ERG STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Midostaurin for treating advanced systemic mastocytosis [ID1573]

Confidential until published

This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR 131327

Completed 15th May 2020

CONTAINS [redacted] AND [redacted] DATA

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Title: Midostaurin for treating advanced systemic mastocytosis [ID1573]

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

Authors: Janette Greenhalgh, Senior Research Fellow (Clinical Effectiveness), LR*i*G, University of Liverpool

Tosin Lambe, Research Associate (Health Economics), LR*i*G, University of Liverpool

Sarah Nevitt, Research Associate (Medical Statistician), LR*i*G, University of Liverpool

James Mahon, Director, Coldingham Analytical Services, Berwickshire

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

Sophie Beale, Senior Associate Researcher, LR*i*G, University of Liverpool

Yenal Dundar, Research Fellow (Information Specialist), LR*i*G, University of Liverpool

Katherine Edwards, Research Associate (Clinical Effectiveness), LR*i*G, University of Liverpool

Rebecca Bresnahan, Research Associate (Clinical Effectiveness), LR*i*G, University of Liverpool

Joanne McEntee, Senior Medicines Information Pharmacist, North West Medicines Information Centre, Liverpool

Mufaddal Moonim, Consultant Histopathologist, Guy's and St Thomas' NHS Foundation Trust, London

Correspondence to: Janette Greenhalgh, Senior Research Fellow, Liverpool Reviews and Implementation Group, University of Liverpool, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 15th May 2020.

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number 131327.

Acknowledgements: The authors would like to thank Dr Clare Oni, Clinical Fellow, Department of Haematology, Guy's and St Thomas' NHS Foundation Trust, London and Dr Rui Duarte, Deputy Director, LR*i*G, University of Liverpool who provided feedback on a draft version of the report.

Rider on responsibility for report: The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

Declared competing interests of the authors: Dr Moonim participated in an event, hosted at Guy's hospital, that was sponsored by Novartis. Dr Oni is involved in running a clinical trial of avapritinib for systemic mastocytosis. The Department of Haematology at Guy's and St Thomas' is in receipt of funding for her salary from the company who sponsors avapritinib (Blueprint Medicines).

This report should be referenced as follows: Greenhalgh J, Lambe T, Nevitt S, Mahon J, Boland A, Beale S, Dundar Y, Edwards K, Bresnahan R, McEntee, J, Moonim MT. Midostaurin for treating advanced systemic mastocytosis [ID1573]: A Single Technology Appraisal. LR*i*G, University of Liverpool, 2020

Contributions of authors:

Janette Greenhalgh	Project lead, critical appraisal of the clinical evidence and supervised the production of the final report
Tosin Lambe	Critical appraisal of the economic evidence and the economic model
Sarah Nevitt	Critical appraisal of the statistical evidence
James Mahon	Critical appraisal of the economic model
Angela Boland	Critical appraisal of the clinical and economic evidence, editorial input
Sophie Beale	Critical appraisal of the clinical and economic evidence, editorial input
Yenal Dundar	Critical appraisal of the adverse event data and cross checking of the company search strategies
Katherine Edwards	Critical appraisal of the clinical evidence
Rebecca Bresnahan	Critical appraisal of the clinical evidence
Joanne McEntee	Critical appraisal of the company submission
Mufaddal Moonim	Clinical advice and critical appraisal of the clinical sections of the company submission

Table of contents

List of tables

List of figures

List of abbreviations

AE	adverse event
AESI	adverse event of special interest
allo-HSCT	allogenic haematopoietic stem cell transplant
AML	acute myeloid leukaemia
ASM	aggressive systemic mastocytosis
CCM	current clinical management
CI	confidence interval
CS	company submission
CSR	clinical study report
DoR	duration of response
ECNM	European Competence Network on Mastocytosis
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
eMIT	electronic Market Information Tool
EPAR	European Public Assessment Report
EQ-5D-3L	EuroQol-5 Dimensions-3 Levels
ERG	Evidence Review Group
FAS	final analysis set
GDI	global distress index
HR	hazard ratio
HRG	Healthcare Resource Group
HRQoL	health-related quality of life
HSCT	haematopoietic stem cell transplant
HST	highly specialised technology
ICER	incremental cost effectiveness ratio
ICU	intensive care unit
ITT	intention-to-treat
IWG-MRT	International Working Group-Myeloproliferative Neoplasms Research and Treatment
K-M	Kaplan-Meier
MCL	mast cell leukaemia
MCS	mental composite score
mL	millilitre
MR	major response
MRI	magnetic resonance imaging
MSAS	Memorial Symptom Assessment Scale
NA	not applicable
NE	not estimated
NR	not reported
NSAID	non-steroidal anti-inflammatory drugs
ONS	Office for National Statistics
ORR	overall response rate
OS	overall survival
OWSA	one-way sensitivity analysis
PAS	Patient Access Scheme
PCS	physical composite score
PD	progressed disease
PEP	primary efficacy population
PF-no-response	progression-free without response to treatment
PF-response	progression-free with response to treatment
PFS	progression-free survival
PHYS	physical symptom subscale
PPS	per protocol set
PR	partial response
PSSRU	Personal Social Services Research Unit

PSYCH	psychological symptom subscale
QALY	quality adjusted life year
RCT	randomised controlled trial
RFS	relapse-free survival
SAE	serious adverse event
SES	safety evaluation set
SF-12	short form-12
SM	systemic mastocytosis
SM-AHN	systemic mastocytosis with an associated haematological neoplasm
SmPC	Summary of Product Characteristics
SoC	standard of care
STA	Single Technology Appraisal
TA	technology appraisal
TKI	tyrosine kinase inhibitor
TMSAS	Total Memorial Symptom Assessment Scale
WHO	World Health Organization
WTP	willingness-to-pay

1 EXECUTIVE SUMMARY

1.1 Scope of the submission

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost effectiveness evidence submitted to the National Institute for Health and Care Excellence (NICE) as part of the Single Technology Appraisal (STA) process. Clinical and economic evidence has been submitted to NICE by Novartis Pharmaceuticals UK Ltd in support of the use of midostaurin for the treatment of advanced systemic mastocytosis (SM). Systemic mastocytosis is a group of rare diseases in which uncontrolled growth and accumulation of mast cells (a type of white blood cell) occur in one or more organs.

Midostaurin was granted marketing authorisation by the European Medicines Agency (EMA) as a monotherapy for the treatment of adult patients with advanced SM in September 2017. It is the only treatment licensed in Europe for the treatment of advanced SM. Further, there are no UK clinical guidelines for the treatment of advanced SM. Clinical advice is that in the NHS treatment is tailored to the symptoms and needs of individual patients.

1.2 Critique of the decision problem in the company submission

1.2.1 Population and intervention

As highlighted in Section 2.5 of this ERG report, the decision problem addressed by the company matches the final scope issued by NICE in terms of intervention (midostaurin) and population (adults with advanced SM). In the company submission (CS), the company has estimated that the number of patients in England eligible for treatment with midostaurin is 174 (Section 2.4 of this ERG report). The company has presented clinical evidence for the whole advanced SM population and separately for the three subtypes: aggressive systemic mastocytosis (ASM), systemic mastocytosis with associated haematological neoplasm (SM-AHN) and mast cell leukaemia (MCL). However, the company highlighted that, due to small numbers, there was considerable uncertainty around the reliability of the clinical results for each subtype. Clinical advice to the company was that an exploratory analysis that combined health outcome data from patients with SM-AHN and MCL in the D2201 trial was reasonable, based on the rationale that patients with SM-AHN and MCL have a much shorter life expectancy than patients with ASM.

The evidence to support the clinical effectiveness of midostaurin (licensed dose) was generated by two single arm, open label, phase II trials (D2201 [n=116] and A2213 [n=26]). Clinical advice to the ERG was that the baseline characteristics of patients participating in these trials were similar to the characteristics of patients treated in the NHS with advanced

SM. However, the results from these trials are difficult to interpret because of the open-label design, the lack of a comparator arm and small patient numbers for each disease subtype. Further, the positioning of midostaurin in the treatment pathway is not explicitly stated by the company; patients recruited to the D2201 trial had received between zero and four prior therapies, whilst those participating in the A2213 trial had received between zero and three prior therapies.

1.2.2 Comparators

Five comparators were listed in the final scope issued by NICE. Clinical advice to the company was that only three of these comparators were relevant (cladribine, interferon alpha and imatinib) and that the other two comparators listed in the final scope issued by NICE (nilotinib and dasatinib) were rarely used in the UK. However, clinical advice to the company was that pegylated-interferon alpha and acute myeloid leukaemia (AML) -like treatments were also relevant comparators. Clinical advice to the ERG supports the clinical advice provided to the company.

1.2.3 Outcomes

The outcomes listed in the final scope issued by NICE are overall survival (OS), progression-free survival (PFS), response rate, adverse events (AEs) and health-related quality of life (HRQoL). Clinical advice to the ERG is that these are important outcomes for patients with advanced SM. Whilst results from the D2201 and the A2213 trials provide information about the effectiveness of midostaurin, there is no randomised clinical trial evidence to support the clinical effectiveness of the three relevant comparators listed in the final scope issued by NICE or the two additional comparators identified by the company. Due to the limited data available, it was not possible for the company to carry out any indirect comparisons.

Overall survival

The company identified two comparisons (by Reiter et al and by Chandesris et al) that generated OS results for patients with advanced SM treated with midostaurin versus patients treated with other (unspecified) drugs.

Reiter et al compared pooled D2201 and A2213 trial data (n=89) with data from a German registry (n=42) and reported comparative (midostaurin versus unspecified treatments) OS hazard ratios (HRs) that favour treatment with midostaurin. However, the ERG has concerns about whether the D2201 and A2213 trial data should have been pooled (Section 3.6 of this ERG report). In addition, the ERG has concerns relating to the inputs (midostaurin data and

German registry data) and notes the differences between the results presented in the published abstract and those provided in the unpublished presentation.

Chandesris et al compared data from a cohort of patients receiving midostaurin in a French compassionate use programme (n=28) versus French registry data (n=44). The ERG highlights the small numbers of patients and the differences between the French cohorts and the patients recruited to the D2201 and A2213 trials (e.g., these trials did not include any patients with mast cell sarcoma or progressive smouldering SM). The ERG also highlights uncertainty (due to a lack of published information) around the methods used by Chandesris et al. As a consequence of these issues, the ERG considers that the results reported by Chandesris et al should not be used to inform decision making

Adverse events

The ERG agrees with the company that it is difficult to establish whether the Grade 3 or 4 haematological AEs reported in the D2201 and A2213 trials were related to treatment with midostaurin or to disease progression (Section 3.5 of this ERG report). Clinical advice to the ERG is that AEs arising from treatment with midostaurin, as with current unlicensed treatments for advanced SM, require careful monitoring by a specialist clinical team with the experience to provide early recognition and management of cytoreductive therapy-related AEs, and that this can place a high burden on NHS staff and systems.

1.2.4 Other issues

The company's economic analyses were carried out in line with the final scope issued by NICE. Midostaurin is currently available to the NHS at a discounted Patient Access Scheme (PAS) price. However, the company [redacted]. Azacitidine (included in the company's basket of comparator treatments) is also available to the NHS at a discounted PAS price. However, this price is confidential and not known to the company.

1.3 Summary of the ERG's critique of the submitted cost effectiveness evidence

In Section 4.6 of this ERG report, for the comparison of treatment with midostaurin versus current clinical management (CCM), the ERG has identified four major areas of concern relating to the company model:

- OS HR
- PFS for midostaurin versus current CCM
- partitioning survival outcomes
- lifetime duration of the treatment effect of midostaurin.

The overall survival hazard ratio

The source of the OS HR used in the company model is the Reiter et al unpublished presentation. The company sought expert clinical advice to help them identify the most appropriate OS HR; however, whilst the Reiter et al mulitivariable result was considered to be the most plausible, the range over which clinical experts considered the true OS HR might lie was very wide.

The company's base case incremental cost effectiveness ratio (ICER) per quality adjusted life year (QALY) gained for the comparison of the cost effectiveness of midostaurin versus CCM was £[redacted]. Results from the company's deterministic sensitivity analyses showed that using the upper and lower 95% confidence interval OS HR estimates (0.319 and 0.839 respectively) generated ICERs per QALY gained for the comparison of the cost effectiveness of midostaurin versus CCM of £[redacted] and £[redacted] respectively, demonstrating that the OS HR is the key driver of cost effectiveness results. Thus, without a robust and accurate OS HR estimate, it is not possible to produce reliable cost effectiveness results. The ERG has not been able to identify a robust source for the OS HR and, therefore, has not been able to generate preferred cost effectiveness results.

As discussed in Section 4.7 of this ERG report, the OS HR is so important to the base case cost effectiveness results that without a level of certainty around this estimate, discussions about other model-related concerns are largely academic. The ERG has indicated whether improved modelling of these effects would be likely to increase or decrease the cost effectiveness of midostaurin versus CCM.

Progression-free survival for midostaurin versus current clinical management

The company was not able to identify any comparative PFS data. Therefore, based on clinical advice, the company set the PFS HR equal to the OS HR. There is no way of knowing whether this assumption is reasonable. However, even if the approach were valid, given the uncertainty around the magnitude of the OS HR, the PFS HR would be similarly uncertain. It is not known whether setting the PFS HR equal to the OS HR leads to an under- or over-estimate of the true PFS experience of patients receiving CCM. Thus, it is not known whether using the OS HR as a proxy for the PFS HR leads to an under- or over-estimate of the ICER per QALY gained for the comparison of midostaurin versus CCM.

Partitioning progression-free survival

The PF health state in the company model is partitioned into a PF-response health state and a PF-no-response health state to reflect the assumption that HRQoL differs between responders and non-responders. The ERG has concerns about the reliability of the overall response rates and duration of response estimates used in the company model and therefore considers that it was not appropriate to use these estimates to partition PFS (nor would it have been appropriate to use them to partition OS).

Lifetime duration of midostaurin treatment effect

In the D2201 trial, only 19% of patients were still receiving midostaurin at 3 years; however, the treatment benefit attributed to receipt of midostaurin was modelled to persist over the 38-year model time horizon. Given the proportionally short time frame over which patients received midostaurin, the ERG considers that it would be clinically implausible to assume patients benefited from this treatment for the whole model time horizon. Rather, the ERG considers that at some point before 38 years, it is likely that the progression and mortality rates of patients initially assigned treatment to midostaurin and CCM would become equal. The effect of equalising progression and mortality rates at some point during the model time horizon on the comparative cost effectiveness of midostaurin versus CCM would increase the size of the ICER per QALY gained.

1.3.1 Summary of company's case for NICE End of Life criteria being met

A technology meets NICE End of Life criteria if (i) life expectancy with standard of care treatments for the target population is under 24 months and (ii) the increase in life expectancy with the technology being appraised is at least 3 months. The company considers that midostaurin meets both criteria and should be considered as an End of Life treatment.

1.3.2 Short life expectancy (normally ≤24 months)

The life expectancy of patients with advanced SM varies significantly across disease subtypes. Published median survival times from diagnosis range from 41 months to 11 years for patients with ASM, from 24 months to 4.4 years for patients with SM-AHN and from 2 months to 9.2 months for patients with MCL. The company has suggested that published life expectancy estimates for patients with SM-AHN may be too high as they include unknown proportions of patients with indolent SM (ISM-AHN). Patients with ISM-AHN have a longer life expectancy than patients with SM-AHN and are not included in the population being considered in this appraisal.

Median OS reported by Reiter et al for a cohort of patients with ASM, SM-AHN and MCL who received treatment other than midostaurin was 19.5 months. Further, clinical advice to the company was that patients with advanced SM would normally have a life expectancy of less than 24 months and life expectancy would be even lower in the SM-AHN+MCL combined subgroup.

1.3.3 Life extension (normally ≥3 months)

Results generated by the company model and by Reiter et al suggest that treatment with midostaurin probably offers a life extension of ≥3 months when used to treat patients with advanced SM and also when used to treat the SM-AHN+MCL combined subgroup (see Table).

Estimated life extension as a consequence of receiving midostaurin

Population	Midostaurin versus CCM: life extension	
	Company model: mean	Reiter et al: median
Advanced SM	[redacted] months	21.9 months
SM-AHN+MCL subgroup	[redacted] months	NR

CCM=current clinical management; MCL=mast cell leukaemia; NR=not reported; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with an associated haematological neoplasm
Source: Company model and Reiter et al

1.4 ERG commentary on NICE End of Life criteria

Due to concerns about the reliability of available evidence, the ERG has not generated any preferred or alternative cost effectiveness results (Section 4.7 of this ERG report). This means that the ERG has been unable to present alternative survival estimates for midostaurin or CCM.

1.4.1 Short life expectancy (normally ≤24 months)

Given disease subtype is determined before treatment commences and life expectancy ranges are wide for each subtype, the ERG considers that applying End of Life criteria to the whole advanced SM population is not appropriate.

The published subtype evidence presented by the company suggests that the short life expectancy criterion is not met for patients with ASM, is possible (but unlikely) to be met for patients with SM-AHN and is met for patients with MCL. Assessing short life expectancy for the combined subgroup (SM-AHN+MCL) is therefore problematic.

1.4.2 Life extension (normally ≥3 months)

Whilst the results presented by the company demonstrate that treatment with midostaurin extends life by ≥3 months compared with treatment with CCM, concerns relating to the OS HR used to generate the overall and combined subgroup results cast considerable doubt over their validity.

1.5 Summary of exploratory and sensitivity analyses undertaken by the ERG

Due to the uncertainty around the company's OS HR estimate (the key driver of model cost effectiveness results), the ERG has not generated any preferred or alternative cost effectiveness results. Also, as the ERG considers that the company's assumption that the effect of treatment with midostaurin on OS and PFS lasts a lifetime is optimistic, this may mean that the company base case cost effectiveness results for the comparison of midostaurin versus CCM are underestimates.

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

The focus of this appraisal is on the use of midostaurin to treat adult patients with advanced systemic mastocytosis (SM).[1]

Mastocytosis encompasses a heterogenous group of rare diseases.[2,3] Advanced SM is the most severe form of mastocytosis and, following diagnosis, patients are classified as having one of three subtypes: aggressive systemic mastocytosis (ASM), systemic mastocytosis with an associated haematological neoplasm (SM-AHN) or mast cell leukaemia (MCL).[3] In Europe, advanced SM is a rare disease with a reported prevalence of 0.06 per 100,000 population across all subtypes; however, the exact prevalence in the UK is not known.[4] The disease mostly occurs in Caucasian adults (males and females) and is most frequently diagnosed in people over 60 years old.[5]

In advanced SM, mast cells infiltrate various tissues and organs, commonly the bone marrow, spleen, liver, lymph nodes and gastrointestinal tract.[3] Advanced SM is associated with a wide range of physical symptoms related to increased mast cell proliferation and activity.[3,6,7] Symptoms can include fatigue, itching, bone or muscle pain, osteoporosis, fractures or anaphylactic reactions. Patients may present with organ dysfunction (for example, organomegaly, organopathy or organ failure) which arises from mast cell accumulation within organs.[8]

For the majority of patients with advanced SM, treatment involves symptom control alongside cytoreductive therapy. Cytoreductive therapy aims to reduce the underlying mast cell burden and alleviate disease-related organ dysfunction.[9] However, UK clinical experts estimate that one third of patients with advanced SM are unable to receive cytoreductive therapy due to frailty or risk of cytopenia.[10] For these patients, disease management involves symptom control alongside supportive or palliative care. For patients with SM-AHN, the treatment pathway is influenced by the nature and severity of the associated haematological neoplasm. Clinical advice to the company was that the associated haematological neoplasm is treated separately and may be prioritised over treatment for the SM component of the disease.[11]

Life expectancy for patients with advanced SM is just under 24 months from time of diagnosis, though this varies by disease subtype.[12] Patients with MCL have the shortest life expectancy, ranging from less than 2 months to 9.2 months.[7,13] For patients with SM-AHN, life expectancy ranges from 24 months to 4.4 years,[4,7] and for patients with ASM life expectancy ranges from

41 months to 11 years.[7,14] However, the company has suggested that published life expectancy estimates for patients with SM-AHN may be too high as they include unknown proportions of patients with indolent SM. Advanced SM also has a negative impact on patient and carer health-related quality of life (HRQoL) due to symptom burden.[15,16]

2.2 Company's overview of current service provision

There is currently no established treatment pathway for patients with advanced SM; however, clinical advice to the company is that local UK-specific guidelines are expected to be published in 2021. In the company submission (CS), the company states that only three centres in England manage patients with advanced SM (CS, p15). The company has devised the treatment algorithm displayed in Figure 1 using the National Comprehensive Cancer Network[17] guidelines and feedback from their clinical advisors.

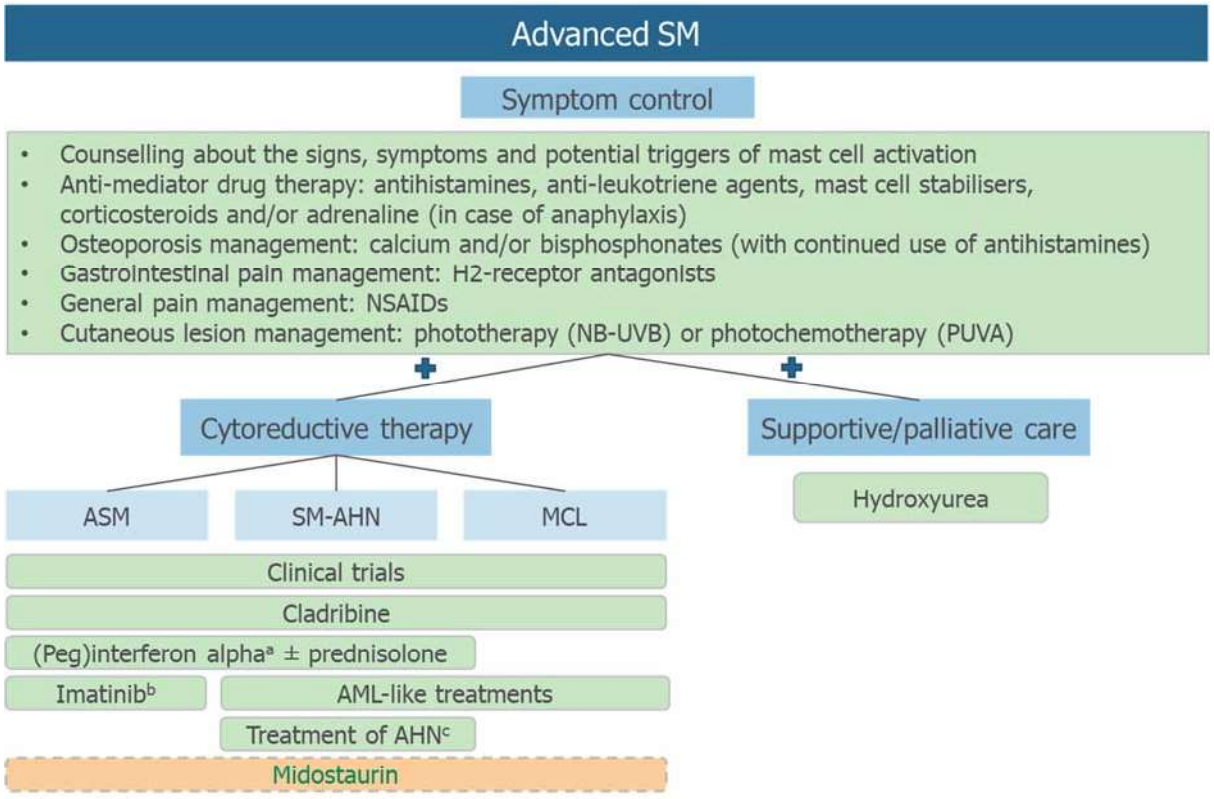

Figure 1 Anticipated pathway of care for patients with advanced SM in the UK

[a] Pegylated or un-pegylated interferon alpha with or without prednisolone
[b] Only if KIT D816V mutation negative or if eosinophilia is present with FIP1L1-PDGFRA fusion gene
[c] For patients with SM-AHN, only if SM component treatment takes precedence over AHN component treatment

AML=acute myeloid leukaemia; ASM=aggressive systemic mastocytosis; HSCT=haematopoietic stem cell transplant; MCL=mast cell leukaemia; NB-UVB=narrowband ultraviolet B; NSAIDs=non-steroidal anti-inflammatory drugs; PDGFRA=platelet dependent growth factor receptor A; PUVA=psoralen plus ultraviolet A; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with associated haematologic neoplasm
Source: CS, Figure 3

In the NHS, treatment is selected based on individual patient symptoms.[17,18] The company highlights that the diagnosis of advanced SM is challenging, given that is it is both a rare disease and has a range of non-specific symptoms. Clinical advice to the Evidence Review Group (ERG) is that time from symptom presentation to diagnosis can vary by subtype and that patients with SM-AHN are often diagnosed more quickly than patients with ASM or MCL due to the recognisable involvement of the associated haematological neoplasm.

