Single Technology Appraisal

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

Contents:

The following documents are made available to stakeholders:

The **final scope** and **final stakeholder list** are available on the NICE website.

1. **Company submission from Novartis:**
   a. Full submission
   b. Summary of Information for Patients (SIP)

2. **Clarification questions and company responses**

3. **Patient group, professional group and NHS organisation submissions** from:
   a. MPN Voice & Leukaemia Care

4. **External Assessment Report** prepared by Southampton Health Technology Assessments Centre

5. **External Assessment Report – factual accuracy check**

6. **Technical engagement response from company**

7. **Technical engagement responses and statements from experts:**
   a. Mark Hill – patient expert, nominated by MPN Voice
   b. Claire Harrison, Professor of myeloproliferative neoplasms and clinical director – clinical expert, nominated by Novartis and MPN Voice
   c. Tim Somervaille, Professor of Haematological Oncology – clinical expert, nominated by Novartis

8. **Technical engagement responses from stakeholders:**
   a. MPN Voice & Leukaemia Care

9. **External Assessment Report critique of company response to technical engagement** prepared by Southampton Health Technology Assessments Centre

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

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Ruxolitinib for treating polycythaemia vera

ID5106

Document B

Company evidence submission

October 2022

File name	Version	Contains confidential information	Date
ID5106 Ruxolitinib_Document B_25May2022_[FULLY REDACTED]	Final	Yes	25th May 2023

Instructions for companies

This is the template for submission of evidence to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Please note that the information requirements for submissions are summarised in this template; full details of the requirements for pharmaceuticals and devices are in the user guide.

This submission must not be longer than 150 pages, excluding appendices and the pages covered by this template. If it is too long it will not be accepted.

Companies making evidence submissions to NICE should also refer to the NICE health technology evaluation guidance development manual.

In this template any information that should be provided in an appendix is listed in a box.

Highlighting in the template (excluding the contents list)

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Grey highlighted text in the footer does not work as an automatic form field, but serves the same purpose – as prompt text to show where you need to fill in relevant details. Replace the text highlighted in [grey] in the header and footer with appropriate text. (To change the header and footer, double click over the header or footer text. Double click back in the main body text when you have finished.)

Contents

Tables

Figures

Abbreviations

Abbreviation	Definition
AE	Adverse event
AML	Acute myeloid leukaemia
AMI	Acute myocardial infarction
ATP	Adenosine triphosphate
AUC	Area under the curve
BAT	Best available therapy
BIC	Bayesian Information Criterion
BID	Twice daily
BNF	British National Formulary
BSH	British Society for Haematology
CEAC	Cost-effectiveness acceptability curve
CFB	Change from baseline
CHMP	Committee for Medicinal Products for Human Use
CHR	Complete haematological remission
CI	Confidence interval
CS	Company submission
CTC	Common terminology criteria
CVD	Cardiovascular disease
DSA	Deterministic sensitivity analyses
DSU	Decision Support Unit
DVT	Deep vein thrombosis
ECG	Electrocardiogram
ED	Emergency department
EFS	Event-free survival
ELN	European LeukemiaNet
EMA	European Medicines Agency
eMIT	Electronic market information tool
EORTC QLQ-C30	European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire
EQ-5D-5L	EuroQol 5 Dimension 5 Level
F2F	Face-to-face
FAS	Full analysis set
GP	General practitioner
HC	Hydroxycarbamide
HCRU	Healthcare resource use
HCT	Haematocrit
HIV	Human immunodeficiency virus
HRG	Healthcare Resource Group
HRQoL	Health-related quality of life

HU	Hydroxyurea
ICER	Incremental cost-effectiveness ratio
ICU	Intensive care unit
IFN	Interferon
IL	Interleukin
IPD	Individual patient level data
IQR	Interquartile range
IR	Incidence rate
IRR	Incidence rate ratio
ITT	Intention-to-treat
JAK	Janus kinase
KM	Kaplan–Meier
LLN	Lower limit of normal
LYG	Life years gained
MAIC	Matching-adjusted indirect comparison
MDS	Myelodysplastic syndrome
MF	Myelofibrosis
MHRA	Medicines and Healthcare products Regulatory Agency
MID	Minimally important difference
MPN	Myeloproliferative neoplasm
MRI	Magnetic resonance imaging
MU	Microgram
NA	Not applicable
NHB	Net health benefit
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMSC	Non-melanoma skin cancer
NR	Not reported
NSI	Nervous System Infections
OS	Overall survival
PAD	Peripheral artery disease
PAS	Patient access scheme
PBT	Phlebotomy
PFS	Progression-free survival
PGIC	Patient Global Impression of Change
PH	Proportional hazard
PPV-MF	Post-polycythaemia vera myelofibrosis
PR	Partial response
PRO	Patient-reported outcome
PSA	Probabilistic sensitivity analysis
PSIS	Pruritus Symptom Impact Scale
PSM	Propensity score matching/partitioned survival model

PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTA	Percutaneous Transluminal Angioplasty
PV	Polycythaemia vera
PVT	Portal vein thrombosis
PY	Patient year
QALY	Quality-adjusted life year
QD	Once daily
RBC	Red blood cell
RCT	Randomised controlled trial
RWE	Real world evidence
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of product characteristic
SRM	Standardised response mean
STAT	Signal transducer of activators of transcription
STM	State-transition model
TA	Technology appraisal
TE	Thromboembolic event
TIA	Transient ischemic attack
TNF	Tumour necrosis factor
TSS	Total symptom score
TTD	Time to treatment discontinuation
UK	United Kingdom
ULN	Upper limit of normal
US	United States
VAS	Visual analogue scale
VAT	Value-added tax
WBC	White blood cell
WHO	World Health Organization
WTP	Willingness-to-pay

B.1 Decision problem, description of the technology and clinical care pathway

Polycythaemia vera (PV) is a haematological disorder associated with a substantial and varied symptom burden, as well as a risk of serious complications[1]

PV is a chronic haematologic cancer characterised by an increased production of red and white blood cells (R/WBCs), and platelets, which leads to patients experiencing an increased risk of serious complications, such as thromboembolic and cardiovascular events, and disease progression.[1] There is also a substantial symptom burden including pruritus (itching) and severe fatigue, which can have a large impact on patient health-related quality of life (HRQoL).[2, 3] Patients with PV may also develop splenomegaly, which is associated with early satiety (being unable to eat full meals), abdominal pain, unintentional weight loss, and portal hypertension.[4] PV typically develops in adulthood, with a median diagnosis of approximately 60 years, and the majority of patients are >40 years of age.[5, 6] In the United Kingdom (UK), prevalence and incidence of PV is estimated to be approximately 6.05 per 100,000 patients and 0.5–2.2 per 100,000 per year, respectively.[7]

First-line treatment for PV consists of phlebotomy and aspirin, with high-risk patients also administered cytoreductive therapy, such as hydroxyurea (HU; also known as hydroxycarbamide [HC]).[1] However, a high proportion of patients with PV develop resistance to or intolerance (R/I) to HC/HU, which limits treatments options and increases the risk of serious complications[8]

The aim of treatment in PV is to control haematocrit (HCT) levels in order to reduce the risk of thromboembolic events and the associated complications which can lead to death. For high-risk patients (i.e. those aged ≥65 years and/or with prior PV-associated arterial or venous thrombosis), best available therapy (BAT) consists of cytoreductive therapy in addition to low-dose aspirin and phlebotomy. HC/HU is predominantly used in the first-line setting; however, approximately, 15.4%–32.2% of patients with PV develop R/I to HC/HU.[8, 9] There are currently no treatments recommended by the National Institute for Health and Care Excellence (NICE) for patients with PV who are R/I to HC/HU. However, the British Society for Haematology (BSH) recommend other cytoreductive therapies in later lines of therapy, including interferon (IFN)-alfa, busulfan, or radioactive phosphorus and continued use of HC/HU.[1]

Current cytoreductive therapies for patients with PV who are R/I to HC/HU do not fully address symptom severity and quality of life (QoL) impairments, and do not offer a tolerable safety profile.[1] Moreover, these patients continue to experience a high incidence rate of myelofibrosis (MF) and acute myeloid leukaemia (AML).[10] As a result, patients who are R/I to HC/HU have a worse life expectancy and prognosis compared with those that are not R/I.12[11, 12] The risk of thromboembolic events in this population has also been associated with high costs and healthcare resource use (HCRU) due to the substantial care and rehabilitation required following an event.[13]

There is an unmet need for an alternative, effective treatment that helps to control HCT and improve the symptom burden of PV in patients who are R/I to HC/HU, thus reducing the risk of progression to MF or AML, thromboembolic events and death and the economic burden associated with managing these events in this population. This unmet need is anticipated to be greatest in high-risk patients, for whom prognosis is particularly poor.

Ruxolitinib is an oral Janus kinase (JAK)1/JAK2 inhibitor that selectively targets the relevant biological pathway implicated in PV and is licensed for use in the treatment of adult patients with PV who are R/I of HC/HU.[14] This is in line with the positioning recommended by the BSH as a treatment option in the second- or third-line for patients with PV who are R/I to HC/HU[1] and the positioning as proposed in this submission. Given the limitations of currently available cytoreductive therapies and the present unmet need, ruxolitinib is expected to provide a targeted effective treatment with substantial clinical benefit to the population of patients PV who are R/I to HC/HU.

B.1.1 *Decision problem*

The objective of this appraisal is to determine the clinical and cost-effectiveness of ruxolitinib in line with its marketing authorisation, for the treatment of adult patients with PV who are R/I to HU therapy (also known as HC).[14] Please note that HC and HU are interchangeable as the same drug, and have been referred to throughout the submission as HC/HU.

The submission covers the technology's full marketing authorisation for this indication. The decision problem addressed in this submission is compared to that specified in the final scope issued by NICE in Table 1.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with PV that is R/I to HC/HU	In line with final scope	N/A
Subgroup	People with and without splenomegaly	In line with final scope	Additional subgroup based on MAJIC-PV population (high-risk PV)
Intervention	Ruxolitinib with established clinical management	In line with final scope	N/A
Comparator(s)	Established clinical practice without ruxolitinib, comprising of treatment with phlebotomy and aspirin, and: • HC/HU • IFN-alfa • anagrelide • busulfan • radioactive phosphorus	**Established clinical practice defined as treatment with phlebotomy and aspirin, and BAT, including:** • HC/HU • IFN-alfa • anagrelide • busulfan	Radioactive phosphorus was listed in the final scope but excluded in the submission as clinical feedback indicated that this is no longer used in the UK.[11]
Outcomes	The outcome measures to be considered include: • CHR (including reporting of HCT, WBC count and platelet count separately) • TTD • mortality • symptom relief (including a reduction in spleen size, itching, fatigue and phlebotomy) • thrombosis • progression to AML or MF • adverse effects of treatment • HRQoL	Key outcomes are: • CHR including reporting of HCT, WBC count and platelet count separately • TTD • OS • symptom relief (including a reduction in spleen size, itching, fatigue and phlebotomy) • thrombosis • safety (including transformation to AML/MF and adverse events) • HRQoL	N/A

Abbreviations: AML: acute myeloid leukaemia; BAT: Best Available Therapy; BSH: British Society for Haematology; CHR: complete haematological remission; HC/HU: hydroxycarbamide/hydroxyurea; HCT: haematocrit; HRQoL: health-related quality of life; IFN: interferon; MF: myelofibrosis; N/A: not applicable; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; OS: overall survival; PV: polycythaemia vera; R/I: resistant or intolerant; TTD: time to treatment discontinuation; WBC: white blood cell.
Source: NICE. Final scope (TA356).[15]

B.1.2 *Description of the technology being evaluated*

A description of ruxolitinib is presented in Table 2. Please refer to Appendix C for the summary of product characteristics (SmPC)[14] and the public assessment report.[16]

Table 2: Technology being appraised

UK approved name and brand name	Ruxolitinib (Jakavi®)
Mechanism of action	PV is associated with aberrant activation of JAK1 and JAK2 signalling pathways. Dysregulation of these pathways are associated with mutations such as *JAK2* V617F, which can be found in >95% of patients with PV,[17] and mutations in exon 12 of *JAK2*, which can be found in a further 3% of patients.[18] *JAK2* V617F is a gain-of-function mutation that results in expression of a constitutively active JAK2 protein. Expression of this mutation can also result in cytokine hypersensitivity and cytokine-independent growth in haematopoietic cells, potentially through the activation of the STAT family of transcription factors.[19] Mutations in *JAK2* exon 12 observed in patients with PV are also gain-of-function, further supporting that constituent activation of JAK2 signalling is central to PV pathogenesis.[19] In turn, these changes have been associated with increased production of RBCs and platelets, with associated complications such as thrombosis, AML/MDS and bleeding.[20] These symptoms, alongside development of secondary MF, are characteristic of PV.[21] Ruxolitinib is a selective inhibitor of the JAK enzymes. It works through the competitive inhibition of the ATP-binding catalytic site on JAK1 and JAK2.[21] JAK1 and JAK2 mediate the signalling of a number of cytokines and growth factors needed for haematopoiesis and immune function, such as IL-2 and IL-6 and TNF-α.[22]
Marketing authorisation/CE mark status[14]	Ruxolitinib has a UK marketing authorisation from the MHRA (originally granted by the EMA) for '*the treatment of adult patients with PV who are R/I of HU*' (date of EMA CHMP opinion: 22nd January 2015).[16]
Indications and any restriction(s) as described in theSmPC[14]	Ruxolitinib is licensed for use in the following indication: the treatment of adult patients with PV who are R/I of HC/HU.[14] This indication will be evaluated in this submission. Ruxolitinib is also currently indicated for:[14] • **MF:** Treatment of disease-related splenomegaly or symptoms in adult patients with primary MF (also known as chronic idiopathic MF), post-PV MF or post-ET MF. This indication has been evaluated by NICE and guidance has been published [TA386][23] • **Graft versus host disease:** Treatment of patients aged 12 years and older with acute/chronic graft versus host disease who have inadequate response to corticosteroids Contraindications include hypersensitivity to ruxolitinib or any of the following excipients:[14] • Cellulose, microcrystalline • Magnesium stearate • Silica, colloidal anhydrous • Sodium starch glycolate (Type A) • Povidone K30

<table>
<tr><td></td><td>

Hydroxypropylcellulose 300 to 600 cps
Lactose monohydrate

Other contraindications also include pregnancy and lactation.[14]
</td></tr>
<tr><td>Method of administration and dosage[14]</td><td>
<u>Posology</u>
The recommended starting dose of ruxolitinib for this indication (adults with PV who are R/I to HC/HU) is 10 mg orally twice daily, with a maximum dose of 25 mg twice daily.[14]

<u>Dose modification</u>
Doses may be titrated based on safety and efficacy:[14]

Treatment should be discontinued for platelet counts less than 50,000/mm^3 or absolute neutrophil counts less than 500/mm^3. Treatment should also be interrupted when haemoglobin is below 8 g/dL. After recovery of blood counts above these levels, dosing may be re-started at 5 mg twice daily and gradually increased based on careful monitoring of CBC count, including a WBC count differential
Dose reductions should be considered if the platelet count decreases below 100,000/mm^3, with the goal of avoiding dose interruptions for thrombocytopaenia. Dose reductions should also be considered if haemoglobin decreases below 12 g/dL and is recommended if it decreases below 10 g/dL. Recommended reductions are outlined in Table 3, with the aim of avoiding dose interruptions for thrombocytopenia
If efficacy is considered insufficient and blood counts are adequate, doses may be increased by a maximum of 5 mg twice daily, up to the maximum dose of 25 mg twice daily
The starting dose should not be increased within the first four weeks of treatment and thereafter no more frequently than at 2-week intervals

Table 3: Dosing recommendation for thrombocytopenia
<table>
<tr><th></th><th colspan="5">Dose at time of platelet decline</th></tr>
<tr><td></td><td>25 mg, twice daily</td><td>20 mg, twice daily</td><td>15 mg, twice daily</td><td>10 mg, twice daily</td><td>5 mg, twice daily</td></tr>
<tr><th>Platelet count</th><th colspan="5">New dose</th></tr>
<tr><td>100,000 to <125,000/mm^3</td><td>20 mg, twice daily</td><td>15 mg, twice daily</td><td>No change</td><td>No change</td><td>No change</td></tr>
<tr><td>75,000 to <100,000/mm^3</td><td>10 mg, twice daily</td><td>10 mg, twice daily</td><td>10 mg, twice daily</td><td>No change</td><td>No change</td></tr>
<tr><td>50,000 to <75,000/mm^3</td><td>5 mg, twice daily</td><td>5 mg, twice daily</td><td>5 mg, twice daily</td><td>5 mg, twice daily</td><td>No change</td></tr>
<tr><td>Less than 50,000/mm^3</td><td>Hold</td><td>Hold</td><td>Hold</td><td>Hold</td><td>Hold</td></tr>
</table>
</td></tr>
<tr><td>Additional tests or investigations[14]</td><td>Before initiating therapy with ruxolitinib, a CBC count, including WBC count differential, must be performed.
Thereafter, a CBC, including a WBC count differential, should be monitored every 2–4 weeks until ruxolitinib doses are stabilised, and then as clinically indicated, as per the ruxolitinib SmPC.</td></tr>
</table>

<table>
<tr><td>List price and average cost of a course of treatment</td><td>Acquisition cost excluding VAT sourced from BNF online (accessed 2022).[24]<table><tr><th></th><th>NHS list price</th></tr><tr><td>56 x 5 mg tablets:</td><td>£1,428.00</td></tr><tr><td>56 x 10 mg tablets:</td><td>£2,856.00</td></tr><tr><td>56 x 15 mg tablets:</td><td>£2,856.00</td></tr><tr><td>56 x 20 mg tablets:</td><td>£2,856.00</td></tr></table></td></tr>
<tr><td>Patient access scheme (if applicable)</td><td>A confidential simple PAS exists for ruxolitinib, agreed as part of the MF submission to NICE, allowing the NHS to procure ruxolitinib at a net price lower than the current list price.
Ruxolitinib is provided to the NHS with a [redacted] discount off the list price, and this PAS price has been included in the economic analysis of this submission.
Based on the need for [redacted] to allow a cost-effective price to be offered for this appraisal. Novartis are working [redacted] (subject to NICE's assessment). [redacted], results are presented in the main body of this submission using the PAS discount agreed for MF in line with the NICE method guide.[25] [redacted]</td></tr>
</table>

Abbreviations: AML: acute myeloid leukaemia; ATP: adenosine triphosphate; BNF: British National Formulary; cps: centipoise; CBC: complete blood cell count; CHMP: Committee for Medicinal Products for Human Use; EMA: European Medicines Agency; ET: essential thrombocythaemia; HC/HU: hydroxycarbamide/hydroxyurea; IL: interleukin; JAKs: Janus Associated Kinases; MDS: myelodysplastic syndrome; MF: myelofibrosis; MHRA: Medicines and Healthcare products Regulatory Agency; NHS: National Health Service; PAS: patient access scheme; PV: polycythaemia vera; RBC: red blood cell; R/I: resistant to or intolerant: SmPC: summary of product characteristics; STAT: signal transducer of activators of transcription; TNF-α: tumour necrosis factor-alpha; VAT: value-added tax; WBC: white blood cell.

B.1.3 *Health condition and position of the technology in the treatment pathway*

B.1.3.1 Disease overview

PV is a chronic haematologic cancer characterised by excessive proliferation of the erythroid, myeloid, and megakaryocytic components in the bone marrow, leading to overproduction of RBCs, WBCs, and platelets, respectively.[26] PV is part of the Philadelphia chromosome-negative group of diseases known as myeloproliferative neoplasms (MPNs), which also includes essential thrombocythaemia and primary myelofibrosis .[27]

Limited data are available regarding the epidemiology of PV in the UK. International data suggests that the annual incidence rates of PV ranges from 0.02–2.8 per 100,000 inhabitants.[28-30] Based on data from a literature review and publicly available registries, the incidence and prevalence of PV in Europe is estimated to be 0.6–2.8 per 100,000 inhabitants and 5.5–50 cases per 100,000 inhabitants, respectively.[30] In the UK, between 2000 and 2012, prevalence and incidence were estimated to be approximately 6.05 per 100,000 patients and 0.5–2.2 per 100,000 per year, respectively.[7]

Age is considered to be the main risk factor for the development of PV,[8] with disease onset typically occurring in adulthood. The median age of diagnosis is approximately 60 years, and around 75% of patients are over 40 years of age.[5, 6]

PV stems from the overactivation of the *JAK2* gene, which causes aberrant regulation of the JAK-signal transducer and activator of transcription (STAT) pathway.[19] This dysregulation may arise due to mutations in the *JAK2* gene. For example, a *JAK2* mutation that causes a valine-to-phenylalanine substitution at position 617 (*JAK2* V617F) is identified in approximately 96% of patients with PV.[18] For patients without *JAK2* V617F, the most commonly carried mutation is in exon 12 of *JAK2,* which is found in a further 3% of patients.[18] These changes have been associated with increased RBC mass (erythrocytosis), elevated WBC and elevated platelet counts, which leads to a high frequency of burdensome symptoms and vascular complications.[26]

Consequently, parameters such as HCT level are critical to diagnosing patients with PV. HCT levels are used to indicate the volume percentage of RBCs in the blood, with levels greater than 60% being associated with an increased RBC mass.[31] It is currently recommended by the BSH that individuals with persistently raised venous HCT (males, >0.52; females, >0.48) should be investigated for PV.[1] In addition, initial observations in patients with PV suggested a higher risk of thrombosis at moderately increased HCT levels, further consolidating the association between increased RBC mass and vascular complications.[32]

Alongside HCT levels, raised WBC counts can also inform on PV diagnosis as well as disease risk stratification.[1] Leucocytosis is typically indicated by increased WBC counts above the normal range. As these WBCs play a central role in the activation of the blood coagulation system, leucocytosis has been found to be an independent risk factor for thrombosis and contributes to the increased risk of cardiovascular events in patients with PV.[33-35]

Symptom burden

PV is associated with a wide variety and high frequency of debilitating symptoms including headaches (often migrainous), visual disturbances, tinnitus, dizziness, difficulty concentrating, transient weakness or pins and needles, burning pains in the extremities, pruritus (itching), and night sweats and fatigue.[36-39] Overproduction of RBCs, WBCs and platelets may also cause an enlarged spleen size, leading to approximately 18% to 38% of patients with PV experiencing splenomegaly.[9, 40, 41] In turn, splenomegaly is associated with early satiety (being unable to eat full meals), abdominal pain, unintentional weight loss, and portal hypertension.[4]

Of the symptoms experienced by patients with PV, pruritus is considered to be particularly burdensome, with specific studies in PV highlighting its profoundly negative effect on HRQoL.[2, 3] Pruritus is characterised by generalised burning, prickling, tingling, or itching sensations and impacts 50–65% of patients with PV.[37-39] Pruritus is frequently observed following high-temperature water contact (aquagenic pruritus) or large temperature shifts, alcohol consumption, and exercise.[42]

In a study of 441 patients with PV, results from the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire (EORTC QLQ-C30) indicated significant reductions in global health status (56 versus 64.1; p=0.0007) and numerous functional (role, emotional, cognitive, and social functioning) and symptom scales (dyspnoea, fatigue, pain, and financial difficulties) for patients with aquagenic pruritus versus those without.[3] Moreover, aquagenic pruritus was classified as “unbearable” in 14.6% of patients. Twenty-four percent of

patients received aquagenic pruritus-specific treatments, with improvement reported in only about 40%. In contrast, 51% of patients reported that their pruritus was the same or worse after PV therapy, suggesting that current treatments, including antihistamines, selective serotonin reuptake inhibitors, steroids, and external therapy options (e.g., oil lotions and ultraviolet-based treatment), have limited effectiveness. Pruritus has also been associated with negative emotions such as aggression, irritability, depression, and suicidal ideation.[42, 43]

In another study using the abbreviated (10-item) Myeloproliferative Neoplasm Symptom Assessment Form (MPN-SAF) total symptom score, fatigue was the symptom that was assigned the highest severity score by all patients. Most patients with PV reported feeling anxious or worried about their disease (78%) and that their symptoms reduced their QoL (~65%). Moreover, many patients reported that their disease interfered with their activities of daily living, impacting their family and/or social life and forcing them to cancel plans.[44]

Fatigue was also identified as the most common symptom amongst UK-based patients with PV (n=78) in the Landmark Health Survey, an international online survey of patients with myeloproliferative neoplasms. This was followed by difficulty sleeping, night sweats, loss of concentration, pruritus, shortness of breath and dizziness.[45]

Disease specific mortality

In addition to burdensome symptoms, patients with PV experience an increased risk of thromboembolic and cardiovascular events. This risk is highest in patients who are older, have a history of thrombosis, or have cardiovascular risk factors (including smoking, poor diet, obesity and high blood pressure).[46] Such complications may include stroke, transient ischemic attack (TIA; 22.1%), acute myocardial infarction (20.7%), pulmonary embolism (6.1%), peripheral arterial thrombosis (12.3%), deep venous thrombosis (14.8%), and haemorrhage.[5, 47, 48] Patients with PV may also progress to leukaemic diseases such as MF (scarring of the bone marrow) or AML/myelodysplastic syndrome (MDS), which occur in approximately 10% and 6% of patients with PV, respectively.[49]

Both vascular events and disease progression substantially elevate the risk of morbidity and death in affected individuals and thus are associated with a poor prognosis in PV.[12, 35] As a result, patients with PV have a 1.6-fold higher risk of death than the general population,[50] with median survival of these patients estimated as 14.1 years. Furthermore, estimated median survival is considered dependent on certain risks factors and lowers to only 8.3 years if patients have a history of thrombosis and are older than 60.[5]

High-risk patients with PV and R/I to HC/HU

Although patients with PV experience a high symptom burden and high-risk of morbidity and death, limited management options exist to treat PV. In order to determine the correct course of treatment (see Section B.1.3.2), patients are typically stratified into low- or high-risk categories based on their history of thrombotic events. Young age and a lack of prior thrombosis define the low-risk category, whereas ages older than 65 years and/or prior thrombosis define the high-risk category.[1, 46] However, other aspects such as cardiovascular risk factors, elevated WBC count and extreme HCT uncontrolled with phlebotomy can also influence risk stratification.[1] Limited UK data are available on the proportion of patients who are high-risk. However, a 2018 US based real-world evidence study suggested that these patients may represent approximately 64% (1,823/2,856) of the population of patients with PV.[13]

The BSH guidelines recommend that all patients with PV receive phlebotomy and low-dose aspirin in the first-line, with high-risk patients also administered cytoreductive therapy. HC/HU represents the mostly commonly prescribed cytoreductive therapy in the first-line.[1] However, a considerable percentage of patients develop R/I to HC/HU; resistance is defined by the European Leukemia Net (ELN) as an insufficient clinical response to HC/HU (at 1.5 g per day for at least four months and without reporting intolerance) as determined by at least one of the criteria outlined in Table 4. The ELN consider patients to be intolerant to HC/HU if they experience Grade 3–4 or prolonged Grade 2 non-haematological toxicity (e.g. symptoms, fever, pneumonitis) and/or haematological toxicity, as described in Table 4. Additionally, patients who develop vascular events (including clinically relevant bleeding, venous thrombosis, or arterial thrombosis) or non-melanoma skin cancers (NMSC) are recommended by the ELN to discontinue HC/HU.[10]

It should be noted that while the ELN criteria for R/I to HC/HU are often used to inform eligibility criteria for clinical trials, UK clinical expert opinion indicated that these criteria are not used in clinical practice.[11] Rather, clinician's review a patient's history and use their own judgement of symptoms such as fatigue and other subtle symptoms to determine whether a patient is R/I to HC/HU.[11]

Table 4: ELN definition of R/I to HC/HU in patients with PV

Resistance – defined as an insufficient clinical response to HC/HU (at =1.5 g per day for at least four months and without reporting intolerance), according to at least one of the following criteria:	
1	Persistent disease-related symptoms: a total symptom score of at least 20 or an itching score of at least ten for at least six months
2	Persistent thrombocytosis: a platelet count >1000 × 109 cells per L, microvascular symptoms, or both, persisting for more than three months
3	Symptomatic or progressive splenomegaly: increased in spleen size by more than 5 cm from the left costal margin in one year
4	Progressive (at least 100% increase if baseline count is <10 × 109 cells per L or at least 50% increase if baseline count is >10 × 109 cells per L) and persistent leukocytosis (leukocyte count >15 × 109 cells per L confirmed at three months
5	Insufficient HCT: need for six or more phlebotomies per year to keep HCT <45%
Intolerance – defined according to the following criteria:	
1	Grade 3–4 or prolonged Grade 2 non-haematological toxicity (e.g., mucocutaneous manifestations, gastrointestinal symptoms, fever, or pneumonitis) at any dose
2	Intolerance to HC/HU because of haematological toxicity (haemoglobin <100 g/L, platelet count <100 × 109 cells per L, or neutrophil count <1 × 109 cells per L) at the lowest dose of HC/HU to achieve a response

Abbreviations: ELN: European Leukemia Net; HCT: haematocrit; HC/HU: hydroxycarbamide/hydroxyurea; PV: polycythaemia vera.
Source: Marchetti et al. 2022.[10]

Limited UK data are available on the proportion of patients who develop R/I to HC/HU, and studies across Europe vary in their estimates. In a study investigating 890 patients from the Spanish Registry of PV, 15.4% of patients met at least one of the ELN criteria* defining R/I to HC/HU (Barosi et al. 2010).[9, 51] However, in another retrospective study in Germany, of the 950 patients evaluable

* Please note the ELN criteria for R/I to HU have since been updated as shown in Table 4.

for R/I to HC/HU, 32.2% were resistant to HC/HU treatment and 17.7% were intolerant according to criteria reviewed by Vannucchi et al. in 2014.[8, 52]

R/I to HC/HU further limits treatment options available to patients with PV, especially for those within the high-risk subgroup. As a result, UK clinical experts indicated that patients who are R/I to HC/HU may have poorer HRQoL and a worse life expectancy and prognosis compared with those who are not.[11] This opinion is further supported by published European data; one study demonstrated that patients who are resistant to HC/HU have a 6.8% increased risk of disease progression to AML or MF (hazard ratio [HR]: 6.8; 95% confidence intervals [CI]: 3.0, 15.4; p<0.001) and a 5.6% increased risk of death (HR: 5.6; 95% CI: 2.7, 11.9; p<0.001) compared to non-resistant patients.[12]

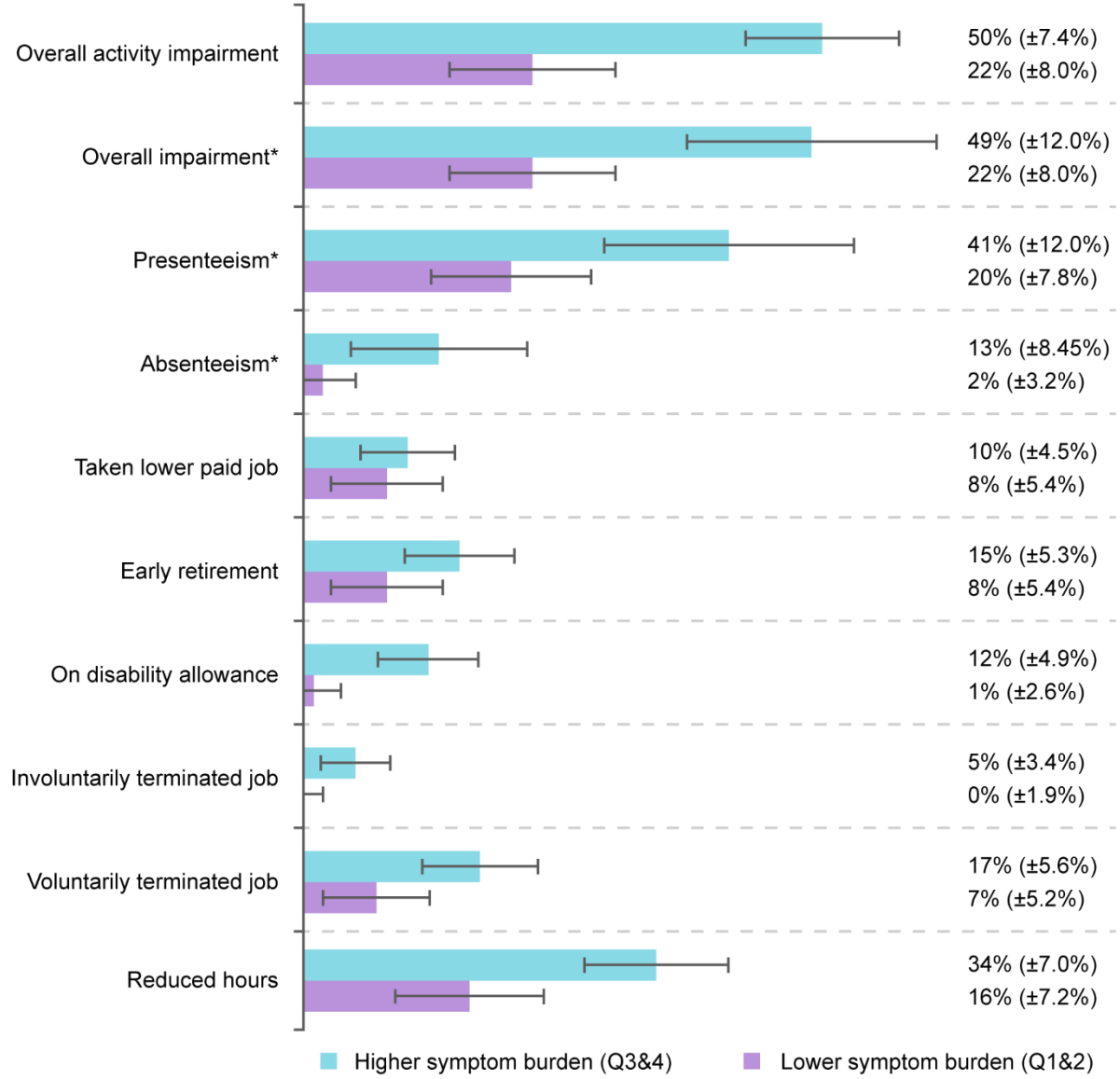
Economic burden and resource use

There are limited data available regarding the economic burden of PV and other myeloproliferative neoplasms. However, the management of symptoms, disease progression, and thromboembolic and cardiovascular events in the PV population is considered to impose a high economic and resource burden on healthcare systems per person. A 2011 study of claims data of more than 25,000 patients found that the medical costs for patients with MPNs are up to six times higher than the medical costs incurred by patients with other non-cancer conditions.[53]

Annual all-cause medical costs have been found to be three times higher in patients with PV treated with HC/HU compared with the total PV population, potentially reflecting the inherently more severe disease in the high-risk patients who are eligible to receive HC/HU. These patients remain at risk of thromboembolic events, which have been associated with substantially high costs and HCRU in patients with PV due to the substantial care and rehabilitation required following an event.[13] Moreover, a European real-world evidence study found that elevated leukocyte counts (a criterion for R/I to HC/HU according to the ELN definition; see Table 4)[10] have been associated with increased HCRU and costs compared with patients without leukocytosis, likely due to the increased rates of myelofibrotic transformation and thromboembolic complications in this population.[54] Therefore, given that the prognosis for patients with PV who are high-risk and/or R/I is particularly poor,[12, 35] the care of these individuals is anticipated to impose a disproportionate economic burden and HCRU compared with the general PV population.

As well as high medical costs, the high symptom burden of PV and its impact on HRQoL result in high indirect costs to both patients and society. In the Landmark Health Survey, a 2016 UK based survey of 286 patients with MPNs (including 78 patients with PV) and 31 treating physicians found that the impact of PV on patients' work productivity and activity impairment was high. This impact was highest in patients with a higher symptom burden compared to patients with a lower symptom burden, as shown in Figure 1.[55] Given that the burden of symptoms is particularly high in patients with PV who are high-risk and/or R/I to HC/HU, the indirect costs associated with the impact of PV on their work and daily living is anticipated to be particularly high.

Figure 1: Patient reported impact ratings (where only 'frequently' is shown) stratified by their quartile classification for overall symptom severity



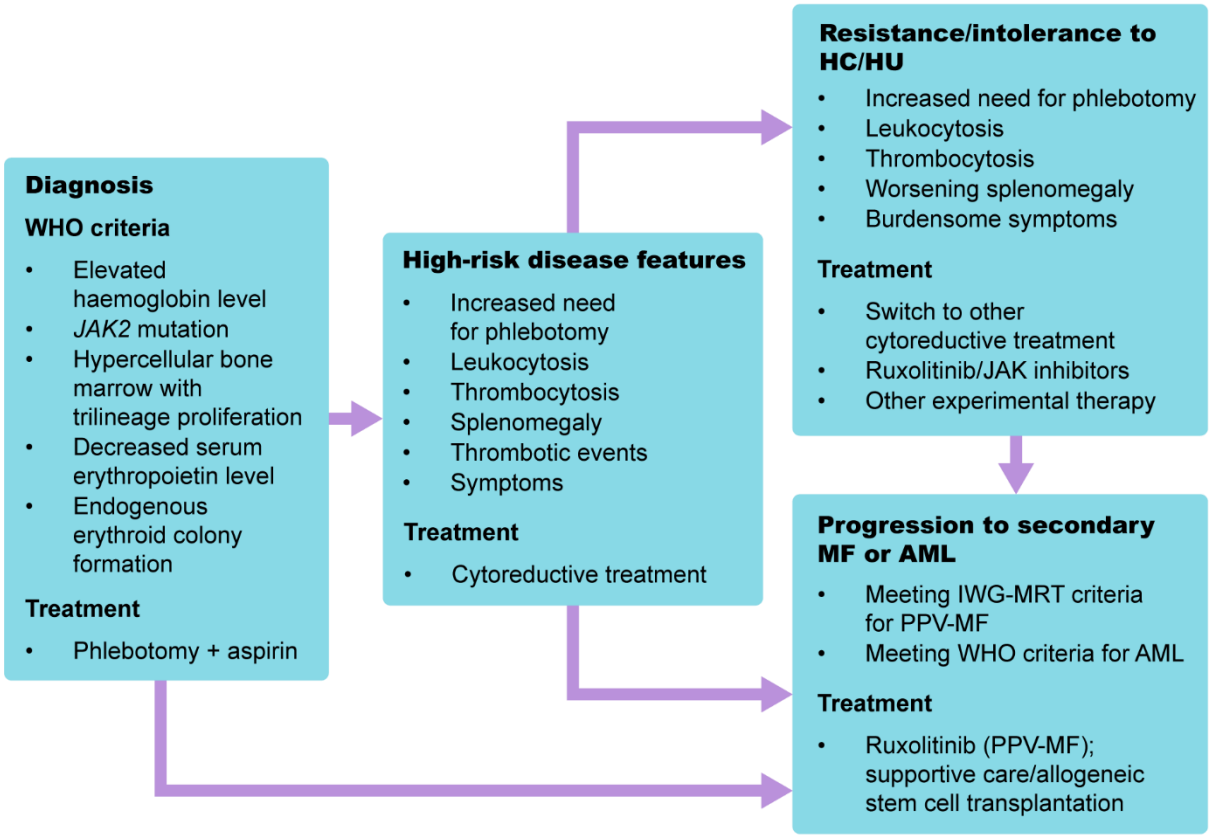
Footnotes: *n = 45 and 61 for lower symptom burden group and higher symptom burden group respectively. Given that the burden of symptoms is particularly high in patients with PV who are R/I to HC/HU, the impact of PV on patients' work productivity and activity impairment may be particularly high in this population.
Source: Harrison et al. 2019[55]

B.1.3.2 Description of the clinical care pathway

The main goals in the treatment of PV are the reduction of thrombosis and haemorrhage risk, minimisation of complications and symptomatology, and the minimisation of risk of transformation to MF and AML/MDS.[1] Disease management is also tailored to individual patients on the basis of their risk stratification (high or low) and/or whether they are R/I to HC/HU (see Section B.1.3.1).[56]

The typical course of progression for PV is summarised in Figure 2.[57]

Figure 2: Progression of PV



Abbreviations: HC/HU: hydroxycarbamide/hydroxyurea; JAK(2): Janus kinase (2); PPV-MF: post-polycythaemia vera myelofibrosis; PV: polycythaemia vera; WHO: World Health Organization.
Source: Reiter and Harrison. 2016.[57]

Current treatment pathway

There are currently no treatments recommended by NICE for patients with PV who are R/I to HC/HU. There are however published guidelines from the BSH from 2018, which are shown in Figure 3 and are described below.[1]

First-line treatment

The BSH guidelines recommend all patients with PV receive phlebotomy and aspirin to maintain HCT at less than 0.45 and reduce the risk of vascular events, with high-risk patients (i.e. those aged ≥65 years and/or with prior PV-associated arterial or venous thrombosis) also administered cytoreductive therapy. Treatment with cytoreductive therapy is also recommended for patients who are not considered to be high-risk but who meet one of several criteria, including uncontrolled HCT or poor tolerability of phlebotomy.

HC/HU or IFN-alfa are recommended as the first-line cytoreductive therapy option. HC/HU is overwhelmingly used as the first-line treatment option for patients with PV in the UK.

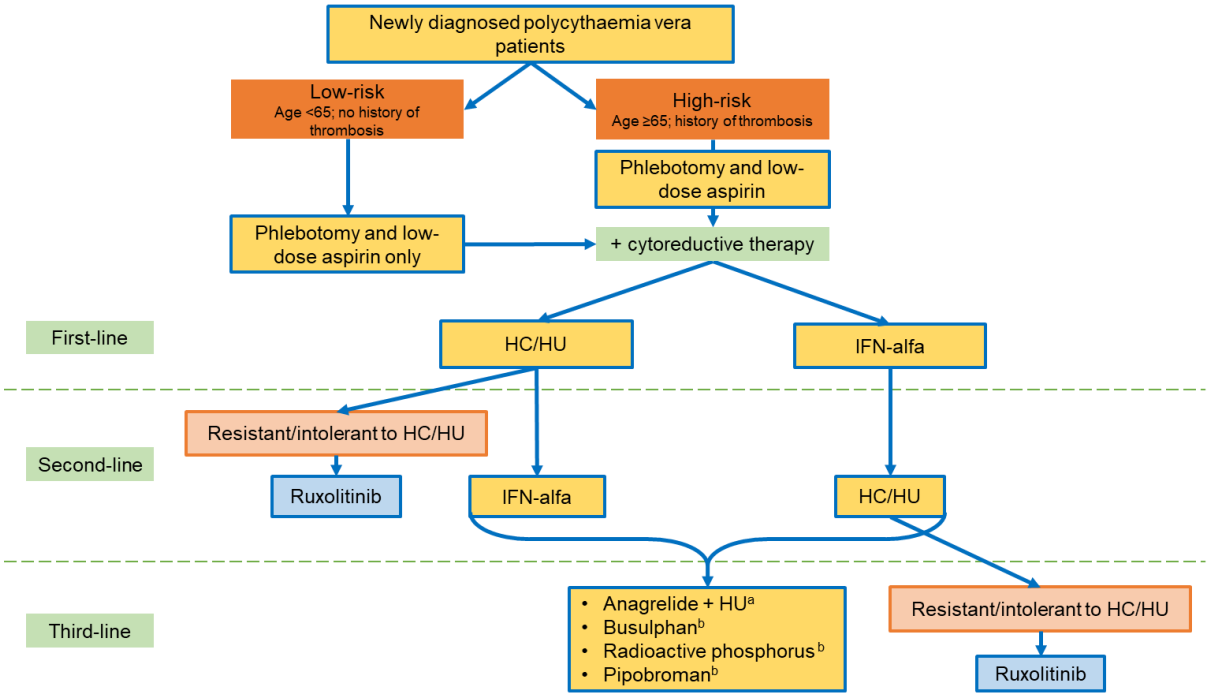
Second-line treatment

For patients treated with HC/HU in the first-line who develop R/I, clinical experts indicated that IFN-alfa would usually be prescribed in the second-line, in line with recommendations from the BSH guidelines.[1, 11] However, a significant proportion of patients do not tolerate or cannot be prescribed IFN-alfa so, despite patients being R/I to HC/HU, a large proportion would also continue to receive HC/HU.[58]

Third-line treatment

The BSH guidelines recommend that busulfan, radioactive phosphorus or pipobroman can be considered in patients with a limited life expectancy in the third-line or later.[1] However, clinical experts have indicated that pipobroman and radioactive phosphorus are no longer used in the UK, meaning busulfan would normally be prescribed in the third-line or later.[11] Although based on limited evidence, anagrelide in combination with HC/HU may also be helpful in patients with platelet control difficulties.[1]

Figure 3: Current treatment pathway for patients with PV who are R/I to HC/HU



[a]May be helpful in those where platelet control is difficult. [b]Only recommended for those with a limited life expectancy.
Abbreviations: HC/HU: hydroxycarbamide/hydroxyurea; IFN: interferon; PV: polycythaemia vera.
Source: McMullin et al. 2019.[1]

Unmet need for treatment of PV in patients with R/I to HC/HU

There is an unmet need for a convenient, alternative, effective treatment that helps to control HCT and improve the high symptom burden of PV, improving HRQoL and reducing the risk of thromboembolic events, progression to MF and AML and death in patients with PV who are R/I to HC/HU

For patients who are R/I to current treatment options, there is an unmet therapeutic need for a therapy that can address the underlying cause of PV and/or its debilitating burden during long-term treatment.[1, 59] This unmet need is anticipated to be greatest in the subgroup of patients who are

defined as high-risk, due to the inherently more severe disease and worse prognosis experienced in this subgroup. Only a small proportion of patients who are R/I to HC/HU will subsequently achieve a response when treated with currently available cytoreductive therapies, as demonstrated in the BAT arms of the RESPONSE and RESPONSE-2 trials, where response rates were 0.9% (HCT control in the absence of phlebotomy eligibility and ≥35% reduction in spleen volume at Week 32) and 19% (HCT control in the absence of phlebotomy eligibility at Week 28), respectively.[58, 60] As a result, patients with PV R/I to HC/HU have a worse life expectancy and increased risk of disease progression compared with patients who are not R/I to HC/HU.[11, 12] Therefore, there is a clear unmet need for effective therapies for patients with PV who are R/I to HC/HU that fully address symptom severity and QoL impairments, while reducing the risk of thromboembolic events, progression to MF and AML and death in this population.

Additionally, currently available cytoreductive therapies for high-risk patients with PV do not offer a tolerable safety profile.[1] For example, IFN-alfa has been associated with high discontinuation rates due to side effects such as fatigue, muscle pain, headache, flu-like symptoms, depression, hair loss and tachycardia, and busulfan has been associated with leukaemic transformation.[1, 61] Moreover, IFN-alfa is a self-administered injection, which may be considered less favourable in terms of convenience compared with an oral formulation.

There is a need to reduce the burden of phlebotomies in patients with PV who are R/I to HC/HU from both a clinical, patient-centric and economic perspective

Long-term phlebotomy can result in noncompliance and intolerance and can cause iron deficiency leading to fatigue and reactive thrombocytosis.[62] Phlebotomies are performed by a healthcare professional and therefore, are both resource intensive and inconvenient for patients. The burden of phlebotomy is anticipated to be particularly high in patients who are high-risk and/or R/I to HC/HU due to the severity of disease and lack of effective treatment options. In these patients, there is a need for a treatment that can achieve HCT control without regular need for phlebotomy, in order to alleviate the burden of these procedures on healthcare resources and patient lives. Following the COVID-19 pandemic, National Health Service (NHS) England waiting times for treatment are the highest they have been since records began.[63] Therefore, the need to reduce the burden on NHS England resources is greater than ever before.

There is a need to reduce the HCRU and direct and indirect costs caused by symptoms, thromboembolic events, disease progression and death in patients with PV who are R/I to HC/HU

A lack of effective treatment options for patients with PV who are R/I to HC/HU means that the care of these individuals imposes a disproportionately high cost and burden on healthcare resources, due to the management of increased rates of myelofibrotic transformation and thromboembolic complications in this population.[54] Furthermore, patients experience debilitating symptoms as a result of PV, particularly within the high-risk population, which can reduce their capacity to work and can result in high activity impairment.[55] There is a high unmet need for an effective treatment that alleviates the burden of thromboembolic events and disease progression in patients with PV who are R/I to HC/HU on healthcare resources. Such a treatment would also help to address the impact of symptoms on patients daily living and work productivity, helping to reduce the indirect cost burden experienced in this population.[55]

Positioning of ruxolitinib relative to the current treatment pathway

Ruxolitinib is indicated for the treatment of adult patients with PV who are R/I of HC/HU.[14] This is in line with the positioning recommended by the BSH as a treatment option in the second- or third-line for patients with PV who are R/I to HC/HU (Figure 3)[1] and the positioning as proposed in this submission. Therefore, BAT consisting of the various cytoreductive therapies available to patients with PV who are R/I to HC/HU is considered to represent the relevant comparator for ruxolitinib in the context of this submission.

Ruxolitinib was recommended by the Scottish Medicines Consortium (SMC) for the treatment of adult patients with PV who are R/I of HC/HU in Scotland in 2019.[64] In England and Wales, ruxolitinib would represent a clinically effective alternative cytoreductive therapy for patients who are R/I to HC/HU and would help to alleviate the substantial burden of symptoms, thromboembolic events, progression to MF and AML and death experienced in this population.

B.1.4 *Equality considerations*

It is not expected that this appraisal will exclude any people protected by equality legislation, nor is it expected to lead to a recommendation that would have a different impact on people protected by equality legislation than on the wider population. Similarly, it is not expected that this appraisal will lead to recommendations that have any adverse impact on people with a particular disability or disabilities.

B.2 Clinical effectiveness

Ruxolitinib offers improved and durable responses, in terms of HCT control and platelet and WBC counts, leading to improvements in symptom burden and HRQoL in patients who are R/I to HC/HU

The efficacy and safety evidence base to support the use of ruxolitinib versus BAT in this indication primarily comes from two randomised controlled trials (RCTs) in adult patients with PV who are R/I to HC/HU: the RESPONSE trial for patients with splenomegaly, and RESPONSE-2 for patients without splenomegaly.[58, 60] Crossover from BAT to ruxolitinib was permitted in both trials (after Week 32 in RESPONSE and Week 28 in RESPONSE-2) and by Week 80 all patients in the BAT arms had discontinued BAT.[65, 66] These results can also be considered applicable to patients in the UK, given the similarity of the BAT administered compared with UK clinical practice and the European centres included (with UK centres specifically for RESPONSE).[58, 60]

In RESPONSE, the primary endpoint (both HCT control and ≥35% reduction in spleen volume) was reached by a significantly higher proportion of patients in the ruxolitinib group versus the BAT group at Week 32 (25/110 [22.7%] versus 1/112 [0.9%]; p<0.001).[67] Complete haematological remission (CHR), which considers platelet and WBC counts as well as HCT control, was achieved by a significantly higher proportion of patients in the ruxolitinib group compared with the BAT group (23.6% versus 8.9%, respectively; p=0.003).[67] In turn, patients treated with ruxolitinib experienced substantial benefits in terms of PV symptoms and HRQoL when compared to patients in the BAT arm.[67]

In RESPONSE-2, the primary endpoint of HCT control at Week 28 was reached by a significantly higher proportion of patients in the ruxolitinib group (42/74 [62%]) compared to the BAT group (14/75 [19%]) (odds ratio [OR] ruxolitinib versus BAT: 7.28; 95% CI: 3.43, 15.45; p<0.0001).[60] The proportion of patients achieving CHR was also significantly higher in the ruxolitinib group compared with the BAT group (23% versus 5%, respectively; OR ruxolitinib versus BAT: 5.58; 95% CI: 1.73, 17.99; p=0.0019). In turn, patients treated with ruxolitinib experienced substantial benefits in terms of PV symptoms and HRQoL when compared to patients in the BAT arm.[60]

Long-term efficacy results for patients receiving ruxolitinib have been reported after a 5-year follow-up (256 weeks) for RESPONSE and 260 weeks for RESPONSE-2. These results provide further evidence to support the primary analysis.[65, 66]

The matching adjusted indirect comparison (MAIC) supports the hypothesis that ruxolitinib is associated with a survival gain, despite limitations to the analysis

Due to the high degree of patient crossover following the primary treatment period in RESPONSE and RESPONSE-2, and the fact that no patients were receiving BAT in either trial from the time of the Week 80 analyses, a MAIC was conducted to estimate the relative efficacy of ruxolitinib versus BAT in terms of overall survival (OS). Using Week 256 data from patients from RESPONSE matched to patients who are R/I to HC/HU and were treated with BAT in the Grupo Español de Enfermedades Mieloproliferativas Filadelfia Negativas (GEMFIN) registry, ruxolitinib was estimated to be associated with a significant improvement in OS versus BAT (post-matching HR: [redacted]; 95% CI: [redacted]).[68] Although there are limitations associated with the MAIC, the observed improvement in OS for ruxolitinib compared to BAT supports the hypothesis that ruxolitinib is associated with a survival advantage.[12, 60, 67]

Ruxolitinib is associated with a tolerable safety profile that is supported by long-term data

Ruxolitinib had a consistent safety profile across both RESPONSE and RESPONSE-2 and adverse events (AEs) were generally manageable with standard clinical monitoring and care.[69] Anaemia was the most common haematological AE experienced by patients treated with ruxolitinib, but was rarely Grade 3 or 4 in severity, and few non-haematological AEs were seen in either treatment group prior to crossover.[60, 67] At the 5-year follow-up in both trials, the exposure-adjusted rate of thromboembolic events was higher in the BAT group compared to the ruxolitinib group.[65, 66]

The efficacy and safety of ruxolitinib is supported by 5-year data from the MAJIC-PV trial in high-risk patients with PV who are R/I to HC/HU

MAJIC-PV (ISRCTN61925716) was an open-label, randomised controlled trial of ruxolitinib versus BAT in high-risk patients with PV who are R/I to HC/HU. Crossover was not permitted throughout this trial as per the protocol, meaning MAJIC-PV provides 5-year supporting efficacy and safety evidence that is not confounded by crossover for ruxolitinib in a subgroup of the licensed indication.[70] In MAJIC-PV, ruxolitinib was associated with improved treatment efficacy versus BAT in terms of CHR (HR: 0.38; 95% CI: 0.24, 0.61, p<0.001) and event-free survival (EFS; HR 0.58; 95% CI: 0.35, 0.94, p=0.03).[70] There was also a trend towards improved OS with ruxolitinib, with the curve starting to diverge after 3.0 years. The patterns of AEs with ruxolitinib were similar to previously reported, with no new events emerging with longer follow-up.[70]

B.2.1 *Identification and selection of relevant studies*

A systematic literature review (SLR) was conducted to identify relevant clinical evidence from RCTs on the efficacy and safety of ruxolitinib or any other pharmacological intervention for the treatment of adult patients with PV who are R/I to HC/HU. Full details of the SLR search strategy, study selection process and results can be found in Appendix D. The SLR included a total of 39 publications, reporting on 8 unique clinical trials.[†]

B.2.2 *List of relevant clinical effectiveness evidence*

RESPONSE and RESPONSE-2

RESPONSE (NCT01243944) and RESPONSE-2 (NCT02038036) provide the main clinical evidence for the efficacy of ruxolitinib for the treatment of patients with PV who are R/I to HC/HU.[60, 67]

RESPONSE is a completed, randomised, open-label, multicentre Phase 3 study which investigated the efficacy and safety of ruxolitinib versus BAT in patients with PV who were R/I to HC/HU.[67] The RESPONSE trial was one of two studies that provided evidence supporting the marketing authorisation application for ruxolitinib for the treatment of PV but only included patients with splenomegaly.[67, 71] Therefore, RESPONSE-2, a second completed randomised, open-label, multicentre Phase 3b trial of ruxolitinib versus BAT, which included patients without palpable splenomegaly, was conducted.[60] Together, these trials were used to support the efficacy and safety of ruxolitinib within the full licensed indication, i.e., the treatment of adult patients with PV who are R/I to HC/HU.[60, 67]

Several published sources of evidence from the RESPONSE[65, 67, 69, 72-82] and RESPONSE-2[60, 66, 83-87] trials were identified in the SLR, including an analysis of the RESPONSE and RESPONSE-2 trials for patients who received IFN-alfa as BAT, and a published analysis based on propensity score matching using data from the RESPONSE trial.[88-90] The data from the RESPONSE trial presented within this submission are largely taken from the following sources:

[†] A Phase II uncontrolled, dose-finding study of ruxolitinib (INCB 18424-256; NCT00726232), in which adult patients with PV or essential thrombocythemia who were R/I to HU were randomised to 10 mg twice daily, 25 mg twice daily, or 50 mg once daily, was also identified in the SLR, but was excluded from the review as it failed to meet the RCT eligibility criterion. The INCB 18424-256 study is also described in the CHMP Assessment Report for the PV indication.

- Week 48 data cut-off (Vannucchi et al. 2015)[58]
- Week 256 published manuscript (Kiladjian et al. 2020)[65]
- Week 256 abstract (Kiladjian et al. 2018)[91]
- Data from Week 80 (Verstovsek et al. 2016)[69] and Week 208 (Kiladjian et al. 2017)[75] have been included in Appendix D for completeness.

The data from the RESPONSE-2 trial presented within this submission are taken from the following sources:

- Week 28 data cut-off (Passamonti et al. 2017)[60]
- Week 260 publication (Passamonti et al. 2022)[66]
- Data from Week 80 (Griesshammer et al. 2018)[84] and Week 156 (Passamonti et al. 2018)[83] have been included in Appendix D for completeness.

Clinical study reports for the RESPONSE and RESPONSE-2 trials have also been used to supplement the information available from the publications. A summary of the RESPONSE and RESPONSE-2 trials is presented in Table 5 below.

Additional sources of RCT evidence identified from the SLR

MAJIC

MAJIC (ISRCTN61925716) is a completed Phase 2 RCT of ruxolitinib versus BAT (without crossover) in patients with ET [MAJIC-ET] and PV [MAJIC-PV] who are R/I to HC/HU.[70] The MAJIC-PV trial population is limited to high-risk patients, who represent a subgroup of the licensed population for ruxolitinib.[14, 70] In the MAJIC trial the definition of high-risk was broad and may be considered to represent the majority of patients with PV who are R/I to HC/HU; patients were considered high-risk if they had significant or symptomatic splenomegaly, previous documented thrombosis (secondary to PV or within 10 years of diagnosis), a platelet count >1000 ×109/L, diabetes or hypertension requiring pharmacological therapy for >6 months and are ≥60 years old.[70] Therefore, the results from MAJIC-PV are presented in this submission to supplement RESPONSE and RESPONSE-2 trial data by providing additional efficacy and safety data for ruxolitinib in a subgroup of patients with PV who are R/I to HC/HU.[70]

Data from MAJIC-PV presented in this submission have been sourced from an unpublished manuscript that has been provided in confidence by the authors for use in this submission and presents Month 60 trial data (Harrison et al.).[70] This publication has been submitted to the New England Journal of Medicine and is undergoing peer review at the time of this submission. This unpublished manuscript is additionally supported by published evidence on MAJIC-PV (Harrison et al. 2018).[92] A summary of the MAJIC-PV trial is additionally presented in Table 5.

RELIEF

RELIEF (NCT01632904) was also identified in the SLR and is a Phase 3b RCT of ruxolitinib versus HC/HU in patients with PV who were well-controlled with a stable dose of HC/HU but reported PV-related symptoms. However, this population is not fully aligned with the licensed indication for ruxolitinib in PV, as patients were not R/I to HC/HU, as per the modified ELN criteria used in the RESPONSE trials. Consequently, this study will not be presented in full as part of this submission.[93, 94]

Table 5: Clinical effectiveness evidence

Study	RESPONSE (NCT01243944)[58]	RESPONSE-2 (NCT02038036)[60]	MAJIC-PV (ISRCTN61925716)[95]
Study design	Randomised, open-label Phase 3 study	Randomised, open-label Phase 3b trial	Randomised, open-label Phase 2 study
Population	Patients aged ≥18 years with PV who are R/I to HC/HU and have splenomegaly	Patients aged ≥18 years with PV who are R/I to HC/HU and have no palpable splenomegaly	Patients aged ≥18 years with high-risk PV who are R/I to HC/HU
Intervention(s)	Ruxolitinib (at a starting dose of 10 mg twice daily)		
Comparator(s)	BAT: BAT was selected at the investigator's discretion and could include HC/HU, IFN-alfa, pipobroman, anagrelide, immunomodulators such as lenalidomide or thalidomide, or no medication		
Indicate if study supports application for marketing authorisation	Yes	Yes	No
Indicate if study used in the economic model	Yes	Yes	Yes
Rationale if study not used in model	*N/A*	*N/A*	*N/A*
Reported outcomes specified in the decision problem	• CHR • OS • Symptom relief • Thrombosis • Rates of transformation to MF or AML • Safety • Health-related quality of life	• CHR • OS • Symptom relief • Thrombosis • Rates of transformation to MF or AML • Safety • Health-related quality of life	• CHR • OS • EFS (composite of major thrombosis, major haemorrhage, transformation or death) • Symptom relief • Thrombosis • Safety • Health-related quality of life
All other reported outcomes	• Change in the frequency of phlebotomy procedures • Change in spleen volume • Change in HCT level over time • Change in WBC and platelet	• Change in phlebotomy eligibility over time • Change in spleen volume • Change in HCT level over time • Change in WBC and platelet counts	• Dose intensity • Histological response: bone marrow biopsy analysis • JAK2 V617F allele burden

Study	RESPONSE (NCT01243944)[58]	RESPONSE-2 (NCT02038036)[60]	MAJIC-PV (ISRCTN61925716)[95]
	counts over time • Transformation-free survival • Rates of MSC and NMSC • JAK2 V617F allele burden	over time • Transformation-free survival	

Abbreviations: AML: acute myeloid leukaemia; BAT: Best Available Therapy; BSH: British Society for Haematology; CHR: complete haematological remission; CTC: common terminology criteria; ECOG: Eastern Cooperative Oncology Group; EFS: event-free survival; ELN: European LeukemiaNet; HCT: haematocrit; JAK: Janus kinase; MF: myelofibrosis; N/A: not applicable; OS: overall survival; PR: partial response; R/I: resistant or intolerant.
Source: Vannuchi et al. 2015;[58] Passamonti et al. 2017;[60] Novartis Data on File (MAJIC Clinical Study Protocol) 2018.[95]

B.2.3 *Summary of methodology of the relevant clinical effectiveness evidence*

B.2.3.1 Trial design

A summary of the trial methodology for RESPONSE and RESPONSE-2 is presented in Table 6.

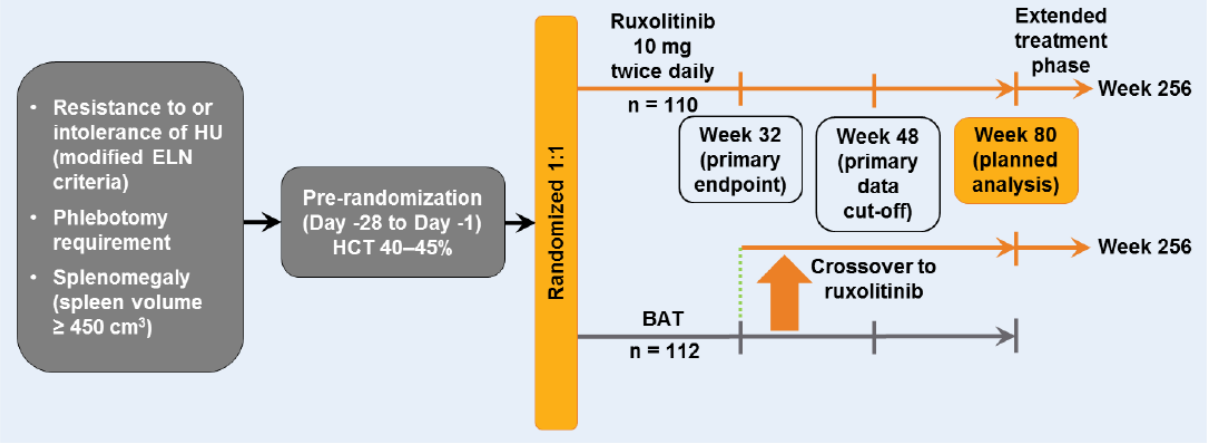
In both trials, ruxolitinib (starting dose of 10 mg, twice daily) was compared to BAT in patients with PV who were R/I to HC/HU (with splenomegaly in RESPONSE and without palpable splenomegaly in RESPONSE-2). Doses could be titrated for safety and efficacy at 5 mg increments, from a minimum of 5 mg once daily and a maximum of 25 mg twice daily.[60, 67] Eligible patients had to meet the definition of HC/HU resistance (an inadequate response to HC/HU treatment) or intolerance (unacceptable side effects from HC/HU treatment) according to modified ELN criteria and were required to have phlebotomy dependency for HCT control.[51, 60, 67] Patients were considered to be phlebotomy dependent if their HCT was 40–45% with two phlebotomies or more spaced at least four weeks apart within 24 weeks before screening, or if their HCT level was higher than 45% with at least one phlebotomy within 16 weeks before screening. Patients who were randomised to BAT could crossover to receive ruxolitinib in both trials if they failed to meet the respective primary endpoints in each trial.[60, 67]

RESPONSE

The study design for RESPONSE is presented in Figure 4. The trial consisted of five phases: screening, pre-randomisation, treatment, extended treatment and survival follow-up.[96]

- **Screening period:** for up to three weeks; Day −49 to Day −29. During this time screening evaluations of spleen volume and HCT were carried out to determine eligibility for progression to study randomisation.[96]
- **Pre-randomisation period:** for up to four weeks; Day −28 to Day −1. Eligible patients with an HCT outside of the values for eligibility could enter an HCT control period. During this time, patients meeting the eligibility criteria for spleen volume and HCT could proceed to randomisation.[96]
- **Treatment period:** Day 1 to Week 80. At screening, patients were stratified by HC/HU intolerance or resistance and randomised 1:1 to ruxolitinib or BAT. At Week 32, patients randomised to BAT could crossover to receive ruxolitinib if they failed to meet the primary endpoint. Patients could also crossover after Week 32 if they did not achieve HCT control or had a spleen volume progression.[96]
- **Extended treatment period:** Week 80 to Week 256. At Week 80, patients receiving ruxolitinib were eligible to continue treatment. Patients receiving BAT were not eligible to continue in the study.[96]
- **Survival follow-up phase:** Until individual Week 256 visit. Applicable for patients who completed or discontinued study treatment prior to Week 256. This follow-up phase continued until the time the Week 256 visit from randomisation would have taken place.[96]

Figure 4: RESPONSE study design



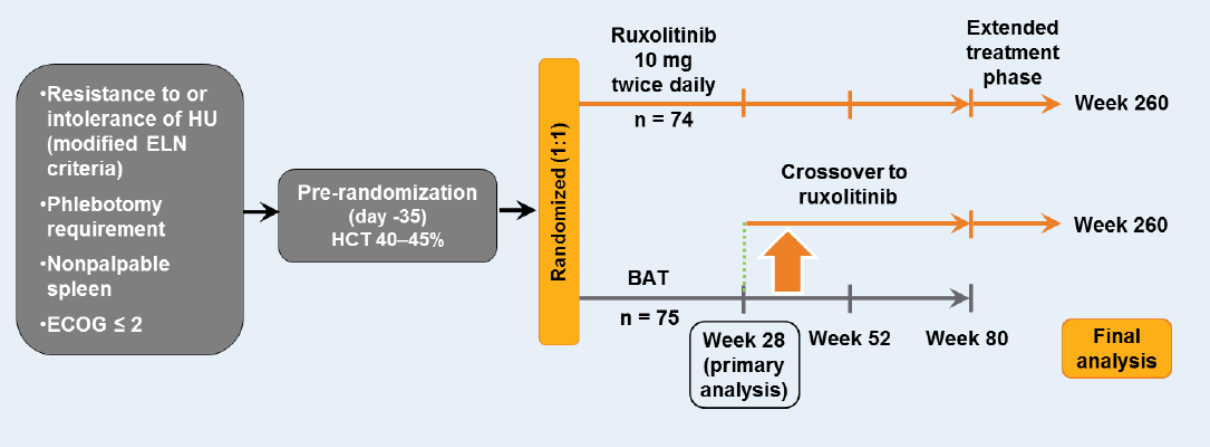
Abbreviations: BAT: Best Available Therapy; ELN; European LeukemiaNet; HCT: haematocrit; HC/HU: hydroxycarbamide/hydroxyurea.
Source: Novartis Data on File (RESPONSE Week 256 CSR) 2018.[97]

RESPONSE-2

The study design for RESPONSE-2 is presented in Figure 5 below. The trial consisted of four phases: screening, core treatment, extended treatment and follow-up.[98]

- **Screening period:** for up to five weeks; Day −35 to Day −1. Screening evaluations were reviewed to determine patient eligibility prior to randomisation. Patients with HCT >45% could enter an HCT control period. Patients with HCT 40–45% within 14 days of the randomisation visit could proceed to randomisation.[98]
- **Core treatment period:** Day 1 to Week 80. Patients were randomised 1:1 to ruxolitinib or BAT.
- **Crossover treatment period**: Week 28 or after for BAT patients only. At Week 28, patients randomised to BAT who did not respond to treatment were eligible to crossover to receive ruxolitinib. Patients crossing over on or after Week 28 had to complete all assessments for the end of treatment visit of the core treatment period followed by the assessments in the crossover treatment period.[98]
- **Extended treatment period:** Week 80 to Week 260. Patients receiving ruxolitinib at Week 80 (including those who crossed over from the BAT group) were eligible to continue treatment until Week 260, continuing the dose that they received at Week 80. Patients receiving BAT at Week 80 were not eligible to enter the extended treatment period.[98]
- **Safety follow-up period:** Patients were followed for safety for 30 days after the last dose of study drug. End of study assessments were carried out post 30 days after the last dose of study drug.[98]
- **Survival and antineoplastic therapies follow-up period:** Until individual Week 260 visit. Patients were followed up for survival and antineoplastic therapies every three months following completion of study treatment or from the time of premature discontinuation until the end of the study.[99]

Figure 5: RESPONSE-2 study design



Abbreviations: BAT: Best Available Therapy; ECOG: Eastern Cooperative Oncology Group; ELN; European LeukemiaNet; HCT: haematocrit; HC/HU: hydroxycarbamide/hydroxyurea.
Source: Novartis Data on File (RESPONSE-2 Week 260 CSR) 2020.[98]

Table 6: Summary of RESPONSE and RESPONSE-2 trial methodology

Trial name	RESPONSE (NCT01243944)	RESPONSE-2 (NCT02038036)
Location	International, multicentre trial with 92 sites across 18 countries: Argentina, Australia, Belgium, Canada, China, France, Germany, Hungary, Italy, Japan, Korea, Netherlands, Russia, Spain, Thailand, Turkey, UK and US.	International, multicentre trial with 48 sites across 12 countries: Australia, Belgium, Canada, France, Germany, Hungary, India, Israel, Italy, Korea, Spain and Turkey.
Trial design	Randomised, open-label Phase 3 study Patients with an HCT of <40% or >45% entered an HCT control period before randomisation. Patients with an HCT of 40–45% within 14 days before day 1 of the study could proceed directly to randomisation. Patients assigned to BAT could crossover to ruxolitinib at Week 32 if the primary endpoint was not met, or later in the case of disease progression (phlebotomy eligibility, progression of splenomegaly, or both).	Randomised, open-label Phase 3b study Patients with an HCT >45% entered an HCT control period before randomisation, to ensure that their HCT was similar and controlled at study initiation. An HCT of 40–45% achieved with phlebotomy was required within 14 days before randomisation. Patients randomised to BAT could crossover to ruxolitinib from Week 28 if they did not meet the primary endpoint, or later if treatment was shown to be ineffective (i.e. HCT >45% or if they received phlebotomy) or for safety-related reasons.
Method of randomisation	Patients were randomised in a 1:1 ratio to receive ruxolitinib or BAT. Randomisation was stratified by status with regards to HC/HU therapy (inadequate response versus unacceptable side effects).	Patients were randomised in a 1:1 ratio to receive ruxolitinib or BAT. Randomisation was stratified by status with regards to HC/HU therapy (inadequate response versus unacceptable side effects).
Eligibility criteria for participants	**The full eligibility criteria for the RESPONSE trial are provided in Appendix D.** **Key eligibility criteria:** • Adults (≥18 years of age) with PV requiring phlebotomy for HCT control[a] • R/I to HC/HU according to modified ELN criteriab • Spleen volume of 450 cm3 or more (as measured by MRI or CT) • No prior treatment with a JAK inhibitor	**The full eligibility criteria for the RESPONSE-2 trial are provided in Appendix D.** **Key eligibility criteria:** • Adults (≥18 years of age) with PV requiring phlebotomy for HCT control[a] • PV diagnosis according to 2008 WHO criteria • R/I to HC/HU (according to modified ELN criteria)b • No palpable splenomegaly • ECOG status of 0,1 or 2 • No prior treatment with a JAK inhibitor • Lack of pregnancy/nursing in female patients • Adequate gastrointestinal, liver or renal function • Platelet count ≥100×10⁹ platelets per L or an absolute

		neutrophil count of ≥1×10⁹ cells per L • Compliance to the protocol • No active malignancy during the previous 5 years • No clinically significant cardiac disease • No history of progressive multifocal leukoencephalopathy
Trial drugs and method of administration	**Treatment groups:** **1. Ruxolitinib (n=110)** **2. BAT (n=112)** • Study medication consisted of ruxolitinib or single-agent BAT as judged by the treating physician • Standard therapy could include HC/HU, IFN-alfa, pipobroman, anagrelide, immunomodulators such as lenalidomide or thalidomide, or no medication. radioactive phosphorus, busulfan and chlorambucil were prohibited • Standard therapy could be changed owing to lack of response or toxic effects requiring drug discontinuation **Starting dose of ruxolitinib: 10 mg twice daily** **Dose adjustments for ruxolitinib: Doses could be titrated for safety and efficacy (5 mg increments; minimum of 5 mg once daily and maximum of 25 mg twice daily)[c]**	**Treatment groups:** 1. Ruxolitinib (n=74) 2. BAT (n=75) • Study medication consisted of ruxolitinib or single-agent BAT as judged by the treating physician • Standard therapy could include HC/HU (at maximum tolerated dose), IFN-alfa, pipobroman, anagrelide, approved immunomodulators such as lenalidomide and thalidomide, or no cytoreductive treatment • Standard therapy could be changed if the patient had an insufficient response to treatment or if therapy-related toxic effects occurred that necessitated drug discontinuation **Starting dose of ruxolitinib: 10 mg twice daily** **Dose adjustments for ruxolitinib: Doses could be titrated for safety and efficacy (5 mg increments; minimum of 5 mg once daily and maximum of 25 mg twice daily)[d]**
Permitted and disallowed concomitant medication	All patients received low-dose aspirin unless it was medically contraindicated	All patients received low-dose aspirin unless it was medically contraindicated
Primary outcome	The proportion of patients achieving a primary response at Week 32 • Primary response was defined as both HCT control and a reduction of ≥35% in spleen volume from baseline • Spleen volume as assessed by means of centrally reviewed MRI or CT studies • HCT control was defined as protocol-specified ineligibility for phlebotomy from Week 8 to 32 and no more than one	The proportion of patients achieving HCT control at Week 28 • HCT control was defined as the absence of phlebotomy eligibility between Week 8 and Week 28, with phlebotomy eligibility occurring only once after randomisation and before Week 8 • Phlebotomy eligibility was defined as an HCT of >45% that was at least three percentage points higher than baseline, or an HCT of >48%, whichever was lower

	instance of phlebotomy eligibility between randomisation and Week 8 • Phlebotomy eligibility was defined as an HCT of >45% that was at least three percentage points higher than baseline, or an HCT of >48%, whichever was lower	
Key secondary outcomes ***Outcomes not presented within this submission**	• Duration of primary response, defined as the time from the first occurrence when both components of the primary endpoint are met until the date of the first documented disease progression • Proportion of patients achieving CHR, defined as HCT control (defined as per the primary outcome) with a platelet count ≤400×109/L and a WBC count ≤10×109/L • Duration of CHR • Change in the frequency of phlebotomy procedures • Change in spleen volume • Change in HCT level over time* • Transformation-free survival • OS • Safety • PROs	• Proportion of patients achieving CHR, defined as HCT control (defined as per the primary outcome) with a platelet count ≤400×109/L and a WBC count ≤10×109/L • Change in phlebotomy eligibility over time • Change in HCT level over time • Change in spleen length* • Change in ECOG status* • Transformation-free survival • OS • Safety • PROs
Duration of study and follow-up	The study was initiated on 27th October 2010 and completed on 9th February 2018. Data presented within this submission are from the 9th February 2018 data cut-off which represents when all patients had completed the Week 256 visit or discontinued as per-specified in the protocol.	The study was initiated on 25th March 2014 and completed on 7th April 2020. Data presented within this submission are from the 2nd July 2020 data-cut off which represents when all patients had completed the Week 260 visit or discontinued as per-specified in the protocol.

[a] Patients were judged to be phlebotomy dependent if their HCT was 40–45% with two phlebotomies or more spaced at least 4 weeks apart within 24 weeks before screening, or if their HCT level was higher than 45% with at least one phlebotomy within 16 weeks before screening. Before randomisation, eligible patients with HCT greater than 45% entered a HCT control period to ensure that their HCT was similar and controlled at study initiation, preventing any potential bias; a HCT between 40–45% achieved with phlebotomy within 14 days before randomisation was required. [b] An inadequate response to HC/HU is defined as a dose ≥2 g/day or a maximum tolerated dose <2 g/day resulting in at least one of the following: Need for phlebotomy to maintain HCT <45%; platelet count >400 × 10^9/L and WBC count >10 × 10^9/L; failure to reduce splenomegaly extending >10 cm below the costal margin by >50%, as measured by palpation. Unacceptable side effects from HC/HU were defined as at least one of the following: absolute neutrophil count <1.0 × 10^9/L; platelet count <100 × 10^9/L or haemoglobin <100 g/L (i.e. 10 g/dL) at the lowest dose of HC/HU required to achieve a response; presence of leg ulcers or other unacceptable HC/HU-related non-haematologic toxicities (such as mucocutaneous manifestations, gastrointestinal symptoms, pneumonitis, or fever at any dose of HC/HU), defined as Common Terminology Criteria for Adverse Events (CTCAE) version 3.0 grade 3–4 or >1 week of CTCAE version 3.0 grade 2, permanent discontinuation of HC/HU, interruption of HC/HU until toxicity resolved, or hospitalisation due to HC/HU toxicity. [c] For RESPONSE, dose increases could occur for patients who met all of the following conditions: inadequate efficacy as demonstrated by at least one of the following HCT ≥45% or <45% but ≥3 percentage points higher than baseline; WBC >ULN; platelet count

>ULN; palpable spleen that is reduced by <25% from baseline at Week 4 or <50% at subsequent study visits; platelet count ≥140×10^9/L; haemoglobin ≥12 g/dL; absolute neutrophil count ≥1.5×10^9/L. Dose reductions or interruptions were required for: specified cytopaenias of ≥ Grade 2; haemoglobin <10.0 g/dL; platelet count <75×10^9/L; Grade 1 anaemia (haemoglobin level < LLN to 10.0 g/dL); platelet count <100×10^9/L if, in the investigator's judgement, the reduction was warranted given the rapidity and magnitude of the haematologic change. [d] For RESPONSE-2, dose increases could occur for inadequate efficacy, that is: HCT increase of ≥3 percentage points from baseline; WBC count >ULN; palpable spleen. Dose reductions were required for: haemoglobin <100 g/L; platelet count <75×10^9/L; dose interruptions were required for: haemoglobin <80 g/L; platelet count <50×10^9/L; absolute neutrophil count <1×10^9/L.

Abbreviations: AML: acute myeloid leukaemia; BAT: Best Available Therapy: CHR: complete haematological remission; CT: computed tomography; CYP3A4: cytochrome P450 3A4; ECOG: Eastern Cooperative Oncology Group; ELN: European LeukemiaNet; HCT: haematocrit; HIV: human immunodeficiency virus; HC/HU: hydroxycarbamide/hydroxyurea; IFN: interferon; LLN: lower limit of normal; MF: myelofibrosis; MRI: magnetic resonance imaging; OS: overall survival; PRO: patient-reported outcome; PV: polycythaemia vera; R/I: resistance to or intolerance to; ULN: upper limit of normal; WBC: white blood cell; WHO: World Health Organization.

