use of this transition probability and believe it is likely to:

- over-estimate the probability of patients discontinuing sunitinib treatment
- over-estimate the rate of disease progression toward the best supportive care health state
- under-estimate quality of life.

It is not clear from the study report whether the reported median number of days "on drug" takes account of the 2-week period within each sunitinib treatment cycle where patients do not take the drug. Table 2 in the journal publication by Demetri and colleagues (2006)[46] indicates the median weeks on sunitinib treatment was 12. The reported median value of 56 days fall short of an average treatment duration of 84 days [based on multiplying the reported median number of cycles (2) by the number of days, on and off-drug, in each cycle (42)]. Even if the monthly transition probability of 0.050 estimated in the MS were correct, it does not seem appropriate to apply it to all cycles in the sunitinib state. The median progression-free survival used to derived the monthly transition probability for discontinuation due to progression (24.1 weeks) is substantially greater than both the reported median number of days "on drug" and the estimated average treatment duration (based on the median of two cycles of sunitinib treatment). Given that patients discontinuing sunitinib treatment in the model progress directly to the BSC state – which is modelled as synonymous with disease progression and hence includes a substantial reduction in quality of life – this approach is likely to over-estimate the probability of patients

discontinuing sunitinib treatment, over-estimate the rate of disease progression toward BSC health state and under-estimate the QALYs.

The probability of progression in patients who discontinue treatment due to adverse events should be captured in the Kaplan-Meier estimates of time to progression and progression-free survival, if the intention-to-treat principle has been followed. Adjusting the model for discontinuation of sunitinib, due to adverse events, would more appropriately be achieved by adjusting the health state cost – or by re-designing the model states.

Mortality in patients receiving imatinib (first line – i.e. non-adjuvant)
The probability of death due to GIST recurrence in patients receiving imatinib for primary recurrence following surgery is based on overall survival reported by Verweij and colleagues.[45] The Kaplan-Meier survival at 1 and 2-years for patients receiving 400 mg imatinib once daily were 85% and 69% respectively. These were transformed to monthly transition probabilities, using appropriate formulae – 0.013 and 0.015 respectively. The value derived from the 1-year survival estimates was used in the model for the base case, with the year 2 value used in a sensitivity analysis. The Kaplan-Meier survival curves reported in Figure 6 of the trial publication by Verweij and colleagues[45] suggest that constant mortality risk (with respect to time) is an appropriate assumption (i.e. an exponential survival function).

Mortality in patients receiving sunitinib (second line)
The probability of death due to GIST recurrence in patients receiving sunitinib (either following primary recurrence during adjuvant imatinib therapy or following recurrence during non-adjuvant imatinib therapy) is based on data reported in the trial Demetri and colleagues (2012).[47] The median survival duration for sunitinib-treated patients was 72.7 weeks. This was transformed for use in the model by first calculating the survival duration in months, assuming an exponential survival function to derive a monthly rate and converting this to a transition probability as described previously.

Mortality in patients receiving BSC
The MS reports that probability of death due to GIST recurrence in patients receiving BSC is based on data reported in three publications: a report of an RCT (Demetri and colleagues (2012)[47]), an epidemiological study using US registry data (Tran and colleagues[48] and a post-hoc sub-group analysis of data from an RCT used in a decision model (Huse and colleagues,

2007[49]). The MS provides no information on how these studies were identified and offers no critical appraisal or discussion of the generalisability of these studies to the modelled population. The three studies vary in terms of the included populations (see Table 22) and provide substantially different estimates of median overall survival duration for adults with GIST (see Table 22). The MS makes no judgement on the comparability of these studies and simply uses the mean value of 0.043 in the model (using the 1-year survival from Tran and colleagues, 2005[48] and values in Table 22 for the other two studies).

Table 22: Overall survival. Data sources and parameter estimates for manufacturer's model

Source	Study population	Extracted data	Transition probability
Demetri et al.[47]	Adults with histologically proven GIST, failed prior imatinib due to resistance or intolerance, ECOG PS ≤ 1 and adequate hepatic/ renal/ cardiac function	Median survival = 39 weeks[a]	0.074
Tran et al.[48]	Diagnosed malignant GIST (histologically confirmed in subjects over 20 years of age)	Median survival = 2.97 years Survival at 1 year = 77%[c] Survival at 5 years = 38%[c]	0.019[b] 0.022 0.016
Huse et al.[49]	Adults with histologically confirmed unresectable or metastatic GIST that expressed CD117, ECOG PS ≤ 3, adequate hepatic/renal/cardiac function	Median survival = 20 months	0.034

[a] Median survival in the placebo arm of the trial, adjusted for cross-over using the rank preserving structural failure time model (RPSFTM). Median survival without adjustment was 64.9 weeks (approximately 15 months). [b] Median survival reported by Tran et al [48] was not included in MS. Included here by ERG for reference. [c] These survival probabilities are for a sub-group of patients in the registry database identified as "white" (see Table 2, 1-Year and 5-Year Observed and Relative Survival Rates of Patients with GIST Diagnosed During 1992 - 2000 (N=1,430), page 165 of Tran et al.[48]

The ERG feel that the manufacturer should have considered these three studies more critically and should have highlighted the possible limitations of these data sources – in particular, the fact that none of the studies appear to identify the proportion of patients at high (or "significant") risk of recurrence, which is the patient group included in the model. The population in Tran and colleagues[48] include subjects with varying stages of disease which vary substantially in survival probability (from 91% 1-year survival with local disease to 49% 1-year survival with distant disease) and in therapy (including some who had undergone no intervention) which makes this study unsuitable for providing a survival estimate for the BSC state (which is treated for quality of life and costing terms as being synonymous with progressive/ end stage disease) in the model. The sub-group used to derive the median survival estimate reported in Huse and colleagues[49] were patients who discontinued imatinib treatment in a RCT of two doses (400mg

and 600mg) of imatinib that had no BSC/placebo control. It is difficult to assess the appropriateness of the analysis or the generalizability of the results as limited methodological information is reported by Huse and colleagues[49] and no information on baseline characteristics in the sub-group of patients analysed.

Of the three studies it seems that Demetri and colleagues[47] may offer the most robust survival estimate for adults with GIST who receive no active treatment. However the reliability of the survival estimate from this study is highly dependent on the approach adopted to adjust for cross-over in the study placebo arm.

Mortality from other causes

Mortality for all causes (other than GIST recurrence) was derived from published life tables for England (2004 - 2006 Interim Life Tables). These were converted from annual rates to monthly probabilities using similar transformations to those described above. The general population mortality rates derived from the life tables have not been adjusted for GIST mortality. As a result there is a risk of double-counting, since excess mortality due to GIST recurrence is also included in the model. Given the low proportion of GIST deaths within the general population mortality statistics, this is unlikely to have a substantial impact on the model results.

The MS used mortality rates starting from age 61 (mean age in the identified clinical trials). However, the sex composition used to derive the mortality probabilities was based on the proportions of males and females in the life table, rather than a population with GIST at high risk of recurrence. This was not discussed in the MS but is unlikely to have a substantial impact on the model results.

Resistance

In the adjuvant setting patients may be inherently resistant to imatinib (primary resistance) in which case they would be expected to derive no benefit from treatment (manifest by recurrence during or after completion of treatment). Patients may also acquire resistance following exposure to imatinib (secondary resistance) where they might still gain a benefit (delay in recurrence) but ultimately recurrence would occur (either on or after treatment – with no response to re-challenge with imatinib in this situation). Because c-kit/PDGFRA (platelet derived growth factor receptor alpha) mutation testing should routinely be performed in all patients being considered for adjuvant imatinib treatment, the majority of patients with primary resistance

(identified by genetic mutation site) will not be offered adjuvant treatment. Therefore the majority of resistance is likely to be secondary.

In the previous NICE appraisal of imatinib for GIST (TA196)[1] the ERG noted there was a lack of data on long-term treatment resistance. Clinical specialists to the NICE appraisal commented that it is plausible that resistance to imatinib may occur increasingly after the first year of adjuvant treatment. The ERG noted that resistance to imatinib had not been incorporated into the manufacturer's base case analysis and it was assumed that patients in the adjuvant imatinib arm do not develop early resistance to imatinib in the adjuvant setting. The manufacturer later investigated the potential development of resistance to imatinib during the Appraisal Consultation stage, providing several scenarios in which a proportion (between 0% and 100%) of patients receiving adjuvant imatinib develop resistance. The ERG noted that the approach used in further exploratory analyses was unclear and the Appraisal Committee concluded that the impact of imatinib resistance means there was high uncertainty in the cost-effectiveness estimates. The Committee noted that until further data become available from ongoing trials, it is unlikely that the possibility of patients developing resistance to imatinib can be fully evaluated.

In the current appraisal the ERG asked the manufacturer to clarify how resistance was defined, and incorporated into the economic model (Clarification question B4). The manufacturer responded that resistance can be primary (no response) or secondary (recurrence following an initial response). They noted that in the adjuvant setting it is difficult to discern between primary and secondary resistance as there is no disease as such to respond to. They clarified that the economic model accounts for patients experiencing recurrence whilst on treatment "representing the rates seen in the trials" and that the response rates from the trials implicitly include resistance. They therefore considered that it is not necessary to take account of resistance in any additional way. The ERG notes that on treatment resistance data is now available for up to 3-years adjuvant imatinib treatment. The journal publication of the SSGXVIII/AIO trial reports that 4 (2%) patients receiving 1-year imatinib experienced recurrence during treatment, compared to 12 (6%) in the 3-year group, describing it as 'infrequent'.[16] In the ACOSOG Z9001 trial the number of early withdrawals from treatment due to recurrence was low for imatinib patients (n=1, <1%) compared to placebo patients (n=41, 12%).