The only curative treatment is allogenic haematopoietic stem cell transplant (allo-HSCT). Only a few patients are suitable for transplant; the vast majority are treated with cytoreductive therapy in combination with symptom management. Current cytoreductive treatment options for all patients with advanced SM are listed in the final scope[1] issued by NICE (i.e., interferon alpha, cladribine, imatinib, nilotinib and dasatinib).

Choice of cytoreductive treatment is largely patient-specific, depending on symptoms. Clinical advice to the ERG is that cladribine is usually the first treatment offered to patients. Imatinib is used only to treat patients with a sensitising mutation or a wild type mutation of advanced SM. As noted by the company, nilotinib and dasatinib are not commonly used in UK clinical practice. None of the comparator treatments listed in the final scope[1] issued by NICE has a European marketing authorisation for the treatment of advanced SM.

Following treatment with a cytoreductive therapy, symptoms are continually reassessed in line with the progressive nature of advanced SM. Treatment failure is defined as a lack of response or disease progression; on treatment failure, patients are reassessed and may receive an alternative form of cytoreductive therapy.

Other treatment options that may be considered for some patients include pegylated interferon alpha and AML-like (acute myeloid leukaemia) treatments. The company states that treatment with pegylated or un-pegylated interferon alpha has led to only minor or partial response, with best response at 1 year or more. Interferon (pegylated or un-pegylated) is also poorly tolerated, leading to toxicities in up to 75% of patients and a high dropout rate.[2,9,19] Peg-interferon alpha is more commonly used in UK clinical practice.[10]

For patients who are unable to be treated with cytoreductive therapies due to frailty or high risk of cytopenia, supportive or palliative care is provided.[10] This may include treatment with hydroxyurea to reduce mast cell burden alongside standard palliative treatments.[10]

2.3 Midostaurin

A summary of the details of the mechanism of action and the European marketing authorisation for midostaurin are shown in Box 1. The company highlights (CS, p15) that midostaurin is the only medicine to currently have a European marketing authorisation for adult patients with advanced SM.

Box 1 Midostaurin

- Midostaurin is an inhibitor to several receptor tyrosine kinases, including a mast/stem cell growth factor receptor (also known as *KIT* or CD117).[20] Up to 96% of patients with advanced SM have an active mutation in the *KIT* gene which results in mast cell proliferation and growth.[21] Midostaurin inhibits signalling in the *KIT* receptor to decrease mast cell production.[22]
- On 18th September 2017, the European Medicines Agency (EMA) granted marketing authorisation to midostaurin (Rydapt®) for use as a monotherapy for the treatment of adult patients with advanced SM (ASM, SM-AHN and MCL).[23]
- Midostaurin has also received orphan designation from the EMA for the treatment of adult patients with advanced SM.[24]
- Midostaurin is an oral therapy with a recommended dose of 100mg twice daily for patients with advanced SM.[20]

ASM=aggressive systemic mastocytosis; EMA=European Medicines Agency; MCL=mast cell leukaemia; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with associated haematologic neoplasm
Source: adapted from CS, Table 2 and CS, p15

2.4 Number of patients eligible for treatment with midostaurin

Given the rarity of the disease, the exact incidence of advanced SM is not known.[25] The company has used Danish incidence data[4] from a cohort of 547 patients diagnosed with advanced SM between 1997 and 2010, along with UK population figures from 2018,[26] to estimate the total number of patients (n=174) with advanced SM in England eligible for treatment with midostaurin. The estimated incidence and prevalence rates of advanced SM in Europe by disease subtype are presented in Table 1, whilst Table 2 shows the estimated total number of patients in England with advanced SM across the three subtypes.

Table 1 Estimated incidence and prevalence rates of advanced SM subtypes in Europe

Disease subtype	Incidence estimates per 100,000	Prevalence estimates per 100,000	Source
ASM	0.01 (0.006 to 0.03)	0.09 (0.03 to 0.21)	Cohen et al (2014)[4]
SM-AHN	0.04 (0.03 to 0.06)	0.31 (0.18 to 0.50)	Cohen et al (2014)[4]
MCL	0.01 (0.003 to 0.02)	0.00	Cohen et al (2014)[4]

ASM=aggressive systemic mastocytosis; MCL=mast cell leukaemia; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with associated haematologic neoplasm.
Source: CS, Table 5

Table 2 Assumption and calculation of the patient population with advanced SM eligible for treatment with midostaurin

Assumption		Value	Reference
Incidence			
1	Incidence of ASM	0.01/100,000	Epidemiology of SM in Denmark[4]
	Incidence of SM-AHN	0.04/100,000	
	Incidence of MCL	0.01/100,000	
	Incidence of advanced SM	**0.06/100,000**	
2	England population (2018)	55,977,000	ONS[26]
3	**England incidence of advanced SM (2018)**	**34**	**Calculation (assumption 1 x assumption 2)**
Prevalence			
4	Prevalence of ASM	0.09/100,000	Epidemiology of advanced SM in Denmark[4]
	Prevalence of SM-AHN	0.31/100,000	
	Prevalence of MCL	0.00/100,000	
	Prevalence of advanced SM	**0.40/100,000**	
5	**England prevalence of advanced SM (2018)**	**224**	**Calculation (assumption 2 x assumption 4)**
Total advanced SM population eligible for treatment			
6	**Advanced SM in England**	**258**	**Calculation (assumption 3 + assumption 5)**
7	Proportion of patients eligible for cytoreductive therapy (e.g., midostaurin)	67%	Clinical opinion (Proportion of patients *ineligible* for cytoreductive therapy=33%)[10]
8	**Patients in England eligible for treatment with midostaurin**	**174**	**Calculation (assumption 6 x assumption 7)**

ASM=aggressive systemic mastocytosis; MCL=mast cell leukaemia; ONS=Office for National Statistics; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with associated haematologic neoplasm
Source: CS, Table 6 (p21)

The company highlights (CS, p25) that NICE did not consider treatment with midostaurin to be eligible for assessment under NICE's Highly Specialised Technology (HST) process, despite the rare incidence of advanced SM. The ERG notes (from NICE's response to comments on the draft scope[27]) that treatment with midostaurin does not meet the criteria for consideration as a HST as it is already used in the NHS to treat a different disease (FLT3-positive AML) and because mastocytosis is not managed within a highly specialised service.

2.5 Critique of company's definition of the decision problem

A summary of the ERG's comparison of the decision problem outlined in the final scope[1] issued by NICE and that addressed within the CS is presented in Table 3. Each parameter is discussed in more detail in the text following the table (Section 2.6 to Section 2.11).

Table 3 Comparison between NICE scope and company's decision problem

Parameter	Final scope issued by NICE (original wording)	Decision problem addressed in the company submission with rationale
Intervention	Midostaurin	As per scope
Population	Adults with ASM, SM-AHN or MCL	As per scope, results were provided individually for each disease subtype
Comparator (s)	Current clinical management including but not limited to: • Interferon alpha • Cladribine • Imatinib • Nilotinib • Dasatinib (These treatments do not currently have a marketing authorisation in the UK for this indication)	Current clinical management including: • Interferon alpha • Cladribine • Imatinib • Pegylated interferon alpha (peg-interferon alpha) • AML-like treatments
Outcomes	The outcome measures to be considered include: • OS • PFS • Response rate • AEs • HRQoL	As per scope
Economic analysis	• The Reference Case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per QALY • The Reference Case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared • Costs will be considered from an NHS and PSS perspective • The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account	As per scope
Subgroups	If evidence allows, subgroup analysis by disease type to include: • ASM • SM-AHN • MCL	Cost effectiveness results were generated for the overall population (ASM+SM-AHN+MCL) and for the combined SM-AHN+MCL population

AE=adverse event; AML=acute myeloid leukaemia; ASM=aggressive systemic mastocytosis; HRQoL=health-related quality of life; MCL=mast cell leukaemia; OS=overall survival; PAS=patient access scheme; PFS=progression-free survival; PSS=Personal Social Services; QALY=quality adjusted life year; SM-AHN=systemic mastocytosis with associated haematological neoplasm
Source: Final scope[1] issued by NICE, CS, Table 1

2.6 Population

In accordance with the final scope[1] issued by NICE, the company has presented clinical effectiveness evidence for patients with ASM, patients with SM-AHN and patients with MCL. The company has also presented clinical effectiveness evidence for an overall population of patients with advanced SM (i.e., ASM+SM-AHN+MCL).

The evidence discussed in the CS is derived from two single arm, open label, phase II trials, the D2201[28-30] trial (n=116) and the A2213[31] trial (n=26). Clinical advice to the ERG was that the baseline characteristics of patients recruited to the D2201 and the A2213 trials were comparable to the baseline characteristics of patients with advanced SM who are treated in the NHS. However, the ERG notes that there are small numbers of patients within each disease subtype in these trials (Table 4).

Table 4 Advanced SM subtypes (D2201 and A2213 trials)

Disease subtype	D2201 (N=116)	A2213 (N=26)
ASM	22 (19%)	3 (12%)
SM-AHN	73 (63%)	17 (65%)
MCL	21 (18%)	6 (23%)

ASM=aggressive systemic mastocytosis; MCL=mast cell leukaemia; SM-AHN=systemic mastocytosis with associated haematological neoplasm
Source: Adapted from CS, Table 10

2.7 Intervention

Midostaurin is a cytoreductive therapy and is a multi-kinase inhibitor administered orally at a dose of 100mg twice daily.[23] The company has presented evidence for the licensed dose of midostaurin. Midostaurin is the only drug licensed in Europe to treat ASM, SM-AHN and MCL (CS, p24).

2.8 Comparators

As noted in Section 2.2 of this ERG report, there are no UK clinical guidelines for the treatment of advanced SM. Clinical advice to the company (CS, p21), and clinical advice to the ERG, is that in the NHS, the treatment of ASM, SM-AHN and MCL is tailored to each patient according to their symptoms.

The comparator treatments listed in the final scope[1] issued by NICE are interferon alpha, cladribine, imatinib, nilotinib and dasatinib. None of the listed comparators has a European marketing authorisation for the treatment of advanced SM. In terms of treatments for advanced SM, clinical advice (to the company [CS, p24] and to the ERG) is that in the NHS: i) nilotinib and dasatinib are rarely used; ii) imatinib is only used to treat the few patients who do not have

the *KIT* D816V mutation, and iii) pegylated interferon alpha and AML-like treatments are used (off licence).

There is no randomised controlled trial (RCT) evidence to support the clinical effectiveness of the five comparators (interferon alpha, cladribine, imatinib, nilotinib and dasatinib) listed in the final scope[1] issued by NICE, or the two comparators (pegylated interferon alpha and AML-like treatments) identified by the company for the treatment of advanced SM. However, the company identified eight[32-39] published studies (single-arm trials or observational studies) that assessed the clinical effectiveness of the individual comparators listed in the final scope[1] issued by NICE. The company also identified three publications[12,40,41] and an unpublished presentation[37] that reported the results of comparisons of health outcomes of patients with advanced SM treated with midostaurin versus patients treated with other (unspecified) drugs using data from European patient registries.

The ERG considers that the Reiter et al abstract[12] and unpublished presentation[12,42] are the most reliable sources of evidence to inform a decision about the comparative effectiveness of midostaurin versus standard of care (SoC) or midostaurin versus comparators. However, the ERG emphasises that there are several areas of uncertainty relating to the Reiter et al[12,42] methods and results (see Section 3.6 of this ERG report). The ERG highlights that only the numbers reported in the presentation[42] are used in the company cost effectiveness analyses (and that some of these unpublished values differ from those reported in the published abstract,[12] albeit the results provided in the unpublished study and subsequently used in the company base case are less favourable than the abstract).

The analysis methods used by Chandesris et al[40,41] are insufficiently described leading to uncertainty around the validity of the presented results. The ERG considers that these results should not be used to inform a decision about the comparative effectiveness of midostaurin versus SoC or midostaurin versus comparators (see Section 3.6 of this ERG report).

2.8 Line of treatment

The EMA's marketing authorisation[23] for treatment with midostaurin does not explicitly state whether midostaurin is to be used as a first-line or subsequent-line treatment; nor is line of treatment explicitly stated by the company in the CS. In the D2201 and A2213 trials, 55% and 19% of patients, respectively, had not received at least one previous treatment for advanced SM before being treated with midostaurin; data from the D2201 trial are used to populate the company's economic model.

2.9 Outcomes

Clinical advice to the ERG is that the health outcomes listed in the final scope[1] issued by NICE and addressed by the company are important outcomes for patients with ASM, SM-AHN and MCL.

No comparative progression-free survival (PFS), adverse event (AE) or HRQoL data have been published for patients with advanced SM. Furthermore, AE data for patients with advanced SM treated with midostaurin are limited to the safety outcomes of the 142 patients who were recruited to the single arm, open label, trials of midostaurin (D2201 and A2213 trials).

In the midostaurin trials and in most of the trials of the comparator treatments, objective response rate (ORR) was assessed using the modified Valent and Cheson[43] criteria. In one[44] of the comparator trials, ORR was measured using the more recently published criteria from the International Working Group-Myeloproliferative Neoplasms Research and Treatment (IWG-MRT) & European Competence Network on Mastocytosis (ECNM), generally referred to as the IWG[45] criteria. Clinical advice to the ERG is that the Valent and Cheson[43] criteria and the IWG[45] criteria are measures of treatment response that are used in the NHS.

2.10 Economic analysis

As specified in the final scope[1] issued by NICE, the cost effectiveness of treatment was expressed in terms of incremental cost per quality adjusted life years (QALYs) gained. Outcomes were assessed over a 38-year time horizon. Costs were considered from an NHS and a Personal Social Services perspective.

Confidential Patient Access Scheme (PAS) discounts are in place for midostaurin and azacitidine (the latter is used in the company's basket of comparator treatments). The PAS price of azacitidine is not known to the company. The actual discounted price of midostaurin and an assumed discounted price of azacitidine were used in the company model. The company highlights (CS, p14) that [redacted]. List prices for all other comparator treatments were used in the company model.

The company has put forward a case for midostaurin to be assessed under NICE's End of Life criteria.[46]

2.11 Other considerations

In the final scope[1] issued by NICE it is stipulated that, if the evidence allows, subgroup analysis by disease subtype (ASM, SM-AHN, MCL) should be considered.

The company considers that separate economic analyses for each of the three disease subtypes (ASM, SM-AHN and MCL) would be appropriate as the clinical activity and prognosis associated with each of these disease subtypes are heterogeneous. However, the company also states (CS, pg67) that results from such analyses would be associated with 'considerable uncertainty' as the only available source of effectiveness data is the D2201 trial (a single arm, open label, trial with very limited patient numbers for each disease subtype).

Instead, the company presented the results of an exploratory cost effectiveness analysis for the combined subgroup of patients with SM-AHN and MCL. Clinical advice to the company was that an exploratory subgroup analysis using the combined health outcome data from patients with SM-AHN and MCL in the D2201 trial was reasonable, based on the rationale that patients with SM-AHN and MCL have a much shorter life expectancy than patients with ASM.

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The full details of the company's search strategy and the methods used to identify and select the clinically relevant evidence of the effectiveness of midostaurin for the treatment of advanced SM are presented in the CS, Appendix D. The ERG did not identify any additional relevant studies to those previously identified by the company. Overall, the ERG considers that the methods used by the company to conduct a systematic review of the clinical effectiveness evidence were mostly satisfactory (Table 5).

Table 5 ERG appraisal of systematic review methods

Review process	ERG response	Note
Was the review question clearly defined in terms of population, interventions, comparators, outcomes and study designs?	Yes	See CS, Appendix D.3, Table D.3.1.
Were appropriate sources searched?	Yes	Sources searched were: MEDLINE, Embase, the Cochrane Library, peer-reviewed journals, and the ClinicalTrials.gov trial registry (including ongoing studies). Manual searches of abstracts from relevant conference proceedings were also conducted.
Was the timespan of the searches appropriate?	Yes	Searches were conducted in October 2019. Databases were searched from inception to the search date. Conference proceedings published from January 2017 to October 2019 were hand-searched in November 2019. Targeted searches were conducted in January 2020 to identify any clinical databases or patient registries for SM.
Were appropriate search terms used?	Yes	No additional ERG comments.
Were the eligibility criteria appropriate to the decision problem?	Yes	No additional ERG comments.
Was study selection applied by two or more reviewers independently?	Yes	No additional ERG comments.
Was data extracted by two or more reviewers independently?	No	Data were extracted by a single reviewer, however, extracted data were then checked and verified by a second reviewer.
Were appropriate criteria used to assess the risk of bias and/or quality of the primary studies?	Partially	The company used a modified version of the Downs and Black checklist [47] for quality assessment and used their own method of interpreting results. See Section 3.2.2 of the ERG report for details.
Was the quality assessment conducted by two or more reviewers independently?	No	Quality assessment was conducted by a single reviewer. A second reviewer confirmed the assessment and highlighted any discrepancies.
Were attempts to synthesise evidence appropriate?	N/A	The company was unable to conduct any robust indirect or mixed treatment comparisons because the evidence for midostaurin and its relevant comparators was derived from single-arm trials and observational studies.

N/A=not applicable; SM=systemic mastocytosis
Source: LRiG in-house checklist

3.2 Critique of trials: company's analysis and interpretation

3.2.1 Included trials

The company identified two trials (the D2201 trial and the A2213 trial) that provided evidence to demonstrate the clinical effectiveness of midostaurin for the treatment of advanced SM. Both trials are open label, phase II trials. In addition, both are single-arm trials and thus neither trial provides direct evidence comparing the effectiveness of midostaurin versus any of the comparators listed in the final scope[1] issued by NICE or identified by the company.

The company identified a further eight studies[32-39] that provide clinical effectiveness evidence for the five comparators listed in the final scope[1] issued by NICE, namely: interferon alpha,[32,35-37] cladribine,[34-37,39] imatinib,[35-37] nilotinib,[33] and dasatinib.[38] Two of the identified studies are single-arm trials,[33,38] two are case series studies,[33,39] and the remaining four are retrospective observational studies.[34-37] It was, therefore, not possible for the company to perform indirect treatment comparisons to estimate the relative effectiveness of midostaurin versus any of the comparators listed in the final scope[1] issued by NICE.

The company search, however, identified two historical control analyses (by Reiter et al[12,42] and Chandesris et al,[40,41]) that compared treatment with midostaurin versus treatment with SoC. The abstract[12] and presentation by Reiter et al,[42] describe results from a comparison of pooled data from the D2201 trial and the A2213 trial versus historical control data obtained from a German registry. The publications by Chandesris et al,[40,41] describe results from a comparison of prospective observational study of patients receiving midostaurin under a French compassionate use programme versus historical control data from a French registry. The Reiter et al[12,42] and Chandesris et al[40,41] publications and presentation provide the only available evidence for the comparison of treatment with midostaurin versus SoC.

Midostaurin trial characteristics

The key characteristics of the D2201 and A2213 trials are summarised in Table 6.

Both trials were non-randomised, single arm, open label, phase II trials. Although both trials were multi-centre trials, only the D2201 trial was conducted internationally; the A2213 trial was conducted across three centres, all of which were based in the USA. The D2201 trial recruited four patients from the UK; the remaining 138 patients were recruited from other European countries, Australia, and the USA. Clinical advice to the ERG is that the management of patients with advanced SM in Europe, Australia and the USA is comparable to the management of patients with advanced SM in the NHS. The ERG notes that the median

duration of follow-up for the primary outcome of ORR, was longer in the A2213 trial (124 months) than in the D2201 trial (26 months).

Table 6 Key characteristics of the D2201 and A2213 trials

Trial parameters	D2201	A2213
Design	• Phase II, multi-centre, open label, single arm, international, N=116 • Adapted Fleming[48] two-stage design	• Phase II, investigator-led, multi-centre, open label, single arm, N=26 • Simon[49] two-stage design
Patient population	• Adults (≥18 years of age) with a diagnosis of ASM, SM-AHN or MCL according to WHO criteria[50,51] • Presence of ≥1 C-findings • ECOG[52] performance status 0-3	• Adults (≥18 years of age) with histologically documented ASM, SM-AHN or MCL • Presence of ≥1 C-findings • ECOG[52] performance status 0-3
Primary outcome	Best response defined as the percentage of participants who classified as confirmed responders (MR or PR within the first 6 treatment cycles and maintained for ≥8 weeks) based on modified Valent and Cheson[43] criteria	Best response defined as the percentage of participants who classified as confirmed responders (MR or PR within the first 2 treatment cycles and maintained for ≥8 weeks) based on published Valent and Cheson[43] criteria
Median length of follow-up for ORR	26 months (range 12 to 54)*	124 months (range 82 to 140)**
ERG comment	• 4 patients recruited in the UK • No control or comparator group	• Small patient numbers • No trial centres in the UK • No control or comparator group

ASM=aggressive systemic mastocytosis; SM-AHN=systemic mastocytosis with associated haematologic neoplasm; ECOG=Eastern Cooperative Oncology Group; MCL=mast cell leukaemia; MR=major response; ORR=overall response rate; PR=partial response; SM-AHN=systemic mastocytosis with associated haematologic neoplasm; WHO=World Health Organisation
Source: Adapted from CS, Table 9 and CS, p48*, CS, p57**

Baseline characteristics of patients participating in the midostaurin trials

Full details of the baseline characteristics of patients participating in the D2201 and A2213 trials are provided in the CS (Table 10), and a summary is presented in this ERG report (Table 7). Clinical advice to the ERG is that the patients recruited to the two trials are generally representative of patients with advanced SM treated in the NHS. However, clinical advice to the ERG is also that the proportions of patients with MCL in the D2201 (18%) and A2213 (23%) trials are greater than the proportion that would be seen in NHS clinical practice with this disease subtype. Compared with patients participating in the D2201 trial, a higher proportion of the patients recruited to the A2213 trial presented with an ECOG performance status of 2 to 3, and a higher proportion presented with three or more C-findings. Data from

the trials suggest that the patients participating in the A2213 trial had greater disease burden than patients participating in the D2201 trial, and the difference is acknowledged by the company (CS, p39). Approximately half of the patients in the D2201 trial (55%) and a fifth of patients in the A2213 trial (19%) were treatment naive.