Source: Vannucchi et al. 2015;[67] ClinicalTrials.gov (NCT01243944);[73] Novartis Data on File (RESPONSE Week 208 CSR) 2017;[96] Passamonti et al. 2017;[60, 100] Passamonti et al. 2018;[83] Passamonti et al. 2022;[66] ClinicalTrials.gov (NCT02038036).[89]

B.2.3.2 Eligibility criteria

An overview of the key eligibility criteria for RESPONSE and RESPONSE-2 is provided in Table 6. The full eligibility criteria for RESPONSE and RESPONSE-2 are provided in Appendix D.

Overall, with the exception of the presence of splenomegaly, the eligibility criteria, and therefore, the patient populations, were similar between RESPONSE and RESPONSE-2.

B.2.3.3 Baseline characteristics

The baseline characteristics for patients evaluated in the RESPONSE and RESPONSE-2 trials are summarised in Table 7 below. Overall, UK clinicians agreed that the populations included in the RESPONSE trials were generally reflective of UK practice. Clinical experts noted that patients included in the RESPONSE trials are likely to have better performance status than patients in UK clinical practice, and the trials included a mix of high-risk and low-risk patients. The proportion of high-risk to low-risk patients was considered by the experts to be broadly reflective of clinical practice. Clinical experts also noted that patients in the RESPONSE trials were required to have at least two phlebotomies in the last 24 weeks, which is not fully representative of the UK population of patients with R/I to HC/HU who may not require phlebotomy.[11]

In RESPONSE, no significant difference between the two treatment groups were observed with respect to baseline characteristics and disease history.[67] In RESPONSE-2, baseline characteristics were generally similar between treatment groups, although differences in median age and sex between the groups were observed.[60]

Table 7: Baseline characteristics for RESPONSE and RESPONSE-2

Characteristic	RESPONSE		RESPONSE-2	
	Ruxolitinib (n=110)	BAT (n=112)	Ruxolitinib (n=74)	BAT (n=75)
Age – years				
Median (range)	62.0 (34–90)	60.0 (33–84)	63 (NR)	67 (NR)
IQR	NR	NR	54–71	61–74
>60 years – n (%)	NR	NR	46 (62)	57 (76)
Sex – n (%)				
Male	66 (60.0)	80 (71.4)	39 (53)	47 (63)
Female	44 (40.0)	32 (28.6)	35 (47)	28 (37)
Time since diagnosis – years				
Median (range)	8.2 (0.5–36)	9.3 (0.5–23)	6.5 (2.9–10.7)	6.7 (3.2–10.6)
Previous lines of antineoplastic therapy				
1	NR	NR	53 (72%)	52 (69%)
>1	NR	NR	21 (28%)	23 (31%)
Duration of prior HC/HU therapy – years				
Median (range)	3.1 (<0.1–20.9)	2.8 (<0.1–20.9)	2.83 (0.57–6.61)[a]	3.55 (0.57–7.03)[a]
ECOG performance status – n (%)[b]				
0	76 (69.1)	77 (68.8)	NR	NR

Characteristic	RESPONSE		RESPONSE-2	
	Ruxolitinib (n=110)	BAT (n=112)	Ruxolitinib (n=74)	BAT (n=75)
1	31 (28.2)	34 (30.4)	NR	NR
2	3 (2.7)	1 (0.9)	NR	NR
Previous HC/HU treatment status – n (%)				
Unacceptable side effects	59 (53.6)	61 (54.5)	44 (59)	45 (60)
Inadequate response	51 (46.4)	51 (45.5)	30 (41)	30 (40)
Prior thromboembolic event				
n (%)	39 (35.5)	33 (29.5)	21 (28)	18 (24)
Presence of *JAK2* V617F mutation				
n (%)	104 (94.5)	107 (95.5)	72 (97)[c]	69 (92)
Allele burden – % ± SD	76.2 ± 17.8	75.0 ± 22.6	NR	NR
Spleen length				
Below costal margin – cm				
Median (range)	7.0 (0–24.0)	7.0 (0–25.0)	NR	NR
<10 cm – n (%)	71 (64.5)	67 (59.8)	NR	NR
>20 cm – n (%)	2 (1.8)	4 (3.6)	NR	NR
Spleen volume – cm^3				
Median (range)	1195 (396–4631)	1322 (254–5147)	NR	NR
Percentage HCT level – %[d]				
Mean ± SD	43.6 ± 2.2	43.9 ± 2.2	42.8 ± 1.46	42.7 ± 1.44
Median (range or IQR)	43.3 (range: 39.2–50.5)	44.0 (range: 37.6–50.5)	43.0 (IQR: 41.7–44.0)	42.7 (IQR: 41.7–44.0)
HCT category – n (%)				
40–45%	79 (71.8)	83 (74.1)	NR	NR
>45%	28 (25.5)	25 (22.3)	NR	NR
WBC count				
× 10^{-9}/L ± SD	17.6 ± 9.6	19.0 ± 12.2	12.0 ± 8.19	13.0 ± 8.06
Platelet count				
× 10^{-9}/L ± SD	484.5 ± 323.3	499.4 ± 318.6	469.5 ± 295.96	471.5 ± 350.38
Phlebotomies within 24 weeks before screening				
≥2 – n (%)	NR	NR	58 (78)	57 (76)
Median (range)	2.0 (1–8)	2.0 (0–16)	NR	NR

[a] Manually converted duration in months from the source to duration in years for consistency. [b] ECOG performance status ranges from 0 to 5, with 0 indicating no symptoms and higher numbers indicating increasing disability. [c] For five patients (ruxolitinib, n=2; BAT, n=3) the *JAK2* V617F mutation was not confirmed by central laboratory assessment. These patients were not included as *JAK2* V617F mutation positive. [d] Value at the end of the HCT control period before randomisation. Patients who had an HCT of 40–45% within 14 days before their day 1 visit could proceed to randomisation; however, the HCT at baseline may have been higher or lower.
Abbreviations: BAT: Best Available Therapy; ECOG: Eastern Cooperative Oncology Group; HCT: haematocrit; HC/HU: hydroxycarbamide/hydroxyurea; IQR: interquartile range; NR: not reported; SD: standard deviation; WBC: white blood cell.
Source: Vannucchi et al. 2015;[67] Passamonti et al. 2017.[60]

B.2.3.4 Patient disposition

Full CONSORT diagrams of the population flow for the RESPONSE and RESPONSE-2 and trials can be found in Appendix D.

RESPONSE

In the REPONSE trial (final data cut-off 9th February 2018), of the 222 enrolled patients, 110 were randomised to ruxolitinib and 112 to BAT.[65]

In the ruxolitinib group, 100% of patients completed the randomisation treatment phase. The primary reasons for discontinuation of treatment before Week 256 or end of extended treatment phase were treatment completed per protocol (72/110; 65.5%), AEs (16/110; 14.5%) and disease progression (12/110; 10.9%).[65]

In the BAT group, at the Week 256 analysis, 100% of patients had either completed the study at Week 80, crossed over to the ruxolitinib group or discontinued.[65] Of the 112 patients initially randomised to BAT, 98 patients crossed over to ruxolitinib, with most crossovers occurring at or immediately after the Week 32 visit.[65, 67] The primary reasons for discontinuation following crossover were treatment completed per protocol (64/98; 65.3%) and AEs (16/98; 16.3%).[65] [redacted] of patients were followed-up for survival.[97]

RESPONSE-2

In the REPONSE-2 trial (final data cut-off 2nd July 2020), of the 149 enrolled patients, 74 were randomised to ruxolitinib and 75 to BAT.[66]

At the Week 260 analysis, 79.7% of patients (59/74) in the ruxolitinib group were still on treatment.[66] Primary reasons for discontinuation of treatment before Week 260 in this group included AEs (7/74; 9.5%), withdrawal of consent (3/74; 4.1%), disease progression (2/74; 2.7%), physician decision (2/74; 2.7%) and death (1/74; 1.4%).[66] Of those randomised to BAT, 58 of 75 patients had crossed over to ruxolitinib (77.3%) and 38 of these were still on treatment at Week 260 (50.7%). Reasons for discontinuation in the crossover group included AEs (9/58; 15.5%), withdrawal of consent (3/58; 5.2%), disease progression (3/58; 5.2%), physician's decision (2/58; 3.4%), deaths (2/58; 3.4%) and lost to follow-up (1/58; 1.7%).[66]

B.2.4 *Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence*

Trial populations

Definitions of the key study populations analysed, and the patient numbers included in each analysis set in the RESPONSE and RESPONSE-2 clinical trials, are presented in Table 8.

The efficacy analysis for the primary and secondary endpoints of RESPONSE and RESPONSE-2 were both performed according to the intention-to-treat (ITT) principle, using data from all patients who underwent randomisation (RESPONSE: N=222; RESPONSE-2: N=149).[67]

Table 8: Trial populations

Analysis set	Definition		Number of patients	
	RESPONSE	RESPONSE-2	RESPONSE	RESPONSE-2
Screened set	Patients assessed for eligibility		342	173
Enrolled set	Patients deemed to meet the eligibility criteria		222	149
Full analysis set[a]	All patients to whom study treatment was assigned by randomisation		222	149
Safety set	All patients who received at least one dose of a study drug, including those who received no drug as standard therapy, if they underwent any post-randomisation procedures or assessments	All patients who received at least one dose of study medication, including those who received no cytoreductive treatment	221	149

[a] For both RESPONSE and RESPONSE-2, the efficacy analysis was conducted according to the ITT principle, including data from all patients randomly assigned to treatment.
Abbreviations: ITT: intention-to-treat.
Source: Vannucchi et al. 2015;[67] Novartis Data on File (RESPONSE Week 208 CSR) 2017;[96] Passamonti et al. 2017.[60]

Statistical analyses

Details for the statistical analysis for the primary analyses in the RESPONSE and RESPONSE-2 trials are presented in Table 6.

Table 9: Statistical methods in RESPONSE and RESPONSE-2 for the primary analysis

Trial name	RESPONSE	RESPONSE-2
Hypothesis objective	The statistical null hypotheses for the primary analysis were: H_0: $\pi_{RUX} = \pi_{BAT}$ versus H_1: $\pi_{RUX} \neq \pi_{BAT}$ where π_{RUX} and π_{BAT} are the responder rates at Week 32 in the ruxolitinib and BAT group, respectively.	For the primary endpoint, the null hypothesis is an OR of 1, with an alternative hypothesis that the OR≠1, where OR is the ratio of odds of proportion of patients with HCT control at Week 28, in the ruxolitinib arm to that in the BAT arm.
Statistical analysis	Responder rates were presented by treatment group along with 95% CIs using Clopper Pearson exact method. The Cochran-Mantel-Haenszel test stratified by the HC/HU status (HC/HU resistant versus HC/HU intolerant) was applied to compare the two treatment arms. The test was two-sided at the 5% significance level. The overall stratum-adjusted OR was used as a measure of association between treatment and response. The OR was presented with 95% Wald confidence limits. However, if the proportion in any group was less than 4% then the stratified exact Cochran-Mantel-Haenszel test was used. In addition, the adjusted proportion difference and its 95% CI were calculated using Cochran-Mantel-Haenszel weight and Wald-type CI or any other appropriate method.	The primary analysis was conducted after all patients completed Week 28 or prematurely discontinued study treatment. In addition, analyses were performed after all patients completed Week 80 (or discontinued prior to Week 80). The final analysis was conducted after all patients in the study completed Week 260 or prematurely discontinued study treatment and had their 30 day follow up visit. A two-sided stratified Cochran-Mantel-Haenszel test was conducted at the 5% level of significance. Rate of HCT control at Week 28 is presented by treatment group with 95% CIs (Clopper-Pearson). The OR is presented with 95% Wald confidence limits.
Sample size, power calculation	Assuming an HCT control rate of 10% in the BAT arm and 30% in the ruxolitinib arm, a sample size of 200 patients was deemed to be required to detect a significant difference with a two-sided test (0.05 significance level and 94% power). HC/HU stratum specific rates for each treatment arm were obtained assuming that the ratio of HC/HU resistance to HC/HU intolerance is 2:1 and that response rates were 20% higher for patients who were intolerant to HC/HU relative to those who were resistant to HC/HU, i.e. the response rate in HC/HU intolerant patients = 1.2 times the response rate in HC/HU resistant patients. Power for the Cochran-Mantel-Haenszel test, stratifying on HC/HU use, was calculated using a large sample normal approximation.	Assuming an HCT control rate of 20% in the BAT arm and 50% in the ruxolitinib arm, a sample size of 116 patients was deemed to be required to detect a significant difference with a two-sided test (0.05 significance level and 90% power). Allowing for an estimated attrition rate of 10%, a total of 130 patients (65 per arm) were anticipated to be required at randomisation.

Data management, patient withdrawals	Patients with missing assessments that prevented the evaluation of the primary endpoint were considered non-responders. A missing HCT assessment that was required to confirm HCT control (absence of phlebotomy eligibility) was presumed to meet phlebotomy eligibility criteria. To be considered as having HCT control (absence of phlebotomy eligibility) between Weeks 8 and 32, patients could have no more than one missing HCT assessment at any of the scheduled visits between Weeks 8 and 32, inclusive. Patients with more than one HCT missing assessment at these scheduled visits between Weeks 8 and 32 were considered to be primary endpoint non-responders. Patients who discontinued the study prior to the completion of the Week 32 visit assessments were considered non-responders.	Patients with missing assessments that prevent evaluation of the primary endpoint were considered non-responders. Patients had to have non-missing HCT data at both Week 8 and Week 28, and could not have >1 missing HCT value among all scheduled visits between these timepoints. Patients who discontinued the trial prior to Week 28 were considered non-responders.

Abbreviations: BAT: Best Available Therapy; CI: confidence interval; HCT: haematocrit; HC/HU: hydroxycarbamide/hydroxyurea; OR: odds ratio.
Source: Novartis Data on File (RESPONSE Week 48 CSR) 2014;[101] Novartis Data on File (RESPONSE-2 Clinical Study Protocol) 2016.[99]

B.2.5 *Critical appraisal of the relevant clinical effectiveness evidence*

A summary of the quality assessment for the primary publications of the RESPONSE (Vannucchi et al. 2015) and RESPONSE-2 (Passamonti et al. 2017) clinical trials are presented in Table 10. The full assessment can be found in Appendix D.[60, 67]

Table 10: Overview of quality assessments of RESPONSE and RESPONSE-2

	Risk of bias	
	RESPONSE	**RESPONSE-2**
Was randomisation carried out appropriately?	Unclear: NR in primary publication	Low risk of bias
Was the concealment of treatment allocation adequate?	Unclear: NR in primary publication	Low risk of bias
Were the groups similar at the outset of the study in terms of prognostic factors?	Low risk of bias	Low risk of bias
Were the care providers, participants and outcome assessors blind to treatment allocation?	High risk of bias: open-label study design	High risk of bias: open-label study design
Were there any unexpected imbalances in drop-outs between groups?	High risk of bias: crossover of treatment arms	High risk of bias: crossover of treatment arms
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Low risk of bias	High risk of bias: outcomes missing in primary publication
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Low risk of bias	Low risk of bias

Abbreviations: NR: not reported.
Source: Adapted from Systematic Reviews: Centre for Reviews and Dissemination's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination).

B.2.6 *Treatment duration and exposure*

B.2.6.1 RESPONSE

A total of 110 patients were randomised to and received at least one dose of ruxolitinib. Of the 112 patients randomised to BAT, 111 received treatment due to one patient withdrawing consent prior to treatment.[91] At or after Week 32, 98 (87.5%) patients randomised to the BAT group crossed over to the ruxolitinib group.[91, 97]

Of those patients in the ruxolitinib group who had an assessment at Week 32 (n=98), 33 patients were receiving a 10 mg twice daily.[67] A total of ten patients were receiving less than 10 mg twice daily, 32 patients were receiving 15 mg twice daily, 15 patients were receiving 20 mg twice daily, and eight patients were receiving 25 mg twice daily.[67] The mean total dose increased over time, with most dose adjustments occurring within the first eight weeks of treatment.[67]

At Week 256 in the ruxolitinib group, median exposure was 255 weeks (interquartile range [IQR]: 158–256).[65] The mean and median dose intensity in the ruxolitinib group was [redacted] mg (standard

deviation [SD]: [redacted]) and 22.5 mg (IQR: 18.7–28.7), respectively.[65, 97] In the BAT group, median exposure was 34 weeks as [redacted] of patients ended their treatment between Week 32 and Week 47.[65, 97] In the crossover population, median duration of exposure to ruxolitinib was 220 weeks (IQR: 135 – 223) at Week 256.[65]

Duration of exposure calculated in patient-years in RESPONSE were reported as 428.4 for the ruxolitinib group, 73.6 for the BAT group and 329.9 for the crossover group.[65]

The duration of treatment with ruxolitinib (Week 256 analysis), in the ruxolitinib group, is presented in Figure 6.

Figure 6: Kaplan–Meier plot for time on treatment in RESPONSE (ruxolitinib group) – Week 256

Source: Novartis Data on file.

B.2.6.2 RESPONSE-2

A total of 74 and 75 patients were randomised to receive ruxolitinib or BAT in the RESPONSE-2 trial.[60] In the ruxolitinib group, 21 (28%) patients had a dose reduction, six (8%) patients had an interruption, and two (3%) patients discontinued treatment at the time of the data cut-off for the primary analysis.[60] Of those patients assigned to the BAT group, 51 (68%) patients crossed over to ruxolitinib at or after Week 28; 45 of these patients completed BAT treatment (28 weeks) before crossing over.[60] By Week 80, a total of 58 (77%) of 75 patients in the BAT group had crossed over to ruxolitinib; no patients continued BAT after week 80 per protocol. 97 patients received ruxolitinib

until Week 260, including 59 (80%) of 74 patients in the ruxolitinib group and 38 (66%) of 58 patients in the crossover groups (see Section B.2.3.4).[66]

At the final data cut-off (Week 260), median exposure was 260 weeks for ruxolitinib (range: [redacted]), 28 weeks for BAT (range: 7–83), and, in crossover patients, 225 weeks for ruxolitinib (range: 3–236).[66] The median dose exposure to ruxolitinib was 20 mg per day in both the ruxolitinib group (IQR range: 17–28) and in the crossover group (IQR range: 19–30).[66] The duration of treatment with ruxolitinib (Week 260 analysis), in the ruxolitinib group, is presented in Figure 7.

Duration of exposure in patient-years for participants of the RESPONSE-2 trials were calculated as 334.27 days in the ruxolitinib group, 53.35 days in the BAT group and 205.96 days in the crossover group.[66]

Figure 7: Kaplan–Meier plot for time on treatment in RESPONSE-2 (ruxolitinib group) – Week 260

Source: Novartis Data on File.

B.2.7 *Clinical effectiveness results*

B.2.7.1 RESPONSE

Data from the RESPONSE trial are primarily presented from the Week 32 (primary) analysis of efficacy outcomes, as reported in Vannucchi et al. 2015, as well as from the final Week 256 analysis.[65] Data from the Week 256 analysis have been used to inform the economic evaluation in Section B.3. Summaries of the efficacy outcomes at Week 80 and Week 208 have been provided

in Appendix D for completeness. All outcomes are presented for the efficacy analysis set (ITT population of all patients who were randomised to a treatment group; n=222).[65, 67]

Primary efficacy outcome: Week 32

In RESPONSE, the composite endpoint was comprised of both HCT control and ≥35% reduction in spleen volume.[67]

- HCT control was defined as protocol-specified ineligibility for phlebotomy from Week 8 to 32 and no more than one instance of phlebotomy eligibility between randomisation and Week 8.[67]
- Phlebotomy eligibility was defined as an HCT of >45% that was at least three percentage points higher than baseline, or an HCT of >48%, whichever was lower.[67]

The primary endpoint was reached by a significantly higher proportion of patients in the ruxolitinib group versus the BAT group (25/110 [22.7%] versus 1/112 [0.9%]; p<0.001)[3].[67, 69]

A numerically higher proportion of patients in the ruxolitinib group also achieved each of the component outcomes of the primary endpoint; HCT control at Week 32 occurred in 60.0% versus 18.8% of patients in the ruxolitinib group compared with the BAT group, respectively, whilst a reduction of ≥35% in spleen volume at Week 32 was seen in 40.0% versus 0.9% of patients in the ruxolitinib group compared with the BAT group, respectively.[67, 69]

Response rates were similar between patients who had unacceptable side effects from HC/HU and those who had an inadequate response (22.0% and 19.6%). No relationship between the primary endpoint and age, sex or baseline spleen volume was observed.[67]

Secondary efficacy outcomes

Duration of primary response

Primary response could occur at the Week 16 visit (if both components of the primary endpoint were met) but no later than the Week 32 visit. Duration of primary response was defined as the time from the first occurrence when both components of the primary endpoint are met until the date of the first documented disease progression.[96]

Progression was defined as the first occurrence of any of the following:

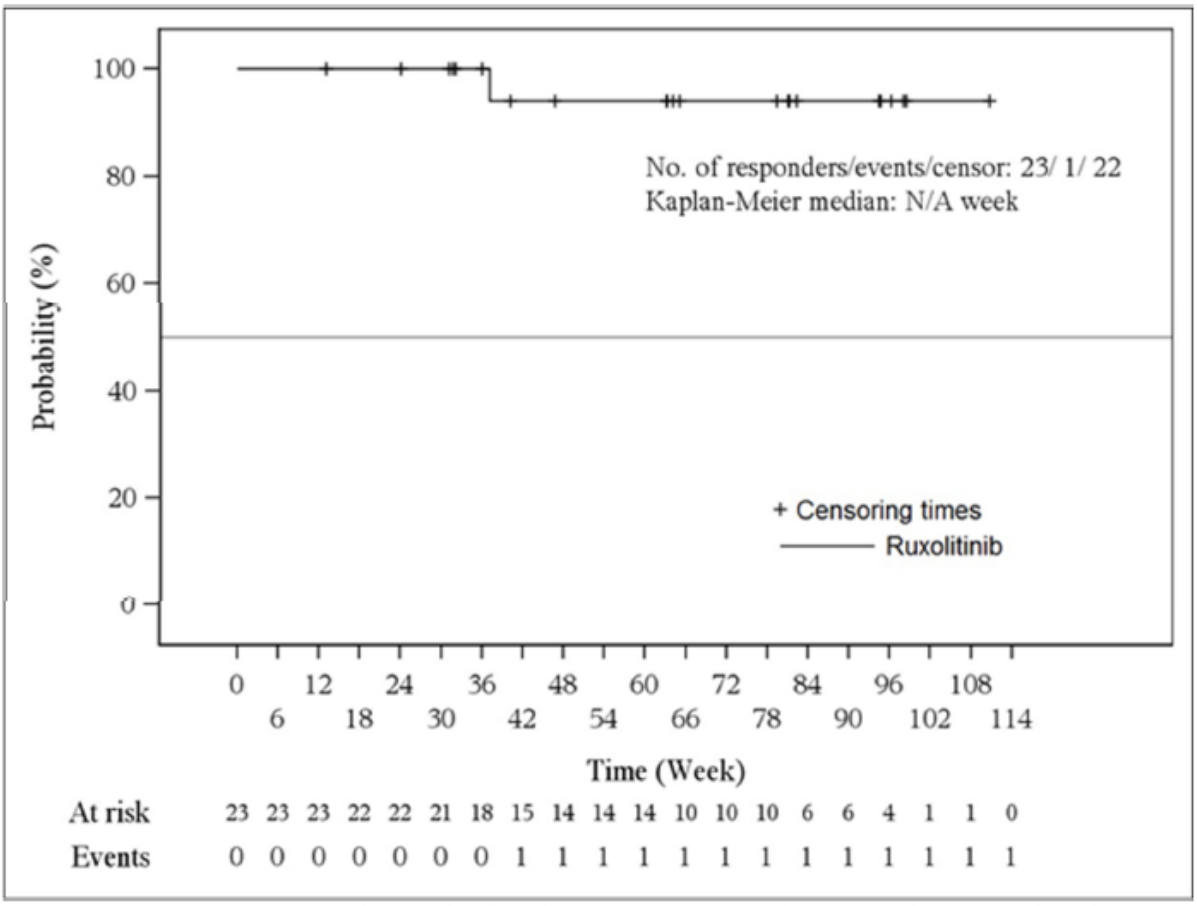
- The first of two consecutive HCT assessments that confirmed phlebotomy eligibility. A missing HCT assessment that was required to confirm phlebotomy eligibility was presumed to meet phlebotomy eligibility criteria;
- A spleen volume assessment by imaging (i.e., magnetic resonance imaging [MRI] or computerised tomography [CT]) that was reduced by <35% from the baseline and that was increased by ≥25% relative to the volume determined at the time of the best documented spleen volume response;

[3] The primary endpoint was originally reported as 20.9% in the ruxolitinib group in Vannucchi et al. 2015. However, following MRI review for the Week 80 analysis, two additional patients in the ruxolitinib group were identified as primary responders. This resulted in 22.7% of patients achieving the primary endpoint, as reported in Verstovsek et al. 2016.

- Death due to any cause;
- Development of MF (as determined by bone marrow biopsy);
- Development of AML (as determined by bone marrow blast counts of ≥20% or peripheral blast counts of ≥20% lasting ≥2 weeks).

The duration of primary response is shown in Figure 8 below. There was a 94% probability that a primary response to ruxolitinib would be maintained for one year from the time of initial response.[67]

Figure 8: Duration of response in the RESPONSE trial



Source: Vannucchi et al. 2015.[67]

Other secondary efficacy outcomes

CHR

CHR was defined as achieving HCT control with a platelet count ≤400×10^9/L and a WBC count ≤10×10^9/L.[96] CHR was observed in a significantly higher proportion of patients in the ruxolitinib group compared with the BAT group (23.6% versus 8.9%, respectively; p=0.0003).[67] At the Week 80 analysis, after correcting for a patient who had a phlebotomy at Week 8, only nine patients

achieved CHR at Week 32 (8.0%; unadjusted, p=0.0016; with adjustment for baseline WBC and platelet status, p=0.0013).[69]

Durable primary response at Week 48

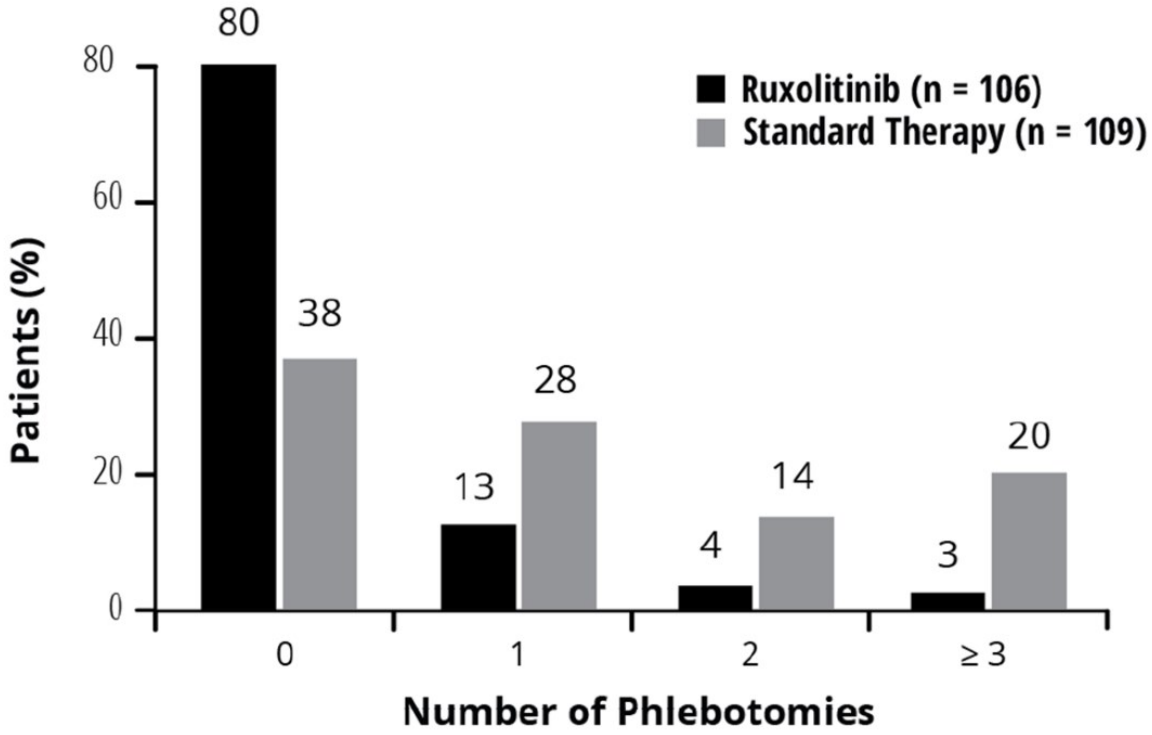
A durable primary response at Week 48 was defined as those patients achieving both components of the primary composite endpoint at Week 32 and maintained this response until Week 48.[96]

In the ruxolitinib group, 19.1% of patients had a primary response at Week 32 that was maintained until Week 48 compared to only 0.9% of patients in the BAT group (p<0.001).[67]

Rate of phlebotomy procedures between Week 8 and Week 32

The rate of phlebotomy procedures between Week 8 and Week 32 was greater in the BAT group as compared to the ruxolitinib group; 19.8% of patients in the ruxolitinib group and 62.4% of patients in the BAT group underwent at least one phlebotomy and 2.8% and 20.2% underwent three or more phlebotomies, respectively (see Figure 9).[67]

Figure 9: Rate of phlebotomy procedures between Week 8 and Week 32 in the RESPONSE trial[a]



[a] Includes patients who did not discontinue randomised treatment prior to Week 8.
Source: Vannucchi et al. 2015 (supplementary appendix).[102]

Symptoms and other patient-reported outcomes (PROs)

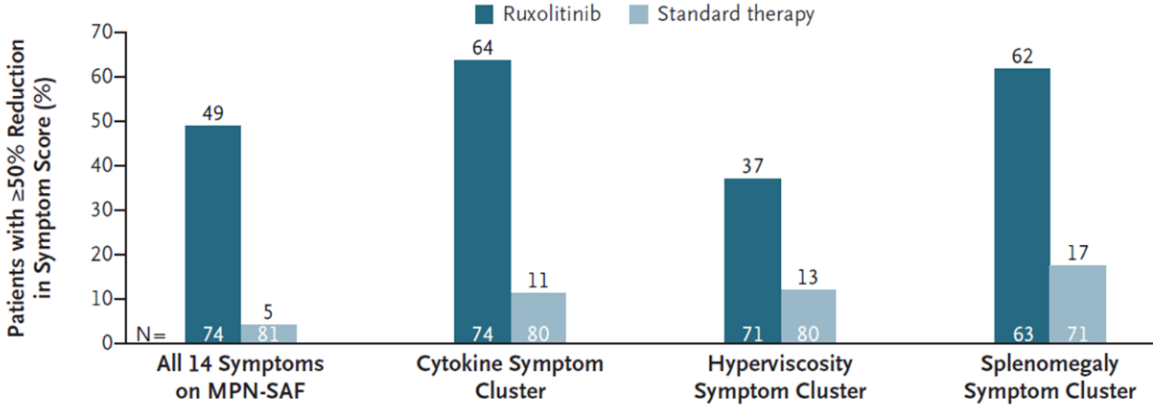
MPN-SAF patient diary

The MPN-SAF patient diary was used to assess 14 PV-related symptoms in three clusters; the cytokine symptom cluster (tiredness, itching, muscle ache, night sweats, and sweating while awake), the hyper-viscosity symptom cluster (vision problems, dizziness, concentration problems, headache, numbness or tingling in the hands or feet, ringing in the ears and skin redness), and

the splenomegaly symptom cluster (abdominal discomfort and early satiety).[67, 100] These symptoms are measured on a scale of 0–10, with higher scores indicating greater severity of symptoms. In RESPONSE, in addition to the MPN-SAF total score (the sum of the scores for the 14 symptoms), scores for individual symptoms and symptom clusters were also determined.[67, 100]

At Week 32, 49% (36/74) patients in the ruxolitinib group and 5% (4/81) patients in the BAT group had ≥50% reduction in the MPN-SAF total score (see Figure 10).[67] A ≥50% reduction in each symptom cluster was also achieved by a higher proportion of patients in the ruxolitinib group versus BAT; 64% versus 11% for the cytokine symptom cluster, 37% versus 13% for the hyper-viscosity symptoms cluster and 62% versus 17% for the splenomegaly symptom cluster (see Figure 10).[67]

Figure 10: Week 32 MPN-SAF total score in the RESPONSE trial



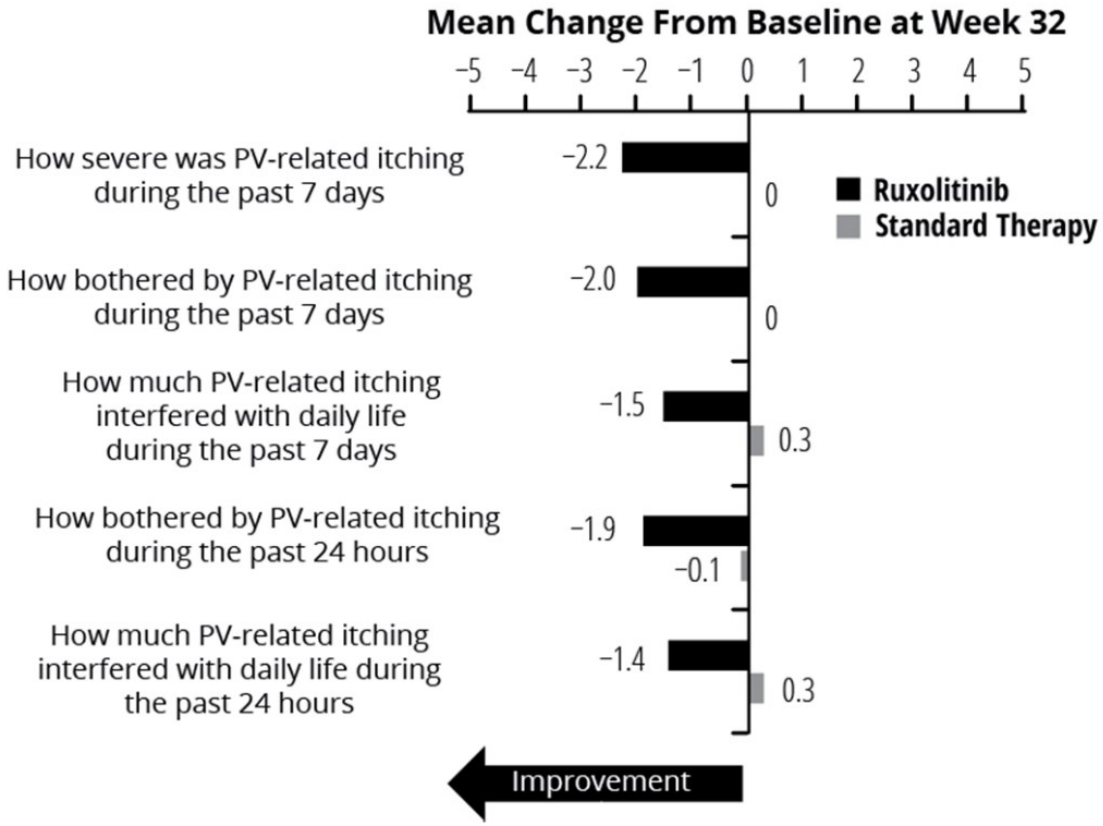
Abbreviations: MPN-SAF: Myeloproliferative Neoplasm Symptom Assessment Form.
Source: Vannucchi et al. 2015.[67]

Pruritus Symptom Impact Scale (PSIS)

The 5-question PSIS survey was used to evaluate pruritus severity and its impact on daily life on a scale from 0 (not at all) to 10 (worst imaginable).[79] The PSIS was completed at baseline and every four weeks from Week 4 to Week 32.[79]

Greater reductions in the PSIS were observed in the ruxolitinib group compared with the BAT group, with improvements in pruritus across all five questions at Week 32 for patients receiving ruxolitinib versus no change or worsening interference with daily life from PV-related itching in patients on BAT (see Figure 11).[67, 102] The improvements associated with ruxolitinib, when compared to the magnitude of the PSIS scale, demonstrate that itching may be alleviated with ruxolitinib treatment, allowing this symptom to have less of an impact on patients' daily lives.

Figure 11: Mean change from baseline on the PSIS at Week 32 in the RESPONSE trial[a]



[a] Includes patients with assessments at both baseline and Week 32. Scale from 0 (not at all) to 10 (worst imaginable).
Abbreviations: PSIS: Pruritus Symptom Impact Scale; PV: polycythaemia vera.
Source: Vannucchi et al. 2015 (supplementary appendix).[102]

EORTC QLQ-C30 and Patient Global Impression of Change (PGIC)

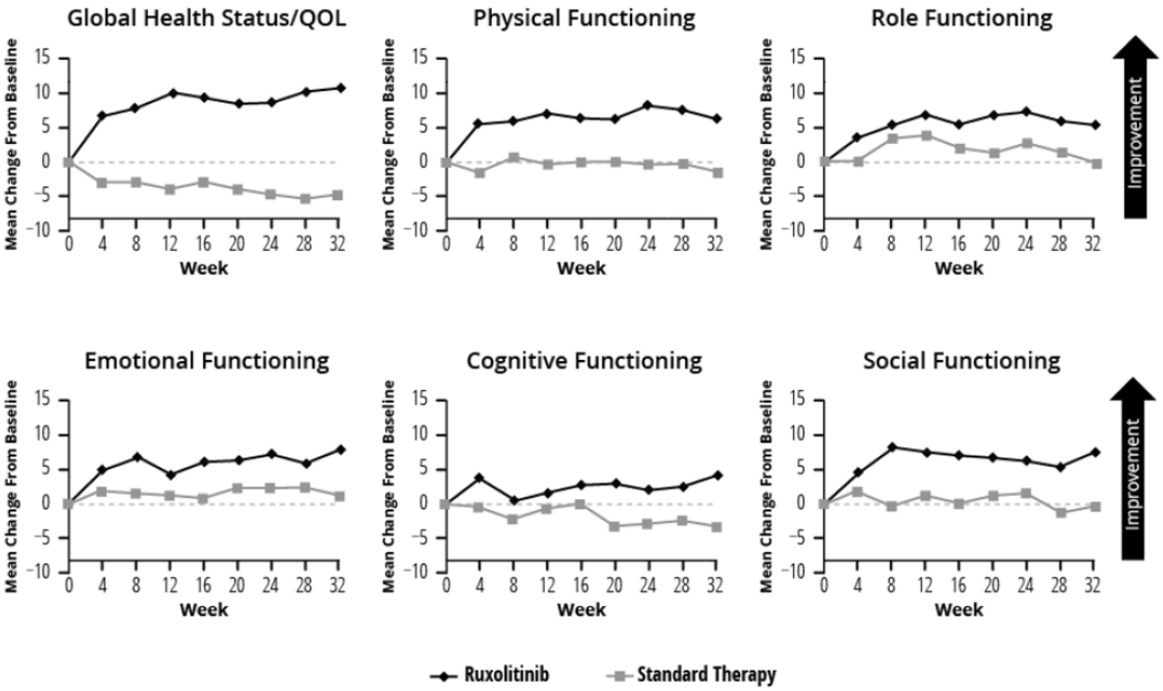
The EORTC-QLQ C30, a 30-item instrument comprising six functional scales (physical functioning, cognitive functioning, emotional functioning, social functioning, role functioning, and global quality of life) as well as nine symptom scales (fatigue, pain, nausea/vomiting, dyspnoea, insomnia, appetite loss, constipation, diarrhoea, and financial difficulties), is used to measure cancer-specific HRQoL. A high score for a functional scale of the EORTC indicates a high level of functioning, whereas a high score for a symptom scale or item represents a high level of symptomology or problems. The PGIC is composed of a single question intended to measure a patient's perspective of improvement or deterioration over time relative to treatment. The PGIC uses a seven-point scale where one equals 'very much improved' and seven equals 'very much worse'.[79, 100]

Patients who received ruxolitinib experienced improvements from baseline compared to Week 32 in EORTC QLQ-C30 global health status/QoL and all functional subscales. In contrast, patients who received BAT experienced functional deterioration from baseline to Week 32, with the exception of the emotional functioning subscale which was slightly improved (see Figure 12). Patients in the ruxolitinib group also experienced improvements in all individual symptoms measured by the EORTC QLQ-C30, including fatigue, insomnia, pain, appetite loss, dyspnoea, financial difficulties, diarrhoea, constipation, and nausea and vomiting. In comparison, individual symptom scores were less improved or worse for patients in the BAT group.[67, 79, 102] These results illustrate the important functional benefits conferred by ruxolitinib when compared to BAT.

A greater proportion of patients in the ruxolitinib group compared with BAT achieved a minimally important difference (MID; ≥10-point improvement from baseline) in global health status/QoL from baseline at each post-baseline study visit through Week 32. By Week 32, 46 patients (44%) in the ruxolitinib group achieved a MID, whereas only ten patients (9%) did so in the BAT group.[79]

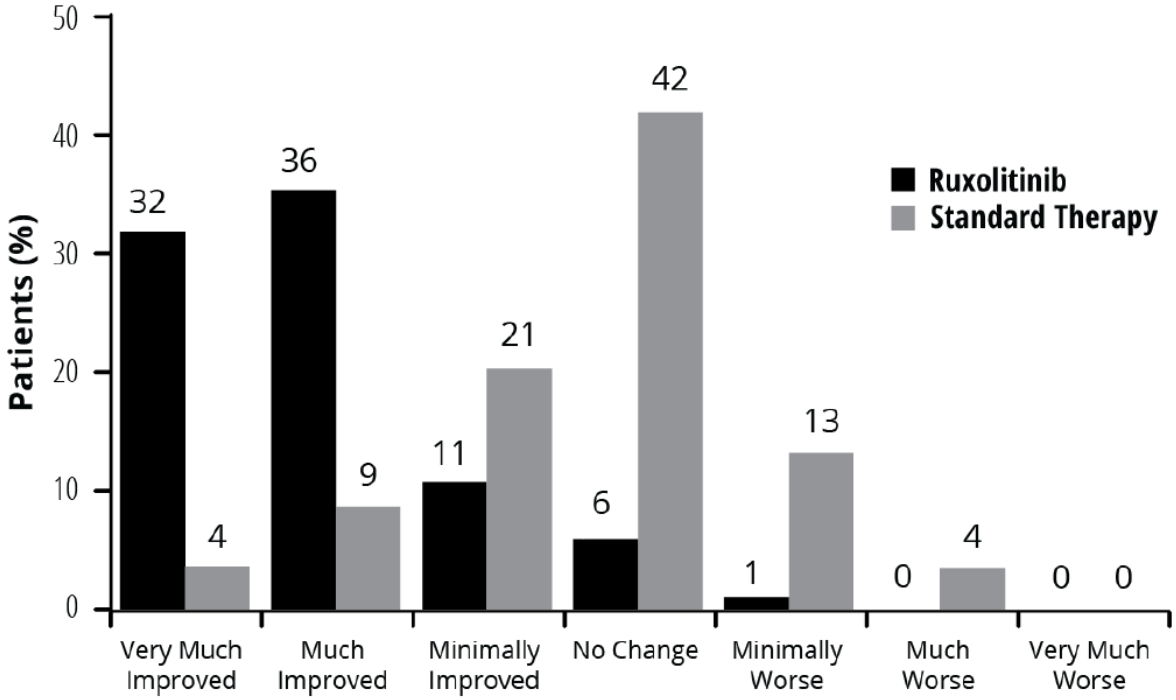
Ruxolitinib-associated improvements in PGIC were rapid and durable, with 46% of patients reporting that their condition was 'much' or 'very much' improved at Week 4 compared with 11% in the BAT group. Compared with the ruxolitinib group, a lower proportion of patients treated with BAT reported that their condition was 'much' or 'very much' improved between Weeks 4 and 32 (see Figure 13).[67, 79]

Figure 12: Mean change from baseline in EORTC QLQ-C30 QOL and Functioning Scores to Week 32 in the RESPONSE trial



Abbreviations: EORTC QLQ-C30 QOL: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30; QOL: quality of life.
Source: Vannucchi et al. 2015 (supplementary appendix).[102]

Figure 13: Patient Global Impression of Change at Week 32 in the RESPONSE trial



Source: Vannucchi et al. 2015 (supplementary appendix).[102]

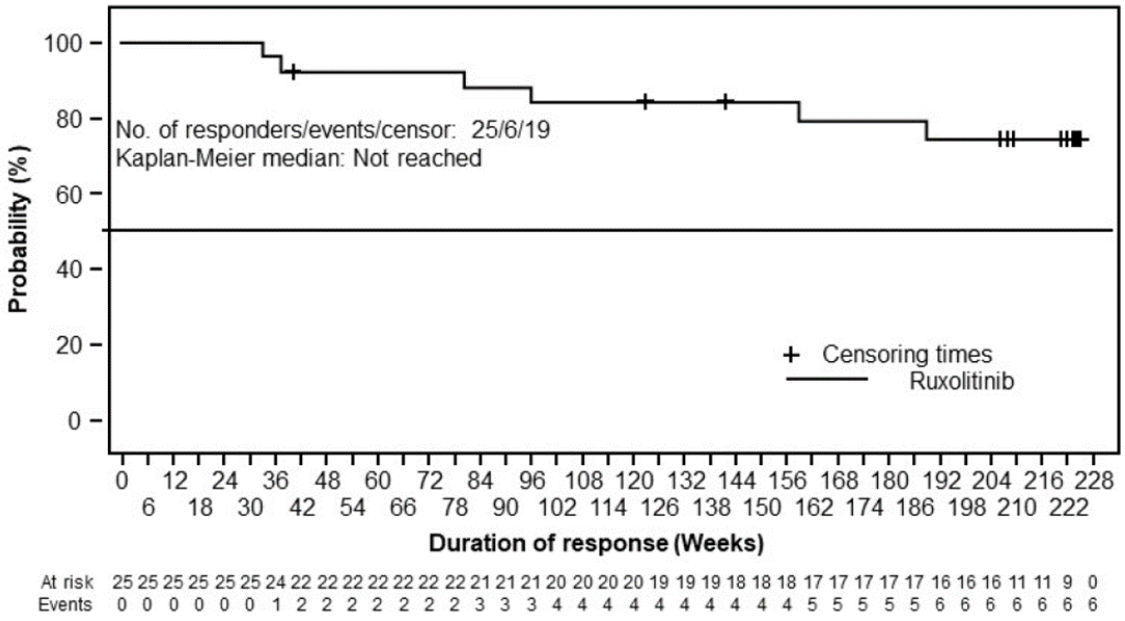
Efficacy outcomes at Week 256

Long-term efficacy results were reported after a 5-year follow-up (256 weeks), representing the final data cut-off from the RESPONSE trial. Due to all remaining patients crossing over from BAT to ruxolitinib at Week 80, long-term results comparing ruxolitinib with BAT should be interpreted with caution, and are predominantly reported for the ruxolitinib group only.[65]

Durability of HCT control (primary response)

The Kaplan–Meier estimated probability of maintaining a primary response from Week 32 for 224 weeks was 0.74 (95% CI: 0.51, 0.88) in the ruxolitinib group, with 6/25 (24.0%) primary responders having progressed by study completion (see Figure 14). Median duration of the primary response was not reached.[65]

Figure 14: Durability of primary response with ruxolitinib in the RESPONSE trial



Source: Kiladjian et al. 2020.[65]

CHR

The probability of maintaining a CHR from Week 32 to Week 256 was 55% (95% CI: 32, 73). Of 26 patients (24%) who achieved CHR at Week 32, ten had progressed by Week 256. Median duration of CHR was not reached.[65]

OS

OS was defined as the time from randomisation to the date of death due to any cause.[97]

During the study or in the survival follow-up phase, ten patients in the ruxolitinib group and nine patients in the BAT group died.[65] Kaplan–Meier estimates for OS at five years (ITT; not accounting for crossover) was 91.9% (95% CI: 84.4, 95.9) in the ruxolitinib group and 91.0% (95% CI: 82.8, 95.4) in the BAT group.[65] However, these results are confounded due to the high degree of crossover (all patients on BAT crossed over by Week 80).[65]

Transformation-free survival

Transformation free survival was defined as the time from the date of randomisation to the date of development of either MF or AML, as evidenced by bone marrow blast counts of at least 20% or peripheral blast counts of at least 20% lasting at least two weeks.[97]

In the ruxolitinib group, the exposure adjusted rates (per 100 patient-years) of transformation to MF and AML were 2.1 and 0.2 (see Section B.2.10).[65] The exposure adjusted rates (per 100 patient-years) of transformations to MF or AML were also low in the BAT group (1.4 and 0.0, respectively) and the crossover group (1.8 and 0.6, respectively).[65] The Kaplan–Meier estimate for transformation-free survival at five years for patients in the ruxolitinib group was [redacted] (95% CI: [redacted]).[97]

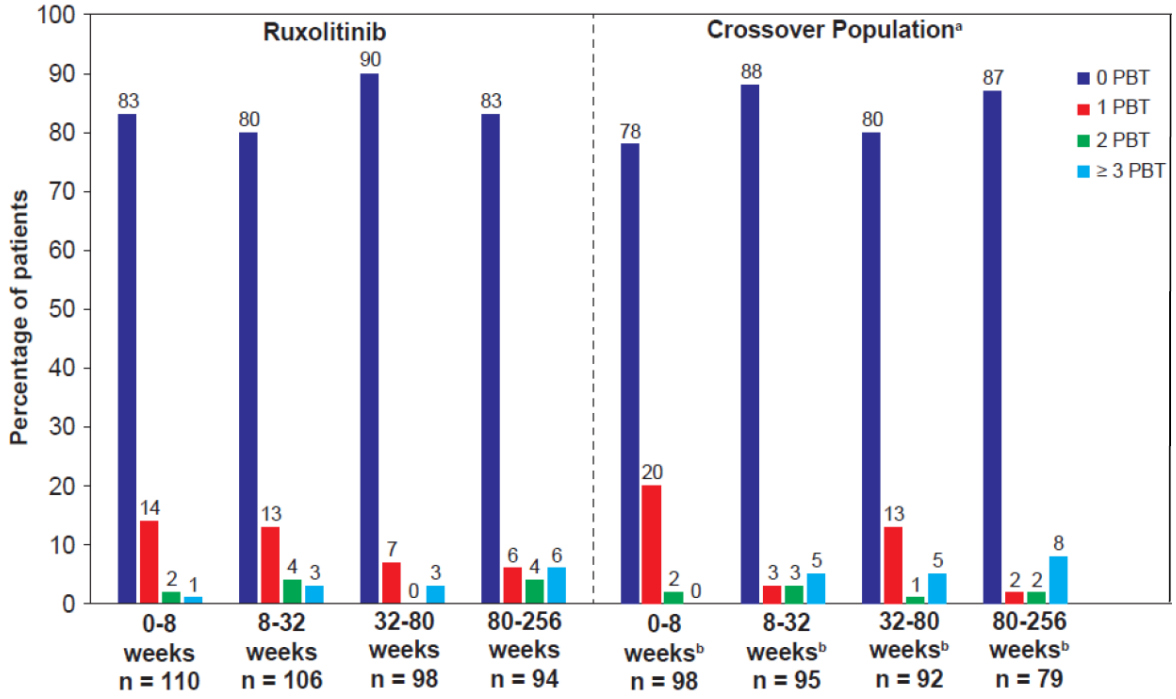
Rate of phlebotomy procedures from Week 80 up to Week 256

In the ruxolitinib group, 78 (83%) of 94 patients, who were evaluable from Week 80 up to Week 256, required no phlebotomies.[65] Only six (6%) of 94 patients needed three or more after Week 80 up to Week 256.[65]

In the crossover group, 69 (87%) of 79 patients, who were evaluable from Week 80 up to Week 256, required no phlebotomies.[65] Only six (8%) of 79 patients needing three or more phlebotomies at Week 224 of crossover.[65]

A summary of the number of phlebotomy procedures over time in ruxolitinib-treated patients is shown in Figure 15.

Figure 15: Number of phlebotomy procedures over time in ruxolitinib-treated patients



[a]All patients who crossed over from BAT. [b]From the time of crossover.
BAT: best available therapy; PBT: phlebotomy.
Source: Kiladjian et al. 2020 (supplementary appendix).[65]

Symptoms and other PROs

EORTC QLQ-C30

For patients in the ruxolitinib group at Week 256, a mean improvement from baseline in global health status score of +9.49 was observed. Improvement from baseline symptom scores was seen for fatigue, nausea, vomiting, pain, dyspnoea, insomnia, appetite loss, constipation and diarrhoea.[65, 97]

PSIS

PV related itching symptoms were reported to be 'very much improved' and 'much improved' in 36 (33%) and 12 (11%) patients, respectively, in the ruxolitinib group at Week 256. As 88% of patients randomised to BAT had crossed over to ruxolitinib therapy by Week 256, no PSIS results are available for the comparator arm.[97]

B.2.7.2 RESPONSE-2

Data from the RESPONSE-2 trial are presented from the Week 28 (primary) and Week 260 (final) analyses. Summaries of the efficacy outcomes at Week 80 and Week 156 have been provided in Appendix D for completeness. All outcomes are presented for the efficacy analysis set (ITT population of all patients who were randomised to a treatment group; n=149).

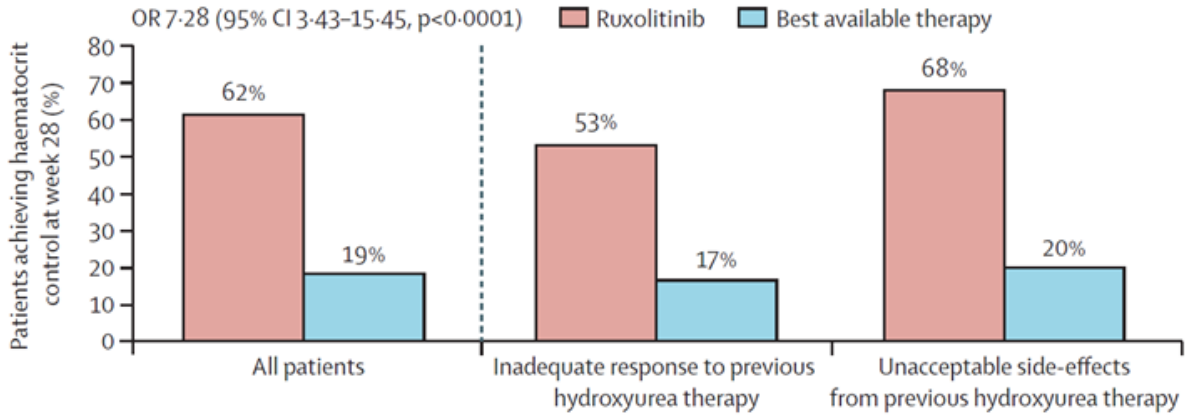
Primary efficacy outcome: Week 28

The primary endpoint in RESPONSE-2 was the proportion of patients who achieved HCT control at Week 28.[60]

- HCT control was defined as the absence of phlebotomy eligibility between Week 8 and Week 28, with phlebotomy eligibility occurring only once after randomisation and before Week 8.[60]
- Phlebotomy eligibility was defined as confirmed HCT level >45% and at least three percentage points higher than baseline, or confirmed HCT level >48%.[60]

HCT control at Week 28 was observed in 46/74 (62%) patients in the ruxolitinib group and 14/75 (19%) patients in the BAT group (OR: 7.28; 95% CI: 3.43, 15.45; p<0.0001; Figure 16).[60]

Figure 16: HCT control at Week 28 in the RESPONSE-2 trial



Abbreviations: OR: odds ratio.
Source: Passamonti et al. 2017.[60]

The primary efficacy results were also consistent across the subgroups assessed (age, sex and risk category [low; no risk factors versus high; one or two risk factors in patients aged >60 years or with thromboembolic history, or both]).[60]

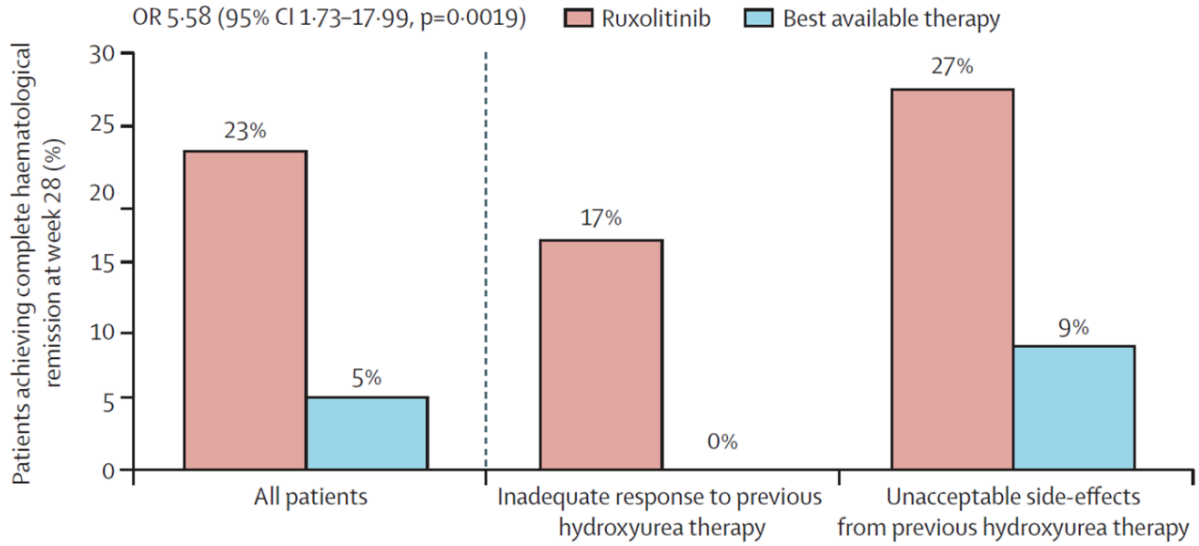
Secondary outcomes: Week 28

CHR

CHR was defined as HCT control, WBC count <10×10^9/L and platelet count ≤400×10^9/L.[60]

The proportion of patients achieving CHR was significantly higher in the ruxolitinib group compared with the BAT group (23% versus 5%, respectively; OR ruxolitinib versus BAT: 5.58; 95% CI: 1.73, 17.99; p=0.0019; see Figure 17). For both the ruxolitinib group and the BAT group, CHR was achieved by a higher proportion of patients who had unacceptable side effects from previous HC/HU treatment, compared to those who had an inadequate response to previous HC/HU treatment (Figure 17).[60]

Figure 17: CHR at Week 28 in the RESPONSE-2 trial

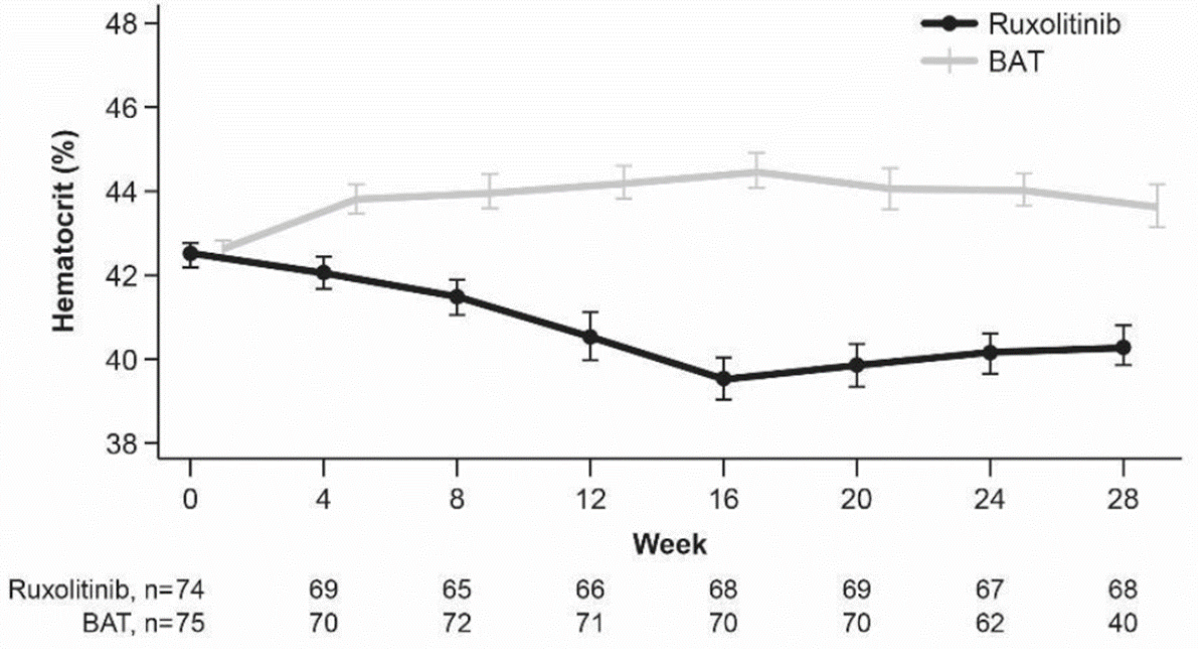


Abbreviations: CHR: complete haematological remission; CI: confidence interval; OR: odds ratio.
Source: Passamonti et al. 2017.[60]

Changes in HCT level up to Week 28

From baseline to Week 28, HCT level decreased in the ruxolitinib group (baseline mean [SD]: 42.8% [1.5]; Week 28 mean [SD]: 40.2% [4.1]). In contrast, in the BAT group, the HCT level increased from baseline to Week 28 (baseline mean [SD]: 42.7% [1.4]; Week 28 mean [SD]: 44.9% [3.8]; see Figure 18).[60]

Figure 18: Mean change in HCT level over time in the RESPONSE-2 trial

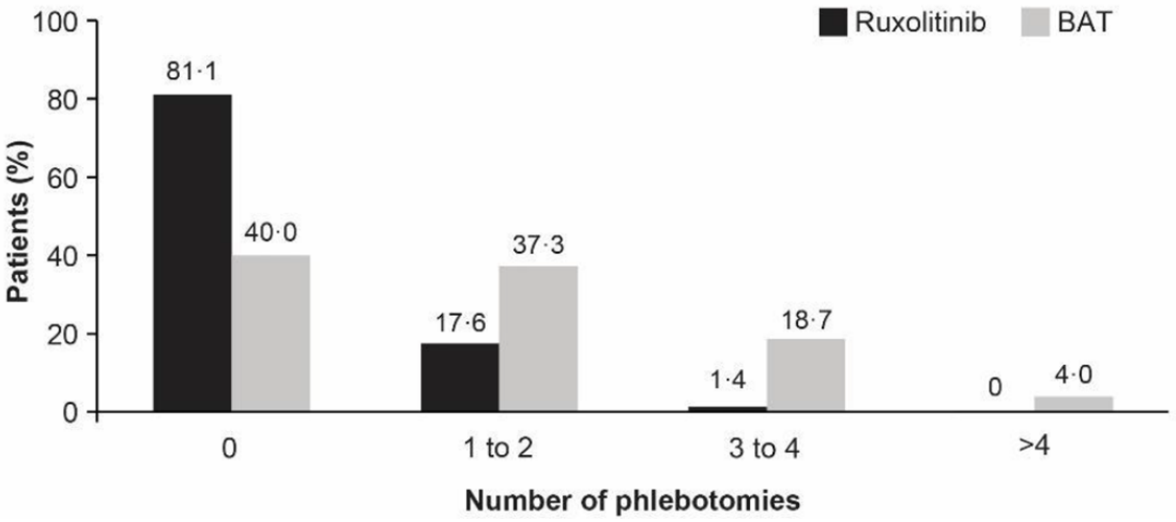


Abbreviations: BAT: Best Available Therapy; HCT: haematocrit.
Source: Passamonti et al. 2017 (supplementary appendix).[100]

Rate of phlebotomy procedures up to Week 28

The rate of phlebotomy procedures at Week 28 was greater in the BAT group as compared to the ruxolitinib group; 19% of patients underwent phlebotomy in the ruxolitinib group, compared with 60% of patients in the BAT group.[60] Rates of phlebotomy procedures up to Week 28 are shown in Figure 19.

Figure 19: Rates of phlebotomy procedures up to Week 28 in the RESPONSE-2 trial



Abbreviations: BAT: Best Available Therapy.
Source: Passamonti et al. 2017 (supplementary appendix).[100]

Symptoms and other PROs

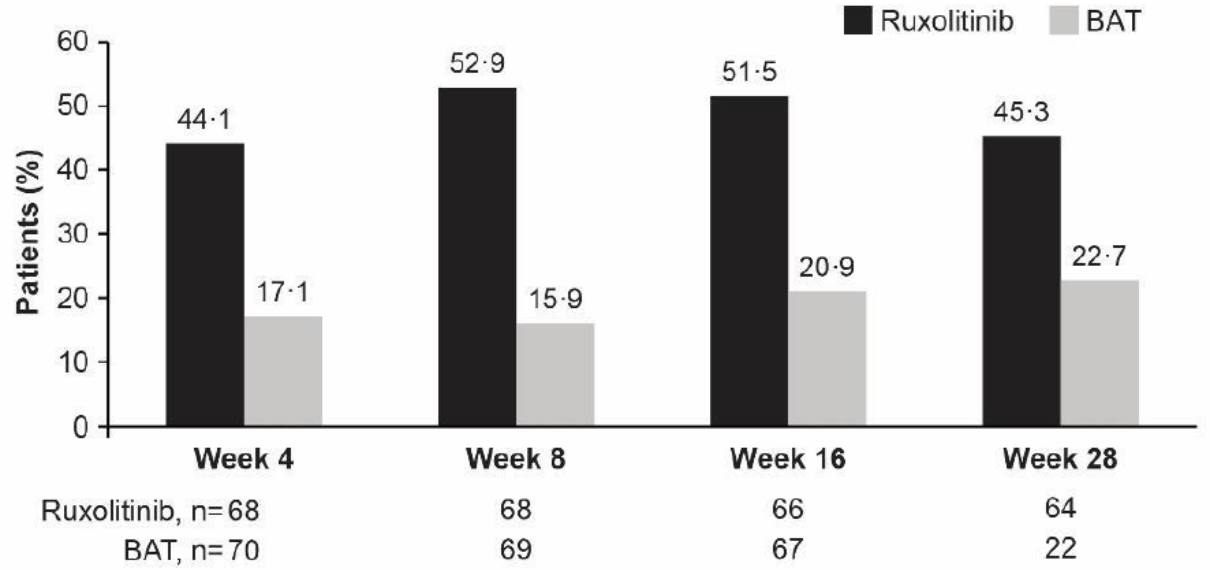
MPN-SAF total symptom score (TSS)

In RESPONSE-2, the MPN-SAF TSS, composed of ten items and measures related to MPN and the severity of nine of the most prevalent associated symptoms, including early satiety, abdominal discomfort, inactivity, concentration, night sweats, itching, bone pain, fever, and weight loss, was used to assess patient-reported PV symptom severity. Each item was scored on a scale ranging from 0 (absent) to 10 (worst imaginable).[100]

Patients in the ruxolitinib group had an improvement in symptoms as measured by MPN-SAF TSS, compared to a deterioration in patients in the BAT group. The median percentage change from baseline was -45.3% (IQR: -82.6, -8.0) for ruxolitinib and +2.4% (IQR: -55.8, 54.6) for BAT at Week 28.[60] Notably, improvements were seen as early as Week 4 in the ruxolitinib group. This was in contrast to the BAT group, in which MPN-SAF TSS increased through to Week 28.[60]

A higher proportion of patients in the ruxolitinib arm achieved a reduction in MPN-SAF TSS of ≥50% compared to the BAT group (see Figure 20).[100]

Figure 20: Proportion of patients achieving a ≥50% reduction in MPN-SAF TSS over time in the RESPONSE-2 trial



Abbreviations: BAT: Best Available Therapy; MPN-SAF TSS: Myeloproliferative Neoplasm Symptom Assessment Form total symptom score.
Source: Passamonti et al. 2017 (supplementary appendix).[100]

PSIS

Patients in the ruxolitinib group were observed to have rapid improvements in severity of pruritus as recorded by the PSIS score. In contrast, patients in the BAT group were observed to have a worsening in severity at most assessments.[60]

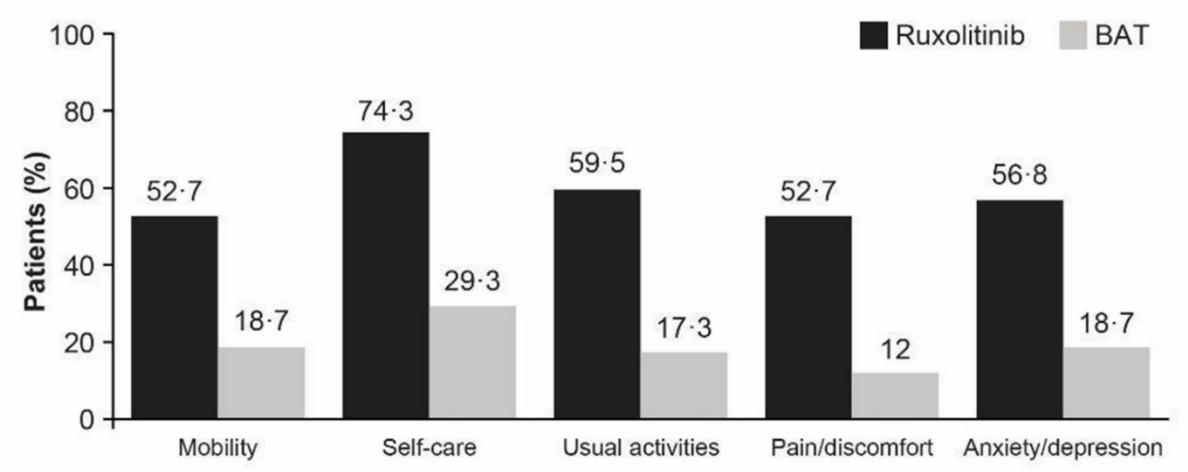
EQ-5D-5L and PGIC

EuroQol 5 Dimension 5 Level (EQ-5D-5L) is a standardised instrument consisting of five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and a visual analogue scale (VAS). The five dimensions are graded in five levels from no problems to extreme problems. The VAS ranges from 'best imaginable health state' to 'worst imaginable health state'. The scores are summarised into a single index score.[100] As mentioned above in the results

for PROs for the RESPONSE trial, PGIC is a single question used to measure a patient's perspective of improvement or deterioration over time relative to treatment, comprising of a seven-point scale from one (very much improved) to seven (very much worse).[100]

At Week 28 of the RESPONSE-2 trial, improvements were recorded in scores from the EQ-5D-5L and PGIC questionnaires in patients in the ruxolitinib group but little to no improvement was observed in those in the BAT group; 44 (60%) patients in the ruxolitinib group rated the change in their overall condition as 'much improved' or 'very much improved' on the PGIC compared with only four (5%) patients on BAT. Additionally, a higher proportion of patients in the ruxolitinib group reported having no problems in all five dimensions of the EQ-5D-5L (Figure 21).[60]

Figure 21: Proportion of patients reporting no problems in the individual domains of the EQ-5D-5L at Week 28 in the RESPONSE-2 trial



Abbreviations: BAT: Best Available Therapy.
Source: Passamonti et al. 2017 (supplementary appendix).[100]

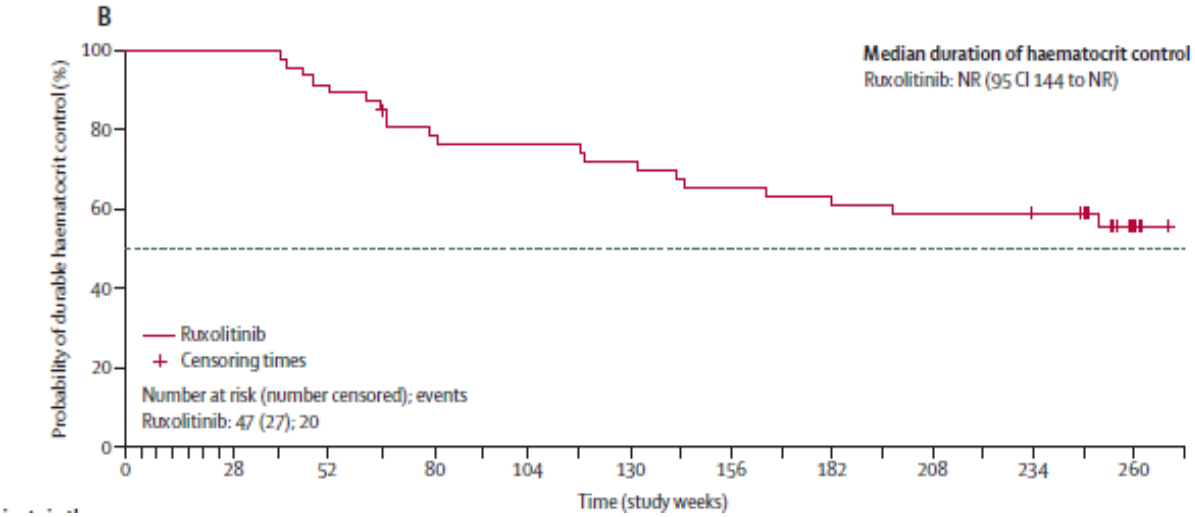
Efficacy outcomes at Week 260

Long-term efficacy results were reported after a 5-year follow-up (260 weeks), representing the final data cut-off from the RESPONSE-2 trial. At Week 260, there were no patients remaining in the BAT group due to crossover to ruxolitinib treatment after Week 80. Therefore, long-term results comparing ruxolitinib with BAT should be interpreted with caution, since they are predominantly reported for the ruxolitinib group only.[66]

Durability of HCT control (primary response)

At Week 260, 16/74 patients (21.6%) had achieved durable HCT control in the ruxolitinib group. The Kaplan–Meier estimated median duration of HCT control was not reached for the ruxolitinib group (Figure 22).[66]

Figure 22: Durability of HCT control with ruxolitinib in the RESPONSE-2 trial

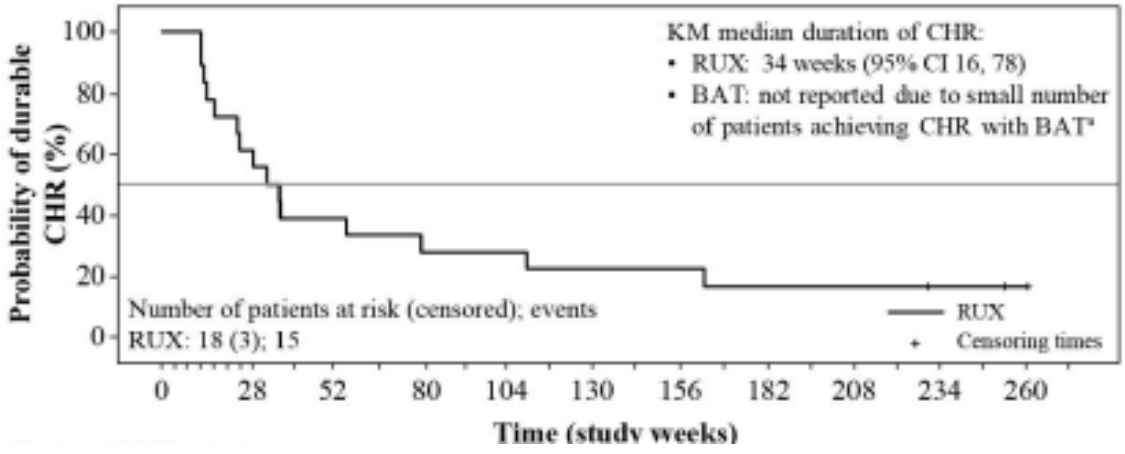


Abbreviations: CI: confidence interval; HCT: haematocrit; NR: not reported.
Source: Passamonti et al. 2022.[66]

CHR

At Week 260, 9/74 patients (12.2%) in the ruxolitinib group had durable CHR. Median duration of CHR for the ruxolitinib group was 34 weeks (95% CI: 16, 78) (from the start of CHR control at Week 28).[66]

Figure 23: Durability of CHR with ruxolitinib in the RESPONSE-2 trial



Abbreviations: BAT: Best Available Therapy; CHR: complete haematological response; KM: Kaplan–Meier; RUX: ruxolitinib.
Source: Passamonti et al. 2022.[66]

OS

OS was defined as the time from randomisation to death occurring before data cutoff.[66]

At Week 260, median OS was not reached. In the ruxolitinib group, three patients died during the study or the survival follow-up phase. The Kaplan–Meier estimate for OS at Week 260 was 96% (95% CI: 87, 99). In the BAT group, six patients died during the study or the survival follow-up

phase, with three deaths occurring after crossover. The Kaplan–Meier estimate for OS at Week 260 was 91% (CI: 80, 96).[66]

Transformation-free survival

The exposure-adjusted rates (per 100 patient-years) of transformation to MF at Week 260 were 0.6 for the ruxolitinib group, 1.9 for the BAT group and 0.5 for the crossover group.[66] Exposure adjusted rates (per 100 patient-years) for transformation to AML at Week 260 were 0 for ruxolitinib, 1.9 for the BAT group, and 0 for the crossover group (see Section B.2.10).[66] The Kaplan–Meier estimate for transformation-free survival for patients in the ruxolitinib group was 94% (95% CI: 85, 98).[66]

Rate of phlebotomy procedures up to Week 260

From baseline to Week 260, 60 phlebotomies were required among the 74 patients in the ruxolitinib group. In comparison, 106 phlebotomies were required among 75 patients in the BAT group at Week 80 (before crossover), which corresponds to a 5.7 times higher rate of phlebotomies compared with the ruxolitinib group (following adjustment for patient numbers and exposure time).[66] In the crossover group, 99 phlebotomies were required among 58 patients over the course of 232 weeks, which corresponds to a 2.4 times higher rate than the ruxolitinib group.[66]

A summary of the number of phlebotomy procedures over the trial period by treatment arm is provided in Table 11.

Table 11: Number of phlebotomy procedures at latest available timepoint by treatment arm

	Ruxolitinib (N=74)	BAT (N=75)	Crossover (N=58)
Total number of phlebotomies from baseline to end of treatment, n			
	60 (within 260 weeks)	106[a] (within 80 weeks)	99[b] (within 232 weeks)
Phlebotomy frequency category, n (%) of patients			
0	51 (69)	27 (36)	16 (28)
1 or 2	12 (16)	29 (39)	23 (40)
>2 to ≤4	7 (10)	16 (21)	16 (28)
>4 to ≤6	4 (5)	2 (3)	2 (3)
>6 to ≤8	0	1 (1)	1 (2)

Safety set (N=149). Phlebotomy eligibility defined as confirmed HCT >45% and ≥3% higher than baseline, or HCT >48%. [a]Includes BAT patients who crossed over to ruxolitinib but includes only phlebotomies that occurred prior to crossover. [b]Includes only phlebotomies that occurred after crossover from BAT to ruxolitinib.
Abbreviations: BAT: best available therapy; HCT: haematocrit.
Source: Passamonti et al. 2022 (supplementary appendix).[66]

B.2.8 *Subgroup analysis*

Results from the primary analyses of the RESPONSE and RESPONSE-2 trials were consistent across all pre-specified subgroups (Figure 24 and Figure 25, respectively).

Figure 24: Subgroup Forest plot of OR of patients achieving primary response at week 32 in RESPONSE (FAS)

Abbreviations: CI: confidence interval; FAS: full analysis set; OR: odds ratio.
Source: Novartis Data on File (RESPONSE Week 32 CSR).

Figure 25: Subgroup Forest plot of OR for patients achieving HCT control at Week 28 in RESPONSE-2 (FAS)

Abbreviations: CI: confidence interval; FAS: full analysis set; HCT: haematocrit; OR: odds ratio.
Source: Novartis Data on File (RESPONSE-2 Week 28 CSR).