Another concern that the NICE Appraisal Committee had was that if adjuvant treatment with imatinib led to development of resistance, this could potentially shorten the duration of benefit

from any subsequent imatinib treatment after disease progression.[1] There is currently little evidence of the longer-term impact of adjuvant imatinib resistance. However, the MS reports a conference abstract by Reichardt and colleagues (2012)[51] on patients from the SSGXVII/AIO trial who were diagnosed with recurrent GIST after having received imatinib in the adjuvant setting. At a median follow-up time of 54 months 84 (42%) and 50 (25%) patients had a recurrent GIST or died in the 1-year and 3-year treatment groups respectively. Fifty four (27.1%) and 27 (13.6%) were re-challenged with imatinib (88% received 400mg/day). Forty-six (56.8%) of these 81 patients were evaluable for response. A clinical benefit rate (complete response + partial response + stable disease) of 84.4% was reported, with no difference between the 1-year and 3-year treatment groups (87.9% vs 76.9%, respectively; p=0.385). The median time to progression after re-challenge was 35.7 months, with no statistically significant difference between trial arms. It was concluded that most patients diagnosed with recurrent GIST in the adjuvant setting re-challenged with imatinib show a response, and the duration of adjuvant treatment does not affect the future response to imatinib. It should be noted that these results apply to the main trial population classified as high risk of recurrence by the NIH consensus criteria, rather than high risk by the Miettinen criteria. A longer follow-up of this sub-group of patients is needed to assess the impact of adjuvant therapy on the time to development of secondary resistance on imatinib re-challenge. The MS also reports a retrospective analysis of re-challenge with imatinib following relapse after 1-year of adjuvant treatment for 23 patients who had a relapse after adjuvant imatinib treatment in the ACOSOG Z9001 trial (MS Table B-30). In the majority of available cases patients responded to re-challenge with 400 mg imatinib. The MS suggests that response rates and duration of response are comparable with those observed in patients who have not been exposed to prior adjuvant imatinib. The ERG notes that the manufacturer's model assumes that all patients re-challenged with imatinib following post-treatment recurrence respond. Given that around 15% of patients in the analysis reported by Reichardt and colleagues[51] did not achieve clinical benefit the ERG has conducted a scenario analysis to assess the impact of this on the ICERs (section 4.3).

Relevant to this discussion is the effect of imatinib re-challenge in patients with advanced GIST. Resistance has been shown to occur to imatinib in advanced GIST at a median time of 18-26 months, and is most commonly caused by acquisition of secondary KIT mutations.[52] A review of recent data from the French Sarcoma Group BFR14 RCT (discussed in TA196 based on Blay et al 2007[31]) examined the impact of interrupting imatinib treatment on disease progression and on resistance.[52] The trial investigated interruption of therapy after 1, 3 or 5 years treatment with

400mg imatinib in patients with advanced GIST. Patients were randomised to remain on imatinib, or to cease imatinib and restart at the same dose on progressive disease. Interruption was associated with a high risk of progression, and tumour response on re-challenge seldom reached that before treatment interruption. Patients receiving continuous imatinib maintained a high rate of tumour control, increasing with longer imatinib treatment. Imatinib-resistant progression free survival was not significantly different between continued and interrupted treatment groups (though caution is required due to small number of patients). It was also reported that patients remaining on continuous therapy were less likely to develop secondary resistance to imatinib.

Given that the patients in the BFR14 RCT had advanced GIST it cannot be assumed that similar findings would be observed in patients in the adjuvant setting. The ERG clinical advisor notes that the likelihood of developing acquired (secondary) resistance is probably related (at least in part) to tumour volume, with the more tumour a patient has the more likely it is that a new mutation will occur somewhere that will lead to resistance. As patients receiving adjuvant therapy have no overt disease they have relatively low volume (microscopic) disease and so would be expected to be less likely to develop acquired resistance than patients with recurrent/advanced disease included in the BFR14 trial. Of note, the EORTC 62024 study of adjuvant imatinib was designed to assess secondary resistance through measuring 'Imatinib failure-free survival'. As reported in section 3.3, the interim results show a non-statistically significant trend in favour of 2-year adjuvant treatment for IFS (5-year IFS 79% vs 73% for observation only, p=0.11).

4.2.5 Patient outcomes

The cost-effectiveness model incorporated the impact of the treatment on HRQoL into QALYs. QALYs associated with each treatment strategy are estimated by applying state-specific utility estimates to patients' life expectancy in each of the model health states, adjusted by decrements to allow for the impact of AEs on HRQoL. Recurrent GIST and development of progressive disease are assumed to have a negative impact on HRQoL, hence progressive disease states are associated with increasingly lower health state utility. In addition (as discussed in the previous section) recurrent GIST with or without treatment is also associated with higher (GIST-specific) mortality risks, hence recurrence and progressive disease are associated with lower life expectancy as well as poorer HRQoL. The HRQoL impact of AEs

while on treatment are incorporated using a fixed decrement and are not directly related to the incidence or severity of AEs reported in the imatinib clinical trials reviewed in section 3 of this report. It is difficult to judge what impact this assumption may have on the results of the model. However applying a HRQoL decrement reported for sunitinib-treated patients to those treated with imatinib may be expected to over-estimate the quality of life impact of imatinib treatment as it is generally considered to have a better adverse event profile than sunitinib.

The utility weights used in the model are reported in section 7.4.8 of the MS. These are all based on published sources as no HRQoL data (or measures that could be mapped to QoL) were collected in the pivotal imatinib RCTs. Section 7.4.5 of the MS reports the searches undertaken to identify studies of HRQoL in patients with GIST. The MS does not explicitly state the inclusion/ exclusion criteria for the review, but the inclusion criteria (based on exclusions reported in Figure B-18 on page 175 of the MS) appear to be studies in a population of adult patients with GIST reporting outcomes in terms on HRQoL or health utility. It does not appear that there were any inclusion criteria relating to specific HRQoL instruments or valuation methods, although the MS states that “a health states utility filter” was included in the search. The MS reports that the systematic searches identified three studies,[53-55] one of which is an economic evaluation included in the manufacturer’s systematic review of economic evaluations (Majer and colleagues[38], section 7.1.2 of the MS). They do not mention the other economic evaluation included in their systematic review (Sanon and colleagues)[39], although that also used utility values from previously published sources. The MS states that only one source, Chabot and colleagues,[53] was identified that reported patient-derived utility values using EQ-5D, in patients with GIST.

The health state utilities used in the model are reported in MS Table B-33 (page 178 -179) and shown below in Table 23. Utilities for recurrence-free health states (A, B and D) are based on age-specific utility values derived from a regression model reported by Ara and Brazier.[56] The model was developed using individual responses to the EQ-5D questionnaire in the 2003 and 2006 Health Survey for England[57;58] which was valued using weights based on time trade-off valuations from the UK general public.[59] For health state B (recurrence-free, receiving adjuvant imatinib) a decrement of 0.081 derived from the study by Chabot and colleagues reporting the on-treatment disutility with sunitinib.[53] In the absence of any other relevant sources this decrement was applied to adjuvant imatinib.

Table 23: Health state utility values used in the economic model

Health state	Value from MS	Source	Calculation/ assumptions
A: recurrence-free	0.822	Ara and Brazier[56]	Calculated for mean age 61 and sex breakdown from UK life table[a]
B: recurrence-free, receiving adjuvant imatinib	0.741	Ara and Brazier[56], Chabot and colleagues[53]	Recurrence-free value minus 0.081
D: recurrence-free, completed adjuvant imatinib	0.822	Ara and Brazier[56]	See above
C: GIST first recurrence. Treated with non-adjuvant imatinib	0.739	Chabot and colleagues[53]	Apply the same assumption as sunitinib – although treatment patterns are different.
E: GIST first recurrence. Treated with sunitinib	0.739	Chabot and colleagues[53]	0.712 + 1/3 * 0.081 [b]
G: GIST recurrence. Second line treatment with sunitinib	0.739	Chabot and colleagues[53]	See above
H: BSC	0.577	Chabot and colleagues[53]	Model uses value reported for progression [c]

[a] Value is then fixed for model run. [b] Sunitinib is provided on a six week cycle, consisting of 4 weeks of treatment and 2 weeks off treatment. Hence the manufacturer added 1/3 * 0.081 (an average improvement over the cycle) to the value reported for patients during the 4 weeks of treatment with sunitinib. [c] Patients reach the BSC state via a range of transitions, some of which are unrelated to recurrence (for example, patients assumed ineligible for sunitinib due to frailty (10%) enter the BSC state, patients discontinuing sunitinib due to adverse events enter the BSC state).

All other utility values were taken from Chabot and colleagues.[53] (The ERG notes that the Chabot and colleagues[53] study was also the main source of utility data in the previous submission to NICE for TA196.[1]) The MS reports that the utility values for patients receiving sunitinib following recurrence are taken directly from this publication. However the reported value of 0.739 does not appear in the publication, but is based on an additional calculation that take into account the fact that sunitinib is provided on a six week cycle including 4 weeks receiving the drug and 2 weeks without treatment. Chabot and colleagues[53] reported a 0.081 improvement in QoL during the off-treatment period. As a result an averaged value for the six week cycle is applied in the model. The same averaged value is applied for patients receiving

imatinib following first GIST recurrence, although the pattern of treatment with sunitinib and imatinib is different. No rationale or justification for this is provided in the MS, other than a discussion of the relative AE profiles of sunitinib and imatinib.

There is very little information in Chabot and colleagues[53] on the population and methods used for the derivation of utilities adopted in the model, other than a statement that the EQ-5D questionnaire was used. There is a brief statement that implies the data were based on responses by patients in the Demetri and colleagues (2006)[47] placebo-controlled sunitinib trial. However the study reports no information on baseline characteristics of respondents, sample size, response rate or the valuation method adopted (there is no indication whether the utility values have been derived using a population tariff or using patient-assessed VAS valuations). The lack of methodological detail and the absence of information on respondents in the study limits the ability to critically appraise these valuations and, while they appear to be the only published set for patients with advanced gastrointestinal stromal tumours who were resistant to/ intolerant of imatinib, the ERG suggest they should be treated with caution. In particular the MS does not discuss the appropriateness of applying a utility value reported for patients with progression (0.577, see Table 23) to the BSC state, despite the fact that patients in the model can reach this state via transitions which are unrelated to recurrence. For example, patients who enter the BSC state as a resulting of discontinuing sunitinib due to adverse events might more appropriately be ascribed a health state value of 0.781 (reported by Chabot and colleagues[53] for non-progressed patients receiving BSC) rather than the value reported for progressive disease.

The MS does not provide a rationale for calculating an age-sex specific utility value for patients in the recurrence-free health states, at the start of the model and keeping this constant over a 50-year time horizon. This would be expected to have the effect of over-estimating the health benefits of patients remaining in the recurrence-free health states, which is likely to bias the analysis in favour adjuvant treatment. However, since none of the other utilities applied in the model are age-specific, recalculating the utility value as patients age would risk introducing illogical values into the model (for example the age-related utility for the recurrence-free health state might reduce to below the value used for the recurrent disease states). There is no discussion of the appropriateness of basing the sex distribution (on which the recurrence-free utility value is calculated) on a general population basis (derived from UK life tables) rather than on the breakdown in the clinical trials or a representative sample of UK patients undergoing surgery for GIST.