Table 7 Baseline characteristics of patients in the D2201 and A2213 trials

Baseline characteristic	D2201 (N=116)	A2213 (N=26)
Age (years)		
Median (range)	63.0 (25–82)	64.5 (24–79)
Sex – n (%)		
Male	76 (66)	15 (58)
Female	40 (34)	11 (42)
ECOG performance status – n (%)		
0	[redacted]	12 (46)
1	[redacted]	
2	[redacted]	14 (54)
3	[redacted]	
Number of previous therapies – n patients (%)		
0	64 (55)	5 (19)
1	29 (25)	8 (31)
2	15 (13)	6 (23)
≥3	8 (7)[a]	7 (27)
Subtype of advanced SM – n (%)		
ASM	22 (19)[b]	3 (12)
SM-AHN	73 (63)[b]	17 (65)
MCL	21 (18)	6 (23)[d]
***KIT* D816 mutation status – n (%)**		
Positive	98 (84)	20 (77)
Negative	13 (11)	5 (19)
Other	5 (4)[c]	1 (4)[e]
Bone marrow mast-cell burden – %		
Median (range)	40 (3–98)	50 (5–95)
Serum tryptase level – µg/L		
Median (range)	200 (2–12,069)	323 (22–1,255)
Number of C-findings per patient – n patients (%)		
1	31 (27)	3 (12)
2	20 (17)	10 (38)
≥3	38 (33)	13 (50)

[a] Therapy in some of these cases was directed toward the AHN component of SM-AHN. [b] These numbers were derived from the EPAR and calculated by subtracting the known number of patients in each category from the total number of patients in the trial. [c] KIT D816 mutation status unknown. [d] Two MCL patients had chronic myelomonocytic leukemia-1 as an AHN.
[e] The patient was positive for the KIT S451C mutation.
ASM=aggressive systemic mastocytosis; MCL=mast cell leukaemia; SM-AHN=systemic mastocytosis with associated haematologic neoplasm.
Source: Adapted from CS, Table 10

3.2.2 Quality assessment of the D2201 and A2213 trials

The company conducted a quality assessment of the D2201 and A2213 trials using a modified version of the Downs and Black[47] checklist (Table 8). To summarise and interpret the responses to the items in the modified checklist, the company used their own method (see Appendix 7.1.1 of this ERG report for ERG comment).

Table 8 Company's modifications to the Downs and Black quality assessment checklist

Company modification	Item description	ERG critique
Item 27 of the checklist	Was the study sufficiently powered to detect clinically important effects where probability value for a difference due to chance is <5%?	The company gives a simplified answer of yes, no or unclear for item 27 (0-1 points) rather than a graded answer (0-5 points). The ERG notes that this modification is widely used and accepted.
Exclusion of item 23 and 24 from the checklist	Item 23. Were study subjects randomised to intervention groups? Item 24. Was the randomised intervention assignment concealed from patients and staff until recruitment was complete?	The guidance from Downs and Black[47] is that the response to Items 23 and 24 should be 'no' for non-randomised studies.
Modification of item 5 from the checklist	Are the distributions of principal confounders in each group of subjects clearly described?	The company gives a simplified response of 'yes', 'no' or 'unclear' for item 5 (0-1 points) rather than a graded answer (0-2 points).

Source: LRiG in-house table

The company assessed the D2201 trial as being of good quality and the A2213 trial as being of reasonable quality (Table D.11.2, Appendix D.11 to the CS). The ERG considers that the company's assessments of the methodological quality of the D2201 and the A2213 trials are not reliable as the assessments are based on non-validated methodology. However, the ERG considers that the D2201 and the A2213 trials appear to be of a higher methodological quality than the studies of the comparator treatments.

The company's risk of bias assessments for the D2201 and A2213 trials, with ERG comments, are presented in Table 9. (The ERG has reinstated the two items [23 and 24]) omitted from the company's assessment). The ERG considers the strengths of the two trials are that they were well reported, patients recruited to the trials were representative of patients treated in the NHS and, that valid outcome measures were used to assess the efficacy of treatment with midostaurin. However, the ERG considers that the weaknesses of the trials are that potential confounding variables are not described or adjusted for in the analyses. The ERG also notes that the statistical methods used in both trials were appropriate, but that the A2213 trial was underpowered to detect differences in treatment responses. The ERG considers that the D2201 and the A2213 trials are of reasonable quality, but highlights that they are single-arm

trials without a control group and their results cannot be considered as reliable or robust as the results of a RCT.

Table 9 Company's quality assessment of the D2201 and A2213 trials with ERG comment

Downs and Black checklist item	D2201	A2213	ERG comment
Reporting			
Q1: Aim of the study clearly described	Yes	Yes	Agree
Q2: Outcomes to be measured clearly described	Yes	Yes	Agree
Q3: Patients characteristics clearly described	Yes	Yes	Agree
Q4: Interventions clearly described	Yes	Yes	Agree
Q5: Principal confounders clearly described	Yes	Yes	Patient characteristics are well described in both trials. However, no potential confounders or treatment effect modifiers are defined for either trial.
Q6: Main findings clearly described	Yes	Yes	Agree
Q7: Random variability for the main outcome provided	Yes	Yes	Agree
Q8: Adverse events reported	Yes	Yes	Agree
Q9: Characteristics of patients lost to follow up reported	No	No	There were minimal losses to follow-up; only two patients were lost to follow-up from the D2201 trial and no losses to follow-up in the A2213 trial
Q10: Actual p-values reported	Yes	Yes	Agree
External validity and bias			
Q11: Sample asked to participate representative of the population from which they were recruited	Unclear. The number of patients screened prior to study entry unclear	Unclear. The number of patients screened prior to study entry unclear	Clinical advice to the ERG is that the patients recruited to the trials are representative of patients treated in the NHS
Q12: Patients who agreed to participate are representative of the population from which they	Unclear. The study has not demonstrated that the distribution of the main confounding factors was the same in the study sample and the source population	Unclear. The proportion of patients screened that then gave consent not reported	Clinical advice to the ERG is that the patients recruited to the trials are representative of patients treated in the NHS
Q13: Staff participating representative of the patient's environment	Unclear. No details provided	Yes. Patients treated at cancer institutes or an academic centre	The ERG agrees that details of the treatment centres in the D2213 trial were not explicitly reported. However, author affiliations suggest that it is likely that patients were treated at cancer institutes or academic centres

Q14: Attempt to blind participants	No	No	Agree
Q15: Attempt to blind assessors	No	No	Central review of patient response to treatment was carried out in D2213
Q16: Any results based on data dredging are clearly stated	No. Subgroup analysis was performed that was not mentioned in methods	No. Extensive subgroup analyses reported in the supplementary appendix were not described in the methods	Agree. See Table 10 of this ERG report for details
Q17: Analysis adjusted for length of follow up	Unclear. No details provided	Unclear. No details provided	Adjustment of follow-up between groups is not necessary for a single-arm trial. Survival analyses of OS and PFS account for different lengths of patient follow-up
Q18: Appropriate statistical tests used?	Yes	Yes	Agree
Q19: Reliable compliance	Unclear	Unclear	Agree
Q20: Reliable and valid outcome measures	Yes	Yes	Agree
Statistical bias and power			
Q21: Were patients in the different intervention groups recruited from the same population?	NA	NA	Agree. Both trials were single-arm trials
Q22: Were patients in the different intervention groups recruited at the same time?	NA	NA	Agree. Both trials were single-arm trials
Q23. Were patients randomised to intervention groups?	This item was excluded from the company's checklist	This item was excluded from the company's checklist	No. Both trials were single-arm trials
Q24. Was the randomised intervention assignment concealed	This item was excluded from the company's checklist	This item was excluded from the company's checklist	Not applicable. Both trials were single-arm trials
Q25: Adequate adjustment in the analyses for confounding variables?	Yes	No. No adjustments reported	No confounding variables were defined (see ERG comment on Q5), therefore it is unclear if confounders have been adequately adjusted for
Q26: Losses of patients to follow up accounted for	Unclear	Unclear	Losses to follow-up were minimal (see ERG comment on Q5). Survival analyses of OS and PFS account for losses to follow-up and patients who discontinue treatment by censoring
Q27: Did the study have sufficient power to detect a clinically important event where the probability value for a difference being due to chance is less than 5%?	Yes	Unclear. No detail provided	The probability of a difference being due to chance is 9.4% in the A2213 trial (A2213 protocol, Section 7.2[53])

NA=not applicable; PFS=progression free survival; OS=overall survival
Source: Adapted from CS, Table 13 with ERG comment

3.2.3 Statistical approach adopted for the trials of midostaurin

Information relevant to the statistical approach used by the company was extracted from the clinical study reports (CSRs) of the D2201 trial (dated 28th June 2016[29] and 27th April 2018[30]) the protocols of the D2201[54] and A2213 trials, [53] and from the CS. A summary of the additional checks made by the ERG in relation to the company's pre-planned statistical approach used to analyse data from the included trials is provided in Table 10.

Table 10 ERG assessment of statistical approaches used in the midostaurin trials

Item	ERG assessment	Statistical approach with ERG comments
Were all analysis populations clearly defined and pre-specified?	Clearly defined: Yes, for both trials Pre-defined: Partly for both trials	The analysis populations (FAS, PEP, SES in both trials and additionally PPS in the D2201 study) are well defined in Table 11 of the CS. Clinical effectiveness results presented in the CS for the final analysis set (FAS) (i.e., all patients to whom study treatment had been assigned according to the ITT principle) for both trials. Clinical effectiveness results for the PEP (i.e., all patients who had measurable C-findings considered related to SM) of the D2201 trial were also reported and these results are used to inform the economic model. The ERG notes that the pre-defined PEPs outlined within the protocols of the trials (Section 10.1 of the D2201 trial protocol and Section 7.1 of the A2213 trial protocol) were different from the PEPs defined in Table 11 of the CS.
Was an appropriate sample size calculation pre-specified?	Yes, for both trials	The two-stage study designs of the D2201 and the A2213 trials are outlined in Figure 4 of the CS and pp37-38 of the CS. These designs and sample sizes were pre-specified in Section 10.4.3 of the D2201 trial protocol and Section 7.2 of the A2213 trial protocol. The ORR achieved by patients enrolled in Stage 1 were significantly greater than the pre-specified thresholds (ORR of 30% in the D2201 trial and 10% in the A2213 trial) for rejection of the null hypothesis in both trials. Therefore, both of the trials continued to enrol patients in Stage 2. The ERG is satisfied that the pre-defined Fleming[55] and Simon[49] two-stage designs are appropriate for the D2201 and A2213 trials respectively and that these designs were implemented appropriately within the trials.
Were all protocol amendments made prior to analysis?	D2201 trial: Mostly A2213: Unknown	Six protocol amendments for the D2201 trial were provided in Section 9.8.1 of the CSRs. [redacted] was issued to [redacted]. The only amendment ([redacted]) to be issued after the first data cut (9th July 2013) was [redacted]. The ERG considers that this amendment is reasonable. Trial protocol version 2.0 for the A2213 was provided. No protocol amendments were listed within this protocol or were available to the company.
Were all primary and secondary efficacy outcomes pre-defined and analysed appropriately?	Partly for both trials	The primary endpoint for both trials was best response (CS, Table 9) and was pre-defined within the trial protocols (Section 10.4.1 of the D2201 trial protocol and Section 7.1.5 of the A2213 trial protocol). For both trials, the secondary efficacy outcomes were OS, PFS, DoR, TTR and histopathologic response based on mast cell infiltration and serum tryptase levels (CS, Table 9). Response assessment based on C-findings (including non-measurable ones) was also a secondary efficacy outcome of the D2201 trial (CS, Table 9). Secondary efficacy outcomes were well defined and pre-defined in Section 10.5 and 10.6 of the D2201 trial protocol. Secondary outcome definitions were well defined within the supplementary documentation to the publication of the A2213 trial[31] but were not pre-specified within the A2213 trial protocol. Appropriate statistical analysis methods for primary and secondary efficacy outcomes of both trials were described in the CS (Table 12) and in the trial publications.[28,31] Limited details of statistical analyses were pre-specified in the

		trial protocols.
Was the analysis approach for PROs appropriate and pre-specified?	D2201 trial: Partly A2213 trial: NA	The analysis approach to PROs in the D2201 trial comprised descriptive statistics of patient responses to the SF-12 and the MSAS at each study visit (Section 10.6.2 of the D2201 trial protocol). The ERG is satisfied that the PRO definitions and analysis approaches are appropriate but notes that the analysis population of evaluable patients (CS, pp54-56) and measurement of median best percentage change from baseline in MSAS and SF-12 scores and statistical testing of PROs using two-sided Wilcoxon two-sample test (CS, Figure 13 and Figure 14) were not pre-specified in the D2201 trial protocol. PROs were not reported within the A2213 trial protocol.
Was the analysis approach for AEs appropriate and pre-specified?	Yes, for both trials	AEs were assessed and graded using the CTCAE version 3.0 classification system within the SES. AEs were estimated as numbers and percentages of patients experiencing events; no formal statistical analyses of AEs were conducted. Summaries of AEs in ≥10% of patients, AEs leading to study drug discontinuation, AEs of special interest, SAEs and death are presented in the CS for both trials (Section B.2.10, pp89-98). The ERG is satisfied that the approach employed for AEs was pre-defined (Section 7.5.1 and Section 8.1 of the D2201 trial protocol and Section 7.1.6 of the A2213 trial protocol) and is appropriate.
Was a suitable approach employed for handling missing data?	Yes, for both trials	The approach to managing missing data for both trials is described in the CS (Table 12). Patients with missing assessments were considered to be non-responders. These approaches were pre-specified in Section 10.4.4 of the D2201 trial protocol and Section 7.1.5 of the A2213 trial protocol. For secondary efficacy outcomes (OS, PFS, DoR, TTR), outcomes were censored at the last available efficacy evaluation for patients with missing data (Section 10.5.2 of the D2201 trial protocol and supplement to the A2213 trial publication).[31]
Were all subgroup and sensitivity analyses pre-specified?	Partly for both trials	Subgroup analyses for both trials were conducted for disease subtype (ASM, SM-AHN and MCL) and *KIT* D816V mutation status (positive and negative or unknown) for both trials (CS, Table 9). Subgroup analysis by number of prior therapies (0 and ≥1) was conducted for the D2201 trial and subgroup analysis for C-findings at baseline (anaemia, thrombocytopenia, neutropenia and non-haematologic C-findings) was conducted for the A2213 trial (CS, Table 9). Subgroup results for ORR of all subgroup analyses described in Table 9 of the CS are presented in Table 17 of the CS for the D2201 trial and in Table 20 of the CS for the A2213 trial. Subgroup analyses by disease subtype for OS, PFS and DoR are presented in Table 18, Table 19 and Figure 23 of the CS for the D2201 trial and in Table 21 of the CS for the A2213 trial. Additional subgroup analyses of OS by response subgroup and by *KIT* D816V mutation status and for ORR by C-findings present at baseline, age, gender, MSAS and SF-12 subscales for the D2201 trial and an analysis of time-to treatment discontinuation by response subgroup are presented in Appendix E to the CS. Only subgroup analyses of disease subtype and number of prior therapies were pre-specified in Section 10.6.5 of the D2201 protocol. Sensitivity analyses of efficacy endpoints in the FAS were pre-specified in Section 10.1 of the D2201 trial protocol. No subgroup or sensitivity analyses were pre-specified in the A2213 trial protocol.

AE=adverse event; ASM=aggressive systemic mastocytosis; CSR=clinical study report; CTCAE=common terminology criteria for adverse events; DoR=duration of response; FAS=full analysis set; ITT=intention-to-treat; MCL=mast cell leukaemia; NA=not applicable; ORR=overall response rate; OS=overall survival; PEP=primary efficacy population; PFS=progression-free survival; PPS=per protocol set; PRO=patient reported outcome; SAE=serious adverse event; SES=safety evaluation set; SF-12=12-item Short-Form health survey; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with associated haematologic neoplasm; TTR=time to response

Source: extracted from the CS, CSRs of the D2201 trial,[29,30] the protocols of the D2201[54] and A2213 trials,[53] the publications of the D2201[28] and A2213 trials,[31] the company's response to the clarification letter and ERG comment

Having carried out these checks, the ERG is satisfied with the pre-planned statistical approach employed by the company and notes that:

- the definitions of the primary efficacy population (PEP) presented in the CS for both the D2201 trial and the A2213 trial were different from those pre-specified in the protocols[53,54]
- subgroup analyses which were not pre-specified in the trial protocols[53,54] are presented in the CS
- limited details of the pre-planned statistical approach were presented in the trial protocols.[53,54]

3.3 Efficacy results from the trials of midostaurin

Efficacy results presented in the CS have been analysed using data from three different D2201 trial data-cuts and from one A2213 trial data-cut (Table 11). The median duration of follow-up is substantially longer in the A2213 trial (124 months) than the median durations of follow-up in the D2201 trial (ranges from 26 months to 76 months). The ERG considers that these differences in median duration of follow-up should be taken into account when comparing results from the D2201 and A2213 trials, particularly for outcomes such as OS, PFS and DoR which are time-dependent.

Table 11 Efficacy outcomes reported in the trials of midostaurin: PEP

Trial	Data cut-off date	Outcomes	Median duration of follow-up	Source
D2201	9th July 2013	ORR, OS, PFS, DoR, TTR	26 months (range 12 to 54 months)	Gotlib et al 2016[28]
	1st December 2014	ORR, OS, PFS, DoR, TTR, ORR and DoR by IWG criteria	43 months (range 29 to 70 months)	EPAR;[23] APAR[56]
	24th August 2017	OS (Final analysis)	[redacted]	D2201 final CSR[30]
A2213	1st March 2017	ORR, OS, PFS, DoR, TTR	124 months (range 82 to 140 months)	De Angelo et al 2018[31]

APAR=Australian public assessment report; CSR=clinical study report; DoR=duration of response; EPAR=European public assessment report; IWG=International Working Group; ORR=overall response rate; OS=overall survival; PEP=primary efficacy population; PFS=progression free survival; TTR=time to response
Source: CS, adapted from Section 2.6

Efficacy results within this section are presented for all patients and by disease subtype (ASM, SM-AHN or MCL). Further efficacy results, including OS, PFS and DoR Kaplan-Meier (K-M) data, can be found in Section 2.6 and Section 2.7 of the CS.

3.3.1 Primary outcome: best overall response

A summary of best overall response and ORR results for all patients and by disease subtype in the PEP of the trials of midostaurin are provided in Table 12. The ORR results are exactly the same in the two data cut-offs reported for the D2201 trial (9th July 2013 and 1st December 2014) and were used to inform the economic model.

Table 12 Summary of ORR results for all patients and by disease subtype in trials of midostaurin: PEP

	Disease subtype			All patients
	ASM	SM-AHN	MCL	
D2201 trial (data cut-off 1st December 2014, median duration of follow-up 43 months [range 29 to 70 months])[a]				
Number of patients	16	57	16	89
Overall response: n (%)	12 (75%)	33 (58%)	8 (50%)	53 (60%)
Major response: n (%)[b]	10 (62%)	23 (40%)	7 (44%)	40 (45%)
Partial response: n (%)[b]	2 (12%)	10 (18%)	1 (6%)	13 (15%)
Stable disease: n (%)	1 (6%)	7 (12%)	3 (19%)	11 (12%)
Progressive disease: n (%)	1 (6%)	6 (11%)	3 (19%)	10 (11%)
Not evaluable: n (%)[c]	2 (12%)	11 (19%)	2 (12%)	15 (17%)
ORR (95% CI)	75% (48 to 93%)	58% (44 to 71%)	50% (25 to 75%)	60% (49 to 70%)
A2213 trial (data cut-off 1st March 2017, median duration of follow-up 124 months [range 82 to 140 months])				
Number of patients	3	17	6	26
Overall response: n (%)	1 (33%)	13 (76%)	4 (67%)	18 (69%)
Major response: n (%)[b]	0 (0%)	11 (65%)	2 (33%)	13 (50%)
Partial response: n (%)[b]	1 (33%)	2 (12%)	2 (33%)	5 (19%)
Stable disease: n (%)	1 (33%)	3 (18%)	1 (17%)	5 (19%)
Progressive disease: n (%)	1 (33%)	1 (6%)	1 (17%)	3 (12%)
ORR (95% CI)	33% (NR to NR)	76% (NR to NR)	67% (NR to NR)	69% (50 to 88%)

[a] Results are exactly the same in the two data cut-offs reported for the D2201 trial (9th July 2013 and 1st December 2014) and results using the latest data cut-off were used to inform the economic model
[b] Numbers and proportions with different categories of major response (complete remission, incomplete remission, pure clinical response) and partial response (good partial response and minor partial response) are presented in Table 14, Table 15 and Table 16 of the CS and Table 2 of the A2213 trial publication[31]
[c] Reasons to explain why patients were not evaluable for response in the D2201 trial were: concurrent use of high-dose glucocorticoids (n=9), not enough time receiving treatment (n=3), death (n=1), red-cell transfusion (n=1), and neutropenia (n=1)
ASM=aggressive systemic mastocytosis; CI=confidence interval; MCL=mast cell leukaemia; NR=not reported; ORR=overall response rate; PEP=primary efficacy population; SM-AHN=systemic mastocytosis with associated haematologic neoplasm;
Source: Extracted and adapted from CS; Table 14, Table 15, Table 16, Table 17 and Table 20; A2213 trial publication[31]

At the time of analysis in the PEP, the ORR in the D2201 trial was 60% (95% confidence interval [CI]: 49 to 70%) and in the A2213 trial was 69% (95% CI: 50 to 88%). The ORR in the full analysis set [FAS] of the D2201 trial (116 enrolled patients) was 46% (95% CI [calculated by the ERG]: 37 to 55%).

In the D2201 trial, patients with ASM had the highest response rate of 75% (95% CI: 48 to 93%) compared to 58% (95% CI: 44 to 71%) and 50% (95% CI: 25 to 75%) for patients with SM-AHN and patients with MCL, respectively. The company notes that, in contrast, patients

with ASM in the A2213 trial had the lowest response rate with an ORR of 33%, compared to 76% and 67% for patients with SM-AHN and patients with MCL, respectively.

The ERG considers that direct comparisons between the numerical results of ORR across the disease subtypes should not be made due to very small numbers of patients and uncertainty in ORR estimates (reflected in the wide CIs of ORR estimates in the D2201 trial and lack of reported CIs in the A2213 trial). As noted in Section 3.3 of this ERG report, the ERG also considers that direct comparisons between the numerical values of the efficacy results of the D2201 trial and the A2213 trial should be not made due to the different median duration of follow-up times in the two trials.

Other subgroup analyses

Aside from disease subtype, other subgroup analyses conducted in the PEP of the D2201 trial (data cut-off: 9th July 2013 and equivalent results for data cut-off 1st December 2014) for ORR based on *KIT* D816V mutation status (positive, negative or unknown) and number of prior therapies (0 or ≥1) are presented in Table 17 and Figure 19 of the CS. Additional subgroup analyses of ORR in the D2201 trial by C-findings present at baseline (data cut-off 9th July 2013) and by age, gender, MSAS category and SF-12 (data cut-off 1st December 2014) are presented in Appendix E to the CS.

Other subgroup analyses conducted in the PEP of the A2213 trial (data cut-off: 1st March 2017) for ORR based on *KIT* D816V mutation status (positive, negative or other), number of prior therapies (0 or ≥1), C-findings present at baseline (anaemia, thrombocytopenia, neutropenia or non-haematologic C-findings) and additional SM-related findings (pleural effusions and increased alkaline phosphatase) are presented in Table 20 of the CS.

Responses were observed in all pre-specified subgroups of the D2201 trial and in all pre-specified subgroups of the A2213 trial, except for transfusion-dependent thrombocytopenia, neutropenia and lytic lesions where the ORR was 0%.

3.3.2 Secondary outcome: OS

A summary of OS results in the trials of midostaurin is provided in Table 13.