Please refer to Appendix E for subgroup analyses of the following from both REPONSE and RESPONSE-2: patients who had received prior IFN-alfa; IFN-alfa as investigator determined BAT or ruxolitinib after crossover from IFN-alfa.

B.2.9 *Indirect and mixed treatment comparisons*

The RESPONSE and RESPONSE-2 trials represent the main sources of clinical evidence for ruxolitinib in this indication. These trials provide direct evidence for ruxolitinib versus the relevant comparator for this submission, BAT. However, given that a high proportion of patients crossed over to ruxolitinib from BAT at Week 32 and Week 28 in the RESPONSE and RESPONSE-2 trials and the number of deaths in these trials was low, crossover adjustment methods were not possible. Therefore, a MAIC was conducted to provide an estimate of OS for ruxolitinib versus BAT that was not confounded by crossover.[68]

In the MAIC, propensity score matching (PSM) was conducted using individual patient level data (IPD) for patients in RESPONSE, and IPD from real-world patients with PV treated with BAT in the GEMFIN registry. The GEMFIN registry is a retrospective survey of patients with PV, which opened in July 2011 for Spanish hospitals affiliated with GEMFIN, and data from patients with R/I to HC/HU have been published previously.[9, 68]

The methods and results of the MAICs using data from the Week 208 analysis of the RESPONSE trial have been presented by Alvarez-Larrán et al. 2018.[103] The MAIC for OS has since been updated using the Week 256 data from the RESPONSE trial; however, there are a number of limitations associated with this analysis (as outlined in Section B.2.9.2).[68]

A summary of the key results and limitations of the MAIC have been presented below, with full details of the methodology and results provided in Appendix D.

B.2.9.1 Results

Baseline characteristics

By October 2016, a total of [redacted] patients from [redacted] hospitals were included in the registry. Of the patients included in the registry, [redacted] were identified as being R/I to HC/HU (and thus being of relevance to the patient population of interest), and of these [redacted] patients were being treated with BAT (confirmed at the last visit) and had at least one follow-up visit since the initiation date.[68]

Prior to matching, there were notable differences between the GEMFIN BAT population (n=[redacted]) and the RESPONSE ruxolitinib population (n=110) in some of the baseline patient characteristics that were used for the PSM: age, cytopaenia at the lowest HC/HU dose, and gender (for OS; see Table 12). In contrast, the matched cohorts (n=[redacted] in each cohort) were balanced for all covariates included in the PSM analysis (see Table 12).[68]

The matched BAT cohort included patients treated with HC/HU ([redacted]%), IFN-alfa ([redacted]%), anagrelide ([redacted]%), busulfan ([redacted]%), radioactive phosphorus ([redacted]%), other therapy ([redacted]%), or no cytoreductive therapy ([redacted]%). Approximately [redacted] of patients received concomitant treatment with more than one therapy.[68]

Table 12: Baseline patient characteristics: before and after matching for GEMFIN and RESPONSE

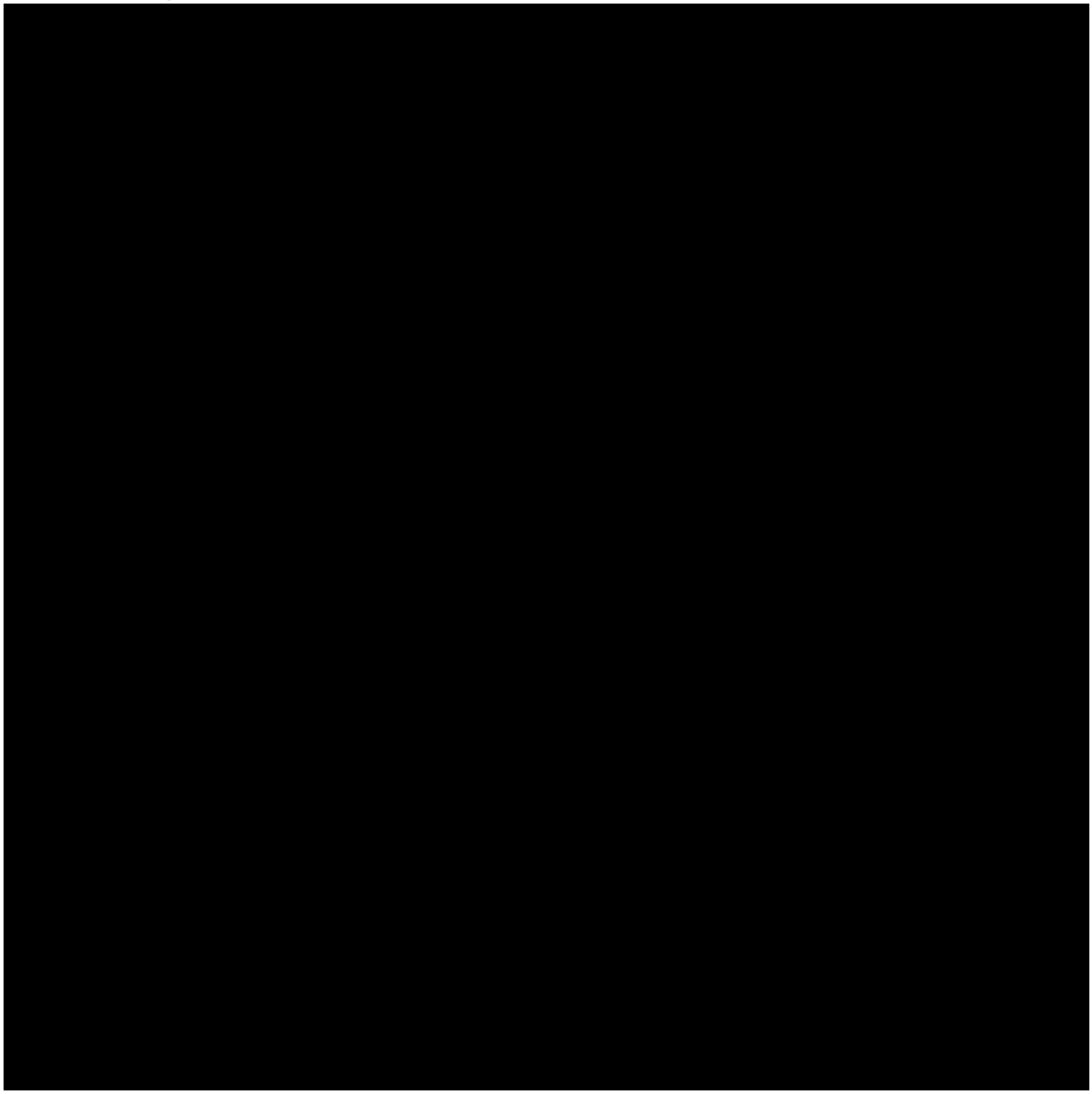
Patient characteristic	Pre-matching			Post-matching		
	GEMFIN (n=[redacted])	RESPONSE (n=110)	Standardised difference	GEMFIN (n=[redacted])	RESPONSE (n=[redacted])	Standardised difference
Age, mean years (SD)	[redacted]	61.16 (10.5)	[redacted]	[redacted]	[redacted]	[redacted]
History of thrombosis at time of R/I, n (%)	[redacted]	39 (35%)	[redacted]	[redacted]	[redacted]	[redacted]
Cytopaenia at the lowest HC/HU dose, n (%)	[redacted]	17 (15%)	[redacted]	[redacted]	[redacted]	[redacted]
Male, n (%)	[redacted]	66 (60%)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HC/HU: hydroxycarbamide/hydroxyurea; R/I: resistance to or intolerance of; SD: standard deviation.
Source: Novartis Data on File (MAIC).[68]

MAIC for OS for the RESPONSE trial (ruxolitinib) versus the GEMFIN registry (BAT)

Before matching, patients treated with ruxolitinib had significantly improved OS (HR: [redacted]; 95% CI: [redacted]) compared with patients treated with BAT (Table 13; please refer to Appendix D for the Kaplan–Meier plot pre-matching). After PSM, patients treated with ruxolitinib maintained a significantly improved OS (HR: [redacted]; 95% CI: [redacted]) compared with patients treated with BAT (Figure 26 and Table 13).[68]

Figure 26: Kaplan–Meier plot for OS for the RESPONSE trial versus the GEMFIN registry – post-matching



* Based on Cox proportional hazards model, only treatment arm (BAT/Ruxolitinib) was used to estimate hazard ratio
** Remaining differences in patient characteristics used in matching are also adjusted for in the Cox proportional hazard model (doubly robust)
Abbreviations: BAT: Best Available Therapy; CI: confidence interval; OS: overall survival.
Source: Novartis Data on File (MAIC).[68]

Table 13: Summary of results for OS for RESPONSE trial versus GEMFIN registry

Analysis	Number of patients		Number of events		HR (95% CI)[a]
	BAT	Ruxolitinib	BAT	Ruxolitinib	
Pre-matching[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-matching[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Analysis	Number of patients		Number of events		HR (95% CI)[a]
	BAT	Ruxolitinib	BAT	Ruxolitinib	
Post-matching[c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] Based on Cox proportional hazards model with a value less than 1 favouring ruxolitinib. [b] Treatment arm (BAT/ruxolitinib) was used to estimate HR. [c] Treatment arm (BAT/ruxolitinib) and covariates used in matching were used to estimate HR.
Abbreviations: BAT: Best Available Therapy; CI: confidence interval; HR: hazard ratio; OS: overall survival.
Source: Novartis Data on File (MAIC).[68]

B.2.9.2 Uncertainties in the indirect and mixed treatment comparisons

A full discussion of the strengths and limitations of the MAIC has been provided in Appendix D. The key sources of uncertainty associated with the MAIC are as follows:

- **Generalisability of the GEMFIN registry:** The GEMFIN registry was chosen as a source of data for BAT as it is one of the largest registries of patients with PV and data from patients who are R/I to HC/HU have been published previously. However, as the registry is Spanish, there is some uncertainty as to whether the population and treatments received by patients are representative of UK clinical practice. While clinical experts agreed that the patient population in GEMFIN is broadly reflective to that of UK clinical practice, it was also noted that the use of IFN-alfa was comparatively low. Furthermore, clinical experts noted that GEMFIN has a historical cohort and consequently, the management of key complications such as AML, MF and thromboembolic events is likely to have improved. Consequently, there is some uncertainty as to whether the MAIC is generalisable to UK clinical practice.[11, 68]
- **The follow-up time for GEMFIN is relatively short:** The median follow-up time for GEMFIN is shorter than RESPONSE (approximately [redacted] years versus 5 years, respectively). Therefore, further analyses using longer follow-up data may be required to validate the findings of the MAIC.
- **RESPONSE-2 trial data were not considered in the MAIC:** At the time of this analysis, the feasibility of conducting MAICs using pooled data from both RESPONSE and RESPONSE-2 has also been explored. However, there was considerable overlap in the number of GEMFIN patients that could be matched to RESPONSE and RESPONSE-2 when these were considered independently, and so matching to the combined population resulted in a poor fit when estimating propensity scores, as these patients could not be double counted. Therefore, this approach (i.e. matching to RESPONSE and RESPONSE-2 pooled) was not deemed to be feasible. Matching to patients in the RESPONSE-2 trial alone was also explored at the time, however, a MAIC for OS versus RESPONSE-2 alone was not performed as no deaths in the ruxolitinib arm had occurred at the time of the MAIC analysis.
- **Matching was only feasible for a limited number of covariates:** PSM is a well-established approach that aims to reduce selection bias when comparing the effectiveness of two treatments across different studies. In the MAIC, a total of [redacted] patients were matched, representing [redacted]% of ruxolitinib-treated patients in the RESPONSE trial.[68] However, the large variation of age between patients in the RESPONSE trial and the GEMFIN registry meant that matching was only feasible for a limited number of covariates. Moreover, potentially important clinical factors could not be used as propensity score covariates. For

example, time since PV diagnosis could not be used because the definitions did not align between RESPONSE and GEMFIN. Additionally, as RESPONSE only included patients with splenomegaly, patients with splenomegaly from RESPONSE were matched to patients without splenomegaly from the GEMFIN registry. The impact of not adjusting for these characteristics and any unobserved differences between RESPONSE and GEMFIN on the HR derived from the MAIC is not known.[68]

B.2.9.3 Conclusions

In order to provide estimates of the relative efficacy of ruxolitinib versus BAT in OS which are not confounded by crossover, the MAIC was conducted using data from the GEMFIN registry and the RESPONSE trial (Week 256 analysis). The results of the MAIC demonstrated that after PSM patients treated with ruxolitinib had significantly improved OS (HR: [redacted]; 95% CI: [redacted]) compared with patients treated with BAT.[68] Although there are limitations associated with the MAIC, the observed improvement in OS for ruxolitinib compared to BAT, supports the hypothesis that ruxolitinib is associated with a survival advantage.[60, 67, 104]

B.2.10 *Adverse reactions*

Safety data for the longest-available follow-up in RESPONSE and RESPONSE-2 has been summarised in the section below. In this submission, these results have predominantly been presented as exposure-adjusted rates to account for the different durations of exposure to the study drugs in the ruxolitinib, BAT and crossover groups in RESPONSE and RESPONSE-2. Please refer to Appendix F supplementary AE data.

B.2.10.1 RESPONSE

Week 256 analysis

AEs

At Week 256, the overall safety profile for ruxolitinib remained consistent with the previous data cut-offs, with no new long-term safety signals identified. Almost all patients in the ruxolitinib (n=[redacted] and crossover groups (n=[redacted] had experienced at least one AE,[97] while [redacted] patients ([redacted] in the BAT arm experienced AEs prior to crossover at Week 32.[101]

As presented in Table 14, in the ruxolitinib group, the most frequent AEs of any grade (exposure-adjusted rate per 100 person-years) were anaemia (8.9), pruritus (7.0), diarrhoea (7.0), weight increased (6.1), headache (5.8), arthralgia (5.6), fatigue (5.1), and muscle spasms (5.1). In the BAT group, the most frequent AEs of any grade (exposure-adjusted rate per 100 person-years) were pruritus (32.6), headache (28.5), fatigue (23.1) and abdominal pain (17.7). In the crossover group, the most frequent AEs of any grade were anaemia (8.8), pruritus (6.1), dizziness (6.1), back pain (5.5) and headache (5.2).[65]

A summary of the most frequent AEs (with an exposure-adjusted rate ≥5 per 100 patient-years) at Week 256 is presented in Table 14.

Table 14: Exposure-adjusted rates of common AEs occurring at a rate of ≥5 per 100 patient-years of exposure in any group[a]

AEs, n (exposure adjusted rate per 100 person-years)[b]	Ruxolitinib rate (n=110)[c]		BAT rate (n=111)[d]		Crossover rate (n=98)[e]	
	All grades	Grade 3-4	All grades	Grade 3-4	All grades	Grade 3-4
Haematological adverse events						
Anaemia	8.9	0.9	5.4	0.0	8.8	0.6
Thrombocytopenia	4.4	1.2	16.3	2.7	1.2	0.3
Non-haematological adverse events						
Pruritus	7.0	0.5	32.6	5.4	6.1	0.0
Diarrhoea	7.0	0.2	12.2	1.4	3.6	0.0
Increased weight	6.1	0.7	1.4	0.0	4.2	0.6
Headache	5.8	0.5	28.5	1.4	5.2	0.0
Arthralgia	5.6	0.2	10.9	1.4	3.3	0.3
Fatigue	5.1	0.2	23.1	4.1	3.9	0.0
Muscle spasms	5.1	0.2	9.5	0.0	3.3	0.0
Pyrexia	4.0	0.2	6.8	0.0	3.3	0.3
Dizziness	4.0	0.0	15.0	0.0	6.1	0.0
Back pain	4.0	0.2	6.8	0.0	5.5	0.3
Hypertension	4.0	0.5	5.4	1.4	4.5	0.9
Abdominal pain	3.7	0.5	17.7	0.0	3.0	0.3
Nausea	3.5	0.2	5.4	0.0	2.1	0.0
Night sweats	3.0	0.0	12.2	0.0	1.8	0.0
Pain in extremity	2.3	0.2	5.4	0.0	3.3	0.0
Decreased appetite	2.1	0.2	8.2	0.0	1.5	0.0
Musculoskeletal pain	1.9	0.2	5.4	0.0	1.8	0.0
Myalgia	1.6	0.0	10.9	0.0	1.2	0.0
Paraesthesia	1.6	0.0	9.5	0.0	2.4	0.3
Vertigo	1.6	0.0	5.4	0.0	1.2	0.0
Abdominal distension	1.4	0.2	5.4	0.0	0.3	0.0
Vomiting	1.4	0.0	5.4	0.0	2.4	0.3
Peripheral neuropathy	1.4	0.0	6.8	1.4	0.6	0.0
Bone pain	0.9	0.0	8.2	1.4	1.2	0.3
Hyperuricaemia	0.7	0.2	6.8	2.7	0.9	0.0
Gout	0.2	0.0	6.8	2.7	0.3	0.0
All infections	18.9	3.5	59.8	4.1	19.1	6.1
Herpes zoster infection	4.7	0.5	0.0	0.0	3.9	0.6
Nasopharyngitis	4.4	0.0	12.2	0.0	4.2	0.0
Bronchitis	3.3	0.0	6.8	0.0	3.9	0.3
Upper respiratory tract infection	2.3	0.0	6.8	0.0	2.4	0.0
Cellulitis	0.2	0.2	5.4	1.4	0.6	0.6

[a]Adverse events occurring at a rate of ≥5 per 100 patient-years of exposure in any group, regardless of relationship to study drug. [b]Adjusted rates were calculated as the number of patients with events per 100 patient-year of exposure. [c]Exposure=428.4 patient-years. [d]Exposure=73.6 patient-years. [e]Exposure=329.9 patient-years.
Abbreviations: AE: adverse event; BAT: best available therapy.
Source: Kiladjian et al. 2020.[65]

As presented in Table 15, in the ruxolitinib group, the most frequent serious adverse events (SAEs; exposure-adjusted rate per 100 person-years) were pneumonia (1.2), squamous cell carcinoma (0.9), atrial fibrillation (0.7) and basal cell carcinoma (0.7).[65] In the BAT group, some of the most frequent AEs of any grade (exposure-adjusted rate per 100 person-years) were pneumonia, atrial fibrillation, cellulitis and diverticulitis, all with a rate of 1.4.[65] For the crossover group, the most frequent SAE was pneumonia at a rate of 1.8.[65]

Table 15: Exposure-adjusted rates of SAEs occurring at a rate of ≥0.5 per 100 patient-years of exposure in any group[a]

AEs, n (exposure adjusted rate per 100 person-years)[b]	Ruxolitinib rate (n=110)[c]		BAT rate (n=111)[d]		Crossover rate (n=98)[e]	
	All grades	Grade 3 or 4	All grades	Grade 3 or 4	All grades	Grade 3 or 4
Pneumonia	1.2	1.2	1.4	1.4	1.8	1.8
Squamous cell carcinoma	0.9	0.9	0.0	0.0	0.3	0.3
Atrial fibrillation	0.7	0.7	1.4	1.4	0.3	0.0
Basal cell carcinoma	0.7	0.5	0.0	0.0	0.3	0.3
Rectal haemorrhage	0.5	0.5	0.0	0.0	0.0	0.0
Chest pain	0.5	0.2	0.0	0.0	0.3	0.3
Metastatic squamous cell carcinoma	0.5	0.5	0.0	0.0	0.0	0.0
Squamous cell carcinoma of skin	0.5	0.5	0.0	0.0	0.0	0.0
Dehydration	0.5	0.2	0.0	0.0	0.0	0.0
Cellulitis	0.2	0.2	1.4	1.4	0.6	0.6
Herpes zoster	0.2	0.2	0.0	0.0	0.6	0.3
Urinary tract infection	0.2	0.2	0.0	0.0	0.6	0.6
Diverticulitis	0.2	0.2	1.4	1.4	0.3	0.3
Malignant melanoma	0.2	0.2	1.4	1.4	0.0	0.0
Prostate cancer	0.2	0.2	0.0	0.0	0.6	0.6
Subdural hematoma	0.0	0.0	1.4	0.0	0.0	0.0
Gout	0.0	0.0	1.4	1.4	0.0	0.0
Pulmonary embolism	0.0	0.0	1.4	1.4	0.0	0.0

Deep vein thrombosis	0.0	0.0	1.4	1.4	0.0	0.0
Bladder disorder	0.0	0.0	1.4	0.0	0.0	0.0
Abdominal Pain	0.0	0.0	0.0	0.0	0.6	0.3
Dyspnoea	0.0	0.0	0.0	0.0	0.6	0.0
Epistaxis	0.0	0.0	0.0	0.0	0.6	0.6
Acute myocardial infarction	0.0	0.0	1.4	1.4	0.0	0.0
Transient ischaemic attack	0.0	0.0	0.0	0.0	0.6	0.6
Varicella zoster virus infection	0.0	0.0	0.0	0.0	0.6	0.6
Gastroenteritis	0.0	0.0	1.4	0.0	0.0	0.0

[a]AEs occurring at a rate of ≥0.5 per 100 patient-years of exposure in any group, regardless of relationship to study drug. [b]Adjusted rates were calculated as the number of patients with events per 100 patient-years of exposure. [c]Exposure=428.4 patient-years. [d]Exposure=73.6 patient-years. [e]Exposure=329.9 patient-years.
Abbreviations: AE: adverse event; BAT: best available therapy; SAEs: serious adverse events.
Source: Kiladjian et al. 2020.[65]

Exposure-adjusted rates of selected AEs are reported in Table 16. Thromboembolic event rates (per 100 patient-years) were lower for the ruxolitinib group (1.2) and the crossover group (2.7), compared to the BAT group (8.2) at Week 256.[65] Rates of transformation (per 100 patient-years) to MF were similar for all three treatment groups: 2.1 for the ruxolitinib group, 1.8 for the crossover group and 1.4 for the BAT group.[65]

Table 16: Exposure-adjusted AEs of interest at Week 256 in the RESPONSE trial

AEs, n (exposure adjusted rate per 100 person-years)	Ruxolitinib	Crossover	BAT
Thromboembolic event	1.2	2.7	8.2
Second malignancies	7.0	4.5	4.1
Non-melanoma skin cancer	5.1	2.7	2.7
Transformation to MF	2.1	1.8	1.4
Transformation to AML	0.2	0.6	0.0

Abbreviations: AML: acute myeloid leukaemia; MF: myelofibrosis.
Source: Kiladjian et al. 2018;[91] Kiladjian et al. 2020.[65]

Deaths

In the ruxolitinib group, there were two on-treatment deaths by the time of the Week 256 analysis; one death due to a gastric adenocarcinoma (investigator suspected event to be related to the study drug) and the other due to a malignant neoplasm (not suspected to be related to the study drug). In the crossover group, there were four on-treatment deaths, unrelated to the study drug. There were no on-treatment deaths amongst patients receiving BAT.[65]

B.2.10.2 RESPONSE-2

Week 260 analysis

Due to the high degree of patient crossover from BAT to ruxolitinib (77.3% at Week 260), long-term results are predominantly reported for the ruxolitinib and crossover groups only.[66]

AEs

At Week 260, the overall safety profile for ruxolitinib remained consistent with previous data cut-offs, with no new long-term safety signals identified.[66] Almost all patients in the ruxolitinib (n=74/74 [100.0%]) and crossover groups (n=57/58 [98.3%]) had experienced at least one AE, while 64/75 patients (85.3%) in the BAT arm experienced AEs prior to crossover at Week 28.[60, 66] The exposure adjusted rate of any grade AE per 100 person-years was 22 for ruxolitinib group, 120 for the BAT group and 28 for the crossover group.[66]

As presented in Table 17, in the ruxolitinib group, the most frequent AEs of any grade (frequency [exposure-adjusted rate per 100 person-years]) were anaemia (27 [8.1]), arthralgia (20 [6.0]) and weight gain (19 [5.7]).[66] In the BAT group, the most frequent AEs of any grade (exposure-adjusted rate per 100 person-years) were pruritus (17 [31.9]), headache (9 [16.9]) and thrombocytopenia (8 [15.0]). In the crossover group, the most frequent AEs of any grade (exposure-adjusted rate per 100 person-years) were anaemia (19 [9.2]) and hypertension (11 [5.3]).

The most common Grade 3–4 AEs for the ruxolitinib group (exposure-adjusted rate per 100 person-years) was hypertension (8 [2.4]). No other Grade 3–4 AEs were reported in >2 patients. In the BAT group, the most common Grade 3–4 AEs were thrombocytosis (4 [7.5]), thrombocytopenia (3 [5.6]) and hypertension (3 [5.6]). In the crossover group, the most common Grade 3–4 AEs were hypertension (6 [2.9]), anaemia (2 [1.0]) and thrombocytosis (2 [1.0]).[66]

A summary of the most frequent (occurring in ≥3% of patients in any arm) on-treatment AEs adjusted for patient-year exposure is provided in Table 17.

Table 17: Most frequent (occurring in ≥3% of patients in any arm) on-treatment AEs adjusted for patient-year exposure

AEs, n (exposure adjusted rate per 100 person-years)	Ruxolitinib (n=74)		BAT (n=75)[a]		Crossover (n=58)[b]	
	Any grade	Grade 3–4	Any grade	Grade 3–4	Any grade	Grade 3–4
Any	74 (22.1)	50 (15.0)	64 (120.0)	22 (41.2)	57 (27.7)	33 (16.0)
Anaemia	27 (8.1)	0	2 (3.7)	1 (1.9)	19 (9.2)	2 (1.0)
Arthralgia	20 (6.0)	2 (0.6)	3 (5.6)	1 (1.9)	6 (2.9)	0
Weight increased	19 (5.7)	2 (0.6)	1 (1.9)	0	9 (4.4)	0
Hypertension	15 (4.5)	8 (2.4)	3 (5.6)	3 (5.6)	11 (5.3)	6 (2.9)
Headache	14 (4.2)	0	9 (16.9)	0	8 (3.9)	0
Fatigue	13 (3.9)	1 (0.3)	6 (11.2)	0	6 (2.9)	1 (0.5)
Constipation	13 (3.9)	0	4 (7.5)	0	8 (3.9)	1 (0.5)
Bronchitis	13 (3.9)	2 (0.6)	2 (3.7)	0	2 (1.0)	0
Pyrexia	13 (3.9)	0	1 (1.9)	0	8 (3.9)	0

AEs, n (exposure adjusted rate per 100 person-years)	Ruxolitinib (n=74)		BAT (n=75)[a]		Crossover (n=58)[b]	
	Any grade	Grade 3–4	Any grade	Grade 3–4	Any grade	Grade 3–4
Pruritus	12 (3.6)	0	17 (31.9)	2 (3.7)	7 (3.4)	0
Pain in extremity	12 (3.6)	1 (0.3)	2 (3.7)	1 (1.9)	4 (1.9)	0
Back pain	12 (3.6)	0	0	0	7 (3.4)	0
Dyspnoea	11 (3.3)	0	2 (3.7)	0	4 (1.9)	0
Abdominal pain	11 (3.3)	0	1 (1.9)	0	9 (4.4)	1 (0.5)
Herpes zoster	11 (3.3)	2 (0.6)	0	0	8 (3.9)	0
Influenza	10 (3.0)	1 (0.3)	5 (9.4)	1. (1.9)	2 (1.0)	0
Oedema peripheral	10 (3.0)	0	2 (3.7)	0	6 (2.9)	0
Haematoma	10 (3.0)	0	1 (1.9)	0	4 (1.9)	0
Cystitis	10 (3.0)	0	0	0	2 (1.0)	1 (0.5)
Asthenia	8 (2.4)	1 (0.3)	6 (11.2)	1 (1.9)	7 (3.4)	0
Thrombocytosis	8 (2.4)	0	4 (7.5)	4 (7.5)	5 (2.4)	2 (1.0)
Dizziness	8 (2.4)	0	5 (9.4)	0	7 (3.4)	0
Nasopharyngitis	8 (2.4)	0	2 (3.7)	0	10 (4.9)	0
Diarrhoea	7 (2.1)	0	7 (13.1)	0	4 (1.9)	0
Cough	7 (2.1)	1 (0.3)	2 (3.7)	1 (1.9)	6 (2.9)	0
Night sweats	6 (1.8)	0	5 (9.4)	0	1 (0.5)	0
Thrombocytopenia	5 (1.5)	1 (0.3)	8 (15.0)	3 (5.6)	3 (1.5)	0
Upper respiratory tract infection	5 (1.5)	0	7 (13.1)	0	5 (2.4)	0
Leucocytosis	5 (1.5)	2 (0.6)	4 (7.5)	1 (1.9)	3 (1.5)	1 (0.5)
Decreased appetite	5 (1.5)	0	4 (7.5)	0	1 (0.5)	0
Abdominal pain upper	5 (1.5)	0	3 (5.6)	0	2 (1.0)	0
Myalgia	5 (1.5)	0	2 (3.7)	0	2 (1.0)	0
Dyspepsia	5 (1.5)	0	2 (3.7)	0	4 (1.9)	0
Nausea	4 (1.2)	0	5 (9.4)	0	4 (1.9)	0
Epistaxis	4 (1.2)	0	2 (3.7)	0	7 (3.4)	0
Oropharyngeal pain	3 (0.9)	0	3 (5.6)	0	0	0
Tinnitus	3 (0.9)	0	2 (3.7)	1 (1.9)	3 (1.5)	0
Toothache	3 (0.9)	0	2 (3.7)	0	2 (1.0)	0
Insomnia	3 (0.9)	0	2 (3.7)	0	2 (1.0)	0
Hyperhidrosis	3 (0.9)	0	2 (3.7)	0	0	0
Erythema	2 (0.6)	0	4 (7.5)	1 (1.9)	0	0
Weight decreased	1 (0.3)	0	4 (7.5)	0	1 (0.5)	0
Aphthous ulcer	1 (0.3)	0	3 (5.6)	1 (1.9)	2 (1.0)	0
Aquagenic pruritus	1 (0.3)	0	3 (5.6)	0	1 (0.5)	0
Abdominal discomfort	1 (0.3)	0	2 (3.7)	0	4 (1.9)	0
Iron deficiency	1 (0.3)	0	2 (3.7)	0	1 (0.5)	0
Rash	1 (0.3)	0	2 (3.7)	0	1 (0.5)	0

AEs, n (exposure adjusted rate per 100 person-years)	Ruxolitinib (n=74)		BAT (n=75)[a]		Crossover (n=58)[b]	
	Any grade	Grade 3–4	Any grade	Grade 3–4	Any grade	Grade 3–4
Palpitations	1 (0.3)	0	2 (3.7)	0	1 (0.5)	0
Early satiety	1 (0.3)	0	2 (3.7)	0	0	0
HCT increased	0	0	5 (9.4)	1 (1.9)	1 (0.5)	0
Mouth ulceration	0	0	3 (5.6)	1 (1.9)	2 (1.0)	0
Cellulitis	0	0	2 (3.7)	1 (1.9)	1 (0.5)	0
Splenomegaly	0	0	2 (3.7)	0	1 (0.5)	0
Skin lesion	0	0	2 (3.7)	0	1 (0.5)	0
Increased blood thyroid stimulating hormone	0	0	2 (3.7)	0	0	0
Stomatitis	0	0	2 (3.7)	0	0	0
Urticaria	0	0	2 (3.7)	0	0	0

A patient with multiple occurrences of an AE was counted only once in AE category. [a]Includes patients randomly assigned to the best available therapy group who crossed over to the ruxolitinib group, but only includes events that occurred before crossover. [b]Only includes events among patients in the best available therapy group that occurred after crossover from the best available therapy group to the ruxolitinib group.
Abbreviations: AE: adverse event; BAT: best available therapy; HCT: haematocrit.
Source: Passamonti et al. 2022.[66]

Exposure-adjusted rates of selected AEs are reported in Table 18. Thromboembolic events (exposure-adjusted rate per 100 patient-years) were 5 (1.5) in the ruxolitinib group, 2 (3.7) in the BAT group and 6 (2.9) in the crossover group.[66] Frequency of transformation to MF (exposure-adjusted rate per 100 patient-years) were 2 (0.6) for the ruxolitinib group, 1 (1.9) for the BAT group, and 1 (0.5) for the crossover group.[66] Similarly, transformation to AML (frequency [exposure-adjusted rate per 100 patient-years]) was more frequent in the BAT group (1 [1.9]) than the ruxolitinib group (0, [0]).[98]

Table 18: Exposure-adjusted AEs of interest at Week 260 in the RESPONSE-2 trial

AEs, n (exposure adjusted rate per 100 person-years)	Ruxolitinib	Crossover	BAT
Thromboembolic event	5 (1.5)	6 (2.9)	2 (3.7)
Non-melanoma skin cancer	9 (2.7)	6 (2.9)	1 (1.9)
Transformation to MF	2 (0.6)	1 (0.5)	1 (1.9)
Transformation to AML	0	0	1 (1.9)

Abbreviations: AML: acute myeloid leukaemia; BAT: Best Available Therapy; MF: myelofibrosis; NR: not reported.
Source: Passamonti et al., 2022.[66]

Death

No treatment-related deaths were reported during the study.[66]

B.2.11 *MAJIC-PV*

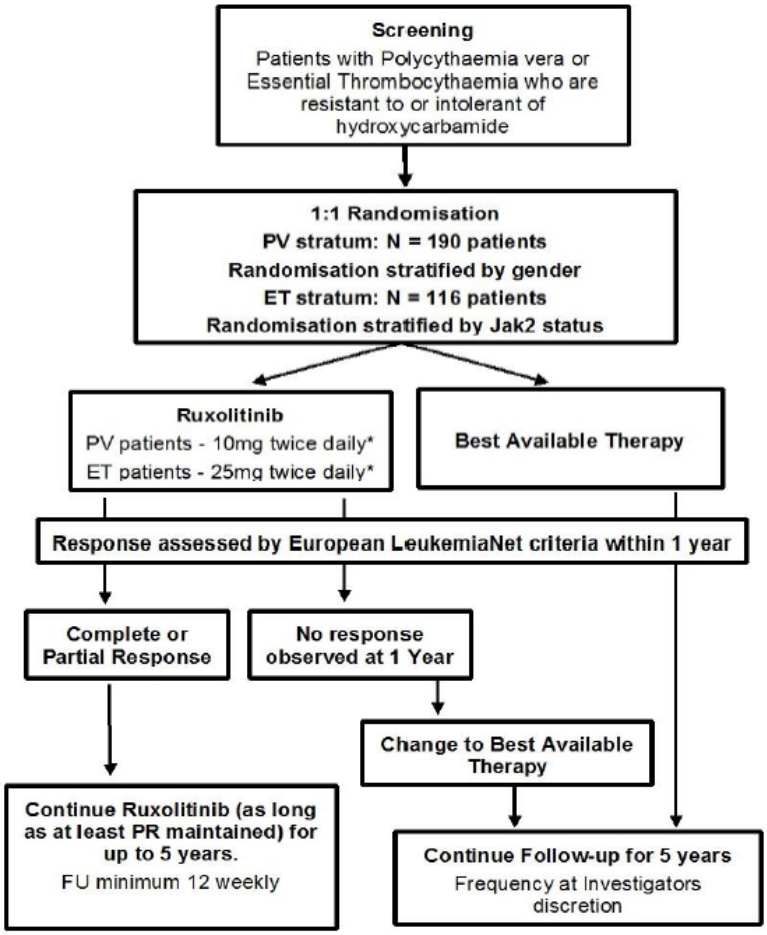
Data on MAJIC-PV has been sourced from an unpublished manuscript provided in confidence by Harrison et al. for use in this submission. This publication has been submitted to the New England Journal of Medicine and is undergoing peer review at the time of this submission. For further details on the MAJIC trial, please refer to this publication provided in the reference pack.

B.2.11.1 Methodology

MAJIC-PV (ISRCTN61925716) was an open-label, randomised controlled trial of ruxolitinib versus BAT conducted at 38 sites in the UK. Patients aged ≥18 with high-risk PV meeting criteria for R/I to HC/HU were recruited over 48 months (2012–2016). In the MAJIC-PV trial the definition of high-risk PV is broader than that provided in the BSH guidelines (see Section B.1.3.2), as it includes patients who have significant or symptomatic splenomegaly, a platelet count >1000 ×109/L, and diabetes or hypertension requiring pharmacological therapy for >6 months in addition to criteria aligned to the BSH guidelines of previous documented thrombosis (secondary to PV or within 10 years of diagnosis) and age of ≥60 years old.[1, 70] Therefore, the MAJIC-PV trial population is anticipated to represent the majority of patients with PV who are R/I to HC/HU.

Patients in the MAJIC-PV trial were stratified by gender and randomised 1:1 to either ruxolitinib (starting 10 mg twice daily or 5 mg twice daily for baseline platelets 100-200 × 109/L) or BAT (Figure 27). Patients were then followed-up for 5 years, with crossover prohibited within the trial as per the protocol.[70] The primary outcome was achievement of CHR (according to ELN guidelines) within one year. Secondary outcomes included partial response (PR), EFS, progression-free survival (PFS), OS, safety, as well as symptom and QoL assessment. Further details on trial design, eligibility criteria, baseline characteristics and patient disposition are provided in Appendix D.[70]

Figure 27: Trial design of MAJIC-PV



*Patients with platelet count between 100 and 200 $\times 10^9$/L were started on a reduced dose
Abbreviations: ET: essential thrombocythaemia; FU: follow-up; PR: partial response; PV: polycythaemia vera.
Source: Harrison et al.[70]

B.2.11.2 Clinical effectiveness results

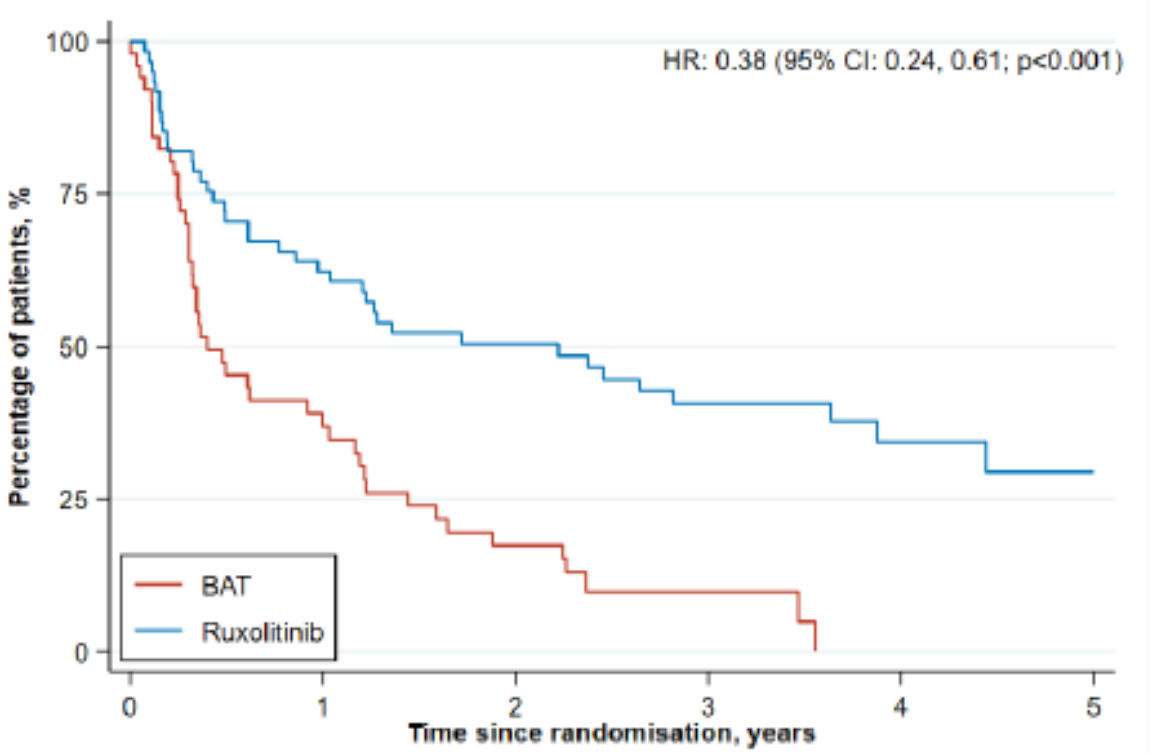
Data from the MAJIC-PV trial are primarily presented from unpublished 5-year data by Harrison et al. (this publication has been submitted to the New England Journal of Medicine and is undergoing peer review at the time of this submission).[70] The median treatment duration on ruxolitinib was 1,568 days and for BAT, it was 1,220 days. The most frequent BATs were HC/HU monotherapy (32%), IFN-alfa monotherapy (15%), and HC/HU/IFN-alfa combination therapy (12%).

Primary endpoint: CHR

CHR was defined according to ELN criteria: HCT <45% without venesection for 3 months; platelets ≤400 × 109/L; WBC count ≤10 × 109/L, and normal spleen size.[70]

CHR was achieved in 40 (43%) patients in the ruxolitinib arm versus 23 (26%) in the BAT arm (gender-adjusted OR 2.12; 90% CI: 1.25, 3.60, p=0.02). Patients receiving ruxolitinib also observed a more durable CHR than those receiving BAT (HR: 0.38; 95% CI: 0.24, 0.61, p<0.001) as shown in Figure 28.

Figure 28: Durability of CHR with ruxolitinib compared to BAT in the MAJIC-PV trial



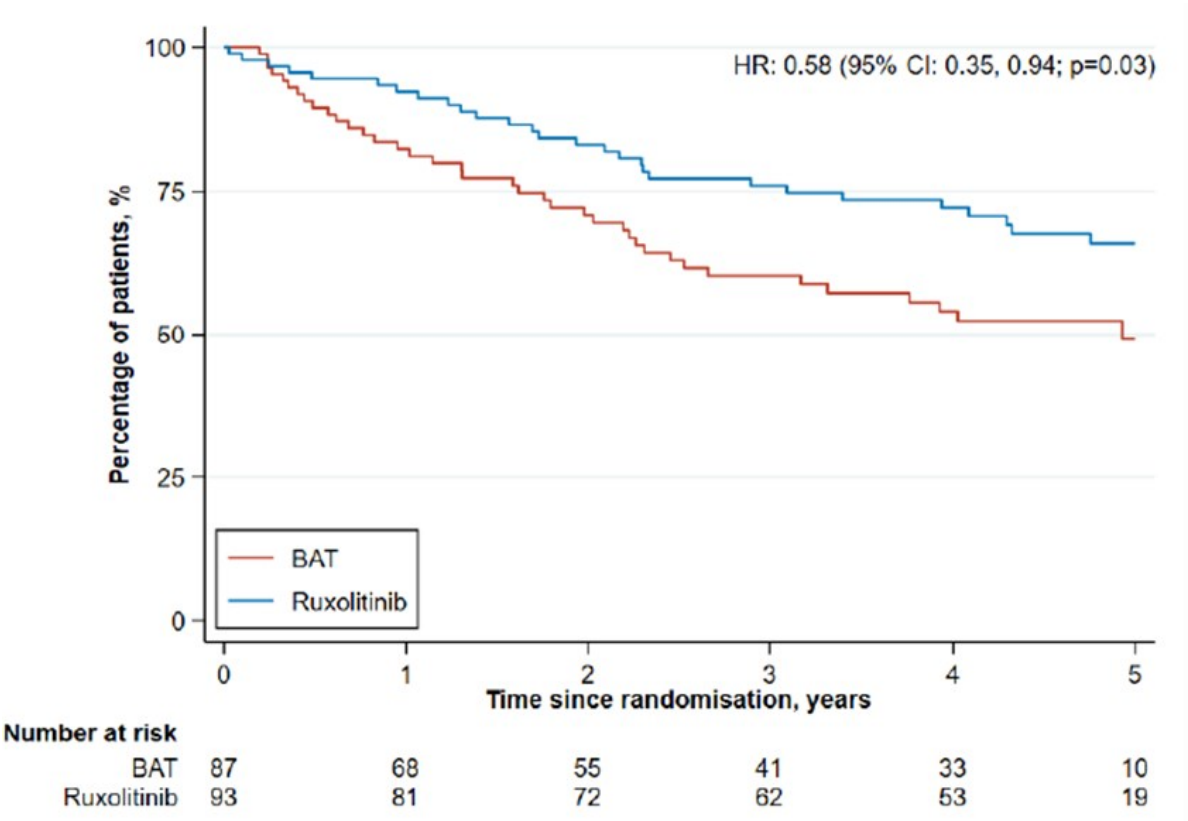
Abbreviations: BAT: Best Available Therapy; CHR: complete haematological remission; CI: confidence interval; HR: hazard ratio.
Source: Harrison et al.[70]

PR was achieved in 50 (54%) patients in the ruxolitinib arm and 58 (67%) in the BAT arm during Year 1. Consequently, overall response rates were 97% and 93% for ruxolitinib and BAT treated patients, respectively.[70]

EFS

EFS was defined as a composite of major thrombosis, major haemorrhage, transformation or death. EFS was found to be superior for patients in the ruxolitinib arm compared with those in the BAT arm (HR 0.58; 95% CI: 0.35, 0.94, p=0.03), as shown in Figure 29.[70]

Figure 29: EFS, stratified by treatment arm

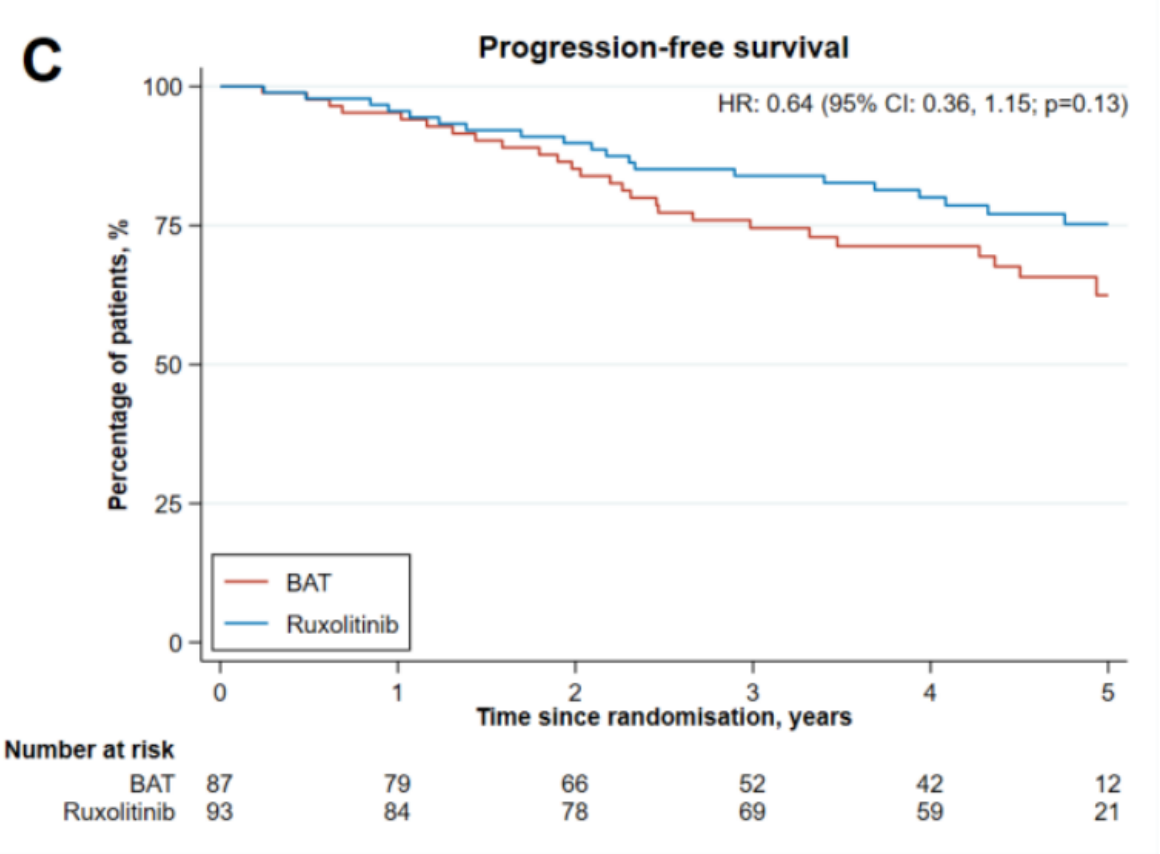


Abbreviations: BAT: Best Available Therapy; CI: confidence interval; EFS: event-free survival; HR: hazard ratio.
Source: Harrison et al.[70]

PFS

Patients receiving ruxolitinib had a 3-year PFS of 84% (95% CI: 74, 90). In comparison, 3-year PFS for BAT was found to be 75% (95% CI: 63, 83), demonstrating a trend for improved PFS with ruxolitinib. The trend for improved PFS with ruxolitinib compared to BAT was maintained at 5-years; however, this difference was not significant (HR: 0.64; 95% CI: 0.36, 1.15; p=0.13; Figure 30).[70]

Figure 30: PFS, stratified by treatment arm

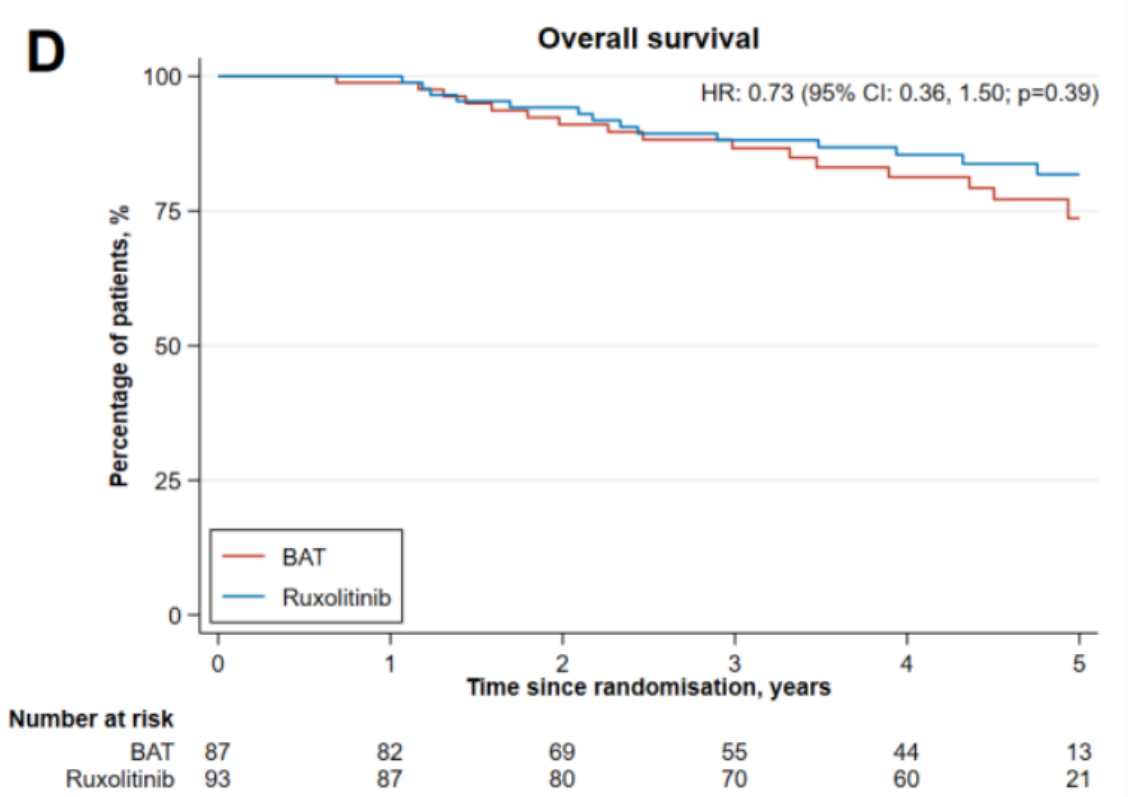


Abbreviations: BAT: Best Available Therapy; CI: confidence interval; HR: hazard ratio; PFS: progression-free survival.
Source: Harrison et al.[70]

OS

There was no difference between 3-year OS in the BAT arm (87%, 95% CI: 77, 93) and the ruxolitinib arm (88%, 95% CI: 79, 93). At 5-years, there was a numerically lower risk of death following ruxolitinib treatment compared to BAT; however, this was not statistically significant (HR: 0.73; 95% CI: 0.36, 1.50, p=0.39; Figure 31).[70]

Figure 31: OS, stratified by treatment arm



Abbreviations: BAT: Best Available Therapy; CI: confidence interval; HR: hazard ratio; OS: overall survival.
Source: Harrison et al.[70]

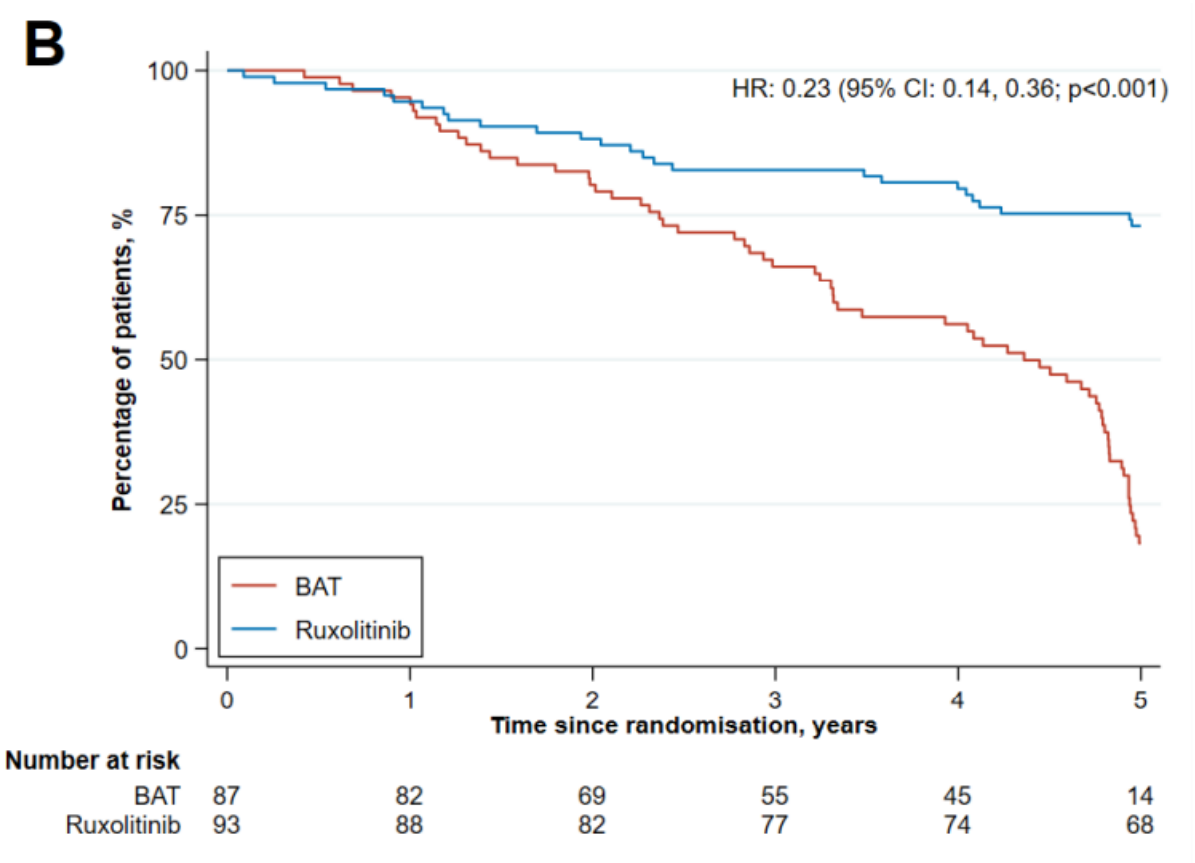
Rate of phlebotomy procedures

Overall, ruxolitinib patients required fewer phlebotomies than the BAT arm ([redacted] versus 307, respectively). Fewer patients in the ruxolitinib arm (29%) had at least one venesection compared with the BAT arm (52%).[70]

Time to treatment discontinuation of first treatment

Patients in the BAT arm were significantly more likely to discontinue first treatment compared to ruxolitinib treated patients (HR: 0.23; 95% CI: 0.14, 0.36; p<0.001), as shown in Figure 32.[70]

Figure 32: Time to discontinuation of first treatment, stratified by treatment arm

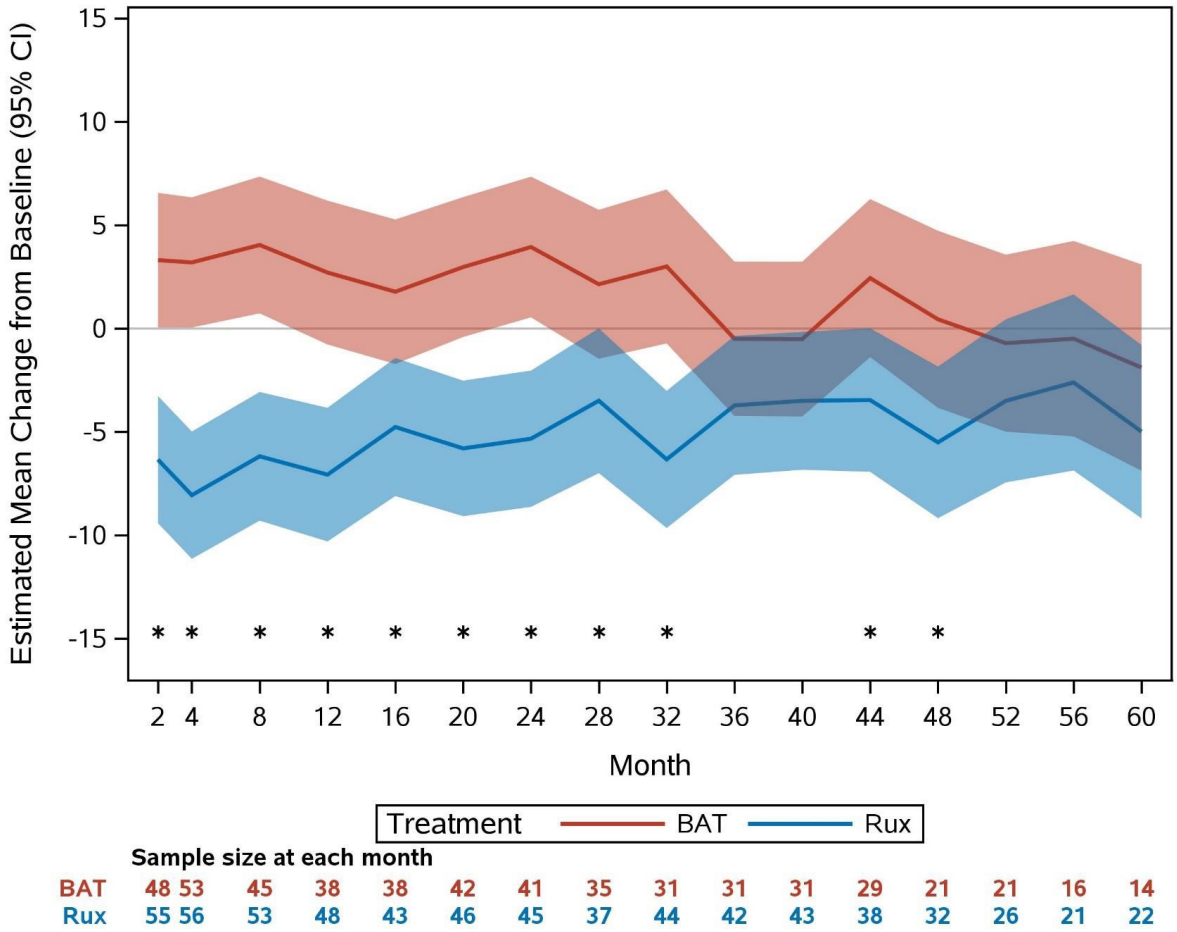


Symptom burden

147 patients completed the baseline symptom assessment (BAT: n=71, ruxolitinib: n=76) and symptom scores were overall similar between treatment arms.[70]

Durable improvements in Myeloproliferative Neoplasm total symptom score (MPN-TSS) were noted for ruxolitinib patients, lasting a mean of 52 months (Figure 33). In contrast, patients receiving BAT experienced a worsening of symptom burden at 56 months from baseline. Of the 115 patients with MPN-SAF TSS scores at baseline and at least one additional timepoint, 17/56 (30%) BAT versus 36/59 (61%) ruxolitinib patients had a TSS reduction of ≥50% at least one time point (p=0.001), indicating that a greater proportion of patients in the ruxolitinib arm experienced an improvement in their symptoms from baseline compared with those in the BAT arm. Moreover, TSS scores for fatigue, early satiety, night-sweats, itching, bone pain and weight loss were observed to significantly reduce for the ruxolitinib arm compared with the BAT arm at over five time points. The overall change in MPN-TSS over 5-years for both treatment arms is shown in Figure 33.[70]

Figure 33: Change in MPN-TSS over 5-years in the MAJIC-PV trial



Abbreviations: BAT: Best Available Therapy; CI: confidence interval; MPN-TSS: Myeloproliferative Neoplasm total symptom score; Rux: ruxolitinib.
Source: Harrison et al.[70]

B.2.11.3 Adverse reactions

Patterns of AEs with ruxolitinib were similar to those previously reported, with no new events emerging with longer follow-up. Grade 3 anaemia occurred in 7% of ruxolitinib patients versus 1% in the BAT arm. Similarly, infections were also more common for ruxolitinib-treated patients (27 versus 12 [BAT] Grade 3/4 events). No Grade 5 AEs were observed in the ruxolitinib arm. However, malignancy and squamous cell skin cancer was reported more commonly in ruxolitinib treated patients (11 versus 0 [BAT] events). A summary of AEs occurring in ≥10% of patients by grade and treatment arm is presented in Appendix F.[70]

Deaths

Overall, there were 32 deaths during the MAJIC-PV trial, 17 were in the BAT group and 15 were in the ruxolitinib group. Of these events, one treatment-related death occurred in the BAT group and one occurred in the ruxolitinib group. Other causes of death included non-cancer related reasons, cancer related-reasons and thrombosis or haemorrhage-related reasons.[70]

B.2.11.4 Conclusions

In MAJIC-PV, ruxolitinib was associated with improved treatment efficacy in terms of CHR, EFS and symptom responses in high-risk patients with PV who are R/I to HC/HU. There was also a trend towards improved OS with ruxolitinib, with the curve starting to diverge after 3.0 years. The analysis of OS was descriptive only and therefore, the study may not have been powered to detect a statistical significance. Furthermore, clinical experts suggested that the absence of (statistical) difference in survival was likely due to the low number of patients and short follow-up period.[11] The patterns of AEs with ruxolitinib were also similar to previously reported, with no new events emerging with longer follow-up. Therefore, these data provide evidence to support the efficacy and safety of ruxolitinib demonstrated in the RESPONSE trials (see Section B.2.7) in a subgroup (high-risk) of patients with PV who are R/I to HC/HU.[70]

B.2.12 *Ongoing studies*

RESPONSE, RESPONSE-2 and MAJIC-PV are all completed and the 5-year data have been analysed or published.[65, 66, 70] There are no ongoing studies of relevance to this appraisal.

B.2.13 *Interpretation of clinical effectiveness and safety evidence*

B.2.13.1 Principle findings from the clinical evidence base

Ruxolitinib offers improved and durable responses, in terms of HCT control and platelet and WBC counts, in patients with PV who are R/I to HC/HU

The efficacy and safety evidence base to support the use of ruxolitinib versus BAT in patients with PV who are R/I to HC/HU primarily comes from two RCTs: the RESPONSE trial for patients with splenomegaly, and RESPONSE-2 for patients without splenomegaly.[60, 67] Crossover from BAT to ruxolitinib was permitted in both trials (after Week 32 in RESPONSE and Week 28 in RESPONSE-2) and by Week 80 all patients in the BAT arms had discontinued BAT.[65, 66] Results from both trials show that ruxolitinib offers improved and durable responses in terms of HCT control, platelet and WBC counts versus BAT for patients R/I to HC/HU. In turn, this leads to reduced reliance on phlebotomy versus current treatments and substantial benefits in terms of PV symptoms and HRQoL.[60, 65-67]

In RESPONSE, the primary endpoint (both HCT control and ≥35% reduction in spleen volume) was reached by a significantly higher proportion of patients in the ruxolitinib group versus the BAT group at Week 32 (25/110 [22.7%] versus 1/112 [0.9%]; p<0.001).[67] CHR, which considers platelet and WBC counts as well as HCT control, was achieved by a significantly higher proportion of patients in the ruxolitinib group compared with the BAT group (23.6% versus 8.9%, respectively; p=0.003).[67]

In RESPONSE-2, the primary endpoint of HCT control at Week 28 was reached by a significantly higher proportion of patients in the ruxolitinib group (42/74 [62%]) compared to the BAT group (14/75 [19%]) (OR ruxolitinib versus BAT: 7.28; 95% CI: 3.43, 15.45; p<0.0001).[60] The proportion of patients achieving CHR was also significantly higher in the ruxolitinib group compared with the BAT group (23% versus 5%, respectively; OR ruxolitinib versus BAT: 5.58; 95% CI: 1.73, 17.99; p=0.0019).[60] Long-term efficacy results for patients receiving ruxolitinib have been reported after

a 5-year follow-up (256 weeks) for RESPONSE and 260 weeks for RESPONSE-2. These results provide further evidence to support the primary analysis.[65, 66]

Ruxolitinib may reduce the risk of complications such as thromboembolic events and progression to MF or AML in patients with PV who are R/I to HC/HU

Ruxolitinib demonstrated a decreased risk of complications in patients with PV who are R/I to HC/HU compared with those who received BAT. At the 5-year follow-up in both trials, the exposure-adjusted rate of thromboembolic events was higher in the BAT group compared to the ruxolitinib group.[65, 66] For transformation to MF or AML, patients receiving ruxolitinib in RESPONSE experienced similar exposure-adjusted rates as those who received BAT at the 5-year follow-up.[65] However, at the 5-year follow-up in RESPONSE-2, patients receiving ruxolitinib had lower exposure-adjusted rates of transformation to MF and AML than the patients receiving BAT.[66]

Ruxolitinib is associated with a tolerable safety profile that is supported by long-term data

Safety was evaluated as a secondary outcome in the RESPONSE and RESPONSE-2 trials.[60, 67] The safety profile of ruxolitinib was consistent across both trials and previous reports with ruxolitinib in other indications and AEs were generally manageable with standard clinical monitoring and care:[69] anaemia was the most common haematological AE experienced by patients treated with ruxolitinib, but was rarely Grade 3 or 4 in severity, and few non-haematological AEs were seen in either treatment group prior to crossover.[60, 67]

The efficacy and safety of ruxolitinib is supported by data from the MAJIC-PV trial in high-risk patients with PV who are R/I to HC/HU

MAJIC-PV (ISRCTN61925716) was an open-label, randomised controlled trial of ruxolitinib versus BAT in high-risk patients with PV who are R/I to HC/HU, in which crossover between treatments was not allowed as per the protocol.[70] In MAJIC-PV, ruxolitinib was associated with improved treatment efficacy versus BAT in terms of CHR (HR: 0.38; 95% CI: 0.24, 0.61, p<0.001), EFS (HR 0.58; 95% CI: 0.35, 0.94, p=0.03) and symptom responses (as measured by MPN-SAF TSS) in high-risk patients with PV who are R/I to HC/HU.[70] The patterns of AEs with ruxolitinib were similar to previously reported, with no new events emerging with longer follow-up.[70] Therefore, these data provide evidence to support the efficacy and safety of ruxolitinib demonstrated in the RESPONSE trials in a subgroup (high-risk) of patients with PV who are R/I to HC/HU.[70]

Both direct and indirect evidence supports the hypothesis that ruxolitinib is associated with a survival gain

In RESPONSE, the Kaplan–Meier estimates for OS at five years were 91.9% (95% CI: 84.4, 95.9) in the ruxolitinib group and 91.0% (95% CI: 82.8, 95.4) in the BAT group.[65] Similarly, in RESPONSE-2 the Kaplan–Meier estimates for OS at five years were 96% (95% CI: 87, 99) in the ruxolitinib group and 91% (95% CI: 80, 96) in the BAT group.[66] Despite the high degree of patient crossover following the primary treatment period in RESPONSE and RESPONSE-2, there was a trend for improved OS with BAT compared with ruxolitinib. However, given that no patients were receiving BAT in either trial from the time of the Week 80 analyses, these results remain uncertain. Therefore, a MAIC was conducted to estimate the relative efficacy of ruxolitinib versus BAT in terms of OS. Although there were limitations associated with the MAIC, the observed improvement in OS for ruxolitinib compared to BAT, supports the hypothesis that ruxolitinib is associated with a survival advantage.[12, 60, 67] This hypothesis is further supported by the MAJIC-PV trial in which

crossover was not permitted; a trend towards improved OS with ruxolitinib compared to BAT was observed at 5-years, with the curve starting to diverge after 3.0 years.[70]

B.2.13.2 Strengths and limitations of the evidence base

Internal validity

The primary clinical evidence base for ruxolitinib in patients with PV who are R/I to HC/HU comes from two well designed and conducted Phase 3 clinical trials: RESPONSE and RESPONSE-2. Data from the RESPONSE trials are also supported by the Phase 2, randomised controlled MAJIC-PV trial in high-risk patients with PV who are R/I to HC/HU.[70] Together, these studies provide high quality, randomised, controlled, long-term evidence for the efficacy and safety of ruxolitinib.

The internal validity of RESPONSE, RESPONSE-2 and MAJIC-PV is supported by several factors, including the use of an ITT analysis that was applied for primary and key secondary endpoints for data from all patients that underwent randomisation and generally similar baseline characteristics between treatment groups in all studies.[60, 65-67, 70] Furthermore, crossover of BAT patients to ruxolitinib was not permitted in MAJIC-PV.[70]

The various treatments and methods of administration for the BAT groups of both trials meant that an open-label study design was necessary for RESPONSE, RESPONSE-2 and MAJIC-PV. This study design (where investigators and patients were not blinded to treatment assignment) may introduce bias to which subjective measures could be particularly susceptible to (e.g. PROs). However, PROs in patients with MF were similar in a double-blind and an open-label study (COMFORT-I and COMFORT-II, respectively), which suggested that the open-label design in a related indication did not cause significant bias.[60] Those responsible for data review, analysis and interpretation, as well as those involved in measuring spleen images, were blinded to treatment assignment to reduce the risk of bias.

External validity

The results of the RESPONSE trials and MAJIC-PV can be generalised to the UK population, given the similarity of the BAT administered compared with UK clinical practice and the European centres included (with UK centres specifically for RESPONSE and MAJIC-PV exclusively based in the UK).[1, 60, 66, 67, 70]

Furthermore, according to expert UK clinical opinion, the population included in the RESPONSE trials were generally reflective of the UK population. However, patients included in the RESPONSE trials were likely to have better performance status than observed in UK clinical practice and were required to have at least two phlebotomies in the last 24 weeks, which is not fully representative of the UK population with R/I to HC/HU as some patients may not require phlebotomy.[11]

The results are well aligned with the decision problem specified in the NICE scope as follows:

- **Population:** The study population of the RESPONSE and RESPONSE-2 trials together covered both patients with and without splenomegaly, and therefore, were representative of the full population of patients with PV who are R/I to HC/HU considered in this submission.[60, 67] The MAJIC-PV trial provided additional supporting efficacy and safety data in the sub-group of high-risk patients (see Section B.2.11).[70] As validated by clinical expert opinion,[11] the RESPONSE and RESPONSE-2 study populations are also relevant to the epidemiology of PV in the UK, and included patients from European sites (with UK

centres specifically for RESPONSE).[60, 67] MAJIC-PV was exclusively based in the UK and therefore, also provided additional data on a highly relevant population.[70]

- **Intervention:** Ruxolitinib was directly evaluated as a treatment option for patients with PV who are R/I to HC/HU by comparing to BAT.[60, 67]
- **Comparators:** In the absence of an established treatment for patients with PV who are R/I to HC/HU, the comparator in both the RESPONSE and RESPONSE-2 trials was BAT, as selected by the treating physician.[60, 67] Concomitant aspirin and phlebotomy were also permitted in the trials, in line with UK clinical practice.[60, 67] Clinical experts noted that as RESPONSE and RESPONSE-2 were international trials, there were some differences in the individual therapies received as part of the BAT arms compared with UK clinical practice. For example, lenalidomide and pipobroman are not used in UK clinical practice and the proportion of patients who received no cytoreductive therapy may be higher than what is typically seen in clinical practice.[11] However, MAJIC-PV was exclusively based in the UK, and clinical experts indicated that the composition of BAT in this trial was more reflective of UK practice compared with the RESPONSE-trials.[11]
- **Outcomes:** The goal of treatment of patients with PV is to alleviate symptoms and thus improve QoL, minimise the risk of complications and to, ultimately, improve survival.[1] Thromboembolic events are a major risk factor for mortality and so treatments aim to reduce this risk, alongside reducing symptom burden.[1] The primary outcome of RESPONSE and RESPONSE-2 was CHR.[60, 67] This endpoint can be considered relevant to UK treatment goals as it has been associated with a reduced risk of thromboembolic events and death.[12] Moreover, other outcomes in line with UK treatment goals (including all outcomes outlined in the scope) such as OS, progression to MF or AML, symptom relief, thrombosis and HRQoL were also evaluated in both trials. Additionally, UK clinical experts have indicated that the results from the RESPONSE trials in terms of durable symptom control generally mirror the outcomes of patients that switch to ruxolitinib in UK clinical practice.[11] Therefore, the results of RESPONSE and RESPONSE-2 trials can be considered relevant to UK patients and clinicians.

Limitations

Patients receiving HC/HU in the RESPONSE and MAJIC-PV trials

In both RESPONSE trials and the MAJIC-PV trial, HC/HU was permitted in the BAT arm (only at a tolerable dose if, in the investigator's opinion, it was likely to provide clinical benefit).[60, 67, 70] Despite patients being R/I of HC/HU, a high proportion of patients in RESPONSE (58.9%), RESPONSE-2 (49%) and MAJIC-PV (32%) received HC/HU as BAT – more than any other permitted therapy for the BAT arm.[60, 67, 70] Following BAT, 'no treatment' was received by the second highest proportion of patients in the RESPONSE trials.[60, 67] Clinical experts noted that the proportion of patients receiving 'no treatment' in the RESPONSE trials may be higher than that typically seen in UK clinical practice. However, this may be due to phlebotomy being considered as a non-cytoreductive therapy in the RESPONSE trials.[11]

The relatively high use of HC/HU or no treatment (i.e. phlebotomy and low-dose aspirin only) in the RESPONSE trials and MAJIC-PV, in which treatment in the BAT arm was based on the investigators' discretion, highlights the unmet need for effective treatments other than HC/HU for patients with PV who are R/I to HC/HU.

Crossover from BAT to ruxolitinib during the RESPONSE trials

A limitation of the RESPONSE and RESPONSE-2 trials was that crossover was permitted after Week 32 and Week 28, respectively. In both trials, crossover was permitted from the BAT group to the ruxolitinib group if the primary endpoint was not met or at later timepoints for disease progression, a lack of treatment effectiveness or for safety reasons.[60, 67]

In RESPONSE and RESPONSE-2, 88% and 77% of patients crossed over to the ruxolitinib group, respectively.[65, 66] The high degree of crossover therefore precludes the comparison of long-term outcomes such as OS due to the fact that the vast majority of patients did not receive BAT past the primary analyses, thus confounding the results.

To address the challenges in interpreting OS data of ruxolitinib compared with BAT, a MAIC for OS was conducted using data for BAT from the GEMFIN registry which was not confounded by crossover. The MAIC provided further evidence to support the efficacy benefits of ruxolitinib in patients with PV who R/I to HC/HU.

Likewise, there was a trend for improved OS with ruxolitinib compared to BAT in the MAJIC trial, in which crossover between treatments was not allowed, which supports the hypothesis that ruxolitinib improves OS.[70]

A discussion of the strengths and limitations of the MAIC is provided in Appendix D.

B.2.13.3 Conclusions

The RESPONSE and RESPONSE-2 trials are well designed and provide high quality evidence to support the efficacy and safety of patients with PV who are R/I to HC/HU.

In these trials, ruxolitinib was shown to offer improved and durable responses in terms of HCT control, platelet and WBC counts versus BAT for patients R/I to HC/HU, as well as substantial benefits in terms of PV symptoms and HRQoL.[60, 67] Observations from the primary analysis were further supported by long-term efficacy and safety data.[65, 66]

Data from the Phase 2, randomised controlled MAJIC-PV trial also provided supporting long-term evidence on the safety and efficacy of ruxolitinib in high-risk patients with PV who are R/I to HC/HU from a UK-based population.[70] These data were not confounded by crossover (as this was not permitted) and demonstrated that ruxolitinib was associated with improved treatment efficacy compared to BAT in terms of CHR, EFS and symptom responses.[70] There was also a trend towards improved OS with ruxolitinib, with the curve starting to diverge after 3.0 years.[70]

Results from RESPONSE, RESPONSE-2 and MAJIC-PV were further supplemented by the MAIC analysis, which provided estimates for OS that were not confounded by crossover and demonstrated an improvement in OS for ruxolitinib versus BAT.[68]

Patients with PV experience a debilitating symptom burden as well as being at high-risk of disease progression, thromboembolic events and death. However, there are very limited effective treatment options for these patients and a clear unmet need for a targeted, effective treatment which helps to control HCT, ultimately improving the substantial symptom burden and reducing the risk of progression to MF or AML, thromboembolic events and death.

The clinical evidence presented demonstrate that ruxolitinib represents a clinically effective treatment for this indication, achieving a high response rate and improvements in MPN-SAF scores as well as reducing the incidence of key events. Ruxolitinib would provide an alternative and targeted effective treatment that helps to control HCT, and reduce the risk of disease progression, thromboembolic events and death in patients with PV who are R/I to HC/HU.

B.3 Cost effectiveness

Summary of the cost-effectiveness analysis

A de novo cost-effectiveness model was developed to assess the cost-effectiveness of ruxolitinib compared with established clinical management (BAT) in England and Wales in adult patients with PV who are R/I to HC/HU.

The economic evaluation is based on the two pivotal Phase 3 randomised trials (RESPONSE and RESPONSE-2) and one Phase 2 study (MAJIC-PV). A state-transition approach is used for the primary analysis based on the RESPONSE-trials (licensed population with and without splenomegaly). A partitioned survival model is used for the subgroup analysis in high-risk PV based on the MAJIC-PV trial. Health states are defined by therapy phases (on ruxolitinib, on BAT, death). The inclusion of PV-related complications (thromboembolic, myelofibrosis, leukaemia, haemorrhage etc) was simplified in the absence of robust data, and the economic model only considered their impact on cost and quality-adjusted life years (QALYs). Utility values are estimated using the MF-8D in the base-case. Healthcare resource use were estimated by clinical experts.

Cost-effectiveness results

Base case deterministic results show that ruxolitinib (when provided with the existing confidential patient access scheme [PAS]) is associated with higher costs but also higher QALYs than current clinical management, with an incremental cost per QALY gained of [redacted] and [redacted] in the licensed population without and with splenomegaly at baseline respectively. The incremental cost-effectiveness ratio (ICER) in the subgroup of high-risk patients was [redacted].

Despite the large unmet need, ruxolitinib does not meet the criteria for a severity weight in this indication.

Based on the need for [redacted] to allow a cost-effective price to be offered for this appraisal, Novartis are working [redacted] (subject to NICE's assessment). [redacted], the ICERs are: [redacted] and [redacted] in the licensed population with and without splenomegaly at baseline respectively and [redacted] in the subgroup of high-risk patients. [redacted]

Probabilistic sensitivity analyses and deterministic sensitivity analyses were conducted and demonstrate that the cost-effectiveness results were robust in most scenario analyses. The key drivers and source of uncertainty were the treatment effect and its duration, and utility values.

Summary

PV is a rare disease, with no other reimbursed treatments in patients who are R/I to HC/HU to fulfil the large unmet medical need. At the current PAS price, the ICERs for ruxolitinib are [redacted] the commonly accepted cost-effectiveness threshold for the NICE STA process.