4.2.6 Resource use

Resource use reported in the MS includes drug costs, on-treatment monitoring and health state costs (primarily related to post-treatment monitoring for recurrence/ disease progression). Standard post-surgical follow-up includes out-patients appointments, CT scans, GP visits as well as full blood counts and liver function tests, primarily monitoring patients for GIST recurrence. The annual frequency of post-surgical monitoring reduces with time, roughly halving (from four out-patients appointments, four CT scans and two GP visits) at 3-years post-surgery and halves again at 5-years. Monitoring during adjuvant imatinib treatment is similar to standard post-surgical follow-up, but is assumed to be less intensive than for patients not receiving adjuvant imatinib. Resource use for each of these is presented in MS Table B-37 (page 191) and MS Table B-38 (page 192). Resource use for management of recurrence is presented in MS Table B-39 (page 192). Drug acquisition and resource use for management of AEs is detailed in MS Table B-40 (page 194).

Drug use in the model is based on standard dosing of imatinib for adjuvant therapy of 400 mg once daily. The expected course of adjuvant treatment is continuous treatment over 1 or 3-years. The model base case does not take account of dose adjustments or interruptions. However mean dose provided in the clinical trials is included in a sensitivity analysis. No additional resource use has been included for administration as imatinib is an oral medication and no additional pharmacy resource has been included.

Patients experiencing recurrence who did not receive adjuvant imatinib or who completed adjuvant treatment prior to recurrence receive non-adjuvant imatinib at the standard dose of 400 mg once daily until further recurrence, discontinuation due to AEs or death. A proportion (15%) of these patients experiencing first recurrence will also undergo further surgery.

Patients experiencing recurrence while receiving adjuvant imatinib or who experience recurrence while receiving non-adjuvant treatment are eligible to receive sunitinib at the standard dose of 50 mg daily for 4 weeks within a six week treatment cycle. Given this six week cycle, the estimation of resource use with sunitinib is complicated in the model, which uses monthly cycles. Sunitinib use was estimated allowing for a 21% probability of discontinuation per month (discussed in section 7.3.2 of MS). This figure was used to derive an estimated mean duration of treatment of 4.82 months (20.87 weeks) or 3.48 cycles. This estimate was based on data on mortality, progression-free survival and discontinuation due to adverse events reported

from two placebo-controlled sunitinib trials[46;47] (see Table 21). The MS does not compare the estimated number of cycles of sunitinib used in the model with those reported for the two trials (both of which report a median of two cycles of treatment with sunitinib). The MS appears to have over-estimated use of sunitinib compared with that reported in the two clinical trials. Using methods similar to those presented in the MS to convert the reported median of two cycles to a discontinuation rate (-1/2*ln(0.5) = 0.3466) and then to mean (by taking the reciprocal of the discontinuation rate, 1/0.34466 = 2.89 cycles). The ERG suggest it would be more appropriate to use this estimate for the number of cycles of sunitinib in the model, as it is derived directly from the trials used to estimate the clinical effectiveness of sunitinib in imatinib-resistant patients with GIST.

A systematic search was undertaken in order to identify publications reporting resource use and costs relevant to GIST in order to populate the model (MS section 7.5). Non-UK studies were excluded from consideration, due to likely differences in resource use between countries. The searches did not identify any primary studies reporting resource use associated with management of GIST in the UK. All non-primary studies identified were STA submissions to NICE and were not considered further.

Resource use assumptions related to on-treatment monitoring and post-treatment follow-up were based on UK clinical guidelines[5] and assumption. The MS does not explicitly indicate whether these assumptions have been subject to discussion with relevant clinical experts or any other clinical validation. These assumptions are in line with the treatment algorithm presented in Figure A-1 (page 26) of the MS (derived from UK clinical guidelines[5]) are summarised in Figure 7, developed by the ERG.

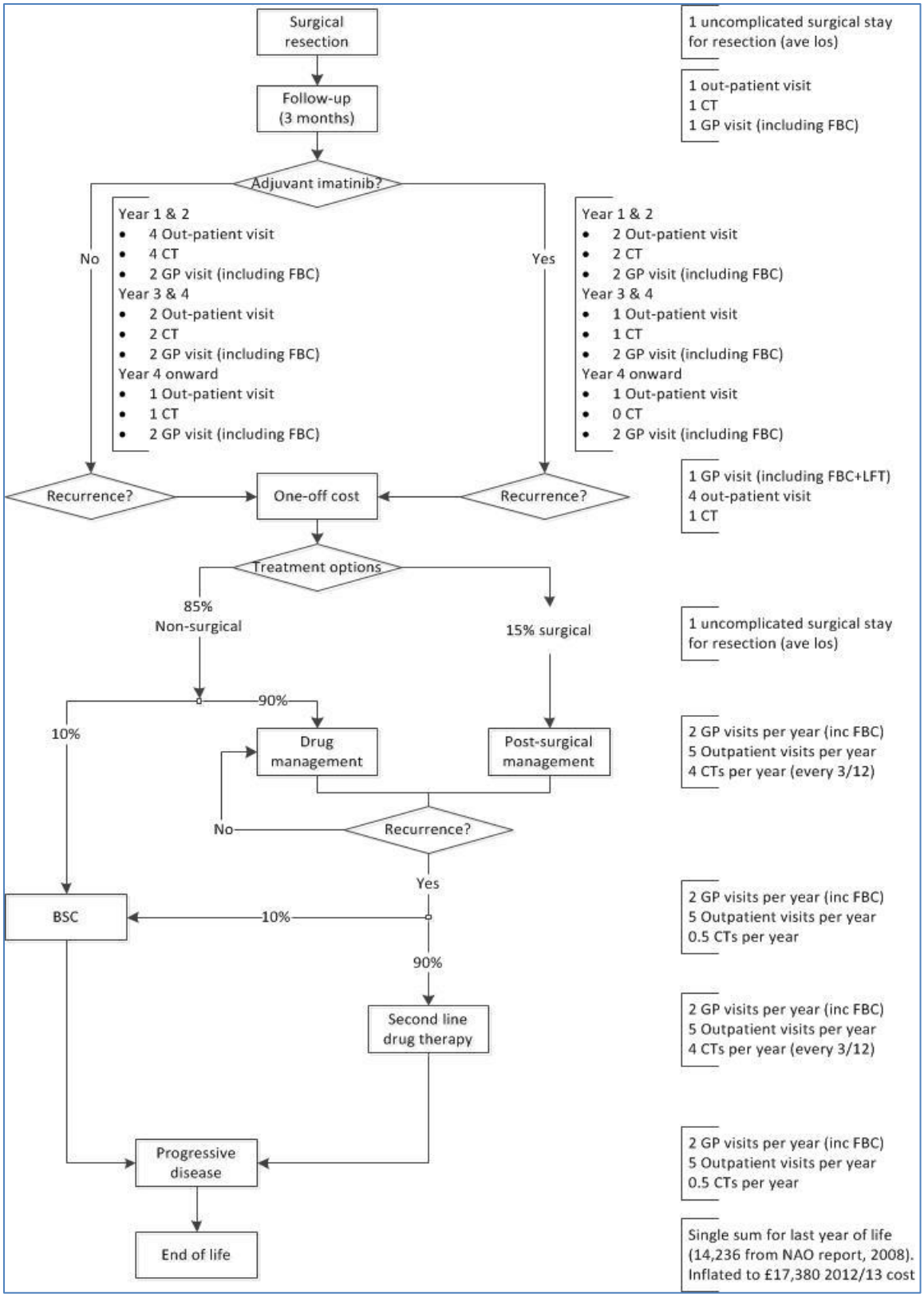


Figure 7: Resource use assumptions included in the model

Clinical advice to the ERG suggested that the assumed frequency of clinical monitoring in secondary care for patients treated with adjuvant imatinib, while in line with clinical guidelines, may be lower than would be expected in routine NHS practice. It may be more appropriate to assume identical frequency of follow-up for both groups of patients following surgery. However, the assumption that no further CT scans are required after 5-years of RFS in patients treated with adjuvant imatinib is likely to agree with current NHS practice.

Health state cost for BSC, as indicated in Figure 7, was a combination of on-going patient monitoring and end of life costs. As with the resource use estimate for sunitinib monthly resource use for this state is estimated outside the model, based on an average duration in this state of 23.12 months (estimated from the probability of death in the BSC state). The MS assumes that resource use for the final twelve months relate to end of life care, while resource use for the preceding 11.12 months relates to standard monitoring. The ERG is concerned that the average time spent in the BSC state is determined outside the model and is entered as a deterministic value and, therefore, not included in the probabilistic sensitivity analysis. Given that the three studies used to derive the death rate applied to the BSC state produced widely different estimated monthly death rates (0.034, 0.022, 0.074), hence widely different estimates of time spent in this health state (28.85, 45.91 and 12.98 months respectively) it would be more appropriate to reflect some of this uncertainty in the model results.

Resource use for adverse events in the model was based on the observed frequency of grade 3 and 4 adverse events in both arms of the SSGXVII/AIO trial. Comparing Table B-40 in the MS (page 194 of the MS) with Table 12-7 (Most frequent grade 3 or 4 adverse events by preferred term for overall treatment period) in the Clinical Study Report indicates that adverse events affecting at least one percent of the trial population were included, although this is not explicitly stated in the MS. The MS assumed that each adverse event would be associated with three out-patient appointments, with an additional assumption that 5% of patients with decreased neutrophil count, decreased white blood cell count, diarrhoea and nausea would be hospitalised. The MS states that these assumptions were based on clinical advice but does not report the number of clinicians approached or the method for eliciting expert opinion. These assumptions are used to calculate an average resource per adverse event (weighted by the occurrence of adverse events observed in the trial). No resource use assumptions for adverse events occurring in patients receiving non-adjuvant imatinib or sunitinib are reported in the MS. Examination of the electronic model indicates that the resource use assumptions used for

patients receiving adjuvant imatinib are applied to patients receiving non-adjuvant imatinib or sunitinib. The MS contains no discussion of the appropriateness of applying the same resource use assumptions for adjuvant imatinib, non-adjuvant imatinib and sunitinib nor does it provide any rationale or explanation for adopting this approach.

4.2.7 Costs

Drug acquisition costs applied in the model are presented in Table 24. Drug costs in the MS were based on BNF 66 (October 2013).[6] Costing for sunitinib is further complicated by the existence of a patient access scheme (PAS) where the first treatment cycle is free to the NHS. The MS took account of the PAS in the base case analysis, but included the full cost of sunitinib is a sensitivity analysis.