Table 13 Summary of OS results for all patients in the trials of midostaurin: PEP and FAS

Population	Number of patients alive: n (%)	Median OS (95% CI): months	Survival rate: (95% CI)	
			3 years	5 years
D2201 trial (data cut-off 9th July 2013, median duration of follow-up 26 months [range 12 to 54 months])				
PEP (n=89)	48 (54%)	28.7 (18.1 to NE)	46% (32 to 58%)	Not reported
FAS (n=116)	48 (41%)	33.9 (20.3 to 45.5)	46% (35 to 57%)	Not reported
D2201 trial (data cut-off 1st December 2014, median duration of follow-up 43 months [range 29 to 70 months])				
PEP (n=89)	35 (39%)	26.8 (17.6 to 34.7)	38.2% (27.5 to 48.8%)	Not reported
FAS (n=116)	35 (30%)	29.9 (20.3 to 42.0)	42.4% (32.6 to 51.8%)	Not reported
D2201 trial (data cut-off 24th August 2017, median duration of follow-up 76 months [range 62 to 103 months])				
PEP (n=89)	[redacted]	[redacted]	[redacted]	[redacted]
FAS (n=116)	[redacted]	[redacted]	[redacted]	[redacted]
A2213 trial (data cut-off 1st March 2017, median duration of follow-up 124 months [range 82 to 140 months])				
PEP (n=26)[c]	4 (15%)	40.0 (27.3 to 52.7)	Not reported	Not reported

[a] 4 patients were known to be alive (ongoing without event), an additional 9 patients (10%) were lost to follow-up early in the trial and an additional 12 patients (13%) were lost to follow-up but known to be alive in the 5 months before data cut-off
[b] 6 patients were known to be alive (ongoing without event), an additional 14 patients (12%) were lost to follow-up early in the trial and an additional 16 patients (14%) were lost to follow-up but known to be alive in the 5 months before data cut-off
[c] The PEP and the FAS were equivalent in the A2213 trial
CI=confidence interval; FAS=full analysis set; NE=not estimated; OS=overall survival; PEP=primary efficacy population
Source: Extracted and adapted from CS: Section 2.6.1 (pp49-51), Section 2.6.2 (p58) and ERG calculation

At the time of the final D2201 trial OS analysis, after a median duration of follow-up of 76 months, the median OS (95% CI) was [redacted] months for the PEP and [redacted] months for the FAS. In the A2213 trial, after a median duration of follow-up of 124 months, the median OS (95% CI) was 40.0 (27.3 to 52.7) months for the PEP. The ERG emphasises that the different durations of median follow-up times in the D2201 and the A2213 trials should be considered when drawing conclusions from the OS trial results.

A summary of OS results in the PEP of the trials of midostaurin by disease subtype is provided in Table 14.

Table 14 Summary of OS results by disease subtype in the PEP of the trials of midostaurin

	Disease subtype			All patients
	ASM	SM-AHN	MCL	
D2201 trial (data cut-off 9th July 2013, median duration of follow-up 26 months [range 12 to 54 months])				
Number of patients	16	57	16	89
Median OS (95% CI), months	NR (28.7 to NE)	20.7 (16.0 to 44.4)	9.4 (7.5 to NE)	28.7 (18.1 to NE)
D2201 trial (data cut-off 1st December 2014, median duration of follow-up 43 months [range 29 to 70 months])				
Number of patients	16	57	16	89
Median OS (95% CI), months	51.1 (28.7 to NE)	20.7 (16.3 to 33.9)	9.4 (7.5 to NE)	26.8 (17.6 to 34.7)
D2201 trial (data cut-off 24th August 2017, median duration of follow-up 76 months [range 62 to 103 months])				
Number of patients	16	57	16	89
Median OS (95% CI), months	[redacted]	[redacted]	[redacted]	[redacted]
A2213 trial (data cut-off 1st March 2017, median duration of follow-up 124 months [range 82 to 140 months])				
Number of patients	3	17	6	26
Median OS (95% CI), months	NR (NR to NR)	40.0 (24.2 to 55.9)	18.5 (0.0 to 62.2)	40.0 (27.3 to 52.7)

ASM=aggressive systemic mastocytosis; CI=confidence interval; MCL=mast cell leukaemia; NE=not estimated; NR=not reached; OS=overall survival; PEP=primary efficacy population; SM-AHN=systemic mastocytosis with associated haematologic neoplasm;
Source: Extracted and adapted from CS: Table 18, Table 19, Table 21 and Section B.2.7.1 (p67)

The company notes that K-M data (CS, Figure 15, Figure 20 and Figure 23) indicate that OS was highest for patients with ASM. The ERG considers that direct comparisons between numerical results for OS across the disease subtypes should not be made due to very small numbers of patients and uncertainty in OS estimates (wide CIs and CIs not reached or not estimated for median OS).

Additional subgroup analyses of D2201 trial OS data by response (data cut-offs 9th July 2013 and 1st December 2014) and by *KIT* D816V mutational status (data cut-off 1st December 2014 and 24th August 2017) are presented in the CS, Appendix E.

3.3.3 Secondary outcomes: PFS and DoR

A summary of PFS and DoR results for all patients and by disease subtype in the PEP of the trials of midostaurin are provided in Table 15.

Table 15 Summary of PFS and DoR results for all patients and by disease subtype in the PEP of the trials of midostaurin

	Disease subtype			All patients
	ASM	SM-AHN	MCL	
D2201 trial (data cut-off 9th July 2013, median duration of follow-up 26 months [range 12 to 54 months])				
Number of patients	16	57	16	89
Median PFS (95% CI), months	28.7 (24.8 to NE)	11.0 (7.4 to 17.0)	11.3 (2.8 to NE)	14.1 (10.2 to 24.8)
Number of responders	12	33	8	53
Median DoR (95% CI), months[a]	NR (24.1 to NE)	12.7 (7.4 to 31.4)	NR (3.6 to NE)	24.1 (10.8 to NE)
D2201 trial (data cut-off 1st December 2014, median duration of follow-up 43 months [range 29 to 70 months])				
Number of patients	16	57	16	89
Median PFS (95% CI), months	NR (NR to NR)	11.0 (7.4 to 17.9)	11.3 (2.8 to NE)	17.0 (10.2 to 24.8)
Number of responders	12	33	8	53
Median DoR (95% CI), months[a]	NR (NR to NR)	12.7 (7.4 to 31.4)	Not reported	31.4 (10.8 to NE)
A2213 trial (data cut-off 1st March 2017, median duration of follow-up 124 months [range 82 to 140 months])				
Number of patients	3	17	6	26
Median PFS (95% CI), months	Not reported	Not reported	Not reported	41.0 (4.4 to 77.6)
Number of responders	1	13	4	18
Median DoR (95% CI), months[a]	Not reported	Not reported	Not reported	132 (NE to NE)

[a] DoR is measured only in those achieving an overall response
ASM=aggressive systemic mastocytosis; CI=confidence interval; DoR=duration of response; MCL=mast cell leukaemia; NE=not estimated; NR=not reached; OS=overall survival; PEP=primary efficacy population; PFS=progression free survival; SM-AHN=systemic mastocytosis with associated haematologic neoplasm;
Source: Extracted and adapted from CS: Table 18 and Table 19, Section 2.6.1 (pp51-53), Section 2.6.2 (pp58-60)

At a median duration of follow-up of 43 months, the median PFS (95% CI) was 17.0 (10.2 to 24.8) months in the D2201 trial; and for the 53 responders, the median DoR (95% CI) was 31.4 (10.8 to NE) months. At a median duration of follow-up of 124 months, the median PFS (95% CI) was 41.0 (4.4 to 77.6) months in the A2213 trial; and for the 18 responders, the median DoR (95% CI) was 132 (NE to NE) months. The company states (CS, Section 2.7) that PFS and DoR results are often highest for the ASM subgroup (where results are reached); the ERG considers that comparisons of numerical results for PFS and DoR across disease subtypes should not be made due to very small numbers of patients.

According to the definitions of OS and PFS used within the A2213 trial,[31] median OS should be longer than PFS. However, the ERG notes that the reported median PFS is longer than the reported median OS (41.0 months versus 40.0 months respectively). These results are reported in all published sources of the A2213 trial[23,31,56] and therefore are unlikely to be a typographical error. The ERG acknowledges that the company has access to only published sources of the A2213 trial data and therefore cannot explain this result. The ERG concludes that the inconsistency in median OS and PFS is likely due to the small numbers of patients enrolled within the A2213 trial as a single OS or PFS event could have a large impact on the survival probabilities, and therefore on the time taken for median OS or PFS to be reached. The uncertainty around these results is also reflected within the large 95% CIs around median OS (Table 13) and median PFS (Table 15).

3.3.4 Other outcomes

Histopathologic response outcomes based on mast cell infiltration and serum tryptase levels are reported in the CS (on pp53-54 for the D2201 trial, and on pp60-61 for the A2213 trial).

3.3.5 Post-hoc analysis of response by IWG criteria in the D2201 trial

A post-hoc analysis using data from the D2201 trial (data cut-off 1st December 2014) was conducted using the more stringent IWG[45] criteria to measure response. The response of patients without measurable C-findings who were excluded from the pre-specified analyses of ORR based on modified Valent[57] and Cheson[58,59] criteria adjudicated by the trial Steering Committee[28] could be assessed according to the IWG criteria. Within the post-hoc analysis using the IWG criteria,[45] the confirmation period for responses was 12 weeks and analyses excluded ascites as C-findings.

A summary of best overall response, ORR and DoR results for all patients and by disease subtype by IWG criteria in the D2201 trial are provided in Table 16.

Table 16 Summary of ORR and DoR results for all patients and by disease subtype by IWG criteria in the D2201 trial (data cut-off 1st December 2014)

	Disease subtype				All patients evaluated
	ASM	SM-AHN	MCL	Subtype unknown	
Number of patients	15	72	21	5	113
Overall response: n	9	15	7	1	32
Complete remission: n (%)[a]	0 (0%)	0 (0%)	1 (5%)	0 (0%)	1 (1%)
Partial remission: n (%)[b]	5 (33%)	8 (11%)	3 (14%)	1 (20%)	17 (15%)
Clinical improvement: n (%)[c]	4 (27%)	7 (10%)	3 (14%)	0 (0%)	14 (12%)
ORR (95% CI)	60% (32 to 84%)	21% (12 to 32%)	33% (15 to 57%)	20% (1 to 72%)	28% (20 to 38%)
Median DoR (95% CI), months	36.8 (10.4 to 36.8)	NR (17.3 to NR)	NR (4.1 to NR)	NR (NR to NR)	NR (27.0 to NR)

[a] Patients with all organ damages in complete remission
[b] Patients with at least one organ damage in partial remission AND no progression on any other organ damage
[c] Patients with at least one organ damage clinically improved AND patient not in complete remission AND patient not in partial remission. A clinical improvement cannot be considered if a progression started before confirmation of clinical improvement in partial remission AND no progression on any other organ damage
ASM=aggressive systemic mastocytosis; CI=confidence interval; DoR=duration of response; IWG=International Working Group; MCL=mast cell leukaemia; NR=not reached; ORR=overall response rate; SM-AHN=systemic mastocytosis with associated haematologic neoplasm;
Source: Extracted and adapted from the CS, Appendix E, Table E.1.9, Table E.1.12

The ERG notes that for all patients, and for all disease subtypes, the ORR was lower when measured by the IWG criteria compared to the original measurement of response based on modified Valent[57] and Cheson[58,59] criteria (see Table 12 of this ERG report). However, the ERG acknowledges that due to the differences in the definitions of response according to the IWG criteria[45] and modified Valent[57] and Cheson[58,59] criteria, numerical results of ORR of the original analysis and of the post-hoc analysis may not be directly comparable.

3.4 Patient reported outcomes from the D2201 trial of midostaurin

While the A2213 trial and the D2201 trial both provided clinical effectiveness evidence for midostaurin for advanced SM, only the D2201 trial included any patient reported outcomes.

During the D2201 trial, data about patients' symptoms were collected using the Memorial Symptom Assessment Scale (MSAS).[60,61] The MSAS provides information about the frequency, severity and distress caused by 32 symptoms commonly reported by patients with cancer.

In the CS (p54), the company presents the total MSAS score (TMSAS) as the average score, across the 32 symptoms, across three domains. The TMSAS thus provides an overall score of zero to four, with a minimally important difference of 0.20–0.45.[28] The company also

presents three subscale scores from the MSAS (CS, p54): i) the physical symptom subscale (PHYS), ii) the psychological symptom subscale (PSYCH), and iii) the global distress index (GDI).

HRQoL data were also collected during the D2201 trial using the Medical Outcomes Study 12-Item Short-Form Health Survey (SF-12).[62] The SF-12 questionnaire comprises 12 items that provide two component scores, a physical component summary (PCS) score and a mental component summary (MCS) score. The mental and the physical component summary scores of the SF-12 both have a range of 0–100 and a minimal important difference of 4 points. The SF-12 scores collected during the D2201 trial were mapped to EQ-5D values and used to inform the values used in the company's base case economic analysis.

Responses to the MSAS and SF-12 questionnaires were collected at baseline, the end of treatment cycles 1-12, every 3 cycles thereafter, and at the end of treatment in the D2201 trial. The patient-reported symptoms and HRQoL results reported in the CS are derived from the earliest data cut-off (9th July 2013).

3.4.1 Summary of MSAS data

Evaluable data were available for 79 patients. The baseline values for TMSAS, PHYS, PSYCH, and GDI were compared to the values obtained at the data point when the best reported total score (TMSAS) compared to baseline was achieved. The best reported TMSAS was selected from all available data points, collected prior to the 9th July 2013 data cut-off.

The most commonly reported symptoms at baseline (CS, p54) were lack of energy (n=68, 86%), feeling drowsy (n=57, 72%), and difficulty sleeping (n=47, 60%). The company showed that the frequency for 30 of the 32 symptoms decreased from baseline to the time of best reported TMSAS (CS, p55). The frequency of two symptoms, nausea and vomiting, increased from baseline to the time of best reported TMSAS; both symptoms are known AEs associated with treatment with midostaurin. The company reported that the median TMSAS and the median scores for all subscales (PHYS, PSYCH, and GDI) were significantly lower at the time of best reported TMSAS compared to baseline (Table 17).

The ERG considers that, in the absence of MSAS data for a comparator arm, and due to the open-label design of the D2201 trial, it is difficult to interpret the results of the MSAS questionnaires. Furthermore, none of other relevant studies, identified by the literature searches,[7,12,32-41] reported any HRQoL outcomes using the MSAS tool.

Table 17 Median best percentage change in MSAS scores from baseline to time of best reported total score in D2201

Median MSAS score	Number of patients included in the analysis	Median best percentage CfB	p-value*
TMSAS	77	-58%	<0.001
PHYS	76	-67%	<0.001
PSYCH	67	-80%	<0.001
GDI	73	-69%	<0.001

* p-values determined using two-sided Wilcoxon two-sample test (t approximation)
CfB=change from baseline; GDI=global distress index; MSAS=Memorial Symptom Assessment Scale; PHYS=physical subscore; PSYCH=psychological subscore; TMSAS=total score on Memorial Symptom Assessment Scale
Source: Adapted from CS, Figure 13

3.4.2 Summary of SF-12 data

Evaluable D2201 trial data were available from 81 patients for the PCS and 80 patients for the MCS. The baseline values for the PCS and MCS were compared to the best values achieved during treatment. The best values were selected from all available data points collected prior to the 9th July 2013 data cut-off.

The company demonstrates (CS, Figure 14) that the median best values for PCS and MCS during treatment were significantly higher than the median scores at baseline ($p<0.001$). The median best percentage change from baseline to treatment with midostaurin for PCS and MCS was were 29% and 26%, respectively.

The ERG considers that in the absence of a comparator arm, and due to the open-label design of the D2201 trial, it is difficult to interpret the relative percentage change in the PCS and MCS scores.

3.5 Safety and tolerability results from midostaurin studies

Safety and tolerability data are presented in the CS (Section B.2.10), with additional information provided in Appendix F. Safety data have been derived from the D2201 and A2213 trials. Data from the D2201 trial are from the Safety Evaluation Set (SES) final OS and safety analysis (24th August 2017 data-cut off). Safety data from the A2213 trial are derived from the 1st March 2017 data cut, or (in the absence of more recent data) the 3rd December 2012 data-cut off. AEs were monitored and graded according to Common Terminology Criteria of Adverse Events (CTCAE) version 3.0. In both trials, data were collected from the first day of midostaurin administration until 28 days after discontinuation of treatment.

The ERG cautions that the differences in patient numbers, trial treatment protocols and duration of follow up should be considered when making comparisons of the safety data from the D2201 and the A2213 trials.

Exposure to study treatment

The D2201 trial and A2213 trial treatment exposure data are summarised in the CS (Table 29). The median duration of treatment exposure was [redacted] in the D2201 trial ([redacted]) than in the A2213 trial (9.8 months).

3.5.1 Adverse events

A summary of the AEs from the D2201 and A2213 trials is presented in

Table 18. All patients in the D2201 and A2213 trials experienced at least one AE of any grade. The proportion of AEs considered to be related to treatment was similar in the D2201 and A2213 trials ([redacted] and 96.2% respectively).

The rates of Grade 3 or 4 AEs were greater in the D2201 trial ([redacted]) than in the A2213 trial (61.5%). Further, a higher proportion of patients in the D2201 trial experienced serious adverse events (SAEs) compared to the proportion of patients in the A2213 trial ([redacted] and 46.2%, respectively)

Table 18 Summary of AEs from the D2201 and A2213 trials

Adverse event, n (%)	D2201 (N=116)	A2213 (N=26)
All causality AEs	[redacted]	26 (100)
Grade 3 or 4	[redacted]	16 (61.5)
Suspected to be drug-related	[redacted]	25 (96.2)
SAEs	[redacted]	12 (46.2)
Grade 3/4	[redacted]	NR
Suspected to be drug-related	[redacted]	4 (15.4)
AEs leading to discontinuation	[redacted]	4 (15.4) | 6 (23.1)[a]
Suspected to be drug-related	[redacted]	1 (3.8)
AEs leading to dose adjustment/interruption	NR	13 (50.0)
AEs leading to dose reduction	[redacted]	NR
AEs leading to dose interruption	[redacted]	NR

[a] More recent data; analysis of A2213 data cut-off: 1st March 2017; Safety Evaluation Set (SES)
Final analysis of D2201 data cut-off: 24th August 2017; SES. Analysis of A2213 data cut-off: 3rd December 2012; SES
AE=adverse event; NR=not reported; SAE=serious adverse event
Source: CS, Table 30

Treatment related adverse events

Treatment-related AEs reported in ≥10% of patients participating in the D2201 and A2213 trials are presented in Table 19. Nausea, vomiting and diarrhoea were the most commonly reported AEs in the two trials. The ERG notes that high rates of nausea and vomiting were reported in both trials, even though prophylaxis for the prevention of nausea and vomiting was recommended to be given to all patients in the D2201 and A2213 trials (CS, p91).

Table 19 Treatment-related AEs in ≥10% patients from D2201 and A2213

Adverse event, n (%)	D2201 (N=116)		A2213 (N=26)	
	Any grade	Grade 3/4	Any grade	Grade 3/4
Nausea	[redacted]	[redacted]	24 (92.3)	0 (0.0)
Vomiting	[redacted]	[redacted]	19 (73.1)	0 (0.0)
Diarrhoea	[redacted]	[redacted]	7 (26.9)	0 (0.0)
Lipase increased	NA	NA	3 (11.5)	2 (7.7)
Thrombocytopaenia	NA	NA	4 (15.4)	1 (3.8)
Fatigue	NA	NA	4 (15.4)	1 (3.8)
Headache	NA	NA	6 (23.1)	0 (0.0)
Anaemia	NA	NA	4 (15.4)	2 (7.7)

Final analysis of D2201 data cut-off: 24th August 2017; SES. Analysis of A2213 data cut-off: 3rd December 2012; Safety Evaluation Set (SES)
NA=not applicable
Source: CS Appendix F, Table F2

Deaths

The [redacted] patient deaths in the D2201 trial (on treatment or within ≥28 days after treatment discontinuation) were attributed to disease progression ([redacted]), cardiac disorders (n=5), multiple organ dysfunction syndrome (n=3), sepsis (n=3), pneumonia (n=1) and acute myeloid leukaemia (n=1).

The 5 patient deaths in the A2213 trial, were attributed to disease progression (n=2), non-neutropenic sepsis (n=2) and bacterial urinary tract infection (n=1).

The company states that none of the deaths were considered to be related to treatment.

Serious adverse events regardless of study drug relationship

SAEs experienced by ≥1% patients in the D2201 and A2213 trials (regardless of study drug relationship) are presented in the CS (Table 32 and Table 33 respectively).

In the D2201 trial, pneumonia (any grade: [redacted]; Grade 3 or 4: [redacted]) and sepsis (any grade: [redacted]; Grade 3 or 4: [redacted]) were the most frequently reported SAEs. No Grade 5 SAEs were reported.

In the A2213 trial, sepsis (n=1), febrile neutropenia (n=1), facial bone fracture due to mechanical fall (n=1), elevated total bilirubin (n=1) and hypercalcaemia (n=1) were experienced as SAEs. Two patients (CS, Table 35) experienced Grade 5 sepsis.

Adverse events leading to treatment discontinuation

A slightly higher proportion of patients in the D2201 trial discontinued treatment due to AEs ([redacted]) compared to the proportion in the A2213 trial (23.1%). In the D2201 trial, nausea

([redacted]) and QT prolongation ([redacted]) were the most commonly reported AEs that caused patients to discontinue treatment.

In the A2213 trial, sepsis was the most common AE that caused patients to stop treatment. One patient experienced sepsis as a Grade 3 or 4 AE and two patients experienced it as a Grade 5 AE.

Adverse events of special interest

The company defined severe infections, leukopenia, pulmonary toxicity, cardiac dysfunction and reproductive and developmental toxicity as AEs of special interest (AESI) and presented data for those events from the D2201 trial.

Severe infection was the most common AESI with a frequency of [redacted] and included viral upper respiratory tract infection ([redacted]), urinary tract infection ([redacted]), pneumonia ([redacted]), and upper respiratory tract infection ([redacted]). Grade 3 or 4 infections occurred in [redacted] of patients and included sepsis ([redacted]) and pneumonia ([redacted]).

Leukopenia was reported in [redacted] of patients and [redacted] of events were categorised as Grade 3 or 4 AEs. Neutropenia was the most common leucopoenia event ([redacted]) and was a Grade 3 or 4 AE for [redacted] of patients.

Pulmonary toxicity was experienced by [redacted] of patients, with [redacted] considered to be Grade 3 or 4 AE. Pleural effusion of any grade was the most frequent AE ([redacted]).

Cardiac dysfunction was reported in [redacted] of patients, of these, [redacted]events were categorised as Grade 3 or 4 AEs. The most commonly reported AE was cardiac failure ([redacted]) and [redacted] were considered to be Grade 3 or 4 AEs.

Reproductive and developmental toxicity was reported in [redacted] of patients, with [redacted] considered to be Grade 3 or 4 AEs.

ERG summary of safety results

Overall, nausea and vomiting were the most frequently reported AEs experienced by patients treated with midostaurin in the D2201 and A2213 trials. The company states that nausea and vomiting events were generally manageable with antiemetics and by administering midostaurin with food.

The ERG agrees with the company (Section B.2.10.3) that it is difficult to establish whether the Grade 3 or 4 haematological AEs reported in the trials were related to treatment with midostaurin or to disease progression.

In the absence of a control arm, the safety and tolerability data from the D2201 and A2213 trials are difficult to interpret. Clinical advice to the ERG is that AEs arising from treatment with midostaurin, as with current unlicensed treatments, require careful monitoring by a specialist clinical team with the experience to provide early recognition and management of cytoreductive therapy-related AEs, and that this can place a high burden on NHS staff and systems.

3.6 ERG critique of the additional and indirect evidence

As described in Section 3.2.1 of this ERG report, the company identified eight studies (two single-arm trials,[33,38] two case-series[32,39] and four retrospective studies[34-37,63]) that provided clinical evidence for the five comparators listed in the final scope[1] issued by NICE (cladribine, interferon alpha, nilotinib, imatinib and dasatinib). Due to the absence of comparative RCT data for midostaurin versus comparators, an indirect treatment comparison was not possible.

The company concluded that the evidence for the comparators from single arm, mostly retrospective studies, is 'much weaker' than the evidence base for midostaurin (CS, p88). The ERG agrees with the company conclusions and also agrees that results from the comparator studies should not be used to inform a decision about the comparative effectiveness of midostaurin versus the comparators. A summary and ERG critique of clinical evidence for the comparators is provided in Appendix 7.1.1 to this ERG report.