Novartis will continue to work [redacted] for ruxolitinib in the PV indication. [redacted] and demonstrate that ruxolitinib has the potential to represent a cost-effective use of NHS resources.

B.3.1 *Published cost-effectiveness studies*

An SLR was conducted in February 2019 and updated in June 2022 to identify published literature on economic evaluations in adult patients with PV. In total, five economic evaluations in PV were identified in the economic SLRs. Full details of the SLR search strategy and results are presented in Appendix G.

B.3.2 *Economic analysis*

A '*de novo*' economic model was developed to inform this NICE submission. The objective of this economic analysis is to assess the cost-effectiveness of ruxolitinib compared with established clinical management (e.g. BAT) without ruxolitinib in adult patients with PV who are R/I to HC (listed in the NICE final scope), also referred to as HU. As highlighted in Section B.1.1, HC and HU are interchangeable as the same drug and is referred to as HC/HU throughout the submission.

In line with the NICE reference case,[25] the analysis is conducted from the perspective of the NHS and Personal Social Services (PSS) and exclusively includes direct medical costs over a lifetime horizon.

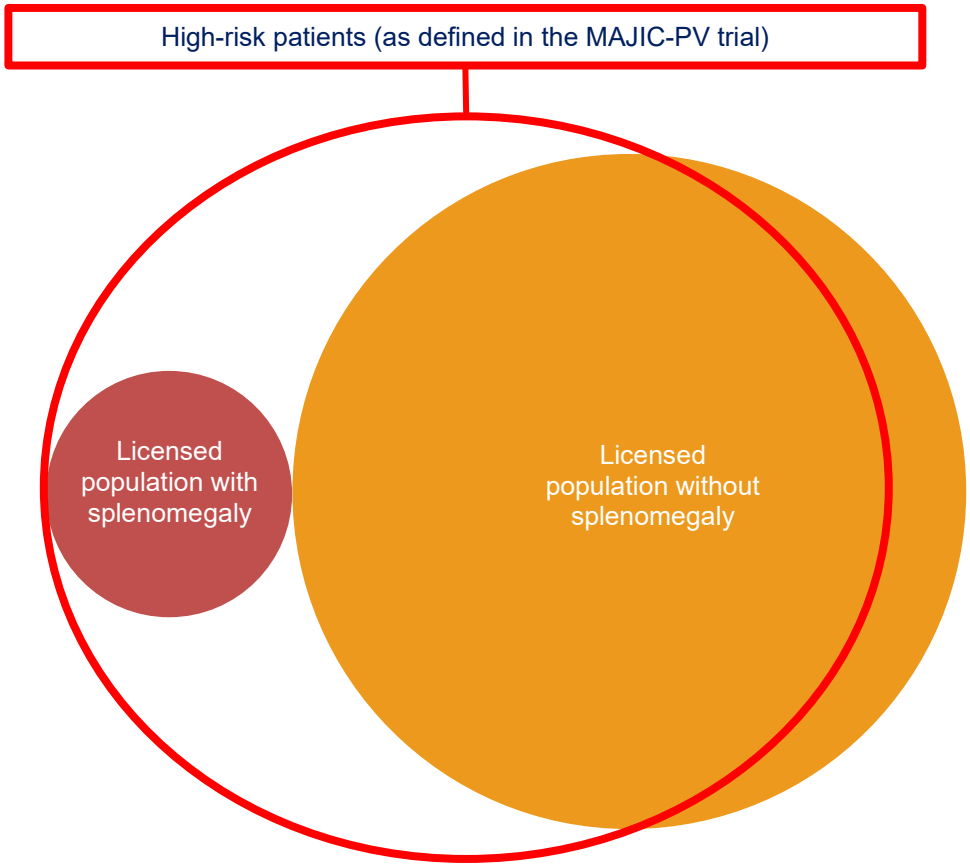
B.3.2.1 Patient population

The population covered in this economic evaluation (Figure 34) are adult patients with PV who are R/I to HC/HU. This is in line with the patient population described in the decision problem, the final scope issued by NICE[15] and covered by the marketing authorisation[16] for ruxolitinib.

The economic evaluation is based on the two pivotal Phase 3 randomised trials that supported the license,[65, 66] and final results from a UK Phase 2 study that are currently undergoing peer review[70]:

- **RESPONSE**: a pivotal Phase 3 RCT in adult patients with PV R/I to HC/HU with splenomegaly (representing approximately 20% of the licensed population).[65]
- **RESPONSE-2**: a pivotal Phase-3b RCT in adult patients with PV R/I to HC/HU without splenomegaly (representing approximately 80% of the licensed population).[66]
- **MAJIC-PV**: a UK Phase 2 investigator-led study in the subgroup of patients with high-risk PV R/I to HC/HU.[70] In this trial, high-risk was defined as ANY of the following:
    - Age >60 years;
    - Previous documented thrombosis (including TIA), erythromelalgia or migraine (severe, recurrent, requiring medications, and felt to be secondary to the MPN) either after diagnosis or within 10 years before diagnosis and considered to be disease related;
    - Significant splenomegaly (>5 cm below costal margin on palpation) or symptomatic (splenic infarcts or requiring analgesia);
    - Platelets > 1000 x 10^9/L;
    - Diabetes or hypertension requiring pharmacological therapy for >6 months.

Figure 34: Patient population considered in the economic model



Collectively, the RESPONSE-trials[65, 66] represent the entire adult patient population with PV who are R/I to HC/HU covered by the licence and described in the decision problem. As the population included in the pivotal Phase 3 RESPONSE and RESPONSE-2 trials[65, 66] are two mutually exclusive sub-populations (patients with PV R/I to HC/HU with and without splenomegaly), separate analyses are conducted for these two populations as described in the NICE final scope[105] to reflect any potential differences between populations. Analysis based on the RESPONSE-trials will be referred to as the primary analysis for the licensed population with and without splenomegaly.

Subgroup analyses of the RESPONSE and RESPONSE-2 trials[65, 66] show that the benefit of ruxolitinib over BAT is observed across all pre-specified subgroups investigated including sex, age (≤60 versus >60 years of age), patients who had unacceptable side effects from HC/HU (intolerant) and those who had an inadequate response to HC/HU (resistant), and risk category (low versus high) (Section B.2.8). Data were not available to conduct subgroup analyses.

In addition to the primary analysis based on the RESPONSE trials,[65, 66] a subgroup analysis is presented for adult patients with high-risk PV R/I to HC/HU based on the final results from the MAJIC-PV trial.[70] These results are currently unpublished, although the manuscript has been submitted for publication and is currently undergoing peer review. While Novartis does not have access to the data, authors were approached to include data from this study as part of this submission, despite the manuscript currently undergoing peer review. This analysis will be referred to as the high-risk PV subgroup. As highlighted in Section B.2.11.1, the definition for high-risk used in the MAJIC-PV trial is broader than that defined by the ELN.

B.3.2.2 Model structure

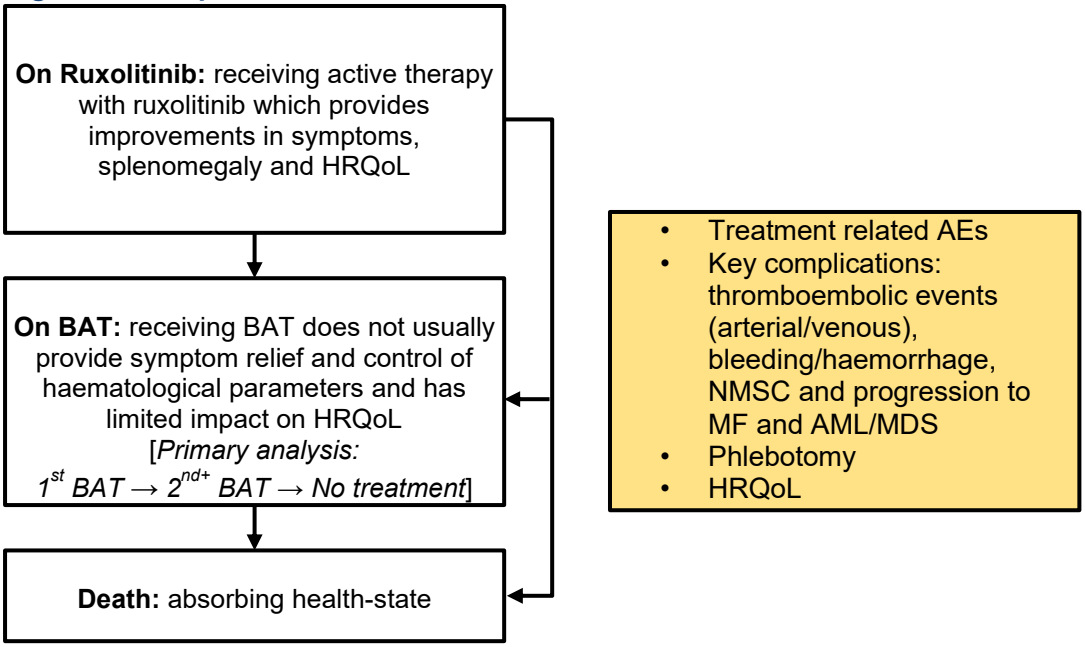
A de-novo economic model was developed in Microsoft Excel® to reflect the natural history and clinical pathway of adult patients with PV R/I to HC/HU in the England and Wales. The conceptual model was developed with the aid of an advisory group composed of five UK haematologists experienced in the treatment and management of PV through a series of interactive meetings (two advisory board meetings),[11, 106] teleconferences and email exchanges, supplemented by the published literature. Clinical experts were selected based on their experience in the management of patients with PV R/I to HC/HU, their experience with ruxolitinib for the treatment of patients with PV, and the different geographical regions covered. The advisory group provided clinical input and opinion on the following topics:

- The natural history of PV in patients that are R/I to HC/HU;
- Key features of PV that need to be captured in the economic model;
- Description of the current standard of care;
- Key benefits (and adverse reactions) expected from the use of ruxolitinib in PV;
- Plausibility of the survival extrapolations;
- Resource utilisation.

Model schematic

A simplified schematic of the model structure is shown below in Figure 35.

Figure 35: Simplified model structure schematic



Abbreviations: AE: adverse event; AML: acute myeloid leukaemia; BAT: best available therapy; HRQoL: health-related quality of life; MDS: myelodysplastic syndrome; MF: myelofibrosis; NMSC: non-melanoma skin cancer.

Health state and PV-related complications captured in the economic analysis

The clinical advisors indicated that PV is characterised by a progressive worsening of symptoms, haematological parameters, spleen size (in patients with splenomegaly) and HRQoL. Key complications that develop as the disease progresses include thromboembolic (TE) events (venous and/or arterial), bleeding/haemorrhage and transformation to leukaemia or MF.

The clinical advisors also indicated that treatment regimens that are part of established clinical practice in England and Wales (e.g. BAT) do not usually provide symptom relief or the symptom relief is short-lived, with patients switching and cycling through different regimens until treatment options are exhausted.[11]

Similar to the approach used in MF in TA386 and TA756, outcomes with respect to HRQoL and costs are largely defined by a patient's phase in the management of the condition.[11] Consequently, the key health states in the model are defined by the therapy phases, namely:

- *On ruxolitinib*: Patients initiating ruxolitinib enter the model in the ruxolitinib health state. Upon discontinuation, patients are able to move to the death health state directly (when the cause for discontinuation is death) or move to the BAT health state with the disease progressively worsening in the primary analysis. In this health state, patients receive active therapy with ruxolitinib which provides improvements in symptoms, reduction in spleen size (in patients with splenomegaly), HRQoL and could be associated with a reduction in key PV-complications including TE, MF and leukaemic events, but an increase in NMSC.
- *On BAT*: In the absence of ruxolitinib or following ruxolitinib discontinuation (due to reasons other than death), patients initiating on BAT enter the model in the BAT health state. While patients remain in this health state until death, patients are allowed to cycle and switch through BAT regimen (1st BAT regimen, 2nd+ BAT regimen sub-health states) or discontinue treatment (no further treatment sub-health state) in the primary analysis prior to death. In this health state patients receive BAT which does not provide symptom relief, or the symptom relief is short-lived, with patients switching and cycling through different regimens with HRQoL deteriorating over-time.
    a. To account for the progressive worsening in HRQoL as patients cycle through BAT regimens, the BAT health-state is further partitioned into three sub-health states in the primary analysis irrespective of the treatment arm. To explore the impact of the structural uncertainty for partitioning the BAT health state in the primary analysis (due to the number of assumptions required), a scenario analysis is conducted assuming a single BAT health-state (i.e. removing the partitions). The impact on the cost-effectiveness results for the primary analysis is minor (see Section B.3.11.3 and Appendix P). The BAT health state is partitioned onto:
        - 1st BAT regimen (first BAT treatment)
        - 2nd+ BAT regimen (second BAT treatment and beyond)
        - Exhaustion of BAT treatment (e.g. no further treatment).
    b. This is in contrast to the high-risk PV subgroup analysis where this was not possible due to the inflexibility of the modelling approach (see Section B.3.3.5).
- *Death*: Absorbing health state.

Defining health states based on therapy phases was employed in TA386[23] and TA756[107] during the assessment of ruxolitinib and fedratinib, respectively, for the treatment of MF patients. However, in contrast with TA386,[23] supportive care is not included explicitly as a separate health state following discussion with clinical experts. The rationale for this decision includes the consideration that PV is more benign, with a more favourable prognosis compared with MF. While a supportive care health state is not explicitly included in the model, in the primary analysis, patients in the BAT health state may discontinue treatment and enter the 'no further treatment' sub-health state where patients have exhausted all BAT regimens (where quality of life is reduced and management cost increased).

Additionally, clinical advisors indicated that patients with PV are at increased risk of complications such as TE (arterial and venous), bleeding/haemorrhage, transformation to leukaemia (acute myeloid leukaemia [AML] and myelodysplastic syndrome) and transformation to MF (post-PV MF).[11] 'Secondary' leukaemic transformation and MF are important aspects of the progression in PV and arterial and venous TE have significant impact on survival. Additionally, clinical advisors noted that although uncertain, there is evidence suggesting that patients on ruxolitinib may experience a higher incidence of non-melanoma skin cancer (NMSC) compared with patients on BAT.

The inclusion of these events in the model structure is particularly challenging. Robust modelling of these events is challenging as demonstrated by the complexity of the economic models used in prior NICE TAs for these conditions[23, 107-111] and would require many assumptions and data that are not available for this population. Further details are available below. Moreover, the clinical effectiveness data used within the model for survival already includes patients with these events; therefore, including these events as a separate health state would lead to the double counting of their impact on survival. Overall, clinical advisors agreed that the benefit of including these events as separate health states[11] would not outweigh the increased uncertainty and complexity due to the large number of assumptions that would be required. Consequently, their inclusion was simplified, and the economic model only considered the impact on cost and QALYs (modelled as an event rather than explicit health state).

Justification for the approach used to model OS

A key challenge when conducting an economic evaluation for a rare disease such as PV is the limited amount of available evidence, limited follow-up and the small patient numbers included in the studies. These challenges need to be recognised in the context of the modelling approach.

Five-year OS data are available for only ruxolitinib in the RESPONSE-trials; due to the cross-over study design, no patients remained on BAT after Week 80.[65, 66] In the MAJIC-PV trial,[70] five-year OS data are reported for both the ruxolitinib and BAT arms of the trial in the manuscript currently under peer review (shared by the authors). However, due to the comparatively good prognosis of patients with PV, these data on OS remain immature. OS is also confounded for BAT in the RESPONSE-trials due to the high and early cross-over (all patients on BAT crossed-over to ruxolitinib by Week 80). There are therefore challenges when modelling OS and the effect of ruxolitinib over BAT on survival.

As highlighted in Section B.1.3, patients with PV are at increased risk of complications such as TE, bleeding/haemorrhage, transformation to leukaemia (AML and MDS) and transformation to MF (post-PV MF); all can have a significant effect on survival. These complications could theoretically be used as a surrogate for OS, with a reduction in the incidence of events leading to an

improvement in survival. NICE and the external assessment group (EAG) during the decision problem meeting suggested that such an approach could be explored if feasible. However, following clinical input,[11, 106] careful consideration of the data available,[65, 66, 70] the need for transparency and robustness,[25] this approach was not considered feasible or robust for the following reasons:

- The total number of TE and bleeding/haemorrhage events by treatment arm is not reported in the unpublished manuscript of the MAJIC-PV trial.[70] The trial is investigator-led and therefore Novartis do not have access to the IPD under the existing MAJIC investigator initiated trial (IIT) contract.
- All events are competing with each other, e.g. patients with a TE event may subsequently transform to AML or MF. Similarly, patients with MF/AML may subsequently develop a TE event. Therefore, a large number of transitions are required, which cannot be estimated from the key trials due to the small number of events reported, short follow-up and absence of IPD from the MAJIC-PV trial.[70]
- None of the trials were designed or powered to detect difference in events, and events take time to manifest. Therefore, large sample sizes and long follow-up are required to robustly assess the incidence and effect of ruxolitinib on these events.
- In the RESPONSE-trials,[65, 66] all patients in the BAT arm crossed-over by Week 80, which is not long enough to robustly capture the incidence of TE, AML/MDS and MF for BAT.
- No AML was reported in RESPONSE-2[66] in the ruxolitinib arm at 5 years. Similarly, the number of TE events in the RESPONSE-trials[65, 66] was low at 5-years (5 in RESPONSE and 8 in RESPONSE-2) and therefore are unlikely to reflect the full distribution of TE events that patients with PV R/I to HC/HU experience. This is important as the prognosis differs according to the type of event and patient's history.
- In the unpublished manuscript of the MAJIC-PV trial,[70] despite 5-year follow-up, no AML events were reported in the BAT arm (while four were reported in the ruxolitinib arm) which is inconsistent with the published literature.[112]
- The baseline risk of mortality for patients with PV R/I to HC/HU without events cannot be robustly estimated from the RESPONSE-trials[65, 66] due to the immaturity of the data. Approximately 91% and 96% of patients on ruxolitinib were still alive at the end of follow-up period in RESPONSE and RESPONSE-2, respectively. Similarly, the baseline risk of mortality (without events) cannot be estimated from the MAJIC-PV trial in the absence of IPD.[70]
- Evidence from the general population suggest that the incidence of TE varies by age with the incidence increasing as patients get older.[109] The incidence by age cannot be estimated from the key trials. The incidence of other events (MF, AML/MDS) is also unlikely to be constant as demonstrated in Alvarez et al. 2022.[112]
- Robustly capturing survival and quality of life for the key events in PV (MF, AML, MDS, TE, bleeding/haemorrhage) would require developing separate models for each of these conditions. Models used in previous TAs for these conditions are typically complex[23, 107-111] and rely on a large number of transitions (which are typically estimated from large clinical trials or registries involving thousands of patients). In particular, there are many types of arterial (myocardial infarction, stroke, peripheral artery disease) and venous (deep vein thrombosis, pulmonary embolism) events that each have a different prognosis, with

the probability of experiencing a future event (the same or a different one) being a function of the previous event (type and history) and age. While it is possible to source some of these transitions from previous TAs, patients with PV are at higher risk of TE events compared with the general population. Patients that are R/I to HC/HC are also likely to behave differently compared with non-R/I patients. Clinical advisors agreed that using data from the general population is unlikely to be appropriate.[106] There are also doubts if using data from patients with PV that is non-R/I would be appropriate.

- Finally, conditions are heterogeneous in nature and require modelling of the entire pathway. As previously highlighted, there are many types of arterial and venous events associated with different prognoses. Modelling TE events as a single event is unlikely to be appropriate. Similarly, the management of MF in England and Wales varies according to the risk category with patients with intermediate-2/high-risk (Int-2/high-risk) MF secondary to PV treated with ruxolitinib or BAT as recommended in TA386,[23] while those with low or intermediate-1 risk (low/int-1 risk) receive BAT (consisting mostly of HC/HU and watch and wait). The management of AML in the UK also varies according to the presence of mutation and eligibility for stem cell transplant. Therefore, to robustly capture survival and quality of life associated with these complications would require developing complex models capturing the entire pathway. While simplifications could be made, these would add to the overall uncertainty.

In summary, while clinical experts recognised conceptually that modelling PV-complications as a surrogate for OS makes sense, they agreed that due to the absence of data, increased complexity and large number of assumptions required, the benefits of this approach did not outweigh the limitations.[106] Clinical experts agreed that a simpler approach using the final endpoint would be more appropriate and transparent, despite having limitations.[106] Clinical experts further agreed that despite uncertainty with the long-term extrapolation, 5-year OS data are available in the MAJIC-PV trial[70] and that using direct data from the MAJIC-PV trial[70] is preferable and would be considered more robust, compared with attempting to link events to survival using data from different populations and unsupported assumptions.[106] Clinical experts agreed that modelling survival directly would be more transparent and simpler to comprehend.[106] Clinical experts further agreed that modelling complications as a surrogate for OS is likely to lead to predictions for OS that are inconsistent with data reported in the MAJIC-PV trial due to the use of external evidence from different populations.[106] Consequently, modelling complications as a surrogate for OS was not considered further following clinical validation[11, 106] and data availability.[65, 66, 70]

Description of the modelling approach

Economic models for anti-cancer therapies that utilise the final endpoint (OS) typically follow one of two approaches:

1. A partitioned survival model (PSM) approach whereby curves are fitted directly to the trial outcomes and deal with health state occupancy (rather than transition between health states). The area under the curve (AUC) is used to estimate proportions of patients occupying each health state at a given time.

2. A state-transition model (STM) approach whereby OS is considered as a function of the time spent in an intermediate health state, with PFS/time to treatment discontinuation (TTD) typically used as an intermediate outcome for OS.

The approaches used for the primary and subgroup analyses are summarised below and were selected based on the evidence available and the ability for the model to capture the key features of the condition and the intervention.

Approach used for the primary analysis for the licensed population

IPD from the RESPONSE-trials[65, 66] are available to Novartis, so the model supporting the primary analysis for the licensed population with and without splenomegaly used a cohort state-transition approach whereby transitions between health states were explicitly modelled (ruxolitinib→BAT→death) with:

- *Patients initiated on ruxolitinib*: OS for ruxolitinib was modelled indirectly based on the extrapolated TTD for ruxolitinib (modelled under a competing-risk framework – Section B.3.1.2) and the time to death following ruxolitinib discontinuation (Section B.3.3.2) using data from the RESPONSE-trials.[65, 66]
- *Patients initiated on BAT*: In contrast, patients initiated on BAT (e.g entering the model in the BAT arm) remain in this health state until death (although they are allowed to cycle through/switch BAT until discontinuation in the primary analysis). Consequently, the survival for patients initiated on BAT is based on the predicted survival for ruxolitinib (estimated using the STM approach) adjusted downward using the treatment effect taken from the MAJIC-PV trial (Section B.3.3.3).

An STM approach with TTD modelled under a competing-risk framework was preferred over an AUC approach where OS and TTD are modelled directly for the following reasons:

- *Challenges in modelling survival directly*: While the RESPONSE-trials[65, 66] provided 5-year data, the survival for ruxolitinib was immature, with approximately 91% and 96% of patients still alive at the end of the follow-up periods for RESPONSE and RESPONSE-2, respectively.[65, 66] This high survival rate reflects both the prognosis of patients with PV and the potential benefits of ruxolitinib on survival. Furthermore, crossover from BAT to ruxolitinib was permitted in both trials (after Week 32 in RESPONSE and Week 28 in RESPONSE-2) and by Week 80 all patients in the BAT arms had discontinued BAT.
- *It enables the incorporation of discontinuation due to death over-time and preserves the correlation between time on treatment and death*: Ruxolitinib is well tolerated and effective. Therefore, patients remain on ruxolitinib for an extended period of time as demonstrated in the RESPONSE-trials[65, 66] and the MAJIC-PV trial[70] and supported by clinical experts.[106] Consequently, plausible separate extrapolations of OS and TTD is challenging. Furthermore, few discontinuations due to death were reported while on ruxolitinib in the RESPONSE-trials[65, 66] and therefore the direct extrapolation of TTD would not account for the increasing likelihood of discontinuation due to death with time. TTD is therefore modelled under a competing-risk framework[113] to account for discontinuations due to death (accounting for the risk of mortality over time) and discontinuation due to reasons other than death (both acting as competing events). This approach allows a more plausible modelling of TTD as it accounts for the increased likelihood of discontinuation due to death as patients age.
- *It enables reflection of the progressive worsening in HRQoL in the primary analysis and management cost to vary with time*: During the conceptual phase of the model development, clinical advisors indicated that quality of life while on BAT would worsen over

time. Clinical advisors considered this was an important feature of PV that needed to be captured if possible. Furthermore, resource use are varying with time (higher the first 6 month). As patients initiated on ruxolitinib move to BAT at different times following discontinuation of their primary treatment, tunnel states are required (progressive worsening of HRQoL as patients cycles through/switch between BAT regimens and varying resource use with time).

Approach used for the high-risk PV subgroup

A PSM approach was used for the subgroup analysis in patients with high-risk PV R/I to HC/HU (as defined in the MAJIC-PV trial) whereby OS and TTD are extrapolated directly from the outcomes reported in the unpublished manuscript of the MAJIC-PV trial.[70] As the trial was investigator-led, IPD from the MAJIC-PV trial do not belong to Novartis and Novartis do not have access to the IPD under the existing MAJIC IIT contract. Therefore, analyses could only be conducted based on the information reported in the unpublished manuscript, limiting the choice of modelling approach. The 5-year KM for OS was available for both the ruxolitinib and BAT arms, with the 5-year survival for patients randomised in the BAT and ruxolitinib arms approximately 72% and 80%, respectively.

A PSM was chosen over an STM primarily due to the absence of IPD. In a PSM, despite the movement of patients between health states, transitions are not explicitly modelled but the proportion of patients in each health state at each cycle is determined by the AUC for OS and TTD for BAT and ruxolitinib.

In the base-case, OS for BAT was extrapolated directly with OS for ruxolitinib estimated by applying a treatment effect to the BAT OS. In this analysis, the BAT health-state was not partitioned due to inflexibility with the PSM approach (compared with STM).

B.3.2.3 Feature of the economic analysis

The key features of the '*de novo*' analysis are summarised in Table 19. The model estimates the cost per QALY which is in line with the NICE methods guide.[25] A 28-day cycle length was used, which was considered short enough to capture the differences in costs and outcomes. The decision model employs a lifetime patient horizon and uses a direct NHS and PSS perspective as recommended by the NICE methods guide.[25] A patient lifetime horizon was used to reflect the chronic nature of the disease and to capture all the relevant costs and benefits associated with the introduction of ruxolitinib in England and Wales. However, given the uncertainty in the long-term extrapolation, shorter time horizons are considered in scenario analyses (see Section B.3.10.3 and Appendix P). The decision model uses a discount rate of 3.5% per annum for both costs and benefits in the base-case as recommended in the NICE methods guide for economic evaluation.[25] Alternative discount rates are explored in scenario analyses (see Section B.3.10.3 and Appendix P). Half-cycle correction is not included due to the short cycle length (28 days).

Table 19: Features of the economic analysis

	Current appraisal	
Factor	**Chosen values**	**Justification**
Cycle Length	28 days	Sufficiently short to capture the differences in costs and outcomes between ruxolitinib and BAT[25]

	Current appraisal	
Factor	**Chosen values**	**Justification**
Perspective	NHS/PSS	NICE reference case[25]
Time horizon	Lifetime (max 600 cycles / 46 years)	Sufficient to capture all meaningful differences in technologies compared[25]
Discounting	3.5%	NICE reference case[25]
Model type	Licensed population: STM High-risk PV subgroup: PSM	An STM was chosen for the primary analysis due to challenges in modelling OS directly (data immaturity) and to account for progressive worsening in HRQoL on BAT and resource use varying with time. A PSM was chosen for the subgroup analysis in the absence of IPD. MF, AML/MDS, bleeding/haemorrhage, NMSC and TE were included as events (costs and QALY impact only)
Treatment waning effect	20 years in the base-case; treatment effect gradually diminishes (linearly) between Year 5 – Year 20	In the base-case, the treatment effect is assumed to stop after 20 years and is assumed to diminish gradually (linearly) between Year 5 and Year 20, to reflect clinical expectations (Section B.3.3.4)
Source of utilities	RESPONSE trials[65, 66] MF-8D and EQ-5D (scenario analysis)	Evidence suggests that the EQ-5D does not sufficiently capture HRQoL in PV and condition-alike (such as MF). Consequently, the MF-8D (a condition specific measure) was used in the base-case as the EQ-5D was not deemed appropriate as explained in the method guide.[25] The EQ-5D was used in a scenario analysis.
Source of costs	NHS reference costs 2020/2021,[114] PPSRU 2021,[115] BNF,[24] eMIT[116]	The sources of cost data are as per the NICE methods guide[25]

Abbreviations: AML: acute myeloid leukaemia; BAT: best available therapy; BNF: British natural formulary; eMIT: electronic market information tool; EQ-5D: euroqol 5-dimensions; HR: hazard ratio; HRQoL: health-related quality of life; MDS: myelodysplastic syndrome; MF: myelofibrosis; NHS: national health service; NICE: National institute for health and care excellence; NMSC: non-melanoma skin cancer; PSS: Personal Social Services; PV: polycythaemia vera; QALY: quality-adjusted life years; OS: overall survival; PSM: partitioned survival model; STM: state-transition model; TE: thromboembolic events; TTD: Time to treatment discontinuation.

B.3.2.4 Intervention technology and comparators

Intervention: ruxolitinib

The economic analysis utilised evidence from the RESPONSE-trials[65, 66] and evidence presented in an unpublished manuscript (currently undergoing peer review) of the MAJIC-PV trial[70] in which ruxolitinib was prescribed in accordance with its licence (recommended starting dose of 10 mg twice daily [BID]) and expected administration in UK practice.[16] No formal stopping rule was considered. This was confirmed by clinical experts.[11]

Comparator: established clinical practice without ruxolitinib

The comparator in the economic model was established current management, also referred to as best available therapy (BAT). This is in line with the comparator defined in the NICE final scope[15] and included in the RESPONSE and MAJIC-PV trials.[65, 66, 70]

The NICE final scope[105] defined the comparators of interest to be established clinical practice without ruxolitinib including HC/HU, interferon-alfa, anagrelide, busulfan and radioactive phosphorus.

In the RESPONSE-trials[65, 66] the most frequently used therapies were HC/HU, IFN-alfa and anagrelide. Clinical advisors[11, 106] noted that:

- A minority of patients (4.5% in RESPONSE and 1.33% in RESPONSE-2) received immunomodulatory drugs (e.g. lenalidomide) that are not used in England and Wales.
- A minority of patients (1.8% in RESPONSE and 6.67% in RESPONSE-2) received pipobroman which is no longer available in England and Wales.
- Radioactive phosphorus is no longer used in England and Wales.
- The proportion of patients receiving no cytoreductive therapy (15.32% in RESPONSE and 28.00% in RESPONSE-2) was higher than expected in UK clinical practice.

Clinical advisors further indicated that the RESPONSE trials[65, 66] were multinational, and that IFN-alfa (Pegasys) is the primary treatment used in England and Wales in patients with PV R/I to HC/HU, unless patients are contraindicated to or cannot tolerate or receive interferons.[11, 106]

Clinical advisors indicated that it was important for the BAT composition in the model to reflect the mix of treatments currently given in England and Wales.[11, 106] Consequently, the BAT composition reported in the unpublished manuscript of the MAJIC-PV trial[70] (Table 20) was used for the purpose of the economic model to reflect clinical practice in England and Wales. Using the BAT composition reported in the unpublished MAJIC-PV manuscript[70] also allows for costs to align with effectiveness data. It should be noted that 10 patients received subsequent ruxolitinib in the BAT arm, the costs of which were excluded from the model to avoid overestimating costs. Furthermore, one patient received pipobroman and three patients received ^{32}P. These treatments have been excluded as pipobroman and ^{32}P are no longer available in England and Wales.

Table 20: BAT composition reported in the unpublished manuscript of the MAJIC-PV trial

BAT drugs given	N	%
^{32}P*	1	1.15%
Anagrelide	3	3.45%
Busulfan	1	1.15%
HC/HU	28	32.18%
Interferon	13	14.94%
^{32}P*, HC/HU	1	1.15%
Anagrelide, HC/HU	9	10.34%
Anagrelide, Interferon	3	3.45%
Anagrelide, Ruxolitinib*	1	1.15%
Busulfan, Interferon	2	2.30%
Busulfan, Ruxolitinib*	1	1.15%
HC/HU, Interferon	10	11.49%
HC/HU, Ruxolitinib	1	1.15%

Interferon, Pipobroman*	1	1.15%
Interferon, Ruxolitinib*	2	2.30%
^{32}P*,Anagrelide, Interferon	1	1.15%
Anagrelide, HC/HU, Interferon	3	3.45%
Anagrelide, Interferon, Ruxolitinib*	1	1.15%
Busulfan, HC/HU, Interferon	1	1.15%
HC/HU, Interferon, Ruxolitinib*	3	3.45%
Anagrelide, HC/HU, Interferon, Ruxolitinib*	1	1.15%

Abbreviations: BAT: best available therapy; HC/HU:hydroxycarbamide/hydroxyurea; ^{32}P: Phosphorus-32.

* excluded from the economic analysis

B.3.3 *Clinical parameters and variables*

The sources for the clinical parameters used in the economic model are summarised below in Table 21 and discussed in turn. It should be noted that the final data-cuts (5-year data) were used for RESPONSE[65] (data cut: February 9th 2018), RESPONSE-2[66] (data cut: April 7th 2020) and the MAJIC-PV trial[70] (data cut: April 2022).

Table 21: Summary of sources of data used in the economic model

Parameter	Licensed population	Subgroup	Reference in Submission
Baseline characteristics	RESPONSE-trials[65, 66]	MAJIC[70]	Section B.3.3.1
TTD on ruxolitinib	RESPONSE-trials[65, 66]	MAJIC[70]	Section B.3.1.2
Rux post-discontinuation survival	RESPONSE-trials[65, 66]	N/A	Section B.3.3.2
OS BAT – MAJIC analysis	N/A	MAJIC[70]	Section B.3.3.3
HR OS	MAJIC[70]		Section B.3.3.4
Partitioning BAT health state	MAJIC[70]	N/A	Section B.3.3.5
UK life table	ONS[117]		Section B.3.3.6
Incidence of adverse events	RESPONSE-trials[65, 66]		Section B.3.3.7
Incidence of events and venesection	RESPONSE-trials[65, 66]	MAJIC[70]	Section B.3.3.8 & B.3.3.9
IRR (MF, AML/MDS, TE, NMSC, venesection)	Pooled RESPONSE-trials[65, 66] and MAJIC[70]		Section B.3.3.8 & B.3.3.9

Abbreviations: AML: acute myeloid leukaemia; BAT: best available therapy; HR: hazard ratio; IRR: incidence rate ratio; MDS: myelodysplastic syndrome; MF: myelofibrosis; N/A: not applicable; NMSC: non-melanoma skin cancer; OS: overall survival; Rux: ruxolitinib; TTD: Time to treatment discontinuation.

B.3.3.1 Baseline characteristics

The baseline characteristics for the modelled cohort of patients for the licensed population were derived from the RESPONSE-trials[65, 66] since the patients included in the trials were deemed representative of patients in England and Wales (Table 22).[11] The baseline characteristics for the high-risk PV subgroup (as defined in the MAJIC-PV trial) were taken from the unpublished manuscript of the MAJIC-PV trial.[70].

The mean/median age and gender distribution were used in the model, in conjunction with UK life tables[117] to incorporate the deterioration of quality of life with age and incorporate general population mortality when extrapolating survival curves in order to avoid the extrapolated hazard being less than the expected hazard of death in the general population.

Table 22: Baseline characteristics at entry

Baseline characteristic	RESPONSE-trials[65, 66]		MAJIC[70]
	with splenomegaly	Without splenomegaly	High-risk PV
Age	[redacted] (SD: [redacted])	[redacted] (SD: [redacted])	66 (28 – 88) - median
% male	[redacted]	[redacted]	58.33%

Abbreviations: N/A: not applicable; PV: polycythaemia vera; SD: standard deviation.
Source: Analysis of the RESPONSE-trials IPD[65, 66], unpublished manuscript of the MAJIC-PV trial.[70]

B.3.1.2 Time to ruxolitinib discontinuation

Primary analysis for the licensed population

In the primary analysis, TTD was modelled under a competing risk framework with the resulting TTD relative to the Kaplan–Meier (KM) presented below in Figure 36 and Figure 37 for the licensed population with and without splenomegaly respectively (after selection of the most appropriate extrapolation). A competing-risk approach[113] was chosen over direct extrapolation of TTD due to the small number of discontinuations due to death observed in the RESPONSE-trials ([redacted] in RESPONSE[65] and [redacted] in RESPONSE-2[66]) and to incorporate the higher likelihood of discontinuation due to death beyond the observed period which would not be possible if discontinuation for all reasons (otherwise TTD) was modelled directly as a single endpoint.

Consequently, discontinuation due to death and discontinuation due to other reasons were modelled as two separate and competing events and subsequently combined under a competing-risk framework (Figure 36 and Figure 37).

Figure 36: Comparison of the KM and predicted TTD for ruxolitinib estimated under the competing-risk framework for the licensed population with splenomegaly (RESPONSE)

Abbreviations: KM: Kaplan–Meier; TTD: time to treatment discontinuation.
Source: Analysis of the RESPONSE-trials IPD.[65, 66]

Figure 37: Comparison of the KM and predicted TTD for ruxolitinib estimated under the competing-risk framework for the licensed population without splenomegaly (RESPONSE-2)

Abbreviations: KM: Kaplan–Meier; TTD: time to treatment discontinuation.
Source: Analysis of the RESPONSE-trials IPD.[65, 66]

Time to ruxolitinib discontinuation due to reasons other than death (e.g. discontinuations due to death are censored)

IPD was obtained from the RESPONSE[65] and RESPONSE-2[66] trials. Discontinuations due to death were censored and only discontinuations due to reasons other than death were considered as events. A total of [redacted] and [redacted] discontinuations due to reason other than death were reported in RESPONSE and RESPONSE-2 trials respectively. The KM curves for the analysis of time to discontinuation (with death censored) is presented in Figure 38 for RESPONSE (n=110) and RESPONSE-2 (n=74).

Figure 38: KM for the time to treatment discontinuation (with death censored)

Abbreviations: KM: Kaplan–Meier.
Source: Analysis of the RESPONSE trials IPD.[65, 66]

In accordance with the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 14,[118] a range of standard parametric distributions (exponential, Weibull, Gompertz, loglogistic, lognormal and generalised gamma) and more flexible models (hazard, normal and odds spline models with up to four knots) were explored in the extrapolation of the clinical trial data beyond the observed period. The spline models (hazard, normal, odds) with one and four knots were estimated in R using the FlexSurv package.[119]

Different parametric models incorporate different hazard functions. For instance, exponential models are only suitable if the observed hazard is approximately constant and positive. Weibull and Gompertz models incorporate monotonic hazards, while the logged model (lognormal) can incorporate non-monotonic hazards but typically has a long tail due to a reducing hazard as time increases beyond a certain point. Spline models, in contrast, are more flexible but can lead to overfitting of data.

NICE TSD 14[118] also recommends that the most appropriate distribution is selected based on consideration of: (a) the visual fit of the predicted models to the observed KM, (b) the statistical goodness-of-fit of the model relative to all other fitted models (measured using the Akaike

Information Criterion [AIC] or Bayesian Information Criterion [BIC]), (c) an assessment of the observed hazards and (d) the plausibility of the long-term extrapolation.

The fit of each parametric function relative to the KM curves is presented in Appendix N:

- For the licensed population with splenomegaly (based on RESPONSE[65]), none of the standard parametric distributions provided a good visual fit to the observed KM. While the spline models with more than one knot provided a better visual fit, the addition of two or more knots led to over-fitting, and therefore were not considered further.
- For the licensed population without splenomegaly (based on RESPONSE-2), the Gompertz distribution was the only standard distribution that provided a good visual to the observed KM. Similarly, while the spline models with more than one knot provided a better visual fit, the addition of two or more knots led to over-fitting, and therefore were not considered further.
- The generalised gamma was unstable and did not converge in both trials.
- The statistical goodness of fit in terms of AIC and BIC was relatively similar between the different distributions (Appendix N), with the Weibull and exponential distribution having the lowest BIC for RESPONSE and RESPONSE-2, respectively.

However, the statistical goodness of fit only provides an indication of the fit to the observed data, therefore assessing the plausibility of the long-term extrapolation beyond the observed period is important. Predictions at 5, 10, 15, 20 and 25 years are shown in Appendix N. Assessment of the long-term extrapolation for the time to ruxolitinib treatment discontinuation (with death censored) was informed by clinical expert opinion.[106] Clinical experts considered that, excluding discontinuation due to death as patients get older, patients are likely to remain on ruxolitinib for an extended period.[106] Clinical experts noted that the rates of discontinuation are highest early on, but once a patient has settled on to treatment they will rarely discontinue. Reasons for discontinuation in the long term are likely due to include death, infection or skin cancer. For the licensed population with splenomegaly, clinical experts considered the exponential to be pessimistic and the Gompertz to be optimistic and suggested that the most appropriate extrapolation would lie somewhere in-between.[106] Clinical experts expected a large proportion of patients to remain on treatment at 30 years (if discontinuation due to death were excluded).[106] Similarly, for the licensed population without splenomegaly, clinical experts considered the Gompertz distribution to be pessimistic (if discontinuation due to death were excluded) and indicated that similar to the licensed population with splenomegaly, a large proportion of patients would remain on treatment at 30 years if discontinuation due to death were removed.

In summary, while clinical experts were not able to select a curve confidently, in the base-case the odd spline model with one knot was selected for both patients with and without splenomegaly based on the visual fit to the KM, statistical goodness of fit, clinical plausibility and consistency between analyses. This curve was selected in the base-case to reflect clinical opinion[106] that the rate of discontinuations is high early on, and reduces with time and this curve also lay in between the curves presented (Appendix N). The fit of the selected parametric function in the base-case relative to the KM are presented in Figure 39.

Figure 39: Comparison of the KM and parametric distribution fits to TTD (death censored) for ruxolitinib for the licensed population

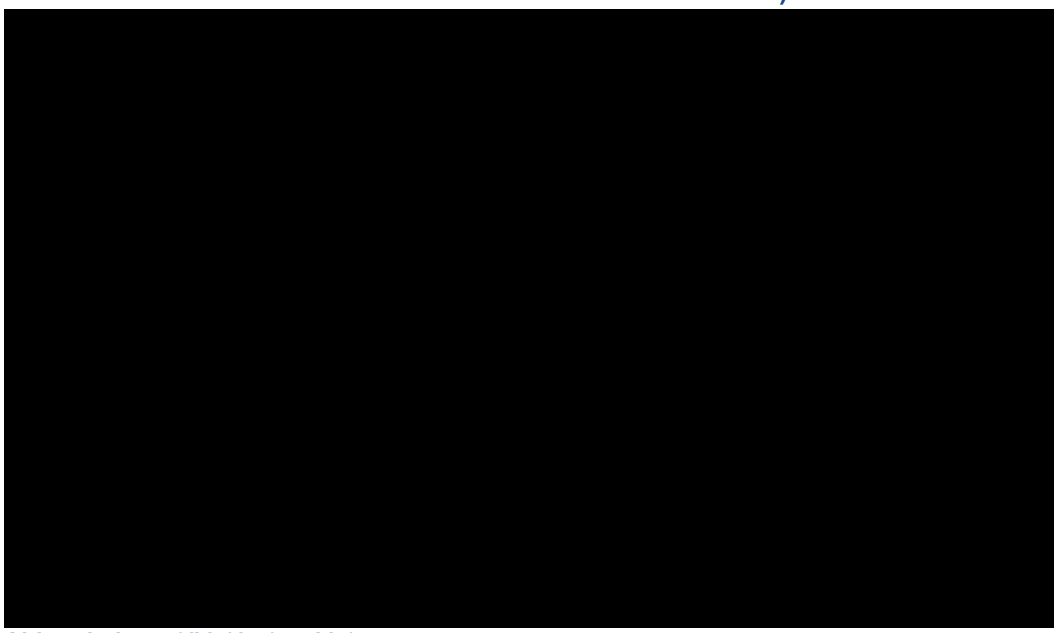
Abbreviations: KM: Kaplan–Meier; TTD: time to treatment discontinuation.
Source: Analysis of the RESPONSE trials IPD.[65, 66]

However, as highlighted, the choice of parametric extrapolation remains uncertain. Therefore, in line with the NICE TSD14,[118] extensive scenario analyses were conducted using alternative distributions and the observed KM followed by parametric extrapolation (see Section B.3.10.3 and Appendix P). Overall, the different plausible extrapolation methods had a modest impact on the cost-effectiveness results (see Section B.3.10.3 and Appendix P). The spline models with two or more knots led to an improvement in the ICER, but these over-fitted the data leading to high discontinuation rates beyond the observed period.

Pre-discontinuation survival (e.g. time to treatment discontinuation due to death [discontinuation due to reasons other than death are censored])

IPD from the RESPONSE-trials[65, 66] were obtained and analysed to estimate the pre-discontinuation survival. The KM curves for the analysis of pre-discontinuation survival (discontinuation due to reason other than death censored) are presented in Figure 40 for RESPONSE (n=110) and RESPONSE-2 (n=74), respectively. Due to the small number of events ([redacted] in RESPONSE and [redacted] in RESPONSE-2), data were pooled (n=184) in the base-case to increase the statistical power and reduce the uncertainty, although the number of events remained very low (n=[redacted]). Therefore, the same pre-discontinuation survival was used in the primary analysis in patients with and without splenomegaly. For transparency and completeness, a scenario analysis was conducted using the pre-discontinuation survival from these two trials separately (see Section B.3.10.3 and Appendix P).

Figure 40: KM for the time to pre-discontinuation survival for ruxolitinib (with discontinuation due to reasons other than death censored)



Abbreviations: KM: Kaplan–Meier.
Source: Analysis of the RESPONSE trials IPD.[65, 66]

Parametric functions were fitted to the data (Appendix N) and the selection process for the extrapolation of pre-discontinuation survival was similar to that described previously for TTD.

In summary, spline models and the generalised gamma did not converge due to the low number of events. The visual fit and long-term extrapolation was similar between the remaining curves examined (Appendix N).

The exponential distribution was used in the base-case as this had the best statistical fit (lowest AIC/BIC – Appendix N). Assuming a constant rate can also be deemed more realistic considering the small number of events. Alternative distributions were explored in scenario analysis, in addition to using the fit to each trial individually. The impact on the cost-effectiveness results was minor (see Section B.3.10.3 and Appendix P). A constraint was added to ensure that the extrapolated hazard of death beyond the observed period is greater than that of the background general population mortality[117] as the risk of discontinuation due to death increases as patients get older. The fit to the data selected in the base-case is presented in Figure 41, before and after including general population mortality.

Figure 41: Comparison of the KM and fit to the pre-discontinuation survival for ruxolitinib

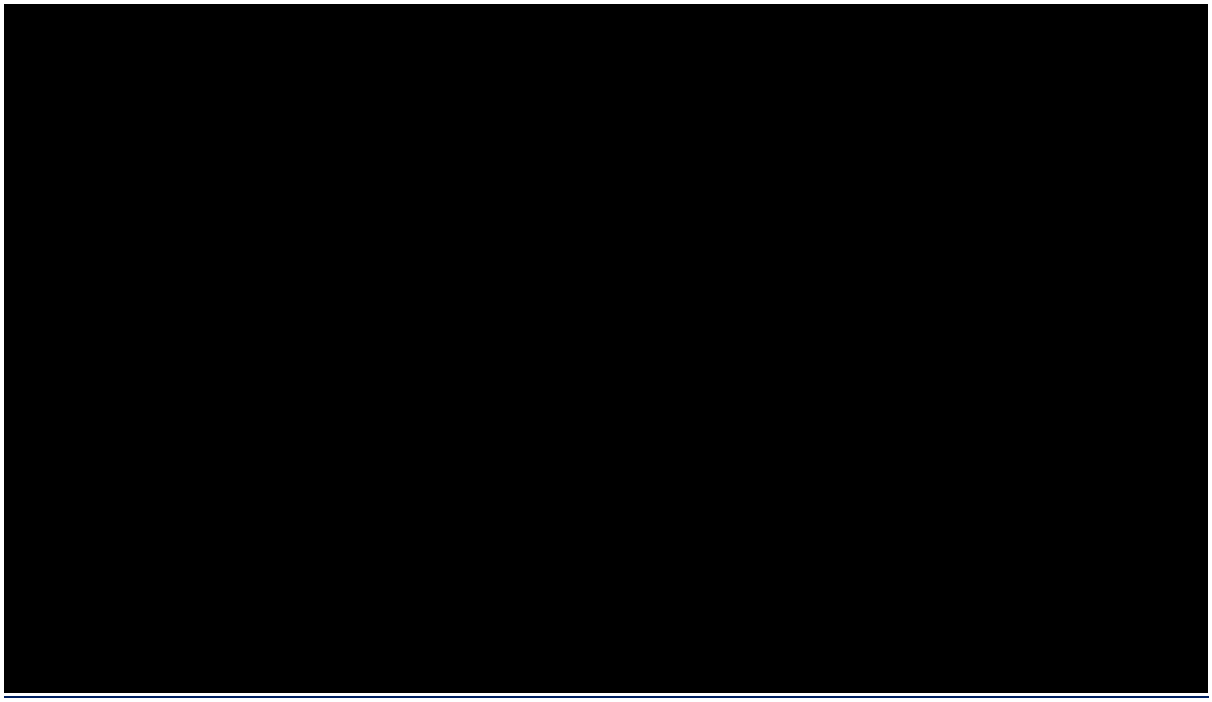
Abbreviations: gen pop: general population; KM: Kaplan–Meier.
Source: Analysis of the RESPONSE trials IPD[65, 66] and life table.[117]

Subgroup analysis in patients with high-risk PV

IPD were not available from the MAJIC-PV trial[70] and therefore it was not possible to model TTD under a competing risk framework. Clinical experts further indicated that the extrapolation for TTD for ruxolitinib needed to be consistent with OS.[106] Pseudo IPD for OS and TTD for ruxolitinib from the MAJIC-PV trial were reconstructed based on the KMs reported in the paper submitted for publication[70] (Figure 42) and a HR was derived ([95% CI:]) and applied the predicted ruxolitinib OS curve to allow consistent extrapolation.

For transparency and completeness, scenario analyses are presented fitting parametric distributions directly to the TTD KM from the MAJIC-PV trial[70] for ruxolitinib. Overall, the different extrapolation methods had a modest impact on results (see Section B.3.10.3 and Appendix P).

Figure 42: Comparison of the KM for ruxolitinib OS and TTD for ruxolitinib in the MAJIC-PV trial (based on the reconstructed pseudo-IPD)



Abbreviations: BAT: best available therapy; IPD: individual patient level data; KM: Kaplan–Meier; OS: overall survival; PV: polycythaemia vera; TTD: time to treatment discontinuation.
Source: Analysis of the reconstructed pseudo-IPD of the MAJIC-PV trial.[70]

B.3.3.2 Post-discontinuation survival (primary analysis only)

Survival following ruxolitinib discontinuation (or post-discontinuation survival) was only used for the primary analysis based on the RESPONSE-trials[65, 66] where an STM was used.

IPD from the RESPONSE-trials[65, 66] were obtained and analysed to estimate the time to death following ruxolitinib discontinuation for the primary analysis. The KM curves for the analysis of time to death following discontinuation are presented in Figure 43 for RESPONSE (n=[redacted]) and RESPONSE-2 (n=[redacted]). Given the small sample size, low number of events ([redacted] in RESPONSE and [redacted] in RESPONSE-2) and the absence of differences (HR: [redacted]; CI: [redacted]; p = [redacted]), data were pooled in the base-case (Figure 43) to increase the statistical power and reduce uncertainty. Therefore, the same post-discontinuation survival was used for the licensed population with and without splenomegaly. For completeness, scenario analyses were conducted using the time to death following discontinuation from these two trials separately. The impact on the cost-effectiveness results was minor (see Section B.3.10.3 and Appendix P).

Figure 43: KM plot for the time to death following ruxolitinib discontinuation for the primary analysis

Abbreviations: KM: Kaplan–Meier.
Source: Analysis of the RESPONSE trials IPD.[65, 66]

The selection process for the extrapolation for the time to death following ruxolitinib discontinuation was similar to that described previously for TTD in Section B.3.1.2. Only the exponential provided a reasonable visual fit to the data and was considered the most clinically plausible (Appendix N) of the curves presented. It also had the best statistical fit (Appendix N). The spline models with four knots overfitted the data, while other distributions predicted that more than 20% of patients would remain alive 10 years following ruxolitinib.[106] The fit of the exponential distribution to the data is presented in Figure 44.

Alternative distributions were explored in scenario analysis. Overall, the different extrapolation methods had a modest impact on results (see Section B.3.10.3 and Appendix P).

A constraint was included in the economic model to ensure that the extrapolation of post-discontinuation survival was consistent with that of the background mortality (based on the time when they stop treatment). Therefore, despite the exponential being used, patients do not remain on an extended time alive following ruxolitinib discontinuation as patients tend to discontinue later on.

Figure 44: Comparison of the KM and parametric distribution fits to the post-discontinuation survival (pooled RESPONSE trial)

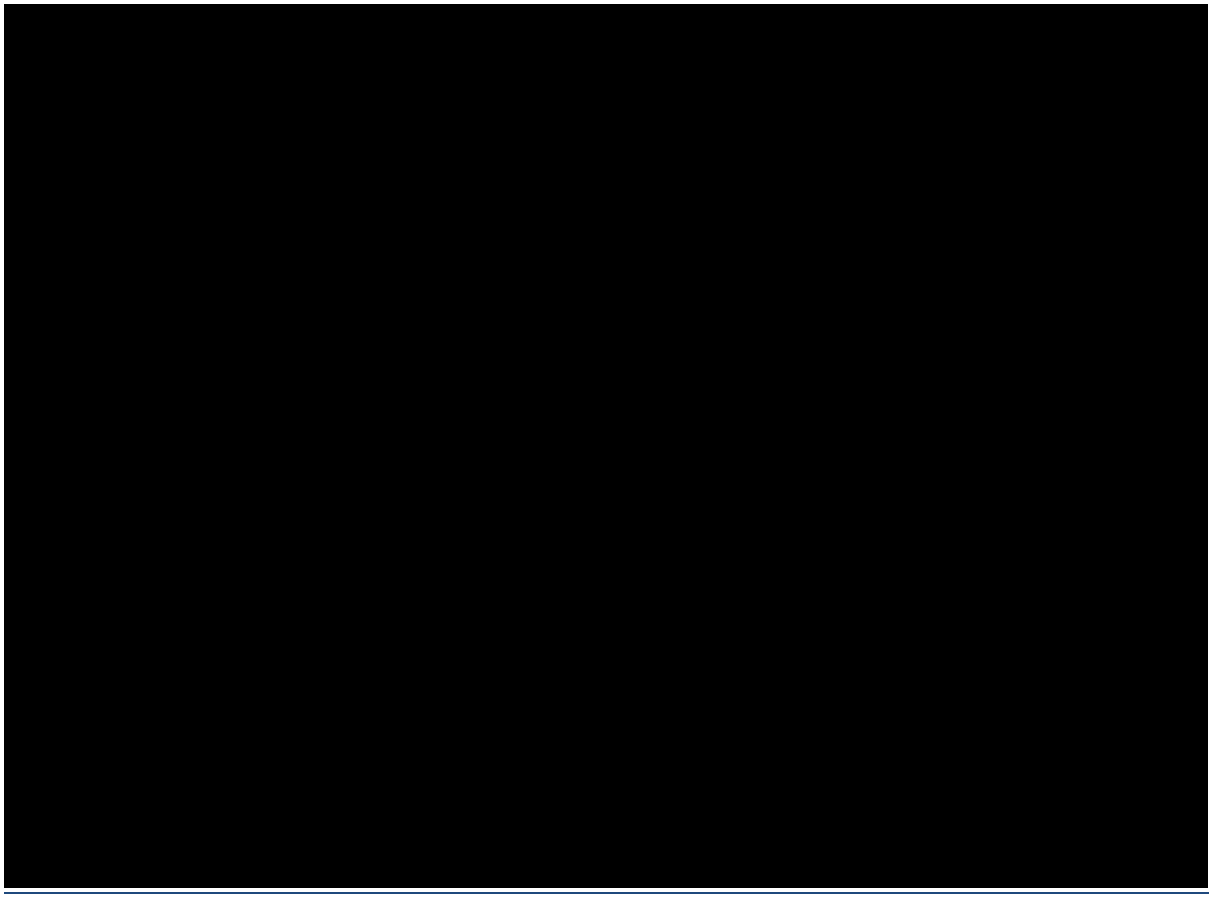
Abbreviations: KM: Kaplan–Meier.
Source: Analysis of the RESPONSE trials IPD.[65, 66]

B.3.3.3 Overall survival for BAT for the high-risk PV subgroup

Overall survival for patients initiated on BAT for the subgroup analysis of patients with high-risk PV R/I to HC/HU (as defined in the MAJIC-PV trial[70]) is extrapolated directly from the reconstructed OS curve from the unpublished manuscript of the MAJIC-PV trial (Figure 45).[70] Pseudo-IPD were generated and parametric functions were fitted to the data.

The selection process for the extrapolation for OS for BAT for the subgroup analysis was similar to that described previously for TTD (Section B.3.1.2). Except for the exponential distribution, all parametric models examined provided a good visual fit to the KM (Appendix N). The spline models with more than one knot overfitted the data and therefore were not considered further. The log-normal distribution had the best statistical fit but predicted that more than 25% of patients would remain alive after 25 years (Appendix N) which was considered optimistic by clinical advisors given the age of the population.[106] Clinical advisors expected that approximately 10–15% of patients would remain alive after 20 years considering the population recruited in the MAJIC-PV trial (high-risk PV).[106] Clinical advisors noted that in Tefferi et al. 2013, approximately 25% of patients with PV were still alive at 20 years after diagnosis when considering the most mature cohort (Mayo clinic; 44% followed to death).[106] However, this study reported outcomes from diagnosis, not in patients R/I to HC/HU and a non-contemporary cohort.

Figure 45: KM for OS for patients BAT in the MAJIC-PV trial (based on the reconstructed pseudo-IPD)



Abbreviations: BAT: best available therapy; IPD: individual patient level datal; KM: Kaplan–Meier; OS: overall survival; PV: polycythaemia vera.
Source: Analysis of the reconstructed pseudo-IPD of the MAJIC-PV trial.[70]

The Gompertz distribution predicted 0% of patients alive at 20 years (Appendix N), which was considered pessimistic.[106] The hazard spline model with one knot predicted 22.9% of patients alive at 20 years and 14% at 25 years. The Weibull distribution predicted 7.1% at 20 years and 2.2% at 25 years. The Weibull distribution is selected in the base-case.

Alternative distributions were explored in scenario analysis. The impact on the cost-effectiveness results was modest (see Section B.3.11.3 and Appendix P). It should be noted that in the economic model, a constraint was also added to ensure that the extrapolation of OS was consistent with that of the background mortality.

Figure 46: Comparison of the KM and parametric distribution fits to OS for BAT for the high-risk PV subgroup

Abbreviations: BAT: best available therapy; IPD: individual patient level datal; KM: Kaplan–Meier; OS: overall survival; PV: polycythaemia vera.
Source: Analysis of the reconstructed pseudo-IPD of the MAJIC-PV trial.[70]

B.3.3.4 Treatment effect for OS

In the primary analysis for the licensed population, OS for patients initiated on ruxolitinib was estimated indirectly from the TTD (Section B.3.1.2) and time to death following discontinuation (Section B.3.3.2) using an STM approach. OS for patients initiated on BAT was derived from the predicted OS for ruxolitinib (estimated using STM) adjusted downward using a time-varying treatment effect derived from the MAJIC-PV trial.[70]

In contrast, for the high-risk PV subgroup, OS for BAT was extrapolated directly (Section B.3.3.3) with OS for ruxolitinib estimated indirectly using a time-varying treatment effect derived from the MAJIC-PV trial.[70]

Treatment effect for OS assumed in the economic analysis

Clinical advisors indicated that it is challenging to demonstrate a survival gain in PV due to the comparably favourable prognosis of patients with PV, the small number of patients typically recruited into clinical studies and their short follow-up.[11] Clinical advisors also noted that none of the studies were powered or designed to evaluate the benefit of ruxolitinib on survival.[11] The advisors indicated that ruxolitinib was more effective compared with BAT in controlling HCT, reducing spleen size (in patients with splenomegaly at baseline), control of haematological parameters, achievement of CHR, reducing the JAK2V617F burden and reduction in TE events and MF.[11] It was further noted that these observed differences are likely to lead to an improvement in survival.[11] The benefits associated with ruxolitinib were also sustained over time.

Clinical advisors noted that the MAJIC-PV trial reported a statistically significant EFS (HR = 0.58; 95% CI: 0.35, 0.94; p = 0.3) and a non-statistically significant PFS (HR = 0.64; 95% CI: 0.36, 1.15; p=0.13).[11] Clinical advisors indicated that while the number of patients randomised and the follow-up in the MAJIC-PV trial was too short to demonstrate a statistically significant treatment effect for OS, the improvement in EFS and PFS seen in the trial is likely to translate into a survival gain.[11]

Clinical advisors further noted the curve for OS started to diverge after approximately 3.0 years in the MAJIC-PV trial (Figure 47), which was in line with their expectation that a survival difference would not manifest immediately.[11] Clinical advisors expected the curve to separate further in the long-term due to the effect of ruxolitinib on haematological parameters and reduction in key complications.[11, 106] Clinical experts further noted that in the MAJIC-PV trial, more patients on ruxolitinib had a molecular response, which is associated with a better survival.[106]

In the base-case, the treatment effect for OS derived from the unpublished manuscript of the MAJIC-PV trial[70] is used. To account for the late separation of the curve (Figure 47) a piecewise cox proportional hazard model was fitted to the reconstructed pseudo-IPD for OS using a cut-off point of [redacted] years. The cut-point for the piecewise Cox models was chosen following assessment of the log-log plots (Appendix O) and the KM curve.

HRs of [redacted] (95% CI: [redacted]) and [redacted] (95% CI: [redacted]) were estimated for the first ([redacted] years) and second ([redacted] years) intervals, respectively. It should be noted that for the primary analysis based on the RESPONSE-trials, as the treatment effect is applied to the predicted ruxolitinib OS curve to derive the survival for BAT, the inverse of the HR is used in the model (e.g. [redacted] and [redacted]).

Figure 47: KM for OS from the MAJIC-PV trial (based on the reconstructed pseudo-IPD)

[redacted]

Abbreviations: BAT: best available therapy; IPD: individual patient level datal; KM: Kaplan–Meier; OS: overall survival; PV: polycythaemia vera.
Source: Analysis of the reconstructed pseudo-IPD of the MAJIC-PV trial.[70]

Due to the uncertainty around the most appropriate cut-off point, scenario analyses were conducted exploring different cut-off points between [redacted] to [redacted] years. The impact on the cost-effectiveness results was modest (see Section B.3.11.3 and Appendix P). Additional scenario analyses were conducted using the constant treatment effect reported in the unpublished manuscript of the MAJIC-PV trial (0.73),[70] the treatment effect estimated from the MAIC (HR=[redacted]),[90] the treatment effect (HR=0.8) reported in Alvarez et al. 2022[112] and the treatment effect estimated from the RESPONSE-trials (presented for transparency) despite high and early cross-over (HR=[redacted]).[65, 66] Unsurprisingly, the treatment effect was a key driver of the cost-effectiveness results (see Section B.3.11.3 and Appendix P).

Duration of the treatment effect beyond the observed period

Treatment waning was included in the economic model in the base-case. A key area of uncertainty is the duration at which the treatment effect will be maintained beyond the observed period. The treatment effect over 5-years is directly available from the MAJIC-PV trial.[70] As patients remain on ruxolitinib for an extended duration, it is expected that the treatment effect will be maintained in the long-term beyond the observed period.[106] However, without additional follow-up, the duration of the treatment effect remains uncertain.

Another area of uncertainty is how the treatment would wane over time. Economic models typically assume a sudden change in the treatment effect, often leading to a sudden change in the hazard and predicted OS. To account for this, the treatment effect is allowed to diminish gradually linearly over time; e.g. the treatment effect diminishes linearly between Year 5 (end of observed period) and the time at which it is assumed to stop. Beyond this, no treatment effect is assumed. Therefore, the treatment effect beyond the observed period is not constant and diminishes gradually (linearly) over time until no treatment effect is assumed. This is likely to be more plausible.

Clinical experts were consulted to understand how long they would expect the treatment effect to last. In summary, clinical experts indicated that it is challenging to estimate how long the treatment effect would last and it will vary over time as only 5-year data is available.[106] Clinical experts indicated that patients on ruxolitinib would remain on treatment for an extended duration and would experience benefits.[106] Clinical experts indicated that assuming no treatment effect after 10 years was likely to be pessimistic, but it was challenging to define when the treatment effect would stop.[106] Clinical experts indicated that they expected approximately almost twice as many patients to be alive on ruxolitinib at Year 20 compared with BAT in the MAJIC-PV trial.[106] Consequently, in the base-case, to reflect clinical expectation, the treatment effect was assumed to stop at Year 20. By then, approximately 14% of patients on ruxolitinib are alive compared with 7% on BAT in the high-risk PV subgroup analysis based on the MAJIC-PV trial. The treatment effect assumed over time in the economic model in the base-case is shown below in Figure 48.

Owing to the uncertainty, extensive scenario analyses were conducted assuming waning at different time points. Unsurprisingly, the assumption around when the treatment effect stops was a key driver of the cost-effectiveness results (see Section B.3.10.3 and Appendix P).

Figure 48: Treatment effect assumed over time in the base-case

Source: Assumption.

B.3.3.5 Approach to partitioning the BAT health state for the primary analysis and data used

As highlighted in Section B.3.2.4, clinical advisors indicated that patients with PV R/I to HC/HU on BAT switch/cycle through BAT regimens over time with HRQoL progressively worsening. To reflect this aspect of the natural history of PV, the economic model splits the BAT health-state into three sub-health states in the primary analysis only (due to inflexibility of PSM in the subgroup analysis).

In the primary analysis, the BAT health state was partitioned (Figure 49) using two inputs:

- The time to first BAT treatment discontinuation (taken from the MAJIC-PV trial[70])
- The time to (all) BAT discontinuation (derived from the MAJIC-PV trial[70]), with:
  - The time in 1st BAT estimated from the AUC for the time to first BAT discontinuation
  - The time in 2nd BAT+ estimated as the difference between the time to first BAT discontinuation and time to BAT discontinuation
  - The time in no treatment (BAT exhaustion) is estimated as the difference between OS on BAT and time to BAT discontinuation.

This structural uncertainty was explored in scenario analysis by removing the partition in the primary analysis (due to the number of assumptions required); e.g. assuming a single BAT health state. The impact on the cost-effectiveness results was minor (see Section B.3.10.3 and Appendix P).

Figure 49: Partitioning the BAT health state

Abbreviations: BAT: best available therapy.

Time to discontinuation of first BAT regimen for patients initiated on BAT

In the primary analysis, the time to 1st BAT regimen discontinuation was taken from the unpublished manuscript of the MAJIC-PV trial.[70] Pseudo-IPD for the time to discontinuation of 1st BAT treatment from the MAJIC-PV trial were reconstructed based on the KM provided in the unpublished manuscript, with the KM curve presented in Figure 50. Parametric functions were fitted to the pseudo-IPD (Appendix N) with the selection process similar to that previously described in Section B.3.1.2.

Figure 50: KM for TTD of 1st BAT treatment in the MAJIC-PV trial (based on the reconstructed pseudo-IPD)

Abbreviations: BAT: best available therapy; IPD: individual patient level datal; KM: Kaplan–Meier; PV: polycythaemia vera; TTD: time to treatment discontinuation.
Source: Analysis of the reconstructed pseudo-IPD of the MAJIC-PV trial.[70]

In summary, although splines models with more than one knot provided a good visual to the KM, they tended to over-fit the data with unrealistic non-monotonic shapes during the observed period and were therefore not considered further (Appendix N). While the Gompertz distribution provided the best visual fit of the standard models and spline models with one knot, the visual fit was suboptimal. The Gompertz distribution also had the best statistical fit among these distributions. As the KM was relatively complete, the KM was used directly during the observed period, followed by the Gompertz. Scenario analyses were conducted using alternative distributions and parametric distributions for the entire period. Overall, the different extrapolation methods had a minor impact on the cost-effectiveness results (see Section B.3.11.3 and Appendix P).

Time to (all) BAT discontinuation for patients initiated on BAT

The KM for TTD for the BAT arm was not reported in the MAJIC-PV trial manuscript (only the KM for the time to discontinuation of first BAT regimen). However, the publication reported 17 deaths (e.g. 20% death) and 23 discontinuations (e.g. 27% discontinuation) in the BAT arm. Assuming no censoring, it is possible to approximate an HR between OS and TTD of [redacted]. It should be noted that this is an approximation based on the limited data reported as some patients were censored and therefore the true OS and TTD probability at 5 years would be lower compared with the probabilities used in this calculation. This was also based on a single time point.

In the economic model, in the primary analysis, the time to BAT treatment discontinuation was therefore estimated from the predicted BAT OS curve adjusted downward using a HR of [redacted]. Due to the uncertainty in this parameter, scenario analyses were conducted varying this parameter. The impact on the cost-effectiveness results was modest (see Section B.3.11.3 and Appendix P).

Partitioning of the BAT health state for patients initiated on ruxolitinib who move to BAT

In the primary analysis, for patients initiated on ruxolitinib who move to BAT, patients were redistributed into the three sub-health states, based on the distribution of patients in each health state at each cycle for patients initiated on BAT (Figure 49).

B.3.3.6 UK Life tables

Age- and gender-specific hazard rates of death were taken from published national life tables for England, using data for 2017–2019 in the base-case.[117] Life tables were used in the model to ensure the monthly hazard rate of mortality never falls below that of the general population. A scenario analysis was conducted using the life table using data for 2018–2020, however this is likely to be biased due to the effect of COVID-19.

For the primary analysis of the licensed population, age- and gender-specific hazard rate of death are used for the following:

- Extrapolation of the time to ruxolitinib discontinuation due to death (Section B.3.1.2).
- Extrapolation for the time to death following ruxolitinib discontinuation (Section B.3.3.2).

For the high-risk PV subgroup age- and gender-specific probabilities of death are used for the following:

- Extrapolation of the time to treatment discontinuation (Section B.3.1.2).
- Extrapolation for BAT OS (Section B.3.3.3).

B.3.3.7 Incidence of adverse events

Results of the Phase 3 RCTs (RESPONSE[65] & RESPONSE-2[66]) demonstrate that ruxolitinib is generally well tolerated in patients with PV. The overall pattern of AEs observed was also consistent across the RESPONSE trials and the MAJIC-PV trial. Exposure-adjusted AEs rates for ruxolitinib (428.4 and 334 patients-year exposure) and BAT (73.6 and 53 patients-year exposure) were reported in both the RESPONSE[65] and RESPONSE-2[66] trials, respectively. AEs occurring at a rate of ≥5 per 100 patient-years of exposure in either arm in RESPONSE and those occurring in ≥3 per 100 patient-years in RESPONSE-2 were included in the economic model as reported in the publications of key pivotal trials. These values (adjusted to the cycle length) were applied in the economic model to all model cycles according to the type of treatments received and the duration of treatment.

To fully capture the impact of AEs on both resource use and quality of life, AEs of any grade were considered in the economic model (with different impact according to Grade 1 and 2 AEs, and Grade 3 and 4 AEs). In the base-case, the rate of AEs from the RESPONSE[65] and RESPONSE-2[66] trials were pooled. These pooled event rates were used irrespective of the population considered (patients without and with splenomegaly) in the economic model (Table 23). Scenario analyses are presented using the rate of AEs from each respective trial. The impact on the cost-effectiveness results was minor (see Section B.3.10.3 and Appendix P).

Grade 1 and 2 AEs were not reported in the unpublished manuscript of the MAJIC-PV trial[70] and only AE categories experienced by ≥10% of patients were reported. The duration of exposure was also not reported. Consequently, in the economic model the pooled incidence of AEs estimated from the RESPONSE-trials[65, 66] were used for the high-risk PV subgroup analysis. Clinical experts considered this was reasonable.[106]

Table 23: Pooled exposure-adjusted rates (per 100 patient-years) of adverse events used in the economic analysis

	Ruxolitinib		BAT	
Adverse event	**G 1 and 2**	**G 3 and 4**	**G 1 and 2**	**G 3 and 4**
Anaemia	8.00	0.52	3.95	0.79
Arthralgia	5.38	0.39	7.11	1.58
Weight increased	5.25	0.66	1.58	0.00
Hypertension	2.89	1.31	2.37	3.16
Headache	4.85	0.26	22.91	0.79
Fatigue	4.33	0.26	15.80	2.37
Constipation	1.71	0.00	3.16	0.00
Bronchitis	3.28	0.26	5.53	0.00
Pyrexia	3.80	0.13	4.74	0.00
Pruritus	5.25	0.26	27.65	4.74
Pain in extremity	2.62	0.26	3.95	0.79
Back pain	3.67	0.13	3.95	0.00
Dyspnoea	1.44	0.00	1.58	0.00
Abdominal pain	3.28	0.26	11.06	0.00
Herpes zoster	3.54	0.52	0.00	0.00
Influenza	1.18	0.13	3.16	0.79
Oedema peripheral	1.31	0.00	1.58	0.00
Haematoma	1.31	0.00	0.79	0.00
Cystitis	1.31	0.00	0.00	0.00
Asthenia	0.92	0.13	3.95	0.79

Thrombocytosis	1.05	0.00	0.00	3.16
Dizziness	3.28	0.00	12.64	0.00
Nasopharyngitis	3.54	0.00	8.69	0.00
Diarrhoea	4.72	0.13	11.85	0.79
Cough	0.79	0.13	0.79	0.79
Night sweats	2.49	0.00	11.06	0.00
Thrombocytopenia	2.36	0.79	11.85	3.95
Upper respiratory tract infection	1.97	0.00	9.48	0.00
Leucocytosis	0.39	0.26	2.37	0.79
Decreased appetite	1.71	0.13	7.90	0.00
Abdominal pain upper	0.66	0.00	2.37	0.00
Myalgia	1.57	0.00	7.90	0.00
Dyspepsia	0.66	0.00	1.58	0.00
Nausea	2.36	0.13	7.11	0.00
Epistaxis	0.52	0.00	1.58	0.00
Oropharyngeal pain	0.39	0.00	2.37	0.00
Tinnitus	0.39	0.00	0.79	0.79
Toothache	0.39	0.00	1.58	0.00
Insomnia	0.39	0.00	1.58	0.00
Hyperhidrosis	0.39	0.00	1.58	0.00
Erythema	0.26	0.00	2.37	0.79
Weight decreased	0.13	0.00	3.16	0.00
Aphthous ulcer	0.13	0.00	1.58	0.79
Aquagenic pruritus	0.13	0.00	2.37	0.00
Abdominal discomfort	0.79	0.13	4.74	0.00
Iron deficiency	0.13	0.00	1.58	0.00
Rash	0.13	0.00	1.58	0.00
Palpitations	0.13	0.00	1.58	0.00
Early satiety	0.13	0.00	1.58	0.00
Haematocrit increased	0.00	0.00	3.16	0.79
Mouth ulceration	0.00	0.00	1.58	0.79
Cellulitis	0.00	0.13	3.16	1.58
Splenomegaly	0.00	0.00	1.58	0.00
Skin lesion	0.00	0.00	1.58	0.00
Increased blood thyroid stimulating hormone	0.00	0.00	1.58	0.00
Stomatitis	0.00	0.00	1.58	0.00
Urticaria	0.00	0.00	1.58	0.00
Muscle spasms	2.75	0.13	5.53	0.00
Musculoskeletal pain	0.92	0.13	3.16	0.00
Paraesthesia	0.92	0.00	5.53	0.00
Vertigo	0.92	0.00	3.16	0.00
Vomiting	0.79	0.00	3.16	0.00
Peripheral neuropathy	0.79	0.00	3.16	0.79
Bone pain	0.52	0.00	3.95	0.79
Hyperuricaemia	0.26	0.13	2.37	1.58
Gout	0.13	0.00	2.37	1.58
All infections	8.66	1.97	32.39	2.37

Abbreviations: AE, adverse events; BAT: best available therapy; G: grade.
Source: Table 1 in RESPONSE[65] and Supp Table 3 in RESPONSE-2[66] publications.

B.3.3.8 Incidence of events (TE, MF, AML/MDS, bleeding/haemorrhage, NMSC) for patients on ruxolitinib and treatment effects for patients on BAT

The treatment effects (applied as incidence-rate ratios [IRR]) were applied to the baseline incidence of events on ruxolitinib to estimate the incidence of events while on BAT. This is because 5-year data were available for ruxolitinib from both the RESPONSE-trials and the MAJIC-PV trial.

Exposure-adjusted incidence rate of events for patients on ruxolitinib

The per-cycle (28 days) incidence rates for the occurrence of TE, MF, AML/MDS, bleeding/haemorrhage and NMSC while on ruxolitinib treatment were calculated based on the number of events reported from the trials adjusted by the duration of exposure to ruxolitinib (or total follow-up time). Incidence rates used in the model for ruxolitinib are summarised in Table 24.

Table 24: Per cycle, adjusted-exposure rate of key events while on ruxolitinib

	MF	AML/MDS	TE	NMSC	Bleeding/haemorrhage
RESPONSE					
Number of events	9	1	5	22	NR
Exposure (PY)	428.4	428.4	428.4	428.4	428.4
IR (per PY)	2.10%	0.23%	1.17%	5.14%	NR
IR (per 28-day cycle)	0.16%	0.02%	0.09%	0.39%	0.73%*
RESPONSE-2					
Number of events	2	0	8	9	32
Exposure (PY)	334.3	334.3	334.3	334.3	334.3
IR (per PY)	0.60%	[redacted]	2.39%	2.69%	9.57%
IR (per 28-day cycle)	0.05%	[redacted]	0.18%	0.21%	0.73%
MAJIC					
Number of events	5	4	NR	14	NR
Exposure (PY)**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IR (per PY)	[redacted]	[redacted]	NR	[redacted]	NR
IR (per 28-day cycle)	[redacted]	[redacted]	0.13%*	[redacted]	0.73%*

*based on RESPONSE-2[66]; ** total follow-up time for OS from reconstructed IPD;[70] *** based on MAJIC-PV
Abbreviations: AML: acute myeloid leukaemia; IR: incidence rate; MDS: myelodysplastic syndrome; MF: myelofibrosis; NMSC: non-melanoma skin cancer; PY: patient year; TE: thromboembolic events.
Source: RESPONSE (Table 3, S5),[65] RESPONSE-2 (Table 2, S4, S6), [66] MAJIC-PV (Table S9B and S5).[70]

The incidence of events was taken from each respective trial, when possible, to reflect possible differences between populations. For instance, clinical advisors indicated that patients with PV with splenomegaly are more likely to develop MF compared with those without splenomegaly.[11]

As previously highlighted, none of the trials were designed or powered to evaluate the incidence of events. Despite 5-year follow-up, this is too short to robustly capture the incidence of events. The MAJIC-PV trial was also an investigator-led trial and therefore only information reported in the submitted manuscript for publication[70] could be used for the purpose of the economic model. Assumptions were required for the following items:

- No AML/MDS were reported in RESPONSE-2 in the ruxolitinib arm at 5 years.[66] This is likely due to the small sample size and short follow-up duration rather than a real effect. Consequently, the incidence estimated from the MAJIC-PV trial was used as a proxy.