Table 24: Drug acquisition costs

Intervention	Dose	Unit cost/ dose (£)	Treatment course	Cost (£)
Adjuvant imatinib	400 mg daily	57.48	1 year 3 years	20,994[a] (1,749.54)[b]
Non-adjuvant imatinib	400 mg daily	57.48	Until recurrence, discontinuation due to AEs or death	
Sunitinib	50 mg daily	112.10	Six week cycle till recurrence, discontinuation due AEs or death	3,138[c] (1615.34)[d] (2,266.91)[e]

[a] Cost per patient year. [b] Cost per month applied in the model. [c] Cost per six week cycle. [d] Cost per month (accounting for PAS). [e] Cost per month (Full cost, i.e. excluding PAS).

As noted in the previous section, the MS appears to have over-estimated use of sunitinib compared with that reported in the two clinical trials which were used estimate the clinical effectiveness of sunitinib in imatinib resistant patients with GIST. The model assumes that 3.48 cycles of sunitinib are provided whereas the ERG estimated an average of 2.89 cycles (based on a median of two cycles reported in both sunitinib RCTs).[46;47] Using an estimate of 2.89 cycles reduces the estimated monthly cost of sunitinib to £1,882.57, excluding PAS, and £1,231.17 (allowing for the PAS).

The ERG is also concerned that the estimated average time on sunitinib treatment, hence the estimated number of treatment cycles in the model, is entered as a deterministic value (determined outside the model) and is not included in the probabilistic sensitivity analysis. Given

that sunitinib discontinuation and GIST mortality while on sunitinib (which were used to estimate the average duration of sunitinib treatment) are both included in the PSA, it is unclear why this derived parameter was excluded.

Costs for the majority of other resource use in the model were taken from NHS Reference Costs 2011/12,[60] with costs for GP visits taken from Unit Costs of Health and Social Care [61] (see Table 25).

Table 25: Unit costs associated with monitoring and follow up visits in the MS

	Mean £	**Source/Comments**
Complete blood count	3.00	NHS reference cost 2011/12. DAP823
Liver function tests	1.00	NHS reference cost 2011/12. DAP841
Routine OP visit	128	NHS reference costs 2011/12 HRG code: 301 Outpatient Follow-up Medical Gastroenterology
CT scan	135	NHS reference costs 2011/12 Tariff RA13Z (three area with contrast)
Surgery (on recurrence)	4,931	NHS reference costs 2011/12 Weighted average of G04 Complex Open Hepatobiliary or Pancreatic procedures (10%), GA05 Very Major Open Hepatobiliary or Pancreatic procedures (17.5%), GA07 Intermediate Open Hepatobiliary or Pancreatic procedures 20%), GA13 Minor Open or Laparoscopic, Hepatobiliary or Pancreatic procedures (2.5%), and FZ12 Major General Abdominal procedures (50%).
GP visit	40	Curtis 2012[61] Per surgery consultation lasting 11.7 minutes (with qualification costs, excluding direct care staff costs) Unit Costs of Health and Social Care

CT, computed tomography; GP, general practitioner; HRG, Health Resource Group; OP, out-patient.

Health states costs were calculated by multiplying the resource use estimates reported in Tables B-37, B-38 and B-39 (pages 191 to192 of MS), summarised in Figure 7 of this report, by the unit costs and then averaging to derive monthly costs. Health state costs are reported in Table B-35 of the MS. For the recurrence-free health state and those where patients receive

active treatment, the majority of the health state costs relate to out-patient follow-up and regular CT scans. The MS contains no discussion of the appropriateness of using the Reference Cost for medical gastroenterology follow up for all outpatient contacts, including those for patients receiving new drug treatment for recurrent disease. The reliance on this single Reference Cost is likely to underestimate the costs of care provided by a multi-disciplinary team in the group of patients experiencing recurrent disease and switching treatments.

Adverse events included in the model were costed using NHS Reference Costs 2011/12.[60] The majority of adverse events were assumed to be treated in out-patients, using the same unit cost (Outpatient follow up medical gastroenterology) as was used for costing standard follow up. The MS contains no discussion of the appropriateness of using this Reference Cost for outpatient contacts for patients experiencing chemotherapy-related adverse events. Additional adverse event costs are included on the assumption that five percent of those experiencing decreased neutrophil count, decreased white blood cell count (costed as fatigue), vomiting and diarrhoea will require hospitalisation (although the basis for these assumptions is not discussed in the MS). Table 26 reports the unit costs applied to AEs in the model.

Table 26: Unit costs for adverse events in the MS

AE	Cost	Source
Neutropenia	£2,372.87	Average (weighted by activity levels) of: WA02W: disorders of immunity without HIV/AIDS without CC WA02Y: disorders of immunity without HIV/AIDS without CC National Schedule of Reference Costs, 2011/12
Fatigue	£328.25	Average (weighted by activity levels) of: WA18V: Admission for unexplained symptoms with Major CC WA18X: Admission for unexplained symptoms with Intermediate CC WA18Y: Admission for unexplained symptoms without CC National Schedule of Reference Costs, 20111/12
Nausea/ vomiting	£663.23	FZ43C: Non-Malignant Stomach or Duodenum Disorders with length of stay 1 day or less National Schedule of Reference Costs, 2011/12
Diarrhoea	£685.21	FZ36F: Intestinal Infectious Disorders with length of stay 1 day or less National Schedule of Reference Costs, 2011/12

The MS contains no discussion of the costs of treating adverse events in sunitinib treated patients, either in terms of the incidence of adverse events or appropriate unit costs in this group of patients. Examination of the electronic model indicates that the monthly unit cost estimate derived for imatinib-treated patients is also used to estimate the cost of treating adverse events in sunitinib treated patients. This does not appear appropriate as these unit costs were developed based on the incidence of adverse events in the SSGXVII/AIO trial. The ERG is also concerned that the overall probability of discontinuing sunitinib (0.167), including both discontinuations due to disease recurrence (0.117) and due to adverse events (0.050), appears to be used in the calculation of adverse events costs for sunitinib-treated patients. A fixed proportion (0.3) is then applied to the overall probability of discontinuing sunitinib – the ERG presumes this is to reduce the over-estimation of adverse events with sunitinib. However, none of this is described or discussed in the MS and no rationale for this approach to calculating adverse event costs for sunitinib is provided.

4.2.8 Consistency/ Model validation

Internal consistency

The electronic model is coded in Microsoft Excel and is fully executable. The model is well presented and documented and user friendly.

The MS states that quality assurance of the model included an independent health economist assessing the internal validity of the model through checking total numbers of patients in the health states for consistency, conducting an empirical validation comparing the costs and effects and a number of alternative scenarios and using a range of extreme parameter values.

The ERG have not undertaken a comprehensive check of all cells in the model, rather, random checking of the model has been done for some of the key equations in the model. Changing the parameter values produced intuitive results and from random checking the ‘wiring’ of the model appears to be accurate, although the ERG has uncovered some minor coding errors as listed below. The ERG was able to replicate the results presented in the MS and the deterministic sensitivity analyses. The ERG views the model as a reasonable approach to modelling the cost effectiveness of adjuvant treatment for GIST.

Calculation errors

There is a minor error for the calculation of health state costs for medical costs which differ from those in the MS:

- Worksheet Costs_QALYs_noNo tx: No recur and no treatment health state costs, the duration is incorrect (see MS Table B-31) £102 0 - 2 years, £94.83 2 - 4 years, £51 4 - 6 years, £29 6+ years.
- Worksheet Costs_QALYadj_1yr: No recur and no treatment health states: duration is incorrect after 4.25-years £102; No recur and adj tx: £51 until 3-years (should be 2.25-years), £18 after 6 years (should be 5-years).
- Worksheet Costs_QALYadj_3yr: No recur and no treatment health states: duration is incorrect £102 for 0 - 4 years, £51 4 - 6 years; £51 until 3-years (should be 2.25-years), £18 after 6 years (should be 5-years).

Error in calculation of utilities, no brackets in the formula:

- Costs_QALYs_Adj1yr and Cost_QALYs_Adj1yr columns F5-16, E5*uGISTnorecS-uImatadj should be E5*(uGISTnorecS-uImatadj).
- Similar error in columns V, AC, and AP. Similar errors in sheet Costs_QALYs_noTx col V, AC and AP.

The ERG corrects these errors in section 4.3.

External consistency

The MS model results and structure have been compared to the two economic models identified in the manufacturer's systematic review of cost effectiveness studies (MS page 234). The MS states that they had similar findings to those in the studies by Sanon and colleagues[39] and Majer and colleagues[38] although those studies were non-UK based and therefore they may not be generalisable to a UK setting. The ERG notes that both studies by Sanon and colleagues[39] and Majer and colleagues[38] were funded by Novartis.

The MS does not compare the results from the current economic model to those for the previous NICE appraisal TA196.[1] The ERG notes that the base case results for the previous appraisal were for a different patient population for patients with significant risk of recurrence (moderate risk and high risk), rather than high risk only as in the current MS. The previous

appraisal presented scenarios for high risk patients for 1-year and 3-year adjuvant imatinib and the ICERs for these groups are £6109 and £19,813 per QALY compared to no treatment. These are less favourable than those presented in the current MS (£3509 and £8390 per QALY). The ERG has not been able to ascertain the specific reasons for the differences between the results, however it notes that the approach taken to model RFS and the HRs used has changed between the appraisals.

The MS has compared the outcomes from the model for RFS and OS to the results of the clinical trials (MS page 205). The MS states that these show a good fit between RFS predicted by the model and the clinical trial results (Table 27).

Table 27: Comparison of MS model results with ACOSOG Z9001 and SSGXVII/AIO clinical trials

RFS (%)									
	Placebo		**1-year arm**				**3-year arm**	
Year	**Model**	ACOSOG *Primary*[a]	**Model**	SSG[b]	ACOSOG *Primary*[a]	ACOSOG *5 yr update*	**Model**	SSG[b]
1	**56**	52.2	**93**	91.3	93.6	94.8	**93**	96.6
2	**35**	39.5	**68**	68.0	74.1	76.0	**88**	89.6
3	**24**	-	**53**	49.6		57.2	**84**	83.8
4	**18**	-	**44**	40.2		44.7	**68**	71.6
5	**14**	-	**37**	35.1		37.9	**57**	58.8
OS (%)								
	Placebo		**1-year arm**				**3-year arm**	
Year	**Model**	ACOSOG *Primary*	**Model**	SSG	ACOSOG *Primary*	ACOSOG *5 yr update*	**Model**	SSG
1	**95**		**98**	98.6		98.9	**98**	100
2	**83**	94.7	**93**	94.1	100	98.9	**94**	97
3	**69**	-	**85**	91.7	-	94.2	**90**	95.5
4	**55**	-	**75**	83	-	86.9	**85**	94.5
5	**42**	-	**65**	74.2	-	82.0	**79**	89.5

[a] AGOSOG primary: RFS is based on efficacy population, high risk, event type =21. [b] SSGXVII/AIO RFS: is based on efficacy population, high risk (modified Miettinen risk classification). ACOSOG 5–year update: taken from clinical study report.