In the absence of head-to-head RCTs or other prospective controlled studies, the only available evidence demonstrating the comparative efficacy of midostaurin versus SoC is provided by two historical control studies:

- Reiter et al[12,42] study: a pooled analysis of 89 patients from the D2201 (n=63) and A2213 (n=26) trials versus historical control data for 42 patients from a German registry
- Chandesris et al[40,41] studies: a prospective observational survey of 28 patients with mastocytosis treated with midostaurin in France under a compassionate use programme, compared with 44 historical controls.

The ERG considers that the best source of indirect evidence that is available for the comparison of midostaurin versus SoC is the Reiter et al[42] presentation. A summary and ERG critique of the Reiter et al[12,42] study is provided in Section 3.6.1 of this ERG report.

The ERG also considers that the results of the Chandesris et al[40,41] study should not be used to inform a decision about the comparative effectiveness of midostaurin versus SoC due to the small number of patients and a lack of published information regarding the methods employed leading to uncertainty around the validity of the results. A summary and ERG critique of the study by Chandesris et al[40,41] is provided in Appendix 7.1.2 to this ERG report.

3.6.1 Analysis of pooled data from the D2201 and A2213 trials compared to German registry data: Reiter et al study

The OS and baseline characteristics data from the D2201 and A2213 trials (from 89 patients with a known date of SM diagnosis) were pooled and compared to a historical control group of 42 patients from a German registry. Data from the German registry were described as being 'contemporary' with the trials of midostaurin (CS, p71).

Baseline characteristics of the midostaurin group and the control group are summarised in Table 22 of the CS. The ERG agrees with the company assessment that baseline patient characteristics were similar across the two groups, with the exceptions of age and time since diagnosis; the control group were older (71% diagnosed with SM over the age of 65 compared to 42% of the midostaurin group) and had a slightly longer time from diagnosis to start of treatment (median 7.3 months since diagnosis compared to 2.2 months since diagnosis in the midostaurin group). Despite differences in trial cut-off dates (pooled data cut-off 1st July 2016 for OS) and the German registry data (data cut-off 9th May 2017 for OS), the median duration of follow-up from time of diagnosis to the data cut-off were comparable; the median duration of follow-up was 79.5 months (range 51.4 to 234 months) in the trials of midostaurin and was 84.2 months (range 22.3 to 176.3 months) in the Germany registry (CS, p70).

Four comparative analyses of OS were conducted; a description and ERG critique of these analyses is provided in Table 20 and the results of these analysis are provided in Table 21.

In response to question A1 of the clarification letter, the company provided a statistical analysis plan (SAP) that described how the pooled D2201 and the A2213 trial data were compared to the German registry data.[64] The ERG is satisfied that outcome definitions, methods of pooling data from the D2201 and A2213 trials and the statistical approaches used in the comparative analyses of OS were appropriate and were mostly pre-specified in the SAP. The ERG notes that aspects of the multivariable OS analysis which were not pre-specified in the SAP (but are summarised in Table 20) do not impact the size of the HR. The ERG also notes that the OS subgroup analyses presented in Table 24 of the CS, with the exception of sex, were not described in the SAP.

Table 20 Comparative OS analyses reported by Reiter et al

Analysis	Description	ERG comment
Primary (unadjusted)	OS was defined as the time of diagnosis to death so only patients with known dates of diagnosis were included. Patients who were alive at the end of follow-up were censored at their last date of contact.	The analysis approach, including date of diagnosis of ASM, SM-AHN or MCL was pre-defined in the SAP (Section 6.1.1 and Section 6.2.1). The ERG considers the analysis approach to be appropriate.
Multivariable[a]	OS was defined as the time of diagnosis to death as in the primary analysis with multivariable adjustment for age group at diagnosis (≤65 years vs >65 years), sex, type of disease (ASM vs SM-AHN, MCL), AHN (Yes, No or unknown), *KIT* D816 mutation status (positive, negative, unknown) and prior lines of therapy (≤1 vs >1).	The analysis approach was pre-specified in the SAP (Section 6.2.2). The ERG considers the analysis approach to be appropriate but notes that the multivariable analyses were pre-defined to be 'exploratory' and that pre-specified analyses described that a Wald two-sided p-value would be presented (SAP, Section 6.2.2). However, a one-sided p-value calculated according to the methods described in response to question A1 of the clarification letter is presented in the CS for the multivariable analysis (Table 23). Results of the multivariable analysis were used within the base case economic analysis (CS, Section 3.3.2 and Section 3.3.6) as the results of the multivariable analysis were consistent with the primary (unadjusted) analysis while adjusting for multiple baseline characteristics.
Propensity score adjusted (matched pair)	OS was defined as the time of diagnosis to death as in the primary analysis in a propensity score matched subset of patients, matched based on age group at diagnosis (≤65 years vs >65 years), sex, type of disease (ASM vs SM-AHN, MCL) and prior lines of therapy (≤1 vs >1).	Analysis approach, including variables for propensity score matching and method of matching were pre-specified in the SAP (Section 6.2.5 to Section 6.2.7). The ERG considers the analysis approach to be appropriate. The ERG acknowledges that this propensity score adjusted analysis was based on a much smaller sample size than the other comparative OS analyses and excludes 53% of patients from the trials of midostaurin.
Sensitivity analysis	OS was defined as the start date of last treatment to death. Patients who were alive at the end of follow-up were censored at their last date of contact.	The analysis approach was pre-specified in the SAP (Section 6.2.4). The ERG considers the analysis approach to be appropriate. The ERG acknowledges that this analysis was conducted to compensate for bias in the selection of the patient populations and the ERG notes this analysis included the largest number of patients from the trials of midostaurin.
Subgroup analysis	Subgroup analyses of OS by age at diagnosis (≤65 years vs >65 years), sex, type of disease (ASM, SM-AHN or MCL and SM with or without AHN) and *KIT* D816 mutation positive. OS defined as in the primary (unadjusted analysis).	With the exception of sex, these subgroup analyses were not pre-specified in the SAP (Section 5).

[a] The comparative OS analyses which adjusted for baseline characteristics were pre-defined in the SAP and described in the CS as 'multivariate' analyses. The ERG considers that a more accurate term for these analyses is 'multivariable.'[65]
ASM=aggressive systemic mastocytosis; MCL=mast cell leukaemia; OS=overall survival; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with associated haematologic neoplasm;
Source: CS, Section 2.9.1 (pp72-74); Reiter et al;[12,42] statistical analysis plan of the pooled analysis compared to German registry data[64]

Table 21 Summary of OS results reported by Reiter et al

Analysis	Midostaurin group			Registry control group		
	Patients (n)	Events (n)	Median OS (95% CI), months	Patients (n)	Events (n)	Median OS (95% CI), months
Primary (unadjusted)	89	56	41.4 (31.0 to 49.1)	42	36	19.5 (13.0 to 35.3)
	HR (95% CI): 0.50 (0.33 to 0.76); p=0.0007					
Multivariable	89	56	41.4 (31.0 to 49.1)	42	36	19.5 (13.0 to 35.3)
	HR (95% CI): 0.52 (0.32 to 0.84); p=0.0075[a]					
Propensity score adjusted (matched pair)	42	31	27.8 (19.3 to 44.6)	42	36	19.5 (13.0 to 35.3)
	HR (95% CI): 0.64 (0.33 to 1.24); p=NR					
Sensitivity analysis	115	75	28.7 (19.2 to 34.7)	39[b]	33	5.7 (2.2 to 11.7)
	HR (95% CI): 0.44 (0.29 to 0.67); p<0.0001					

[a] One-sided p-value
[b] Three patients in the German registry were not treated.
CI=confidence interval; HR=hazard ratio; NR=not reported; OS=overall survival
Source: Extracted and adapted from the CS, Table 23

The ERG agrees with the company that the results from the multivariable analysis are consistent with results from the primary (unadjusted) analysis, and also benefit from including adjustments for multiple baseline characteristics. The ERG notes that the propensity score adjusted analysis excludes 53% of patients from the trials of midostaurin.

The ERG notes that very limited information is available regarding the Germany registry data (SAP, Section 3), including the source and recruitment methods of patients into the registry. The ERG also highlights concerns raised by the company as the rationale for not using the A2213 trial results in the model:

"Treatment and study not reflective of UK clinical practice in that per the study protocol, treatment was discontinued for non-responders." (CS, Table 7)

Due to these differences in study protocols, the ERG, therefore, questions whether the pooling of data from the D2201 trials and A2213 trials was appropriate. A comparative analysis based on the D2201 trial alone, which is reflective of UK practice, may have shown different results and may have been considered a more relevant comparison to inform UK practice.

3.6.2 ERG conclusions: indirect clinical evidence

The ERG considers that the most reliable source of evidence to inform a decision about the comparative effectiveness of midostaurin versus SoC is from the multivariable analysis reported in the Reiter et al[42] presentation.

However, the ERG emphasises the following uncertainties: very limited information is available regarding the historical German registry data and, therefore, the comparability of these patients with those in the trials of midostaurin is unknown; pooling of the data from the trials of midostaurin may have been inappropriate due to differences in study protocols; and all comparative analyses of OS are based on very small numbers of patients.

3.7 Conclusions of the clinical effectiveness section

3.7.1 Direct evidence

The company provided direct clinical evidence from two single arm, open label, phase II trials of midostaurin in patients with advanced SM, the D2201 trial and the A2213 trial. The D2201 and the A2213 trials appear to be of a higher methodological quality than the studies of the comparator treatments discussed in the CS.

The results from the D2201 and A2213 trials are available for patients with ASM, SM-AHN and MCL disease subtypes and for the overall advanced SM patient population. Neither of the midostaurin trials provides direct evidence comparing the effectiveness of midostaurin versus any of the comparators listed in the final scope[1] issued by NICE.

The outcomes available from the D2201 and A2213 trials match the outcomes specified in the final scope[1] issued by NICE; however, the results from the trials are difficult to interpret given the lack of a comparator arm, the open-label design of the trials, and the small numbers of patients within each disease subtype.

3.7.2 Indirect evidence

There is no randomised clinical evidence available for the use of any of the comparators listed in the final scope[1] issued by NICE for treating patients with advanced SM. The company identified eight studies[32-39] (observational or retrospective) that reported outcomes for patients treated with the comparators. The heterogeneous designs of the comparator studies and the small and heterogeneous patient populations meant that the company was unable to conduct indirect treatment comparisons.

The only available evidence demonstrating the comparative efficacy of midostaurin is provided by two historical control studies.[12,40-42] Reiter et al[12,42] compared pooled D2201 and A2213 trial data (n=89) with data from a German registry (n=42). The ERG considers that the most reliable source of evidence to inform a decision about the comparative effectiveness of midostaurin versus SoC is from the multivariable analysis reported in the Reiter et al[42] presentation. However, the ERG emphasises that there are several areas of uncertainty relating to the Reiter et al[12,42] methods and results.

Chandesris et al[40,41] compared data from a cohort of patients receiving midostaurin in a French compassionate programme (n=28) versus French registry data (n=44). The analysis methods used by Chandesris et al[40,41] are insufficiently described leading to uncertainty around the validity of the presented results; the ERG considers that these results should not be used as a basis for decision making.

4 COST EFFECTIVENESS EVIDENCE

This section provides a structured critique of the economic evidence submitted by the company in support of the use of midostaurin for treating advanced SM. The two key components of the economic evidence presented in the CS are (i) a systematic review of the relevant literature and (ii) a report of the company's de novo economic evaluation. The company has provided an electronic copy of its economic model, which was developed in Microsoft Excel.

4.1 Published cost effectiveness evidence

4.1.1 Objective of the company's literature searches

The company undertook systematic and targeted searches to identify studies evaluating the cost effectiveness of midostaurin and other relevant interventions for the treatment of advanced SM.

4.1.2 Search strategy

The searches were carried out on 30 October 2019 and were updated on 14 November 2019. Relevant electronic databases (MEDLINE, Embase, Health Technology Assessment Database [HTAD] and the National Health Service Economic Evaluations Database [NHS-EED]) were searched and the search terms used included combinations of index terms and free text words. Searches of conference abstracts were also conducted to identify relevant abstracts published during the 2 years prior to the database searches. Abstracts relating to the following organisations/conferences were searched:

- American Society of Clinical Oncology (ASCO)
- American Society of Haematology (ASH)
- Annual Congress of the European Haematology Association (EHA)
- European Organisation for Research and Treatment of Cancer (EORTC)
- European Society for Medical Oncology (ESMO)
- International Society of Pharmacoeconomic and Outcomes Research (ISPOR): European and International meetings.

In addition, the websites of UK and international health technology appraisal (HTA) agencies were searched to identify appraisals or assessments of relevant therapies used to treat advanced SM that included descriptions of cost effectiveness models.

4.1.3 Eligibility criteria used in study selection

The eligibility criteria were designed to identify cost effectiveness studies that had been developed to estimate the cost effectiveness of midostaurin, interferon alpha, cladribine, imatinib, nilotinib or dasatinib versus any comparator for the treatment of advanced SM.

Two researchers independently screened all publications according to their title and abstract content. Any discrepancies in terms of inclusion/exclusion decisions between the researchers were resolved through discussion or the involvement of a third researcher. The same procedure was repeated when determining eligibility of the full-length articles selected during the title and abstract screening process, and for the data extraction process.

4.1.4 Findings from the company's cost effectiveness review

The company's selection strategy identified two economic evaluations.[66,67] One publication contained details about a model-based evaluation, but details about costs were not provided.[66] The other publication presented the findings from a cost effectiveness analysis conducted from an Australian health care system perspective, but the company identified that the details reported about the model were too limited for that model to be relevant to this appraisal.[67]

4.1.5 ERG comments

The ERG has updated the company's searches and is satisfied with the company's cost effectiveness literature search, study selection methods and search results. The ERG also agrees with the company's conclusion that the information in the two publications[66,67] that were identified is not sufficient to inform the development of a model for this appraisal.

The searches used by the company to identify cost effectiveness models were also used to identify HRQoL, resource and cost information that could be used to populate the company economic model. The study selection process for identifying these types of data differed only slightly from those used to identify the cost effectiveness studies.

4.2 ERG critique of the company model

4.2.1 NICE Reference Case checklist

Table 22 NICE Reference Case checklist

Element of health technology assessment	Reference Case	Does the company model adhere to the Reference Case?
Definition of the decision problem	The scope developed by NICE	Yes
Comparator(s)	As listed in the scope issued by NICE (namely, current clinical management including but not limited to: interferon alpha, cladribine, imatinib, nilotinib and dasatinib)	Partly. A basket of drugs was used to represent current clinical management. Clinical advice to the company (which was reflected by clinical advice to the ERG) was that nilotinib and dasatinib were not relevant comparators, but that peg-interferon alpha and AML-like treatments were relevant comparators.
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Yes
Perspective on costs	NHS and PSS	Yes
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	Yes
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes
Synthesis of evidence on health effects	Based on systematic review	Yes
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D-3L is the preferred measure of health-related quality of life in adults.	Yes
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Yes
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Yes
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Yes

AML=acute myeloid leukaemia; EQ-5D=EuroQol-5 dimensions; EQ-5D-3L=EuroQol-5 dimensions-3 levels; PSS=Personal Social Services; QALYs=quality adjusted life years
Source: NICE Guide to the Methods of Technology Appraisal[46]

4.2.2 Drummond checklist

Table 23 Drummond checklist for the company's economic analysis completed by the ERG

Question	Critical appraisal	ERG comment
Was a well-defined question posed in answerable form?	Yes	-
Was a comprehensive description of the competing alternatives given?	Yes	-
Was the effectiveness of the programme or services established?	No	There is no direct evidence comparing the effectiveness of treatment with midostaurin versus current clinical management. The multivariable OS HR generated by comparing pooled data from two single-arm midostaurin trials (D2201 trial and A2213 trial) with historical German registry data [42] was a key driver of cost effectiveness.
Were all the important and relevant costs and consequences for each alternative identified?	Yes	-
Were costs and consequences measured accurately in appropriate physical units?	Yes	-
Were the cost and consequences valued credibly?	Yes	-
Were costs and consequences adjusted for differential timing?	Yes	-
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	-
Was allowance made for uncertainty in the estimates of costs and consequences?	Yes	-
Did the presentation and discussion of study results include all issues of concern to users?	Yes	-

Source: Drummond checklist[68]

4.3 ERG summary of the company model

The company developed a de novo economic model to compare the cost effectiveness of midostaurin versus current clinical management (CCM) in the UK as a first-line treatment for advanced SM. The primary outcomes from the company model are incremental cost effectiveness ratios (ICERs) per quality adjusted life year (QALY) gained.

4.3.1 Populations

The modelled population is adults with advanced SM. The population comprises three disease subtypes (ASM, SM-AHN and MCL). The modelled population is consistent with the D2201 trial population and that described in the final scope[1] issued by NICE. Within the final scope[1] issued by NICE, it is stated that analysis for each subtype should be explored if there is sufficient relevant available evidence. The company has carried out analyses relating to the overall advanced SM population and to the combined SM-AHN+MCL combined subgroup.

4.3.2 Structure of the company model

The company model structure (a partitioned survival model) is shown in Figure 2. It comprises four mutually exclusive health states: progression-free with sustained response (PF-response) and progression-free with lack/loss of response (PF-no-response), progressed disease (PD) and death. The modelled population enters the model in either the PF-response health state or the PF-no-response health state depending on the presumed ORR. A fixed ORR is applied throughout the time horizon of the model. At the end of each 28-day cycle, patients in the PF-response and PF-no-response health states can remain in their respective health state or experience disease progression and move to the PD health state. In addition, patients in the PF-response health state can lose previously achieved response without having a progressed disease and thereby transit to the PF-no-response health state. Patients in the PD health state can, at the end of each cycle, remain in that health state but they cannot return to the PF-response health state or to the PF-no-response health state. Transitions to the death health state can occur from any of the other health states. Death is an absorbing health state from which transitions to other health states are not permitted.

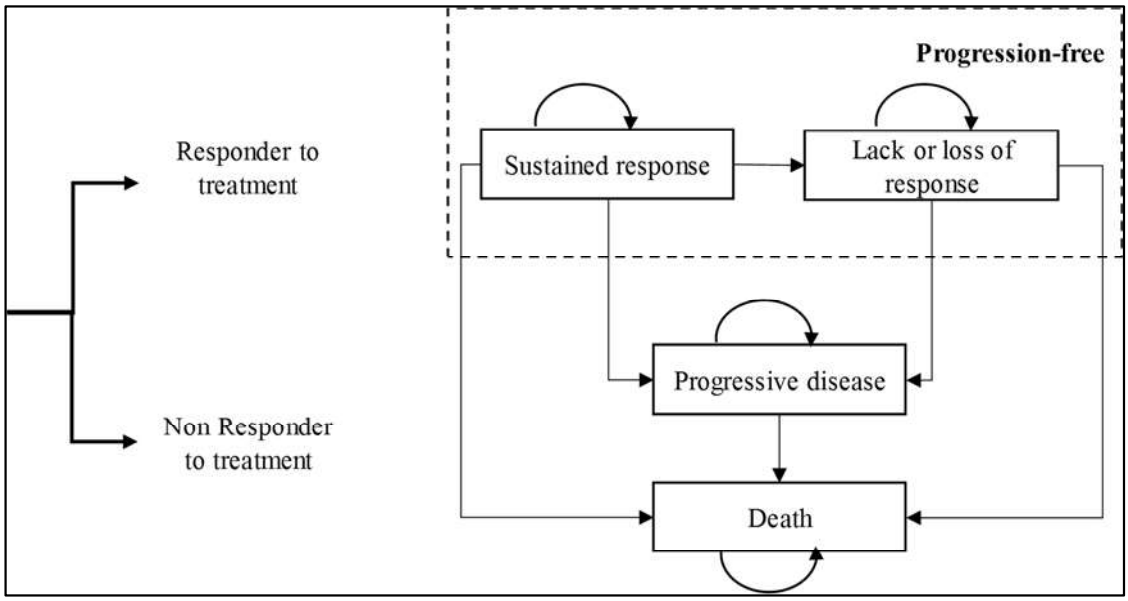

Figure 2 Structure of the company model

Source: CS, Figure 27

4.3.3 Interventions and comparators

Intervention

The intervention is midostaurin. In line with the summary of product characteristics (SmPC)[20] and the D2201 trial, the modelled dose was 100mg twice daily.

Comparator

There are no therapies licensed for the treatment of advanced SM in the UK. The comparator used in the model is CCM, which is costed as the weighted average of the five treatment options that the company considers are in use in the UK (Table 24). The weights for combining the treatment options are the proportions of patients receiving each treatment option (i.e., treatment mix), which the company estimated from discussions with five clinicians.

Table 24 Treatment mix of the five treatment options that constitute current clinical management in the model based on clinical advice to the company

Comparator	Dosing schedule	Clinical expert estimates of proportion used in the UK*	
		Advanced SM	SM-AHN+MCL
Listed in the final scope[1] issued by NICE			
Cladribine	0.14mg/kg at day 1 to day 5 of 28 day cycle[69]	[redacted]	[redacted]
Interferon alpha (Roferon-A)	3, 4, or 5 million units 3 times a week[70]	[redacted]	[redacted]
Imatinib	400mg daily[71]	[redacted]	[redacted]
Identified by UK clinical experts			
Peg-interferon alpha (Pegasys)	180 mcg per week[72]	[redacted]	[redacted]
AML-like treatments, defined as treatment typically used to treat AML	Based on TA552[73]	[redacted]	[redacted]

* Weights used to estimate costs
AML=acute myeloid leukaemia; CCM=current clinical management; kg=kilogram; mcg=microgram; MCL=mast cell leukaemia; mg=milligram; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with an associated haematological neoplasm; TA=technology appraisal
Source: CS, Table 40

Dasatinib and nilotinib are listed as relevant comparators in the final scope[1] issued by NICE. Clinical advice to the company is that these two treatments are rarely used in UK clinical practice due to the limited clinical evidence base available. The company has, therefore, not included these drugs as comparators in their base case analysis; however, the company model does include a switch that enables dasatinib and nilotinib to be included as treatment options.

4.3.4 Perspective, time horizon and discounting

The company states that costs were considered from the perspective of the NHS and Personal Social Services. The model cycle length was 28 days; a half-cycle correction was not applied as the company considered that this was not required given the cycle length. The model time horizon was set at 38.3 years and costs and outcomes were discounted at 3.5% per annum.

4.3.5 Treatment effectiveness and extrapolation in the base case

Treatment effectiveness was modelled using OS, PFS, ORR and DoR. The company fitted parametric functions to OS, PFS and DoR data from the D2201 trial to model the experience of patients treated with midostaurin. To obtain survival estimates for patients treated with CCM, the 'Reiter et al'[42] OS HR was applied to OS, PFS and DoR midostaurin data. For patients treated with midostaurin, the ORR estimate was obtained from the D2201 trial, whilst ORR estimates relating to treatment with CCM were calculated using values from the literature and assumptions.

Modelling overall survival for patients treated with midostaurin

The company fitted six parametric functions (exponential, gamma, Gompertz, log-normal, log-logistic and Weibull) and two spline hazard functions (one knot and two knots) to the OS K-M data from the D2201 trial (24th August 2017 data cut). The company identified the spline hazard one knot function as being the most appropriate function to use to represent OS. This conclusion was reached by examining goodness-of-fit statistics (Akaike Information Criterion and Bayesian Information Criterion), visual inspection and clinical opinion. In the company base case analysis, the spline hazard one knot function was used for the entire model time horizon to represent the experience of patients treated with midostaurin.

Modelling overall survival for patients receiving comparator treatments

The comparator treatment, CCM, comprised several different treatments (cladribine, interferon alpha, peg-interferon alpha, imatinib and AML-like treatments); however, in the company model there is a single representation of OS for patients receiving CCM (i.e., OS is the same irrespective of comparator treatment).