- The total number of TE events was not reported in the MAJIC-PV manuscript.[70] Consequently, their incidence was assumed to be the same as RESPONSE-2.[66]
- The total number of bleeding/haemorrhage events were not reported in either the RESPONSE-trial[65] or the MAJIC-PV manuscript.[70] Consequently, their incidence was assumed to be the same as RESPONSE-2.[66].
- The number of events for patients on ruxolitinib was reported across the entire duration of the study, rather than only during the ruxolitinib treatment period as reported in the MAJIC-PV trial manuscript.[70] The exposure time was also not reported in the unpublished manuscript of the MAJIC-PV trial.[70] Therefore, the exposure time was approximated from the reconstructed pseudo IPD for OS for ruxolitinib from the MAJIC-PV trial publication.[70]

Treatment effects (incidence rate ratio) for key events for patients on BAT versus ruxolitinib

The treatment effects assumed in the economic model are summarised in Table 25. Due to the different follow-up durations, small number of events and the absence of clinical rationale for different treatment effects according to the population considered, the treatment effects (IRRs) were estimated by pooling the number of events from the RESPONSE-trials and the MAJIC-PV trial (adjusted for exposure time/follow-up) to increase the statistical power and reduce the uncertainty. Clinical experts considered this approach to be reasonable.[106]

Table 25: Treatment effects assumed for the key events

	MF	AML/MDS	TE	NMSC	Bleeding/haemorrhage
Incidence-rate ratios (BAT versus ruxolitinib)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AML: acute myeloid leukaemia; BAT: best available therapy; MDS: myelodysplastic syndrome; MF: myelofibrosis; NMSC: non-melanoma skin cancer; TE: thromboembolic events.
Source: RESPONSE (Table 3, S5),[65] RESPONSE-2 (Table 2, S4, S6), [66] MAJIC-PV (Table S9B and S5).[70]

B.3.3.9 Therapeutic venesection (phlebotomy)

In the base-case, the per-cycle 28-day rate of therapeutic venesection (also commonly referred to as phlebotomy) for ruxolitinib was derived from each trial ([redacted]% in RESPONSE, [redacted]% in RESPONSE-2 and [redacted]% in MAJIC).

The exposure time was not reported in the unpublished manuscript of the MAJIC-PV trial.[70] Furthermore, the total number of phlebotomy was reported during the entire study period, rather than during ruxolitinib treatment only; therefore, the total follow-up time estimated from the pseudo-IPD for OS was used.

A treatment effect was applied to the rate of venesection on ruxolitinib to calculate the per-cycle rate of venesection for patients on BAT. The treatment effect (IRR of [redacted]) was calculated by pooling the number of venesections across trials and exposure time for ruxolitinib and BAT to increase the sample size and statistical power.

B.3.4 *Measurement and valuation of health effects*

B.3.4.1 Health-related quality-of-life data from clinical trials

HRQoL was assessed using the EORTC QLQ-C30 in RESPONSE,[65] EQ-5D-5L in RESPONSE-2[66] and the MPN-SAF in both RESPONSE and RESPONSE-2.[65, 66] The MPN-SAF and EQ-5D-5L were collected in the MAJIC-PV trial.[70]

Lack of appropriateness of the EQ-5D in PV

The lack of appropriateness of the EQ-5D in some cases is recognised within the NICE methods guide.[25] It is stated that: "*In some circumstances the EQ-5D may not be the most appropriate measure. To make a case that the EQ-5D is inappropriate, provide qualitative empirical evidence on the lack of content validity for the EQ-5D, showing that key dimensions of health are missing. This should be supported by evidence that shows that EQ-5D performs poorly on tests of construct validity (that is, it does not perform as would be expected) and responsiveness in a particular patient population. This evidence should be derived from a synthesis of peer-reviewed literature. In these circumstances alternative health-related quality-of-life measures may be used. These must be accompanied by a carefully detailed account of the methods used to generate the data, their validity, and how these methods affect the utility values*."

Evidence from psychometric analysis in MF[23, 120, 121] indicates that the EORTC QLQ-C30 and EQ-5D do not adequately capture the key symptoms that impact quality of life in this condition including fatigue, early satiety, abdominal discomfort, inactivity, concentration, night sweats, itching, bone pain, fever and weight loss. As such, a condition preference-based measure (the MF-8D) was accepted by the Evidence Review Group and used as the basis for decision making by the NICE committee in TA386[23] and subsequently in TA756[107] for the assessment of ruxolitinib and fedratinib in MF.

Patients with PV experience several symptoms that are similar to MF, in particular fatigue, early satiety, abdominal discomfort, inactivity, concentration, night sweats, itching and bone pain. Given the similarity in symptoms between MF and PV,[122] the EQ-5D is likely to not be appropriate in adequately capturing the impact of PV on HRQoL.

To determine whether the EQ-5D is appropriate in PV, exploratory psychometric analyses were conducted using RESPONSE-2. Throughout the RESPONSE-2 trial, data on both the EQ-5D (5 levels) and MPN-SAF were collected at set intervals (baseline, weeks 4, 8, 16, 28, 52 and 80). The appropriateness of the EQ-5D was examined in terms of psychometric criteria of convergent validity, ceiling thresholds and responsiveness relative to the MPN-SAF in patients randomised to ruxolitinib (n=75). The specific tests examined whether the EQ-5D was related to PV-specific symptoms (convergent validity) and reflected changes in symptoms over time (responsiveness).

In summary, a large proportion ([redacted]%) of patients reported no problems in all 5 EQ-5D dimensions at baseline. The MPN-SAF total score did not show a comparable ceiling effect ([redacted]%). Similarly, the analysis suggests that the EQ-5D construct validity as measured by convergence is inconsistent across MPN-SAF domains at baseline. The EQ-5D preference based showed very strong convergence with the total MPN-SAF score, fatigue, inactivity, and problems with concentration at baseline; however, weak or moderate convergence was observed with respect to abdominal discomfort, fever, early satiety, night sweats, and weight loss. Correlation between the

self-care dimension of the EQ-5D and MPN-SAF dimension was weak to moderate. Regarding tests on whether the EQ-5D reflected changes in symptoms over time, the standardised change from baseline was calculated at Week 4, 8, 16, 28, 52 and 80 using the EQ-5D and MPN-SAF. In summary, although the EQ-5D captured some changes, these changes were much smaller than when assessed using the MPN-SAF. The standardised response mean (SRM) for the MPN-SAF total score was medium to large (>|0.5|) over time indicating that participants had large to medium improvement in PV key symptoms. In contrast, the SRM remained small to very small for the EQ-5D indicating that the EQ-5D is less appropriate at capturing changes in symptoms relative to the MPN-SAF (Figure 51).

Figure 51: Standardised response mean for the change in EQ-5D and MPN-SAF in RESPONSE-2

Notes: The analysis included patients randomised to the ruxolitinib arm who had complete EQ-5D-5L and MPN-SAF data at each visit. The following thresholds were used to categorise the standardised CFB: small: |≥0.2 to <0.5|, medium: |≥0.5 to <0.8|, and large: ≥|0.8|.
Abbreviations: CFB = change from baseline; EQ-5D-5L = EuroQol Five-Dimensional Questionnaire Five Level; MPN-SAF = Myeloproliferative Neoplasm - Symptom Assessment Form; TSS = total symptom score.

Overall, the exploratory psychometric analysis in RESPONSE-2 provides supportive evidence that the EQ-5D performs poorly on tests of construct validity and responsiveness in PV (smaller changes assessed using the EQ-5D). This is in line with the findings in MF[23] that supported the use of a condition specific-preference based measure (the MF-8D) in TA386 and TA756.[23, 107]

Derivation of the MF-8D

The MF-8D includes eight dimensions:

- Three dimensions from the EORTC QLQ-C30:
  - Physical functioning (ability to take walks);
  - Emotional functioning (worry);
  - Fatigue.
- Five dimensions from the MF-SAF:

- Pain under ribs on the left side;
- Abdominal discomfort;
- Night sweats;
- Pruritus (itchiness);
- Bone or muscle pain.

Whereas the MF-8D was developed using the MF-SAF, the RESPONSE trial collected data using the MPN-SAF. However, these questionnaires are very similar[123, 124] and therefore it was assumed that the MPN-SAF could be substituted for the MF-SAF with the following assumptions. These assumptions were validated by clinical experts.

- The "pain under ribs on the left side" dimension of the MF-SAF was equivalent to "abdominal pain" from the MPN-SAF;
- The "bone or muscle pain" dimension of the MF-SAF was equivalent to "bone pain" from the MPN-SAF.

The EORTC QLQ-C30 and MPN-SAF data from ruxolitinib- and BAT-treated patients in the RESPONSE trial (at baseline and Week 32) were used to estimate treatment-specific utility values using the MF-8D using the recommended algorithm (RE MLE consistent model) described in Mukuria et al. 2015.[121]

The following utility values were estimated at baseline and Week 32 using regression analysis including treatment and baseline MF-8D as covariates:

- Baseline = [redacted]
- Ruxolitinib = [redacted] (e.g. change from baseline of [redacted])
- BAT ('Off ruxolitinib') = [redacted] (e.g. change from baseline of [redacted]).

Although there is evidence to suggest that the EQ-5D performs poorly in PV and conditions-alike (such as MF) on tests of construct validity and responsiveness when compared with condition-specific questionnaires,[23, 120, 121] a scenario analysis was also conducted in which health-state utility values were derived using EQ-5D-5L data from the RESPONSE-2 trial for transparency and completeness (ruxolitinib utility = [redacted] and BAT utility = [redacted]).

B.3.4.2 Mapping

The MF-8D derived from RESPONSE[65] was used in the base-case, with the EQ-5D derived from RESPONSE-2[66] used in scenario analysis.

In accordance with the NICE methods guide,[25] EQ-5D-5L data were converted to EQ-5D-3L using the algorithm published by Hernandez et al. 2019.[125]

B.3.4.3 Health-related quality-of-life studies

As detailed in Appendix H.1.3, two studies were identified in the economic SLRs that reported utility estimates for patients with PV using the EQ-5D. However, these were not specific to patients with PV R/I to HC/HU, and utility values were reported for patients with PV with (0.8) and without

(0.8) aquagenic Pruritus in Poland[126] or by JAK2 status (<50; ≥50) in the PROUD study[127] (0.881 versus 0.876).

As quality of life was collected in the RESPONSE-trials, utility values for the economic model were taken directly from estimates from the trial, as detailed in Section B.3.4.1.

B.3.4.4 Adverse reactions

The impact of AEs on HRQoL are likely to be short lived and therefore unlikely to be captured in the utility values used in the economic evaluation assessed at a single time point. Consequently, in the base-case the impact of AEs on HRQoL are captured in the model separately. The health disutility associated with a particular AE was calculated based on the health utility decrement of a particular AE and duration of the impact of the AE on quality of life.

For simplicity, no health disutility was assumed for Grade 1 or 2 AEs. Disutilities for Grade 3 or 4 AEs and their duration were sourced from the literature or values used in previous NICE appraisals when available and appropriate. A disutility of −0.075 lasting seven days was assumed for those Grade 3 or 4 AEs that could not be sourced from the literature or previous NICE appraisals based on results from a multivariate regression model used in NICE TA772.[128] The health disutility associated with a particular AE and its duration used in the economic model are summarised below in Table 26.

Table 26: Adverse events disutilities and durations

	Disutilities	Duration (days)	Source
Anaemia	−0.08	14.78	NICE TA772 (Table 121 CS[128])
Arthralgia	−0.07	18.70	NICE TA722 (Table 45 CS[129])
Weight increased	−0.075	7.00	NICE TA772 (Table 121 CS[128])
Hypertension	−0.075	7.00	Assumption - AE covariate from the multivariate model used in NICE TA772[128]
Headache	−0.085	2.625	Assumed to be the same fatigue
Fatigue	−0.085	2.625	NICE TA722 (Table 45 CS[129])
Bronchitis	−0.075	7.00	Assumption - AE covariate from the multivariate model used in NICE TA772[128]
Pyrexia	−0.075	7.00	Assumption - AE covariate from the multivariate model used in NICE TA772[128]
Pruritus			
Pain in extremity			
Back pain			
Abdominal pain			
Herpes zoster	−0.09	8.30	Assumed to be the same infection
Influenza			
Asthenia	−0.075	7.00	Assumption - AE covariate from the multivariate model used in NICE TA772[128]
Thrombocytosis	−0.108	15.94	NICE TA772 (Table 121 CS[128])
Diarrhoea	−0.063	5.53	NICE TA772 (Table 121 CS[128])
Cough	−0.075	7.00	Assumption - AE covariate from the multivariate model used in NICE TA772[128]
Thrombocytopenia	−0.108	15.94	NICE TA772 (Table 121 CS[128])

Leucocytosis	−0.108	15.94	Same as thrombocytopenia
Decreased appetite	−0.075	7.00	Assumption - AE covariate from the multivariate model used in NICE TA772[128]
Nausea	−0.075	7.00	NICE TA772 (Table 121 CS[128])
Tinnitus	−0.075	7.00	Assumption - AE covariate from the multivariate model used in NICE TA772[128]
Erythema			
Aphthous ulcer			
Abdominal discomfort			
Haematocrit increased			
Mouth ulceration			
Cellulitis			
Muscle spasms			
Musculoskeletal pain			
Peripheral neuropathy	−0.33	76.00	NICE TA772 (Table 121 CS[128])
Bone pain	−0.075	7.00	Assumption - AE covariate from the multivariate model used in NICE TA772[128]
Hyperuricaemia			
Gout			
All infections	−0.085	8.30	NICE TA722 (Table 45 CS[129])

Abbreviations: AE: adverse event; CS: company submission; NICE: national institute of health; TA: technology appraisal.

Scenario analyses were conducted removing the impact of AEs on quality of life. As expected, the impact on the cost-effectiveness results was modest (see Section B.3.10.3 and Appendix P).

B.3.4.5 HRQoL data used in the cost-effectiveness analysis

General population utility values

In line with the NICE methods guide, age-specific multipliers were included to account for the reduction in quality of life as patients get older based on the utility values by age and gender reported by Hernandez et al. 2022.[130]

QALY losses associated with the key events

The calculation of the QALY losses for events considered in the economic model (TE, AML/MDS, MF, NMSC, bleeding/haemorrhage) are summarised below in Table 27. Simplifications were made due to the challenge in accurately estimating QALY losses for these events.

Due to the uncertainty, QALY losses were halved and double in sensitivity analysis. The impact on the cost-effectiveness results was minor.

QALY losses associated with venesection

A QALY loss of −0.0000103 per venesection was assumed, based on a decrement in utility of −0.037 taken from Matza et al. 2013 and an assumption that the decrement lasts a day.

Table 27: QALY loss assumed in the economic evaluation

	AML/MDS	MF*	TE	Bleeding/haemorrhage	NMSC
QALYs losses	−0.197	−0.063	−0.043	−0.007	0
Approach	Calculated based on the survival of patients with 'secondary AML' and the difference in utility value between patients with AML' and those with PV without AML;	Calculated based on % of patients with int-2/high-risk MF, % treated with ruxolitinib for MF, the survival for MF patients on BAT, the utility values in MF, and incremental QALYs with ruxolitinib	Calculated based on decrement in utility values for cardiovascular complications.	Calculated based on % and QALY loss for minor/major bleeding	Gorry et al. 2018 reported there was no evidence of NMSC impacting utility.[131]
Data source1	**Survival AML:** Mean survival of 10 months assumed based on median survival (7.0 months) reported in Dai chichara et al. 2016[132] and assume exponentially distributed	**% int-2/high-risk & % treated with ruxolitinib:** Informed by Mead et al. 2022.[133] 57% of MF events are int-2/high-risk and 23% receive ruxolitinib.[133]	Briggs et al. 2017[134] reported a decrement in utility value of −0.057 <3 months, −0.043 3–6 month and −0.035 between 6–12 month.	**% major bleeding:** Taken from RESPONSE-2 (18%).	
Data source2	**Utility value AML:** Utility value of 0.53 taken from Tolley et al. 2010[135]	**Survival BAT int2/high-risk:** 4.04 years[23] **Incremental QALYs ruxolitinib versus BAT:** 2.45 QALYs[23]		**QALY loss:** Taken from Doble et al. 2018[136] for minor −0.0025 and major bleed −0.0297	
Data source3		**Utility value for MF:** 0.71 based on Mesa et al. 2021.[137] Assumed to be the same irrespective of risk categories			
Data source4		**Survival MF low/int-1 risk:** median of 17.5 and 7.8 years based on Tefferi et al. 2012[138] – mean (exponential)			

*It should be noted the management of MF in the UK varies according to the risk category with patients with intermediate-2/high-risk (Int-2/high-risk) MF secondary to PV treated with ruxolitinib as recommended in TA386,[23] while those with low or intermediate-1 risk (low/int-1 risk) receive BAT (consisting mostly of HC/HU and watch and wait). Fedratinib is recommended after ruxolitinib failure, but currently in the Cancer Drugs Fund and therefore not considered.[107]
Abbreviations: AML: acute myeloid leukaemia; BAT: best available therapy; int-1: intermediate-1; int-2: intermediate-2; MDS: myelodysplastic syndrome; MF: myelofibrosis; NMSC: non-melanoma skin cancer; QALY; quality adjusted life year; TE: thromboembolic events.

Summary of utility values in the cost-effectiveness analysis

The utility values used in the cost-effectiveness model are presented in Table 28. As highlighted in Section B.3.3.5, the BAT health state was partitioned into three sub-health states in the primary analysis. A reduction in utility was assumed as patients switch treatment in the primary analysis and move to the 2nd BAT sub-health state. For patients moving to the no treatment health state, a small decrement on -0.05 was assumed to reflect that patient are no longer treated and experience a decline in HRQoL. In TA756, a decrement in utility of approximately -0.2 was assumed for patients in supportive care (utility value of 0.53). This structural uncertainty was explored in scenario analysis by removing the partition in the primary analysis. The impact on the cost-effectiveness results was minor (see Section B.3.10.3 and Appendix P).

Furthermore, the MF-8D derived from RESPONSE[65] was used in the base-case, with the EQ-5D derived from RESPONSE-2[66] used in scenario analysis given problems of the EQ-5D to adequately capture symptoms in PV. Unsurprisingly, utility values were a key driver of the cost-effectiveness results.

Table 28: Summary of utility values for cost-effectiveness analysis

State	Utility value: mean (SE)	95% CI	Reference in submission (section and page number)	Justification
Main health state				
Baseline utility value	[redacted]	[redacted]	Section B.3.4.1	Change in utility value using MF-8D from baseline calculated from the RESPONSE trial using regression.[65] EQ-5D not appropriate in PV.
On ruxolitinib: Change from baseline	[redacted]	Estimated from a regression model (Variance covariance matrix used in PSA)	Section B.3.4.1	
On 1st BAT (1st BAT sub-health state): Change from baseline	[redacted]		Section B.3.4.1	
On 2nd BAT+ (2nd BAT sub-health state): Change from baseline	[redacted]		Assumption	Primary analysis only: Additional decrements in utility values are assumed as patients move through 2nd BAT+ sub health states. Assumption for no treatment
No treatment (3rd BAT sub-health state)	–0.05			
QALY loss for key events				
AML/MDS	−0.197	NA	Table 27	
MF	−0.063	NA	Table 27	
TE	−0.043	NA	Table 27	
Bleeding/haemorrhage	−0.007	NA	Table 27	
NMSC	0	NA	Table 27	
Venesection	−0.0000103	NA	Table 27	

Abbreviations: AML: acute myeloid leukaemia; BAT: best available therapy; CI: confidence interval; MDS: myelodysplastic syndrome; MF: myelofibrosis; NMSC: non-melanoma skin cancer; PV: polycythaemia vera; QALY; quality adjusted life year; SE: standard error; TE: thromboembolic events.

B.3.5 *Cost and healthcare resource use identification, measurement and valuation*

Costs considered in the economic model included treatment costs, costs associated with the management of PV/monitoring associated with treatments, costs associated with the management of key events, costs associated with the management at the end of life, and the costs associated with the management of AEs. Treatment costs included both drug acquisition and administration costs.

B.3.5.1 Intervention and comparators' costs

Drug acquisition and administration costs for treatments included in this economic evaluation are summarised in Table 29. The 4-weekly treatment costs (including administration) for patients treated with ruxolitinib were estimated to be £[redacted] for RESPONSE, [redacted] for RESPONSE-2 and [redacted] for the subgroup of high-risk patients (assumed to be the same as RESPONSE-2 in the absence of data on dosage distribution in the MAJIC-PV trial publication) using the current PAS for MF ([redacted]The 4-weekly treatment costs for patients treated with BAT (including administration) was estimated at £226.48 for all analyses. A one-off administration cost of £24.71 was also assumed to reflect training for the first administration of IFN-alfa.

Drug acquisition costs

The list price for the intervention (ruxolitinib) and treatments that are part of BAT (IFN-alfa (peg-interferon alfa [Pegasys] – subcutaneous form) and busulfan (oral form)) were taken from the British Natural Formulary (BNF)[24] (Table 29). Since HC/HU and anagrelide are available to the NHS as a generic medicine, costs were calculated from the Electric Market Information Tool (eMIT) based on the number of prescriptions. Prednisolone is often given in patients receiving IFN-a. The cost of prednisolone (as well as that of aspirin) was not considered within the economic analysis as it is minimal.

Ruxolitinib is currently provided to the NHS at a confidential discount off the current NHS list price ([redacted]%) for the treatment of MF. Based on the need for [redacted] to allow a cost-effective price to be offered for this appraisal, Novartis are working [redacted] (subject to NICE's assessment). [redacted]results are presented in the main body of this submission using the PAS discount agreed for MF in line with the NICE method guide.[25][redacted]

Dosing schedule assumed in the economic model

The dosing schedules assumed for treatments included in the economic model are presented in Table 29. This was based on discussion with clinical experts and/or published sources.

Table 29: Summary of treatment costs used in the economic model

	Dosing schedule	Vial/pack concentration and volume	Number of tablets/vial	Cost per vial/pack (NHS list price)	Cost per vial/pack (PAS)	Number of packs/vials required per 28 days cycle	Cost per 28 days cycle	Admin costs	Source
Ruxolitinib	Daily based on dose in RESPONSE Trials	5 mg	56	£1,428	[redacted]	Dose based on RESPONSE trials[65, 66]	N/A	£0	BNF RESPONSE trials[65, 66]
		10 mg		£2,856	[redacted]				
		15 mg		£2,856	[redacted]				
		20 mg		£2,856	[redacted]				
IFN-alfa (pegasys)	One weekly 128.7 MU[139]	135 MU / 0.5ml	1	£107.76	N/A	4 syringes	£431.04	£15.55 (help with injection) £24.71 (on-off training)	BNF; Yacoub et al. 2019[139]
HC/HU	Daily 1,250 mg (expert opinion)	500 mg	100	£9.54	N/A	0.7 packs	£6.68	£0	eMIT[24, 116]
Anagrelide	2 mg daily	0.5 mg	100	27.16	N/A	1.12 packs	£30.42	£0	eMIT[116]; NHS Thames Valley[140]
Busulfan	1 mg daily**	2	25	£41.73	N/A	0.56 packs	£23.37	£0	BNF[24]

* using confidential PAS discount for MF; ** as busulfan may not be given continuously, the dose (2mg daily) was halved to avoid over estimating costs (busulfan is given as start and stopped)

Abbreviations: admin: administration; eMIT: electronic market information tool; HC/HU: hydroxycarbamide/hydroxyurea; IFN-a: interferon alfa; MF: myelofibrosis; MU: microgram; N/A: not applicable; PAS: patient access scheme; SD: standard deviation; TA: technology appraisal.

Treatment received

As highlighted in Section B.3.2.4, the BAT composition reported in the MAJIC-PV trial manuscript[70] was used to reflect both clinical practice in England and Wales and align costs with effectiveness data used (Table 20). Pipobroman and Phosphorus-32 (p32) were excluded as they are no longer used in England and Wales. Similarly the cost for subsequent ruxolitinib given in the BAT arm were excluded to avoid over-estimating costs.

Dose intensity/reduction

Dose intensity/dose reduction was included in the economic model when possible:

- The relative dose intensity and relative dose interruption for ruxolitinib was calculated using the same approach that was used for decision making in TA386,[23] based on the number of days patients were treated with different doses in the RESPONSE-trials and the cost per day according to dosage (Table 30). This approach accounts for both dose reductions and interruptions and reflects the dosage used in the RESPONSE-trials from which efficacy data is taken from. A scenario analysis was conducted assuming no dose reduction/interruption, where ruxolitinib is given according to the recommended licensed dose (10 mg BID). The impact on the cost-effectiveness was modest. The dosage distribution was not reported in the MAJIC-PV manuscript,[70] therefore in the base-case, the dosage distribution from RESPONSE-2 was used as a proxy.
- Pegasys® comes in the form of 90 MU (syringe for single patient use), 135 MU (syringe for single patient use) and 180 MU (four syringes). Clinical experts indicated that Pegasys® syringes cannot be re-used.[106] Therefore, if the prescribed dosage is lower than the size of the syringe, the rest of the syringe needs to be discarded. Clinical experts noted that the dosage of Pegasys® varies between 45 MU to 180 MU.[106] In this economic evaluation, patients treated with IFN-afa were assumed to receive Pegasys® 135 MU once weekly based on the mean dosage (128.7 MU once weekly) reported in Yacoub et al. 2019[139] in patients with PV R/I to HC/HU. Clinical experts considered this was a reasonable proxy.[106]

Table 30: Number of days treated with different dosage in the RESPONSE Trial and assumption on costing used in the economic model

Daily Dose	Total Daily	Number of	Frequency	Cost per day
0 mg	0	[redacted]	[redacted]	[redacted]
5mg BID	10	[redacted]	[redacted]	[redacted]
5mg BID	10	[redacted]	[redacted]	[redacted]
10mg BID	20	[redacted]	[redacted]	[redacted]
10mg BID	20	[redacted]	[redacted]	[redacted]
15mg BID	30	[redacted]	[redacted]	[redacted]
15mg BID	30	[redacted]	[redacted]	[redacted]
20mg BID	40	[redacted]	[redacted]	[redacted]
25mg BID	50	[redacted]	[redacted]	[redacted]
5mg QD	5	[redacted]	[redacted]	[redacted]
10mg QD	10	[redacted]	[redacted]	[redacted]
10mg + 5mg	15	[redacted]	[redacted]	[redacted]
10mg + 15mg	25	[redacted]	[redacted]	[redacted]
15 QD	15	[redacted]	[redacted]	[redacted]
20 QD	20	[redacted]	[redacted]	[redacted]

15mg + 20mg	35	[redacted]	[redacted]	[redacted]
20mg + 25mg	45	[redacted]	[redacted]	[redacted]
25 QD	25	[redacted]	[redacted]	[redacted]

Abbreviations: BID: twice daily; mg: milligram; QD: once daily.
Source: analysis of RESPONSE IPD.[65]

Table 31: Number of days treated with different dosage in the RESPONSE-2 Trial and assumption on costing used in the economic model

Daily Dose	Total Daily	Number of days	Frequency	Cost per day
0 mg	0	[redacted]	[redacted]	[redacted]
2.5mg BID	5	[redacted]	[redacted]	[redacted]
2.5mg QD	2.5	[redacted]	[redacted]	[redacted]
5mg BID	10	[redacted]	[redacted]	[redacted]
5mg QD	5	[redacted]	[redacted]	[redacted]
7.5mg BID	15	[redacted]	[redacted]	[redacted]
7.5mg QD	7.5	[redacted]	[redacted]	[redacted]
10mg BID	20	[redacted]	[redacted]	[redacted]
10mg QD	10	[redacted]	[redacted]	[redacted]
12.5mg BID	25	[redacted]	[redacted]	[redacted]
15mg BID	30	[redacted]	[redacted]	[redacted]
15mg QD	15	[redacted]	[redacted]	[redacted]
17.5mg BID	35	[redacted]	[redacted]	[redacted]
20mg BID	40	[redacted]	[redacted]	[redacted]
20mg QD	20	[redacted]	[redacted]	[redacted]
22.5mg BID	45	[redacted]	[redacted]	[redacted]
25mg BID	50	[redacted]	[redacted]	[redacted]
25mg QD	25	[redacted]	[redacted]	[redacted]
30mg QD	30	[redacted]	[redacted]	[redacted]
40mg QD	40	[redacted]	[redacted]	[redacted]

Abbreviations: BID: twice daily; mg: milligram; QD: once daily.
Source: analysis of RESPONSE-2 IPD.[66]

Drug administration costs

Ruxolitinib is an oral treatment; therefore, no administration cost was assumed.

Most patients treated with IFN-alfa (92.5%) were assumed to self-inject (or with the help of a family member) and therefore did not incur any administration costs. Clinical experts noted that between 5% to 10% of patients with PV require nurse administration. Unit costs associated with injection administration was taken from the NHS reference cost [district nurse – face to face N02AF]. Clinical experts further noted that patients would require one or two nurse visits or general practitioner (GP) appointments at the start to train patients on how to self-inject. To reflect this, a one-off cost £24.71 is applied upfront, assuming [redacted]% of patients receive IFN-alfa based on the proportion of patients receiving IFN-alfa in the MAJIC-PV trial and the unit cost associated with a district nurse visit.

B.3.5.2 Health-state unit costs and resource use

Management/monitoring associated with PV

The economic SLR did not identify any UK cost studies or NICE appraisals for PV (Appendix G, Appendix I). Therefore, resource utilisation data for the management of PV in the UK are not available.

To address this data gap, five UK clinical experts with expertise in the management of patients with PV were asked to provide estimates for 6-month intervals i.e. during months 1–6 of treatment, during months 7–12 of treatment, and 13+ months. As the economic model already separately includes the resource use associated with venesection (Section B.3.5.2), the management of adverse events (Section B.3.5.3), the management of TE, MF, AML/MDS, NMSC and haemorrhage separately (Section B.3.5.2), clinical experts were asked to estimate the frequency of resource use associated exclusively with the management and monitoring of PV. For simplicity, the same resource use was assumed for all populations.

Five out of five questionnaires were correctly completed by the clinical experts. As expected, there was some variation between responses given the rare nature of the disease and different practice. This economic analysis used the average of the frequency of estimates provided by the experts (Table 32). Scenario analyses are conducted using the resource use estimated by each expert individually. The impact on the cost-effectiveness results was modest.

Unit costs were derived from the NHS reference costs 2020/2021,[114] and PSSRU[115] published costs. The expected calculated per-28-day-cycle cost was estimated to be [redacted] and [redacted] for patients receiving ruxolitinib (ruxolitinib health state) between 0–6 months, 7–12 months and 13 months+, respectively. The calculated per-28-days cycle cost for patients treated with BAT (initiated on BAT or after discontinuation from ruxolitinib) was estimated to be [redacted], and [redacted]. For patients moving to the "no treatment" sub-health state, a management cost of [redacted] was assumed (doubled that of patients on BAT). This is to reflect the higher management expected when patients stop treatment and the disease is uncontrolled. This structural uncertainty was explored in scenario analysis by removing the partition in the primary analysis (due to the number of assumptions required); The impact on the cost-effectiveness results was minor (see Section B.3.10.3 and Appendix P).

Cost associated with terminal care

A one-off cost of £6,774 for terminal/palliative care was applied within the model at the point of death taken from the cost for health and social care reported in Round et al. 2015 and inflated to 2021 costs using the PSSRU inflation indices.[115, 141]

Cost associated with therapeutic venesection

A unit cost of £316 per venesection was assumed based on the weighted cost of a day case (SA07G-J) taken from the NHS reference cost 2020/2021.[114]

Management costs for key events included in the economic model

The cost associated with the management of key events is summarised below in Table 33.

Table 32: Estimated per cycle resource use and unit costs

Resource	On ruxolitinib			On BAT			Unit cost	Source
	0–6 mth	7–12 mth	13–18 mth	0–6 mth	7–12 mth	13–18 mth		
Primary/Community care visits								
GP - Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£39.23	PSSRU (2021).[115] Per surgery consultation lasting 9.22 minutes. p120
GP - Home visit	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£90.83	PSSRU (2021).[115] Assume 12 minutes for travel. p120
Community nurse	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51.84	NHS ref cost 2020/2021.[114] district nurse (N02AF)
Cancer nurse visit	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£90.49	NHS ref cost 2020/2021.[114] Specialist Nursing, Cancer Related, Adult, Face to face (N10AF)
Pain and symptom management	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£126.12	NHS ref Costs 2020-2021.[114] Community Health Services, (N21AF): specialist nursing,
Depression management	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£115.22	NHS ref Costs 2020-2021.[114] Community Health Services, Allied Health Professionals (A06A1):
Hospitalisation ED/ICU Outpatient visits								
Outpatient visit	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£214.56	NHS ref Costs 2020-2021.[114] Outpatient attendance data, Consultant Led (face to face - Follow up)
ED use	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£296.88	NHS ref Costs 2020-2021.[114] Emergency medicine. VB01Z, VB04Z, VB05Z, VB07Z, VB08Z
Hospitalisation days	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£311.98	NHS Reference Costs 2020-2021.[114] SA07G, SA07H, SA07J (regular day/night admission)
Imaging/Tests								
Bone marrow biopsy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£584.89	NHS ref Costs 2020-2021.[114] SA33Z: Diagnostic Bone Marrow Extraction (outpatient procedure)
ECG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£222.95	NHS ref Costs 2020-2021.[114] EY50Z: Complex Echocardiogram (outpatient procedure)
Ultrasound scan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£80.80	NHS ref Costs 2020-2021.[114] RD14Z: Ultrasound Scan with duration of less than 20 minutes, with Contrast
Blood test	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£3.63	NHS ref Costs 2020–2020.[114] Directly Accessed Pathology Services, Haematology, DAPS05

Abbreviations: CT: computerised tomography; ECG: electrocardiogram; ED: emergency department; F2F: face to face; GP: general practitioner; HRG: Healthcare Resource Group; ICU: intensive care unit; MRI: magnetic resonance imaging; mth: month; NHS: national health service; PSSRU: Personal Social Services Research Unit.

Table 33: Management cost assumed for key events

	MF	AML	TE	NMSC	Bleeding/haemorrhage
Cost per event	£63,920	£44,903	£1,302	£1,058	£1,929

Abbreviations: AML: acute myeloid leukaemia; MDS: myelodysplastic syndrome; MF: myelofibrosis; NMSC: non-melanoma skin cancer; TE: thromboembolic events.

The impact on the cost-effectiveness results was minor.

Management costs of MF events

A total cost of £63,920 was assumed for the management of MF events.

The cost associated with management of intermediate-2/high-risk MF (57.3%) was derived from TA386[23] based on the total costs reported for patients treated with ruxolitinib (£128,403) or BAT (£36,095), including confidential discounts and the proportion of patients who would receive ruxolitinib reported in Mead et al. 2022.[133]

No data are available on the management costs for patients with low/intermediate-1-MF (assumed to be 42.7% of MF events). Although these patients may require less resource use per unit of time compared with intermediate-2/high-risk MF patients, the prognosis of low/intermediate-1-MF patients is more favourable and therefore overall resource use is likely to be higher due to the increased duration for which these patients receive care. In the base-case it was arbitrarily assumed that the management cost for low/intermediate-1 risk was double that of intermediate-2/high-risk treated with BAT in TA356 (e.g. £72,190). This uncertainty was explored in sensitivity analysis and the impact on the cost-effectiveness was minor.

Management costs of AML/MDS events

The cost associated with the management of AML/MDS in the UK was taken from results of a probabilistic decision model in AML by Wang et al. 2014,[142] used in NICE TA386[23] and more recently NICE TA756.[107] Medical costs were calculated using a micro-costing approach and included costs associated with treatment, hospitalisations, diagnostic tests, transfusions and associated complications. The authors estimated the 5-year medical costs for the management of AML in the UK to range between £8,170 and £81,636 and the life expectancy per patient to range between 3.03 to 34.74 months. In the base case, a one-off cost of £44,903 (middle range of the cost reported) was used in line with TA386.[23]

Management costs for TE events

The cost associated with the management of a TE event was assumed to be £1,302 based on the distribution of TE events reported in RESPONSE and RESPONSE-2 in the ruxolitinib arm, grade level and respective unit costs. This is a simplification as the distribution of TE events in the trials is unlikely to reflect the true distribution of events due to the small number of events.

The cost associated with one ER visit (£182)[114] was assumed for the management of all Grade 1–2 TE events. Unit costs for the management of Grade 3-4 thromboembolic events (Table 34) were sourced from the NHS reference costs 2020/21.[114]

Table 34: Unit costs for the management of grade3/4 thromboembolic events

	Cost	Source
CI	£3,128	NHS Reference cost 2020/2021[114]: HC Disorders (AA23C-G)
AMI/PAD	£1,596	NHS Reference cost 2020/2021[114]: Actual or Suspected MI (EB10A-E)
CA	£2,726	NHS Reference costs 2020/2021[114]:CA,NSI or Encephalopathy(AA22C-G)
PVT	£3,692	NHS Reference cost 2020/2021[114]: PTA (YR10A-15C)
PE	£1,498	NHS Reference cost 2020/2021[114]: Pulmonary Embolus (DZ09J-Q)

Abbreviations: AMI: Acute myocardial infarction; CA: Cerebrovascular accident; CC: complexity and comorbidity; CI: Cerebral infarction; HC: Haemorrhagic Cerebrovascular; MI: myocardial infarction; NHS: national health service; NSI: Nervous System Infections; PAD: peripheral artery disease; PE: Pulmonary embolism; PTA: Percutaneous Transluminal Angioplasty; PVT: Portal vein thrombosis.

Management costs of NMSC events

The cost associated with the management of NMSC was assumed to be £1,058 per case based on Vallejo-Torres et al. 2013.[143] The authors estimated the mean cost per case of NMSC in England to be £889 and £1,226, using a bottom-up and top-down approach, respectively.

Management costs for bleeding/haemorrhage events

The cost associated with the management of a major bleed was assumed to be £9,788 based on Carthorne et al. 2018.[144] The cost associated with the management of a minor bleed was assumed to be the cost of one ER visit (£182).[114]

B.3.5.3 Adverse reaction unit costs and resource use

The cost of two GP e-consultations was assumed for the management of Grade 1 and 2 AEs taken from the PSSRU 2021.[115] Costs associated with the management of Grade 3 and 4 AEs (Table 35) were sourced from the NHS reference costs 2020/21.[114].

Table 35: Adverse events costs

AEs	Unit cost	Source
Anaemia	£1,699	NHS Reference cost 2019/2020: Haemolytic Anaemia (SA03G-H)
Arthralgia	£1,366	NHS Reference cost 2020/2021[114]: Musculoskeletal Signs or Symptoms (HD26D-G)
Weight increased	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
Hypertension	£919	NHS Reference cost 2020/2021[114]: Hypertension (EB04Z)
Headache	£889	NHS Reference cost 2020/2021[114]: Headache, Migraine or Cerebrospinal Fluid Leak (AA31C-E)
Fatigue	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
Bronchitis	£993	Assumed to be same as cough
Pyrexia	£1,495	NHS Reference cost 2020/2021[114]: Fever of Unknown Origin with or without interventions (WJ07A-D)
Pruritus	£2,069	NHS Reference cost 2020/2021[114]: Skin Disorders with or without interventions (JD07A-K)
Pain in extremity	£1,366	NHS Reference cost 2020/2021[114]: Musculoskeletal Signs or Symptoms (HD26D-G)
Back pain	£1,296	NHS Reference cost 2020/2021[114]: Low Back Pain with or without interventions (HC32G-K)

Abdominal pain	£868	NHS Reference cost 2020/2021[114]: Abdominal Pain with or without interventions (FD05A-B)
Herpes zoster	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
Influenza	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
Asthenia	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
Thrombocytosis	£1,070	NHS Reference cost 2020/2021[114]: Thrombocytopenia (SA12G-K)
Diarrhoea	£2,011	NHS Reference cost 2019/2020: Non-Malignant Gastrointestinal Tract Disorders (FD10A-M)
Cough	£993	NHS Reference cost 2020/2021[114]: Other Respiratory Disorders with or without interventions (DZ19H-M)
Thrombocytopenia	£1,070	NHS Reference cost 2020/2021[114]: Thrombocytopenia (SA12G-K)
Leucocytosis	£1,070	same as thrombocytopenia
Decreased appetite	£2,094	NHS Reference cost 2020/2021[114]: Nutritional Disorders with or without interventions (FD04A-E)
Nausea	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
Tinnitus	£1,374	Assumed to be same as mouth ulceration
Erythema	£2,069	NHS Reference cost 2020/2021[114]: Skin Disorders with or without interventions (JD07A-K)
Aphthous ulcer	£1,374	NHS Reference cost 2020/2021[114]: Non-Malignant, Ear, Nose, Mouth, Throat Disorders (CB02A-F)
Abdominal discomfort	£868	NHS Reference cost 2020/2021[114]: Abdominal Pain with or without interventions (FD05A-B)
Haematocrit increased	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
Mouth ulceration	£1,374	NHS Reference cost 2020/2021[114]: Non-Malignant, Ear, Nose, Mouth, Throat Disorders (CB02A-F)
Cellulitis	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
Muscle spasms	£868	NHS Reference cost 2020/2021[114]: Abdominal Pain with or without interventions (FD05A-B)
Musculoskeletal pain	£1,366	NHS Reference cost 2020/2021[114]: Musculoskeletal Signs or Symptoms (HD26D-G)
Peripheral neuropathy	£734	NICE TA772[128]
Bone pain	£1,627	NHS Reference cost 2020/2021[114]: Non-Inflammatory, Bone or Joint Disorders (HD24D-H)
Hyperuricaemia	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
Gout	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)
All infections	£297	Assumption: NHS Reference costs 2020/2021[114]: Emergency medicine (VB01-09Z)

Abbreviations: AE: adverse event; CC: complexity and comorbidity; NHS: national health service; NICE: national health and care excellence; TA: technology appraisal.

Miscellaneous unit costs and resource use

No miscellaneous unit costs or resource use were included.

B.3.6 *Severity*

The total life expectancy for the modelled population was calculated using population mortality data from the ONS for 2017–2019.[117] The total life expectancy was quality-adjusted using UK population norm values for EQ-5D by age and sex as reported by Hernández Alava et al. 2022 through the NICE DSU.[130] Despite the large unmet need, ruxolitinib does not meet the criteria for a severity weight in this indication (Table 36).

Table 36: Summary of QALY shortfall analysis

Analysis	Expected total QALYs for the general population	Total QALYs that people living with a condition would be expected to have with current treatment	Absolute QALY shortfall	Proportional QALY Shortfall
RESPONSE	12.60	6.97	5.63	0.45
RESPONSE-2	11.13	7.80	3.32	0.30
MAJIC-PV	10.55	6.11	4.45	0.42

Abbreviations: QALY: quality-adjusted life year.

B.3.7 *Uncertainty*

PV is a rare condition, with a relatively favourable prognosis when compared with MF. As highlighted in Section B.3.3.4, to be able to demonstrate a statistically significant survival benefit and reduction in key events, a large sample size and long follow-up is required which is not possible due to the rarity of the condition. The effectiveness of ruxolitinib in PV over 5 years is already supported by two phase III trials and one UK phase II trial. We therefore urge the committee to consider the strength of the data package presented despite the nature of the condition impacting on the ability to generate evidence.

B.3.8 *Summary of base-case analysis inputs and assumptions*

B.3.8.1 Summary of base-case analysis inputs

A summary of the base case model inputs is provided below in Table 37.

Table 37: Summary of variables applied in the economic model

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
NICE reference case			
Time horizon	600 cycles (46 years)	Not varied	Section B.3.2.3
Discount rate for costs	3.5%	Not varied	Section B.3.2.3
Discount rate for benefits	3.5%	Not varied	Section B.3.2.3
Baseline characteristics			
Age	See Table 22	Normal	Section B.3.3.1
% male (male %)	See Table 22	Beta	Section B.3.3.1
Weight	See Table 22	Normal	Section B.3.3.1
Treatment effect for OS (HR ruxolitinib versus BAT)			
HR for OS (0-3.00 years)	[redacted] (95% CI: [redacted])	Lognormal	Section B.3.3.4

HR for OS (3.00+ years)	[redacted] (95% CI: [redacted])	Lognormal	Section B.3.3.4
Per cycle rate of events (MF, AML/MDS, NMSC, bleeding/haemorrhage)			
Rux per cycle rate of events	See Table 24	Beta	Section B.3.3.8
IRR (BAT versus rux)	See Table 25	Beta	Section B.3.3.8
HR TTD			
HR TTD versus OS – BAT	[redacted]	Beta	Section B.3.3.5
HR TTD versus OS - Rux	[redacted]	Lognormal	Section B.3.1.2
HRQoL (MF-8D) – Baseline and CFB			
Baseline	[redacted]	Beta	Section B.3.4.1
On Ruxolitinib – CFB	[redacted]	Multivariate normal	Section B.3.4.1
On BAT (1st) – CFB	[redacted]	Multivariate normal	Section B.3.4.1
No treatment	[redacted]	Beta	Section B.3.4.1
QALY loss for the key events			
MF	−0.063	Beta*	Section B.3.4.5
AML/MDS	−0.197	Beta*	Section B.3.4.5
TE	−0.043	Beta*	Section B.3.4.5
NMSC	0	Beta*	Section B.3.4.5
Bleeding/haemorrhage	−0.007	Beta*	Section B.3.4.5
Venesection	−0.0000103	Beta*	Section B.3.4.5
BAT composition			
BAT distribution	Table 20	Dirichlet	Section B.3.2.4
Drug costs assumptions			
Dosage for ruxolitinib	Table 30 & Table 31	Dirichlet	Section B.3.5.1
Drug acquisition and administration costs per cycle unless stated			
Ruxolitinib	[redacted]	RDI varied	Section B.3.5.1
BAT / on off adm	£226.48 / £24.71	BAT comp varied	Section B.3.5.1
Management costs (per 28-day cycle)			
On Rux (0–6 month)	[redacted]	Unit costs varied using Gamma*	Section B.3.5.2
On Rux (7–12 month)	[redacted]		
On Rux (13+ month)	[redacted]		
On BAT (0–6 month)	[redacted]		
On BAT (7–12 month)	[redacted]		
On BAT (13+ month)	[redacted]		
Management costs for the key events			
MF	£63,920	Gamma*	Section B.3.5.2
AML/MDS	£44,903	Gamma*	Section B.3.5.2
TE	£1,731	Gamma*	Section B.3.5.2
NMSC	£973	Gamma*	Section B.3.5.2
Bleeding/haemorrhage	£2,023	Gamma*	Section B.3.5.2
Venesection	£316	Gamma*	Section B.3.5.2
Other costs			
End of life	£6,774	Gamma*	Section B.3.5.2

* SE assumed to be 10%; ** RESPONSE; ***RESPONSE-2; **** MAJIC-PV; *****primary analysis only.
Abbreviations: AML: acute myeloid leukaemia; BAT: best available therapy; CFB: change from baseline; CI: confidence interval; HC: Hydroxycarbamide; HR: hazard ratio; HRQoL: health-related quality of life; HC/HU: hydroxycarbamide/hydroxyurea; MDS: myelodysplastic syndrome; IFN: interferon; IRR: incidence rate ratio; MF: myelofibrosis; NICE: National institute for health and care excellence; NMSC: non-melanoma skin cancer; p^{32} radiophosphorus; PV: polycythaemia vera; QALY: quality-adjusted life years; OS: overall survival; R/I: resistant to or intolerant: rux: ruxolitinib; SE: standard error; TE: thromboembolic events; TTD: Time to treatment discontinuation.

B.3.8.2 Assumptions

The assumptions used in the base case analysis are described in Table 38, with a description of the scenarios conducted to explore the potential impact of these assumptions, where appropriate.

Table 38: List of assumptions for the base case analysis model

Assumption	Description of assumption for the base case	Justification	Addressed in scenario analysis
Population			
Analyses for RESPONSE[65] and RESPONSE-2[66] are presented separately	Separate analyses are presented for these two populations (licensed population) as described in the NICE final scope[105].	Collectively, the RESPONSE-trials[65, 66] represent the entire patient population with PV who are R/I to HC/HU covered by the licence[16] and described in the decision problem. As the population included in the pivotal RESPONSE and RESPONSE-2 trials[65, 66] are two mutually exclusive sub-populations (patients with and without splenomegaly) separate analyses are presented for these two populations as described in the NICE final scope.[105]	N/A
The RESPONSE[65] and RESPONSE-2[66] trials are generalisable to England and Wales	Baseline characteristics (age, gender, weight) of patients who would receive ruxolitinib in clinical practice are reflective of those included in the RESPONSE-trials	Clinical experts deemed the trial to be representative of UK practice[11] but noted that patients had to require two phlebotomies in the past 6-month prior randomisation which may have excluded some patients	N/A
The MAJIC-PV trial[70] is generalisable to England and Wales	Baseline characteristics (age, gender) for the subgroup of patients with high-risk PV are reflective of those included in the MAJIC-PV trial.[70]	The MAJIC-PV trial[70] was conducted in the UK and therefore reflects UK practice.	N/A
Comparators			
The comparator in the economic case is current clinical UK management (including a basket of treatments)	The BAT composition is taken from the MAJIC-PV trial[70] for all analyses.	There are currently no licensed or recommended treatments for PV in the UK. Treatment primarily consists of HC/HU and IFN-a, with a minority of patients receiving busulfan and anagrelide.[105] The MAJIC-PV trial[70] is a UK based trial and therefore reflects the composition of treatments given in England and Wales. Relative treatment effectiveness for OS was also taken from the MAJIC-PV trial.[70]	N/A
Modelling not based on surrogacy			

Model based on survival rather than events or surrogacy	OS is not modelled as a function of key events	None of the trials[65, 66, 70] were designed or powered to detect a difference in incidence of key events (TE, MF, AML/MDS, haemorrhage, NMSC). Although five-year data is available for ruxolitinib, due to cross-over the follow-up time in the BAT arm of the RESPONSE-trials[65, 66] is too short to capture the incidence of these events which take time to manifest. In order to model OS as a function of these key events, a significant number of assumptions would be required.[106] Data from the general population would need to be utilised,[106] with models for DVT, CVD, MF and AML needing to be constructed. As shown in previous appraisals these models are complex and the choice of data and health states are often heavily scrutinised.[23, 107-111] Additionally, a number of treatments are available to treat these conditions, therefore, to model these events accurately would require assessing all possible interventions. Assumptions would also be required for the baseline survival in people without events, however, there are no data to inform a robust analysis. Further description is available in Section B.3.3.4.	N/A
Modelling of OS and TTD in the primary analysis (licensed population)			
Use of a STM	OS for patients on ruxolitinib in the primary analysis is estimated indirectly using a STM approach based on time on treatment and time to death following discontinuation taken from the RESPONSE-trials.[65, 66] A HR is applied to the predicted OS for ruxolitinib to derive BAT OS.	There are challenges to model survival directly in the RESPONSE-trials[65, 66] due to (a) the immaturity of OS data for ruxolitinib and (b) early and high levels of cross-over rates in BAT arm. A STM approach compared with a PSM approach allows for the capture of progressively worsening HRQoL which is associated with PV.	N/A
OS in patients initiating current clinical management	Estimated by applying a HR to the baseline OS curve for ruxolitinib A time-varying HR ([redacted] years versus [redacted] years+) estimated from the MAJIC-PV trial[70] is used in the base-case.	In the MAJIC-PV trial manuscript,[70] the curves for OS started to diverge after approximately 3.0 years, this is in line with clinical expectation that a survival difference would not manifest immediately.[11] UK Clinical advisors expected the curve to separate further in the long-term because of the effect	Alternative cut-off points are examined in scenario analyses.

		ruxolitinib on haematological parameters and reduction in thromboembolic events and myelofibrosis.[11, 106]	
TTD for ruxolitinib modelled under a competing-risk framework	Time to discontinuation due to death and reason other than death modelled separately.	Availability of IPD from the RESPONSE-trials[65, 66] to account for the low number of discontinuations due to death and long-term extrapolation required.	N/A
Treatment effect waning			
No treatment effect after 20 years	In the base-case, the treatment effect is assumed to stop after 20 years and diminish gradually (linearly) between Year 5 – 20.	Cut-off selected to reflect clinical expectation that approximately twice the number of patients would be alive at Year 20.[106]	Alternative duration of treatment effects are assumed
Modelling of OS and TTD for the high-risk PV subgroup			
Use of a PSM	OS for BAT and TTD for ruxolitinib are extrapolated directly from the trial (reconstructed pseudo-IPD of the MAJIC-PV trial[70]) using a PSM approach	Absence of IPD of the MAJIC-PV trial[70] and more mature data compared with the RESPONSE-trials.[65, 66] The KM for OS and TTD are available in the unpublished manuscript of the MAJIC-PV trial[70] and therefore IPD could be reconstructed for these two outcomes separately.	N/A
TTD for ruxolitinib	TTD is modelled using a HR applied to ruxolitinib OS from MAJIC-PV trial.[70]	To ensure consistency between OS extrapolation and ruxolitinib TTD.[106]	Scenario analyses are conducted extrapolating TTD directly
OS in patients initiating ruxolitinib	A time-varying HR ([redacted] years versus [redacted] years+) estimated from the MAJIC-PV trial[70] is used in the base-case applied to BAT arm	Treatment effect from the MAJIC-PV trial.[70]	Scenario analyses are conducted extrapolating OS directly for ruxolitinib
Selection of parametric functions for primary analysis (licensed population)			
TTD for ruxolitinib (with death censored)	Odd spline with one knot used in the base-case for the primary analysis.	Selected following (1) visual fit, (2) statistical goodness of fit and (3) long-term plausibility.	Alternative distributions used in scenario analysis
Discontinuation due to death	Data from RESPONSE and RESPONSE-2 are pooled.[65, 66] Exponential used in the base-case. The maximum hazard of death between	Data is pooled to increase the statistical power and reduce the uncertainty. Exponential distribution selected following (1) visual fit, (2) statistical goodness of fit and (3) long-term plausibility.	Alternative distributions used in scenario analysis

	general population mortality (life table[117]) and extrapolation is used.	Constraint included to avoids the hazard of death in the model being lower than death from general causes.	
Time to death following ruxolitinib discontinuation	Data from RESPONSE and RESPONSE-2 are pooled.[65, 66] Exponential used in the base-case. The maximum hazard of death between general population mortality (life table[117]) and extrapolation is used.	Data is pooled to increase the statistical power and reduce the uncertainty. Exponential distribution selected following (1) visual fit, (2) statistical goodness of fit and (3) long-term plausibility. Constraint included to avoids the hazard of death in the model being lower than death from general causes.	Alternative distributions used in scenario analysis
Section of parametric functions for the high-risk PV subgroup & time to first BAT discontinuation (for partition in licensed population)			
OS for BAT	Weibull is used in the base. The maximum hazard of death between general population mortality (life table[117]) and extrapolation is used.	Selected following (1) visual fit, (2) statistical goodness of fit and (3) long-term plausibility. Constraint included to avoids the hazard of death in the model being lower than death from general causes.	Alternative distributions used in scenario analysis
TTD for BAT	KM used followed by Gompertz extrapolation	None of the parametric provided a good visual to the data.	Alternative distributions used in scenario analysis
Partitioning of the BAT health state – primary analysis			
BAT health state partition	The BAT health state is partitioned onto three sub-health states (1st BAT, 2nd BAT+, no treatment) in the primary analysis only	This is to account for the progressive worsening in HRQoL as patients cycle through BAT regimens. Assumption are made about the decrement in utility and management costs as patient switch and stop BAT. It was not possible to partition the BAT health for the subgroup analysis due to the inflexibility of PSM.	Structural uncertainty explored by assuming a single BAT health state (e.g. removal of partition) in the primary analysis
Time to BAT discontinuation	Approximated by applying a HR to BAT OS. HR approximated using number of death and discontinuation reported in the unpublished MAJIC-PV trial manuscript.[70]	Time to 'all' BAT discontinuation is not reported in the MAJIC-PV trial manuscript.[70]	Structural uncertainty explored by removing "no treatment" sub-health state
Inclusion of complications (MF, AML/MDS, NMSC, bleeding/haemorrhage)			
Cost and QALY impact of events	The inclusion of complications (MF, AML/MDS, NMSC, haemorrhage) in the data analysis model was simplified and limited to the cost and QALY impact.	Including MF, AML/MDS, NMSC and haemorrhage as separate health states would require a large number of assumptions, lead to double counting, and add considerable complexity.[11]	Scenario analyses are conducted removing the impact of events on costs and quality of life

Analyses-specific incidence of events for ruxolitinib	Incidence of events while on ruxolitinib are taken from each trial respectively when available.	5-year data is available for ruxolitinib from both the RESPONSE and the MAJIC-PV trials.[65, 66, 70] Data is not pooled to reflect differences in the incidence of events between patients with PV with splenomegaly, without splenomegaly and high-risk, notably for MF	N/A
Incidence of events on BAT derived using a treatment effect applied to incidence of events on ruxolitinib	A treatment effect (incidence rate ratio) is applied to the incidence of events on ruxolitinib. The treatment effect is pooled across trials.	Early and high levels of cross-over in the RESPONSE-trials.[65, 66] Long follow-up is required to assess incidence of events. Treatment effect (incidence rate ratio) is pooled across trials[65, 66, 70] to increase sample size and statistical power. Clinical experts considered this was reasonable.[106]	N/A
HRQoL			
MF-8D used in the base-case	Changes in HRQoL are measured using the MF-8D (a condition-specific measure) derived from RESPONSE.[65]	Change in HRQoL is taken from a condition-specific measure (the MF-8D) given the lack of sensitivity of the EQ-5D in capturing changes in symptoms and their impact on HRQoL in PV and condition alike such as MF.	EQ-5D explored in scenario analysis.
Cumulative decrement in HRQoL on BAT	A cumulative decrement in HRQoL is assumed as patients move through the BAT sub-health states in the primary analysis only.	UK clinical experts indicated that HRQoL progressively worsens as the disease progresses. Switching BAT used as a proxy in the primary analysis.	Structural uncertainty explored assuming a single health state in the primary analysis.
Adverse events			
Adverse events	The effect of AEs (all grades) on costs and HRQoL is included.	The impact of AE on costs and quality of life is included in the base-case to reflect the NICE reference case.[25]	A scenario analysis is conducted removing the effect of AEs.
Resource use			
Resource estimates	The type and frequency of resource use was estimated from the average resource use estimated by five clinical experts	In the absence of evidence for NHS resource use for patients with PV R/I to HC/HU, five UK clinical experts were asked to complete a resource utilisation questionnaire. The average frequency is used.	Resource use estimate from each expert

Abbreviations: AE: adverse event; AML: acute myeloid leukaemia; BAT: best available therapy; BNF: British natural formulary; CVD: cardiovascular disease; eMIT: electronic market information tool; EQ-5D: euroqol 5-dimensions; HC: Hydroxycarbamide; HR: hazard ratio; HRQoL: health-related quality of life; HC/HU: hydroxycarbamide/hydroxyurea; IPD: individual patient level data; MDS: myelodysplastic syndrome; MF: myelofibrosis; N/A: not applicable; NHS: national health service; NICE: National institute for health and care excellence; NMSC: non-melanoma skin cancer; PSS: Personal Social Services; PV: polycythaemia vera; QALY: quality-adjusted life years; OS: overall survival; PSM: partitioned survival model; R/I: resistant to or intolerant: SE: standard error; SmPC: summary of product characteristics; STM: state-transition model; TE: thromboembolic events; TTD: Time to treatment discontinuation.

B.3.9 *Base-case results*

Results of the economic analysis for the primary analysis for the licensed population are presented below in Section B.3.9.1 with the results for the subgroup of patients with high-risk PV resented in Section B.3.11. Results are presented in the main body of this submission using the PAS discount agreed for MF [redacted]

B.3.9.1 Base-case incremental cost-effectiveness analysis results – primary analysis

Table 39 presents the base case results of the economic evaluation for the primary analysis for the licensed population. Clinical outcomes from the cost-effectiveness model, the proportion of the cohort in each health state over time (Markov trace), and the disaggregated results of the base case incremental cost-effectiveness analysis are reported in Appendix J. The net health benefit is presented in Table 40.

Table 39: Base-case incremental cost-effectiveness results: Primary analysis for the licensed population

Technologies	Total costs (£)	Total LYG*	Total QALYs	Incr. costs (£)	Incr. LYG*	Incr. QALYs	ICER (£/QALY)
Licensed population without splenomegaly							
Current clinical management	£86,809	10.46	7.80	-	-	-	-
Ruxolitinib	[redacted]	12.25	[redacted]	[redacted]	1.79	[redacted]	[redacted]
Licensed population with splenomegaly							
Current clinical management	£92,017	9.28	6.97	-	-	-	-
Ruxolitinib	[redacted]	11.45	[redacted]	[redacted]	2.17	[redacted]	[redacted]

Note: all results presented are discounted unless otherwise stated. *undiscounted
Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; LYGs: life years gained; QALYs: quality-adjusted life years.

Table 40: Net health benefits: Primary analysis for the licensed population

Technologies	Total costs (£)	Total QALYs	Incr. costs (£)	NHB at £20,000	NHB at £30,000
Licensed population without splenomegaly					
Current clinical management	£86,809	7.80	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Licensed population with splenomegaly					
Current clinical management	£92,017	6.97	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: Incr.: incremental; NHB, net health benefit; QALYs: quality-adjusted life years.

Licensed population without splenomegaly

The base-case incremental cost-effectiveness results for the licensed population without splenomegaly show that over a lifetime time horizon, the total costs associated with ruxolitinib were estimated to be [redacted] compared with £86,809 for patients treated with current clinical management in the UK (an incremental cost of [redacted]).

The total QALYs for patients receiving ruxolitinib were estimated to be [redacted] compared with 7.80 for patients treated with current clinical management in the UK (an incremental QALY gain of [redacted]) in adult patients without splenomegaly, resulting in an ICER of £[redacted] per QALY gained.

Licensed population with splenomegaly

The base-case incremental cost-effectiveness results for the licensed population with splenomegaly show that over a lifetime time horizon, the total costs associated with ruxolitinib were estimated to be [redacted] compared with £92,017 for patients treated with current clinical management in the UK (an incremental cost of [redacted]).

The total QALYs for patients receiving ruxolitinib were estimated to be [redacted] compared with 6.97 for patients treated with current clinical management in the UK (an incremental QALY gain of [redacted]) in adult patients with splenomegaly, resulting in an ICER of £[redacted] per QALY gained.

B.3.10 *Exploring uncertainty*

B.3.10.1 Probabilistic sensitivity analysis – Primary analysis

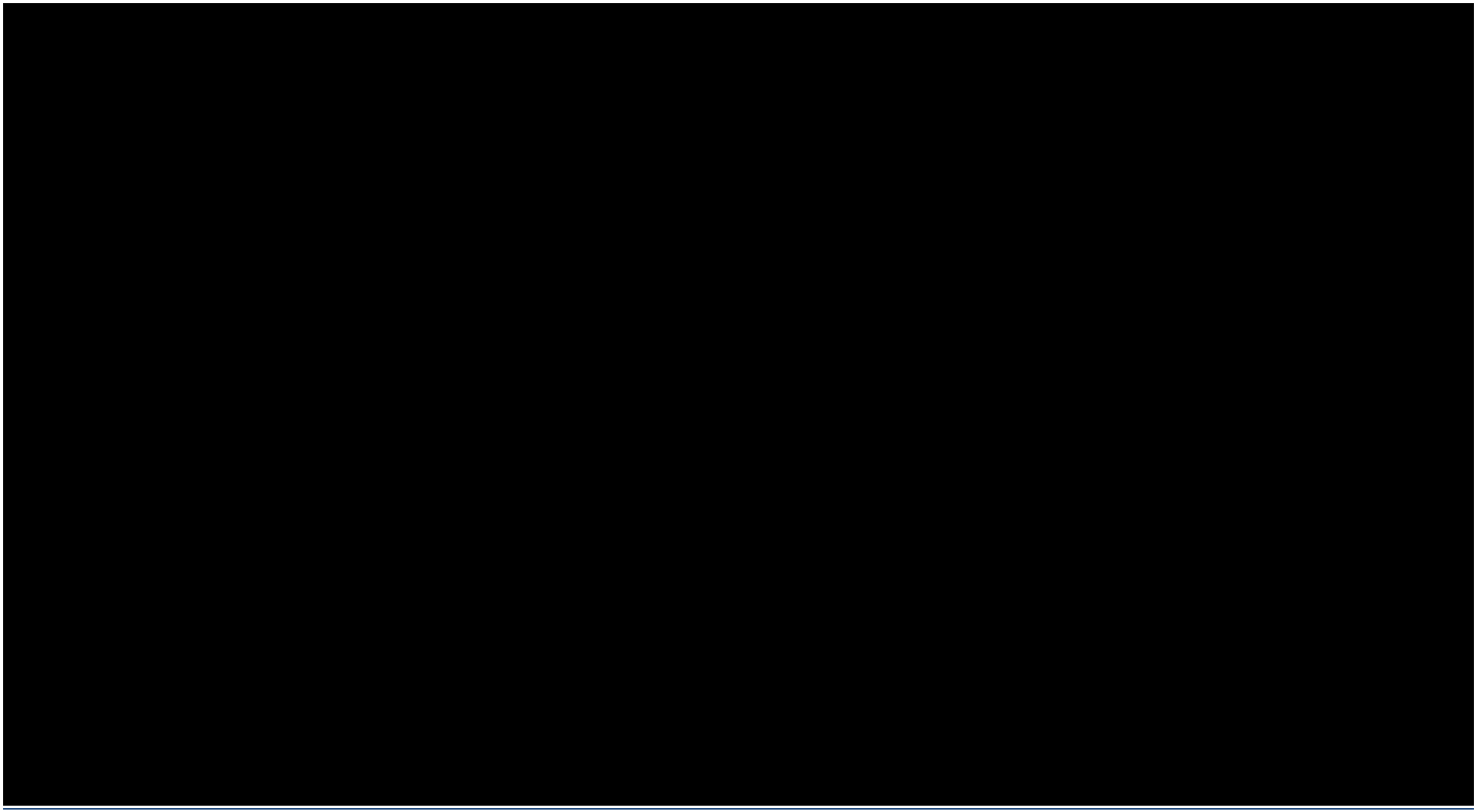
A probabilistic sensitivity analysis (PSA) was conducted in order assess the simultaneous effect of uncertainty in the different model parameters. A Monte-Carlo simulation with 2,000 iterations was performed and, in each iteration, model inputs were randomly sampled from the specified probability distributions described in Table 37. An arbitrary SE of 10% around the mean was assumed when the SE or 95% CI was not available. Survival distribution and regression models were varied using multivariate normal distributions. Proportions were varied using a Dirichlet distribution or beta distribution (when binary). Costs and utility values were varied using a gamma and beta distribution respectively. Treatment effect (hazard ratio) were varied using a lognormal distribution. The results of the PSA are presented in Table 41 with the cost-effectiveness (CE) plane and cost-effectiveness acceptability curves (CEAC) resulting from the PSA in Figure 52.

Table 41: PSA results: primary analysis for the licensed population

Technologies	Total costs (£)	Total QALYs	Incr. costs (£)	Incr. QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Licensed population without splenomegaly						
BAT	£85,779	7.61	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Licensed population with splenomegaly						
BAT	£91,580	6.86	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a]The probability of ruxolitinib being cost-effective versus current clinical management in the UK at a WTP threshold of £20,000/QALY gained. Note: all results presented are discounted unless otherwise stated.
Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; QALYs: quality-adjusted life years.

Figure 52: PSA cost-effectiveness plane and CEAC: Primary analysis for the licensed population



Note: all results presented are discounted unless otherwise stated.
Abbreviations: CE: cost-effectiveness, CEAC: cost-effectiveness acceptability curve; ICER: incremental cost-effectiveness ratio; PSA: probabilistic sensitivity analysis; QALYs: quality-adjusted life years; WTP: willingness-to-pay threshold.

Licensed population without splenomegaly

Results of the PSA (Table 41) show that in the licensed population without splenomegaly, over a lifetime time horizon, ruxolitinib is associated with greater QALYs ([redacted]), at a greater cost ([redacted]) compared with current clinical management in the UK (7.61 QALYs and £85,779 respectively). As such, the average PSA ICER was estimated to be £[redacted] per QALY gained, with a [redacted]% probability of ruxolitinib being a cost-effective treatment option at a £20,000/QALY gained willingness-to-pay (WTP) threshold. The cost-effectiveness plane and cost-effectiveness acceptability curves resulting from the PSA for the analysis in adult patients with PV R/I to HC/HU without splenomegaly are presented in Figure 52.

Licensed population with splenomegaly

Results of the PSA (Table 41) show that in the licensed population with splenomegaly, over a lifetime time horizon, ruxolitinib was associated with greater QALYs ([redacted]), at a greater cost ([redacted]) compared to current clinical management in the UK (6.86 QALYs and £91,580 respectively). As such, the average PSA ICER was estimated to be £[redacted] per QALY gained, with a [redacted]% probability of ruxolitinib being a cost-effective treatment option at a £20,000/QALY gained WTP threshold.

B.3.10.2 Deterministic sensitivity analysis – Primary analysis

In order to assess the robustness of the base case cost-effectiveness results, deterministic sensitivity analyses (DSA) were conducted by varying one model input at a time to assess which parameters had the most impact on the ICER. Parameters were varied within their 95% CI where available (or possible to calculate) or within a reasonable range (+/- 20%).

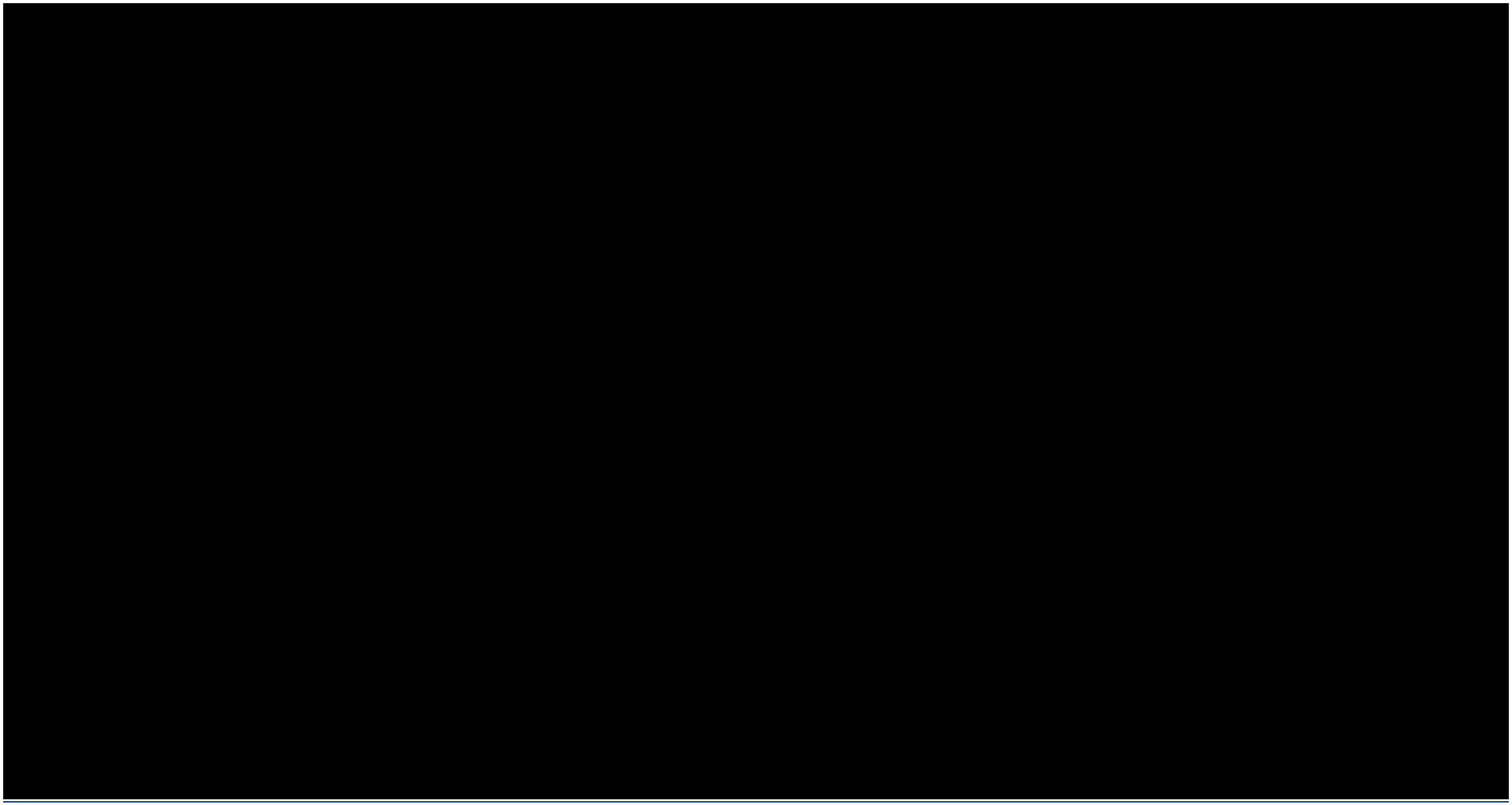
The results for the ten most influential parameters assessed in the DSA and the ICERs calculated at the upper and lower bounds are shown graphically in the tornado plot in Figure 53, sorted from the widest to narrowest range of ICER values to highlight the parameters with the strongest influence on the cost-effectiveness results. Unsurprisingly, the results of the DSA show that results were most sensitive to assumptions around the treatment effect for OS, discount rate for both cost and benefits and assumptions around utility values.

B.3.10.3 Scenario analysis

In addition to the DSA, extensive scenario analyses were conducted altering important variables in the cost-effectiveness model. Results of the top 20 scenario analyses that most significantly impacted the ICER are presented below in Figure 54.

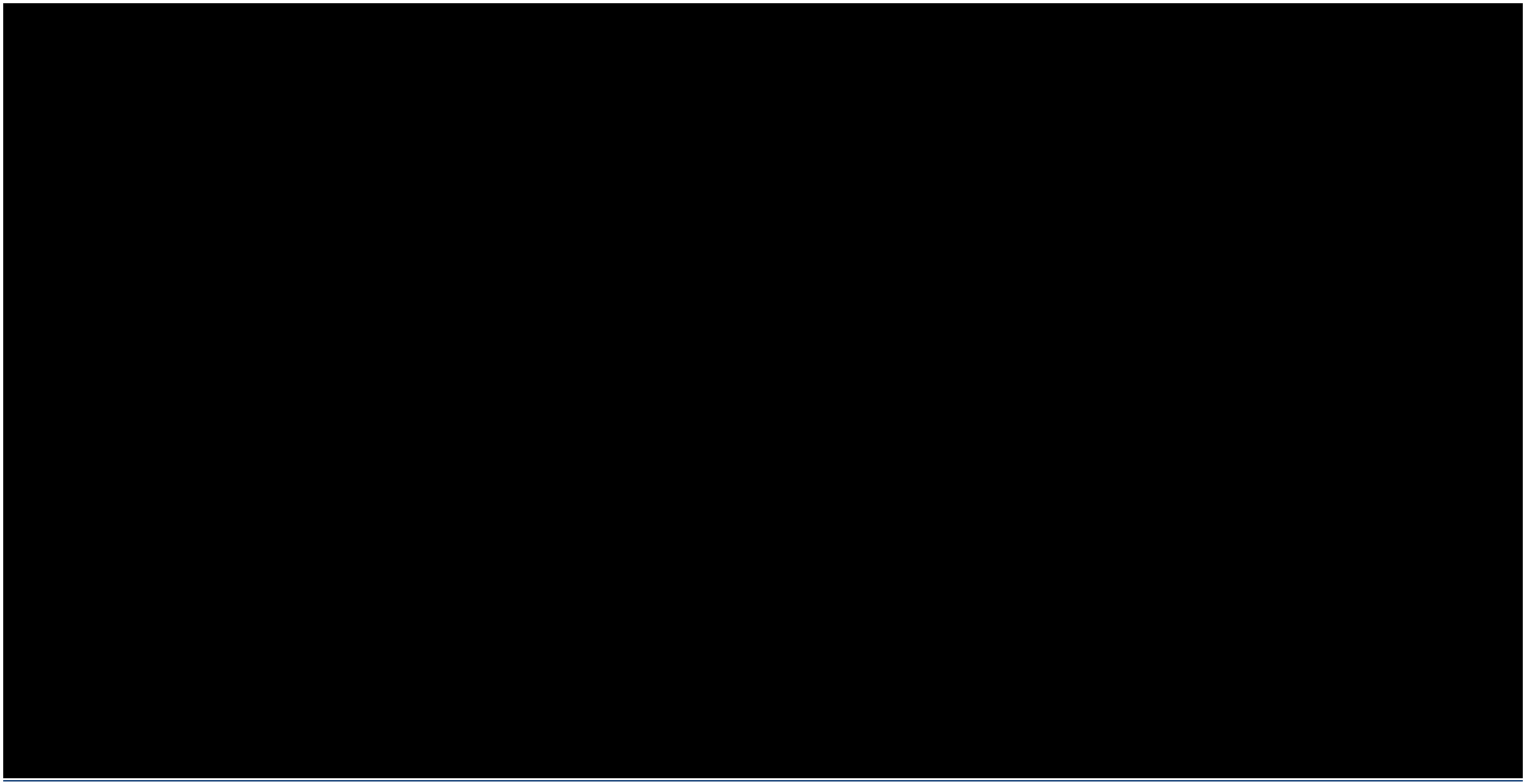
The scenarios that result in the largest impact on the ICER are those around the source of the treatment effect for OS, assumption of treatment effect waning, time horizon and dose reduction/interruption. The ICER was also sensitive to the choice of extrapolation for TTD, notably the use of spline models with two or more knots. However, as described in Section B.3.1.2, the spline models over-fitted the data and lead to high discontinuation rates beyond the observed period were not considered realistic.

Figure 53: Tornado diagram based on DSA results: Primary analysis for the licensed population



Note: all results presented are discounted unless otherwise stated.
Abbreviations: CI: confidence interval; DSA: deterministic sensitivity analysis; HR: hazard ratio; ICER: incremental cost-effectiveness ratio; OS: overall survival; PAS: patient access scheme; QALY: quality-adjusted life year.

Figure 54: Scenario analysis results: Primary analysis for the licensed population



Note: all results presented are discounted unless otherwise stated.
Abbreviations: AE: adverse event; AML: acute myeloid leukemia; EQ-5D-3L: EuroQol 5 Dimensions 3 Levels; HR: hazard ratio; ICER: incremental cost-effectiveness ratio; LYs: life years; NHS: national health system; OS: overall survival; QALYs: quality-adjusted life year; PAS: patient access scheme; TTD: time to treatment discontinuation.

B.3.11 *Subgroup analysis*

B.3.11.1 Deterministic results

Table 42 presents the base case results for the subgroup of patients with high-risk PV (based on the MAJIC-PV trial[70]). The net health benefit is presented in Table 42.

Table 42: Base-case incremental cost-effectiveness results: subgroup of adult patients with high-risk PV

Technologies	Total costs (£)	Total LYG*	Total QALYs	Incr. costs (£)	Incr. LYG*	Incr. QALYs	ICER (£/QALY)
Current clinical management	£83,317	8.02	6.11				
Ruxolitinib	[redacted]	9.65	[redacted]	[redacted]	1.63	[redacted]	[redacted]

Note: all results presented are discounted unless otherwise stated. *Undiscounted.
Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; LYGs: life years gained; QALYs: quality-adjusted life years.

Table 43: Net health benefit: subgroup of adult patients with high-risk PV

Technologies	Total costs (£)	Total QALYs	Incr. costs (£)	NHB at £20,000	NHB at £30,000
Current clinical management	£83,317	6.11	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Note: all results presented are discounted unless otherwise stated.
Abbreviations: Incr.: incremental; NHB, net health benefit; QALYs: quality-adjusted life years.

The base-case incremental cost-effectiveness results for the subgroup of patients with high-risk PV R/I to HC/HU show that over a lifetime time horizon, the total costs associated with ruxolitinib were estimated to be [redacted] compared with £83,317 for patients treated with current clinical management in the UK, representing an incremental cost of [redacted].

The total QALYs for patients receiving ruxolitinib were estimated to be [redacted] compared with 6.11 for patients treated with current clinical management in the UK representing an incremental QALY gain of [redacted], resulting in an (probabilistic) ICER of [redacted] per QALY gained.

B.3.11.2 Probabilistic results

The results of the PSA are presented in Table 44.