At 5-years, the model predicted that 57% of patients receiving 3-year adjuvant imatinib were recurrence-free, compared with 37% of patients receiving 1-year adjuvant imatinib and 14% of patients treated with surgical resection only. These are similar to those reported for the clinical trials (see section 3.3 of this report). The model estimates for the no treatment arm are compared to the ACOSOG Z9001 trial, however data is only available for 2-years. The MS

states that the primary analysis did not provide RFS out to five years due to limited follow-up. The ERG notes that there are data for the 5-year follow-up analysis, however these data was confounded by the placebo crossing over to adjuvant imatinib after study unblinding. The MS notes that 5-year RFS predicted by the model in the no treatment group (14%) are lower than seen in other observational studies (20%).

The model underestimates OS for patients receiving adjuvant imatinib, compared with the clinical trial data for 1-year and 3-year treatment and for the no treatment arm (only data available for the 2-year time point for the no treatment arm). The ERG considers that the OS data is more difficult to interpret, as there is no long term data for the no treatment group. The MS reported a scenario analysis with longer survival in post recurrence health states in order to provide a better fit for OS (as requested by the ERG, see section 4.3).

The MS reports the disaggregated results for each health state in MS Table B-51 - Table B-53. These are summarised in this report in Table 28. For treatment with adjuvant imatinib for 1-year, patients spend an additional 3.24 years in the no recurrence health states (A, B and D) and accrue an additional 2.66 QALYs. They also spend less time in the post recurrence health states. The additional cost of 1-year adjuvant imatinib is offset by the reduced cost of post recurrence imatinib (-£9,283), due to the short duration of time spent in health state C.

Table 28: No adjuvant treatment vs 1 year (reproduced from MS B-51 – B-53)

	Incremental: No tx vs. 1yr			**Incremental: 3yrs vs 1yr**		
	LYG	QALY	Cost	LYG	QALY	Cost
State A No recurrence and no adjuvant treatment	-2.141	-1.759	-£1,870	0.6845	0.5626	£254
State B No recurrence and on imatinib adjuvant therapy	0.905	0.741	£19,780	1.4784	1.1521	£31,976
State C Post-recurrence and on imatinib 400 mg once daily	-0.401	-0.296	-£9,283	-0.3220	-0.2380	-£7,464
State D No recurrence and completed adjuvant imatinib therapy	4.473	3.677	£1,591	0.0406	0.0334	-£427
State E Post-recurrence and on sunitinib	0.021	0.016	£514	0.0253	0.0187	£611
State G Sunitinib second-line treatment	-0.064	-0.047	-£1,326	-0.0510	-0.0377	-£1,066
State H Best supportive care	-0.166	-0.096	-£1,561	-0.1014	-0.0585	-£953
Total (discounted)	2.629	2.235	£7,844	1.75	1.43	£22,931

Tx, treatment.

The results are similar for 3-year treatment versus 1-year treatment, however here there is slightly lower gain in terms of QALYs and life years than for adjuvant imatinib for 1-year versus no treatment but with a large additional cost of adjuvant imatinib (£31,976).

4.2.9 Assessment of Uncertainty

The manufacturer has assessed uncertainty within the model by conducting sensitivity and scenario analysis for structural assumptions and parameter input values.

One-way sensitivity analyses

A description of the variables subjected to sensitivity analysis is given in MS section 7.6 (page 197). These include: treatment effect HR, probability of death from GIST in different health states, probability of discontinuation of imatinib treatment, utility values, sunitinib costs, health state management costs, cost of treating adverse events and the cost of BSC. The MS also includes sensitivity analyses for different time horizons, and alternative parametric survival models for RFS. The manufacturer provides detailed justification for the ranges used in the sensitivity analysis. Where possible the manufacturer has varied within the lower and upper confidence interval. ERG considers the parameters varied and the ranges chosen for the sensitivity analyses to be appropriate and comprehensive.

Results of the one-way sensitivity analyses (MS Table B-55 page 223) indicate that the ICER is most sensitive to the length of the time horizon, and the treatment HR for the on and off treatment phase. For a time horizon of 5-years, adjuvant 1-year imatinib treatment has an ICER of £7,368 per QALY versus no treatment, and adjuvant 3-year adjuvant imatinib treatment has an ICER of £89,182 per QALY versus 1-year adjuvant imatinib treatment. For changes to the off treatment HR, adjuvant 1-year imatinib varies between a dominant strategy (adjuvant 1-year imatinib cheaper and more effective than no treatment) and having an ICER of £21,498 per QALY compared to no treatment. Varying the on treatment HR, the cost effectiveness of 3-year adjuvant imatinib varies between £13,917 and £26,878 per QALY compared with 1-year adjuvant imatinib. The model results were fairly robust for changes to all other parameter values.

Scenario Analysis

The MS reports scenario analyses for alternative parametric distributions to extrapolate RFS (exponential, Weibull, log-logistic and gamma), dose escalation to 800 mg imatinib following

recurrence (as included in the previous NICE submission TA196), change to the proportion of patients moving to BSC (progressive disease) following recurrence and extended survival after recurrence. The results of the scenario analyses are shown in MS Table B-58 (page 231). In general, the model results were fairly robust to these scenarios. The scenario with the greatest impact on the model results is changing the parametric distribution to a gamma distribution. For this scenario adjuvant 1-year imatinib treatment has an ICER of £9,886 per QALY versus no treatment, and adjuvant 3-year adjuvant imatinib treatment has an ICER of £19,239 per QALY versus 1-year adjuvant imatinib treatment.

The scenario of assuming longer survival in post recurrence health states estimates survival using the lower confidence intervals of the GIST mortality for each of the post recurrence health states. In this scenario, the 5-year OS for the 3-year, 1-year and no treatment arms are 85%, 77% and 63% compared to 89.5% and 74.2% survival for the 3-year and 1-year arms respectively from the SSGXVII/AIO study.

The ERG considers that other scenarios could have been explored in the MS. For example, the MS could have investigated the impact of a waning effect on the off treatment hazard ratio, and the effect of changing the post-recurrence GIST mortality between treatment arms, such that the mortality probability is higher in the adjuvant imatinib arms than the control arm. The ERG requested the manufacturer provide these scenarios (section 4.3.2).

Probabilistic Sensitivity Analysis

The PSA uses 1000 iterations and takes about 4 minutes to run. Variables included in the PSA are reported in MS Table B-31 (page 161). The ERG considers that the PSA includes most of the variables within the model but not that the MS did not include variation around the cost of imatinib or the proportion receiving sunitinib or BSC after recurrence.

The PSA results (MS Table B-56 (page 227) are similar to the deterministic sensitivity analysis results. A cost effectiveness acceptability curve shows 41.7% and 58.3% likelihood that 1-year and 3-year adjuvant imatinib is a cost effective strategy when using Willingness To Pay (WTP) of £20,000 per QALY and 30.9% and 69.1% using the £30,000 threshold (Figure 8 below, MS Figure B-24).

The ERG considers that the probability distributions are correctly applied and the methods of assessment of parameter uncertainty are appropriate.

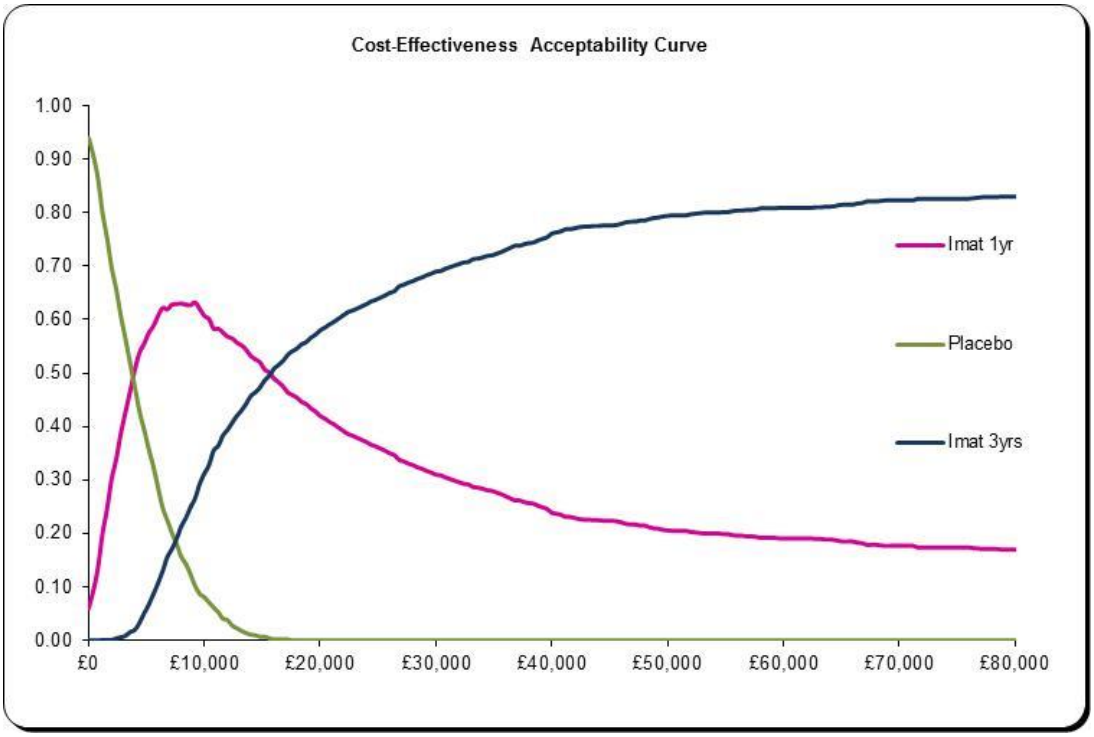


Figure 8: Cost-effectiveness acceptability curves (MS Figure B-24)

4.2.10 Comment on validity of results with reference to methodology used

The structure adopted for the economic model is reasonable and consistent with current clinical understanding of GIST and previous economic evaluations of treatments for GIST. The methods of analysis are appropriate and conform to NICE methodological guidelines. The parameters used for the model are generally appropriate.

The MS has provided disaggregated results for each health states that show that the main difference in costs between the arms are from the adjuvant imatinib treatment, offset by the reduced time spent on post recurrence imatinib. The increase in life years is largely due to the additional time spent in the no recurrence health states.