Reiter et al[42] reported results from OS analyses comparing pooled D2201 and A2213 trial data versus historical German registry data. Four different analyses were carried out:

- Unadjusted analysis (midostaurin: n=89, registry: n=42): HR=0.500 (95% CI:0.33 to 0.76)
- Propensity score matched-pair approach (midostaurin: n=42, registry: n=42): HR=0.636 (95% CI: 0.326 to 1.244)
- Multivariable approach (midostaruin: n=89, registry: n=42): HR=0.517 (95% CI: 0.319 to 0.839)
- Sensitivity analysis: time from last treatment to death (midostaurin: n=115, registry: n=39): HR=0.44 (95% CI: 0.29 to 0.67)

Chandesris et al[40,41] also reported HRs generated from the comparison of a historical cohort of patients that received treatments other than midostaurin (CEREMAST database, n=44) with

those who received midostaurin as part of a French compassionate use programme (n=28). The HRs generated by the approach taken by Chandesris et al[40,41] were lower than those reported by Reiter et al[42] (univariable [matched] approach HR=0.447, multivariable [matched] approach HR=0.333).

The company interpreted the clinical advice it received to mean that the multivariable OS HR from the study by Reiter et al[42] was the most appropriate for use in the base case analysis. This OS HR was applied to the midostaurin spline hazard one knot function OS estimates to generate OS estimates for patients treated with CCM.

Modelling progression-free survival for patients receiving midostaurin

The method used by the company to identify an appropriate parametric distribution for modelling PFS for patients receiving midostaurin was the same as the method used to identify an appropriate distribution for modelling OS. In brief, parametric functions and spline hazard functions were fitted to D2201 trial (1 December 2014 data cut) PFS K-M data. The suitability of these functions was assessed based on goodness-of-fit statistics, visual inspection and clinical opinion. Based on these assessments, the two-knot spline hazard function was selected to model the PFS experience of patients treated with midostaurin.

Modelling progression-free survival for patients receiving comparator treatments

The company did not identify any PFS data relating to patients treated with CCM. Clinical advice to the company was that, in the absence of any PFS data, it was reasonable to assume that, the OS HR used to adjust midostaurin OS estimates to represent the experience of patients treated with CCM, could be used to adjust midostaurin PFS data to represent the PFS experience of patients treated with CCM. As per the OS representation for patients receiving CCM, PFS is the same irrespective of comparator treatment.

Overall response rates

The ORR for patients treated with midostaurin (59.6%) was obtained directly from the D2201 trial, whilst the ORR for patients treated with CCM was calculated using published ORRs. Where available, subgroup (ASM, SM-AHN, MCL) ORRs were used in the model and weighted according to the population included in the D2201 trial. The ORRs used in the company model are presented in Table 25.

Table 25 Overall response rates used in the company model

Treatment	Overall response rate used in the economic model		Source		
	Overall advanced SM population	SM-AHN + MCL	ASM	SM-AHN	MCL
Midostaurin	59.5%	56.2%	D2201 trial[28]	D2201 trial[28]	D2201 trial[28]
Cladribine	[redacted]	[redacted]	Barete et al[34] Lim et al[35]	Barete et al[34]	Jawhar et al[39]
Interferon-based regimens	[redacted]	[redacted]	Lim et al[35] Hauswirth et al[32]	Hauswirth et al[32] Pardanani et al[37]	Derived from ASM and SM-AHN
Imatinib	[redacted]	[redacted]	Lim et al[35] Pardanani et al[37]	Lim et al[35] Pardanani et al[37]	Pardanani et al[37]
AML-like treatments	[redacted]	[redacted]	Assumption (same as cladribine)		

AML=acute myeloid leukaemia; ASM=aggressive systemic mastocytosis; MCL=mast cell leukaemia; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with an associated haematological neoplasm
Source: CS, Table 46

Duration of response

For treatment with midostaurin, the company obtained the DoR estimate used in the company model from the D2201 trial. The company noted that the DoR K-M data were flat for the first three cycles and, therefore, fitted parametric functions to the DoR K-M data after the third cycle. The method used to identify the most appropriate parametric distribution was the same as that used to identify the most appropriate OS and PFS distributions. The one-knot spline hazard function was selected as the preferred function. In the company model, a constraint was added to ensure than the extrapolation of DoR was consistent with that of PFS.

The company assumed that DoRs differed by comparator drug. To estimate the DoRs for the comparator drugs, the company applied the DoR HR for each drug to the parametric distribution that was used to reflect the DoR experience (one-knot hazard function) of patients receiving midostaurin. The DoR HRs for the comparator drugs were calculated from median DoRs reported in published studies (Table 26).

Table 26 Duration of response hazard ratios used in the company model

Comparator treatment	Median DoR for comparator	DoR probability for midostaurin when median reached	Derived HR	Source
Cladribine	11.0 months	[redacted]	[redacted]	Derived from Lim et al[35] and the D2201 trial
Interferon alpha/peg-interferon alpha	12.0 months	[redacted]	[redacted]	
TKIs (imatinib, nilotinib, dasatinib)	19.6 months	[redacted]	[redacted]	
AML-like treatments	11.0 months	[redacted]	[redacted]	Assumed to be the same as cladribine

AML=acute myeloid leukaemia; DoR=duration of response; HR=hazard ratio: TKI=tyrosine kinase inhibitor
Source: CS, Table 48

4.3.6 Time on treatment

Time on treatment (ToT) K-M data from the D2201 trial were used in the company model to calculate midostaurin treatment costs. As complete data were available there was no need to extrapolate the available ToT K-M data.

Based on the treatment regimen reported by Barete et al,[34] the company assumed that all patients treated with cladribine received 3.68 cycles of treatment and that 14.7% of these patients remained on treatment for an additional two cycles. The treatment cost for these cycles (3.68+0.29 cycles) was applied as a one-off cost to the first cycle.

Patients receiving interferon-based treatments (interferon alpha and peg-interferon alpha) and imatinib were assumed to be treated until disease progression, i.e., it was assumed that ToT could be modelled using PFS estimates.

4.3.7 Adverse events

Rates of Grade ≥3 AEs occurring in ≥5% of patients in the D2201 trial and A2213 trial were used to represent the experience of patients treated with midostaurin. Rates for those treated with CCM were obtained from the SmPC[69] for cladribine and from a published study[34] that evaluated the long term efficacy and safety of cladribine in patients with mastocytosis. The modelled AE rates and unit costs (obtained from previous NICE technology appraisals [TA400[74] and TA460[75]]) used in the company model are presented in Table 27.

Table 27 Adverse events (Grade 3/4) included in the company model: prevalence and unit costs

Adverse event	Unit cost		Midostaurin		CCM	
	Cost	Source	Prevalence	Source	Prevalence	Source
Nausea	£182.00	TA400[74]	[redacted]	Pooled D2201 trial and A2213 trial	16.50%	Cladribine SmPC[69]
Vomiting	£182.00	Assumed =nausea	[redacted]		7.00%	Cladribine SmPC[69]
Diarrhoea	£182.00	TA400[74]	[redacted]		7.50%	Cladribine SmPC[69]
Anaemia	£211.73	TA460[75]	[redacted]		14.00%	Cladribine SmPC[69]
Fatigue	£91.68	TA460[75]	[redacted]		25.00%	Cladribine SmPC[69]
Thrombocytopaenia	£280.28	TA460[75]	[redacted]		20.79%	Cladribine SmPC[69]
Dyspnoea	£422.41	TA460[75]	[redacted]		6.34%	Assumed same as midostaurin
Neutropenia	£808.28	TA460[75]	[redacted]		47.06%	Barete et al[34]
Infection	£517.68	TA460[75]	[redacted]	D2201 trial[28]	22.06%	Barete et al[34]
Lymphopenia	£808.28	Assumed =neutropenia	[redacted]		82.35%	Barete et al[34]

CCM=current clinical management; SmPC=Summary of Product Characteristics
Source: CS, Table 49 and Table 58

4.3.8 Health-related quality of life

Patients in the D2201 trial completed the SF-12 questionnaire[62] at baseline (on Day1, cycle 1) and then on the last day of each 28-day treatment cycle until cycle 12. Trial participants also completed the questionnaire at study completion or discontinuation if this occurred before cycle 12. Patient responses to the SF-12 questionnaire (MCS and PCS) were mapped onto EQ-5D-3L scores using the Gray et al algorithm.[76] This approach is consistent with the methods recommended in the NICE Reference Case.[46] A regression equation with progression status and response to midostaurin as covariates was then used to estimate health utility values. The equation accounted for multiple observations per patient and a manual adjustment was made to ensure that the utility value for the PF-response health state and the PF-no-response health state were higher than that the PD health state utility value. The health state utility values used in the economic model are shown in Table 28.

Table 28 Utility values used in the company model

Health state	Treatment arm	Utility value
PF-response	Midostaurin and CCM	[redacted]
PF-no-response	Midostaurin and CCM	[redacted]
Progressive disease	Midostaurin and CCM	[redacted]

CCM=current clinical management; PF=progression-free
Source: CS, Table 52

The company model also applied utility decrements to account for the discomfort associated with subcutaneous (interferon-based treatments), or intravenous (cladribine and AML-like treatments) administration routes, and the frequency of administration: cladribine (-0.041 applied to first cycle only), interferon alpha (-0.003 applied to every cycle), peg-interferon alpha (-0.002 applied to every cycle) and AML-like treatments (-0.041 applied to first cycle only). Midostaurin and imatinib were not associated with a utility decrement as these treatments are administered orally.

4.3.9 Resources and costs

Three main categories of costs were included in the company model (Section B.3.5):

- drug costs
- health state costs
- AEs.

Drug acquisition and administration costs

Confidential PAS discounts are available for midostaurin and azacitidine (an AML-like treatment). However, the PAS discount for azacitidine is not known to the company. In the CS, the company assumed that this PAS discount for azacitidine was 85%. The dosing schedules used in the company model for midostaurin and CCM drugs are reported in Section 4.3.3 of this report.

Midostaurin and imatinib are administered orally and do not have any administration costs. All other drugs are administered either subcutaneously or intravenously. Treatment with cladribine and AML-like treatments are assumed to have a one-off administration cost to the first cycle, in line with the method that was used to estimate ToT for these drugs. The administration costs associated with interferon-based treatments are applied to the ToT estimate to each model cycle. Details of intervention and comparator drug costs, including administration costs, are presented in Table 29.

The costs of AML-like treatments were applied as a one-off cost in the first cycle. These costs were obtained from the NICE STA of cytarabine+daunorubicin for previously untreated AML (TA552) and reflect use as a second-line treatment.[73]

Table 29 Drug acquisition costs (list price) and administration cost used in the company model

Drug	Vial/pack information (units per pack)	Cost per vial/pack	Vials/ packs per cycle	Cost per 28-day cycle	Admin cost	Source
Per cycle costs						
Midostaurin	25mg (56 tablets)	£5,609.94	4	£22,439.76	£0.00	D2201 trial[28]
Interferon alpha	6MU/0.5ml (1 vial)	£28.37	12	£340.44	£69.22	Lim et al[35] (single use syringe)
Peg-interferon alpha	90mcg/0.5ml (1 vial)	£76.51	4	£306.04	£23.07	Expert opinion (single use syringe)
Imatinib▲	400mg (30 tablets)	£172.29 (generic) £1,933.21 (Glivec)	1	£506.67	£0.00	eMIT (calculated)
One-off costs						
Cladribine	10mg/5ml (1 vial)	£159.50	N/A	£3,173.33	£8,634.10	Barete et al[34]
AML-like treatments (azacitidine)	-	-	-	£3,842.40*	£14,135	TA552[74]
AML-like treatments (others)	-	-	-	£6,8	£18,327	TA552[74]

▲ 79% receive generic drug and 21% receive branded drug
* 85% Patient Access Scheme discount assumed;
Admin=administration; eMIT=electronic market information tool; mcg=microgram; mg=milligram; ml=millilitre; MU=million unit
Source: CS, Table 55

The company model also included a one-off disease progression cost (of £11,807) to represent subsequent treatment costs. This cost was based on the cost of treatment with cladribine and, in the absence of evidence, this assumption was considered reasonable. In the company base case, 50% of patients were assumed to receive a subsequent therapy.

Resource use by health state

The company considered that the per-cycle cost incurred by patients who are progression-free varied over time. The company was unable to identify any UK cost study or NICE appraisal relating to advanced SM. The company, therefore, asked clinical experts (n=5) to estimate the resources used by patients. Unit costs were applied to the mean resource use estimates provided by these clinical experts. The company then calculated the per-cycle costs during the first 6 months ([redacted]). These costs were greater than the costs incurred during the following 6 months ([redacted]). The costs incurred after the first year of treatment were [redacted] per cycle. The per-cycle cost of the PD health state was estimated to be [redacted]. The resource use assumptions underpinning the health state cost calculations

are summarised in Table 30.

Table 30 Company resource use, unit costs and health state cost (per cycle)

Resource use	Unit cost	HRG code/Source	Health states resource use			
			Progression-free			PD
			0-6M	6-12M	>12M	
GP visit - surgery	£39.00	PSSRU (2019)[77]	[redacted]	[redacted]	[redacted]	[redacted]
GP visit - home visit	£100.62	PSSRU (2019)[77]	[redacted]	[redacted]	[redacted]	[redacted]
District nurse visit	£38.45	NHS Ref Cost (2017/2018)[78]: N02AF	[redacted]	[redacted]	[redacted]	[redacted]
Cancer nurse visit	£25.65	Assumed: 66.7% of community nurse cost (£38.45) as per TA400[74]	[redacted]	[redacted]	[redacted]	[redacted]
Pain and symptom management	£104.17	TA181[79] NHS Ref Cost (2017/2018)[78]: N21AF	[redacted]	[redacted]	[redacted]	[redacted]
Depression assessment and management	£81.31	TA181[79] NHS Ref Cost (2017/2018)[78]:A06A1	[redacted]	[redacted]	[redacted]	[redacted]
Outpatient visit	£194.39	NHS Ref Cost (2017/2018)[78]: face to face Clinical Oncology follow-up	[redacted]	[redacted]	[redacted]	[redacted]
ED use	£253.67	TA460[75] NHS Ref Cost (2017/2018)[78]: VB01Z, VB04Z, VB05Z, VB07Z, VB08Z	[redacted]	[redacted]	[redacted]	[redacted]
Hospitalisation days	£666.28	NHS Ref Cost (2017/2018)[78]: SA08G, SA08H, SA08J	[redacted]	[redacted]	[redacted]	[redacted]
ICU	£1,602.04	TA460[75] NHS Ref Cost (2017/2018)[78]: XC01Z - XC07Z	[redacted]	[redacted]	[redacted]	[redacted]
Bone marrow biopsy	£272.94	TA460.[75] NHS Ref Cost (2017/2018)[78]: SA33Z	[redacted]	[redacted]	[redacted]	[redacted]
ECG	£264.80	NHS Ref Cost (2017/2018)[78]: EY50Z	[redacted]	[redacted]	[redacted]	[redacted]
CT scan	£106.88	NHS Ref Cost (2017/2018)[78]: RD24Z	[redacted]	[redacted]	[redacted]	[redacted]
Chest X Ray	£106.88	Assume same as CT scan (assumption in TA400[74])	[redacted]	[redacted]	[redacted]	[redacted]
US scan	£89.08	NHS Ref Cost (2017/2018)[78]: RD24Z	[redacted]	[redacted]	[redacted]	[redacted]
MRI scan	£202.64	NHS Ref Cost (2017/2018)[78]: RD05Z	[redacted]	[redacted]	[redacted]	[redacted]
Blood test	£2.51	NHS Ref Cost (2017–2018): Directly Accessed Pathology Services, Haematology, DAPS05 (98)[78]	[redacted]	[redacted]	[redacted]	[redacted]
Bone densitometry	£71.72	NHS Ref Cost (2017/2018)[78]: RD50Z	[redacted]	[redacted]	[redacted]	[redacted]
Total cost per cycle			[redacted]	[redacted]	[redacted]	[redacted]

CT=computed tomography; ECG=electrocardiogram; ED=emergency department; GP=general practitioner; HRG=healthcare resource group; ICU=intensive care unit; MRI=magnetic resonance imaging; PD=progressed disease; PSSRU=Personal Social Services Research Unit; NHS Ref Cost=NHS Reference Cost; US=ultrasound
Source: CS, Table 57

Adverse event costs

The costs associated with AEs were obtained from recent NICE technology appraisals (TA400[74] and TA460[75]). The company estimated the cost of AEs for treatment with midostaurin and treatment with CCM to be £411 and £1,354 respectively. These costs were applied to the first cycle only.

4.3.10 Results

The company base case ICERs per QALY gained for the comparison of treatment with midostaurin versus CCM are shown in Table 31. The costs and QALYs associated with treatment with CCM were calculated using the weighted average (by treatment mix shown in Table 24) of the costs and QALYs associated with treatment with the five most common comparator drugs used in the NHS.

The company used a confidential PAS discount price when costing treatment with midostaurin and assumed that treatment with azacitidine had an 85% discount applied (azacitidine is an AML-like treatment that is available to patients in basket of comparator treatments). List prices have been used for all other treatments.

Table 31 Base cost effectiveness results for the overall advanced SM population (discounted prices for midostaurin and azacitidine)

Treatment	Total cost	Total LYG	Total QALYs	Incremental			Incremental cost per QALY gained
				Cost	LYG	QALYs	
Midostaurin	[redacted]	[redacted]	[redacted]				
CCM	£39,189	1.90	1.10	[redacted]	[redacted]	[redacted]	[redacted]

CCM=current clinical management; LYG=life years gained; QALY=quality adjusted life year; SM=systemic mastocytosis
Source: CS, Table 61

The company also presented results for treatment with midostaurin versus the individual treatment strategies that comprised CCM in scenario analyses (CS, Section B.3.8.3) as shown in Table 33 of this ERG report.

4.3.11 Subgroup analysis

The company conducted cost effectiveness analyses for the SM-AHN+MCL combined subgroup. The company base case cost effectiveness results for the comparison of treatment with midostaurin versus CCM for this subgroup are shown in Table 32.

Table 32 Base case cost effectiveness results for the SM-AHN+MCL combined subgroup (discounted prices for midostaurin and azacitidine)

Treatment	Total cost	Total LYG	Total QALYs	Incremental			Incremental cost per QALY gained
				Cost	LYG	QALYs	
Midostaurin	[redacted]	[redacted]	[redacted]				
CCM	£37,836	1.46	0.85	[redacted]	[redacted]	[redacted]	[redacted]

CCM=current clinical management; LYG=life years gained; QALY=quality adjusted life year; SM=systemic mastocytosis; SM-SM-AHN=systemic mastocytosis with an associated haematological neoplasm
Source: CS, Table 66

4.3.12 Sensitivity analyses

Deterministic sensitivity analyses

The company states that the choice of deterministic one-way sensitivity analyses (OWSAs) parameters was made a priori. Results from the company's OWSAs showed that variation (upper and lower 95% CI) of the OS HR had the greatest impact on the magnitude of the cost effectiveness results (see Figure 3).

Figure 3 Tornado diagram showing OWSA results for the comparison of treatment with midostaurin versus CCM in patients with advanced SM

CI=confidence interval; gen pop=general population; CCM=current clinical management; OS=overall survival; OWSA=one-way sensitivity analysis; PFS=progression-free survival; QALY: SM=systemic mastocytosis
Source: CS, Figure 39

Probabilistic sensitivity analyses

The results from the company's probabilistic sensitivity analysis are reproduced in Figure 4. Using the discounted price of midostaurin and azacitidine, the mean probabilistic ICER (£[redacted] per QALY gained) is [redacted] the deterministic ICER (£[redacted] per QALY gained) for the comparison of treatment with midostaurin versus CCM. The company estimated that the probability of midostaurin being a cost effective treatment option at a willingness-to-pay threshold of £50,000 per QALY gained was [redacted]% (see Figure 5).

[redacted]Figure 4 Scatter plot of the cost effectiveness of treatment with midostaurin versus CCM in patients with advanced SM (1,000 iterations)

CCM=current clinical management; PSA=probabilistic sensitivity analysis; QALYs=quality adjusted life years; SM=systemic mastocytosis; WTP=willingness-to-pay threshold
Source: CS, Figure 37

Figure 5 Cost effectiveness acceptability curve of treatment with midostaurin versus CCM at a willingness-to-pay threshold of £50,000 per additional QALY gained in patients with advanced SM

CCM=current clinical management; QALYs=quality adjusted life years; SM=systemic mastocytosis; WTP=willingness-to-pay
Source: CS, Figure 38

The company highlights [redacted].

4.3.13 Scenario analyses

The company explored several alternative scenarios for the comparison of treatment with midostaurin versus CCM. Table 33 shows selected results relating to various approaches for extrapolation of OS and PFS data. None of the scenario analyses undertaken by the company generated ICERs below £100,000 per QALY gained.

Table 33 Selected scenario analysis results – with discounted prices for midostaurin and azacitidine and list prices for other drugs

Scenario	Incremental			ICER (£/QALY)
	Costs	Life years (undiscounted)	QALYs	
Base case	[redacted]	[redacted]	[redacted]	[redacted]
Individual comparators				
Comparator=cladribine	[redacted]	[redacted]	[redacted]	[redacted]
Comparator=interferon alpha	[redacted]	[redacted]	[redacted]	[redacted]
Comparator= peg-interferon	[redacted]	[redacted]	[redacted]	[redacted]
Comparator=imatinib	[redacted]	[redacted]	[redacted]	[redacted]
Comparator=AML-like treatments	[redacted]	[redacted]	[redacted]	[redacted]
Use of piecewise extrapolation				
K-M+extrapolation	[redacted]	[redacted]	[redacted]	[redacted]
Parametric extrapolation for OS				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Lognormal	[redacted]	[redacted]	[redacted]	[redacted]
Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]
Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]
Spline2	[redacted]	[redacted]	[redacted]	[redacted]
Parametric extrapolation for PFS				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Lognormal	[redacted]	[redacted]	[redacted]	[redacted]
Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]
Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]
Spline1	[redacted]	[redacted]	[redacted]	[redacted]
HR for OS for comparator				
Reiter et al[42] - Matched	[redacted]	[redacted]	[redacted]	[redacted]
Reiter et al[42]- Unmatched	[redacted]	[redacted]	[redacted]	[redacted]
Reiter et al[42]- From last treatment	[redacted]	[redacted]	[redacted]	[redacted]
Chandesris et al[40,41] - Univariable (matched)	[redacted]	[redacted]	[redacted]	[redacted]
Chandesris et al[40,41] - Multivariable (matched)	[redacted]	[redacted]	[redacted]	[redacted]
Barete et al[34] – derived	[redacted]	[redacted]	[redacted]	[redacted]

AML=acute myeloid leukaemia; HR=hazard ratio; ICER=incremental cost effectiveness ratio; OS=overall survival; PFS=progression-free survival; QALYs=quality adjusted life years; spline1=spline hazard model with one knot; spline2=spline hazard model with two knots
Source: CS, Table 65

4.3.14 Model validation

The company states that input from clinical experts was sought during model development. Additionally, an independent reviewer stress-tested the model, assessed the model for coding errors and validated the model.

4.4 ERG critique of the company models

4.4.1 Introduction

The company provided an MS Excel model that is easy to understand and accurately represents the model structure described in the CS. The company has made significant efforts to make the best use of the available data from the D2201 trial and other relevant trials to estimate the cost effectiveness of treatment with midostaurin versus CCM. The ERG confirms that the parameter values used in the company model match the values presented in the CS and that the model algorithms are error free. The approaches used by the company to value and incorporate utility weightings, estimate resource use for different health states and summarise the AEs associated with different treatments are appropriate for use in decision making.

Data to populate the midostaurin arm of the company model were available from two single arm, open label, phase II trials, the D2201 and A2213 trials. The ERG considers that both these trials were of reasonable methodological quality and clinical advice to the ERG is that the baseline characteristics of the patients included in these trials were similar to those of NHS patients. However, the A2213 trial was very small (n=26), and the trial protocol did not reflect NHS practice as treatment for non-responders was discontinued. The company, therefore, focused on using results from the D2201 in their model. The quality of the D2201 and A2213 trials is discussed in Section 3.2.2 of this ERG report and results from these trials are presented in Section 3.3 of this ERG report.