Table 44: PSA results: subgroup of adult patients with high-risk PV

Technologies	Total costs (£)	Total QALYs	Incr. costs (£)	Incr. QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Current clinical management	£83,804	6.07				
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a]The probability of ruxolitinib being cost-effective versus current clinical management in the UK at a WTP threshold of £20,000/QALY gained. Note: all results presented are discounted unless otherwise stated.

Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; QALYs: quality-adjusted life years; PSA: probabilistic sensitivity analysis; WTP: willingness-to-pay.

Results of the PSA (Table 44) show that in the subgroup of patients with high-risk PV, over a lifetime time horizon, ruxolitinib was associated with greater QALYs ([redacted]), at a greater cost ([redacted]) compared with current clinical management in the UK (6.07 QALYs and £83,804 respectively). As such, the average PSA ICER was estimated to be £[redacted] per QALY gained, with a [redacted]% probability of ruxolitinib being a cost-effective treatment option at a £20,000/QALY gained WTP threshold. The CE plane and cost-effectiveness acceptability curve for the subgroup of high-risk patients are presented in Figure 52.

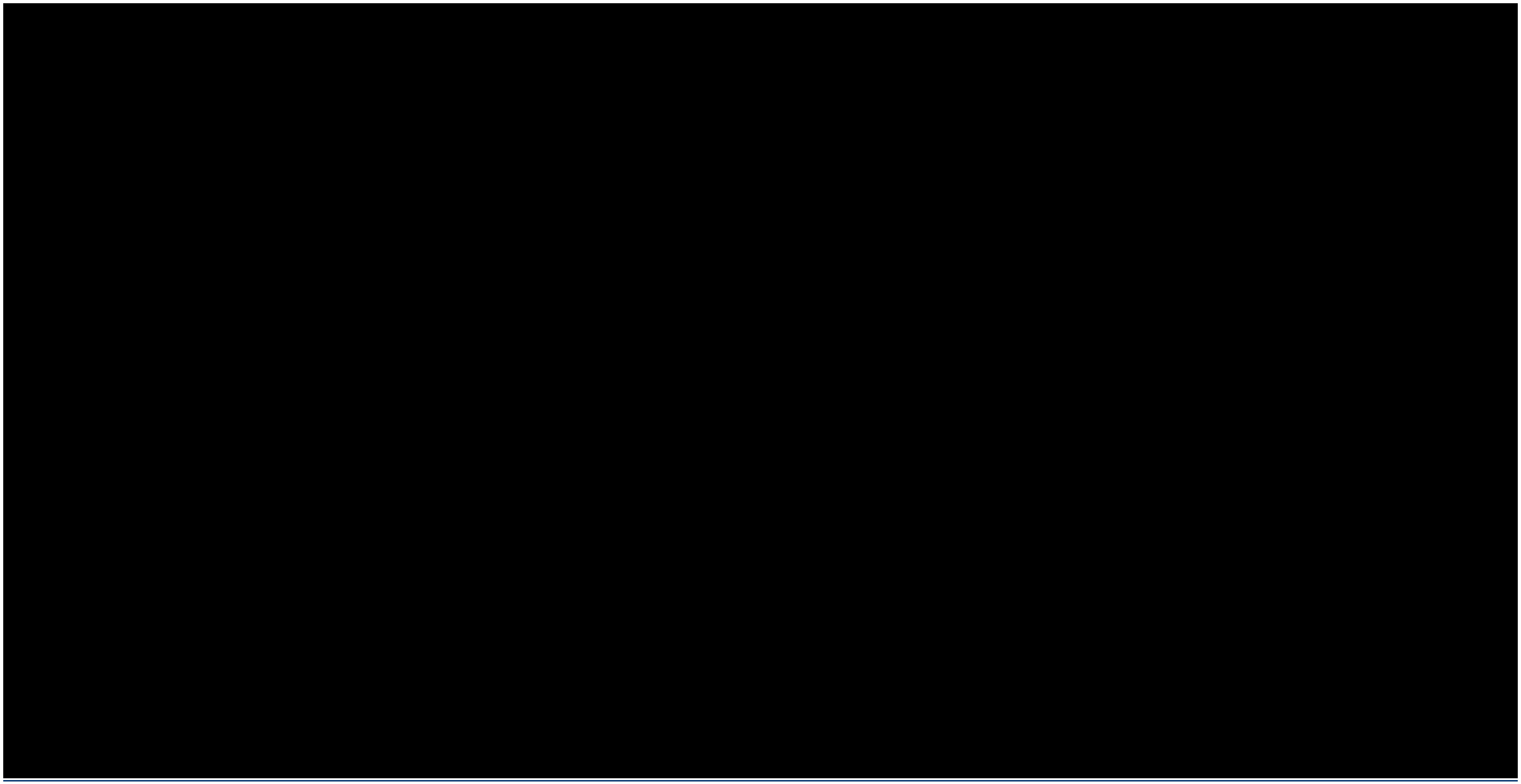
Figure 55: PSA cost-effectiveness plane and CEAC: subgroup of adult patients with high-risk PV

Note: all results presented are discounted unless otherwise stated.
Abbreviations: CE: cost-effectiveness, CEAC: cost-effectiveness acceptability curve; ICER: incremental cost-effectiveness ratio; PSA: probabilistic sensitivity analysis; QALYs: quality-adjusted life years; WTP: willingness-to-pay threshold

B.3.11.3 One-way SA and scenario analyses

Results from the DSA and scenario analysis for subgroup of patients with high-risk PV are presented in Figure 56.

Figure 56: One way SA and scenario analysis: subgroup of adult patients with high-risk PV



Note: all results presented are discounted unless otherwise stated.
Abbreviations: AE: adverse event; AML: acute myeloid leukemia; EQ-5D-3L: EuroQol 5 Dimensions 3 Levels; HR: hazard ratio; ICER: incremental cost-effectiveness ratio; LYs: life years; NHS: national health system; OS: overall survival; PFS: progression-free survival; QALYs: quality-adjusted life year; PAS: patient access scheme; TTD: time to treatment discontinuation.

In line with the sensitivity analysis results for the primary analysis (Section B.3.10 and Appendix P), the sensitivity analysis results for this subgroup were most sensitive to the treatment used for OS, the duration of treatment effect, discount rates and utility values. In addition to those, the ICER was also sensitive to the parametric distribution for OS for BAT and HR for TTD for ruxolitinib.

B.3.12 *Benefits not captured in the QALY calculation*

The economic analysis has attempted to capture all the potential benefits related to ruxolitinib within the QALY calculation. There are, however, several potential benefits of treatment with ruxolitinib which are not captured within the assessment, specifically:

- The positive impact of an oral treatment. No decrement in utility was assumed for oral vs. treatments that are self-administered.
- Supply issues with IFN-alfa within the NHS;
- The benefit on NHS capacity through the reduction in patients requiring venesection, amid the current backlogs faced by the NHS. Avoiding hospital visits reduces the financial and administrative strain on NHS capacity. While direct (cost of venesection) costs are captured, keeping patients away from hospital and alleviating some burden on NHS staff and infrastructure (i.e. human and physical capital) are important elements to consider at a time when the NHS continues to face significant backlogs from the pandemic.
- In addition to helping alleviate capacity issues within the NHS, reducing hospital visits will also have a positive impact on patient quality of life as patients may experience increased anxiety and stress, in particularly due to the risk of contracting COVID-19 during their hospital visit.
- The likely positive impact on patient well-being associated a treatment that selectively targets the relevant biological pathway implicated in PV.

Given the above, it is plausible that additional potential benefits of ruxolitinib are not captured in the QALY (and ICER) calculation, and we would ask the Appraisal Committee to consider these factors when generating their recommendations.

B.3.13 *Validation*

Clinical validation was sought for support with this submission, consisting of individual interviews with one clinical expert, in addition to two advisory board meetings with five clinical experts.[11, 106] The five clinical experts were leading medical and clinical oncologists with experience in the management of patients with PV R/I to HC/HU, selected from a range of centres across England and Wales in order to provide a variety of expert perspectives. Clinical experts selected were also involved in the MAJIC-PV trial,[70] a UK Phase 2 investigator-led trial evaluating ruxolitinib versus established clinical management.

The following key aspects were discussed and validated:

- The natural history of PV;
- The population and relevance of the RESPONSE and MAJIC trials;
- The model structure and appropriateness to the decision problem;

- The comparator and distribution of treatments that make up BAT;
- Extrapolation of survival beyond the observed period;
- Validity of model inputs such as costs and utilities.

In addition to clinical validation of model inputs, the cost-effectiveness model was quality assured by a health economist not involved in the model building who reviewed the model for coding errors, inconsistencies, and plausibility of inputs. The model was also subject to stress testing of extreme scenarios to test for known modelling errors and questioning of results.

Long-term data for patients R/I to HC/HU are not available, therefore long-term predictions could not be compared against external data. However, despite using different model structures and inputs, the predicted life years for the licensed population without splenomegaly was higher compared with that of the licensed population with splenomegaly, which aligns with clinical expectations. Similarly, the predicted life years for the subgroup analysis in patients with high-risk PV were lower compared with that from the primary analysis, which reflect the poorer prognosis of this group of patients. The prediction for the subgroup analysis was also in line with that observed in Alvarez et al (2022).

Furthermore, despite different structure used between the primary (STM) and subgroup analysis (PSM), the predicted time alive post-ruxolitinib was similar (2.81 to 3.56 years in the primary analysis based on RESPONSE trials vs. 2.77 years in the subgroup analysis based on the MAJIC-PV trial).

B.3.14 *Interpretation and conclusions of economic evidence*

The deterministic results of the base case economic analysis show that ruxolitinib was associated with higher costs but also higher QALYs than current clinical management in the UK, with an incremental cost per QALY gained in the primary analysis of [redacted] (with ruxolitinib at the current agreed PAS price for MF) in the licensed population without splenomegaly (based on RESPONSE-2[66]) and [redacted] in the licensed population without splenomegaly (based on RESPONSE[65]). In the subgroup of patients with high-risk PV based on the population recruited in the MAJIC-PV trial,[70] the incremental cost per QALY gained was [redacted].

Despite the large unmet need, this intervention does not meet the criteria for a decision modifier for severity of disease. Across all analyses, the absolute shortfall was less than six years while the proportional shortfall was less than 0.45. Despite the nature of the condition impacting on the ability to generate evidence, the effectiveness of ruxolitinib in PV is available from 2 phase III RCTs and one phase II UK study.

Sensitivity and scenario analyses indicated the ICER to be robust to plausible changes, with the exception of assumptions around the treatment effect and its duration, utility values and time horizon. The ICER was also sensitive to the choice of parametric extrapolation, but these tended to lead to an improvement in the ICERs.

Strengths of the economic analysis include:

- The economic analysis is underpinned by two well-designed Phase 3 RCT (RESPONSE[65]; RESPONSE-2[66]) and one Phase 2 RCT (the MAJIC-PV trial[70]) that are representative of the population expected to be treated with ruxolitinib in England and Wales.
- The economic analysis includes all the relevant evidence available that are likely to arise during this appraisal. The MAJIC-PV trial is a UK Phase 2 investigator-led trial and while Novartis does not have access to the data, authors were approached to include data from this study as part of this submission, despite the manuscript currently undergoing peer review.
- The model structure and assumptions were developed with input from five key UK clinical experts specialising in the treatment of patients with PV R/I to HC/HU and involved in the MAJIC-PV trial.[70]
- Uncertainty in the model inputs and assumptions has been explored in a large number of scenario and sensitivity analyses that demonstrate the robustness of the model results to most assumptions and inputs.
- Medium-term evidence (5-year follow-up data) for ruxolitinib is available in patients with PV R/I to HC/HU from both the RESPONSE-trials[65, 66] and the MAJIC-PV trial.[70] The relative effectiveness of ruxolitinib over BAT (without cross-over) is also available in the MAJIC-PV trial.[70]
- Whilst acknowledging the data are immature, final endpoints are modelled directly rather than using surrogacy (which would rely on a large number of unsupported assumptions and increase uncertainty).
- Utility values are estimated from the trial and use a condition preference based-measure due to the limitation of the EQ-5D in adequately capturing change in quality of life in PV and condition alike.

Limitations of the analysis include:

- While the RESPONSE-trials[65, 66] provided 5-year data, the survival data for ruxolitinib are immature, with approximately 91% and 96% of patients still alive at the end of follow-up period for RESPONSE and RESPONSE-2, respectively.[65, 66] This high survival rate reflects both the prognosis of patients with PV and the potential benefits of ruxolitinib on survival. Furthermore, crossover from BAT to ruxolitinib was permitted in both trials (after Week 32 in RESPONSE and Week 28 in RESPONSE-2) and by Week 80 all patients in the BAT arms had discontinued BAT.
- Novartis do not have access to the IPD under the existing MAJIC IIT contract and therefore analyses had to be based on data that are contained in the unpublished manuscript of the MAJIC-PV trial, restricting the choice of structure and inputs.
- None of the trials were designed or powered to evaluate the effectiveness of ruxolitinib on events or survival. A long follow-up and large sample size is required for this.[11]
- Extrapolations for TTD and OS remain uncertain as data are immature. While the choice of parametric extrapolation for TTD for ruxolitinib and OS for BAT had an impact on the cost-effectiveness results in the scenario analyses presented, these were the scenarios using curves that were not plausible.

- The relative treatment effect of ruxolitinib versus BAT on survival beyond 5 years is highly uncertain and was a key driver for the cost-effectiveness results. In the base-case, no treatment effect was assumed after 20 years, with the treatment effect gradually diminishing between Years 5–20 to reflect clinical experts' expectation of approximately twice as many patients alive with ruxolitinib compared with BAT.[106] Extensive scenario analyses were conducted for transparency, and unsurprisingly, the ICERs were less favourable when the treatment effect was assumed to stop earlier, but the ICER improved if the treatment effect was assumed to stop beyond 20 years.
- No data exist on the management of PV in the UK. In the absence of UK data, the type and frequency of resource use was based on the average resource use estimated by five clinical experts; this is therefore uncertain.

Concluding remarks

There is no reimbursed standard therapy for treating patients with PV R/I to HC/HU in England and Wales. While IFN-alfa is the recommended second-line treatment in patients with PV that are R/I to HC/HU, a large proportion of patients do not tolerate IFN-alfa or cannot be prescribed IFN-a. Therefore, despite being R/I to HC/HU, a large proportion of patients continue to receive HC/HU in the absence of alternative therapy. Consequently, there is a high unmet medical need for a well-tolerated and effective therapy to reduce disease burden, improve survival rates, reduce disease transformation and improve HRQoL.

The cost-effectiveness analysis showed that at a willingness-to-pay threshold of £20,000/QALY, ruxolitinib [redacted] cost-effective treatment option compared with current clinical management. Based on the need for [redacted] to allow a cost-effective price to be offered for this appraisal, Novartis are working [redacted] (subject to NICE's assessment). [redacted], the ICERs are:

- [redacted] for the licensed population without splenomegaly;
- [redacted] for the licensed population with splenomegaly;
- [redacted] for the subgroup of patients with high-risk PV (as defined in MAJIC).

Novartis will continue to work with [redacted]

References

1. McMullin MF, Harrison CN, Ali S, et al. A guideline for the diagnosis and management of polycythaemia vera. A British Society for Haematology Guideline. Br J Haematol 2019;184:176-191.
2. Diehn F, Tefferi A. Pruritus in polycythaemia vera: prevalence, laboratory correlates and management. British journal of haematology 2001;115:619-621.
3. Siegel FP, Tauscher J, Petrides PE. Aquagenic pruritus in polycythemia vera: characteristics and influence on quality of life in 441 patients. American Journal of Hematology 2013;88:665-9.
4. Chapman J, Goyal A, Azevedo AM. Splenomegaly. StatPearls. Treasure Island (FL): StatPearls Publishing., 2022.
5. Tefferi A, Rumi E, Finazzi G, et al. Survival and prognosis among 1545 patients with contemporary polycythemia vera: an international study. Leukemia 2013;27:1874-81.
6. Griesshammer M, Gisslinger H, Mesa R. Current and future treatment options for polycythemia vera. Annals of hematology 2015;94:901-910.
7. Moulard O, Mehta J, Olivares R, et al. Epidemiology of Myelofibrosis (MF), Polycythemia Vera (PV) and Essential Thrombocythemia (ET) in the European Union (EU). Blood 2012;120:1744.
8. Jentsch-Ullrich K, Eberhardt J, Zeremski V, et al. Characteristics and treatment of polycythemia vera patients in clinical practice: a multicenter chart review on 1476 individuals in Germany. Journal of Cancer Research and Clinical Oncology 2016;142:2041-2049.
9. Alvarez-Larran A, Kerguelen A, Hernandez-Boluda JC, et al. Frequency and prognostic value of resistance/intolerance to hydroxycarbamide in 890 patients with polycythaemia vera. British Journal of Haematology 2016;172:786-93.
10. Marchetti M, Vannucchi AM, Griesshammer M, et al. Appropriate management of polycythaemia vera with cytoreductive drug therapy: European LeukemiaNet 2021 recommendations. The Lancet Haematology 2022;9:e301-e311.
11. Novartis. Data on File. Novartis UK Polycythemia Vera Advisory Board. 24th June 2022.
12. Alvarez-Larran A, Pereira A, Cervantes F, et al. Assessment and prognostic value of the European LeukemiaNet criteria for clinicohematologic response, resistance, and intolerance to hydroxyurea in polycythemia vera. Blood 2012;119:1363-9.
13. Parasuraman SV, Shi N, Paranagama DC, et al. Health Care Costs and Thromboembolic Events in Hydroxyurea-Treated Patients with Polycythemia Vera. Journal of Managed Care & Specialty Pharmacy 2018;24:47-55.
14. Medicines and Healthcare products Regulatory Agency. Jakavi® 10 mg tablets: Summary of Product Characteristics. Available at: https://mhraproducts4853.blob.core.windows.net/docs/d61f7a73d3447fbfcf6d3936bed37a80946f480f [Last Accessed 08 September 2022].
15. National Institute for Health and Care Excellence (NICE). Ruxolitinib for treating polycythaemia vera (review of TA356): Final scope. Available at: https://www.nice.org.uk/guidance/gid-ta11079/documents/final-scope-2. [Last Accessed 28 August 2022].
16. European Medicines Agency (EMA): Jakavi® European Public Assessment Report (EPAR) - Product Information. Available at: https://www.ema.europa.eu/en/documents/overview/jakavi-epar-summary-public_en.pdf [Last Accessed 10th June 2022].
17. Blum S, Martins F, Alberio L. Ruxolitinib in the treatment of polycythemia vera: patient selection and special considerations. J Blood Med 2016;7:205-215.
18. Tefferi A. Novel mutations and their functional and clinical relevance in myeloproliferative neoplasms: JAK2, MPL, TET2, ASXL1, CBL, IDH and IKZF1. Leukemia 2010;24:1128-38.
19. Levine RL, Pardanani A, Tefferi A, et al. Role of JAK2 in the pathogenesis and therapy of myeloproliferative disorders. Nat Rev Cancer 2007;7:673-83.
20. Lu X, Chang R. Polycythemia Vera. StatPearls. Treasure Island (FL): StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC., 2022.
21. Ajayi S, Becker H, Reinhardt H, et al. Ruxolitinib. Recent Results Cancer Res 2018;212:119-132.
22. Haq M, Adnan G. Ruxolitinib. StatPearls. Treasure Island (FL): StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC., 2022.

23. NICE. TA386: Ruxolitinib for disease-related splenomegaly or symptoms in adults with myelofibrosis (review of TA289). Available at: https://www.nice.org.uk/guidance/ta386 [Last Accessed 05 August 2022].
24. BNF. British National Formulary. Volume 2022: Joint Formulary Committee, 2022.
25. NICE. NICE health technology evaluations: the manual. Available at: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation [Last Accessed 05 August 2022].
26. Stein BL, Moliterno AR, Tiu RV. Polycythemia vera disease burden: contributing factors, impact on quality of life, and emerging treatment options. Ann Hematol 2014;93:1965-76.
27. Barbui T, Barosi G, Birgegard G, et al. Philadelphia-negative classical myeloproliferative neoplasms: critical concepts and management recommendations from European LeukemiaNet. Journal of Clinical Oncology 2011;29:761.
28. Anía BJ, Suman VJ, Sobell JL, et al. Trends in the incidence of polycythemia vera among Olmsted County, Minnesota residents, 1935–1989. American journal of hematology 1994;47:89-93.
29. Chaiter Y, Brenner B, Aghai E, et al. High incidence of myeloproliferative disorders in Ashkenazi Jews in northern Israel. Leukemia & lymphoma 1992;7:251-255.
30. Moulard O, Mehta J, Fryzek J, et al. Epidemiology of myelofibrosis, essential thrombocythemia, and polycythemia vera in the E uropean U nion. European journal of haematology 2014;92:289-297.
31. McMullin MF. Idiopathic erythrocytosis: a disappearing entity. Hematology Am Soc Hematol Educ Program 2009:629-35.
32. Pearson T, Wetherley-Mein G. Vascular occlusive episodes and venous haematocrit in primary proliferative polycythaemia. The Lancet 1978;312:1219-1222.
33. Barbui T, Carobbio A, Rambaldi A, et al. Perspectives on thrombosis in essential thrombocythemia and polycythemia vera: is leukocytosis a causative factor? Blood 2009;114:759-63.
34. Landolfi R, Di Gennaro L, Barbui T, et al. Leukocytosis as a major thrombotic risk factor in patients with polycythemia vera. Blood 2007;109:2446-2452.
35. Bonicelli G, Abdulkarim K, Mounier M, et al. Leucocytosis and thrombosis at diagnosis are associated with poor survival in polycythaemia vera: a population-based study of 327 patients. British journal of haematology 2013;160:251-254.
36. Spivak JL. Polycythaemia vera. Hematology 2013;18:244-245.
37. Mesa RA, Niblack J, Wadleigh M, et al. The burden of fatigue and quality of life in myeloproliferative disorders (MPDs) An international internet-based survey of 1179 MPD patients. Cancer 2007;109:68-76.
38. Scherber R, Dueck AC, Johansson P, et al. The Myeloproliferative Neoplasm Symptom Assessment Form (MPN-SAF): international prospective validation and reliability trial in 402 patients. Blood 2011;118:401-408.
39. Emanuel RM, Dueck AC, Geyer HL, et al. Myeloproliferative neoplasm (MPN) symptom assessment form total symptom score: prospective international assessment of an abbreviated symptom burden scoring system among patients with MPNs. Journal of Clinical Oncology 2012;30:4098.
40. Cerquozzi S, Barraco D, Lasho T, et al. Risk factors for arterial versus venous thrombosis in polycythemia vera: a single center experience in 587 patients. Blood Cancer J 2017;7:662.
41. Accurso V, Santoro M, Raso S, et al. Splenomegaly Impacts Prognosis in Essential Thrombocythemia and Polycythemia Vera: A Single Center Study. Hematology Reports 2019;11:8281.
42. Hensley B, Geyer H, Mesa R. Polycythemia vera: current pharmacotherapy and future directions. Expert Opin Pharmacother 2013;14:609-17.
43. Steinman HK, Greaves MW. Aquagenic pruritus. Journal of the American Academy of Dermatology 1985;13:91-96.
44. Mesa R, Miller CB, Thyne M, et al. Impact of myeloproliferative neoplasms (MPNs) on patients' overall health and productivity: results from the MPN LANDMARK SURVEY in the United States: American Society of Hematology Washington, DC, 2014.
45. Harrison C, Pemberton-Whiteley Z, Ali S, et al. The International Landmark Health Survey – MPNs. UK Results – Patient perceptions on the symptomatic and economic burden of Myeloproliferative Neoplasms. Presented at BSH 2017.

46. Finazzi G, Barbui T. Evidence and expertise in the management of polycythemia vera and essential thrombocythemia. Leukemia 2008;22:1494-1502.
47. Landolfi R, Nicolazzi MA, Porfidia A, et al. Polycythemia vera. Intern Emerg Med 2010;5:375-84.
48. Stuart BJ, Viera AJ. Polycythemia vera. Am Fam Physician 2004;69:2139-44.
49. Tefferi A. Essential thrombocythemia, polycythemia vera, and myelofibrosis: current management and the prospect of targeted therapy. Am J Hematol 2008;83:491-7.
50. Passamonti F, Rumi E, Pungolino E, et al. Life expectancy and prognostic factors for survival in patients with polycythemia vera and essential thrombocythemia. The American journal of medicine 2004;117:755-761.
51. Barosi G, Birgegard G, Finazzi G, et al. A unified definition of clinical resistance and intolerance to hydroxycarbamide in polycythaemia vera and primary myelofibrosis: results of a European LeukemiaNet (ELN) consensus process. Br J Haematol 2010;148:961-3.
52. Vannucchi AM. How I treat polycythemia vera. Blood 2014;124:3212-20.
53. Price G, Pohl G, Xie J, et al. A retrospective observational study of annual healthcare costs for patients with forms of myeloproliferative neoplasms (MPN). Blood (ASH Annual Meeting Abstracts) 2011;118.
54. Kosunen M, Kurki S, Snicker K, et al. Resource Use, Health Care Costs And Burden Of Disease In Patients With Polycythemia Vera In Finland: Evaluation Based On Leukocyte Counts Using Finnish Real World Data From Auria Biobank. Value in Health 2017;20:PA499.
55. Harrison C, Mathias J, Campbell-Drew M, et al. UK results from the myeloproliferative neoplasms (MPN) landmark survey on the symptom, emotional and economic burden of MPN. British Journal of Haematology 2019;186:e1-e4.
56. Tefferi A. Polycythemia vera and essential thrombocythemia: 2013 update on diagnosis, risk-stratification, and management. American journal of hematology 2013;88:507-516.
57. Reiter A, Harrison C. How we identify and manage patients with inadequately controlled polycythemia vera. Current hematologic malignancy reports 2016;11:356-367.
58. Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera. N Engl J Med 2015;372:426-35.
59. Kiladjian J-J, Chevret S, Dosquet C, et al. Treatment of polycythemia vera with hydroxyurea and pipobroman: final results of a randomized trial initiated in 1980. Journal of Clinical Oncology 2011;29:3907-3913.
60. Passamonti F, Griesshammer M, Palandri F, et al. Ruxolitinib for the treatment of inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): a randomised, open-label, phase 3b study. The Lancet Oncology 2017;18:88-99.
61. Samuelsson J, Hasselbalch H, Bruserud O, et al. A phase II trial of pegylated interferon α-2b therapy for polycythemia vera and essential thrombocythemia: feasibility, clinical and biologic effects, and impact on quality of life. Cancer 2006;106:2397-2405.
62. Najean Y, Rain J-D. Treatment of polycythemia vera: the use of hydroxyurea and pipobroman in 292 patients under the age of 65 years. Blood, The Journal of the American Society of Hematology 1997;90:3370-3377.
63. Griffin S. Covid-19: Waiting times in England reach record highs. BMJ 2020;370:m3557.
64. Scottish Medicines Consortium (SMC): ruxolitinib phosphate 5mg, 10mg, 15mg, 20mg tablets (Jakavi®). Available at: https://www.scottishmedicines.org.uk/media/4942/ruxolitinib-jalkavi-final-november-2019-for-website.pdf [Last Accessed 10th June 2022].
65. Kiladjian J-J, Zachee P, Hino M, et al. Long-term efficacy and safety of ruxolitinib versus best available therapy in polycythaemia vera (RESPONSE): 5-year follow up of a phase 3 study. The Lancet Haematology 2020;7:e226-e237.
66. Passamonti F, Palandri F, Saydam G, et al. Ruxolitinib versus best available therapy in inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): 5-year follow up of a randomised, phase 3b study. The Lancet Haematology 2022;9:E480-E492.
67. Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera. New England Journal of Medicine 2015;372:426-435.
68. Novartis. Data on File: Propensity Score Matching Analysis: Overall Survival at Week 256. Ruxolitinib (RESPONSE) versus BAT (GEMFIN).
69. Verstovsek S, Vannucchi AM, Griesshammer M, et al. Ruxolitinib versus best available therapy in patients with Polycythemia Vera: 80 Week follow up from the RESPONSE trial. Haematologica 2016:haematol. 2016.143644.

70. Harrison C NJ, Boucher R, et al. Ruxolitinib versus Best Available Therapy for PV intolerant or resistant to Hydroxycarbamide in a Randomized trial (Manuscript in Development), 2022.
71. European Medicines Agency (EMA): Jakavi® Committee for Medicinal Products for Human Use (CHMP) Assessment Report. Available at: https://www.ema.europa.eu/documents/variation-report/jakavi-h-c-2464-ii-0016-epar-assessment-report-variation_en.pdf [Last Accessed 09 September 2022].
72. Vannucchi AM, Verstovsek S, Guglielmelli P, et al. Ruxolitinib reduces JAK2 p.V617F allele burden in patients with polycythemia vera enrolled in the RESPONSE study. Ann Hematol 2017;96:1113-1120.
73. ClinicalTrials.gov: Study of Efficacy and Safety in Polycythemia Vera Subjects Who Are Resistant to or Intolerant of Hydroxyurea: JAK Inhibitor INC424 (INCB018424) Tablets Versus Best Available Care (RESPONSE; NCT01243944). Available at: https://clinicaltrials.gov/ct2/show/NCT01243944 [Last Accessed 09 September 2022].
74. Kirito K, Suzuki K, Miyamura K, et al. Ruxolitinib is effective and safe in Japanese patients with hydroxyurea-resistant or hydroxyurea-intolerant polycythemia vera with splenomegaly. Int J Hematol 2018;107:173-184.
75. Kiladjian J-J, Verstovsek S, Griesshammer M, et al. Results from the 208-week (4-Year) follow-up of RESPONSE trial, a phase 3 study comparing ruxolitinib (Rux) with best available therapy (BAT) for the treatment of polycythemia vera (PV): Am Soc Hematology, 2017.
76. Verstovsek S, Harrison CN, Kiladjian J-J, et al. Markers of iron deficiency in patients with polycythemia vera receiving ruxolitinib or best available therapy. Leukemia research 2017;56:52-59.
77. Kiladjian J-J, Guglielmelli P, Griesshammer M et al. European Hematology Association (EHA) 2016; Abstract E1360. Ruxolitinib treatment following interferon in patients with polycythemia vera: An analysis from the RESPONSE trial. Available at: https://bit.ly/2Ucv9Fn [Last Accessed 09 September 2022].
78. Samuelson BT, Vesely SK, Chai-Adisaksopha C, et al. The impact of ruxolitinib on thrombosis in patients with polycythemia vera and myelofibrosis: a meta-analysis. Blood Coagul Fibrinolysis 2016;27:648-52.
79. Mesa R, Verstovsek S, Kiladjian JJ, et al. Changes in quality of life and disease-related symptoms in patients with polycythemia vera receiving ruxolitinib or standard therapy. Eur J Haematol 2016;97:192-200.
80. Barosi G. Ruxolitinib for patients with polycythemia vera who have had an inadequate response or are intolerant to hydroxyurea: A critical appraisal. Clinical Investigation 2015;5:643-651.
81. Harrison CN, Griesshammer M, Miller C, et al. Comprehensive haematological control with ruxolitinib in patients with polycythaemia vera resistant to or intolerant of hydroxycarbamide. British Journal of Haematology 2018;182:279-284.
82. Verstovsek S, Kiladjian J-J, Wroclawska M, et al. Ruxolitinib for Patients (Pts) with Polycythemia Vera: Responders Vs Non-Responders As Defined in the Response Trial. Blood 2019;134:2947.
83. Passamonti F, Palandri F, Saydam G et al. American Society of Hematology (ASH) 2018; Abstract 1754. Ruxolitinib for the Treatment of Inadequately Controlled Polycythemia Vera without Splenomegaly: 156-Week Follow-up from the Phase 3 RESPONSE-2 Study.
84. Griesshammer M, Saydam G, Palandri F, et al. Ruxolitinib for the treatment of inadequately controlled polycythemia vera without splenomegaly: 80-week follow-up from the RESPONSE-2 trial. Ann Hematol 2018;97:1591-1600.
85. Passamonti F, Griesshammer M, Palandri F et al. European Hematology Association (EHA) 2016; Oral Presentation S112. Ruxolitinib proves superior to best available therapy in patients with polycythemia vera (PV) and a non-palpable spleen: results from the phase IIIb RESPONSE-2 study. Available at: https://bit.ly/2FSAIPw [Last Accessed 09 September 2022].
86. Griesshammer M, Saydam G, Palandri F et al. European Hematology Association (EHA) 2017; Oral Presentation S784. Ruxolitinib for the treatment of inadequately controlled polycythemia vera without splenomegaly: 80-week follow-up from the RESPONSE-2 trial. Available at: https://bit.ly/2I6ga8I [Last Accessed 09 September 2022].
87. Passamonti F, Palandri F, Saydam G, et al. Long-term effect of ruxolitinib (RUX) in inadequately controlled polycythemia vera (PV) without splenomegaly: 5-year results from the phase 3 response-2 study. Blood 2020;136:40-41.

88. Kiladjian JJ, Guglielmelli P, Griesshammer M, et al. Efficacy and safety of ruxolitinib after and versus interferon use in the RESPONSE studies. Ann Hematol 2018;97:617-627.
89. ClinicalTrials.gov: Ruxolitinib Efficacy and Safety in Patients With HU Resistant or Intolerant Polycythemia Vera vs Best Available Therapy (RESPONSE-2; NCT02038036). Available at: https://clinicaltrials.gov/ct2/show/NCT02038036 [Last Accessed 09 September 2022].
90. A Alvarez-Larran SV, M Perez-Encinas, JJ Kiladjian, F Ferrer-Marin, A Vannucchi, J Martinez; C Besses, C Cameron, A Situ, A Charter, JC Hernandez-Boluda. Comparison Of Ruxolitinib And Real-World Best Available Therapy in Terms Of Overall Survival And Thrombosis In Patients With Polycythemia Vera Who Are Resistant Or Intolerant To Hydroxyurea. European Hematology Association Congress. Stockholm, Sweden, 2018.
91. Kiladjian J-J, Zachée P, Hino M et al. American Society of Hematology (ASH) 2018; Abstract 1753. Long-Term Efficacy and Safety (5 Years) in RESPONSE, a Phase 3 Study Comparing Ruxolitinib (rux) with Best Available Therapy (BAT) in Hydroxyurea (HU)-Resistant/Intolerant Patients (pts) with Polycythemia Vera (PV).
92. Harrison C, McMullin MF, Panchal A, et al. PF625: Ruxolitinib compared with best available therapy for polycythaemia vera patients resistant or intolerant to hydroxycarbamide in MAJIC - an investigator-led randomised trial. 23rd European Hematology Association Congress. Stockholm, Sweden, 2018.
93. Mesa R, Vannucchi AM, Yacoub A, et al. The efficacy and safety of continued hydroxycarbamide therapy versus switching to ruxolitinib in patients with polycythaemia vera: a randomized, double-blind, double-dummy, symptom study (RELIEF). British journal of haematology 2017;176:76-85.
94. ClinicalTrials.gov: Randomized Switch Study From Hydroxyurea to Ruxolitinib for RELIEF of Polycythemia Vera Symptoms: the Relief Study (NCT01632904). Available at https://clinicaltrials.gov/ct2/show/NCT01632904 [Last Accessed on 09 September 2022]
95. Novartis Data on File. MAJIC Clinical Study Protocol. 2018.
96. Novartis. Data on File. Interim Clinical Study Report Week 208: Randomized, open label, multicenter phase III study of efficacy and safety in polycythemia vera subjects who are resistant to or intolerant of hydroxyurea: JAK inhibitor INC424 tablets versus best available care (The RESPONSE Trial).
97. Novartis. Data on File. Final Clinical Study Report (Week 256): Randomized, open label, multicenter Phase III study of efficacy and safety in polycythemia vera subjects who are resistant to or intolerant of hydroxyurea: JAK inhibitor INC424 tablets versus best available care (The RESPONSE Trial).
98. Novartis. Data on File. Final Clinical Study Report (Week 260): Randomized, open-label, multi-center phase IIIb study evaluating the efficacy and safety of ruxolitinib versus best available therapy in patients with polycythemia vera who are hydroxyurea resistant or intolerant (RESPONSE-2).
99. Novartis. Data on File. Protocol CINC424B2401: Randomized, open label, multicenter phase IIIb study evaluating the efficacy and safety of ruxolitinib versus best available therapy in patients with Polycythemia Vera who are hydroxyurea resistant or intolerant (RESPONSE-2).
100. Passamonti F, Griesshammer M, Palandri F, et al. Ruxolitinib for the treatment of inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): a randomised, open-label, phase 3b study (supplementary appendix). The Lancet Oncology 2017;18:88-99.
101. Novartis. Data on File. Interim Clinical Study Report (Week 48): Randomized, open label, multicenter phase III study of efficacy and safety in polycythemia vera subjects who are resistant to or intolerant of hydroxyurea: JAK inhibitor INC424 tablets versus best available care (The RESPONSE Trial).
102. Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera (supplementary appendix). New England Journal of Medicine 2015;372:426-435.
103. Alvarez-Larran A, Verstovsek S, Pérez-Encinas M, et al. PF628: Comparison of ruxolitinib and real-world best available therapy in terms of overall survival and thrombosis in patients with polycythemia vera who are resistant or intolerant to hydroxyurea. 23rd European Hematology Association Congress. Stockholm, Sweden, 2018.

104. Alvarez-Larrán A, Pereira A, Cervantes F, et al. Assessment and prognostic value of the European LeukemiaNet criteria for clinicohematologic response, resistance, and intolerance to hydroxyurea in polycythemia vera. Blood 2012;119:1363-9.
105. NICE. Ruxolitinib for treating polycythaemia vera (review of TA356): Final scope. Available at: https://www.nice.org.uk/guidance/gid-ta11079/documents/final-scope-2 [Last Accessed 26 August 2022].
106. Novartis. Data on File. Novartis UK Polycythemia Vera Advisory Board. 08th August 2022.
107. NICE. TA756: Fedratinib for treating disease-related splenomegaly or symptoms in myelofibrosis. Available at: https://www.nice.org.uk/guidance/ta756 [Last Accessed 05 August 2022].
108. NICE. TA393: Alirocumab for treating primary hypercholesterolaemia and mixed dyslipidaemia. Available at: https://www.nice.org.uk/guidance/ta393 [Last Accessed 05 August 2022].
109. NICE. TA394: Evolocumab for treating primary hypercholesterolaemia and mixed dyslipidaemia. Available at: https://www.nice.org.uk/guidance/ta394 [Last Accessed 05 August 2022].
110. NICE. TA341: Apixaban for the treatment and secondary prevention of deep vein thrombosis and/or pulmonary embolism. Available at: https://www.nice.org.uk/guidance/ta341 [Last Accessed 05 August 2022].
111. NICE. TA523: Midostaurin for untreated acute myeloid leukaemia. Available at: https://www.nice.org.uk/guidance/ta523 [Last Accessed 05 August 2022].
112. A Alvarez-Larrán, M Garrote, F Ferrer-Marín, et al. Real-world analysis of main clinical outcomes in patients with polycythemia vera treated with ruxolitinib or best available therapy after developing resistance/intolerance to hydroxyurea. Cancer 2022;128:2441-2448.
113. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med 2007;26:2389-430.
114. Department of Health. National Schedule of Reference Costs 2020-2021. Available at: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/#ncc1819 [Last Accessed 05 August 2022].
115. Curtis LB, Amanda. Unit Costs of Health and Social Care 2021. Available at: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/ [Last Accessed 08 September 2022].
116. GOV.UK Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT). Available at: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [Last Accessed 08 September 2022].
117. Office for National Statistics National life tables – life expectancy in the UK: 2018 to 2020. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2018to2020#:~:text=Across%20the%20UK%2C%20life%20expectancy,years%20for%20females%20in%20Northern [Last Accessed 05 August 2022].
118. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data, 2011.
119. Jackson CM, Paul; Amdahl, Jordan. flexsurv: Flexible Parametric Survival and Multi-State Models v2.2. cran.r-project, 2022.
120. Rafia R, Mukuria CW. PCN127 - Exploratory Psychometric Analysis of The Eq-5d In A Myelofibrosis Population. Value in Health 2015;18:A211.
121. Mukuria C, Rowen D, Brazier JE, et al. Deriving a Preference-Based Measure for Myelofibrosis from the EORTC QLQ-C30 and the MF-SAF. Value in Health 2015;18:846-855.
122. Mesa R, Miller CB, Thyne M, et al. Myeloproliferative neoplasms (MPNs) have a significant impact on patients' overall health and productivity: the MPN Landmark survey. BMC Cancer 2016;16:167.
123. Mesa RA SS, Radia D, Cheville A, Hussein K, Niblack J, Pardanani AD, Steensma DP, Litzow MR, Rivera CE, Camoriano J, Verstovsek S, Sloan J, Harrison C, Kantarjian H, Tefferi A. The Myelofibrosis Symptom Assessment Form (MFSAF): an evidence-based brief inventory to measure quality of life and symptomatic response to treatment in myelofibrosis. Leuk Res. 2009 33.

124. Incyte. Myeloproliferative Neoplasm Symptom Assessment Form. Available at: https://www.jakafi.com/pdf/mpn-symptoms-form.pdf [Last Accessed 09 September 2022]. 2022.
125. Hernández Alava MP, Steve; Wailoo, Allan. Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from an English Population Study.: Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, UK, 2020.
126. Lelonek E ML, Wrobel T, et al. . Burden of Aquagenic Pruritus in Polycythaemia Vera. Acta Derm Venereol 2018;98:185-190.
127. Scottish Medicines Consortium (SMC). Ropeginterferon alfa-2b 250 micrograms/0.5 mL solution for injection in pre-filled pen (Besremi®). Avaliable at: https://www.scottishmedicines.org.uk/medicines-advice/ropeginterferon-alfa-2b-besremi-full-smc2421/. [Last Accessed 08 September 2022].
128. NICE. TA772: Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma after stem cell transplant or at least 2 previous therapies. Available at: https://www.nice.org.uk/guidance/ta772 [Last Accessed 05 August 2022].
129. NICE. TA722: Pemigatinib for treating relapsed or refractory advanced cholangiocarcinoma with FGFR2 fusion or rearrangement. Available at: https://www.nice.org.uk/guidance/ta722 [Last Accessed 05 August 2022].
130. Wailoo MHASPA. Estimating EQ-5D by age and sex for the UK: The University of Sheffield, 2022.
131. Gorry C. Cost Effectiveness of Treatment Strategies for Skin Cancer in the Irish Healthcare Setting. Trinity's Access to Research Archive: Trinity College Dublin, 2020.
132. Chihara D, Kantarjian HM, Newberry KJ, et al. Survival Outcome of Patients with Acute Myeloid Leukemia Transformed from Myeloproliferative Neoplasms. Blood 2016;128:1940-1940.
133. Mead AJ, Butt NM, Nagi W, et al. A retrospective real-world study of the current treatment pathways for myelofibrosis in the United Kingdom: the REALISM UK study. Therapeutic Advances in Hematology 2022;13:20406207221084487.
134. Briggs AH, Bhatt DL, Scirica BM, et al. Health-related quality-of-life implications of cardiovascular events in individuals with type 2 diabetes mellitus: A subanalysis from the Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus (SAVOR)-TIMI 53 trial. Diabetes Res Clin Pract 2017;130:24-33.
135. Tolley K, Oliver N, Miranda E, et al. Cost effectiveness of deferasirox compared to desferrioxamine in the treatment of iron overload in lower-risk, transfusion-dependent myelodysplastic syndrome patients. J Med Econ 2010;13:559-70.
136. Doble B, Pufulete M, Harris JM, et al. Health-related quality of life impact of minor and major bleeding events during dual antiplatelet therapy: a systematic literature review and patient preference elicitation study. Health Qual Life Outcomes 2018;16:191.
137. Mesa RA, Schaap N, Vannucchi AM, et al. Patient-reported Effects of Fedratinib, an Oral, Selective Inhibitor of Janus Kinase 2, on Myelofibrosis-related Symptoms and Health-related Quality of Life in the Randomized, Placebo-controlled, Phase III JAKARTA Trial. Hemasphere 2021;5:e553.
138. Tefferi A, Lasho TL, Jimma T, et al. One thousand patients with primary myelofibrosis: the mayo clinic experience. Mayo Clin Proc 2012;87:25-33.
139. Yacoub A, Mascarenhas J, Kosiorek H, et al. Pegylated interferon alfa-2a for polycythemia vera or essential thrombocythemia resistant or intolerant to hydroxyurea. Blood 2019;134:1498-1509.
140. Thames Valley Strategic Clinical Networks. Anagrelide. Available at: http://nssg.oxford-haematology.org.uk/myeloid/protocols/ML-69-anagrelide.pdf [Last Accessed 09 September 2022].
141. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: A modelling study. Palliat Med 2015;29:899-907.
142. Wang H-I, Aas E, Howell D, et al. Long-Term Medical Costs and Life Expectancy of Acute Myeloid Leukemia: A Probabilistic Decision Model. Value in Health 2014;17:205-214.
143. Vallejo-Torres L, Morris S, Kinge JM, et al. Measuring current and future cost of skin cancer in England. J Public Health (Oxf) 2014;36:140-8.
144. Crathorne L, Reimer E, Harris S, et al. PSY79 - The Cost Of Major Bleeding Events Associated With Anticoagulation Treatment - An Estimation Of The Contribution Of The

Reversal Agents To The Overall Cost Of Major Bleeding In A UK Setting. Value in Health 2018;21:S449.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Ruxolitinib for treating polycythaemia vera

ID5106

Summary of Information for Patients (SIP)

September 2022

File name	Version	Contains confidential information	Date
ID5106 Ruxolitinib_SIP_15Sep2022	Final	No	15th September 2022

Summary of Information for Patients (SIP):

The pharmaceutical company perspective

What is the SIP?

The Summary of Information for Patients (SIP) is written by the company who is seeking approval from NICE for their treatment to be sold to the NHS for use in England. It is a plain English summary of their submission written for patients participating in the evaluation. It is not independently checked, although members of the public involvement team at NICE will have read it to double-check for marketing and promotional content before it is sent to you.

The **Summary of Information for Patients** template has been adapted for use at NICE from the Health Technology Assessment International – Patient & Citizens Involvement Group (HTAi PCIG). Information about the development is available in an open-access IJTAHC journal article

SECTION 1: Submission summary

1a) Name of the medicine (generic and brand name):

Generic name: Ruxolitinib **Brand name:** Jakavi®

1b) Population this treatment will be used by. Please outline the main patient population that is being appraised by NICE:

Adults with polycythaemia vera (PV) that are resistant or intolerant (R/I) to hydroxyurea (HU), which is also sometimes called hydroxycarbamide (HC). If someone is resistant to HC/HU, that means that HC/HU does not work anymore to treat PV. If they are intolerant to HC/HU, that means that they experience unacceptable side effects when taking the treatment.

1c) Authorisation: Please provide marketing authorisation information, date of approval and link to the regulatory agency approval. If the marketing authorisation is pending, please state this, and reference the section of the company submission with the anticipated dates for approval.

In 2015, the European Medicines Agency granted marketing authorisation (approval) for ruxolitinib for the treatment of adult patients with PV who are R/I of HC/HU. The marketing authorisation approval is available through the following web-link: https://www.ema.europa.eu/en/medicines/human/EPAR/jakavi.[1]

1d) Disclosures. Please be transparent about any existing collaborations (or broader conflicts of interest) between the pharmaceutical company and patient groups relevant to the medicine. Please outline the reason and purpose for the engagement/activity and any financial support provided:

Novartis has a long-standing partnership with the MPN Voice mainly directed on patient support and disease awareness projects around myelofibrosis. The nature of collaboration is concentrated around counselling on Novartis projects, joint non-promotional disease awareness activities and funding through grants and sponsorship for the patient advocacy groups.

SECTION 2: Current landscape

2a) The condition – clinical presentation and impact

Please provide a few sentences to describe the condition that is being assessed by NICE and the number of people who are currently living with this condition in England.

Please outline in general terms how the condition affects the quality of life of patients and their families/caregivers. Please highlight any mortality/morbidity data relating to the condition if available. If the company is making a case for the impact of the treatment on carers this should be clearly stated and explained.

What is PV?

PV is usually caused by a change in the genes called Janus kinase-1 and 2 (JAK1 and 2), which causes the bone marrow cells to produce too many red blood cells. This leads to a high concentration of red blood cells in the blood. This makes the blood thicker and less able to travel through blood vessels and organs. The affected bone marrow cells can also develop into other cells found in the blood, which means that people with PV may also have high numbers of both platelets and white bloods cells.[2]

Symptoms of PV

Many of the symptoms of PV are caused by this slow flow of blood. These symptoms include:[2]

- Headaches
- Blurred vision
- Red skin – particularly in the face, hands and feet
- Tiredness
- High blood pressure
- Dizziness
- Discomfort in the tummy
- Confusion
- Bleeding problems – such as nosebleeds and bruising
- Gout – which can cause joint pain, stiffness and swelling
- Itchy skin – especially after a bath or shower
- An enlarged spleen[3]
- Weight loss[3]

The symptoms of PV can have a large impact on the quality of life of patients.[4, 5] Itching skin can be very difficult to tolerate and has been linked to negative emotions such as aggression, irritability and depression.[6, 7]

Complications of PV

PV leads to an increased risk of complications, such as blood clots, heart attacks, myelofibrosis (scarring of the bone marrow) and acute myeloid leukaemia (a type of blood cancer).[8, 9] These complications lead to an increased risk of death compared with the general population.[10] The average length of life after diagnosis is approximately 14 years.[11]

How common is PV?

PV is rare and normally develops in adulthood. The average age at diagnosis is 60 years.[11, 12] In the UK, prevalence (frequency) is 6.05 cases per 100,000 people.[13] The incidence (new cases diagnosed each year) is 0.5–2.2 per 100,000 people each year.[13]

2b) Diagnosis of the condition (in relation to the medicine being evaluated)

Please briefly explain how the condition is currently diagnosed and how this impacts patients. Are there any additional diagnostic tests required with the new treatment?

PV can be diagnosed by carrying out a blood test to:

- Measure the number of red blood cells in your blood (red blood cell count) and the amount of space the red blood cells take up in the blood (haematocrit [HCT] level). A high concentration of red blood cells suggests you may have PV.[2]
- Look for the changed JAK2 gene.[2]

2c) Current treatment options:

The purpose of this section is to set the scene on how the condition is currently managed:

- What is the treatment pathway for this condition and where in this pathway the medicine is likely to be used? Please use diagrams to accompany text where possible. Please give emphasis to the specific setting and condition being considered by NICE in this review. For example, by referencing current treatment guidelines. It may be relevant to show the treatments people may have before and after the treatment under consideration in this SIP.
- Please also consider:
  - if there are multiple treatment options, and data suggest that some are more commonly used than others in the setting and condition being considered in this SIP, please report these data.
  - are there any drug–drug interactions and/or contraindications that commonly cause challenges for patient populations? If so, please explain what these are.

Treatment guidelines

In the UK, treatment of people with PV is directed by guidelines from the British Society for Haematology (BSH).[14]

What are the aims of treatment?

Medicine is prescribed to slow down the production of red blood cells. Slowing the production of red blood cells reduces the risk of complications, such as blood clots, heart attacks and blood cancer development.[15, 16]

What treatments are used?

The BSH guidelines recommend all patients with PV receive aspirin and phlebotomy to keep the HCT level (the space the red blood cells take up in the blood) normal.[14]

People who are aged 65 years or older or have experienced a blood clot in the vein or artery related to PV (arterial or venous thrombosis) are considered 'high-risk'. These patients will be prescribed best available therapy (BAT) in addition to aspirin and phlebotomy.[14]

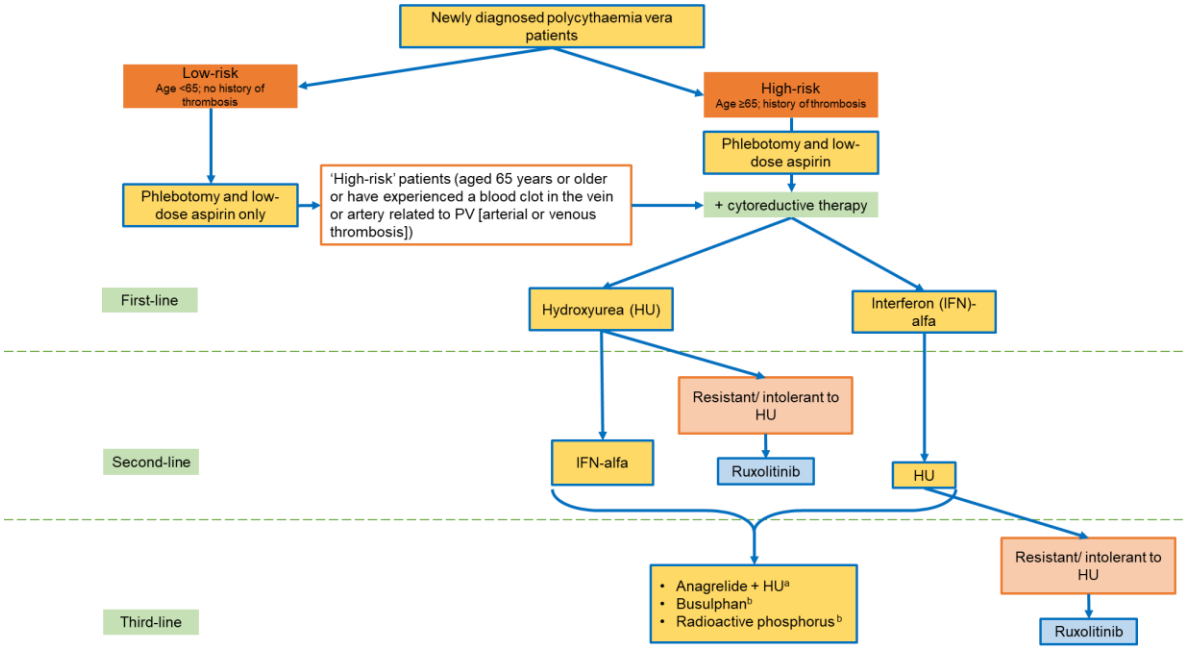
BAT is a combination of different medicines. The medicines that may be included in BAT are cytoreductive therapies, including HC/HU, interferon-alfa (IFN-alfa), anagrelide, and busulfan.[14]

HC/HU is usually the first treatment given to patients with PV who are 'high-risk'. Sometimes IFN-alfa is given instead. Between 15% and 32% of patients with PV develop R/I to HC/HU.[15, 16]

There are currently no treatments recommended by NICE for patients with PV who are R/I to HC/HU. However, the BSH guidelines recommend other cytoreductive therapies in later lines of therapy, including IFN-alfa, busulfan, and continued use of HC/HU.[14] These treatments do not fully help with symptoms or the impact on quality of life. These treatments also have unpleasant side effects.[14] Patients who receive these treatments may also continue to have additional complications, including blood cancer.[17] Therefore, patients who are R/I to HC/HU have a worse life expectancy compared with patients who are not R/I.[18, 19]

A summary of the treatment pathway for patients with PV who are R/I to HC/HU is shown in Figure 1.

Figure 1 Current treatment pathway for patients with PV who are R/I to HC/HU



[a]May be helpful in those where platelet control is difficult. [b]Only recommended for those with a limited life expectancy.
Abbreviations: HC/HU: hydroxycarbamide/hydroxyurea; IFN: interferon; PV: polycythaemia vera.
Source: McMullin et al. 2019.[14]

2d) Patient-based evidence (PBE) about living with the condition

Context:

- **Patient-based evidence (PBE)** is when patients input into scientific research, specifically to provide experiences of their symptoms, needs, perceptions, quality of life issues or experiences of the medicine they are currently taking. PBE might also include carer burden and outputs from patient preference studies, when conducted in order to show what matters most to patients and carers and where their greatest needs are. Such research can inform the selection of patient-relevant endpoints in clinical trials.

In this section, please provide a summary of any PBE that has been collected or published to demonstrate what is understood about **patient needs and disease experiences**. Please include the methods used for collecting this evidence. Any such evidence included in the SIP should be formally referenced wherever possible and references included.

The symptoms associated with PV have a significant impact on the quality of life experienced by the patient. However, few studies have been carried out to understand what is important to patients in regards to their treatment.

The International Landmark Health Survey was an online survey of patients with myeloproliferative neoplasms (MPNs – cancers originating in the bone marrow, of which PV is one) and treating physicians. The study was previously conducted in the US. Questions asked were related to disease burden, emotional and economic impact, disparities between patients and physicians on disease treatment and management. This study aimed to explore the impact of symptoms on patients' lives, particularly in terms of their work productivity and quality of life.[20]

Most patients experienced a reduction in quality of life, including those with low symptom burden or low-risk scores. A substantial proportion of patients reported impairment at work and in overall activity. This study also revealed a lack of alignment between physician and patient perceptions relating to communication and disease management, with patients often having different treatment goals than physicians. Overall, the study suggested that therapies that reduce symptom burden and improve quality of life in patients with MPNs are crucial in minimising the disease impact on patients' daily lives.[20]

SECTION 3: The treatment

3a) How does the new treatment work?

What are the important features of this treatment?

Please outline as clearly as possible important details that you consider relevant to patients relating to the mechanism of action and how the medicine interacts with the body

Where possible, please describe how you feel the medicine is innovative or novel, and how this might be important to patients and their communities.

If there are relevant documents which have been produced to support your regulatory submission such as a summary of product characteristics or patient information leaflet, please provide a link to these.

How does ruxolitinib work?

Ruxolitinib is a type of medicine called a kinase inhibitor. Ruxolitinib blocks JAK1 and JAK2 from working, which stops the signals that cause blood cells to multiply. This helps to keep the number of red blood cells, white blood cells and platelets at a normal level. This reduces the symptoms of the disease.[1]

How is this medicine important to patients and their communities?

There are currently no treatments recommended by NICE for patients with PV who are R/I to HC/HU. However, the BSH recommend other cytoreductive therapies in later lines of therapy, including IFN-alfa, busulfan, and continued use of HC/HU.[14] As described above, these treatments do not fully help with symptoms or the impact on quality of life. These treatments also have unpleasant side effects.[14] Patients who receive these treatments may also continue to have complications, including blood cancer.[17] Therefore, patients who are R/I to HC/HU have a worse life expectancy compared to patients that are not R/I.[18, 19]

Ruxolitinib is recommended by the BSH guidelines for second and third line treatment for patients who are RI to HC/HU.[14] Approval of ruxolitinib by NICE would provide patients with a treatment that would help to address the underlying causes of disease, leading to improved control of symptoms, and reduced risk of blood clots, heart attacks, blood cancer and death.[21-24]

Supporting documents

The summary of product characteristics is available here: https://www.ema.europa.eu/en/documents/product-information/jakavi-epar-product-information_en.pdf.

3b) Combinations with other medicines

Is the medicine intended to be used in combination with any other medicines?

- Yes / No

If yes, please explain why and how the medicines work together. Please outline the mechanism of action of those other medicines so it is clear to patients why they are used together.

If yes, please also provide information on the availability of the other medicine(s) as well as the main side effects.

If this submission is for a combination treatment, please ensure the sections on efficacy (3e), quality of life (3f) and safety/side effects (3g) focus on data that relate to the combination, rather than the individual treatments.

The BSH guidelines recommend all patients with PV receive aspirin and phlebotomy to keep the HCT level (the space the red blood cells take up in the blood) normal.[14]

Aspirin

- Aspirin is an antithrombotic medicine, which means that it reduces the formation of blood clots. Therefore, aspirin is used to reduce the risk of vascular complications, such as heart attacks and blood clots.[14]
- In clinical trials, the most frequently reported side effects of long-term aspirin use in patients with PV were gastrointestinal intolerance (2.8% of patients) and bleeding (4.4% of patients).[25]

Phlebotomy

- Phlebotomy is the removal of blood from the vein with a medical instrument. Phlebotomy is used to reduce the 'thickness' of the blood and therefore, reduce the symptoms of PV.[14]
- How often phlebotomy is needed will be different for each person. At first, patients may need the treatment every week, but once PV is under control patients may only need it every six to 12 weeks or less.[2]
- Phlebotomies are performed by a healthcare professional, at the hospital, on a day case basis. They are therefore inconvenient for patients. Treatment with ruxolitinib has been linked to a reduction in the need for phlebotomy.[21-24]
- Most people carry on as normal following phlebotomy, but a few people do report feeling tired for a few days following the procedure. People may have bruising to the site of phlebotomy. However, some more serious side effects may occur, such fainting. Long-term phlebotomy can cause iron deficiency, leading to fatigue and an increased number of platelets in the blood.[26]

3c) Administration and dosing

How and where is the treatment given or taken? Please include the dose, how often the treatment should be given/taken, and how long the treatment should be given/taken for.

How will this administration method or dosing potentially affect patients and caregivers? How does this differ to existing treatments?

How is ruxolitinib taken?

- Ruxolitinib is available as tablets taken orally twice a day.[27]
- Ruxolitinib treatment may be continued as long as the clinical benefits outweigh the risks.[27]

What is the recommended dose for ruxolitinib?

- The recommended dose depends on how well patients are responding to treatment. The dose should be reduced or the treatment should be stopped if certain side effects occur.[27]
- The recommended starting dose of ruxolitinib is 10 mg orally twice daily, with a maximum dose

of 25 mg twice daily.[27]

How does this differ to existing treatments?

The BSH guidelines recommend cytoreductive therapies. Some of these therapies are given by injection, which may be considered less convenient for patients than an oral formulation.

- IFN-alfa is a self-administered injection.
- High-dose bulsulfan is given via intravenous infusion (a drip). Low dose bulsulfan is given as a tablet taken orally.[28]

3d) Current clinical trials

Please provide a list of completed or ongoing clinical trials for the treatment. Please provide a brief top-level summary for each trial, such as title/name, location, population, patient group size, comparators, key inclusion and exclusion criteria and completion dates etc. Please provide references to further information about the trials or publications from the trials.

RESPONSE and RESPONSE-2

The main effectiveness and safety evidence for the use of ruxolitinib compared with BAT in patients with PV who are R/I to HC/HU comes from two randomised controlled clinical trials (RCTs): the RESPONSE trial for patients with splenomegaly (enlarged spleen), and RESPONSE-2 for patients without splenomegaly.[21, 22] A summary of the design of these trials is provided in Table 1.

Table 1: Summary of RESPONSE and RESPONSE-2 trial methodology

Trial name	RESPONSE (NCT01243944)	RESPONSE-2 (NCT02038036)
Location	International, multicentre trial with 92 sites across 18 countries: Argentina, Australia, Belgium, Canada, China, France, Germany, Hungary, Italy, Japan, Korea, Netherlands, Russia, Spain, Thailand, Turkey, UK and US.	International, multicentre trial with 48 sites across 12 countries: Australia, Belgium, Canada, France, Germany, Hungary, India, Israel, Italy, Korea, Spain and Turkey.
Trial design	Randomised, open-label Phase 3 study	Randomised, open-label Phase 3b study
Eligibility criteria for participants	**Key eligibility criteria:** • Adults (18 years of age or older) with PV requiring phlebotomy for HCT control • R/I to HC/HU • Splenomegaly (enlarged spleen)	**Key eligibility criteria:** • Adults (18 years of age or older) with PV requiring phlebotomy for HCT control • R/I to HC/HU • No splenomegaly (enlarged spleen)
Trial drugs and method of administration	**Treatment groups:** 1. Ruxolitinib (n=110) 2. BAT (n=112) • BAT could include HC/HU, IFN-alfa, pipobroman, anagrelide, immunomodulators such as lenalidomide or thalidomide, or no medication. radioactive phosphorus, busulfan and chlorambucil were prohibited • BAT could be changed due to lack of response or side effects requiring drug discontinuation	**Treatment groups:** 1. Ruxolitinib (n=74) 2. BAT (n=75) • BAT could include HC/HU (at maximum tolerated dose), IFN-alfa, pipobroman, anagrelide, approved immunomodulators such as lenalidomide and thalidomide, or no cytoreductive treatment • BAT could be changed due to lack of response or side effects requiring drug discontinuation

	• Patients who were randomised to BAT could crossover to receive ruxolitinib if they failed to meet the primary endpoint **Starting dose of ruxolitinib:** 10 mg twice daily **Dose adjustments for ruxolitinib:** Doses could be reduced for safety and efficacy (by 5 mg at a time; minimum of 5 mg once daily and maximum of 25 mg twice daily)[c]	• Patients who were randomised to BAT could crossover to receive ruxolitinib if they failed to meet the primary endpoint **Starting dose of ruxolitinib:** 10 mg twice daily **Dose adjustments for ruxolitinib:** Doses could be reduced for safety and efficacy (by 5 mg at a time; minimum of 5 mg once daily and maximum of 25 mg twice daily)[d]
Duration of study and follow-up	The study was started on 27th October 2010 and completed on 9th February 2018.	The study was started on 25th March 2014 and completed on 7th April 2020.

Abbreviations: AML: acute myeloid leukaemia; BAT: Best Available Therapy: HCT: haematocrit; HC/HU: hydroxycarbamide/hydroxyurea; IFN: interferon; PV: polycythaemia vera; R/I: resistance to or intolerance to.
Source: Vannucchi et al. 2015;[22, 29] ClinicalTrials.gov (NCT01243944);[30] Novartis Data on File (RESPONSE Week 208 CSR) 2017;[31] Passamonti et al. 2017;[21, 32] Passamonti et al. 2018;[33] Passamonti et al. 2022;[24] ClinicalTrials.gov (NCT02038036).[34]

MAJIC-PV trial

MAJIC (ISRCTN61925716) is a UK-based, completed Phase 2 RCT of ruxolitinib versus BAT in 'high-risk' patients with PV who are R/I to HC/HU.[35] In the MAJIC-PV trial, the definition of 'high-risk' was broad and may be considered to represent the majority of patients with PV who are R/I to HC/HU.[35] The MAJIC-PV trial data provides additional evidence for the efficacy and safety for ruxolitinib to support data from the RESPONSE and RESPONSE-2 trials.[35] The study was started on August 2016 and completed in April 2022.[35]

3e) Efficacy

Efficacy is the measure of how well a treatment works in treating a specific condition.

In this section, please summarise all data that demonstrate how effective the treatment is compared with current treatments at treating the condition outlined in section 2a. Are any of the outcomes more important to patients than others and why? Are there any limitations to the data which may affect how to interpret the results? Please do not include academic or commercial in confidence information but where necessary reference the section of the company submission where this can be found.

The symptoms and complications of PV are caused by a high concentration of red blood cells, white blood cells and platelets in your blood.

- HCT is a measure of the concentration of red blood cells in the blood. Achieving HCT control is linked to improved symptoms and reduced risk of complications.
- Complete haematological remission (CHR) is a measure of the concentration of red blood cells, white blood cells and platelets in the blood. Achieving CHR is linked to improved symptoms and reduced risk of complications.

Ruxolitinib offers improved and long-term responses, in terms of HCT control and platelet and white blood cell counts compared to BAT

- In RESPONSE, all patients in the trial had splenomegaly (enlarged spleen). At Week 32 after starting treatment, 23% of patients who received ruxolitinib achieved HCT control and a meaningful reduction in the volume of their spleen, compared with 1% of patients who received BAT.[22]
- 24% of patients who received ruxolitinib achieved CHR compared with 9% of patients who received BAT.[22]
- In RESPONSE-2, no patients had splenomegaly. At Week 28 after starting treatment, 62% of

patients who received ruxolitinib achieved HCT control compared to 19% of patients who received BAT.[21]
- 23% of patients who received ruxolitinib achieved CHR compared to 5% of patients who received BAT.[21]

Ruxolitinib treatment leads to a reduction in the frequency of phlebotomy required, which reduces the burden this places on patients' lives

- In RESPONSE, 20% of patients who received ruxolitinib underwent at least one phlebotomy over 25 weeks compared with 62% of patients who received BAT.[22]
- In RESPONSE-2, 19% of patients who received ruxolitinib underwent at least one phlebotomy over 25 weeks compared with 60% of patients who received BAT.[21]

These results from RESPONSE and RESPONSE-2 are supported by 5-year long-term data for ruxolitinib, with data available for up to 256 weeks in RESPONSE and 260 weeks in RESPONSE-2.[23, 24]

Data from the MAJIC trial support the conclusions from the RESPONSE trials, with ruxolitinib shown to be effective in high-risk patients with PV who are R/I to HC/HU.

- 49% of patients who received ruxolitinib achieved CHR within one year compared with 27% of patients who received BAT.
- Responses to treatment were also achieved more quickly and were more durable with ruxolitinib compared with BAT. These responses were also associated with improved symptoms.[36]

Evidence from the clinical trials supports the idea that treatment with ruxolitinib might lead to a longer life expectancy

Evidence of the impact of ruxolitinib on survival compared with BAT is available up to 5 years from the MAJIC-PV trial. Ruxolitinib is well tolerated and patients remain on ruxolitinib for an extended period of time. The improvements seen in controlling HCT, reducing spleen size, controlling haematological (blood-related) outcomes, achieving CHR, and reducing complications are likely to lead to an improvement in survival.

3f) Quality of life impact of the medicine and patient preference information

What is the clinical evidence for a potential impact of this medicine on the quality of life of patients and their families/caregivers? What quality of life instrument was used? If the EuroQol-5D (EQ-5D) was used does it sufficiently capture quality of life for this condition? Are there other disease specific quality of life measures that should also be considered as supplementary information?

Please outline in plain language any quality of life related data such as **patient reported outcomes (PROs).**

Please include any **patient preference information (PPI)** relating to the drug profile, for instance research to understand willingness to accept the risk of side effects given the added benefit of treatment. Please include all references as required.

In the RESPONSE, RESPONSE-2 and MAJIC-PV trials, quality of life was measured using a number of different tools. These types of tools are often referred to as 'patient-reported outcomes':[21, 22]

- **Myeloproliferative Neoplasm Symptom Assessment Form (MPN-SAF patient diary):** The MPN-SAF patient diary was used to assess 14 PV-related symptoms in three groups:
  - The cytokine symptom cluster (tiredness, itching, muscle ache, night sweats, and sweating while awake).
  - The hyper-viscosity symptom cluster (vision problems, dizziness, concentration problems, headache, numbness or tingling in the hands or feet, ringing in the ears and skin redness).

- The splenomegaly symptom cluster (abdominal discomfort and feeling full from food early).[22, 32]

- These symptoms are measured on a scale of 0–10, with higher scores relating to greater severity of symptoms.[22, 32]
- **Pruritus Symptom Impact Scale (PSIS):** The 5-question PSIS survey was used to evaluate pruritus severity (itchiness) and its impact on daily life on a scale from 0 (not at all) to 10 (worst imaginable).[37]
- **EORTC QLQ-C30:** The EORTC-QLQ C30, a 30-item survey comprising six functional scales (physical functioning, cognitive functioning, emotional functioning, social functioning, role functioning, and global quality of life) as well as nine symptom scales (fatigue, pain, nausea/vomiting, dyspnoea, insomnia, appetite loss, constipation, diarrhoea, and financial difficulties), is used to measure cancer-specific health-related quality of life. A high score for a functional scale of the EORTC indicates a high level of functioning, whereas a high score for a symptom scale or item represents a high level of symptomology or problems.
- **Patient Global Impression of Change (PGIC):** The PGIC is composed of a single question intended to measure a patient's perspective of improvement or deterioration over time relative to treatment. The PGIC uses a seven-point scale where one equals 'very much improved' and seven equals 'very much worse'.[32, 37]
- **EuroQol 5 Dimension 5 Level (EQ-5D-5L)** is a standardised instrument consisting of five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and a visual analogue scale (VAS). The five dimensions are graded in five levels from no problems to extreme problems. The VAS ranges from 'best imaginable health state' to 'worst imaginable health state'. The scores are summarised into a single index score.[32]

Across quality of life measures, treatment with ruxolitinib led to improvements in symptom burden and quality of life compared to treatment with BAT in patients who are R/I to HC/HU.

- Patient reported outcomes revealed a substantial benefit with ruxolitinib versus BAT. For example:
  - In the RESPONSE trial, 49% of the patients receiving ruxolitinib had a greater than 50% reduction in the MPN-SAF total score, which acts as a measure of symptom severity. This is compared with only 5% of the patients receiving BAT.
  - In the MAJIC-PV trial, itching, fatigue, night sweats, feeling full early, weight loss, bone pain, inactivity and concentration during the first 12 months were all measured by the MPN-SAF to be significantly lower in patients receiving ruxolitinib compared with those receiving BAT.[36]
  - Ruxolitinib patients also experienced improvements in the EORTC QLQ-C30 scores whereas patients who received BAT experienced worsening in all areas except for emotional functioning.[21, 22]

3g) Safety of the medicine and side effects

When NICE appraises a treatment, it will pay close attention to the balance of the benefits of the treatment in relation to its potential risks and any side effects. Therefore, please outline the main side effects (as opposed to a complete list) of this treatment and include details of a benefit/risk assessment where possible. This will support patient reviewers to consider the potential overall benefits and side effects that the medicine can offer.

Based on available data, please outline the most common side effects, how frequently they happen compared with standard treatment, how they could potentially be managed and how many people had treatment adjustments or stopped treatment. Where it will add value or context for patient readers, please include references to the Summary of Product Characteristics from regulatory agencies etc.

Ruxolitinib was generally well tolerated by patients in RESPONSE and RESPONSE-2 and MAJIC-PV and this evidence is supported by long-term data.[23, 24, 36]

Ruxolitinib had a consistent safety profile across both RESPONSE and RESPONSE-2 trials and side effects were generally manageable with standard clinical monitoring and care.[38]

- Anaemia was the most common blood-related side-effect experienced by patients treated with ruxolitinib. Anaemia occurs when you lack enough healthy red blood cells to carry enough oxygen around your body. However, it was rarely classed as severe, undesirable or life threatening (Grade 3 or 4) in the trials.[21, 22]
- Few non-blood-related side-effects were seen in either treatment group.[21, 22]

At the 5-year follow-up in both trials, the rate of blood clots when adjusted for the different durations of exposure to treatment was lower for patients receiving ruxolitinib compared with those receiving BAT.[23, 24] Similarly, ruxolitinib reduced the likelihood of patients with PV subsequently developing other blood cancers such as acute myeloid leukaemia and myelofibrosis.

No new safety signals were identified in MAJIC-PV compared with RESPONSE and RESPONSE-2.[36]

3h) Summary of key benefits of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key benefits of the treatment for patients, caregivers and their communities when compared with current treatments.
- Please include benefits related to the mode of action, effectiveness, safety and mode of administration

Patients with PV often experience symptoms that substantially impact on their day-to-day lives. They are also at high risk of their disease progressing to other forms of blood cancer or developing blood clots, as well as having a shorter life expectancy than the general population. There are currently very limited effective treatment options for these patients and a substantial unmet need for a targeted, effective treatment to improve the burden that symptoms have on patients' lives.

Ruxolitinib selectively targets a specific part of the biology of PV, JAK1 and 2. Ruxolitinib blocks JAK1 and 2 from working, which stops the signals that cause blood cells to multiply. This means ruxolitinib can provide patients with a targeted treatment that directly impacts part of the processes involved in PV. This is in contrast to the currently available cytoreductive treatments, such as HC/HU and IFN-alfa, that are non-specific to PV.

As well as this, ruxolitinib has been shown to lead to improved and durable improvements in symptoms and quality-of-life for patients compared with BAT. White blood cell and platelet counts were reduced and ruxolitinib provided improved HCT control, which is linked to improved symptoms and reduced risk of complications. These improvements lead to patients reducing their reliance on phlebotomy and the practical burdens caused by needing to attend such procedures. In patient-reported outcomes, treatment with ruxolitinib led to improvements in symptom burden for patients and improved quality of life compared with BAT.

Overall, ruxolitinib offers a convenient, targeted and effective treatment that improves the symptom burden and quality of life of patients who currently have limited treatment options available to them.

3i) Summary of key disadvantages of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key disadvantages of the treatment for patients, caregivers and their communities when compared with current treatments. Which disadvantages are most important to patients and carers?
- Please include disadvantages related to the mode of action, effectiveness, side effects and mode of administration
- What is the impact of any disadvantages highlighted compared with current treatments

All medicines have the potential to cause side effects. There are still some side effects that might be experienced by patients who take this new medicine.

The side effects that patients taking this new medicine may experience are described above in Section 3g and are considered manageable by clinicians.

3i) Value and economic considerations

Introduction for patients:

Health services want to get the most value from their budget and therefore need to decide whether a new treatment provides good value compared with other treatments. To do this they consider the costs of treating patients and how patients' health will improve, from feeling better and/or living longer, compared with the treatments already in use. The drug manufacturer provides this information, often presented using a health economic model.

In completing your input to the NICE appraisal process for the medicine, you may wish to reflect on:

- The extent to which you agree/disagree with the value arguments presented below (e.g., whether you feel these are the relevant health outcomes, addressing the unmet needs and issues faced by patients; were any improvements that would be important to you missed out, not tested or not proven?)
- If you feel the benefits or side effects of the medicine, including how and when it is given or taken, would have positive or negative financial implications for patients or their families (e.g., travel costs, time-off work)?
- How the condition, taking the new treatment compared with current treatments affects your quality of life.

How the economic model reflects the condition

- The economic model for this submission compared the costs and benefits for patients receiving ruxolitinib against those receiving current treatment, over the entire lifespan of a patient. Current treatment comprises a mix of treatments, also called BAT.
- Three analyses were conducted to reflect the populations recruited in the main trials supporting this submission:
  - (1) patients with PV with an enlarged spleen
  - (2) patients with PV without an enlarged spleen
  - (3) patients with high-risk PV.
- A mathematical model was developed for each population and included three stages (or health states) in the patient's pathway:
  - (1) receipt of ruxolitinib where quality of life is generally improved and there is a reduction in complications associated with PV

    - (2) receipt of BAT (either as a starting treatment or following ruxolitinib discontinuation) where quality of life is not controlled, and patients are at an increased risk of complications
    - (3) patients who have died.
- Within each of these health states, the economic model captures the costs for each patient (including cost of treatment and healthcare resource use), and quality adjusted life years (QALYs). QALYs are a health outcome measure that considers both the length and the quality of life provided by a treatment e.g., one year spent in perfect health represents one QALY. The QALYs in the model captured the benefits the treatments provide as well as the impact of side effects, complications, and health-related quality of life.

Modelling how much a treatment extends life

- Ruxolitinib is more effective compared with BAT in controlling HCT, reducing spleen size, controlling haematological (blood-related) outcomes, achieving CHR, and reducing the risk of complications. These benefits are likely to lead to an improvement in survival.
- The effect of ruxolitinib on survival over BAT is taken from the MAJIC-PV trial, that reported survival up to 5 years for patients receiving ruxolitinib or BAT. The survival benefit is carried forward beyond 5 years and assumed to reduce gradually until no survival benefit is assumed after 20 years.

Modelling how much a treatment improves quality of life

- Ruxolitinib leads to an improvement in the quality of life for patients compared with those receiving BAT. This is due to an improvement in PV related symptoms such as itchiness, a decreased need for phlebotomy and a decrease in treatment related side effects.
- Quality of life data utilised in the model comes from the MF-8D questionnaire in the RESPONSE-trials. The MF-8D is a condition preference-based measure that captures the impact of key symptoms in conditions like PV.
- The EQ-5D was used in a scenario analysis only as this is a generic instrument and there is evidence that it does not adequately capture the key symptoms of PV.

Modelling how the costs of treatment differ with the new treatment

- The model shows that the total costs associated with ruxolitinib are higher compared with BAT. This is due to higher drug costs.
- Resource use associated with phlebotomy, management of side effects and complications, and management of the condition were lower for ruxolitinib compared with BAT.

Uncertainty

- The key uncertainty relates to the extent and duration of the benefit of ruxolitinib on survival. Different assumptions are presented in the submission to reflect this uncertainty.
- Evidence of the impact of ruxolitinib on survival is available up to 5 years from the MAJIC-PV trial. Ruxolitinib is well tolerated, and patients remain on ruxolitinib for an extended period of time. The improvement in controlling HCT, reducing spleen size, controlling haematological (blood-related) outcomes, achieving CHR, and reducing complications are likely to support the assumption that the benefit in survival will be sustained beyond 5 years.