The MS has provided validation of the model results compared against the clinical trials for RFS and OS. These provide a reasonable fit for RFS against the clinical trials at 5-years for 1 and 3-year adjuvant imatinib treatment and for no treatment at 2-years. The fit for OS is less good and

the model underestimates OS for 1 and 3-year adjuvant imatinib treatment at 5-years and no treatment at 2-years. There is uncertainty around the estimation of long term extrapolation of RFS and the long term RFS differs widely according to the parametric distribution chosen. The ERG notes that the parametric distribution chosen by the manufacturer produces the most favourable ICER for adjuvant imatinib treatment.

4.3 Additional work undertaken by the ERG

4.3.1 Corrected base case

The ERG has corrected the errors identified in the manufacturer's model for utility and management costs, as described in Section 4.2.8. The corrected base case results are shown below in Table 29 and are similar to the MS base case.

Table 29: Base case cost effectiveness results

	Total Per Patient:		Incremental		
	Costs	**QALYs**	**Costs**	**QALYs**	**ICER (Cost/QALY gained)**
No treatment	£46,962	3.83			
Adjuvant imatinib 1-year treatment	£54,780	6.00	£7,819	2.16	£3,612
Adjuvant imatinib 3-year treatment	£77,708	7.37	£22,928	1.38	£16,663

4.3.2 Additional scenarios conducted by the manufacturer

The ERG requested the manufacturer to conduct scenario analyses that assumes treatment effect declines over time for the off treatment period. The manufacturer provided these scenarios as part of their response to the ERG's clarification questions (see clarification request B10). The ERG has not been able to check these analyses as the manufacturer did not provide the electronic model with these changes, and there is only limited information provided on the changes made to the model.

Scenarios were conducted where the off treatment HR was reduced after 5-years to 75%, 50% and 25%. The MS states that results for the SSGXVIII/AIO trial are reported for a median of follow-up of 54 months and therefore offer robust data for 5 to 6 years. The results are shown in Table 30 to Table 32. These show that the ICER increase varies between £4569 and £14,079 per QALY for off treatment HR of 75% and 25% respectively for 1-year adjuvant treatment

versus no treatment and between £18,242 and £34,683 per QALY for 3-year adjuvant treatment versus 1-year treatment. The ERG notes that the results for ACOSOG Z9001 are reported for a median follow-up of 19.7 months and therefore offer robust data for less than two years.

Table 30: Scenario A: Off treatment HR reduced to 75% after 5 years

	Cost/QALY Gained	Cost/LY Saved
Adjuvant Imatinib vs Placebo	**£4 569**	**£3 880**
Adjuvant Imatinib 3 years vs Adjuvant Imatinib 1 year	**£18 242**	**£14 818**
Adjuvant Imatinib 3 years vs Placebo	**£9 952**	**£8 302**

Table 31: Scenario B: Off treatment HR halved after 5 years:

	Cost/QALY Gained	Cost/LY Saved
Adjuvant Imatinib vs Placebo	**£6 831**	**£5 783**
Adjuvant Imatinib 3 years vs Adjuvant Imatinib 1 year	**£22 735**	**£18 277**
Adjuvant Imatinib 3 years vs Placebo	**£13 210**	**£10 951**

Table 32: Scenario C: Off treatment HR reduced to 25% after 5 years:

	Cost/QALY Gained	Cost/LY Saved
Adjuvant Imatinib vs Placebo	**£14 079**	**£11 821**
Adjuvant Imatinib 3 years vs Adjuvant Imatinib 1 year	**£34 683**	**£27 153**
Adjuvant Imatinib 3 years vs Placebo	**£22 939**	**£18 678**

The ERG requested a scenario analysis in which the post-recurrence GIST mortality is different between treatment arms, such that the mortality probability is higher on the adjuvant imatinib arms than the control arm. The ERG suggested the manufacturer repeat the sensitivity analysis carried out in their previous submission (i.e. submission for TA196[1]), as described in the West Midlands ERG report on page 69: 'the monthly probability of death in the recurrent state was changed so as to be greater in the adjuvant arm than the control arm.' In the current model this is state C (post recurrence and on imatinib 400mg). The manufacturer conducted this analysis by increasing the mortality 4-fold in the adjuvant arm for patients in the 'post recurrence and on imatinib 400 mg' health state – Health State C. The results are presented in Table 33.

Table 33: Scenario analysis increasing post recurrence death rate in treatment arms

	Total Per Patient:		Incremental:		ICERs
	Costs	QALYs	Costs	QALYs	Cost/QALY Gained
No treatment	£47 292	3.83			
Adjuvant imatinib 1 year treatment	£41 829	5.55	-£5 463	1.72	Dominant
Adjuvant imatinib 3 years treatment	£69 123	7.15	£27 294	1.60	£17 036

In this scenario, the life years in the 1-year and 3-year treatment arms are reduced from 7.71 and 9.46 to 6.95 and 8.95, respectively. In the 1-year arm the costs savings from avoided costs on treatment in the metastatic setting (imatinib, sunitinib and BSC) reduces the overall costs to be below those in the no adjuvant treatment arm so the 1-year treatment arm becomes dominant. In the 3-year arm the higher costs in the adjuvant setting mean the impact on overall costs is reduced.

4.3.3 Additional scenarios undertaken by the ERG

This section details the ERG's further exploration of the issues and uncertainties raised in the review and critique of the MS cost effectiveness analyses. These analyses concern alternative assumptions regarding the off treatment effect of adjuvant imatinib, the parametric distribution used for modelling recurrence-free survival, resistance to imatinib and the mortality estimates used for the recurrence health states. These analyses are shown in Table 35 separately and combined together.

The ERG noted that there is uncertainty around the continuation of the off treatment effect. The MS assumes that there is a continued off treatment effect beyond the reported trial follow-up. However the ERG considers this assumption may be optimistic and therefore presents an analysis that assumes there is no long term off treatment benefit beyond the reported follow-up of the clinical trial (i.e. after 5-years for 1-year and 3-year adjuvant imatinib from the SSGXVIII/AIO trial, and after 2-years for no treatment from the ACOSOG Z9001 trial). The probability of recurrence in the placebo arm has been set to be the same as the 1-year treatment arm after 2-years and the rate of recurrence for the 3-year treatment arm has been set to be the same as for the 1-year arm after 5-years (Table 34). The results are not changed significantly by changing this assumption.

Table 34: Probability of recurrence at selected time points for MS base case and ERG analysis with no extended off treatment benefit

	MS base case			**ERG analysis with no extended off treatment benefit**		
Time (months)	No treatment	1 year adjuvant imatinib	3 year adjuvant imatinib	No treatment	1 year adjuvant imatinib	3 year adjuvant imatinib
1	5.19%	0.59%	0.59%	5.19%	0.59%	0.59%
13	4.17%	2.73%	0.47%	4.17%	2.73%	0.47%
25	3.35%	2.19%	0.38%	2.19%[a]	2.19%	0.38%
37	2.69%	1.76%	1.82%	1.76%	1.76%	1.82%
61	1.73%	1.13%	1.17%	1.13%	1.13%	1.13%[b]
90	1.01%	0.65%	0.68%	0.66%	0.66%	0.66%

[a] Probability of recurrence for no treatment assumed to be equal to 1 year adjuvant imatinib after 24 months
[b] Probability of recurrence for 3-year adjuvant imatinib assumed to be equal to 1 year adjuvant imatinib after 60 months

The ERG expressed concern over the parametric distribution used for recurrence free survival and has presented the results for the corrected model using the exponential distribution as a plausible alternative. As a result of this change, the ICERs for 1-year adjuvant treatment versus no treatment increase to £9,386 per QALY and for 3-year treatment versus 1-year increase to £18,741 per QALY.

The MS varied the HR in the sensitivity analysis for the off treatment phase within the calculated confidence interval by varying both HRs for 1-year adjuvant imatinib versus no treatment and 3-year versus 1-year treatment together although they may vary independently of each other. The ERG investigated the effect of varying the off treatment HR for 1-year adjuvant treatment vs. no treatment whilst not varying the HR for 3-year versus 1-year treatment. The analysis was run for HR 95% upper confidence interval of base case estimate (HR = 0.906, HR 3-year imatinib off treatment 0.601), and also for the 5-year update unadjusted HR estimate (question C1 of manufacturer's clarifications; HR = 0.727, HR 3-year imatinib off treatment 0.482). The model results were very sensitive to changes in the off treatment HR for the 1-year adjuvant treatment versus no treatment analysis. For the 5-year update unadjusted HR estimate, the ICER for 1-year adjuvant treatment versus no treatment increases to £10,489 per QALY.

The ERG noted that the manufacturer's model underestimated OS and there was a closer fit to the clinical trial results using lower mortality rates. The ERG varied the mortality rate, using the lower confidence interval estimates for GIST mortality in the post recurrence health states. The analysis has the effect of a slight improvement in the ICER for the 1-year adjuvant imatinib

versus no treatment analysis.

The ERG investigated the effect on the model results of resistance to imatinib therapy. The model assumes that all patients re-challenged with imatinib upon recurrence respond. However, the ERG assumed that 15% of patients initially treated with adjuvant imatinib and re-challenged upon recurrence would not respond, based upon analysis of patients in the SSGXVIII/AIO trial by Reichardt and colleagues[51] (section 4.2.4). Upon non-response to re-challenge, these patients would progress to be treated with sunitinib. Accounting for this non-response produces marginal changes to the ICERs (Table 35).

Table 35: ERG additional analyses

	No treatment		**1-year adjuvant imatinib**		**3-year adjuvant imatinib**		**ICER (cost per QALY)**		
	Costs	**QALYs**	**Costs**	**QALYs**	**Costs**	**QALYs**	**1 year vs no treatment**	**3 years vs 1 year**	**3 years vs no treatment**
Corrected base case	£46,962	3.83	£54,780	6.00	£77,708	7.37	£3,612	£16,663	£8,684
No extended benefit	£44,823	4.30	£54,780	6.00	£77,466	7.41	£5,854	£15,995	£10,465
Exponential dist. used for RFS	£49,546	3.26	£62,464	4.63	£86,211	5.90	£9,386	£18,741	£13,871
HR off treatment 0.906 [a]	£46,962	3.83	£61,675	4.44	£84,484	6.02	£24,252	£14,393	£17,123
HR off treatment 0.727[a]	£46,962	3.83	£59,255	5.00	£81,865	6.55	£10,489	£14,582	£12,820
Imatinib resistance[b]	£46,962	3.83	£52,394	5.93	£76,279	7.33	£2,591	£17,018	£8,377
Lower mortality rates[c]	£65,251	4.76	£68,328	6.69	£87,715	7.90	£1,595	£16,112	£7,171

Combined analysis[d]	£67,603	4.53	£79,205	5.49	£102,087	6.25	£12,122	£29,966	£20,041

[a] HR shown for the off treatment period for 1-year adjuvant imatinib versus no treatment. [b] 15% of patients, initially treated with adjuvant imatinib, develop resistance [c] Using lower confidence interval estimates for post recurrence health states. [d] Analysis has no off treatment benefit after the end of trial, exponential distribution for the RFS, and lower mortality rates.