The company has generated model cost effectiveness results for the overall population with advanced SM and for the SM-AHN+MCL combined subgroup. The comparator arm of the model is CCM, which has been modelled to comprise the weighted average of the five treatment options that the company considered were used in the UK (cladribine, interferon alpha, imatinib, peg-interferon alpha and AML-like treatments). The weight applied to each treatment was equal to the proportion of patients receiving each treatment option (as estimated from discussions with five UK clinical experts [CS, Section B.3.2.3]).

In the company model, OS and PFS for patients treated with midostaurin was modelled by fitting curves to D2201 trial data. An OS HR was used to adjust the midostaurin OS and PFS data to model the experience of patients receiving CCM. This OS HR is the key driver of cost effectiveness. This parameter is so important to the company's base case cost effectiveness results that, without a level of certainty around this estimate, discussions about other model-related concerns are largely academic. Nevertheless, the ERG has identified three additional

major concerns relating to the parameter values that were used to populate the company model:

- PFS for midostaurin versus CCM
- Partitioning survival data
- Lifetime duration of the treatment effect of midostaurin.

The ERG cautions that any cost effectiveness results generated by making changes to model parameters relating to these areas of concern could be misleading as the magnitude of their impact might be very different if used in combination with a reliable OS HR. Therefore, the ERG has only indicated (where this has been possible) whether suggested changes to these parameter values would be likely to increase or decrease the company base case cost effectiveness results.

4.4.2 Overall survival hazard ratio

As midostaurin OS data were not available from the D2201 trial for the whole model time horizon, the company carried out standard procedures to fit a range of parametric curves to the trial data. The distribution selected by the company, which was used for the whole of the model time horizon, was the spline hazard one-knot distribution. The ERG is satisfied that the distribution selection process was carried out appropriately and, whilst there is always uncertainty around projections of OS data, the distribution chosen by the company provided a good visual fit to the D2201 trial OS K-M data.

The experience of patients receiving CCM was modelled by applying a HR to midostaurin OS estimates. The OS HRs considered by the company were generated by Reiter et al,[42] Chandesris et al[41] and, by the company, using data from the D2201 trial and Barete et al[34] (Table 34). The ERG has some concerns about the reliability of the Reiter et al[42] results; nevertheless, the ERG considers that given the limited options available to the company, the results generated by Reiter et al[42] provide the most reliable estimates. Summary details relating to the ERG's critiques of the Reiter et al[42] analyses, the Chandesris et al[41] analyses and the analyses undertaken using information published by Barete et al[34] are provided in Box 2. Full details of the ERG's critiques of the Reiter et al[42] and Chandesris et al[41] analyses are provided in Section 3.6.1 and Appendix 7.1.2 respectively.

Box 2 Summary of the ERG critique relating to the derivation of the overall survival hazard ratio analyses considered for inclusion in the company cost effectiveness analyses

ERG critique of Reiter et al[42] analyses

The ERG concerns:

- differences between the D2201 and A2213 trial treatment protocols (patients in the A2213 trial who did not respond, discontinued treatment) mean that it is unclear whether it was appropriate to pool data from these two trials
- small numbers of patients (the primary [unadjusted] and multivariable analyses included data from 89 patients receiving midostaurin [D2201 and A2213 trials] and 42 patients receiving treatment other than midostaurin)
- no information about the type of treatments received by German registry patients
- no details relating to the source or recruitment methods of patients into the German registry; nor is it known whether data from all German registry patients were included in the analyses
- the partial results reported in the abstract do not completely match those described in more detail in the powerpoint presentation (which does not appear to have been peer-reviewed), albeit the results provided in the unpublished study[42] and used in the company's base case are less favourable to the cost-effectiveness case than the published study.[12]

Given the limited options available to the company, the ERG considers it was appropriate for the company to use the multivariable OS HR in their base case analysis.

ERG critique of Chandesris et al[41] analyses

The ERG concerns:

- the small numbers of patients (French compassionate programme: n=28, French registry: n=44) included some patients with SM subtypes that were not relevant to the decision problem
- that the source and recruitment methods of patients were unclear
- that the process used to match intervention and control group patients was unclear (group sizes are not proportional)
- that no statistical methods for any of the OS analyses were reported
- the OS the univariate HR estimate calculated by the company relates to a single time point (20 years), rather than being a comparative measure relating to a period of time
- the methods used to estimate the HR were not reported in detail. This adds uncertainty to the reliability of the result.

The ERG considers that analysis methods were insufficiently described and that this had led to uncertainty around the validity of the presented results. The ERG, therefore, considers that these results should not be used to inform decision making.

ERG critique of analyses carried out using data reported by Barete et al[34]

The ERG concerns:

- the analysis only included data relating to patients with ASM, and numbers were very small (D2201 trial: 16 patients, cladribine: 14 patients)
- the HR estimate related to a specific time point rather than being a comparative measure over a period of time.

The ERG considers that this OR HR result should not be used to inform decision making.

ASM=aggressive systemic mastocytosis; CS=company submission; ERG=Evidence Review Group; HR=hazard ratio; OS=overall survival; SM=systemic mastocytosis. Source: ERG report Section 3.61 and Appendix 7.1.2

The company sought advice from clinical experts to help identify the most appropriate OS HR for the overall advanced SM population and for the SM-AHN+MCL combined subgroup. For both the overall advanced SM population and for the SM-AHN+MCL combined subgroup, clinical experts considered that the OS predictions generated by the Reiter et al[42] matched pair analysis OS HR (0.636) were optimistic, but that the OS HRs lying between this OS HR and the Chandesris et al[41] matched multivariable analysis OS HR (0.333) were reasonable. They also considered that the predictions generated by the Reiter et al[42] multivariable analysis OS HR (0.517) were the most plausible. Consequently, the multivariable OS HR of 0.517 reported by Reiter et al[42] was used in the company base case analysis to model relative effectiveness for both the overall advanced SM population and for the SM-AHN+MCL combined subgroup. The ERG highlights that the range over which clinical experts considered the OS HR might be plausible is very wide.

Table 34 Published overall survival hazard ratios considered for inclusion in the company economic analyses

Source	Analysis	Number		HR		
		Midostaurin	Registry	Mean	LCI	UCI
Reiter et al[42]*	Primary analysis (unmatched)	89	42	0.500	0.330	0.760
	Propensity scoring (matched pair)	42	42	0.636	0.326	1.244
	Multivariable	89	42	0.517	0.319	0.839
	From last treatment	115	39	0.440	0.290	0.670
Chandesris et al[41]	Univariable (matched)	28	44	0.447	NR	NR
	Multivariable (matched)	28	44	0.333	NR	NR
Barete et al[34]	Unmatched (derived)	16	14	0.22	NE	NE

* These numbers are from the unpublished presentation and not from the published abstract
HR=hazard ratio; LCI=lower confidence interval; NE=not estimated; NR=not reported; OS=overall survival; UCI=upper confidence interval
Source: CS, Table 44

Results from sensitivity analyses carried out by the company show that the company's cost effectiveness estimates are very sensitive to changes in the OS HR. The company's base case ICER per QALY gained for the comparison of the cost effectiveness of midostaurin versus CCM was £[redacted]. Results from the company's deterministic sensitivity analyses showed that using the upper and lower 95% CI OS HR estimates (0.319 and 0.839 respectively) generated ICERs per QALY gained for the comparison of the cost effectiveness of midostaurin versus CCM of £[redacted] and £[redacted] respectively. Company deterministic sensitivity analysis results demonstrated that the OS HR was the key driver of cost

effectiveness results. The ERG considers that without a robust and accurate OS HR estimate, it is not possible to produce reliable cost effectiveness results. The ERG has not been able to identify a reliable OS HR and, therefore, has not generated preferred cost effectiveness results.

4.4.3 Progression-free survival for midostaurin versus current clinical management

The company was not able to identify any comparative PFS data. Therefore, based on clinical advice, the company set the PFS HR equal to the OS HR. However, it is unclear whether this assumption is reasonable and, even if it were reasonable, there is no way of knowing if the uncertainty around the OS HR would extend to the PFS estimate. The ERG could not identify any clinical evidence to support the company's PFS HR estimate or identify an alternative estimate that would be more clinically plausible. It is not known whether setting the PFS HR equal to the OS HR leads to an under- or over-estimate of the true value. Therefore, the likely direction of the impact on the ICER per QALY gained from the uncertainty around the PFS HR is also not known. However, the results of the sensitivity analysis around the PFS HR undertaken by the company suggested that, even if the PFS HR was five times higher than had been assumed in the base case analysis, it would only reduce the ICER per QALY gained for the comparison of treatment with midostaurin versus CCM by 3.6% (CS, Table 65). As such, it is unlikely that even if the true PFS HR were known, it would make a big difference to the company's cost effectiveness results.

4.4.4 Partitioning survival data

The PF health state in the company model is partitioned into a PF-response health state and a PF-no-response health state to reflect the company assumption that HRQoL differs between responders and non-responders. The data provided by the company in response to the ERG's clarification request (question B1) suggest that, whilst the D2201 trial was not powered to show a difference in PFS or OS between responders and non-responders, PFS and OS results for responders differ to those for non-responders (Figure 6). As the decision was made to partition PFS, the ERG considers that it was inconsistent not to have also partitioned OS.

Figure 6 Overall survival and progression-free survival for responders and non-responders (D2201 trial)

PFS=progression-free survival; OS=overall survival SM=systemic mastocytosis
Source: Company clarification response question B1

The company used ORR and DoR estimates to partition the PFS health state into responders and non-responders. This approach effectively used DoR as a proxy for PFS in the responder group and led to an over-estimate of PFS after 18 cycles (Figure 7). However, the ERG has concerns about the reliability of the ORR and DoR estimates used in the company model and therefore considers that it was not appropriate to use these estimates to partition the PFS health state (nor would it have been appropriate to use them to partition OS). The ERG's critique of the sources for these estimates is provided in Section 3.6.1 of this ERG report. If the D2201 trial PFS K-M curves had been stratified by response status, the issues relating to the reliability of DoR and ORR results would have been avoided.

Figure 7 Progression-free survival and duration of response for responders and non-responders (D2201 trial)

DoR=duration of response; PFS=progression-free survival; SM=systemic mastocytosis
Source: Company clarification response question B1

The effect of removing partitioning from the company model on the cost effectiveness of treatment with midostaurin versus CCM is to decrease incremental QALYs, and thus increase the size of the ICER per QALY gained. For example, using the company model, the ERG has estimated that, if the average utility value across the PF-response and PF-no-response health states (0.652) were applied to both health states, the ICER per QALY gained for the comparison of midostaurin versus CCM would increase by £[redacted] to £[redacted] per QALY gained.

4.4.5 Lifetime duration of treatment effect of midostaurin

In the D2201 trial, the median time on treatment for patients receiving midostaurin was less than 1 year and only 19% of patients were still on treatment at 3 years, yet the treatment benefit attributed to midostaurin persisted for the whole 38-year model time horizon. The ERG considers that even if the OS and PFS HRs suggested by the company were reliable, it is unlikely that treatment with midostaurin would deliver a lifetime benefit (i.e., mortality and disease progression rates for patients treated with midostaurin would be lower than the same rates for patients treated with CCM for the whole of the model time horizon). The ERG considers that it is more likely that, at some point before 38 years, the progression and

mortality rates for patients initially receiving midostaurin and those initially receiving CCM would become equal.

The effect of equalising progression and mortality rates at some point during the model time horizon on the comparative cost effectiveness of midostaurin versus CCM would be to decrease incremental QALYs and thus increase the size of the ICER per QALY gained. For example, in an ERG scenario where the progression and mortality rates of treatment with midostaurin become equal to those of CCM after 3 years, the ICER per QALY gained for the comparison of treatment with midostaurin versus CCM would increase by £[redacted] to £[redacted] per QALY gained

4.5 Conclusions of the cost effectiveness section

Concerns around the reliability of the midostaurin clinical effectiveness data (the D2201 trial is a small, open label, single arm, trial) underpin the uncertainties around the company's cost effectiveness results. The key driver of cost effectiveness results is the uncertainty associated with the OS HR used to adjust midostaurin survival data to represent the experience of patients receiving CCM. Neither the company nor the ERG were able to generate reliable cost effectiveness results for this comparison. The assumption that the treatment effect of midostaurin on OS and PFS compared to the treatment effect of CCM would last a lifetime suggests that, whilst accurate cost effectiveness results cannot be calculated, the company base case ICER per QALY gained may be underestimated.

5 END OF LIFE CRITERIA

A technology meets NICE End of Life criteria[46] if (i) the treatment is indicated for patients with a short life expectancy, normally less than 24 months and (ii) there is sufficient evidence to indicate that the treatment offers an extension to life, normally of a least an additional 3 months compared with current NHS treatment.

The ERG considers that as the advanced SM subtypes can be determined at the time of treatment commencement, the End of Life criteria should be assessed, independently, for each subtype and not for the overall population with advanced SM. However, there are insufficient data to generate reliable survival estimates for each advanced SM subtype.

Short life expectancy

The company has generated model cost effectiveness results for the overall population with advanced SM and for the SM-AHN+MCL combined subgroup. The company base case model mean OS estimate for patients treated with CCM was [redacted] months for the overall advanced SM population and [redacted] months for the SM-AHN+MCL combined subgroup. The note of caution here is that these mean OS estimates depend on the validity of the OS HR in the company model, which is currently unknown.

The median OS for German registry patients (advanced SM population) who provided data included in the analysis carried out by Reiter et al[42] was 19.5 months. However, the company quotes published data (CS, p21) that suggest that median survival from diagnosis differs substantially between the different subtypes of advanced SM. Median survival is estimated to be between 41 months[7] and 11 years[14] for patients with ASM, between 24 months[7] and 4.4 years[4] for patients with SM-AHN, and between 2 months[7] and 9.2 months[13] for patients with MCL. The company advises that published life expectancy estimates for patients with SM-AHN may be too high as they include unknown proportions of patients with indolent SM. The ERG considers that, given the paucity of data, these estimates are reasonable and highlights that the only subtype for which the short life expectancy End of Life criterion is definitely met is patients with MCL, although it may also be met for patients with SM-AHN. Assessing short life expenctancy for the combined subgroup (SM-AHN+MCL) is, therefore, problematic.

Life extension

Company model base case results for the overall population with advanced SM, suggest that mean OS for patients treated with midostaurin is 31.8 months longer than that for patients treated with CCM. Company base case model results for the SM-AHN+MCL combined subgroup, suggest that mean OS for patients treated with midostaurin is 22.6 months longer than that for patients treated with CCM. However, these results are uncertain given the lack of robust clinical effectiveness data available to describe the size of the benefit of midostaurin over CCM.

Median OS results for the overall population with advanced SM presented by Reiter et al[42] suggest that, compared with the historical control group in that study, treatment with midostaurin extends life by 21.9 months. However, there is some uncertainty around the reliability of the OS results generated by Reiter et al.[42] This uncertainty relates to the inputs (midostaurin data and German registry data) and to the differences between the results presented in the published abstract[12] and those provided in the unpublished presentation.

6 REFERENCES

1. National Institute for Health and Care Excellence (NICE). Midostaurin for treating advanced systemic mastocytosis. Final scope. Issue date: March 2020. 2020; Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10503. Accessed 4 March 2020.
2. Valent P, Akin C, Metcalfe DD. Mastocytosis: 2016 updated WHO classification and novel emerging treatment concepts. Blood. 2017; 129:1420-7.
3. Horny HP, Metcalfe DD, Escribano L, Bennett JM, Valent P. Mastocytosis. In: Swerdlow SHCE, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, editors. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Lyon: International Agency for Research on Cancer.2008. p. 54-63.
4. Cohen SS, Skovbo S, Vestergaard H, Kristensen T, Moller M, Bindslev-Jensen C*, et al.* Epidemiology of systemic mastocytosis in Denmark. Br J Haematol. 2014; 166:521-8.
5. Orphanet. Systemic Mastocytosis. 2019; Available from: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?Expert=2467. Accessed April 2020.
6. NHS. Mastocytosis symptoms. 2019; Available from: https://www.nhs.uk/conditions/mastocytosis/symptoms/. Accessed 02/10/19.
7. Lim KH, Tefferi A, Lasho TL, Finke C, Patnaik M, Butterfield JH*, et al.* Systemic mastocytosis in 342 consecutive adults: survival studies and prognostic factors. Blood. 2009; 113:5727-36.
8. Ustun C, Arock M, Kluin-Nelemans HC, Reiter A, Sperr WR, George T*, et al.* Advanced systemic mastocytosis: from molecular and genetic progress to clinical practice. Haematologica. 2016; 101:1133-43.
9. Pardanani A. Systemic mastocytosis in adults: 2019 update on diagnosis, risk stratification and management. Am J Hematol. 2019; 94:363-77.
10. Novartis Pharmaceuticals. Advanced Systemic Mastocytosis Advisory Board Meeting. 29th January 2020.
11. Novartis Pharmaceuticals. Advanced Systemic Mastocytosis Advisory Teleconference. 12th December 2019.
12. Reiter A, Kluin-Nelemans H, George T, Akin C, DeAngelo DJ, Hermine O*, et al.* Pooled survival analysis of midostaurin clinical study data (D2201+A2213) in patients with advanced systemic mastocytosis compared with historical controls. European Hematology Association; 2017; Madrid. Haematologica.
13. Budnik J, Milano MT. A registry-based analysis of survival outcomes in mast cell leukemia. Leuk Res. 2019; 78:24-8.
14. Jawhar M, Schwaab J, Álvarez-Twose I, Shoumariyeh K, Naumann N, Lübke J*, et al.* MARS: Mutation-Adjusted Risk Score for Advanced Systemic Mastocytosis. J Clin Oncol. 2019; 37:2846-56.
15. Jennings SV, Slee VM, Zack RM, Verstovsek S, George TI, Shi H*, et al.* Patient perceptions in mast cell disorders. Immunol Allergy Clin North Am. 2018; 38:505-25.
16. Siebenhaar F, von Tschirnhaus E, Hartmann K, Rabenhorst A, Staubach P, Peveling-Oberhag A*, et al.* Development and validation of the mastocytosis quality of life questionnaire: MC-QoL. Allergy. 2016; 71:869-77.
17. Gotlib J, Gerds AT, Bose P, Castells MC, Deininger MW, Gojo I*, et al.* Systemic Mastocytosis, Version 2.2019, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2018; 16:1500-37.
18. Valent P, Akin C, Hartmann K, Nilsson G, Reiter A, Hermine O*, et al.* Advances in the classification and treatment of mastocytosis: Current status and outlook toward the future. Cancer Res. 2017; 77:1261-70.

19. Casassus P, Caillat-Vigneron N, Martin A, Simon J, Gallais V, Beaudry P, *et al.* Treatment of adult systemic mastocytosis with interferon-α: results of a multicentre phase II trial on 20 patients. Br J Haematol. 2002; 119:1090-7.
20. (EMA) EMA. Rydapt Summary of Product Characterisics (SmPC). 2017; Available from: https://www.ema.europa.eu/en/documents/product-information/rydapt-epar-product-information_en.pdf. Accessed March 2020.
21. Garcia-Montero AC, Jara-Acevedo M, Teodosio C, Sanchez ML, Nunez R, Prados A, *et al.* KIT mutation in mast cells and other bone marrow hematopoietic cell lineages in systemic mast cell disorders: a prospective study of the Spanish Network on Mastocytosis (REMA) in a series of 113 patients. Blood. 2006; 108:2366-72.
22. Jawhar M, Gotlib J, Reiter A. Tyrosine Kinase Inhibitors in Systemic Mastocytosis. Mastocytosis2020. p. 257-65.
23. European Medicines Agency. Rydapt European Public Assessment Report. 2017; Available from: https://www.ema.europa.eu/en/documents/assessment-report/rydapt-epar-public-assessment-report_en.pdf. Accessed March 2020.
24. European Medicines Agency. Recommendation for maintenance of orphan designation at the time of marketing authorisation. 2017; Available from: https://www.ema.europa.eu/en/documents/orphan-review/recommendation-maintenance-orphan-designation-time-marketing-authorisation-rydapt-midostaurin_en.pdf. Accessed April 2020.
25. Valent P. Mastocytosis: a paradigmatic example of a rare disease with complex biology and pathology. Am J Cancer Res. 2013; 3:159-72.
26. Office for National Statistics (ONS). Population esimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2018. 2020; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/latest. Accessed 11/02/20.
27. National Institute for Health and Care Excellence (NICE). Midostaurin for treating advanced systemic mastocytosis. Response to consultee and commentator comments on the draft remit and draft scope (pre-referral) 2019; Available from: https://www.nice.org.uk/guidance/gid-ta10503/documents/scope-consultation-comments-and-responses. Accessed May 2020.
28. Gotlib J, Kluin-Nelemans HC, George TI, Akin C, Sotlar K, Hermine O, *et al.* Efficacy and safety of midostaurin in advanced systemic mastocytosis. N Engl J Med. 2016; 374:2530-41.
29. Novartis Pharmaceuticals. D2201 Full Clinical Study Report (26th Jun 2016)2016.
30. Novartis Pharmaceuticals. D2201 Full Clinical Study Report (27th April 2018)2018.
31. DeAngelo DJ, George TI, Linder A, Langford C, Perkins C, Ma J, *et al.* Efficacy and safety of midostaurin in patients with advanced systemic mastocytosis: 10-year median follow-up of a phase II trial. Leukemia. 2018; 32:470-8.
32. Hauswirth AW, Simonitsch-Klupp I, Uffmann M, Koller E, Sperr WR, Lechner K, *et al.* Response to therapy with interferon alpha-2b and prednisolone in aggressive systemic mastocytosis: report of five cases and review of the literature. Leukemia research. 2004; 28:249-57.
33. Hochhaus A, Baccarani M, Giles FJ, le Coutre PD, Muller MC, Reiter A, *et al.* Nilotinib in patients with systemic mastocytosis: analysis of the phase 2, open-label, single-arm nilotinib registration study. J Cancer Res Clin Oncol. 2015; 141:2047-60.
34. Barete S, Lortholary O, Damaj G, Hirsch I, Chandesris MO, Elie C, *et al.* Long-term efficacy and safety of cladribine (2-CdA) in adult patients with mastocytosis. Blood. 2015; 126:1009-16.
35. Lim KH, Pardanani A, Butterfield JH, Li CY, Tefferi A. Cytoreductive therapy in 108 adults with systemic mastocytosis: Outcome analysis and response prediction during

treatment with interferon-alpha, hydroxyurea, imatinib mesylate or 2-chlorodeoxyadenosine. Am J Hematol. 2009; 84:790-4.
36. Pagano L, Valentini CG, Caira M, Rondoni M, Van Lint MT, Candoni A*, et al.* Advanced mast cell disease: an Italian Hematological Multicenter experience. Int J Hematol. 2008; 88:483-8.
37. Pardanani A, Lim KH, Lasho TL, Finke C, McClure RF, Li CY*, et al.* Prognostically relevant breakdown of 123 patients with systemic mastocytosis associated with other myeloid malignancies. Blood. 2009; 114:3769-72.
38. Verstovsek S, Tefferi A, Cortes J, O'Brien S, Garcia-Manero G, Pardanani A*, et al.* Phase II study of dasatinib in Philadelphia chromosome-negative acute and chronic myeloid diseases, including systemic mastocytosis. Clin Cancer Res. 2008; 14:3906-15.
39. Jawhar M, Schwaab J, Meggendorfer M, Naumann N, Horny HP, Sotlar K*, et al.* The clinical and molecular diversity of mast cell leukemia with or without associated hematologic neoplasm. Haematologica. 2017; 102:1035-43.
40. Chandesris MO, Damaj G, Canioni D, Brouzes C, Lhermitte L, Hanssens K*, et al.* Midostaurin in advanced systemic mastocytosis. N Engl J Med. 2016; 374:2605-7.
41. Chandesris MO, Damaj G, Lortholary O, Hermine O. Clinical potential of midostaurin in advanced systemic mastocytosis. Blood Lymphat Cancer. 2017; 7:25-35.
42. Reiter A, Kluin-Nelemans HC, George T, Akin C, DeAngelo DJ, Hermine O*, et al.* Pooled survival analysis of midostaurin clinical study data (D2201 + A2213) in patients with advanced systemic mastocytosis compared with historical controls. Unpublished conference presentation. European Haematological Society; Madrid2017.
43. Valent P, Akin C, Sperr WR, Escribano L, Arock M, Horny HP*, et al.* Aggressive systemic mastocytosis and related mast cell disorders: current treatment options and proposed response criteria. Leuk Res. 2003; 27:635-41.
44. Jawhar M, Schwaab J, Naumann N, Horny HP, Sotlar K, Haferlach T*, et al.* Response and progression on midostaurin in advanced systemic mastocytosis: KIT D816V and other molecular markers. Blood. 2017; 130:137-45.
45. Gotlib J, Pardanani A, Akin C, Reiter A, George T, Hermine O*, et al.* International Working Group-Myeloproliferative Neoplasms Research and Treatment (IWG-MRT) & European Competence Network on Mastocytosis (ECNM) consensus response criteria in advanced systemic mastocytosis. Blood. 2013; 121:2393-401.
46. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. Process and methods [PMG9]. NICE; 2013; Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword. Accessed April 2020.
47. Downs S, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998; 52:377-84.
48. Fleming TR. One-sample multiple testing procedure for phase II clinical trials. Biometrics. 1982; 38:143-51.
49. Simon R. Optimal two-stage designs for phase II clinical trials. Control Clin Trials. 1989; 10:1-10.
50. Valent P, Horny HP, Escribano L, Longley BJ, Li CY, Schwartz LB*, et al.* Diagnostic criteria and classification of mastocytosis: a consensus proposal. Leuk Res. 2001; 25:603-25.
51. Vardiman JW, Harris NL, Brunning RD. The World Health Organization (WHO) classification of the myeloid neoplasms. Blood. 2002; 100:2292-302.
52. ECOG-ACRIN cancer research group. ECOG Performance Status. 2020; Available from: https://ecog-acrin.org/resources/ecog-performance-status. Accessed April 2020.
53. Novartis Pharmaceuticals. A2213 Trial Protocol2004.