Cost effectiveness results

- Cost-effectiveness results for ruxolitinib versus current clinical management can be found in Section B.3.9 of the Company Submission.
- In summary, across all different analyses, ruxolitinib was associated with an improvement in survival (between 2–2.5 years) and improvement in quality of life (between 2–3 QALYs). ICERs were above currently accepted thresholds using the current net price offered to the NHS for ruxolitinib in an earlier indication. Novartis is working with NHS England to offer ruxolitinib at a significantly lower price for patient with PV to ensure ruxolitinib represents a

good use of NHS resources.

Additional factors

- Patients receiving ruxolitinib experience improvements in their symptoms and require fewer hospital visits associated with phlebotomy. In addition to helping alleviate capacity issues within the NHS, reducing hospital visits will also have a positive impact on patient quality of life as patients may experience increased anxiety and stress when attending hospital appointments, due in particular to the risk of contracting COVID-19 during their hospital visit.

3j) Innovation

NICE considers how innovative a new treatment is when making its recommendations.
If the company considers the new treatment to be innovative please explain how it represents a 'step change' in treatment and/ or effectiveness compared with current treatments. Are there any QALY benefits that have not been captured in the economic model that also need to be considered (see section 3f)

Ruxolitinib is a medicine designed to specifically target JAK 1 and 2. Ruxolitinib blocks JAK 1 and 2 from working, which stops the signals that cause blood cells to multiply. This helps to keep the number of red blood cells, white blood cells and platelets at a normal level. This reduces the symptoms of the disease.[1] No other targeted treatments are available in England for patients with PV.

Patients receiving ruxolitinib experience improvements in their symptoms and therefore require fewer hospital visits. Hospital appointments can often cause increased anxiety and stress for patients as well as the practical implications of needing to travel to and attend these appointments. Reducing the number of appointments can also have positive benefits for the NHS by freeing up additional appointment slots for other patients and reducing the burden on capacity.

3k) Equalities

Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged.
Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics

More information on how NICE deals with equalities issues can be found in the NICE equality scheme
Find more general information about the Equality Act and equalities issues here

No equality issues are expected in this appraisal.

SECTION 4: Further information, glossary and references

4a) Further information

Feedback suggests that patients would appreciate links to other information sources and tools that can help them easily locate relevant background information and facilitate their effective contribution to the NICE assessment process. Therefore, please provide links to any relevant online information that would be useful, for example, published clinical trial data, factual web content, educational materials etc.
Where possible, please provide open access materials or provide copies that patients can access.

The following websites provide useful information relating to polycythaemia vera and ruxolitinib:

- Cancer Research UK. Polycythaemia vera (PV): https://www.cancerresearchuk.org/about-

cancer/other-conditions/polycythaemia-vera

- NHS. Polycythaemia: https://www.nhs.uk/conditions/polycythaemia/
- Macmillan Cancer Support. Polycythaemia vera (PV): https://www.macmillan.org.uk/cancer-information-and-support/blood-cancer/polycythaemia-vera-pv
- Macmillan Cancer Support. Ruxolitinib (Jakavi®): https://www.macmillan.org.uk/cancer-information-and-support/treatments-and-drugs/ruxolitinib
- National Organization for Rare Diseases (NORD). Polycythaemia vera for patients and families: https://rarediseases.org/rare-diseases/polycythemia-vera/
- Myeloproliferative Neoplasms (MPN) Voice: https://www.mpnvoice.org.uk/about-mpns/questions/polycythaemia-vera/ Leukaemia Care. Polycythaemia vera (PV) - https://www.leukaemiacare.org.uk/support-and-information/information-about-blood-cancer/blood-cancer-information/about-myeloproliferative-neoplasms-mpn/polycythaemia-vera/

Further information on NICE and the role of patients can be found at the following links:

- Public Involvement at NICE Public involvement | NICE and the public | NICE Communities | About | NICE
- NICE's guides and templates for patient involvement in HTAs Guides to developing our guidance | Help us develop guidance | Support for voluntary and community sector (VCS) organisations | Public involvement | NICE and the public | NICE Communities | About | NICE
- EUPATI guidance on patient involvement in NICE: https://www.eupati.eu/guidance-patient-involvement/
- EFPIA – Working together with patient groups: https://www.efpia.eu/media/288492/working-together-with-patient-groups-23102017.pdf
- National Health Council Value Initiative. https://nationalhealthcouncil.org/issue/value/
- INAHTA: http://www.inahta.org/
- European Observatory on Health Systems and Policies. Health technology assessment - an introduction to objectives, role of evidence, and structure in Europe: http://www.inahta.org/wp-content/themes/inahta/img/AboutHTA_Policy_brief_on_HTA_Introduction_to_Objectives_Role_of_Evidence_Structure_in_Europe.pdf

4b) Glossary of terms

- **Aspirin:** an anti-thrombotic medicine that reduces the formation of blood clots.
- **Best available therapy (BAT):** a combination of different medicines that represent the current standard of care for patients with PV. The medicines that may be included in BAT are cytoreductive therapies, including HC/HU, interferon-alpha (IFN-alfa), anagrelide and busulfan.
- **Complete haematological response (CHR):** measure of the concentration of red blood cells, white blood cells and platelets in the blood. Achieving CHR is linked to improved symptoms of PV and reduced risk of complications.
- **Cytoreductive therapy:** a medicine that works to control the blood cell count without allowing the blood cell count to go up and down.
- **EQ-5D-5L:** A self-assessed, health related, quality of life questionnaire. The scale measures quality of life on a five-component scale including mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.
- **Haematocrit (HCT) levels:** the amount of space red blood cells take up in the blood. Achieving HCT control is linked to improved symptoms of PV and reduced risk of complications.
- **High-risk PV:** patients with high-risk PV are those who are aged 65 years or older or have experienced a blood clot in the vein or artery that is related to PV.
- **Hydroxyurea/hydroxycarbamide:** a medicine used to treat PV. It works by slowing or

stopping the growth of cancer cells in the body.

- **Intolerant to HC/HU:** occurs when a patient experiences unacceptable side effects when taking HC/HU treatment.
- **Janus kinase-1/2:** genes in the body that when changed can cause the bone marrow cells to produce too many red blood cells, resulting in PV.
- **Myeloproliferative neoplasm symptoms assessment form (MPN-SAF):** a patient-completed diary that is used to assess symptoms related to PV. Symptoms are measured on a scale of 0–10, with higher scores relating to greater severity of symptoms.
- **Phlebotomy:** the removal of blood from the vein with a medical instrument. It is often used to reduce the 'thickness' of the blood.
- **Pruritus:** itchy skin.
- **Resistant to HC/HU:** occurs when HC/HU does not work anymore to treat PV.
- **Ruxolitinib:** this is the medicine that is under evaluation by NICE. It is a type of medicine called a kinase inhibitor. It blocks JAK1 and JAK2 from working, which stops the signals that cause blood cells to multiply.
- **Splenomegaly:** enlarged spleen.
- **Thromboembolic event:** a broad term for the occurrence of different types of blood clots.
- **Thrombosis:** a broad term for different types of blood clots.

4c) References

Please provide a list of all references in the Vancouver style, numbered and ordered strictly in accordance with their numbering in the text:

1. European Medicines Agency (EMA): Jakavi® European Public Assessment Report (EPAR) - Product Information. Available at: https://www.ema.europa.eu/en/documents/overview/jakavi-epar-summary-public_en.pdf [Last Accessed 10th June 2022].
2. National Health Service (NHS). Polycythaemia. Available at: https://www.nhs.uk/conditions/polycythaemia/ [Last Accessed 11 August 2022].
3. Chapman J, Goyal A, Azevedo AM. Splenomegaly. StatPearls. Treasure Island (FL): StatPearls Publishing., 2022.
4. Diehn F, Tefferi A. Pruritus in polycythaemia vera: prevalence, laboratory correlates and management. British journal of haematology 2001;115:619-621.
5. Siegel FP, Tauscher J, Petrides PE. Aquagenic pruritus in polycythemia vera: characteristics and influence on quality of life in 441 patients. American Journal of Hematology 2013;88:665-9.
6. Steinman HK, Greaves MW. Aquagenic pruritus. Journal of the American Academy of Dermatology 1985;13:91-96.
7. Hensley B, Geyer H, Mesa R. Polycythemia vera: current pharmacotherapy and future directions. Expert Opin Pharmacother 2013;14:609-17.
8. Tefferi A. Essential thrombocythemia, polycythemia vera, and myelofibrosis: current management and the prospect of targeted therapy. Am J Hematol 2008;83:491-7.
9. Finazzi G, Barbui T. Evidence and expertise in the management of polycythemia vera and essential thrombocythemia. Leukemia 2008;22:1494-1502.
10. Passamonti F, Rumi E, Pungolino E, et al. Life expectancy and prognostic factors for survival in patients with polycythemia vera and essential thrombocythemia. The American journal of medicine 2004;117:755-761.
11. Tefferi A, Rumi E, Finazzi G, et al. Survival and prognosis among 1545 patients with contemporary polycythemia vera: an international study. Leukemia 2013;27:1874-81.
12. Griesshammer M, Gisslinger H, Mesa R. Current and future treatment options for polycythemia vera. Annals of hematology 2015;94:901-910.
13. Moulard O, Mehta J, Olivares R, et al. Epidemiology of Myelofibrosis (MF), Polycythemia Vera (PV) and Essential Thrombocythemia (ET) in the European Union (EU). Blood 2012;120:1744.

14. McMullin MF, Harrison CN, Ali S, et al. A guideline for the diagnosis and management of polycythaemia vera. A British Society for Haematology Guideline. British journal of haematology 2019;184:176.
15. Alvarez-Larran A, Kerguelen A, Hernandez-Boluda JC, et al. Frequency and prognostic value of resistance/intolerance to hydroxycarbamide in 890 patients with polycythaemia vera. British Journal of Haematology 2016;172:786-93.
16. Jentsch-Ullrich K, Eberhardt J, Zeremski V, et al. Characteristics and treatment of polycythemia vera patients in clinical practice: a multicenter chart review on 1476 individuals in Germany. Journal of Cancer Research and Clinical Oncology 2016;142:2041-2049.
17. Marchetti M, Vannucchi AM, Griesshammer M, et al. Appropriate management of polycythaemia vera with cytoreductive drug therapy: European LeukemiaNet 2021 recommendations. The Lancet Haematology 2022;9:e301-e311.
18. Novartis. Data on File. Novartis UK Polycythemia Vera Advisory Board. 24th June 2022.
19. Alvarez-Larrán A, Pereira A, Cervantes F, et al. Assessment and prognostic value of the European LeukemiaNet criteria for clinicohematologic response, resistance, and intolerance to hydroxyurea in polycythemia vera. Blood 2012;119:1363-9.
20. Harrison C, Pemberton-Whiteley Z, Ali S, et al. The International Landmark Health Survey – MPNs. UK Results – Patient perceptions on the symptomatic and economic burden of Myeloproliferative Neoplasms. Presented at BSH 2017.
21. Passamonti F, Griesshammer M, Palandri F, et al. Ruxolitinib for the treatment of inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): a randomised, open-label, phase 3b study. The Lancet Oncology 2017;18:88-99.
22. Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera. New England Journal of Medicine 2015;372:426-435.
23. Kiladjian JJ, Zachee P, Hino M, et al. Long-term efficacy and safety of ruxolitinib versus best available therapy in polycythaemia vera (RESPONSE): 5-year follow up of a phase 3 study. Lancet Haematol 2020;7:e226-e237.
24. Passamonti F, Palandri F, Saydam G, et al. Ruxolitinib versus best available therapy in inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): 5-year follow up of a randomised, phase 3b study. The Lancet Haematology 2022;9:E480-E492.
25. Landolfi R, Marchioli R, Kutti J, et al. Efficacy and Safety of Low-Dose Aspirin in Polycythemia Vera. New England Journal of Medicine 2004;350:114-124.
26. Najean Y, Rain J-D. Treatment of polycythemia vera: the use of hydroxyurea and pipobroman in 292 patients under the age of 65 years. Blood, The Journal of the American Society of Hematology 1997;90:3370-3377.
27. Medicines and Healthcare products Regulatory Agency. Jakavi® 10 mg tablets: Summary of Product Characteristics. Available at: https://mhraproducts4853.blob.core.windows.net/docs/d61f7a73d3447fbfcf6d3936bed37a80946f480f [Last Accessed 08 September 2022].
28. Macmillan Cancer Support. Busulfan. Available at: https://www.macmillan.org.uk/cancer-information-and-support/treatments-and-drugs/busulfan [Last Accessed 18 August 2022].
29. Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera (supplementary appendix). New England Journal of Medicine 2015;372:426-435.
30. ClinicalTrials.gov: Study of Efficacy and Safety in Polycythemia Vera Subjects Who Are Resistant to or Intolerant of Hydroxyurea: JAK Inhibitor INC424 (INCB018424) Tablets Versus Best Available Care (RESPONSE; NCT01243944). Available at: https://clinicaltrials.gov/ct2/show/NCT01243944 [Last Accessed 09 September 2022].
31. Novartis. Data on File. Interim Clinical Study Report Week 208: Randomized, open label, multicenter phase III study of efficacy and safety in polycythemia vera subjects who are resistant to or intolerant of hydroxyurea: JAK inhibitor INC424 tablets versus best available care (The RESPONSE Trial).
32. Passamonti F, Griesshammer M, Palandri F, et al. Ruxolitinib for the treatment of inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): a randomised, open-label, phase 3b study (supplementary appendix). The Lancet Oncology 2017;18:88-99.

33. Passamonti F, Palandri F, Saydam G et al. American Society of Hematology (ASH) 2018; Abstract 1754. Ruxolitinib for the Treatment of Inadequately Controlled Polycythemia Vera without Splenomegaly: 156-Week Follow-up from the Phase 3 RESPONSE-2 Study.
34. ClinicalTrials.gov: Ruxolitinib Efficacy and Safety in Patients With HU Resistant or Intolerant Polycythemia Vera vs Best Available Therapy (RESPONSE-2; NCT02038036). Available at: https://clinicaltrials.gov/ct2/show/NCT02038036 [Last Accessed 09 September 2022].
35. Harrison C, Nangalia J, Boucher R, et al. Ruxolitinib versus Best Available Therapy for PV intolerant or resistant to Hydroxycarbamide in a Randomized trial (Manuscript in Development).
36. Harrison C, McMullin MF, Panchal A, et al. PF625: Ruxolitinib compared with best available therapy for polycythaemia vera patients resistant or intolerant to hydroxycarbamide in MAJIC - an investigator-led randomised trial. 23rd European Hematology Association Congress. Stockholm, Sweden, 2018.
37. Mesa R, Verstovsek S, Kiladjian JJ, et al. Changes in quality of life and disease-related symptoms in patients with polycythemia vera receiving ruxolitinib or standard therapy. Eur J Haematol 2016;97:192-200.
38. Verstovsek S, Vannucchi AM, Griesshammer M, et al. Ruxolitinib versus best available therapy in patients with Polycythemia Vera: 80 Week follow up from the RESPONSE trial. Haematologica 2016:haematol. 2016.143644.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

Clarification questions

October 2022

File name	Version	Contains confidential information	Date
ID5106 ruxolitinib clarification questions to PM for company_25May2023_[FULLY REDACTED]	Final	Yes	25th May 2023

Notes for company

Highlighting in the template

Square brackets and [redacted] highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [redacted] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

Clinical effectiveness and safety evidence

A1. The criteria defining resistance/intolerance to hydroxycarbamide (HC) reported in the RESPONSE trial protocol (citing Barosi et al. 2009) are different to those reported in in company submission (CS) Table 4 (citing Marchetti et al. 2022) which are in turn different to those stated in footnote b of CS Table 6. Please explain these discrepancies. Please explain what "modified ELN criteria" refers to.

Please accept our apologies for the confusion. The original ELN criteria (Barosi et al. 2010) states that the dose of HU needed to be 2 g/day for at least 3 months, with patients still requiring phlebotomy.[1] However, 2 g/day is not a tolerable dose for the majority of patients. Therefore, after clinical expert input at the time of writing the RESPONSE protocols, the criteria for 2 g /day was changed to maximum tolerated dose of HU in the resistance criterion of the phlebotomy requirement. This modification is more reflective of real-world practice.

In January 2021, the ELN promoted an international project to update the clinical indications for the use of cytoreductive drugs in the treatment of polycythaemia vera. An expert panel of 14 haematologists from ELN centres that had actively participated in previous ELN developed a list of clinical questions, and a methodologist established three patient, intervention, comparator, outcome (PICO) questions and systematically reviewed the evidence. Recommendations were approved by six Delphi consensus rounds and two virtual meetings (on Jan 26, 2021, and June 24, 2021). This updated guidance is reflected in Marchetti et al. (2022).[2]

A2. PRIORITY QUESTION. The unpublished MAJIC-PV trial manuscript refers to supplementary material which contains relevant information for interpreting the company's submission including the economic analyses. Please provide the supplementary material.

The supplementary appendix of the MAJIC-PV publication has been included as part of this response.

A3. Please provide the statistical analysis plan for the RESPONSE trial.

The statistical analysis plan for the RESPONSE trial has been included as part of this response.

A4. Please provide the interim clinical study report for week 32 of the RESPONSE-2 trial.

Please note that the primary endpoint was assessed at Week 28 in the RESPONSE-2 trial. The interim clinical study report for Week 28 of the RESPONSE-2 trial has been included as part of this response.

A5. Please provide the clinical study report and (if different from that stated in the protocol) the statistical analysis plan for the MAJIC-PV trial.

As highlighted in our submission, MAJIC-PV is an investigator-led trial and was not conducted by Novartis. As such, we do not currently have access to, and so cannot provide, the clinical study report or statistical analysis plan for this trial.

A6. The EPAR provided in CS Appendix C is for graft versus host disease. Please provide the latest EPAR for polycythaemia vera.

The correct EPAR has been included as part of this response.

A7. Please explain the order of adverse events reported in CS Table 23. Please clarify whether any events were double-counted, since there appear to be some non-independent categories of events reported in the table (e.g. abdominal pain / abdominal pain upper; aphthous ulcer / mouth ulceration).

Please accept our apologies for the confusion. We confirm that there is no double counting. AEs in Table 23 were pooled from the list of AEs reported from the RESPONSE and RESPONSE-2 manuscripts.[3, 4] AEs with the exact same preferred term were pooled (added together), while those with different preferred terms were not. While the expert assessment group (EAG) has highlighted that some of these AEs are associated with the same organ class (e.g. abdominal pain, abdominal pain upper) these were recorded with slightly different wording across the RESPONSE trials and

therefore, the company has taken a consistent approach in the economic analysis in order to accurately reflect the clinical data and prevent any inaccurate pooling of data.

Indirect treatment comparison

A8. PRIORITY QUESTION. As noted in CS section B.2.9.2, the GEMFIN registry has several limitations. Please justify the choice of the GEMFIN registry for matching in the indirect treatment comparison. Was a search conducted for UK-relevant real-world studies or registries? If so, please provide details of the search and studies identified. If not, please justify why no search was conducted.

The Grupo Español de Enfermedades Mieloproliferativas Filadelfia Negativas (GEMFIN) registry was selected as it is one of the largest registries of PV patients, with 1,495 patients available as of October 2016 (when initial study conception and secondary data collection began). The GEMFIN registry is a comprehensive database with sufficient information on baseline characteristics, treatment patterns, and outcomes to conduct an analysis and long-term follow-up. Individual patient data required to conduct the matching could also be obtained.[5] It is unclear if alternative sources were considered at the time the analysis was originally conducted, and we do not believe that a systematic search for real-world registries was performed. Although a potential limitation, it should be noted that the GEMFIN registry was the only registry that published data on outcomes for patients with PV that are R/I to HC/hydroxyurea (HU) (Alvarez et al. 2012) and therefore likely to represent the most appropriate source of evidence at the time the analysis was conducted.[6] It should also be noted that the MAIC is a supportive analysis and presented for transparency and completeness.

A9. CS section B.2.9.2 provides some information on the company's attempt to conduct an indirect treatment comparison on the combined RESPONSE and RESPONSE-2 trials which resulted in a "poor fit" when estimating propensity scores. Please provide the full methods and results of this analysis including the measures of statistical fit that were employed.

An exploratory analysis was performed to assess the feasibility of combining RESPONSE and RESPONSE-2 into a single indirect comparison when the initial matching was conducted. The feasibility assessment involved finding GEMFIN controls who could independently be matched with RESPONSE and RESPONSE-2. Since only [redacted] patients from GEMFIN were well-matched with both trials, it was determined that an insufficient number of control patients would be available to support estimation of indirect comparisons. As this exploratory analysis was conducted during initial study conception (prior to the update in 2019), the results from this exploratory analysis could not be

located. Therefore, we are not able to provide additional details or measures of statistical fit as part of this response.

A10. Please explain the differences in the numbers of patients from GEMFIN used in propensity score matching from Alvarez-Larran et al. 2018 (N=191) and the present company submission (N=184).

The dataset used in Alvarez-Larran et al. (2018) included 7 patients without follow-up beyond the date of being identified as R/I to HC/HU.[7] Subsequent analyses excluded these patients, leaving 184 (versus 191) patients available for PSM. Additionally, the propensity scores matched GEMFIN sample decreased from 90 to 89 patients after this change.

A11. PRIORITY QUESTION. CS Appendix D.1.6 states that the "top eight" prognostic factors and/or treatment effect modifiers were ranked, but the rationale for this approach requires clarification.

(a) Please explain what these eight covariates were and why they were considered "the top" eight covariates.

The top eight covariates are tabulated below in Table 1. These variables were considered the top 8 in terms of prognostic strength and effect modification based on the input of two clinical experts involved in the study.

Table 1: Top 8 prognostic factors for OS

[redacted]

* not included in analysis because definitions differed between the GEMFIN registry and RESPONSE Trial
Abbreviations: GEMFIN: Grupo Español de Enfermedades Mieloproliferativas Filadelfia Negativas; HU: hydroxyurea; OS: overall survival; PV: polycythaemia vera; R/I: resistance/intolerance.

(b) Please provide baseline characteristics for these eight covariates and any other covariates for which data are available and not already reported in the CS and publications.

Please find the baseline characteristics below in Table 2.

Table 2: Baseline characteristics

Abbreviations: HU: hydroxyurea; JAK: Janus kinase; PV: polycythaemia vera; R/I: resistance/intolerance; SD: standard deviation.

(c) What criteria did the experts use for ranking the covariates?

Clinical experts were asked to rank order covariates with respect to their prognostic strength or effect modifying status for overall survival (OS), considering their clinical experience and background knowledge of the literature.

(d) Were the experts who ranked the covariates the same experts referred to in the Alvarez-Larran et al. 2018 poster?

We confirm they were the same experts.

(e) Please explain why, after ranking the covariates, only the four highest-ranked covariates were used for propensity score matching.

The top four ranked characteristics with sufficient data were included in the PSM analysis. For the OS PSM analysis this included age, history of thrombosis, cytopenia at the lowest HC/HU dose to reach response, and gender. Of the remaining top eight covariates:

- Two variables (duration of PV diagnosis and uncontrolled myeloproliferation) were excluded because the variables were defined differently between the GEMFIN registry and RESPONSE trial.
    - For example, duration since PV diagnosis was defined as time from diagnosis to HC/HU resistance/intolerance in GEMFIN versus time from diagnosis to randomisation in RESPONSE.
    - Since no patient from RESPONSE was classified as having uncontrolled myeloproliferation, this variable was excluded.
- Diabetes was ranked low by the clinical experts and displayed low imbalances between comparators, therefore it was also excluded.

(f) Why was PK2 mutation status included in the Alvarez-Larran et al. 2018 analysis but not in the company submission? Please conduct a scenario analysis including PK2 mutation as a covariate.

PK2 mutation status was used in the Alvarez-Larran et al. (2018) analysis of thrombosis but was not used to derive PSM comparisons of OS.[7] Clinical experts were asked to provide separate variable rankings for thrombosis and OS. Clinical experts did not rank PK2 mutation status highly as a prognostic factor or effect modifier for OS. Unfortunately, additional analyses including PK2 cannot be conducted as data from GEMFIN do not belong to Novartis. While no additional analysis could be provided as part of this response, it should be noted there was little difference in Janus kinase 2 (JAK2) mutational status between the GEMFIN registry (█%) and the ruxolitinib arm of RESPONSE (█%).

A12. PRIORITY QUESTION. TSD17 recommends conducting sensitivity analysis around variables included in the matching, and use of interaction/ polynomial terms. Please conduct such analyses or justify why these have not been conducted.

As the PSM sample included only a moderate sample of 174 patients, the dataset was considered insufficient to support further matching on lower ranked prognostic factors, flexible polynomials, and/or interaction terms at the time the analysis were conducted. This is in line with the NICE TSD17 that recognises that flexibility will depend on the size of the dataset.[8]

A13. Please provide the statistical code for the indirect treatment comparison, including descriptions of the data variables used.

The R code that was used to generate the primary results has been included as part of this response. The code was used to estimate unadjusted hazard ratios (HRs) and PSM HRs for ruxolitinib versus best available therapy (BAT).

Section B: Clarification on cost-effectiveness data

B1. CS sections B.3.2.2, B.3.3.5, and B.3.8.2. The CS reports that patients "cycle" or "switch" through BAT regimens in a number of places in the submission, which seems to imply that patients are able to move back and forth between the sub-health states within the BAT health state. However, this does not seem to be implemented within the model; that is, patients that discontinue a treatment

and leave a sub-health state are forbidden from returning to that sub-health state. Please clarify the meaning of the terms "cycling/switching" in the CS.

Please accept our apologies for the confusion. During the advisory board meetings clinical experts advised that many patients with PV who are R/I to HC/HU often receive more than one BAT over their lifetime, and that this change in BAT is often due to inadequate symptom control or unacceptable side effects. Clinical experts noted that patients may 'cycle' or 'switch' back and forth between a number of BATs, however, it was not possible to incorporate this into the economic model as these data were not reported in the MAJIC-PV manuscript.

We acknowledge that the current wording in our description of the BAT health state ("allowed to cycle through/switch BAT") could be misinterpreted as patients moving back and forth between BATs in the economic model. Although this is what occurs in clinical practice, this is not included in the economic model due to lack of data. The economic model only considers the time to first BAT discontinuation and the time to all BAT discontinuation, with the same BAT composition used in both patients in the 1st BAT and 2nd+ BAT sub-health state. This is because only the BAT composition across all line of therapies was reported in the MAJIC-PV manuscript.

Treatment effect

B2. PRIORTY QUESTION. The company's base case analysis uses a time-varying treatment effect, estimated from reconstructed MAJIC-PV OS KM data (CS section B.3.3.4). It is stated that scenario analysis was conducted using the constant treatment effect reported in the unpublished manuscript of the MAJIC-PV trial: 0.73 (95% CI: 0.36 to 1.50) (CS Figure 31). However, this scenario is not included in the CS (Figures 53 and 56, or in Appendix P). Please add this scenario to the economic model and report the results.

Thank you for highlighting this omission. This scenario is now included in the automated list of scenarios in the updated economic model sent as part of our response to the clarification questions. Results for this scenario are also presented below for convenience.

Table 3: Results for scenario analysis using the constant HR for OS from the MAJIC PV trial (using the PAS discount agreed for MF)

	RESPONSE-2	RESPONSE	MAJIC
Base case	[redacted]	[redacted]	[redacted]
HR OS - MAJIC PV-trial (constant)	[redacted]	[redacted]	[redacted]

Abbreviations: HR: hazard ratio; MF: myelofibrosis; OS: overall survival; PAS: patient access scheme; PV: polycythaemia vera..

Survival extrapolations

B3. The KM results for time to treatment discontinuation (TTD) and overall survival in the ruxolitinib arms of the RESPONSE and RESPONSE-2 trials are summarised in the 'KMs' sheet of the economic model. Please provide the numbers of events and numbers at risk for each timepoint used to calculate these results.

The KM, number of events and number of censored patients are now provided in the updated economic model sent as part of this response in the sheet named "KM_EAGrequest".

B4. PRIORITY QUESTION. The base case prediction of pre-discontinuation survival is only adjusted for general population mortality after the five-year period of trial observation (column AE in the 'Trace_Rux' sheet). This results in better predicted survival while patients remain on ruxolitinib than for the general population, which seems implausible (see Figure 1 below).

Please add an option to the model to adjust pre-discontinuation survival for general population mortality over the entire time horizon.

Figure 1 Pre-discontinuation survival extrapolation for ruxolitinib
Source: Produced by EAG from graph on 'pre_disS_com_RESPONSEtrials' tab of the model

Thank you for highlighting this discrepancy. The constraint was added to reflect data from the trial during the observed period. However, as highlighted by the EAG this led to a mismatch between the rate of death from the general population used in the model and that from the pre-discontinuation survival during the observed period of the trial. This discrepancy could be explained by the small sample size in the RESPONSE-trial and/or the use of an average starting age in the model, despite an underlying distribution of age in the trial.

As requested by the EAG, an option is now included in the updated economic model sent as part of this response to adjust the pre-discontinuation survival for the general population mortality over the entire time horizon. Results for this scenario are presented in Table 4 for convenience.

Table 4: Results for scenario analysis adjusting pre-discontinuation for general population mortality for the entire period (using the PAS discount agreed for MF)

	RESPONSE-2	RESPONSE	MAJIC
Base case	[redacted]	[redacted]	[redacted]
Adjustment for the entire period	[redacted]	[redacted]	[redacted]

Abbreviations: MF: myelofibrosis; PAS: patient access scheme.

B5. PRIORITY QUESTION. It is stated in the CS (section B.3.1.2, page 108) that time to discontinuation of ruxolitinib in the high-risk subgroup was predicted by applying a hazard ratio for TTD versus OS from the ruxolitinib arm of the MAJIC-PV trial ([redacted]) to the predicted ruxolitinib OS curve. However, this parameter is reported as [redacted] in the model (cells Efficacy_OS&Baseline_C!I59-K59), and the latter value is also reported in CS Table 37. Please clarify the correct value of this parameter and confirm how it is used in the model.

Thank you for highlighting this transcription error in the CS. Please accept our apologies for the confusion. We confirm that the HR for TTD versus OS for ruxolitinib estimated from the MAJIC-PV manuscript is [redacted], in line with that used in the economic model.[9]

B6. It is stated in CS section B.3.3.5 (page 118) that the time to discontinuation of all BAT was estimated from the reported numbers of deaths (17) and discontinuations (23) from the MAJIC-PV trial. Please explain where these numbers and the related denominators are located in the MAJIC-PV trial manuscript (Harrison et al.).

The number of deaths and total discontinuations were taken from the supplementary appendix of the MAJIC-PV manuscript (Table S6 and Table S4 respectively). The supplementary appendix has been included as part of the response to question A2.

Utilities

B7. PRIORITY QUESTION. Please report MF-8D and EQ-5D utility estimates (means, 95% confidence intervals and numbers of observations) at baseline and all available assessment times from the RESPONSE and RESPONSE-2 trials by

study arm. Please also report any EQ-5D results that are available from the MAJIC-PV study.

Please find below the MF-8D estimated in RESPONSE (Table 5) and mapped EuroQol-5 Dimension 3 Level (EQ-5D 3L) utility estimates from RESPONSE-2 (Table 6) at baseline and all available assessment times.

MAJIC-PV is an investigator-led trial and was not conducted by Novartis. As such, we do not have access to and cannot provide an analysis plan for this trial. Additionally, we could not identify any EQ-5D data in the submitted manuscript (yet to be published) or any published abstracts of the MAJIC-PV trial.

Table 5: MF-8D utilities in RESPONSE at baseline and follow-up time points by study arm

[redacted]

Footnotes: *Number of patients with both baseline and post-baseline data at the time point.; ** Number of patients with non-missing baseline data as of their Day 1 visit.
Abbreviations: BAT: best available therapy; CI: confidence interval; MF-8D: Myelofibrosis-8 Dimension; RUX: ruxolitinib; SD: standard deviation; SE: standard error.
Source: RESPONSE IPD analysis

Table 6: Mapped (EQ-5D-5L to 3L) utilities in RESPONSE-2 at baseline and follow-up time points by study arm

Footnotes: *Number of patients with both baseline and post-baseline data at the time point; ** Number of patients with non-missing baseline data as of their Day 1 visit.
Abbreviations: BAT: best available therapy; CI: confidence interval; RUX: ruxolitinib; SD: standard deviation; SE: standard error.
Source: RESPONSE-2 IPD analysis

B8. PRIORITY QUESTION. Please provide a full report of the methods of analysis and results of the 'exploratory psychometric analyses' on EQ-5D for RESPONSE-2, which are summarised in CS section B.3.4.1 (page 123).

Due to the nature of the exploratory analysis in addition to the short time constraint to conduct this analysis in time for the submission, only a short description of the methods and results are available in a PowerPoint presentation. All details related to this exploratory analysis are included as part of this response.

B9. Please provide more detail on the methods and results of the regression analyses of MF-8D and EQ-5D utility (CS section B.3.4.3, page 125). Please explain the rationale for the choice of functional form and covariates, and report confidence intervals for the coefficients and predicted values used in the model, and measures of fit for the regressions. Why did you not use a repeated measures approach to incorporate data from other timepoints?

The main goal of the regression analyses was to predict the mean utility (MF-8D and EQ-5D) by randomised study arm at Week 32/28, conditional on baseline utility. Therefore, a multivariable regression model was fitted using Week 32/28 mapped utility as the dependent variable, an intercept, and baseline mapped utility as a covariate. Model parameters were derived using ordinary least squares. Analyses were performed in R using the stats::lm function.

For the MF-8D analysis, a repeated measures approach would include observations at Week 4, 16, and 32. For the EQ-5D analysis, a repeated measures approach would include observations at Week 4, 8, 16, and 28. Since symptom-related improvement could take time to occur and a lifetime horizon is used in the economic model, treatment effects evaluated at Week 32/28 were deemed most appropriate for economic modelling.

As requested, please find below in Table 7 the confidence intervals for the coefficients and measure of fit for the regressions (AIC and BIC), and the predicted utility values used in the economic model.

Table 7: Baseline adjusted regression models for MF-8D and EQ-5D and predicted utility values used in the economic model

Footnotes:
**Estimates calculated using patients with MF-8D utilities at baseline and at Week 32; **Estimates calculated using patients with mapped EQ-5D-5L utilities at baseline and at Week 28
Abbreviations: AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; CI: confidence interval; MF-8D: Myelofibrosis-8 Dimension; SD: standard deviation; SE: standard error.

B10. How were the sources for disutility and life expectancy used in the calculation of 'QALY loss' for the key events in CS Table 27 identified and selected?

The sources utilised to calculate QALY losses associated with the key events were not identified systematically. The sources were identified through known sources and search engines. While we acknowledge that the NICE method guide recommends inputs to be identified in a systematic manner, these inputs have a limited impact on the results.[10]

B11. CS Table 26 states that abdominal discomfort has a disutility of -0.075 and a duration of 7 days; the model uses a disutility of -0.0375 and a duration of 9.8 days (AE!R135:S135). Please identify the correct set of values and revise the model if appropriate.

We believe that the EAG refers to the discrepancy for aphthous ulcer rather than abdominal discomfort. We confirm that the economic model uses a disutility of -0.075 for a duration of 7 days for abdominal discomfort (labelled correctly in CS Table 26). There is, however, a discrepancy between the values used in the economic model and that reported in CS Table 26 for aphthous ulcer. The correct value is that used in the economic model (-0.0375 for 9.8 days) and is taken from TA722 (utility value and duration for stomatitis).[11]

B12. PRIORITY QUESTION. It appears that the 'QALY losses' for the key events in CS Table 27 include utility lost during expected survival following the event, but not years of life lost due to the event. If so, please consider whether the model may overestimate survival and QALYs, given that key events occurring within the trial may result in some deaths after the five-year follow up, and that the incidence of the key events is likely to increase with age.

Patients with PV are at increased risk of complications such as thromboembolic events (TEs), bleeding/haemorrhage, transformation to leukaemia (acute myeloid and myelodysplastic syndrome) and transformation to myelofibrosis; all can have a significant effect on survival. As highlighted in our submission, modelling events as a surrogate of survival was not feasible and there are no data to robustly construct a model based on surrogacy. Therefore, in line with expert advice a simpler approach was used whereby OS was extrapolated beyond the trials. While simplistic and not without limitations, the extrapolation implicitly accounts for the increased incidence of deaths over time associated with these complications. Therefore, including years of life lost due to an event is likely to lead to double counting. As OS is modelled directly for an average cohort and extrapolated over time irrespective of the cause of death, it is unknown what the contribution of death from these complications and other reasons are. While the lack of long-

term data make it challenging to validate outcomes, model predictions and extrapolations were validated by clinical experts with experience in the management of patients with PV who are R/I to HC/HU, providing reassurance that predictions from the model reflect their clinical expectation.

While we recognise the limitation and simplicity of the chosen approach, data are unfortunately lacking for an alternative approach based on surrogacy that could have allowed a more accurate modelling of the impact of these events on survival.

Resource use and costs

B13. PRIORITY QUESTION. CS Table 33. The CS states that TE incurs a management cost of £1,302, but the model implements a cost of £1,865. This seems to be due to a difference in cost for Grade 1-2 TE events, which the CS reports as £182. However, in the model this Grade 1-2 cost is given as £297. In addition, the unit costs for the management of Grade 3-4 TE events (CS Table 34) do not match the costs provided in the model. Note also that the management cost for TE provided in CS Table 37 is reported as £1,731, matching neither Table 33 nor the model. Please identify the correct costs to be used in the model.

Thank you for highlighting this discrepancy and apologies for the confusion. We confirm that the cost associated with the management of a TE event was estimated at £1,865 (the cost is incorrectly labelled in CS Table 37 and Table 33). We confirm that the cost for Grade 1-2 AE was assumed to be £297 (cost for associated with an emergency room (ER) visit based on the NHS reference cost 2020/21). We also confirm that the costs in CS Table 34 for individual events are incorrectly labelled and those currently used in the economic model are the correct ones.

B14. PRIORITY QUESTION. CS Table 33 states that bleeding/haemorrhage incurs a management cost £1,929; the model implements a cost £2,023. This seems to be due to a difference in cost for the management of a minor bleed, which the CS reports as £182. However, in the model this cost is given as £297. Furthermore, the CS (page 137) states that the management cost of a major bleed is assumed to be £9,788. Note that CS Table 37 reports a cost of £2,023 for the management of bleeding/haemorrhage. Please identify the correct costs to be used in the model.

Thank you for highlighting this discrepancy. Please accept our apologies for the confusion. We confirm that the weighted cost associated with a bleeding event (including both minor [82%] and

major [18%] bleeding) was estimated at £2,023 and is correctly reported in CS Table 37. We confirm that the cost associated with the management of a minor bleed was incorrectly labelled at £182 when it should have been £297 (cost of an ER visit based on the NHS reference cost 2020/21). The cost associated with a major bleed is £9,788.

B15. PRIORITY QUESTION. CS sections B.3.5.2 and B.3.8.1. and Table 33 report a cost of £1,058 for the management of NMSC, while the corresponding cost in CS Table 37 is given as £973. Please identify the correct cost to be used in the model.

Thank you for highlighting this discrepancy and apologies for the confusion. We confirm that the cost associated with the management of a nonmelanoma skin cancer (NMSC) event is £1,058, calculated as the average cost between the costs estimated from the top (£889) and bottom-down (£1,226) approach reported in Vallejo-Torres et al (2013). The cost associated with the management of NMSC was incorrectly labelled as £973 in CS Table 37.

B16. PRIORITY QUESTION. CS section B.3.5.2 states that the management cost associated with patients receiving ruxolitinib after 13 months is [redacted]; however, CS Table 37 reports a cost of [redacted]. Please indicate which cost is correct and should be used in the model.

Thank you for highlighting this discrepancy and apologies for the confusion. We confirm that the management cost for patients on ruxolitinib after 13 months (currently used in the economic model) is £[redacted] and was mislabelled in CS Table 37.

B17. CS section B.3.5.1 states that although prednisolone is often prescribed for patients receiving IFN-alfa, the costs for prednisolone and aspirin are not considered in the economic model as the costs are minimal. Please provide the costs and frequency of these two medications and provide further justification on the omittance of these medications. If possible, please provide a scenario analysis including the costs to show the negligible effect on results.

While the proportion of patients receiving aspirin in the RESPONSE trial is unclear, Verstovsek et al (2016) report that low dose aspirin was administered to all patients unless contraindicated in the RESPONSE trial.[12] The proportion of patients that would receive prednisolone alongside IFN-alfa is less clear. An extreme scenario analysis is conducted to demonstrate the negligible impact on the incremental cost-effectiveness ratio (ICER; Table 8).

In this scenario it is assumed that all patients (100%) on ruxolitinib and BAT receive aspirin (assumption of one 75 mg tablet daily). It is further assumed that 100% of patients receiving IFN-alfa receive prednisolone alongside (assumption of 20 mg daily).

Table 8: Results for the scenario analysis including aspirin and prednisolone (using the PAS discount agreed for MF)

	RESPONSE-2	RESPONSE	MAJIC
Base case	[redacted]	[redacted]	[redacted]
Utility value using repeated measurement	[redacted]	[redacted]	[redacted]

Abbreviations: MF: myelofibrosis; PAS: patient access scheme.

B18. We found some small discrepancies between the unit costs for managing and monitoring PV cited in CS Table 33 and NHS Cost Collection 2020/21 data:

- **Consultant-led outpatient attendance for clinical haematology is £199.38**

We confirm that the unit cost for a consultant led outpatient visit (WF01A: non-admitted face to face attendance, follow-up) of £214.56 used in the economic model (and reproduced in Table 32) is correct (matches the cost reported in NHS reference cost 2020/21, sheet named "CL", cell F69).

- **Emergency medicine VB01Z, VB04Z, VB05Z, VB07Z, VB08Z weighted mean £354.58**

Please accept our apologies for the confusion. We confirm that the weighted unit cost for an emergency visit of £296.90 used in the economic model (and reproduced in CS Table 32) is correct. However, the Healthcare Resource Group (HRG) codes used were incorrectly labelled in CS Table 32. The weighted unit cost for an emergency visit was calculated using the following codes (VB01Z-09Z, VB11Z). Costs related to VB10Z were excluded from the calculation of the weighted cost as they are for emergency dental treatment.

- **Not clear why A06A1 (occupational therapist) is appropriate for depression management.**

The cost associated with an occupational therapist (A06A1) was used for depression management to be consistent with the cost used for depression management in TA728.[13]

Please check these costs and clarify the correct code or correct the costs in the model if appropriate.

Please see responses above.

B19. CS section B.3.2.4 (page 100) states that pipobroman and phosphorus-32 have been removed from the subsequent treatments included in the model.

However, we note that combination treatments phosphorus-32 + hydroxycarbomide, phosphorus-32 + Anagrelide + Interferon, and interferon + pipobroman are still included in the model (Drug_costs!K47:K67). Please confirm that these combination treatments should also be removed.

As highlighted by the EAG, the costs associated with pipobroman and phosphorus-32 have been removed as these treatments are no longer available in England. We believe this approach is conservative as the effectiveness associated with these treatments are implicitly considered but not the costs. We do not believe that removing the cost for the entire combination to be appropriate as this would remove the cost for other treatments that are used in UK clinical practice despite their effectiveness being implicitly considered.

For transparency, please find below an analysis removing the cost for the entire combination. The impact on the ICER is negligible.

Table 9: Results for the scenario analysis removing the cost for the entire combination when including pipobroman or P32 (using the PAS discount agreed for MF)

	RESPONSE-2	RESPONSE	MAJIC
Base case	[redacted]	[redacted]	[redacted]
Remove cost entire combination	[redacted]	[redacted]	[redacted]

Abbreviations: MF: myelofibrosis; PAS: patient access scheme.

Section C: Textual clarification and additional points

C1. In CS Appendix M.2.1 (baseline characteristics of participants in MAJIC-PV) for palpable spleen length by ultrasound please explain the variance units that are reported in brackets after the median (there are 3 parameters within brackets, not the minimum and maximum that would be expected corresponding to the range of the median)

Thank you for highlighting this discrepancy. This appears to be a reporting error in the MAJIC-PV manuscript. The values within the brackets of 73, 77 and 150 for the BAT arm, ruxolitinib arm or overall population appear to be number of patients from which measurements were available.

We believe the median and range should read as follow: 14 (9 – 30) for BAT, 14 (9 – 26) for ruxolitinib and 14 (9 – 30) for the overall population.

C2. In the June 2022 Advisory Board meeting minutes please explain what "int-2/high risk" means.

Thank you for highlighting this. "int2/high risk" is abbreviated for "intermediate-2/high-risk".

Patients with myelofibrosis are often categorised onto risk categories; low risk, intermediate-1 risk, intermediate-2 risk and high-risk.

Patients with intermediate-2 and high-risk are eligible to receive ruxolitinib in line with TA386.[14]

C3. The QALY losses associated with phlebotomy cited in CS section B.3.4.5 (page 127) and CS Table 28 are wrong. The model uses the correct value −0.0001013 = −0.037/365.25. Please provide the full citation for the reference cited for this disutility (Matza et al. 2013), which is not included in the CS reference list, was not provided with the submission and we cannot find it online.

Thank you for highlighting this error in CS Table 28 and confirming that the model uses the correct value. Please find enclosed the Matza reference.

C4. CS Appendix P. Please note that the scenario analyses results in Table 69 for RESPONSE and RESPONSE-2 are in the wrong order; that is, the results in the RESPONSE column are actually for the RESPONSE-2 population, and vice versa.

Thank you for highlighting this error. We confirm that the columns were incorrectly labelled.

References

1. Barosi G, Birgegard G, Finazzi G, et al. A unified definition of clinical resistance and intolerance to hydroxycarbamide in polycythaemia vera and primary myelofibrosis: results of a European LeukemiaNet (ELN) consensus process. Br J Haematol 2010;148:961-3.
2. Marchetti M, Vannucchi AM, Griesshammer M, et al. Appropriate management of polycythaemia vera with cytoreductive drug therapy: European LeukemiaNet 2021 recommendations. The Lancet Haematology 2022;9:e301-e311.
3. Kiladjian J-J, Zachee P, Hino M, et al. Long-term efficacy and safety of ruxolitinib versus best available therapy in polycythaemia vera (RESPONSE): 5-year follow up of a phase 3 study. The Lancet Haematology 2020;7:e226-e237.
4. Passamonti F, Palandri F, Saydam G, et al. Ruxolitinib versus best available therapy in inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): 5-year follow up of a randomised, phase 3b study. The Lancet Haematology 2022;9:E480-E492.
5. Novartis. Data on File: Propensity Score Matching Analysis: Overall Survival at Week 256. Ruxolitinib (RESPONSE) versus BAT (GEMFIN).
6. Alvarez-Larran A, Pereira A, Cervantes F, et al. Assessment and prognostic value of the European LeukemiaNet criteria for clinicohematologic response, resistance, and intolerance to hydroxyurea in polycythemia vera. Blood 2012;119:1363-9.
7. A Alvarez-Larran SV, M Perez-Encinas, JJ Kiladjian, F Ferrer-Marin, A Vannucchi, J Martinez; C Besses, C Cameron, A Situ, A Charter, JC Hernandez-Boluda. Comparison Of Ruxolitinib And Real-World Best Available Therapy in Terms Of Overall Survival And Thrombosis In Patients With Polycythemia Vera Who Are Resistant Or Intolerant To Hydroxyurea. European Hematology Association Congress. Stockholm, Sweden, 2018.
8. National Institute for Health and Care Excellence (NICE). NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data. May 2015. Available at: https://www.sheffield.ac.uk/nice-dsu/tsds/full-list (accessed 25 October 2022).
9. Harrison C, Nangalia J, Boucher R, et al. Ruxolitinib versus Best Available Therapy for PV intolerant or resistant to Hydroxycarbamide in a Randomized trial (Manuscript in Development).
10. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual (PMG36). January 2022. Available at: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation (accessed 25 October 2022).
11. National Institute for Health and Care Excellence (NICE). Pemigatinib for treating relapsed or refractory advanced cholangiocarcinoma with FGFR2 fusion or rearrangement (TA722). Available at: https://www.nice.org.uk/guidance/ta722 (accessed 25 October 2022).
12. Verstovsek S, Vannucchi AM, Griesshammer M, et al. Ruxolitinib versus best available therapy in patients with Polycythemia Vera: 80 Week follow up from the RESPONSE trial. Haematologica 2016:haematol. 2016.143644.
13. National Institute for Health and Care Excellence (NICE). Midostaurin for treating advanced systemic mastocytosis (TA728). Available at: https://www.nice.org.uk/guidance/ta728 (accessed 25 October 2022).
14. National Institute for Health and Care Excellence (NICE). Ruxolitinib for disease-related splenomegaly or symptoms in adults with myelofibrosis (TA386). Available at: https://www.nice.org.uk/guidance/ta386 (accessed 05 August 2022).

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

Patient Organisation Submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you

1.Your name	[redacted] **and** [redacted]
2. Name of organisation	MPN Voice and Leukaemia Care
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	MPN Voice is the patient support organisation for people with Myeloproliferative Neoplasms (MPNs) in the UK. MPN Voice's mission is to provide clear and accurate information and emotional support to everyone who has been diagnosed with a myeloproliferative neoplasm and their families/friends. MPN Voice has members across the UK and in many other countries throughout the world. MPN Voice offers a website (http://www.mpnvoice.org.uk), patients' forums around the UK during the year, and a peer support programme to allow people with MPNs to contact others in similar circumstances. MPN Voice also has a moderated online forum at HealthUnlocked which is a supportive and informative online forum where patients and carers can ask questions about anything related to MPNs, and get replies from people who really understand the challenges of living with an MPN. In addition, MPN Voice produces information leaflets and a newsletter for people with MPNs so that patients are better informed and have more confidence dealing with the management of their condition. MPN Voice also raises money to fund research towards a cure and advocacy for patients. MPN Voice's work is primarily funded by donations from the public, through a wide range of fundraising activities. MPN Voice also accepts financial support from pharmaceutical companies for specific activities (see below) Leukaemia Care is a national blood cancer charity, founded in 1969. We are dedicated to ensuring that anyone affected by blood cancer receives the right information, advice and support. Approximately 85-90% of our income comes from fundraising activities – such as legacies, community events, marathons etc.

	Leukaemia Care also receives funding from a wide range of pharmaceutical companies, but in total those funds are less than 15% of our annual income. Leukaemia Care has undertaken a voluntary commitment to adhere to specific policies that regulate our involvement with the pharmaceutical industry set out in our code of practice here: https://media.leukaemiacare.org.uk/wp-content/uploads/Leukaemia-CARE-Code-of-Practice-pdf.pdf.
4b. Has the organisation received any funding from the company bringing the treatment to NICE for evaluation or any of the comparator treatment companies in the last 12 months? [Relevant companies are listed in the appraisal stakeholder list.] **If so, please state the name of the company, amount, and purpose of funding.**	MPN Voice: Novartis NI: Nov 2021 - £5,916, support for Dublin patient forum Novartis UK: Feb 2022 - £9,000, support for HealthUnlocked administration Bristol-Meyers Squibb: Oct 2021 - £10,000, support for website and patient forums Leukaemia Care: Novartis - £1,887.95 (£292.95 ASH video and £1,595 honorarium) Takeda - £25,000 core funding
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	Both: No
5. How did you gather information about the experiences of patients and carers to include in your submission?	Data supporting this submission has been gathered from a range of sources: -Patients' forums which take place throughout the UK. These meetings welcome patients' carers and families too. These provide a place where people affected by MPNs can meet one another. They are an excellent

	platform for sharing information. In our last full year before the pandemic, over 1000 patients and their loved ones attended these meetings. -The MPD Voice website: people with MPNs can submit their own stories of diagnosis, treatment and life impact. See https://www.mpnvoice.org.uk/living-with-mpns -Patient questionnaires were compiled to gather information on the effects/changes that taking ruxolitinib has made to the lives of patients with myelofibrosis. A separate carer/family questionnaire was devised. This asked what changes they had noted in the patient since ruxolitinib therapy was started and the affect this has had on the carer/family. These were made available for download via the MPN Voice website and distributed via post, email and social media. We gathered information from our online resources (HealthUnlocked, Facebook), individual patient interviews, a Survey Monkey questionnaire (270 answers), discussions with key UK experts, our discussions with patients (more than 1000 patients and family members with a myeloproliferative neoplasm of which PV is one of the commonest attended face to face meetings with our team in 2019.) We also recently interviewed three PV patients who are currently being treated with Ruxolitinib, after discontinuing other treatments, to understand their experience of living with PV and their comparison of the impact of the different treatments. MPN Voice is a founding member of MPN Advocates Network (MPNAN), a global coalition of MPN Patient groups. In 2019 MPNAN began the largest survey of MPN patient needs to date, with over 1700 responses at the time of writing. 640 responses have been received from PV patients. Evidence has also been taken from two MPN Landmark studies, the original US-based one in 2016 and a subsequent international study. The 2016 study had 816 respondents, of which 380 were PV patients. The international study had 223 responses from PV patients (78 from the UK), and provides information on patient reported quality of life and productivity. (Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5569657/)

	This submission is also informed by a patient experience survey of xx adults diagnosed with Polycythaemia Vera, carried out by Leukaemia Care in 2016. This was part of a wider survey of over 2500 blood cancer patients. MPN Voice continually gathers information through our support services (helpline, support groups, conferences, communications with our membership) and one to one discussion with patients. Lastly, we have used data the MAJIC PV study, information from which will be made available to the NICE committee.

Living with the condition

6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	The clinical symptoms of PV are extremely well documented in the literature. Polycythaemia Vera, while not being as life-threatening as Myelofibrosis, has nearly the same impact on patients' lives in terms of day-to-day symptoms. It can be an extremely debilitating illness that affects not only patients but also their families and carers. The most common symptoms that patients experience are: bleeding and/or clotting; fatigue; shortness of breath; chest pain; redness of skin; blurred vision and headaches; severe skin itchiness; joint pain or gout; dizzy spells; night sweats; unexplained weight loss; fullness/bloating in the left upper abdomen due to enlarged spleen Additionally, the Landmark study documented the burden of PV using a standardised symptom scoring system (MPN_SAF). It reported that 73% of PV patients reported fatigue, 55% pruritis (severe itching), 45% night sweats. The mean Total Symptom Score for PV patients was 17.4, not much lower than that for Myelofibrosis (21.2). These symptoms have a significant impact on patients' quality of life, as some of the answers to our survey illustrate: *"I tire very easily after a few hours of being at, for an example a party. Then the next day I spend most of it in bed, and I'm not talking about staying out all night getting drunk! Just a meal out with friends can completely knock me out the next day"* *"Tiredness means long social events are a thing of the past. A keen swimmer and sailor, I now have to plan my activities and my capacity for involvement carefully. I can no longer undertake long solo car journeys as the driver or do household gardening or decorating for long periods. This effect is slowly worsening. The future does not look too bright"* *"I have needed to cut back on my previous activities as get too exhausted and it's difficult to plan outings/activities more than few days ahead as don't know if I will be feeling well enough to go. I have delayed or cancelled several meetings with friends and recently missed a wedding reception because I ws feeling too exhausted and low"*

	"I went from being an independent professional with loads of self-confidence to a person who now hardly leaves the house. Low energy levels affect my desire to maintain a social life" *"The disease has greatly affected my quality of life. I have had to cancel so many events that I'm invited to, as I'm too unwell to attend. I can be good one day but the next day I'm in so much pain I have to rest as it's the only way I can carry on. I cannot book anything in advance. I have booked holidays, flights, etc. and lost money because I've ended up in hospital"* In addition, because of this symptom burden, patients often need a lot of support from family and friends. Here we have included some quotes from family members and carers that illustrate this. *"Family members help with his general housework and chores. I help with the big household jobs/projects/major clear-outs. And he has a lady to do his ironing for him as he still likes to look as smart as he can. I have to attend all hospital appointments with him to stay on top of his condition as he can't always take all the information in for each Doctor/Consultant."* *"He doesn't drive anymore because he is frequently dizzy. He is unsteady on his feet, so he waits for me to get him to shower, in case he falls. His vision is worse, he doesn't shave hardly ever anymore. He gets confused with simple instructions."* *"[Care is] mainly provided by her husband. Help with housework, children and daily tasks. On the worst days, she needs help bathing and dressing."* *"We use a handyman for simple household repairs that he used to do; as described above, I have taken on all the usual household chores that he did before."* *"I need to deal with some of the daily activities (eg shopping) as spouse can be too fatigued to do them"* *"I cook all meals and I administer his injectable medication. If I'm not around then he will have ready meals so can survive independently but things are more of a struggle. He is unable to go shopping*

	without a mobility aid, such as a scooter (which we don't currently have), so we have some shopping delivered and the rest I do, including collecting medication. He is currently working from home as going into the office is difficult."

Current treatment of the condition in the NHS

7. What do patients or carers think of current treatments and care available on the NHS?	All of the patients we surveyed were being treated with one or more of the available treatments, so their quotes reflect the fact that, at best, these treatments reduce or mitigate the effects of the disease. However, about 25% of patients treated with first line therapies become intolerant to treatment owing to their side effects or the effectiveness of the therapy declines. A minority of patients have difficult symptoms which are not well controlled, the most common being severe itching or pruritus. There are other treatments for this but some patients (we would estimate 1-2% of the total) have very bad itching which can be significantly helped with ruxolitinib but not with other therapies. The most common first line treatments for PV are: • **Phlebotomy** – patients go to hospital to have an amount of blood removed. The procedure itself is relatively well tolerated but has common aftereffects of dizziness and tiredness; patients usually need carers or family members to accompany them. In parts of the country, the distances involved in these regular trips to hospital can be very burdensome and expensive. One patient told us: "*I was always taking time off work to attend hospital and the venesections were difficult and distressing because of my small veins"* • **Hydroxycarbamide** – a form of chemotherapy that has a number of unpleasant and potentially dangerous side effects, including increased risk of other cancers, notably skin cancer. Hydroxycarbamide's effectiveness decreases over time for many patients and it has shown no potential to modify the disease. Prof Harrison estimates that 25% of patients treated in this way need to discontinue treatment because it has become ineffective or because the side effects become intolerable. Importantly, 20% of PV patients are under the age of 40, meaning that the increased cancer risk of Hydroxycarbamide would be born for many decades. 1/3 of PV patients are women. Hydroxycarbamide is not recommended for women of child-bearing age because it reduces fertility. Because of the known risks of medium- and long-term Hydroxycarbamide treatment, many patients' illness is not treated for many years until more severe symptoms develop and

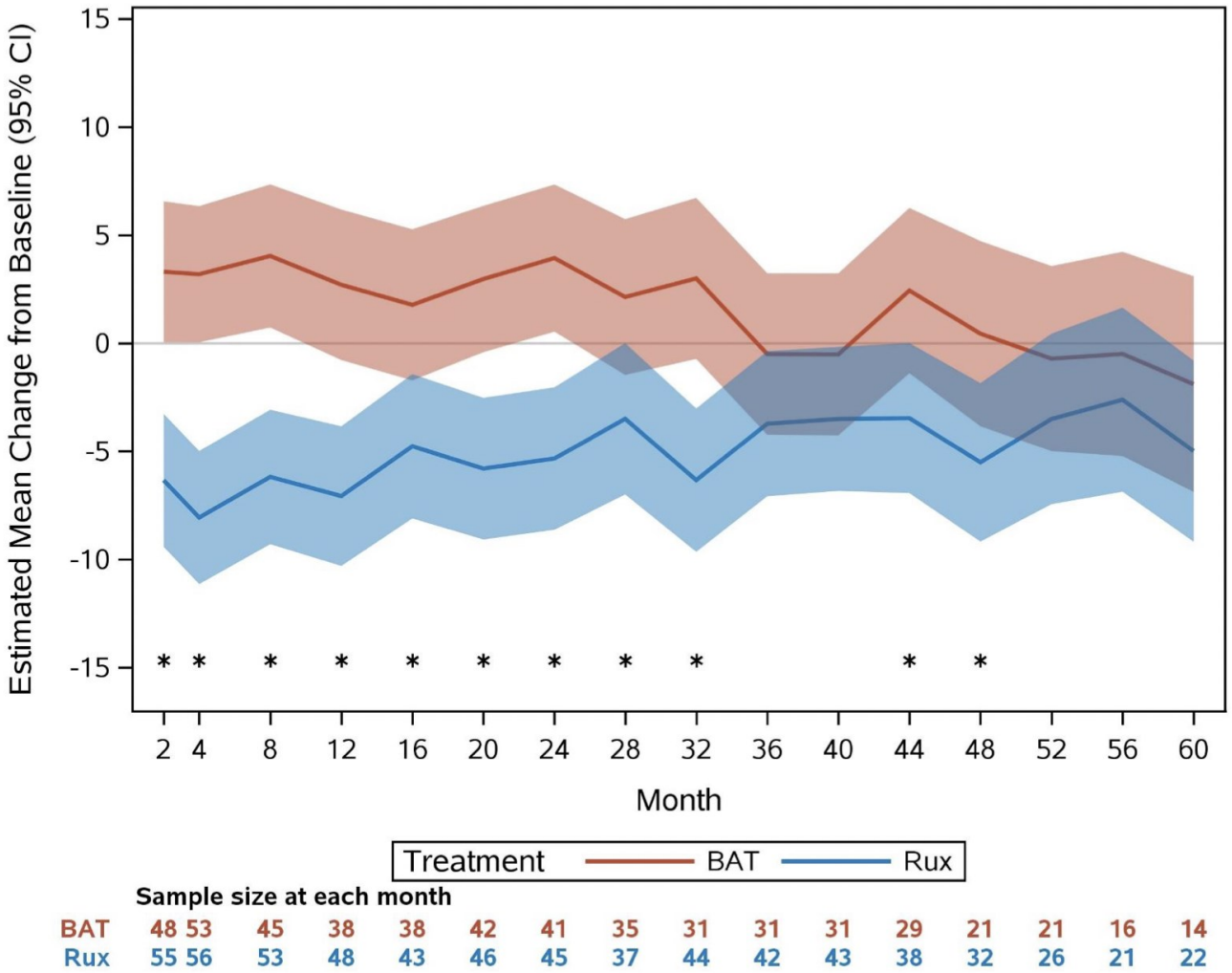
	the risk of disease progression has increased. • **Pegasys** (peginterferon alfa-2a) – this form of interferon is used to treat a small proportion of PV patients. This drug is used 'off-label' so access to it is not secure and NICE approved treatments would be preferable. Yet it is effective in controlling blood counts for most patients, but it has many side effects, including flu-like symptoms, fatigue, nausea and diarrhoea. The drug is not always effective – it gave no response to one of the patients we spoke to, and this patient went on to develop fibrosis of her bone marrow. Another one of the patients we spoke to discontinued Pegasys treatment after she experienced fatigue and problems with her vision. More seriously, some patients experience psychological problems such as low mood and depression. Overall, these problems make Pegasys impossible to use for up to 30% of patients. The production of Pegasys has now been discontinued by the manufacturer, Roche, and the remaining stock of the drug is dispensed as pre-filled syringes – this creates a significant issue of wastage for patients whose dose is only a fraction of the amount in the syringe. A license to re-start the manufacture of Pegasys has been acquired by another company, Pharma& but, at this time, there has been no commitment to its ongoing supply. Ropeginterferon alfa-2b (Besremi) has been shown to be effective in treating PV, but is not get available for NHS patients. There are other, less commonly used therapies including melphalan, busulphan and radioactive phosphorous but these drugs have significant side effects and increase the risk of developing more severe disease. Some quotes from patients illustrate the limitations of the currently available therapies: *"I initially had a year on Pegasys but that was discontinued due to problems with very low mood and flu-like side-effects. Then I had another year on HU but that was discontinued due to side-effects including recurring infections and I was unhappy with being on long-term chemotherapy"* *"I was always taking time off work to attend hospital and the venesections were difficult and distressing because of my small veins"*

8. Is there an unmet need for patients with this condition?	As the patient testimonies have shown, existing therapies for PV have significant limitations and/or problematic side effects. We have clearly seen that currently available treatments do not adequately reduce patients' fatigue, bone pain or pruritis. PV patients usually need frequent venesections to control blood counts, which is highly disruptive to their and their carers' lives, as well as being a significant cost to the NHS. Additionally, patients have told us that the side effects of hydroxycarbamide (mouth and leg ulcers, diarrhoea or constipation) increase the burden of PV. Interferon is often not tolerated by patients because of its side effects, such as flu-like symptoms, reduced white cell counts, nausea, headache, diarrhoea. Interferon can cause depression in some patients. So, even if the currently available therapies had long-lasting effectiveness in controlling blood counts and reducing the risk of clotting (which they do not), their limitations in terms of symptom alleviation and their side effects mean that there is a significant unmet need for a drug that is more effective in controlling symptoms and has fewer and more tolerable side effects. However, we also know that the currently available therapies commonly decline over time in terms of their effectiveness. In these situations, patients have very few options and face a bleak and uncertain future. There is a very important unmet need for a further therapeutic option for these patients too.

Advantages of the technology

9. What do patients or carers think are the advantages of the technology?	Patients who have been treated with ruxolitinib said: *"Prior to being prescribed ruxolitinib I was perpetually exhausted. As a single person, I had to work and the effort this took left nothing for me. Prior to developing PV, I was very active socially – out with family and friends, going to the gym 3 times a week, horse riding. All of this stopped with PV and my previous medication. Now that I am on Ruxolitinib, I have started socialising again and I'm thinking of re-joining the gym. Ruxolitinib has transformed my life and I am truly grateful for it. It is my hope that it will soon be offered to all suffers of diseases that may benefit from it."* *"After I start taking ruxolitinib: 1. I have clearly an improvement in the splenomegaly – the pain and discomfort have diminished considerably 2. The bone pain has reduced significantly; 3. Although I still been affected by fatigue, I can say that I have noted an improvement."* *"Fatigue significantly less, no night sweats. Still need to be careful not to work too many hours or be over stressed. And ensure good water intake. But now living a more or less normal life with ruxolitinib whereas previously I was often very tired and inactive. Hydroxycarbamide was very bad for me: awful mouth ulcers and terrible fatigue. Interferon led to serious weight loss and still didn't reduce the fatigue. Ruxolitinib has been good though I've gained weight and I've had shingles for which I now take additional meds."* *"Since taking ruxolitinib I feel I have got my life back – the difference between Rux and hydroxycarbamide is huge. On Hydroxy for a year, I was in bed most days, my pruritis was unbearable and I was unable to function on just about any level. On Rux for 2 years now I can live a life which is meaningful as long as I take regular rests and monitor my time so that I do not overdo things. I had numerous symptoms from taking Hydroxy which have gone completely since changing to Rux."* *"Hydroxycarbamide gradually became less effective, ruxolitinib has been life changing* *I have felt so much better since I started taking ruxolitinib; it has has given me my life back. I do not experience pruritis so I can bath or shower now without any pain and distress. Previously, I had not been able to bath or shower for 5 years. This was an extremely unsatisfactory situation. I have not suffered with cellulitis infections in more than 4 years. These infections were becoming more frequent and more severe before I started taking ruxolitinib. Fatigue is much improved since taking this medication."* *"Ruxolitinib really helped lower my platelet and haematocrit levels to normal limits. I was unable to use hydroxyurea due to history of melanoma, anagrelide had no effect on my blood levels and although*

	Interferon improved my blood results a little, I felt unwell on it . I was having regular venesections (approximately every couple of weeks) prior to ruxolitinib. I have not needed any venesections in the 2 years since commencing ruxolitinib" The carers said: *"Prior to the ruxolitinib our social life was severely curtailed as my wife did not have the energy for non-essential activities. Since the ruxolitinib a normal social can be sustained."* *"Ruxolitinib has made a big difference to this condition. My wife can carry out daily activities without any assistance"* *"Appetite is much better. Sleeps better and is not so anxious. Reduced spleen size so more comfortable, less itching, less bad days His blood counts are better, very seldom needs phlebotomies."* *"The fatigue is greatly improved, and my wife has much greater energy and can do so much more without my support. The burning sensation following contact with water/cold/heat has been eliminated and the bone pain has gone."* *"My wife is less tired, and therefore more sociable. She does more, works hard and travels more."* *"Ruxolitinib has eased the fatigue; it's nowhere near as bad as previous years"* *"Seems to control the condition much better; ruxolitinib has given him his life back"* The benefits of ruxolitinib are well illustrated by these quotes. The reduction in symptoms has also been captured quantitatively in the MAJIC PV study. The following figure clearly shows a significant and sustained reduction in symptoms:


Estimated Mean Change from Baseline (95% CI)
15
10
5
0
-5
-10
-15
* * * * * * * * * * *
2 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60
Month
Treatment BAT Rux
Sample size at each month
BAT 48 53 45 38 38 42 41 35 31 31 31 29 21 21 16 14
Rux 55 56 53 48 43 46 45 37 44 42 43 38 32 26 21 22

	In preparing this submission, we have interviewed patients who lived with PV for many years and had found that existing treatments did not work for them. Ruxolitinib was prescribed for them as part of the MAJIC or MITHRIDATE trials or under a Compassionate Use protocol. We have summarised their experiences here: **Patient A: a 50 year old psychologist:** *"I was diagnosed with PV 15 years ago. I was treated with regular and frequent venesections During that time, I suffered severely from migraine headaches and visual disturbances, sometimes 2-3 times a week. I would have to take days off work, couldn't drive, and was confined to the house."* *"These symptoms were unresponsive to treatment I received and, 5 years ago, I suffered multiple thromboses in my liver, stomach and gall bladder. We tried Hydroxyurea, but it seriously affected my liver and Interferon resulted in neurological effects."* *"Four and a half years ago, I was prescribed ruxolitinib. There were some side effects at first but since then, I have had no migraines, my blood counts are finally back in normal range, I have no itching and my spleen has shrunk"* *"Prior to the new treatment, I felt my life was constantly precarious, and it was not just my worry – I knew my poor health was something that never really left my daughters' minds"* *"I will have to live with the permanent effects of the thromboses, but I feel that ruxolitinib has halted my disease and I am hugely reassured that it is now under control"* **Patient B: a 64 year old college vice-principal:** *"I was diagnosed in 2004 and , for a few years. I was treated with 3-monthly venesections and aspirin.In 2010 I was enrolled in the vorinostat trial but I was soon suffering from fatigue, rosacea, and enlarged spleen and thinning hair "* *"We then tried Hydroxyurea, but that failed to work for me, with no reduction in my symptoms and I still needed frequent venesections. My illness was having a serios effect on my social life, relationships and my work and causing significant worries for me and my family."*

	"In 2012 I enrolled on the MAJIC PV trial and was prescribed ruxolitinib. The impact was really dramatic - almost immediately, my spleen shrank, and I have had no venesections since taking the drug. I have suffered no side effects and Ruxolitinib has allowed me to start living a normal life again." *"I felt I was at the end of the line until the Ruxolitinib trial came up. My PV has taken a back seat now and doesn't seem terminal anymore."* **Patient C: a 70 year old insurance broker** *''I was diagnosed with PV in 2017 and started having weekly venesections, but it was soon apparent that I needed more treatment. I was not keen on HU because of all the side effects I was aware of, so I agreed with my specialist to try Interferon. It was initially effective, but I then suffered from a rapid deterioration in my blood counts, possibly caused by COVID, or another blood condition I suffer from."* *"I had severe fatigue, and constant bone pain. My brain felt foggy and I had to give up all the things I liked doing like sports, my voluntary work. The frequent venesections gave me anaemia and were a constant drain on my energy. Even walking any distance was hard work – I'd completely lost what it felt like to lead a normal life.""* *" A year ago, I enrolled on the MITHRIDATE trial and was prescribed ruxolitinib. The effect has been rapid and amazing – I feel like I have my old life back. My counts are now stable and in the normal range and I can resume all the things I used to do, both physically and mentally. I have not had a single venesection since I started on the drug."*

Disadvantages of the technology

10. What do patients or carers think are the disadvantages of the technology?	Compared with current standard treatments, ruxolitinib seems to be well tolerated. We are aware that there are risks of lowered blood cell counts (in contrast to PV which is associated with high blood cell counts). In discussion with key opinion leaders, we have ascertained that the risks of infection in particular shingles is commoner with this medication as shown in the RESPONSE trial (approximately 6% of patients). Other infections are also commoner and there are risks of reactivation of hepatitis and TB which are not present with other medication. When we spoke to our clinicians, they felt these risks were mitigated by pre-treatment screening, counselling and for shingles prophylaxis where needed. We were also told about risks of skin cancer especially as this is seen with patients treated with hydroxycarbamide (indeed all patients in the clinical trials with this drug have received hydroxycarbamide and all patients in our survey had also received this drug) then the risks when ruxolitinib is added are greater. In the MAJIC PV study, skin cancer only occurred in patients being treated with ruxolitinib. In addition, ruxolitinib can cause significant weight gain and increases in blood pressure and cholesterol. This medication needs to be taken twice a day (other medication less frequently) and it does interact with more medications but again our patient community (>2000 members with these conditions) did not raise this as a concern, nor did the 35 family members and carers who directly responded to our survey.

Patient population

11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	The primary focus of this appraisal is the significant proportion of PV patients for whom existing treatments have either failed or their side effects are intolerable. These patients are in serious need of further treatment options unless they are to face many years of debilitating symptoms. The availability of a new therapeutic option is particularly important for younger PV patients - 20% of PV patients are under the age of 40, meaning that the increased cancer risk of Hydroxycarbamide would be born for many decades. Because of the known risks of medium- and long-term Hydroxycarbamide treatment, many patients' illness is not treated for many years until more severe symptoms develop and the risk of disease progression has increased. However, we would like to see all PV patients have access to this treatment, as all have significant unmet needs in terms of future options when they cannot tolerate existing treatments, as well as need to improvement quality of life.

Equality

12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	No

Other issues

13. Are there any other issues that you would like the committee to consider?	No
14. **Is splenomegaly a clinically relevant subgroup for polycythaemia vera?**	Enlarged spleen is one of a range of debilitating symptoms that PV patients have to endure. Diagnosis and treatment choice is determined by a number of considerations but we do not believe that spleen size is a particularly dominant factor compared to other symptoms.

Key messages

15. In up to 5 bullet points, please summarise the key messages of your submission.	• MPN Voice strongly supports the use of ruxolitinib for patients with polycythaemia vera who are intolerant or resistant to hydroxycarbamide. • This drug more effectively treats the often disabling symptoms of polycythaemia vera where other drugs have failed and reduces the need for frequent venesections. • An additional treatment option is particularly important for younger PV patients for whom long-term use of Hydroxycarbamide presents unacceptable risk of developing other cancers • There is evidence from clinical trials of a benefit in reducing thrombosis for ruxolitinib treated patients. • The majority of patients with polycythaemia have very limited treatment options and some have severe symptoms in particular itching which is not at all controlled with standard treatments.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - YES or NO

For more information about how we process your personal data please see our privacy notice.

CONFIDENTIAL UNTIL PUBLISHED

External Assessment Group Report commissioned by the NIHR Evidence Synthesis Programme on behalf of NICE

Ruxolitinib for treating polycythaemia vera (ID5106)

Produced by	Southampton Health Technology Assessments Centre (SHTAC)
Authors	Joanne Lord, Professorial Research Fellow, Health Economics Lois Woods, Senior Research Assistant, Evidence Synthesis and Information Specialist Asyl Liyakat Hawa, Research Fellow, Health Economics David Alexander Scott, Principal Research Fellow, Statistics Geoff Frampton, Senior Research Fellow, Evidence Synthesis
Correspondence to	Dr Geoff Frampton Southampton Health Technology Assessments Centre (SHTAC) School of Healthcare Enterprise and Innovation Alpha House Enterprise Road, University of Southampton Science Park Southampton SO16 7NS www.southampton.ac.uk/shtac
Date completed	21st November 2022 Updated following factual accuracy check 8th December 2022

Source of Funding: This report was commissioned by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme as project number 135699.

Acknowledgements

We thank the following for providing expert clinical advice to the project team and for reading and commenting on a draft of the report:

- Dr Andrew Innes, Consultant Haematologist, Imperial College Healthcare NHS Trust and Honorary Clinical Senior Lecture, Imperial College London.
- Dr Srinivasan Narayanan, Consultant Haematologist, University Hospital Southampton NHS Foundation Trust.

We also thank Dr Joanna Picot, Southampton Health Technology Assessments Centre (SHTAC), for reading and commenting on a draft of this report for quality assurance.

Declared competing interests of the authors and advisors

- The authors declare none.
- Dr Innes received an honorarium from Novartis (manufacturer of ruxolitinib) to cover the registration fee for the European Haematology Association Annual Meeting, June 2022. He declares that this did not involve any consultancy work for Novartis.
- Dr Narayanan received an honorarium from Novartis (manufacturer of ruxolitinib) to attend a virtual educational masterclass (10-11 February 2022) covering mixed haematology topics. He declares that he did not attend any sessions relevant to ruxolitinib or polycythaemia vera. Dr Narayanan also received honoraria from Bristol-Myers Squibb (manufacturer of hydroxycarbamide) to attend advisory board meetings on myelodysplastic syndrome and multiple myeloma (ongoing from March 2022); to attend the European Haematology Association meeting (online; June 2022); and to attend a Myeloproliferative Neoplasms Preceptorship virtual meeting covering general topics in myeloproliferative disorders (6-7 October 2022). He is also an investigator on a study of low-risk myelodysplasia unrelated to polycythaemia vera (MIRAS project ID 313756) sponsored by Bristol-Myers Squibb.

Rider on responsibility for report

The view 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.

This report should be referenced as follows:

Lord J, Woods L, Hawa AL, Scott DA, Frampton G. Ruxolitinib for treating polycythaemia vera: A Single Technology Appraisal. Southampton Health Technology Assessments Centre, 2022.

Contributions of authors

Joanne Lord critically appraised the health economic systematic review, critically appraised the economic evaluation, and drafted the report; Lois Woods critically appraised the clinical effectiveness systematic review and drafted the report; Asyl Liyakat Hawa critically appraised the health economic systematic review, critically appraised the economic evaluation, and drafted the report; David Alexander Scott critically appraised the indirect treatment comparison and the clinical effectiveness systematic review and drafted the report; Geoff Frampton critically appraised the clinical effectiveness systematic review, drafted the report and is the project co-ordinator and guarantor.

1 Contents

LIST OF TABLES

LIST of FIGURES

LIST OF ABBREVIATIONS

AE	Adverse event
AIC	Academic in confidence
AML	Acute myeloid leukaemia
BAT	Best available therapy
BNF	British National Formulary
BSH	British Society for Haematology
CHR	Complete haematological remission
CI	Confidence interval
CIC	Commercial in confidence
CRD	Centre for Reviews and Dissemination
CS	Company submission
CQ	Clarification question
CSR	Clinical study report
DSU	Decision Support Unit
EAG	External Assessment Group
EFS	Event-free survival
ELN	European LeukemiaNet
EMA	European Medicines Agency
EMC	Electronic Medicines Compendium
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire
EPAR	European Public Assessment Report
EQ-5D-3L	European Quality of Life Working Group Health Status Measure 3 Dimensions, 3 Levels
EQ-5D-5L	European Quality of Life Working Group Health Status Measure 5 Dimensions, 5 Levels
EQ-VAS	EuroQol Visual Analogue Scale
ESMO	European Society for Medical Oncology
GEMFIN	Grupo Español de Enfermedades Mieloproliferativas Filadelfia Negativas
HC	Hydroxycarbamide (this is synonymous with hydroxyurea)
HCT	Haematocrit
HMRN	Haematological Malignancy Research Network

HR	Hazard ratio
HRG	Healthcare Resource Group
HRQoL	Health-related quality of life
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
IFN	Interferon
IPD	Individual patient level data
IRR	Incidence-rate ratios
ITC	Indirect treatment comparison
ITT	Intent(ion) to treat
JAK	Janus-associated Kinase
KM	Kaplan-Meier
MAIC	Matched-adjusted indirect comparison
MDS	Myelodysplastic syndrome
MF	Myelofibrosis
MF-8D	Myelofibrosis 8 dimensions health outcome measure
MHRA	Medicines and Healthcare products Regulatory Agency
MPN(s)	Myeloproliferative neoplasm(s)
MPN-SAF TSS	Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score
NA	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMSC	Non-melanoma skin cancer
NR	Not reported
OR	Odds ratio
OS	Overall survival
PAS	Patient Access Scheme
PFS	Progression-free survival
PSA	Probabilistic sensitivity analysis
PSM	Partitioned survival model
PSS	Personal Social Services
PV	Polycythaemia vera
R/I	Resistant and/or intolerant
QALY	Quality-adjusted life year

QoL	Quality of life
RCT	Randomised controlled trial
RR	Relative risk/risk ratio
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of product characteristics
STM	State transition model
TA	Technology appraisal
TE	Thromboembolic event
TEAE	Treatment-emergent adverse event
TFS	Transformation-free survival
TSD	Technical Support Document
TTD	Time to discontinuation
UK	United Kingdom
US	United States
VAS	Visual analogue scale
WBC	White blood cell(s)

EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the external assessment group (EAG) as being potentially important for decision making. It also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.6 explain the key issues in more detail. Background information on the condition, health technology, evidence and information on the issues are in the main EAG report.