The combined analysis was run with a combination of the analyses already undertaken: i.e. no treatment benefit after the end of trial, exponential distribution for RFS, and for lower mortality rates. For the combined analysis the ICERs for 1-year adjuvant treatment versus no treatment increase to £12,122 per QALY and for 3-year treatment versus 1-year increase to £29,966 per QALY. The large increases in costs for 1-year and 3-year adjuvant imatinib are due to the increased time spent in the post recurrence health states due to the lower mortality rate and using the exponential distribution for RFS.

4.4 Summary of uncertainties and issues

There is substantial uncertainty over the methods used to derive clinical effectiveness parameters to populate the economic model. The MS has adopted methods to estimate the baseline risk of recurrence and relative treatment effects for adjuvant imatinib that avoid confounding by cross-over in the placebo arm of the ACOSOG Z9001 trial. However these methods required a great deal of post-hoc re-organisation of the trial data and it is uncertain whether these may have introduced other biases into the estimated effects. It may be more appropriate to use the cross-over-adjusted recurrence-free survival estimates to derive clinical effectiveness parameters for the ACOSOG Z9001 trial now that these are available.

There is also substantial uncertainty over the most appropriate assumptions for extrapolating the effectiveness of adjuvant imatinib beyond the trial data. Maximum follow-up in the RCTs providing clinical data (baseline and relative treatment effects for adjuvant imatinib) incorporated in the model is around nine years. However the model extrapolates these effects over a lifetime (40 year) horizon. The ERG has shown that choice of parametric form for the survival function and assumption over duration of benefit following adjuvant treatment have an impact on the cost effectiveness results. The ERG remains concerned that the MS has applied treatment effects derived using a semi-parametric model to fully parametric survival functions.

The manufacturer’s model, while generally appropriate, has defined health states on the basis of treatment and does not explicitly model disease progression. As a result some of the later

progressions in the model do not seem appropriate (for example, patients discontinuing treatment due to adverse events may transition to best supportive care (synonymous in the model with disease progression) without experiencing disease recurrence).

The ERG was unable to critically appraise the majority of the utility data included in the model, which comes from a single trial in patients with advanced GIST treated with sunitinib. The study reports no information on respondents, sample size, response rate or the valuation method adopted. The only information provided is that the EQ-5D questionnaire was used. The ERG cannot judge whether the utility values were derived using a population tariff or the VAS method and suggest, given the lack of methodological detail and the absence of information on respondents, that these data should be treated with caution.

5 End of life

NICE end of life treatment criteria were not included in the MS.

6 Innovation

The manufacturer notes that imatinib is a signal-transduction inhibitor designed to selectively inhibit certain classes of tyrosine kinase, including the receptor for stem cell factor coded for by the c-KIT proto-oncogene, which is expressed in more than 90% of GIST tumours. Imatinib binding to c-KIT protein that affects cell signalling, inhibits proliferation and induces apoptosis (MS section 4). It is suggested that the treatment can be beneficial in the group of patients considered at high risk of recurrence who would not otherwise be offered any adjuvant therapy.

7 DISCUSSION

7.1 Summary of clinical effectiveness issues

- The three included RCTs were generally well designed and executed, though in two of them there were changes in the primary outcome measure after randomisation, and two were open-label. None of the trials were conducted in the UK and their applicability to NHS practice and to the UK GIST population could be questioned. However, the ERG clinical

advisor did not consider there to be any clinically important differences between the trials and the UK population.

- The treatment effects for high risk patients in the MS are based on retrospective sub-population analyses, varying in the proportion of randomised patients classed as Miettinen high risk (the lowest being 28%), and are most likely underpowered. Differences between the arms of the trials at baseline were more pronounced in the Miettinen high risk sub-populations than the full populations, indicating selection bias.
- Five year RFS was longer for patients treated with 3-years adjuvant imatinib than for patients treated for 1-year. Clinical opinion suggests that the standard duration of adjuvant imatinib treatment in practice is now 3-years.[52] Some clinical guidelines recommend this duration, though the UK guidelines (last updated in 2009) don't currently recommend any specific treatment length. Clinical opinion also suggests that, based on the results of the clinical trials, adjuvant treatment generally delays recurrence rather than prevents it. There has therefore been interest in the effectiveness of longer-term adjuvant treatment, though there is no published RCT evidence yet to support treatment duration longer than 3-years. The phase II PERSIST-5 trial of 5-years adjuvant imatinib treatment is in progress and will complete in 2018 (see section 3.1.3).
- There were mixed results across the clinical effectiveness trials in terms of effects on OS. In the ACOSOG Z9001 trial there were few deaths overall, and at 2 and 5-years follow-up there was no statistically significant difference in OS. The manufacturer suggests that the 5-year data are confounded by the high degree of cross-over to imatinib by recurrence-free placebo patients when the study became unblinded. Additional analyses using different statistical methods for adjusting for cross-over in trials produced lower HRs for OS but the difference between trial arms remained non-statistically significant. The SSGXVIII/AIO trial was relatively smaller but evaluated a longer treatment period. The trial reported comparatively more deaths and at 5-years follow-up there was statistically significantly longer OS associated with 3-year imatinib treatment compared to 1-year. Neither of the trials was statistically powered for OS.

7.2 Summary of cost effectiveness issues

- The MS includes evidence on the cost-effectiveness of 1-year and 3-year adjuvant imatinib treatment compared to no treatment. The model structure and methods adopted for the economic evaluation are reasonable and generally appropriate. The model structure is consistent with the clinical disease pathways and available clinical trial evidence. However

the model structure, defining health states on the basis of treatment, results in some later progressions that do not seem appropriate. The MS provides evidence that the model has been validated against clinical trial data.

- There is uncertainty relating to the methods used to derive clinical effectiveness parameters to populate the economic model and the long term extrapolation of RFS in the manufacturer's model. The MS has assumed that patients continue to benefit from adjuvant imatinib after treatment has finished. Furthermore the parametric curves chosen for RFS assume that many patients remain recurrence-free after 20 years. These assumptions appear optimistic and are likely to produce results favourable to adjuvant imatinib.
- The majority of the utility data included in the model comes from a single trial, in patients with advanced GIST treated with sunitinib, which provides no information on participants, sample size, response rate or the valuation method adopted to derive the utilities. The only information provided is that the EQ-5D questionnaire was used. The ERG suggest, given the lack of methodological detail and the absence of information on respondents, that these data should be treated with caution.

8 REFERENCES

(1) National Institute for Health and Care Excellence. Imatinib for the adjuvant treatment of gastrointestinal stromal tumours. NICE technology appraisal guidance TA196. 2010.

(2) Blay JY, Le CA, Cassier PA, Ray-Coquard IL. Gastrointestinal stromal tumors (GIST): a rare entity, a tumor model for personalized therapy, and yet ten different molecular subtypes. *Discov Med* 2012; 13(72):357-367.

(3) Ahmed I, Welch NT, Parsons SL. Gastrointestinal stromal tumours (GIST) - 17 years experience from Mid Trent Region (United Kingdom). *Eur J Surg Oncol* 2008; 34(4):445-449.

(4) MacDuff E, Walsh S, Reid R, The Royal College of Pathologists. Standards and datasets for reporting cancers. Dataset for gastrointestinal stromal tumours (GISTs). http://www.rcpath.org/Resources/RCPath/Migrated%20Resources/Documents/G/G103_DatasetforGISTS_Feb12.pdf . 2012. 25-2-2014.

(5) Reid R, Bulusu R, Carroll N, Eatock M, Geh I, Judson I et al. Guidelines for the management of gastrointestinal stromal tumours (GIST). http://www augis org/pdf/GIST_Management_Guidelines_180809 pdf [2009

(6) Joint Formulary Committee. British National Formulary 61. London: British Medical Association and Royal Pharmaceutical Society of Great Britain; 2011.

(7) Fletcher CD, Berman JJ, Corless C, Gorstein F, Lasota J, Longley BJ et al. Diagnosis of gastrointestinal stromal tumors: A consensus approach. *Hum Pathol* 2002; 33(5):459-465.

(8) Miettinen M, Lasota J. Gastrointestinal stromal tumors: pathology and prognosis at different sites. *Semin Diagn Pathol* 2006; 23(2):70-83.

(9) Muguruma K, Kurokawa Y, Tsujinaka T, Fujita J, Nakai T, Takabayashi A et al. Randomized phase II study of 6 versus 12 months of adjuvant imatinib for patients with intermediate-or high-risk GIST. *Journal of Clinical Oncology* 2013; Conference(var.pagings):4.

(10) Yamamoto K, Kurokawa Y, Tsujinaka T, Muguruma K, Nakai T, Fujita J et al. Multicentre randomized phase II trial of adjuvant imatinib for 6 versus 12 months in patients with intermediate or high risk GIST: Interim analysis results. *European Journal of Cancer* 2013; 49:S625.

(11) Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. Third edition. 2009. York Publishing Services Ltd., CRD.

(12) DeMatteo RP, Ballman KV, Antonescu CR, Maki RG, Pisters PW, Demetri GD et al. Adjuvant imatinib mesylate after resection of localised, primary gastrointestinal stromal tumour: a randomised, double-blind, placebo-controlled trial. *Lancet* 2009; 373(9669):1097-1104.

(13) Corless CL, Ballman KV, Antonescu C, Blanke CD, Blackstein ME, Demetri GD et al. Relation of tumor pathologic and molecular features to outcome after surgical resection of localized primary gastrointestinal stromal tumor (GIST): Results of the intergroup phase III trial ACOSOG Z9001. J Clin Oncol 28[15], (Suppl; abstr 10006). 2010.

(14) Data On File: Clinical Study Report 5-Year Follow-up for the ACOSOG trial. 2014.

(15) ACOSOG Z9001 Tables and Listings for High Risk Patients, Analysis performed by Novartis. Commercial in Confidence. 2014.