54. Novartis Pharmaceuticals. D2201 Trial Protocol2014.
55. Fleming TR. One-sample multiple testing procedure for phase II clinical trials. Biometrics. 1982; 38:143-51.
56. Australian Government. Australian Public Assessment Report for Midostaurin. 2019; Available from: https://www.tga.gov.au/sites/default/files/auspar-midostaurin-190103.pdf. Accessed April 2020.
57. Valent P, Akin C, Sperr WR, Escribano L, Arock M, Horny HP, *et al.* Aggressive systemic mastocytosis and related mast cell disorders: current treatment options and proposed response criteria. Leuk Res. 2003; 27:635-41.
58. Cheson BD, Bennett JM, Kantarjian H, Pinto A, Schiffer CA, Nimer SD, *et al.* Report of an international working group to standardize response criteria for myelodysplastic syndromes. Blood. 2000; 96:3671-74.
59. Cheson BD, Greenberg PL, Bennett JM, Lowenberg B, Wijermans PW, Nimer SD, *et al.* Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. Blood. 2006; 108:419-25.
60. National Palliative Care Research Center. Memorial Symptom Assessment Scale Available at: http://www.npcrc.org/files/news/memorial_symptom_assessment_scale.pdf; Accessed April 2020.
61. Portenoy RK, Thaler HT, Kornblith AB, McCarthy Lepore J, Friedlander-Klar H, Kiyasu E, *et al.* The Memorial Symptom Assessment Scale: an instrument for the evaluation symptom prevalence, characteristics and distress. Eur J Cancer. 1994; 30.
62. Ware J, Kosinski M, Keller S. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. Med Care. 1996; 34:220-33.
63. Lim K-H, Tefferi A, Lasho TL, Finke C, Patnaik M, Butterfield JH, *et al.* Systemic mastocytosis in 342 consecutive adults: survival studies and prognostic factors. Blood. 2009; 113:5727-36.
64. Novartis Pharmaceuticals. Comparison of the overall survival data from studies CPKC412D2201 and CKC412A2213 with the overall survival data from a registry on rare myeloproliferative neoplasms – additional analyses for EHA. Statistical Analysis Plan 2017.
65. Hidalgo B, M. G. Multivariate or multivariable regression? . Am J Public Health. 2013 Jan; 103:39-40.
66. Cariou C, Tremblay G, Dolph M, Brandt PS, Forsythe A. Psy217 - Incremental quality-adjusted survival analysis when no head to head data are available: A case study of midostaurin (mido) versus standard of care (SoC) in patients with advanced systemic mastocytosis (ASM). Value Health. 2018; 21.
67. Pharmaceutical Benefits Advisory Committee (PBAC). Public summary document - Imatinib mesylate, tablets, 100 mg and 400 mg (base), Glivec®. 2008; Available from: http://www.pbs.gov.au/info/industry/listing/elements/pbac-meetings/psd/2008-03/pbac-psd-imatinib-mesylate-mar08. Accessed April 2020.
68. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996; 313:275-83.
69. European Medicines Agency. Litak. Summary of product characteristics (SmPC). 2018; Available from: https://www.ema.europa.eu/en/documents/product-information/litak-epar-product-information_en.pdf. Accessed.
70. eMC. Summary of product characteristics (SmPC). Roferon-A 4.5 million international units (IU) solution for injection in pre-filled syringe. 2018; Available from: https://www.medicines.org.uk/emc/search?q=Roferon. Accessed April 2020.

71. eMC. Imatinib 400mg film-coated tablets. Summary of product characterstics (SmPC). 2019; Available from: https://www.medicines.org.uk/emc/product/2297/smpc. Accessed April 2020.
72. eMC. Pegasys 180mcg solution for injection in Pre-filled Syringe. 2019; Available from: https://www.medicines.org.uk/emc/product/1697/smpc. Accessed April 2020.
73. National Institute for Health and Care Excellence (NICE). Liposomal cytarabine–daunorubicin for untreated acute myeloid leukaemia. Technology appraisal guidance [TA552]. NICE; 2018; Available from: https://www.nice.org.uk/guidance/ta552. Accessed 2020 April.
74. National Institute for Health and Care Excellence (NICE). Nivolumab in combination with ipilimumab for treating advanced melanoma. Technology appraisal guidance [TA400]. NICE; 2016; Available from: https://www.nice.org.uk/guidance/ta400. Accessed April 2020.
75. National Institute for Health and Care Excellence (NICE). Adalimumab and dexamethasone for treating non-infectious uveitis. Technology appraisal guidance [TA460]. NICE; 2017; Available from: https://www.nice.org.uk/guidance/ta460. Accessed April 2020.
76. Gray AM, Rivero-Arias O, Clarke PM. Estimating the association between SF-12 responses and EQ-5D utility values by response mapping. Med Decis Making. 2006; 26:18-29.
77. Personal Social Services Research Unit (PSSRU). Unit Costs of Health and Social Care 2019. 2019; Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/. Accessed April 2020.
78. NHS. 2017/18 Reference Costs. NHS; 2020; Available from: https://improvement.nhs.uk/resources/reference-costs/. Accessed April 2020.
79. National Institute for Health and Care Excellence (NICE). Pemetrexed for the first-line treatment of non-small-cell lung cancer. Technology appraisal guidance [TA181]. NICE; 2009; Available from: https://www.nice.org.uk/guidance/TA181. Accessed April 2020.

7 APPENDICES

7.1 Appendix 1: Additional evidence presented by the company

7.1.1 Summary of clinical evidence: comparators

Characteristics of the studies of the comparators

A summary of the disease subtypes of patients included in the trials of midostaurin and studies of comparators is provided in Table 35.

Table 35 Proportions of patients with different subtypes of SM in trials of midostaurin and studies of the comparators

Intervention	Study[b]	Disease subtype: n (% of total advanced SM)				Non-advanced SM
		ASM	SM-AHN[c]	MCL	Total advanced SM	
Midostaurin	D2201 (PEP)[28]	16 (18%)	57 (64%)	16 (18%)	89	0
	A2213[31]	3 (12%)	17 (65%)	6 (23%)	26	0
Cladribine	Barete et al[34]	14 (44%)	17 (53%)	1 (3%)	32	36
	Jawhar et al[39]	0 (0%)	0 (0%)	6 (100%)	6	0
	Lim et al[35,63]	3 (15%)	13 (85%)	0 (0%)	16	10
	Pagano et al[36]	NR[d]	NR[d]	NR[d]	3	0
	Pardanani et al[37]	0 (0%)	11 (100%)	0 (0%)	11	0
Interferon alpha[a]	Hausworth et al[32]	3 (60%)	2 (40%)	0 (0%)	5	0
	Lim et al[35,63]	14 (39%)	22 (61%)	0 (0%)	36	11
	Pagano et al[36]	NR[d]	NR[d]	NR[d]	8	0
	Pardanani et al[37]	0 (0%)	23 (100%)	0 (0%)	23	0
Imatinib	Lim et al[35,63]	4 (21%)	14 (74%)	1 (5%)	19	8
	Pagano et al[36]	NR[d]	NR[d]	NR[d]	17	0
	Pardanani et al[37]	0 (0%)	21 (100%)	0 (0%)	21	0
Nilotinib	Hochhaus et al[33]	37 (90%)	1 (2%)	3 (7%)	41	20
Dasatinib	Verstovsek et al[38]	9 (60%)	6 (40%)	0 (0%)	15	18

[a] Patients in Lim et al,[35,63] Pardanani et al[37] and four out of five patients in Hausworth et al[32] received interferon alpha plus prednisolone
[b] 'Some' patients (exact numbers not specified) in Lim et al,[35,63] Pagano et al[36] and Pardanani et al[37] received multiple treatments
[c] Unclear for studies of the comparators whether the SM-AHN disease subgroup also included patients with ISM, SSM or MCL with AHN
[d] Numbers of patients within each subtype receiving each intervention were not reported in Pagano et al.[36] For all 24 patients included in the study, 12 (50%) had ASM, 4 (17%) had SM-AHN and 8 (33%) had MCL
ASM=aggressive systemic mastocytosis; ISM=indolent systemic mastocytosis; MCL=mast cell leukaemia; NR=not reported; PEP=primary efficacy population; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with associated haematologic neoplasm; SSM=smouldering systemic mastocytosis
Source: Extracted and adapted from CS, Table 26

In terms of the total number of patients with advanced SM, data were available for 115 patients receiving midostaurin,[28,31] 72 patients receiving interferon alpha (63 of whom received interferon alpha plus prednisolone),[32,35-37,63] 68 patients receiving cladribine,[34-37,39,63] 57 patients receiving imatinib,[35-37,63] 41 patients receiving nilotinib[33] and 15 patients receiving

dasatinib.[38] Four of the studies included interventions which were not relevant comparators listed in the final scope[1] issued by NICE; midostaurin–cladribine 'mix',[39] hydroxyurea[35,37,63] and allogenic haematopoietic stem cell transplant, chemotherapy, steroids or radiotherapy.[36] Furthermore, within three of these studies,[35-37,63] it was reported that patients had received multiple interventions.

The proportion of patients with each disease subtype varied across the trials of midostaurin and across the comparator studies, and the number of patients receiving each of the comparator treatments was very small. Furthermore, four of the studies[33-35,38,63] included patients with non-advanced ASM and, within the comparator studies, it was unclear how the SM-AHN subgroup was defined. In particular, whether patients with non-advanced types of mastocytosis with AHN (such as ISM-AHN and SSM-AHN) were included within the SM-AHN disease subtype.

Characteristics of patients (age, gender and *KIT* D816V mutational status) within the comparator studies are reported in the CS (Table 26 and Appendix D, Table D.4.1.2). The company judges that, compared to the proportion in the D2201 trial, a similar proportion of *KIT* D816V positive patients was included in Barete et al[34] and Verstovsek et al.[38] Similarly, patients with a similar median age and gender distribution were included in Jawhar et al,[39] Lim et al,[35,63] Pagano et al,[36] and Pardanani et al.[37] The ERG agrees with this assessment but considers that direct comparisons between the study populations should not be made as characteristics were mostly reported only for the entire cohort of the study. The cohort may have included patients with non-advanced SM and/or patients receiving other interventions which are not relevant comparators listed within the final scope.[1]

Quality assessment of the studies of the comparators

Quality assessment of the comparator studies was undertaken using the Downs and Black checklist.[47] Assessments of each checklist item and an overview of study quality are presented in Table D.11.1 and Table D.11.2 (see CS, Appendix D).

The ERG agrees with the company assessments relating to characteristics of patients within the comparator studies and notes that the main findings of the studies and main outcome measures were clearly reported. The ERG also agrees with the company assessments that many items of the checklist were unclear or not present within the comparator studies; for example, very limited details of any individuals lost to follow-up were presented, limited adverse event results were available and for all of the studies of comparators, it is unclear how

the patients were recruited and whether the included patients within the studies were representative of the entire population from which they were recruited.

In the response to question A8 of the clarification letter (Table 2), the company explained that overall judgements of quality presented in Table 26 of the CS (good quality, reasonable quality, poor quality) were made 'crudely and qualitatively' based on the number of items marked 'yes', 'unclear' or 'no' on the checklists. It is unclear to the ERG how items marked as 'N/A' in Table 2 of the clarification letter contributed to the overall judgements, or how many items marked as 'yes' or 'no' corresponded to a judgement of good, reasonable or poor quality. Therefore, the ERG does not consider this qualitative assessment of overall quality to be appropriate.

Despite the ERG concerns regarding the assessment of overall quality, particularly relating to whether patients within the comparator studies were representative of the population of advanced SM, the ERG agrees with the company conclusion (CS, p88) that the evidence base provided by the comparator studies is weaker than the evidence provided by the trials of midostaurin.

Clinical efficacy results from the studies of comparators: overall response rate

A summary of the ORR results for all patients and by disease subtypes (where available) in the trials of midostaurin and studies of comparator treatments is provided in Table 36.

Table 36 Summary of ORR results in trials of midostaurin and studies of the comparators

Intervention	Study[b]	Overall response rate: responders / total patients, (%)			
		Disease subtype			Total advanced SM
		ASM	SM-AHN[d]	MCL	
Midostaurin	D2201 (PEP)[28]	12/16 ORR: 75%	33/57 ORR: 58%	8/16 ORR: 50%	53/89 ORR: 60%
	A2213[31]	1/3 ORR: 33%	13/17 ORR: 76%	4/6 ORR: 67%	18/26 ORR: 69%
Cladribine	Barete et al[34]	6/14 ORR: 43%	10/17 ORR: 59 %	0/1 ORR: 0%	16/32 ORR: 50%
	Jawhar et al[39]	No patients ORR: NA	No patients ORR: NA	1/6 ORR: 17%	1/6 ORR: 17%
	Lim et al[35,63]	1/2[c] ORR: 50%	6/11[c] ORR: 55%	No patients ORR: NA	NR/16[c] ORR: NR
	Pagano et al[36]	NR	NR	NR	3/3 ORR: 100%
	Pardanani et al[37]	No patients ORR: NA	6/11 ORR: 55%	No patients ORR: NA	6/11 ORR: 55%
Interferon alpha[a]	Hausworth et al[32]	1/3 ORR: 33%	2/2 ORR: 100%	No patients ORR: NA	3/5 ORR: 60%
	Lim et al[35,63]	6/10[c] ORR: 60%	9/20[c] ORR: 45%	No patients ORR: NA	NR/36[c] ORR: NR
	Pagano et al[36]	NR	NR	NR	3/8 ORR: 38%
	Pardanani et al[37]	No patients ORR: NA	11/23 ORR: 48%	No patients ORR: NA	11/23 ORR: 48%
Imatinib	Lim et al[35,63]	2/4[c] ORR: 50%	1/11[c] ORR: 9%	NR/1 ORR: NR	NR/19[c] ORR: NR
	Pagano et al[36]	NR	NR	NR	5/17 ORR: 29%
	Pardanani et al[37]	No patients ORR: NA	11/21 ORR: 52%	No patients ORR: NA	11/21 ORR: 52%
Nilotinib	Hochhaus et al[33]	8/37 ORR: 22%	NR	NR/3 ORR: NR	NR/41 ORR=NR
Dasatinib	Verstovsek et al[38]	3/9 ORR: 33%	2/6 ORR: 33%	No patients ORR: NA	5/15 ORR: 33%

[a] Patients in Lim et al,[35,63] Pardanani et al[37] and four out of five patients in Hausworth et al[32] received interferon alpha plus prednisolone

[b] 'Some' patients (exact numbers not specified) in Lim et al,[35,63] Pagano et al[36] and Pardanani et al[37] received multiple treatments

[c] ORR results are not reported for all patients in the ASM and SM-AHN subgroups and therefore ORR results are not available for the total number of patients with advanced SM in Lim et al[35,63]

[c] Unclear for studies of the comparators whether the SM-AHN disease subgroup also included patients with ISM, SSM or MCL with AHN.

ASM=aggressive systemic mastocytosis; ISM=indolent systemic mastocytosis; MCL=mast cell leukaemia; NA=not applicable; NR=not reported; ORR=overall response rate; PEP=primary efficacy population; SM=systemic mastocytosis; SM-AHN=systemic mastocytosis with associated haematologic neoplasm; SSM=smouldering systemic mastocytosis

Source: Extracted and adapted from CS, Table 27

ORR results reported in the CS were limited and varied across studies and interventions. For all patients with advanced SM, ORR ranged from 60% to 69% for midostaurin, 17% to 100% for cladribine, 38% to 60% for interferon alpha, 29% to 52% for imatinib, 33% for dasatinib and ORR for all patients with advanced SM were not reported within the study of nilotinib.[33] ORR results reported by disease subtype were also limited, particularly for the MCL disease subtype, and varied across studies and interventions.

The criteria used and the timepoints for response for the trials of midostaurin are summarised in Table 6 of this ERG report. Across the comparator studies and compared to the trials of midostaurin, the criteria used and the timepoints of response varied. Six studies used Valent criteria[33,34,36,37] or modified Valent criteria,[32,35,63] one study[39] used IWG-MRT and ECNM criteria[39] and one study[38] used Valent criteria to assess responses for patients with ASM and 'descriptive criteria' to assess responses for patients with SM-AHN. Five of the studies considered overall response as the timepoint,[32,34-37,63] one study[33] considered responses for a minimum of 4 weeks and another study[38] for a minimum of three treatment cycles, and one study[39] considered responses by month 6.

The ERG considers that direct comparisons between the ORR results for midostaurin and the comparators (and also between the comparators) should not be made:

- due to the uncertainty around the definitions of disease subtypes and the interventions patients received in the studies of the comparators (these are described in 'Characteristics of the studies of the comparators' section of this ERG report)
- due to variability of response criteria definitions and timepoints for response assessments
- due to very small numbers of patients contributing data to the ORR results, particularly for the disease subtypes.

Other clinical efficacy and safety evidence from the studies of the comparators

Other clinical efficacy and safety evidence reported in the studies of the comparators were very limited (CS, Table 27 and Table 28). In summary:

- Median (range) DoR was reported for one study[35,63] of cladribine, interferon alpha (plus prednisolone) and imatinib and the range of DoR was reported for one study of dasatinib.[38] Within these two studies, DoR results for each intervention was available only for a mixed population of patients with advanced SM and non-advanced SM.
- Very limited OS data for all patients with advanced SM or for the different disease subtypes were reported in the comparator studies. The company estimated median OS from graphical figures reported in the publications for three studies[34,36,37] of cladribine, interferon alpha (plus prednisolone) and imatinib and the number of deaths occurring were reported in two studies of nilotinib[33] and dasatinib[38] (CS, Table 28).

- None of the studies of the comparators reported PFS data. One study of cladribine reported relapse-free survival (RFS) and the company estimated the median RFS from graphical figures reported in the publication.[34]
- Adverse events (Grade 3 or 4 reported in at least 5% of patients, or events described as 'major toxicities'[32] or 'substantial side effects'[35,63]) were reported in five of the comparator studies. Adverse events were reported from mixed populations of patients with advanced SM and non-advanced SM.

The ERG considers that direct comparisons between other clinical and safety results for midostaurin and the comparators (and also between the comparators) should not be made as:

- numerical results are very limited or estimated,
- results are mainly based on mixed populations of patients with advanced SM and non-advanced SM who may be receiving multiple interventions
- results reflect very small numbers of patients.

7.1.2 Comparison of midostaurin with historical control data: the Chandesris et al study

The studies by Chandesris et al[40,41] were conducted by the French National Reference Centre for Mastocytosis; therefore, only published information relating to this study was available to the company.[40,41] Twenty-eight patients with advanced SM who received midostaurin 100mg twice daily (number of cycles unclear) under a compassionate transitory-use authorisation programme were included in the study. These midostaurin patients were compared to a control group of 44 patients who did not receive midostaurin, matched for age at diagnosis and subtype of mastocytosis via a 'logistic regression method' of propensity score matching. Patient characteristics of the midostaurin group and control group are summarised in Table 25 of the CS. A small number of patients included in the Chandesris et al[40,41] studies had subtypes of SM not relevant to the decision problem; mast-cell sarcoma (one patient in the midostaurin group and two patients in the control group) and progressive smouldering SM (two patients in each of the midostaurin and control groups). The control group were reported to have received a median of 2 (range 1 to 4) previous therapies. However, it is unclear what these therapies were and how many of the control group were receiving treatment when their data were included within the analysis compared to midostaurin patients.

The ERG considers that the source of the control group data is unclear; for example, it is not explicitly stated whether the control group patients were from the same hospital or clinic as the midostaurin treated patients and/or if they were recruited over the same time period. Furthermore, the ERG does not understand how the control group has been matched to the midostaurin group as matching should result in a control group size that is proportional to the

intervention group size. In other words, if one patient in the control group was matched to each midostaurin patient, the control group would comprise 28 patients, and if two patients in the control group were matched to each midostaurin patient then the control group would comprise 56 patients).

ORR and DoR, with treatment response assessed according to modified Valent and Cheson criteria, as in the D2201 trial, were only reported for the midostaurin group. The ORR for all midostaurin patients was 71% (median DoR 17 months, range 5 to 32 months). The ORR for the different disease subtypes were: three out of four patients with ASM (75%), thirteen out of eighteen patients with SM-AHN (72%) and two out of three patients with MCL (66%).

A comparative OS analysis was reported. The OS rate in the midostaurin group was 42.7% (95% CI: 18 to 100%) compared with 14.9% (95% CI: 6 to 36%) in the control group, corresponding to a two-fold higher hazard of death in the control group compared to the midostaurin group (HR 2.20, 95% CI: 1.08 to 4.47, p=0.02). The authors[40] also reported that in a multivariable analysis, age of diagnosis, signs of organ dysfunction and midostaurin treatment 'significantly affected OS' (p2026). However, the direction of these effects is unclear.

The findings from the Chandesris et al[40,41] studies also suggested that OS may be significantly higher for patients with ASM and MCL compared to the OS of patients with SM-AHN. The authors[41] note that the 'poor prognosis of MCL and ASM appears to be reversed by midostaurin' (p29). The ERG considers that this interpretation of the OS results is not appropriate due to the very small numbers of patients with each subtype and the lack of comparison with control data for this analysis of OS by disease subtype.

Furthermore, no statistical methods for any of the analysis of OS were reported for the Chandesris et al[40,41] studies. The ERG has been unable to verify whether the statistical approach to this comparative analysis of OS was appropriate. It is, therefore, not clear whether the results of the Chandesris et al[40,41] studies are reliable.

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

ERG report – factual accuracy check

Midostaurin for treating advanced systemic mastocytosis [ID1573]

You are asked to check the ERG report to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies, you must inform NICE by **5pm on Thursday 28 May** using the below comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

The factual accuracy check form should act as a method of detailing any inaccuracies found and how and why they should be corrected.