(16) Joensuu H, Eriksson M, Sundby HK, Hartmann JT, Pink D, Schutte J et al. One vs three years of adjuvant imatinib for operable gastrointestinal stromal tumor: a randomized trial. *JAMA* 2012; 307(12):1265-1272.

(17) Heintz D, Noronha-Jackson A, Green S, Scheuer N, Calderwood J, Gallagher N. Short (12 months) versus long (36 months) duration of adjuvant treatment with the tyrosine kinase inhibitor imatinib mesylate of perable GIST with a high risk for recurrence: A randomised Phase III study. Clinical Study Report STI571BF103, editor. 2011.

(18) Casali PG, Le Cesne A, Poveda Velasco A, et al. Imatinib failure-free survival (IFS) in patients with localized gastrointestinal stromal tumors (GIST) treated with adjuvant imatinib (IM): The EORTC/AGITG/FSG/GEIS/ISG randomized controlled phase III trial. ASCO Annual Meeting. Clin Oncol 31[(Suppl; abstr 10500)]. 2013.

(19) Conely AP, Guerin A, Sasane M, Gauthier G, Schwiep F, Hunt Keir C et al. Risk of disease recurrence and mortality in gastrointestinal stromal tumor (GIST) patients receiving short-term versus long-term imatinib adjuvant therapy: A chart review analysis. J Clin Oncol 30[(Suppl; abstr e20511)]. 2012.

(20) Hatoum HT, Lin SJ, Sasane M, Trent JC. Effectiveness of adjuvant imatinib in patients with gastrointestinal stromal tumor: results of a population-based, matched-cohort study. *Curr Med Res Opin* 2012; 28(5):805-814.

(21) Wu XJ, Zhang RX, Chen G, Fang YJ, Ding PR, Li LR et al. Prognostic factors in surgically managed gastrointestinal stromal tumours. *J Int Med Res* 2012; 40(5):1904-1911.

(22) McCarter MD, Antonescu CR, Ballman KV, Maki RG, Pisters PW, Demetri GD et al. Microscopically positive margins for primary gastrointestinal stromal tumors: analysis of risk factors and tumor recurrence. *J Am Coll Surg* 2012; 215(1):53-59.

(23) Nilsson BE, Sjolund K, Nilsson EP, Wangberg B, Nilsson O, Ahlman H. Neoadjuvant, adjuvant, and palliative treatment with imatinib in patients with gastrointestinal stromal tumors. J Clin Oncol 15[(Suppl; abstr 10087)]. 2014.

(24) Kang YK, Kang BW, Im SA, Lee JL, Park SR, Kang WK et al. Two-year adjuvant imatinib mesylate after complete resection of localized, high-risk GIST with KIT exon 11 mutation. *Cancer Chemother Pharmacol* 2013; 71(1):43-51.

(25) Li J, Gong JF, Wu AW, Shen L. Post-operative imatinib in patients with intermediate or high risk gastrointestinal stromal tumor. *Eur J Surg Oncol* 2011; 37(4):319-324.

(26) Li J, Gong J, Li J, Wu WA, Shen L. Adjuvant therapy with imatinib in gastrointestinal stromal tumor (GIST) patients with intermediate or high risk: Analysis from a single-center contrast study. J Clin Oncol 27[(Suppl; abstr 10556)]. 2009.

(27) Jiang WZ, Guan GX, Lu HS, Yang YH, Kang DY, Huang HG. Adjuvant imatinib treatment after R0 resection for patients with high-risk gastrointestinal stromal tumors: a median follow-up of 44 months. *J Surg Oncol* 2011; 104(7):760-764.

(28) Nishida T, Kanda T, Wada N, Kobayashi O, Yamamoto M, Sawaki A et al. Phase II trial of adjuvant imatinib mesylate after resection of localized, primary high risk gastrointestinal stromal tumour (GIST) in Japan. Eur J Cancer 7[(Suppl; abstr 9413)]. 2009.

(29) Yalcin S, Buyukberber S, Yilmaz U, Lee P-H, Srimuninnimit V, Lin P-W et al. Phase II study of adjuvant imatinib mesylate (IM) in patients after resection of primary gastrointestinal stromal tumor (GIST): Efficacy and safety at one year. J Clin Oncol 30[15], (Suppl; abstr e20514). 2012.

(30) Kanda T, Nishida T, Wada N, Kobayashi O, Yamamoto M, Sawaki A et al. Adjuvant therapy with imatinib mesylate after resection of primary high-risk gastrointestinal stromal tumors in Japanese patients. *Int J Clin Oncol* 2013; 18(1):38-45.

(31) Blay JY, Le CA, Ray-Coquard I, Bui B, Duffaud F, Delbaldo C et al. Prospective multicentric randomized phase III study of imatinib in patients with advanced gastrointestinal stromal tumors comparing interruption versus continuation of treatment beyond 1 year: the French Sarcoma Group. *J Clin Oncol* 2007; 25(9):1107-1113.

(32) SCHOENFELD D. Partial residuals for the proportional hazards regression model. *Biometrika* 1982; 69(1):239-241.

(33) GRAMBSCH PM, THERNEAU TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994; 81(3):515-526.

(34) Latimer N, Ward S. Adjusting for treatment switching in a Phase 3 trial of imatinib for the adjuvant treatment of adult patients who are at high risk of relapse following resection of kit (CD117)-positive gastrointestinal stromal tumours (GIST). 2014. School of Health and Related Research (ScHARR), University of Sheffield.

(35) Robins JM, Tsiatis AA. Correcting for Noncompliance in Randomized Trials Using Rank Preserving Structural Failure Time Models. *Communications in Statistics-Theory and Methods* 1991; 20(8):2609-2631.

(36) Branson M, Whitehead J. Estimating a treatment effect in survival studies in which patients switch treatment. *Stat Med* 2002; 21(17):2449-2463.

(37) Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* 2000; 56(3):779-788.

(38) Majer IM, Gelderblom H, van den Hout WB, Gray E, Verheggen BG. Cost-effectiveness of 3-year vs 1-year adjuvant therapy with imatinib in patients with high risk of gastrointestinal stromal tumour recurrence in the Netherlands; a modelling study alongside the SSGXVIII/AIO trial. *J Med Econ* 2013; 16(9):1106-1119.

(39) Sanon M, Taylor DC, Parthan A, Coombs J, Paolantonio M, Sasane M. Cost-effectiveness of 3-years of adjuvant imatinib in gastrointestinal stromal tumors (GIST) in the United States. *J Med Econ* 2013; 16(1):150-159.

(40) Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996; 313(7052):275-283.

(41) National Institute for Health and Care Excellence. Imatinib for the treatment of unresectable and/or metastatic gastro-intestinal stromal tumours. NICE technology appraisal guidance TA86. 2010. London, NICE.

(42) Novartis Pharma GmbH. Summary of product characteristics - Glivec 50 mg hard capsules. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000406/WC500022207.pdf . 2014. 4-4-2014.

(43) Rutkowski P, Bylina E, Wozniak A, et al. Validation of the Joensuu risk criteria for primary resectable gastrointestinal stromal tumour—the impact of tumour rupture on patient outcomes. *Eur J Surg Oncol* 2011; 37:890-896.

(44) Joensuu H. Risk stratification of patients diagnosed with gastrointestinal stromal tumor. *Hum Pathol* 2008; 39:1411-1419.

(45) Verweij J, Casali P, Zalcberg J. Progression-free survival in gastrointestinal stromal tumours with high-dose imatinib: randomised trial. *Lancet* 2004; 364:1127-1134.

(46) Demetri G, van Oosterom A, Garrett C. Efficacy and safety of sunitinib in patient with advanced gastrointestinal stromal tumour after failure of imatinib: a randomised controlled trial. *Lancet* 2006; 368:1329-1338.

(47) Demetri GD, Garrett C, Schoffski P, et al. Complete longitudinal analysis of the randomised, placebo-controlled, phase III trial of sunitinib in patients with gastrointestinal stromal tumour following imatinib failure. *Clinical Cancer Research* 2012; 18(3):3170-3179.

(48) Tran T, Davilla J, El-Serag H. The epidemiology of malignant gfastrointestinal stromal tumors: ana analysis of 1,458 cases from 1992 to 2000. *American Journal of Gastroenterology* 2005; 100:162-168.

(49) Huse D, von Mehren M, Lenhart G, Joensuu H, Blanke C, Feng W et al. Cost effectiveness of imatinib mesylate in the treatment of advanced gastrointestinal stromal tumours. *Clin Drug Invest* 2007; 27(2):85-93.

(50) Miettinen M, Sobin LH, Lasota J. Gastrointestinal stromal tumors of the stomach: a clinicopathologic, immunohistochemical, and molecular genetic study of 1765 cases with long-term follow-up. *Am J Surg Pathol* 2005; 29(1):52-68.

(51) Reichardt P, Hartmann JT, Sundby Hall K. Response to Imatinib Rechallenge of GIST That Recurs Following Completion of Adjuvant Imatinib Treatment - the First Analysis in the SSGXVIII/AIO Trial Patient Population. *European Journal of Cancer* 2011; 47(Supp 2:15).

(52) Blay JY, Perol D, Le CA. Imatinib rechallenge in patients with advanced gastrointestinal stromal tumors. *Ann Oncol* 2012; 23(7):1659-1665.

(53) Chabot I, LeLorier J, Blackstein M. The challenge of conducting pharmacoeconomic evaluations in oncology using crossover trials: the example of sunitinib for gastrointestinal stromal tumours. *European Journal of Cancer* 2008; 44:972-977.

(54) Wilson J, Connock M, Song F. Imatinib for the treatment of patients with unresectable and/or metastatic gastrointestinal stromal tumours: systematic review and economic evaluation. *Health Technology Assessment* 2005; 9:1-142.

(55) Sanon M, Taylor D, Parthan A, Coombs J, Paolantonio M, Sasane M. Cost-effectiveness of 3-years of adjuvant imatinib in gastrointestinal stromal tumors (GIST) in the United States. *J Med Econ* 2013; 16:150-159.

(56) Ara R, Brazier J. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010; 13:509-518.

(57) Joint Health Surveys Unit of Social and Community PlanningResearch and University College London. Health Survey for England. 2003. 2004. HMSO.

(58) Joint Health Surveys Unit of Social and Community PlanningResearch and University College London. Health Survey for England. 2006. 2007. HMSO.

(59) Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. *Health Econ* 1996; 5:141-154.

(60) Department of Health. NHS Reference Costs 2011/2012. https://www gov uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012

(61) Curtis L. Unit Costs of Health and Social Care 2012. http://www.pssru.ac.uk/project-pages/unit-costs/2012/