All issues identified represent the EAG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

Table 1 Summary of key issues

Issue number	Headline description	EAG report section
1	Relevance of the trial populations for modelling UK practice	4.2.3
2	Modelling the relative treatment effect for overall survival	4.2.6.2.1
3	Waning of the treatment effect	4.2.6.2.1
4	Modelling approach: state-transition or partitioned-survival	4.2.2.3
5	Model structure: health states and events	4.2.2.3
6	Extrapolation of time to ruxolitinib discontinuation	4.2.6.1.1
7	Source for utility estimates: MF-8D or EQ-5D	4.2.7.2

The key differences between the company's preferred assumptions and the EAG's preferred assumptions are :

- Use of the general population mortality constraint for survival prior to discontinuation of ruxolitinib throughout the time horizon, rather than only post-trial.
- Partitioning of the best available treatment (BAT) state into substates for first BAT, second or subsequent BAT and no further BAT.
- Estimates for the hazard ratio (HR) for ruxolitinib compared with BAT from the MAJIC-PV trial, constant or time-varying HR.

- The distribution used for extrapolation of the time to ruxolitinib discontinuation.
- Source for estimates of utilities for ruxolitinib and BAT: EQ-5D values from RESPONSE-2 trial data or MF-8D values from the RESPONSE trial.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Lower mortality rates while patients are on ruxolitinib than with standard therapies.
- Better health-related quality of life (utility) while patients are on ruxolitinib than during treatment with standard therapies alone.
- Small overall increase in utility due to reduced incidence of myelofibrosis, thromboembolism, haemorrhage, adverse reactions and therapeutic phlebotomy.

Overall, the technology is modelled to affect costs by:

- The high cost of ruxolitinib compared with standard drug treatments.
- Savings due to reduced use of therapeutic phlebotomy and reduced follow-up and monitoring after the first six months of treatment with ruxolitinib.
- Savings due to reduced need for treatment of myelofibrosis, haemorrhage, thromboembolism and adverse reactions.
- Some additional costs for treatment of non-melanoma skin cancer, acute myeloid leukaemia and myelodysplastic syndrome.

The modelling assumptions that have the greatest effect on the ICER are:

- The hazard ratio for overall survival with ruxolitinib compared with best available therapy.
- Assumptions about waning of the treatment effect for overall survival.
- The distribution used for extrapolation of time to discontinuation of ruxolitinib.
- Use of EQ-5D or MF-8D utility estimates for ruxolitinib and best available therapy.

1.3 The decision problem: summary of the EAG's key issues

The EAG have not identified any key issues with the decision problem. Other issues relating to the decision problem are discussed in section 1.6 below.

1.4 The clinical effectiveness evidence: summary of the EAG's key issues

The EAG have not identified any key issues with the clinical effectiveness evidence. Other issues relating to the clinical effectiveness evidence are discussed in section 1.6 below.

1.5 The cost-effectiveness evidence: summary of the EAG's key issues

Issue 1 Relevance of the trial populations for modelling UK practice

Report section	4.2.3
Description of issue and why the EAG has identified it as important	There is some uncertainty over whether the MAJIC-PV trial or the company's RESPONSE and RESPONSE-2 trials provide a better basis for modelling survival for the relevant patient population in UK practice. This issue is important because cost-effectiveness estimates differ for versions of the model based on the three trial populations. The EAG considers that, as MAJIC-PV was a wholly UK based trial, it is more obviously relevant for the UK PV population and clinical context. This reflects the view of clinical experts consulted by the EAG. The company have put forward the view that the patients recruited to MAJIC-PV represent a 'high-risk' subgroup of the licensed indication for ruxolitinib. In their 'primary' model, the company use survival extrapolations fitted to RESPONSE and RESPONSE-2 data. Alongside this, they report a 'subgroup model' with extrapolations fitted to MAJIC-PV data.
What alternative approach has the EAG suggested?	We consider that the MAJIC-PV trial population is likely to provide a more appropriate basis for modelling outcomes in UK practice. But we also report cost-effectiveness results based on the RESPONSE and RESPONSE-2 populations, as these provide a comparison for the subgroups with and without splenomegaly.
What is the expected effect on the cost-effectiveness estimates?	ICER estimates are lower for the MAJIC-PV population. With the company's base case assumptions, the ICERs are [redacted], [redacted] and [redacted] per QALY for the RESPONSE, RESPONSE-2 and MAJIC-PV populations, respectively. With the EAG preferred assumptions, these ICERs are [redacted], [redacted] and [redacted] respectively.

What additional evidence or analyses might help to resolve this key issue?	Further expert opinion and evidence on the relevance of the three trial populations to UK practice.

Issue 2 Modelling the relative treatment effect for overall survival

Report section	4.2.6.2.1, Table 22, Table 27 and Figure 5 and Figure 6 below show the KM data with the company's choice of distribution for TTD for ruxolitinib due to reasons other than death in comparison with the selected scenario distributions from Table 27 above for the licensed population with and without splenomegaly.  **Figure 5 Comparison of selected scenario distributions for TTD for ruxolitinib for the licensed population with splenomegaly** Abbreviations: KM: Kaplan-Meier; TTD: time to treatment discontinuation; OS: overall survival. Source: Reproduced from CS Appendix N Figure 18 using selected distributions.

	 **Figure 6 Comparison of selected scenario distributions for TTD for ruxolitinib for the licensed population without splenomegaly** Abbreviations: KM: Kaplan-Meier; TTD: time to treatment discontinuation; OS: overall survival. Source: Reproduced from CS Appendix N Figure 19 using selected distributions. Table 28 below shows cost-effectiveness results for selected company scenarios for the MAJIC-PV population analysis. Again, from the many scenarios conducted by the company, we have selected scenarios that relate to key uncertainties and that have an impact on the ICERs. Table 28
Description of issue and why the EAG has identified it as important	Cost-effectiveness is highly sensitive to the relative treatment effect on overall survival. The company use results from the MAJIC-PV trial to inform estimates for their base case analyses. We agree with this decision as cross-over within the RESPONSE and RESPONSE-2 trials means that estimates of treatment effects from these trials are highly confounded. The EAG are not aware of any other data that would provide a more robust analysis. Other sources of evidence regarding the effect of ruxolitinib on survival, including the company's ITC and an analysis of Spanish registry data are less robust. The currently unpublished manuscript for the MAJIC-PV reports a hazard ratio for overall survival (ruxolitinib compared with best available treatment) of 0.73 (95% CI 0.36 to 1.50; p=0.39). However, the company use a time-varying estimate of the hazard ratio, which they estimated with a piecewise Cox proportional hazards model using reconstructed Kaplan-Meier data from MAJIC-PV. This includes a bigger treatment effect (lower HR) from year 3.0 onwards: 0.91 (95% CI

	0.38 to 2.18) 0 to 3 years and 0.45 (95% CI 0.13 to 1.61) 3 to 5 years. The company justifies this approach based on expert advice and visual inspection and analysis of the MAJIC-PV KM results.
What alternative approach has the EAG suggested?	The EAG prefer the constant HR estimate from MAJIC-PV due to uncertainty over the statistical validity of the company's post hoc analysis. However, we report a scenario results with the company's time-varying HR, as this may be considered clinically plausible.
What is the expected effect on the cost-effectiveness estimates?	The HR for OS has a large impact on the ICER. The company's base case estimates increase from [redacted], [redacted] and [redacted] (RESPONSE, RESPONSE-2 and MAJIC-PV populations respectively), to [redacted], [redacted] and [redacted]
What additional evidence or analyses might help to resolve this key issue?	Further expert opinion on the plausibility of an increasing relative effect on survival over time. The economic analyses for subgroups with and without splenomegaly currently use the same estimates of treatment effects, estimated form the MAJIC-PV trial. Further analysis should be conducted to update these analyses if subgroup analysis of MAJIC-PV data by splenomegaly status.

Issue 3 Waning of the treatment effect

Report section	4.2.6.2.1
Description of issue and why the EAG has identified it as important	In their base case analyses, the company assume that the treatment effect diminishes linearly from the end of trial follow-up (5 years) and stops at 20 years (HR=1). This was based on clinical expert judgement that approximately twice the number of patients would be alive at 20 years with ruxolitinib compared with current treatment (see CS section B.3.3.4). The company note uncertainty over these assumptions, and report scenario analysis with the period of waning varied from 5 to 50 years.
What alternative approach has the EAG suggested?	We have not changed the company's waning assumptions in EAG preferred analysis, as the assumption of waning might be seen to mitigate against uncertainty over the treatment effect. However, we note that it might be appropriate to use a longer waning period, or to remove waning from the model, when used in combination with the more conservative fixed HR estimate.
What is the expected effect on the cost-effectiveness estimates?	The duration of waning has a big impact on the ICER. For example, the company's base case ICER for the MAJIC-PV population is [redacted] with a loss of effect at 10 years, and [redacted] with loss of effect at 30 years.
What additional evidence or analyses might help to resolve this key issue?	Further expert opinion on the plausibility of waning from a biological and clinical perspective.

Issue 4 Modelling approach: state-transition or partitioned-survival

Report section	4.2.2.3
Description of issue and why the EAG has identified it as important	It is not clear if different results from the company's state-transition model (STM) for the RESPONSE and RESPONSE-2 populations and their partitioned-survival model (PSM) for the MAJIC-PV population relate to differences in the modelling technique or to the different populations. This adds structural uncertainty to the interpretation of the economic evaluation results. NICE Decision Support Unit Technical Support Document 19 reports that STM and PSM models can give very different results, and that it is not clear which approach is more reliable. TSD19 therefore recommends parallel development of STM and PSM models to verify the plausibility of PSM extrapolations.
What alternative	Comparison of alternative modelling approaches (STM and

approach has the EAG suggested?	PSM) within the same dataset.
What is the expected effect on the cost-effectiveness estimates?	Unknown
What additional evidence or analyses might help to resolve this key issue?	Development of a PSM for the RESPONSE and RESPONSE-2 populations to enable comparison with results from the STM model. It is not possible for the company to develop an STM for the MAJIC-PV population, as they do not have access individual patient data. However, we would encourage the MAJIC-PV investigators to consider appropriate economic evaluation based on the trial data, or to make the data available for such an analysis.

Issue 5 Model structure: health states and events

Report section	4.2.2.3
Description of issue and why the EAG has identified it as important	The EAG also has concerns over the structure of the company's models, as they do not reflect the natural history of PV, and therefore may not reflect long-term impacts of the condition on survival and quality of life. The model 'health states' are based on treatment phases (before and after discontinuation of ruxolitinib) rather than on stages of disease. Although discontinuation of ruxolitinib is likely to be related to long-term survival, other intermediate outcomes such as progression-free survival or event-free survival are likely to be more strongly prognostic. Another problem with the current structure, is that the best available therapy (BAT) arm is modelled with a single health state, with three substates for first-line, second and subsequent line, and discontinuation of all BAT. EAG clinical advisors have suggested that this progression between lines of therapy does not reflect current practice. Furthermore, the decrements in utility for the latter two substates are based on assumption, rather than evidence. We also have concerns that the company's model structure does not reflect increasing risks of key complications of PV, such as myelofibrosis, and major thromboembolic or haemorrhagic events with age. The use of fixed incidence annual rates for these events is not realistic.

What alternative approach has the EAG suggested?	Consideration of an alternative model structure based on a measure of disease progression and a simplified approach to modelling the subsequent types of event.
What is the expected effect on the cost-effectiveness estimates?	Unknown
What additional evidence or analyses might help to resolve this key issue?	Exploration of an alternative model structure to better reflect the natural history of PV.

Issue 6 Extrapolation of time to ruxolitinib discontinuation

Report section	4.2.6.1.1
Description of issue and why the EAG has identified it as important	The results for the company's primary analysis based on the RESPONSE and RESPONSE-2 trials were moderately sensitive to the distribution used for the time to treatment discontinuation. The company used an odd spline model with one knot for the extrapolation of TTD for ruxolitinib due to reasons other than death in the primary analysis. The same distribution was used for both RESPONSE and RESPONSE-2 trial data. The EAG note that, in the primary analysis, pre- and post-discontinuation survival for ruxolitinib make use of pooled RESPONSE and RESPONSE-2 data, as few deaths were observed in the trial, whereas data from the two trials are used separately for TTD for ruxolitinib due to reasons other than death.
What alternative approach has the EAG suggested?	The EAG have selected the Weibull distribution as a preferred assumption for TTD for ruxolitinib, a parametric distribution which has a better fit the RESPONSE trial data more appropriately. The Weibull distribution has a similar fit for the RESPONSE-2 trial data.
What is the expected effect on the cost-effectiveness estimates?	Implementing a Weibull distribution in place of an odds spline model in the company base case reduces the ICER for the licensed population with splenomegaly to [redacted] per QALY and increases the ICER for the licensed population without splenomegaly to [redacted] per QALY.
What additional evidence or analyses might help to resolve this key issue?	Additional scenario using pooled IPD from RESPONSE and RESPONSE-2 trials for TTD for ruxolitinib due to reasons other than death.

Issue 7 Source for utility estimates: MF-8D or EQ-5D

Report section	4.2.7.2
Description of issue and why the EAG has identified it as important	There is uncertainty over the most appropriate instrument to estimate utilities for the economic model. This has a large impact on the ICER. Utilities are available from two sources: EQ-5D-5L data from the RESPONSE-2 trial, and estimates from data collected in the RESPONSE trial and valued using the MF-8D, which is a disease-specific utility measure developed for myelofibrosis. The company argue that the EQ-5D is not appropriate for PV, based on psychometric evidence and precedent for myelofibrosis (TA386 and TA756), and the similar nature of symptoms for PV and MF. They also report an exploratory psychometric analysis comparing RESPONSE-2 data for the EQ-5D and a PV symptom score (the MPN-SAF). This provides some evidence in favour of the MF-8D, including greater responsiveness and lower susceptibility to ceiling effects. However, the MF-8D was not developed for use in PV, and the company had to make assumptions to substitute the PV symptom score for the myelofibrosis symptom score used in the MF-8D. There is also a lack of direct evidence validating the EQ-5D and MF-8D in a PV population.
What alternative approach has the EAG suggested?	We use EQ-5D utilities in the EAG preferred analysis. This follows the NICE preference for use of the EQ-5D when available from relevant clinical trials and improves consistency across NICE appraisals. There is some evidence in favour of the MF-8D measure, but also uncertainty about its transferability from MF to PV.
What is the expected effect on the cost-effectiveness estimates?	Replacing MF-8D with EQ-5D utilities in the company's base case increases the ICER for the MAJIC-PV population [redacted] to [redacted] per QALY. Increases are similar in the RESPONSE and RESPONSE-2 populations.
What additional evidence or analyses might help to resolve this key issue?	Further evidence that the EQ-5D is not appropriate for people with PV. Comparative evidence for the psychometric performance of MF-8D and EQ-5D utilities for a population with PV

1.6 Other issues

The company have excluded radioactive phosphorus from their decision problem although this is stated as a relevant comparator in the NICE scope. As explained in section 2.3.2

below, we believe the exclusion of radioactive phosphorus is appropriate and unlikely to influence validity of the cost-effectiveness results.

The results of the company's indirect treatment comparison (ITC) for overall survival are highly uncertain, primarily due to limited adjustment for imbalances in prognostic factors between the treatment groups (section 3.4). However, the EAG are not aware of alternative data sources that would enable a more robust ITC analysis to be conducted. Overall survival estimates from the ITC are not used in the company's economic analysis base case but do inform scenario analyses (section 4.2.6.2.1).

All three randomised controlled trials included by the company are at high risk of bias, due to the open-label nature of the trials, confounding of long-term outcomes by crossover in the RESPONSE and RESPONSE-2 trials, selective reporting of HRQoL outcomes, and the handling of missing data for HRQoL outcomes in all three trials. For MAJIC-PV there is additionally a lack of clarity around the randomisation process and there are some differences in patient characteristics between the treatment arms (section 3.2.3). Limitations of the existing data and reporting mean that the clinical efficacy outcomes are subject to uncertainty that would be difficult to resolve unless new evidence (and clearer reporting of studies) becomes available. The high risk of bias means that variance estimates from the three RCTs such as 95% confidence intervals would underestimate the uncertainty present.

The survival extrapolations used in the company's base case incorporate a constraint to ensure that the mortality rate cannot be less than that in the general population (adjusted for age and gender). This constraint is applied through the time horizon, except for survival prior to discontinuation of ruxolitinib in the company's primary model, for which the general population mortality constraint was only applied after the trial period (5 years). In response to clarification question B4, the company provided a scenario analysis including the mortality constraint throughout the time horizon and a revised version of their model with an option to apply this scenario. We consider this to be a correction to the company's model and have applied it in EAG preferred analyses.

Other issues that have a limited impact on ICERs are: the EAG adjustment to the cost of managing grade 1 and 2 thromboembolic events; and use of the partition of the BAT state to model first line BAT, second and subsequent line BAT and no further BAT substates.

1.7 Summary of EAG's preferred assumptions and resulting ICER

We made the following changes to the company's base case analyse in the EAG preferred analysis:

- Correction to apply the general population mortality constraint for survival prior to discontinuation of ruxolitinib throughout the time horizon
- The partition of the BAT health state was not used
- Constant HR for overall survival from the MAJIC-PV trial
- Weibull extrapolation for time to ruxolitinib discontinuation in the primary model
- EQ-5D utility values estimated from the RESPONSE-2 trial
- Additional costs for management of Grade 1-2 thromboembolic events

Table 2 Summary of cost-effectiveness results

Scenario	Incremental cost	Incremental QALYs	ICER (change from company base case)
Company's base case			
RESPONSE trial population (with splenomegaly)	[redacted]	[redacted]	[redacted]
RESPONSE-2 trial population (without splenomegaly)	[redacted]	[redacted]	[redacted]
MAJIC-PV trial population	[redacted]	[redacted]	[redacted]
EAG's preferred base case			
RESPONSE trial population (with splenomegaly)	[redacted]	[redacted]	[redacted]
RESPONSE-2 trial population (without splenomegaly)	[redacted]	[redacted]	[redacted]
MAJIC-PV trial population	[redacted]	[redacted]	[redacted]

Modelling errors identified and corrected by the EAG are described in section 5.3.3. For further details of the exploratory and sensitivity analyses done by the EAG, see section 6.1.

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

This report is a critique of the company's submission (CS) to NICE from Novartis on the clinical effectiveness and cost effectiveness of ruxolitinib for treating polycythaemia vera (PV). It identifies the strengths and weaknesses of the CS. Clinical experts were consulted to advise the external assessment group (EAG) and to help inform this report.

Clarification on some aspects of the CS was requested from the company by the EAG via NICE on 10th October 2022. A response from the company via NICE was received by the EAG on 27th October 2022 and this can be seen in the NICE committee papers for this appraisal.

2.2 Background

2.2.1 Background information on polycythaemia vera

Polycythaemia vera (PV) is a type of myeloproliferative neoplasm (MPN) characterised by overproduction of blood cells and platelets in the bone marrow, particularly red blood cells (erythrocytosis).[1 2] The uncontrolled nature of the proliferation of blood cells defines PV as a cancer.[3]

CS section B.3.1.1 provides a clear overview of the disease including: a brief description; epidemiology; relevance of the Janus-associated Kinase (JAK) 2 mutation; diagnosis (blood cell counts, and the haematocrit which is the proportion of red blood cells in a volume of blood, usually expressed as a percentage); symptoms (the most significant being splenomegaly, pruritus (itching), and fatigue); mortality associated with thromboembolic events, cardiovascular events and disease progression; and a discussion of the definitions of high-risk disease and resistance or intolerance to hydroxycarbamide (also discussed in section 3.2.1 of this report).

CS section B.1.3.1 notes the association of increased haematocrit (HCT) levels, i.e. an increased red blood cell mass with vascular complications. This is consistent with the British Society for Haematology (BSH) guidelines that show an increased HCT of >0.45 is a risk factor for thrombosis which in turn is a risk factor for overall survival, hence HCT control is a key goal of therapy.[4]

The EAG note that up to date incidence and prevalence data for PV specifically for England are not available. Data for the UK are available from the Haematological Malignancy Research Network (HMRN) which gives a crude estimate for incidence as 1.7 per 100,000, a prevalence of 1.9 per 100,000, and 1130 expected UK cases per year.[5] These figures sit within the ranges estimated from European registry data and other sources provided in the CS (CS section B.1.3.1).

The current treatment pathway is discussed in CS section B.1.3.2 and covers treatment goals, the course of disease progression, first-, second- and third-line treatments, unmet need, and the safety profile of other cytoreductive therapies.

- As stated in the CS, the main goals of treatment are to reduce the incidence of thrombotic and haemorrhagic complication and the long-term risk of transformation to myelofibrosis (MF) or acute myeloid leukaemia (AML).[4 6]

- European and UK guidelines exist: the European Society for Medical Oncology (ESMO) clinical practice guidelines for Philadelphia chromosome-negative chronic MPNs (which includes PV)[7]; the Pan-London Haemato-Oncology Clinical Guidelines for MPNs;[6] and the British Society for Haematology (BSH) guideline for the diagnosis and management of PV.[4] These guidelines are generally similar and have authors in common, the CS refers to the BSH guideline only which is appropriate as it is recent (2019) and applies to the whole of England.

- Cytoreductive therapy is appropriate in certain low-risk patients, for example if white blood cell (WBC) levels are high or if phlebotomy tolerability is poor. This means that such otherwise low-risk patients would join the high-risk pathway shown in CS Figure 3 (although this reason for joining the high-risk pathway is not shown fully in CS Figure 3). Therefore, not all patients who receive hydroxycarbamide may have necessarily met the criteria for high-risk based on their age or prior thrombosis.

- CS Figure 3 accurately represents the BSH recommendations for management options in high-risk patients, that is first-line treatment with either hydroxycarbamide or interferon-alfa, and second-line treatment switching to whichever of hydroxycarbamide or interferon-alfa they did not receive first-line.[4] The EAG's clinical experts agree that for first- and second-line treatments this is a good representation of clinical practice except that two of the treatments listed for third-line, pipobroman

and radioactive phosphorus, are no longer used (see decision problem section 2.3.2 below). One clinical expert noted that the diagram does not show that in clinical practice patients often cycle on and off hydroxycarbamide, or between hydroxycarbamide and interferon-alfa, to manage side-effects.

- CS Figure 3 refers to interferon-alfa but we note that, according to the BNF[8] and British PV guidelines, interferon-alfa has been superseded by peginterferon-alfa,[6] or is recommended in preference to interferon-alfa.[4] One of the EAG's clinical experts commented that pegylated interferon-alfa may be offered to patients who cannot tolerate interferon-alfa or hydroxycarbamide, but tolerance remains relatively poor so extensive monitoring is still required. The company's economic analysis uses costs for peginterferon-alfa (section 4.2.8.2) which the EAG agree is appropriate.

- Not all patients respond to or can tolerate hydroxycarbamide, hence the population group for the licensed indication. The CS refers to the updated ELN consensus criteria for resistance to or intolerance of hydroxycarbamide for use in clinical trials published in 2022 (CS Table 4),[9] and also states that these criteria are not always used in clinical practice, confirmed by the EAG's clinical experts (see the decision problem discussion for the population in section 2.3.1). However, the original ELN consensus criteria for resistance to or intolerance of hydroxycarbamide are relevant here as they applied at the time the studies included in the CS were conducted. Those criteria are published in Barosi et al. 2010 and duplicated in Table 3 below.[10]

Table 3 ELN definition of resistance/intolerance to hydroxycarbamide in patients with PV from Barosi et al. 2010[10]

Definition of resistance/intolerance to hydroxycarbamide in patients with polycythaemia vera	
1	Need for phlebotomy to keep haematocrit <45% after 3 months of at least 2 g/day of Hydroxycarbamide, OR
2 a	Uncontrolled myeloproliferation, i.e. platelet count >400 x 10^9/l AND white blood cell count >10 x 10^9/l after 3 months of at least 2 g/day of Hydroxycarbamide, OR
3	Failure to reduce massive[a] splenomegaly by more than 50% as measured by palpation, OR failure to completely relieve symptoms related to splenomegaly, after 3 months of at least 2 g/day of Hydroxycarbamide, OR
4	Absolute neutrophil count <1.0 x 10^9/l OR platelet count <100 x 10^9/l or haemoglobin <100 g/l at the lowest dose of Hydroxycarbamide required to achieve a complete or partial clinico-haematological response[b], OR
5	Presence of leg ulcers or other unacceptable Hydroxycarbamide-related non-haematological toxicities, such as mucocutaneous manifestations, gastrointestinal symptoms, pneumonitis or fever at any dose of Hydroxycarbamide

[a] Organ extending by more than 10 cm from the costal margin.
[b] Complete response was defined as: haematocrit <45% without phlebotomy, platelet count ≤400 x 109/l, white blood cell count ≤10 x 109/l, and no disease related symptoms. Partial response was defined as: haematocrit <45% without phlebotomy, or response in three or more of the other criteria (Barosi et al, 2009).
Table sourced directly from: Barosi et al. 2010[10]

2.2.2 Background information on ruxolitinib

A description of ruxolitinib, brand name Jakavi®, is provided in CS section B.1.2. Ruxolitinib is a JAK1 and JAK2 protein kinase inhibitor that inhibits dysfunctional signalling pathways caused by JAK gene mutations, reducing the excessive production of red blood cells which is characteristic of PV. Ruxolitinib aims to reduce symptoms and control HCT levels in order to reduce the risk of thromboembolic events and the associated complications which can lead to death.

Ruxolitinib is licensed for the treatment of adult patients with PV who are resistant to or intolerant of hydroxycarbamide. European Medicines Agency (EMA) marketing authorisation was granted in January 2015 and UK marketing authorisation was granted in January 2021.[11] Ruxolitinib is also licensed for use in myelofibrosis and graft versus host disease.

A summary of product characteristics (SmPC) for the 10 mg tablet of ruxolitinib is provided in CS Appendix C. Ruxolitinib is taken orally in tablet form with a starting dose for PV of 10 mg twice daily. The SmPC provided in CS Appendix C specifies a 10 mg tablet only, but dosage information in CS Table 2 outlines 5 mg increments for titration based on safety and efficacy up to a maximum of 25 mg twice daily. The MHRA website lists all SmPCs for each of the 5,

10, 15 and 20 mg tablets.[11-14] Doses may be increased if efficacy is insufficient and blood counts are adequate, and they may be decreased or discontinued if blood counts fall below specified thresholds.[11] Therefore, complete blood cell counts should be evaluated prior to treatment with ruxolitinib and regularly thereafter as advised in the SmPC.[11]

2.2.3 The position of ruxolitinib in the treatment pathway

CS section B.3.1.2 proposes ruxolitinib as an alternative cytoreductive therapy as a treatment option for patients with PV who are resistant to or intolerant of hydroxycarbamide which they may have received either first-line or second-line. This is in line with positioning in the scope of this appraisal and as recommended by the BSH.[4]

One of the EAG's two clinical experts suggested that ruxolitinib might be used second-line after interferon-alfa because some patients receive interferon-alfa as their first cytoreductive therapy due to hydroxycarbamide not being suitable (e.g. younger age/family planning). However, those reasons (younger age/family planning) are not part of the definition of resistance to or intolerance of hydroxycarbamide so those patients would not be in the licensed indication. The other clinical expert said there are no data to support ruxolitinib use after interferon-alfa as first line therapy. They explained that as patients often cycle between hydroxycarbamide and interferon-alfa therapies that could create a circumstance for use of ruxolitinib third-line according to CS Figure 3.

The EAG's clinical experts indicated that they are familiar with using ruxolitinib, at higher doses, in myelofibrosis (MF) patients for whom the drug was recommended in 2016 according to NICE guideline TA386.[15] Ruxolitinib was also used in 38 UK centres as part of the MAJIC-PV randomised controlled trial (RCT) between 2012 and 2022 for PV.[16] Therefore, the NHS has experience of using ruxolitinib to treat myeloproliferative diseases.

EAG conclusions

The company's description of the care pathway appears appropriate, although in relation to the positioning of ruxolitinib in the pathway, there was a difference of opinion between the EAG's clinical experts about whether treatment with ruxolitinib might follow treatment with first-line interferon-alfa.

2.3 Critique of the company's definition of the decision problem

Table 4 compares the company's decision problem to the final scope for this appraisal issued by NICE. The ERG considers that the decision problem adheres to the NICE scope but with the following caveats relating to the population and comparators.

2.3.1 Population

The populations stated in the NICE scope and company decision problem are consistent. However, the EAG's clinical experts commented that definitions of hydroxycarbamide resistance and intolerance are not standardised in clinical practice so there is some uncertainty as to how well the definitions used in the clinical trials would match those used in clinical practice. The definition of intolerance can be somewhat subjective (e.g. reliant on judging the tolerability of a skin rash, leg ulcer or fatigue). One expert commented that the criteria defining hydroxycarbamide resistance and intolerance are more stringent than would be used in clinical practice. Note that the European LeukemiaNet (ELN) have recently published a consensus-based update of the definition of hydroxycarbamide resistance and intolerance (CS Table 4) (Marchetti et al. 2022[9]) but the clinical trials were completed prior to this definition being approved (clarification response A1).

2.3.2 Comparators

The EAG's clinical experts concurred that hydroxycarbamide and interferon-alfa are the most relevant comparators, with anagrelide, busulfan and radioactive phosphorus used rarely if at all:

- Radioactive phosphorus is specified in the NICE scope but excluded from the company's decision problem as the company argue that it is no longer used in practice (CS Table 1). One of the EAG's clinical experts commented that radioactive phosphorus has highly variable availability and is used very rarely. It is a one-off treatment that covers 6 months so may be of benefit for elderly frail patients unable to tolerate frequent treatments. However, it does increase the risk of leukaemia. The other expert stated that radioactive phosphorus is generally unavailable and not used. British PV guidelines suggest that radioactive phosphorus is only suitable for people with limited life expectancy.[4 6] The company have not included radioactive phosphorus among the best available therapy (BAT) treatments in their economic analysis (section 4.2.8.2) which the EAG believe is appropriate.
- Anagrelide / busulfan: Both clinical experts said they would rarely use these therapies. One commented that anagrelide increases the risk of transformation to myelofibrosis or acute myeloid leukaemia (AML) and has a poor side-effects profile especially for elderly

people. British PV guidelines suggest that anagrelide is rarely used as it is relatively platelet-specific, but it may be used in combination with hydroxycarbamide for people with difficult platelet control.[4 6] Busulfan increases the risk of transformation to leukaemia and is only used for people with limited life expectancy.[4 6]

- The NICE scope and company decision problem refer to interferon-alfa. As noted in section 2.2.1 above, interferon-alfa has largely been replaced in practice by peginterferon-alfa which has a relatively better tolerability.

Table 4 Summary of the decision problem

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
Population	Adults with PV that is resistant or intolerant to hydroxycarbamide	In line with final scope	Not applicable	The scope and decision problem are consistent. However, the EAG's clinical experts noted that there is no single standard definition of hydroxycarbamide resistance or intolerance in clinical practice and definitions of intolerance may be subjective (section 2.3.1).
Intervention	Ruxolitinib with established clinical management	In line with final scope	Not applicable	The scope and decision problem are consistent.
Comparators	Established clinical practice without ruxolitinib, comprising of treatment with phlebotomy and aspirin, and: • hydroxycarbamide • IFN-alfa • anagrelide • busulfan • radioactive phosphorus	Established clinical practice defined as treatment with phlebotomy and aspirin, and BAT, including: • hydroxycarbamide • IFN-alfa • anagrelide • busulfan	Radioactive phosphorus was listed in the final scope but excluded in the submission as clinical feedback indicated that this is no longer used in the UK (CS Table 1)	The EAG's clinical experts commented that hydroxycarbamide and IFN-alfa (or pegylated IFN-alfa) are the main comparators; the other therapies are used rarely if at all. The EAG agree with the exclusion of radioactive phosphorus (section 2.3.2)
Outcomes	The outcome measures to be considered include: • CHR (including reporting of HCT, WBC count and platelet count	Key outcomes are: • CHR including reporting of HCT, WBC count and platelet count separately • TTD	Not applicable	The company's outcomes are consistent with those specified in the NICE scope (NB the scope does not explicitly mention overall survival but it's

	separately) • TTD • mortality • symptom relief (including a reduction in spleen size, itching, fatigue and phlebotomy) • thrombosis • progression to AML or MF • adverse effects of treatment • HRQoL	• OS • symptom relief (including a reduction in spleen size, itching, fatigue and phlebotomy) • thrombosis • safety (including transformation to AML/MF and adverse events) • HRQoL		inclusion in the decision problem is appropriate). Mortality is not listed in the decision problem but is reported by the company trials and CS. Note that itching and fatigue are assessed by HRQoL instruments whilst thrombosis is reported as an adverse event.
Subgroups	People with and without splenomegaly	In line with final scope	Additional subgroup based on MAJIC-PV population (high-risk PV)	Each subgroup (with splenomegaly, without splenomegaly, and high-risk patients) is represented by a separate clinical trial.

Source: CS Table 1 with modifications. AML: acute myeloid leukaemia; BAT: best available therapy; CHR: complete haematological remission; HCT: haematocrit; HRQoL: health-related quality of life; IFN: interferon; MF: myelofibrosis; OS: overall survival; TTD: time to discontinuation; WBC: white blood cells

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The company carried out a systematic literature review (SLR) that aimed to identify RCTs on the clinical efficacy and safety of any treatment in PV patients. The SLR was generally well-conducted and the EAG believe all relevant RCTs have been identified.

CS Appendix D.1.3 states the SLR identified eight unique clinical studies but only four are included in the submission. The four excluded studies had been identified according to the SLR eligibility criteria (CS Appendix D Table 8) which includes any intervention and any comparator and so the criteria are broader than both the NICE scope and the company decision problem. The reasons for exclusion are not given, but the EAG believe the studies were excluded appropriately:

- ARD12042:[17] a randomised phase 2 dose-finding study of fedratinib. This treatment is not a comparator.
- NCT00928707 (UCT1):[18] a randomised phase 2 dosing study of givinostat. This treatment is not a comparator.
- NCT00726232:[19] a randomised phase 2 dose-finding study of ruxolitinib. There was no best available therapy (BAT) arm therefore the evidence is inferior to the pivotal trials. Discussed in a footnote in CS section B.2.2.
- RELIEF (NCT01632904):[20 21] RCT for ruxolitinib versus hydroxycarbamide plus placebo. Discussed in CS section B.2.2 and excluded as the population was not resistant to or intolerant of hydroxycarbamide according to the modified ELN criteria. Study details are in CS Appendix D.1.3. The EAG note that the population "had been receiving a stable dose of hydroxycarbamide and were generally well controlled but still reported disease-associated symptoms". The EAG's clinical experts both agreed that the population in RELIEF is not reflective of patients resistant to or intolerant of hydroxycarbamide in the NHS PV population.

The SLR only searched for RCTs and indirect comparisons (referred to as matched-adjusted indirect comparisons, MAICs) but not observational studies or real-world evidence due to the use of an RCT study design filter in the searches. An indirect comparison comparing the ruxolitinib arm of RESPONSE against BAT data from a real-world registry (GEMFIN) was included and is used in the company's ITC (section 3.3). However, it is not transparent whether the GEMFIN registry is the only source of relevant comparator evidence suitable for

use in the company's ITC analysis (see section 3.3.2 for the critique of studies included in the ITC).

The three RCTs included by the company have been completed and are summarised below (section 3.2). Details of the EAG's full appraisal of the company SLR are provided in Appendix 9.1.

ERG conclusions on the methods of review

The company SLR appropriately identified all relevant RCTs. However, the way in which the GEMFIN registry study that informed the ITC was identified and selected is unclear, and no systematic search was conducted for other relevant observational studies.

3.2 Included studies

The three RCTs included in the CS are:

- RESPONSE: [22] a randomised comparison of ruxolitinib against BAT among patients with PV resistant or intolerant to hydroxycarbamide who had splenomegaly. Crossover from BAT to ruxolitinib occurred from week 32.
- RESPONSE-2:[23] a randomised comparison of ruxolitinib against BAT among patients with PV resistant or intolerant to hydroxycarbamide without palpable splenomegaly. Crossover from BAT to ruxolitinib occurred from week 28.
- MAJIC-PV[16] a randomised comparison of ruxolitinib against BAT among "high risk" patients with PV resistant or intolerant to hydroxycarbamide either with or without splenomegaly. Crossover was only permitted to the BAT arm (Table 5 below).

Primary clinical effectiveness analyses were conducted at weeks 32, 28 and 52 in the RESPONSE, RESPONSE-2 and MAJIC-PV trials respectively. In MAJIC-PV overall survival was reported up to 5 years after randomisation. Due to substantial crossover in RESPONSE and RESPONSE 2, long-term outcomes for these trials were reported primarily for the ruxolitinib arm excluding crossovers, up to 5 years. Extensive information on RESPONSE and RESPONSE-2 is available in the CS and in a series of clinical study reports (CSRs) provided for each main assessment timepoint in each trial (except the week 32 CSR which was not provided by the company). In contrast, relatively limited information on the MAJIC-PV trial is available, provided in the CS and an unpublished manuscript.[16]

3.2.1 Characteristics of the included studies

Details of the RCTs are reported for RESPONSE and RESPONSE-2 in CS section B.2.3.1, and for MAJIC-PV in CS section B.2.11.1, with further methodological details, including outcomes for all the trials in CS Appendix Table 11, CONSORT flow diagrams in CS Appendix D.2, and eligibility criteria in CS Appendix M.1. The main trial characteristics are summarised in Table 5 below.

Table 5 Summary characteristics of the included RCTs

Study characteristic	RESPONSE[22]	RESPONSE-2[23]	MAJIC-PV[16]
Funding	Company-sponsored	Company-sponsored	Investigator-led; funded by Leukaemia & Lymphoma Research (UK)
Study design	Open label phase 3 RCT: ruxolitinib vs BAT	Open label phase 3 RCT: ruxolitinib vs BAT	Open label phase 2 RCT: ruxolitinib vs BAT
Country	International, multi-centre 3 UK sites,[24] unknown number of UK patients	International, multi-centre No UK sites	UK-wide, multi-centre 38 UK sites
Population	Patients with polycythaemia vera R/I to HC [a] with splenomegaly	Patients with polycythaemia vera R/I to HC [a] without palpable splenomegaly	Patients with high-risk [b] polycythaemia vera R/I to HC [a] (with or without splenomegaly)
Randomisation	1:1; stratified according to resistance versus intolerance to HC	1:1; stratified according to resistance versus intolerance to HC	1:1; stratified according to gender
Number of participants	Ruxolitinib arm: n=110 BAT arm: n=112	Ruxolitinib arm: n=74 BAT arm: n=75	Ruxolitinib arm: n=93 BAT arm: n=87
Crossover	BAT arm only: patients failing to meet the primary outcome at week 32 were eligible to crossover to receive ruxolitinib	BAT arm only: patients failing to meet the primary outcome at week 28 were eligible to crossover to receive ruxolitinib	No crossover to the ruxolitinib arm was allowed. Ruxolitinib arm: if no response was observed at year 1 (primary outcome) patients changed to receive BAT
Duration	2010-2018; study is complete; data cut-off represent all patients who completed week 256 or discontinued according to protocol	2014-2020; study is complete; data cut-off represent all patients who completed week 260 or discontinued according to protocol	2012-2022; study is complete; data represent all 5 years of follow-up

BAT: best available therapy; HC: hydroxycarbamide; RCT: randomised controlled trial: R/I: resistant or intolerant; UK: United Kingdom.
[a] R/I to HC defined according to ELN consensus criteria,[10] described above in section 2.2.1.

[b] High-risk defined according to trial protocol, described below in section 3.2.1.

The company trials RESPONSE and RESPONSE-2 are open label RCTs providing evidence for the indicated population split across two trials: one for patients with splenomegaly and one for patients without splenomegaly. However, crossover to the ruxolitinib arm was introduced early, after 32 weeks in RESPONSE and after 28 weeks in RESPONSE-2, which confounds longer-term results after the primary outcome analyses. Therefore, evidence from the MAJIC-PV trial, also an open label RCT, is used to inform hazard ratios for overall survival, overall survival in the BAT population, and several subgroup analyses. Data used in the economic model are outlined in CS section B.3.3 Table 21 and in section 3.2.4 of this report.

The RESPONSE and RESPONSE-2 trials followed the criteria for resistance/intolerance outlined in Table 3 above, with a minor exception relating to hydroxycarbamide dose (explained in clarification response A1). MAJIC-PV followed different "modified criteria" for resistance/intolerance (not separated) which are clearly listed in Table S1 of the trial manuscript[16] but lack an explanation for their source or selection. The MAJIC-PV criteria for resistance/intolerance appear to be stricter than the current (2022) guideline criteria reported in CS Table 4. However, as noted in section 2.3.1 above, definitions of hydroxycarbamide resistance/intolerance are not standardised in clinical trials or clinical practice.

The population in the MAJIC-PV trial is a broadly defined high-risk population compared to high-risk as defined in the BSH guidelines ($\geq$65 and/or prior thrombosis – as outlined in CS Figure 3 of the treatment pathway)[4]. In MAJIC-PV the age threshold is lowered to $\geq$60 and additional criteria can also indicate high-risk including significant or symptomatic splenomegaly, platelet count >1000 x 10^9/L, diabetes or hypertension requiring pharmacological therapy for >six months.[16] It is not obvious from the trials' baseline characteristics (Appendix 9.2 of this report) that the MAJIC-PV population is higher-risk than those included in the RESPONSE and RESPONSE-2 trials, as there is overlap of median age, % with prior thrombosis, median platelet counts and other characteristics between trials. However, the mortality rate was substantially higher in MAJIC-PV than the other trials (section 3.2.8 below), which is consistent with the population being at higher risk.

As MAJIC-PV includes patients with and without splenomegaly it covers more of the population in the licensed indication than either of the RESPONSE or RESPONSE-2 trials individually. Additionally, the MAJIC-PV trial contributes a wholly UK population, and with more stringent outcomes (outcomes assessment section 3.2.4), that is relevant to NHS

clinical practice compared to the company trials where only the RESPONSE trial has three UK sites and an unknown number of UK participants. CS section B.2.11.1 argues that the MAJIC-PV trial population is anticipated to represent the majority of patients with PV who are resistant to or intolerant of hydroxycarbamide which the EAG and our clinical experts agree is reasonable.

Limitations

The three included RCTs are limited by being open label (discussed in the risk of bias section of this report, section 3.2.3). The RESPONSE and RESPONSE-2 trials are limited by early crossover, however the MAJIC-PV trial should provide sufficient unconfounded evidence for longer-term outcomes. There is limited data available for the MAJIC-PV trial as it has only recently completed. There is no clinical study report or statistical analysis plan available for verification of study details or results in MAJIC-PV (clarification response A5), and individual level patient data could not be made available to the company because it was an investigator-led trial.

3.2.2 Patients' baseline characteristics in the included RCTs

Patients' baseline characteristics for RESPONSE and RESPONSE-2 are reported together in CS Table 7, and for MAJIC-PV in CS Appendix M.2.1. The EAG have combined key patient baseline characteristics from all three trials in Appendix 9.2 of this report.

Patient characteristics are similar for the RESPONSE and RESPONSE-2 trials, with the exception that participants in RESPONSE-2 did not have splenomegaly according to the trial eligibility criteria.

MAJIC-PV participants are slightly older on average than those in the company trials, but the age range is the same. The proportion of males, ECOG performance status, and percentage haematocrit (HCT) level, are similar. The MAJIC-PV BAT arm had more participants who had a prior thromboembolic event than in the company trials although the proportion of prior thromboembolic events in the ruxolitinib arm is similar to the company trials. Some characteristics in the MAJIC-PV trial are reported differently to the way in which they are reported in the two company trials, such as for white blood cell and platelet counts, JAK2 mutation status, including an extra category for patients who are both resistant *and* intolerant, and spleen size is measured differently, which makes it difficult to compare them with the characteristics in the company trials.

The EAG's clinical experts agreed that the patients' baseline characteristics in all the included trials are generally reflective of patients with PV who are resistant to or intolerant of hydroxycarbamide in the UK. However, the experts noted the following exceptions:

- The median age in MAJIC-PV is slightly higher than in the RESPONSE and RESPONSE-2 trials and is probably more reflective of that seen in clinical practice, although there is heterogeneity both in the trials and in practice.
- One clinical expert expected 15- 20% of patients would have had a prior PV-related thromboembolic event whereas the frequencies in the trials were higher than this (Appendix 9.2). There is also an imbalance within the MAJIC-PV trial for one of the indicators of high-risk for PV (proportion of patients who had a prior thromboembolic event) where the BAT arm is more at risk than the ruxolitinib arm.

EAG conclusions on the included RCTs

All relevant RCTs (n=3) are included in the CS, with each containing up to five years of data from relevant populations, and all are complete. The trials reflect different subgroups of the licensed indication (patients with or without splenomegaly, or a combination). The MAJIC-PV trial is most likely to reflect UK clinical practice and is not confounded by crossover to the ruxolitinib arm, although the data available from the trial are limited.

3.2.3 Risk of bias assessment

Company and EAG risk of bias assessments for the RESPONSE, RESPONSE-2 and MAJIC-PV trials are shown in Appendix 9.3.

All three trials were judged by both the company and EAG to be at high risk of one or more types of bias.

Patient care, recording of outcomes, especially patient reported outcomes which involve subjective judgements, and analysis of outcomes could have been influenced by patients' and investigators' knowledge of the treatment allocation groups, due to the open-label designs of the trials. Additionally, some HRQoL outcomes including the MPN-SAF TSS were reported without any indication of sample sizes and variances. Analyses of HRQoL outcomes excluded missing data but did not specify the amount of missing data and/or reasons for data being missing.

In MAJIC-PV the randomisation process is unclear and the open-label trial design may have allowed patients to circumvent randomised therapy (Figure S2 in the draft trial manuscript[16] shows that some patients “did not want to be in the BAT arm” after randomisation). In all trials there appears to have been selective reporting of HRQoL outcomes (including protocol-specified EQ-5D results not being reported for MAJIC-PV). For further details see Appendix 9.3.

After weeks 32 and 28 respectively, outcomes in the RESPONSE and RESPONSE-2 trials would be confounded by crossover if analysed according to the originally randomised ruxolitinib and BAT groups. This confounding is acknowledged by the company: following crossover, the trial results are generally reported in the CS as single cohorts (the originally-randomised ruxolitinib arm, and the crossover cohort), rather than parallel randomised arms, which is appropriate. The comparative evidence for ruxolitinib versus BAT is limited to 32 and 28 weeks respectively in these trials.

Longer-term comparative evidence is available from the MAJIC-PV trial (52 weeks) which was not subject to crossover from BAT to ruxolitinib, although crossover from ruxolitinib to BAT was permitted for patients who did not achieve a complete or partial response of the primary outcome after 1 year. However, crossovers are not reported transparently: (i) The timing of crossovers from ruxolitinib to BAT is not reported (the EAG assume all occurred after 1 year as per the trial protocol, but reasons for crossover in Figure S2 of the draft manuscript included non-compliance, and hydroxycarbamide resistance and toxicity, which would seem unlikely to obey a 1-year assessment timescale. (ii) The draft trial manuscript states that 10 patients “received ruxolitinib on the BAT arm”, two of whom received ruxolitinib within one year of randomisation (Table S2 of the draft manuscript). The CS and draft trial manuscript do not discuss the implications of the crossovers to the BAT arm or the receipt of ruxolitinib on the BAT arm. It is unclear whether the patients in question would have had a better or worse prognosis than the other patients in each arm and hence the risk of bias associated with these two aspects of participant flow is unclear. The draft trial manuscript[16] states that supporting analyses were performed censoring at the time the BAT patients began ruxolitinib and these analyses did not affect the conclusions from the modified ITT analysis. However, results of these analyses are not reported.

A consequence of all three trials being at high risk of bias is that uncertainty around the outcomes is not fully captured in the variance measures such as 95% confidence intervals, where reported.

EAG conclusions on risk of bias assessment

Overall, the EAG consider the trials to be at high risk of bias due to the open-label nature of all three trials, potential imbalances between groups in the RESPONSE and RESPONSE-2 trials after crossover at 32 and 28 weeks respectively, selective reporting of HRQoL outcomes, and the handling of missing data for HRQoL outcomes in all trials. For MAJIC-PV there is additionally a lack of clarity around the randomisation process, there are some differences in patient characteristics between the treatment arms, and the implications of crossovers from ruxolitinib to BAT, and of receipt of ruxolitinib by some patients in the BAT arm, are not fully clear.

3.2.4 Outcomes assessment

A large number of outcomes was assessed in the included trials (listed in CS Appendix Table 11), and these are reported in various degrees of detail in the CS, CS Appendices, trial publications and, for the RESPONSE and RESPONSE-2 trials, also in several CSRs provided by the company for different assessment timepoints. We have prioritised those outcomes relevant to the NICE scope and decision problem as summarised in Table 6. The outcomes are briefly explained in the sections below.

Table 6 Summary of the outcomes presented in this report

Outcome type	Summary	Where results reported
Primary trial outcomes (see section 3.2.4.1 below)	RESPONSE trial: HCT control & spleen size reduction (composite outcome) at week 32	Section 3.2.6.1
	RESPONSE-2 trial: HCT control (assessed as absence of phlebotomy ineligibility) at week 28	Section 3.2.6.2
	MAJIC-PV trial: Complete haematological remission (ELN criteria) (composite outcome) at 1 year	Section 3.2.6.3
Key secondary trial outcomes (see section 3.2.4.2 below)	Two “key” secondary outcomes were specified by the company: complete haematological remission in RESPONSE and RESPONSE-2; and durability of the primary outcome of RESPONSE beyond week 32	Section 3.2.6.4
Individual components of the primary outcomes	HCT level	Section **Error! Reference source not found.**
	Phlebotomy ineligibility	Section 3.2.6.6
	Spleen size	Section 3.2.6.7

Survival outcomes	Overall survival is a key outcome for the economic analysis (other survival outcomes are also presented where reported)	Section 3.2.6.8
HRQoL outcomes	Numerous measures are reported in the trials; we have prioritised the EQ-5D, MPN-SAF, EORTC QLQ-C30 and PSIS as explained in section 3.2.4.3 below	Section 3.2.6.9
Safety outcomes	Safety outcomes specified in the decision problem and identified as important by the EAG's clinical experts are presented where reported (section 3.2.4.4 below)	Section 3.2.8
ELN: EuropeanLeukemiaNet; HCT: haematocrit. Abbreviations for HRQoL instruments are explained in **Error! Reference source not found.** below.		

3.2.4.1 Primary efficacy outcomes

The primary efficacy outcomes do not inform the economic model but are important to demonstrate clinical efficacy.

HCT control. This is a key target of therapy for PV. HCT control can be measured directly as the haematocrit per volume of blood (target <45%) or indirectly via measures of phlebotomy, such as phlebotomy ineligibility (or absence of phlebotomy eligibility) which are indicative of adequate HCT control. The primary outcomes of the trials either assessed HCT control alone (RESPONSE-2) or included HCT control as a part of broader composite outcomes (RESPONSE, MAJIC-PV). HCT control was also included as a separate secondary outcome in RESPONSE-2 and MAJIC-PV.

The primary outcome of RESPONSE-2 was the proportion of patients achieving HCT control at 28 weeks, measured (according to ELN criteria) as absence of phlebotomy eligibility, where phlebotomy eligibility is defined as HCT of >45% that was at least three percentage points higher than baseline, or an HCT of >48%, whichever was lower.

HCT control and spleen size reduction. This was the composite primary outcome of RESPONSE, assessed at 32 weeks and defined as the proportion of patients achieving HCT control according to modified ELN response criteria (as above for RESPONSE-2) and a ≥35% reduction in spleen size. HCT control and spleen size were also reported as separate secondary outcomes. The EAG's clinical experts noted that assessment of spleen volume (i.e. using imaging techniques rather than palpation) is not very practical and not always assessed in practice.

Complete haematological remission (CHR) according to ELN criteria. This was the composite primary outcome of MAJIC-PV, assessed at one year and defined as the proportion of patients achieving all of the following: HCT <45% without phlebotomy for 3 months; platelets ≤400 × 10^9/L; WBC count ≤10 × 10^9/L, and normal spleen size. It requires fulfilment of all the ELN criteria for complete clinico-haematological response (CLHR) except for resolution of disease-related symptoms[25] and is therefore the most stringent primary outcome reported across the trials. CHR is clinically meaningful to report but it is not used in the economic model.

There is little evidence that stringent achievement of the ELN criteria contributes to improved outcomes apart from the HCT target,[4 26] and one of the EAG's clinical experts said that absence of phlebotomy, by aiming to maintain HCT levels below 45%, is the most critical outcome. Therefore, although the RESPONSE and RESPONSE-2 trials use less stringent combinations of criteria than MAJIC-PV, each primary outcome fulfils the most important aspect of the minimum reported criteria for response, i.e. HCT control.

3.2.4.2 Secondary efficacy outcomes

Complete haematological remission (CHR) is another composite outcome, considered a key secondary outcome in the RESPONSE and RESPONSE-2 trials. It comprises the modified ELN HCT control criteria, platelet counts and WBC counts. NB the definition of CHR in the RESPONSE and RESPONSE-2 trials differs from the CHR definition for the primary outcome in the MAJIC-PV trial mentioned above (which uses original ELN criteria for HCT control and includes spleen size).

The NICE scope indicates that WBC and platelet counts should be considered for reporting separately. These are included as haematological events in CS Appendix F and are taken into account in the summary of safety (section 3.2.8).

Survival outcomes. Overall survival at 5 years, reported in all three trials, is a secondary outcome informing the economic analysis. Transformation-free survival was also reported in RESPONSE and RESPONSE-2. Other survival outcomes, including progression-free survival and event-free survival, were reported for MAJIC-PV, but as hazard ratios for the ruxolitinib comparison rather than median point estimates.

3.2.4.3 HRQoL outcomes

The wide range of HRQoL measures used in the trials is summarised in Table 7 below. Results are reported in section 3.2.6.9 of this report for those measures highlighted in bold: EQ-5D (from RESPONSE-2), MPN-SAF (from all trials), EORTC-QLQ-C30 (from RESPONSE) and PSIS (from RESPONSE and RESPONSE-2). These HRQoL measures have been prioritised by the EAG as they inform the economic analysis and/or were considered clinically relevant by the EAG's experts. Full names of these instruments are given in Table 7 below.

EQ-5D data from RESPONSE-2 are used in a scenario analysis in the economic model (discussed further in section 4.2.7.2 below).

MPN-SAF and **EORTC QLQ-C30** results from RESPONSE are used in the economic model base case (see section 4.2.7.2 below), mapped to MF-SAF using assumptions validated by clinical experts advising the company, to form MF-8D utility values (a preference-based measure for myelofibrosis) (CS section B.3.4).

MPN-SAF is a myeloproliferative disease-specific instrument which has three versions reported in the trials (Table 7): MPN-SAF, MPN-SAF TSS (total symptom score) and MPN-10 (10 item version). These instruments have all been validated for mixed populations with myeloproliferative diseases that include PV.[27 28] The EAG's clinical experts confirmed that the MPN-10 is the version most used in clinical practice, and it includes dimensions for fatigue and itching. All trials measured the proportion of patients achieving $\geq$50% reduction in total symptom score which the EAG's clinical experts confirmed is a clinically meaningful change.

PSIS: This symptom-specific instrument assesses itching which is a bothersome symptom for many patients with PV. PSIS does not inform the economic analysis. The EAG have reported this outcome alongside the other HRQoL instruments to illustrate the effect of ruxolitinib at controlling PV symptoms. However, the company do not explain whether the PSIS has been validated or what the minimum clinically important change is for this instrument.

Table 7 HRQoL outcomes for the RESPONSE, RESPONSE-2, and MAJIC-PV trials

Source of PROs	RESPONSE[29]	RESPONSE-2[30]	MAJIC-PV[31]
PROs reported in the CS	**MPN-SAF** [a] at Week 32 **EORTC QLQ-C30** [a] at Week 32 and Week 80 and Week 256 PSIS at Week 32 and Week 256 PGIC at Week 4 and Week 32	Change from baseline to Week 28 for **MPN-SAF TSS**, **EQ-5D-5L**, PSIS and PGIC	**MPN-SAF TSS** over 5 years
PROs specified in the protocol	As above, plus MPN-PAF (RESPONSE Protocol section 6.2.4.1)	As above, plus WPAI (RESPONSE-2 Protocol section 10.5.5)	MPN-SAF, MDASI and EQ-5D (MAJIC Protocol section 8)
PROs listed in CS Appendix Table 11	As above for 'PROs reported in the CS', plus ECOG score.	As above for 'PROs reported in the CS', plus WPAI.	As above for 'PRO specified in the protocol', with different terminology: MPN10, MDASI and EQ-5D

Sources: CS section B.2.7; CS section B.2.11.2; CS Appendix Table 11; RESPONSE protocol; RESPONSE-2 protocol; MAJIC protocol.
ECOG: Eastern Cooperative Oncology Group Performance Status Scale; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; MDASI: MD Anderson Symptom Inventory; MPN-PAF: Myeloproliferative Neoplasm Pruritus Assessment Form; MPN-SAF: Myeloproliferative Neoplasm Symptom Assessment Form; MPN-SAF TSS: Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score (abridged MPN-SAF with 11 factors); MPN-10: abridged MPN-SAF TSS with 10 factors; PGIC: Patient Global Impression of Change; PRO: patient reported outcome; PSIS: Pruritis Symptom Impact Scale; WPAI: Work Productivity And Impairment.
[a] Three dimensions from EORTC QLQ-C30 and five dimensions from MPN-SAF (mapped to MF-SAF) were combined to form MF-8D utility values; MF-8D was not measured in the trials.

As noted in the risk of bias section (section 3.2.3), there appears to be selective reporting among the HRQoL outcomes:

- There are several HRQoL outcomes specified in the trial protocols for which results are not reported in the CS, Appendices, or trial publications (MPN-PAF and WPAI in RESPONSE and RESPONSE-2, and EQ-5D, MDASI and MPN-10 in MAJIC-PV) (Table 7). This might reflect selective reporting, particularly the lack of EQ-5D results for MAJIC-PV (though the remaining outcomes were considered less important by the EAG's clinical experts).
- It is unclear which MPN-SAF tool the MAJIC-PV trial used or if the terminology (MPN-SAF/MPN-SAF TSS/MPN-10) has been used interchangeably in MAJIC-PV.

3.2.4.4 **Safety outcomes**

The range of adverse events reported by the company (CS sections B.2.10 and B.2.11.3, and CS Appendix F) is appropriate. Adverse events of special interest are reported and relevant to PV (thromboembolic events, second malignancies, non-melanoma skin cancer, transformation to MF, and transformation to AML) (CS Table 16). Transformation to MF and transformation AML are outcomes in the NICE scope and are also reported as efficacy outcomes in CS sections B.2.7.1 and B.2.7.2 as transformation-free survival. The EAG's clinical experts agreed that malignancies, particularly non-melanoma skin cancer (NMSC) are important. One expert commented that there may also be risk of lymphoma from ruxolitinib treatment. Another expert emphasised that infections, particularly herpes zoster reactivation, are important due to the immunosuppressive characteristics of ruxolitinib.

The trials use different frequency thresholds making it difficult to compare the rates between trials: RESPONSE reports adverse events occurring at a rate of ≥5 per 100 patient-years; RESPONSE-2 reports adverse events occurring in ≥3% of patients adjusted for patient-year exposure; and MAJIC-PV reports descriptive proportional statistics (n, %) for adverse events occurring in ≥10% of patients. The trials report the number of adverse events occurring at different CTCAE (Common Terminology Criteria for Adverse Events)) grades differently: MAJIC-PV reports adverse events (except for infections and malignancies) for all grades for the ruxolitinib and BAT arms combined, and Grades 3, 4 and 5 are reported separately, whereas the RESPONSE and RESPONSE-2 trials report adverse events for any grade for each arm, and Grades 3-4 are combined.

EAG conclusions on outcomes assessment

All reported outcomes are relevant to the disease, particularly HCT control for clinical effectiveness and the reporting of relevant adverse events of specific interest. Some outcomes are reported inconsistently across the trials, e.g. different complete haematological response outcomes, and thresholds for reporting of adverse events differed between trials. A wide range of HRQoL measures were used but reporting appears to be selective.

3.2.5 Statistical methods of the included studies

The CS reports statistical methods only for the primary outcomes. A summary of the EAG's assessment of statistical methods in the trials is provided in Table 8, with information for

secondary and other outcomes sourced from the trial protocols, CSRs and publications. The full assessment is provided in Appendix 9.4.

Table 8 Statistical methods of the RESPONSE, RESPONSE-2 and MAJIC-PV trials

	RESPONSE	RESPONSE-2	MAJIC-PV
Analysis populations	Appropriate for the primary and two key secondary outcomes (full analysis set), and safety outcomes (safety set). Unclear for the remaining secondary outcomes and HRQoL measures.	Appropriate for the primary and key secondary outcomes (full analysis set), and safety outcomes (safety set). Unclear for the remaining secondary outcomes and HRQoL measures.	Limited details of the analysis populations are reported; analysis populations for HRQoL outcomes are unclear. Potential for bias due to unaccounted for missing data (see Appendix 9.3).
Sample size and power calculations	Trial appears to be adequately powered for the primary outcome and probably also the two key secondary outcomes. Adequacy of the sample size for detecting treatment effects in the remaining secondary outcomes is uncertain.	Trial appears to be adequately powered for the primary outcome and key secondary outcome. Adequacy of the sample size for detecting treatment effects in the remaining secondary outcomes is uncertain.	Trial appears to be adequately powered for the primary outcome. Adequacy of the sample size for detecting treatment effects in the remaining secondary outcomes is uncertain.
Methods to account for multiplicity	The type I error control procedure is appropriate but only three outcomes are included. The likelihood of type I error in testing the remaining secondary outcomes is uncertain.	The type I error control procedure is appropriate but only two outcomes are included. The likelihood of type I error in testing the remaining secondary outcomes is uncertain.	No information available. The likelihood of nonsignificant treatment effects being declared significant is uncertain. Reliance on the statistical test results alone for inference is therefore inadvisable.
Analysis of outcomes	The statistical methods appear generally appropriate. The CS does not state whether the analyses were checked or validated.	The statistical methods appear generally appropriate. The CS does not state whether the analyses were checked or validated.	The statistical methods appear generally appropriate. NB alpha =0.1 and 80% confidence intervals are applied for the primary outcome (stated in the trial protocol) giving a relatively high chance of nonsignificant findings being declared significant. No information on whether analyses were checked.
Handling of missing data	Appropriate for primary and secondary	Appropriate for primary and secondary	Overall missing data were not accounted for,

	outcomes. Missing data were not accounted for in analyses of HRQoL and other exploratory outcomes. Number and reasons for missing data not fully reported.	outcomes. Missing data were not accounted for in analyses of HRQoL and other exploratory outcomes. Number and reasons for missing data not fully reported.	and the amount of missing data and reasons for data being missing were not reported.
Subgroup analyses	The pre-specified subgroup analysis method is appropriate. A post-hoc subgroup analysis of patients who received interferon-alfa, pooled from RESPONSE and RESPONSE-2, had small sample sizes ranging from 13 to 30 participants.	The pre-specified subgroup analysis method is appropriate. A post-hoc subgroup analysis of patients who received interferon-alfa, pooled from RESPONSE and RESPONSE-2, had small sample sizes ranging from 13 to 30 participants.	No subgroup analysis method or results are reported.

EAG conclusions on study statistical methods

The primary and key secondary outcomes of RESPONSE and RESPONSE-2 were adequately powered and accounted for multiple testing; however, remaining outcomes were mainly summarised descriptively and could be subject to type I errors. Missing data and multiple testing were not adequately accounted for in the MAJIC-PV trial so the results should be interpreted with caution. Where reported (RESPONSE and RESPONSE-2), subgroup analyses were appropriate but in some cases subject to small sample sizes.

3.2.6 Efficacy results of the intervention studies

As noted in section 3.2.4, many outcomes were assessed in the included trials. We have prioritised the following outcomes in this report, as explained above (Table 6).

3.2.6.1 Primary outcome in RESPONSE (composite of phlebotomy ineligibility and spleen volume reduction)

HCT control as defined by phlebotomy ineligibility and reduction of ≥35% in spleen volume from baseline at week 32 was the primary outcome in the RESPONSE trial and is referred to as the “primary response”. The odds of achieving the primary response at week 32 statistically favoured ruxolitinib over BAT (odds ratio >1.0). However, the majority of patients did not achieve a primary response (Table 9). Due to crossover, results after week 32 are reported for the randomised ruxolitinib arm of the trial, i.e. a single non-comparative cohort.

Of those originally randomised to ruxolitinib who achieved a primary response at week 32, nearly all had maintained the key secondary outcome of response at week 48. The estimated probability of maintaining the primary response from week 32 to week 256 in the ruxolitinib arm (a secondary outcome) was 74% but with a relatively wide 95% confidence interval (51% to 88%).

Table 9 Primary outcome in the RESPONSE trial

Outcome	Ruxolitinib	BAT	Difference	Source
Primary response at week 32 (primary outcome)	23/110; (20.9%) [a] 25/110 (22.7%) [b]	1/112 (0.9%)	20.02 (95% CI 12.22 to 27.82) p<0.001 OR 28.6 (95% CI 4.5−1206)	CS section B.2.7.1 and Table 11-5 in week 48 CSR
Secondary outcomes related to the primary outcome				
Durable primary response (response at week 32 maintained at week 48)	21/110 (19.1%)	1/112 (0.9%)	18.2 %−points; [c] p<0.001	CS section B.2.7.1
Probability of maintaining primary response for ≥1 year	94%	NA	NA	CS Figure 8
Probability of maintaining primary response for ≥80 weeks	92% (ITT) [d] 89% [d]	NA	NA	CS Appendix M.3.1
KM estimated probability of maintaining primary response at 208 weeks	73% (95% CI 49%−87%)	NA	NA	CS Appendix M.3.1
KM estimated probability of maintaining primary response from week 32 for 224 weeks	74% (95% CI 51% to 88%)	NA	NA	CS section B.2.7.1
Median duration of primary response	Not reached	Not reached	NA	CS section B.2.7.1
ITT: intention to treat population; KM: Kaplan-Meier; NA: Not applicable (due to patient crossover); OR: odds ratio. [a] Initial results reported by Vannucchi et al. 2015; [b] updated results from week 80 analysis reported in CS section B.2.7.1 which identified 2 further week 32 responders; [c] calculated by reviewer; [d] ITT population includes crossovers; 89% refers to patients randomised to ruxolitinib.				

3.2.6.2 Primary outcome in RESPONSE-2 (absence of phlebotomy eligibility)

HCT control as defined by phlebotomy ineligibility at week 28 was the primary outcome of the RESPONSE-2 trial. The trial did not include patients with palpable splenomegaly and so the primary outcome for RESPONSE-2 does not include spleen size. The odds of achieving HCT control at week 28 statistically favoured ruxolitinib over BAT (odds ratio >1.0). In the ruxolitinib arm 62% of patients achieved the primary outcome, compared to 19% in the BAT

arm. Due to crossover, results after week 28 are reported for the randomised ruxolitinib arm of the trial, i.e. a single non-comparative cohort (secondary outcomes). Among the patients randomised to the ruxolitinib arm, 21.6% had achieved durable HCT control to week 260 (Table 10).

HCT control as defined by the absence of phlebotomy was also assessed in the RESPONSE trial, as a secondary outcome, and shows a similar picture to that of RESPONSE-2: Of those who received ruxolitinib in RESPONSE, 60.0% achieved HCT control after 24 weeks' treatment (at the week 32 analysis) compared to 19.6% in the BAT arm.[22] The proportion in the ruxolitinib arm with durable HCT control was not reported for the RESPONSE trial, but the estimated probability of maintaining HCT control from week 32 to week 256 was 73% (95% CI 60% to 83%).[24] The median duration of HCT control was not reached in either trial (CS Appendix M.3.1 and M.3.2).

Table 10 Primary outcome in the RESPONSE-2 trial

Outcome	Ruxolitinib (N=74)	BAT (N=75)	Difference	Source
HCT control at week 28 (primary outcome)	46/74 (62%)	14/75 (19%)	OR 7.28 (95% CI 3.43 to 15.45); p<0.0001	CS section B.2.7.2
Secondary outcomes related to the primary outcome				
Proportion maintaining HCT control from week 28 to 52	[redacted]	[redacted]	OR [redacted] (95% CI [redacted]) P<0.0001	Table 11-2 in week 80 CSR
Proportion maintaining HCT control from week 28 to 80	35/74 (47.3%)	2/75 (2.7%)	44.6 %−points [a] OR [redacted] (95% CI [redacted])	CS Appendix M.3.2 Week 80 CSR
Durable HCT control at week 156	30/74 (40.5%) [b]	NA	NA	CS Appendix M.3.2
Durable HCT control at 5 years (week 260)	16/74 (21.6%)	NA	NA	CS section B.2.7.2
NA: not applicable; OR: odds ratio. [a] calculated by reviewer; [b] patients originally randomised to ruxolitinib (i.e. excluding crossovers)				

3.2.6.3 Primary outcome in MAJIC-PV (composite of HCT control, WBC, platelet, and spleen volume thresholds by ELN criteria)

The primary outcome in MAJIC-PV, referred to as "complete haematological remission" according to ELN criteria[25] is a composite of HCT control [comprising HCT <45% with phlebotomy ineligibility], WBC counts, platelet counts, and spleen volume thresholds. The odds of achieving complete haematological remission at 1 year statistically favoured ruxolitinib over BAT (odds ratio >1.0), although fewer than half the patients receiving

ruxolitinib achieved a complete remission (Table). Nearly all of those who did not achieve a compete haematological remission at year 1 achieved a partial haematological remission, giving high overall response rates in both the ruxolitinib and BAT groups.

Table 11 Primary outcome in the MAJIC-PV trial (complete haematological remission)

Outcome	Ruxolitinib (N=93)	BAT (N=87)	Difference	Source
Proportion with complete haematological remission (ELN criteria) in year 1	40/93 (43%)	23/87 (26%)	Adjusted [a] OR 2.12 (90% CI 1.25 to 3.60); p=0.02	CS section B.2.11.2 and unpublished trial manuscript[16]
Secondary outcomes related to the primary outcome				
Proportion with partial haematological remission (ELN criteria) in year 1	50/93 (54%)	58/87 (67%)	−13 %−points [b]	
Overall response rate in year 1	97%	93%	4 %−points [b]	
OR: odds ratio; ELN: European EukemiaNet. [a] adjusted for gender. [b] calculated by reviewer.				

3.2.6.4 Key secondary outcomes

Complete haematological remission (composite of HCT control assessed as phlebotomy ineligibility; together with WBC and platelet count thresholds) was specified as a key secondary outcome in the RESPONSE and RESPONSE-2 trials. Note that this outcome differs from the complete haematological remission outcome of the MAJIC-PV trial reported above (which used ELN criteria that include a more stringent definition of HCT control [HCT <45% without phlebotomy] and a normal spleen size). In both trials the proportion achieving complete haematological remission statistically favoured the ruxolitinib arm after weeks 28 and 32, but was relatively low, not exceeding 24% (Table 12). Median duration of complete haematological remission was not reached in the RESPONSE trial (CS section B.2.7.1). In RESPONSE-2 the KM estimate of median duration of complete haematological remission from week 28 to week 260 (i.e. 5 years) was 34.0 weeks (95% CI 16 to 78 weeks) (CS section B.2.7.2).

Table 12 Complete haematological remission in the RESPONSE and RESPONSE-2 trials

Outcome	Ruxolitinib	BAT	Difference	Source
Proportion achieving CHR at week 32 in RESPONSE	26 [a] /110 (23.6%)	8 [a] /112 (8.9%)	14.7 %−points [a] p=0.003 [b]	CS section B.2.7.1
Proportion achieving CHR at week 28 in RESPONSE-2	17/74 (23%)	4/75 (5%)	OR 5.58 (95% CI 1.73 to 17.99); p<0.0019	CS section B.2.7.2 Week 28 CSR
CHR: complete haematological remission; OR: odds ratio; [a] calculated by reviewer; [b] Vannucchi et al. 2015[22] report p=0.003, CS reports p=0.0003				

Durability of the primary outcome (HCT control and spleen volume reduction) at week 48 in the ruxolitinib arm was specified as a key secondary outcome in the RESPONSE trial. This is reported alongside the primary outcome in Table 9 above.

3.2.6.5 HCT measurements

HCT control is included as a component of the primary outcomes of all three included RCTs (sections 3.2.6.1 to 3.2.6.3). HCT levels are also reported separately in RESPONSE-2 and in MAJIC-PV.

In RESPONSE-2 the baseline and week 28 HCT levels were below the HCT control threshold of <45% for PV. At week 28 the HCT level had decreased in the ruxolitinib arm and increased in the BAT arm, confirming the cytoreductive action of ruxolitinib (Table 13).

Table 13 HCT levels in the RESPONSE-2 trial

Outcome	Ruxolitinib	BAT	Difference	Source
Baseline HCT, mean (SD)	42.8% (1.5%)	42.7% (1.4%)	0.1 %−points [a]	CS section B.2.7.2
Week 28 HCT, mean (SD)	40.2% (4.1%)	44.9% (3.8%)	−4.7 %−points [a]	
Change in HCT from baseline to week 28, mean (SD)	−2.6% [a]	2.2% [a]	4.8 %−points [a]	
[a] calculated by reviewer				

In MAJIC-PV, HCT levels in the ruxolitinib and BAT arms are shown visually in the supplement to the unpublished manuscript (Figure S4 in Harrison et al.[16]) over 54 weeks. Estimates of mean counts are not reported. Following randomisation, the mean HCT count in the ruxolitinib arm initially decreased and then remained below 0.375 whilst the HCT count in the BAT arm remained approximately constant, around 0.400, through the 54 weeks. These differences were significantly different, indicated by non-overlapping 95% CIs.

3.2.6.6 Phlebotomy rates

The trials reported the proportions of patients who underwent different numbers of phlebotomy procedures, as well at the proportions who had any or no phlebotomies. Here we summarise the proportions who had no phlebotomies as this is an indicator of HCT control.

The proportion of patients who had no phlebotomies in the RESPONSE and RESPONSE-2 trials (before crossover) and MAJIC-PV trial was consistently higher in the ruxolitinib arm of each trial than in the BAT arm (Table 14).

As the data in Table 14 show, 34% to 48% of patients in the BAT arms (prior to crossover) did not require phlebotomy. Overall, ruxolitinib increased the proportion who did not require phlebotomy by 23 to 41 percentage points relative to BAT, depending on the trial and assessment time.

Over the 5-year follow-up period, the proportion without phlebotomies in the ruxolitinib arm (excluding crossovers in RESPONSE and RESPONSE-2) was:

- 83% during weeks 80-256 in RESPONSE (CS Figure 15)
- 69% up to week 260 in RESPONSE-2 (CS Table 11)
- 71% in MAJIC-PV (Table 14 below).

Table 14 Proportion without phlebotomy in the RCTs

Outcome	Ruxolitinib	BAT	Difference	Source
Proportion with no phlebotomies in weeks 8-32 in RESPONSE [a]	80/110 (72.7%) [b]	38/112 (33.9%) [b]	38.8 %−points [b]	CS Figure 9 [c]
Proportion with no phlebotomies up to week 28 in RESPONSE-2	81.1%	40%	41.1 %−points [b]	CS Figure 19 [d]
Proportion with no phlebotomies up to 5 years in MAJIC-PV [e]	66/93 (71%) [b]	42/87 (48%) [b]	23% %−points [b]	Unpublished trial manuscript[16]

[a] patients who did not discontinue randomised therapy prior to week 8; [b] calculated by reviewer; [c] CS Figure 9 reports sample sizes less than the full analysis set, EAG calculations use the full analysis set (i.e. ITT analysis); [d] CS Figure 19 does not report the sample size, so unclear whether this is an ITT analysis; [e] assessment time not reported but EAG assume this was 5 years (since adjacent outcomes in the trial manuscript supplementary appendix were reported for 5 years)

3.2.6.7 Spleen measurements

Spleen size is included as a component of the primary composite outcome of the RESPONSE trial (section 3.2.6.1 above). Spleen measurements are also reported separately for RESPONSE, and some limited information on spleen size is also available for RESPONSE-2 (spleen volume measurements are not reported for MAJIC-PV[16]).

In RESPONSE, 40% of patients in the ruxolitinib arm and 0.9% in the BAT arm achieved a ≥35% reduction in spleen volume after 24 weeks of treatment (week 32 analysis) according to CS section 2.7.1, but the trial publication[22] and week 48 CSR report 38.2% in the ruxolitinib arm; the EAG are unclear which is correct. In the ruxolitinib arm, excluding crossovers, the estimated probability of maintaining a ≥35% reduction in spleen volume from week 32 to week 224 was 72% (95% CI 34% to 91%).[24]

In RESPONSE-2, according to the week 260 CSR, nine patients in the ruxolitinib arm had a palpable spleen, with the mean palpable spleen length at week 260 being 0.10 cm. In the BAT arm, nine patients had a palpable spleen but very few patients were assessed (n=5) at week 80, and the mean palpable spleen length was 0 cm (data for the remaining four patients are not reported). These findings suggest splenomegaly during long-term follow up was negligible in RESPONSE-2.

3.2.6.8 Survival outcomes

Survival outcomes reported in CS and trial publications are summarised below. The MAJIC-PV trial manuscript reports that 3-year overall survival did not differ between the trial arms: 87% (95% CI 77% to 93%) for BAT and 88% (95% CI 79% to 93%) for ruxolitinib. Hazard ratios comparing 5-year overall survival for ruxolitinib against BAT are also provided (see below); it is unclear why the 3-year and 5-year outcomes are not reported consistently (CS section B.2.11.2).

Overall survival at 5 years: KM estimates of OS at 5 years are reported for the ruxolitinib arm, excluding crossovers, in the RESPONSE trial (N=110) and RESPONSE-2 trial (N=74), and as a hazard ratio for the comparison of ruxolitinib (N=93) versus BAT (N=87) in MAJIC-PV:

- RESPONSE: 91.9% (95% CI 84.4% to 95.9%) (CS section B.2.7.1); median OS not reached (not reported in the CS, publications or CSRs - stated in the company's Factual Accuracy Check document)
- RESPONSE-2: 96% (95% CI 87% to 99%); median OS not reached (CS section B.2.7.2)
- MAJIC-PV: Median OS not reached;[16] OS hazard ratio, ruxolitinib versus BAT 0.73 (95% CI 0.36 to 1.50; p=0.39 (CS section B.2.11.2).

Transformation-free survival at 5 years: KM estimates of TFS at 5 years for the ruxolitinib arm, excluding crossovers, for the RESPONSE trial (N=110) and RESPONSE-2 trial (N=74) were:

- RESPONSE: [redacted] (95% CI [redacted]) (CS section B.2.7.1)
- RESPONSE-2: 94% (95% CI 85% to 98%) (CS section B.2.7.2).

Other survival outcomes at 5 years: The following hazard ratios based on KM estimates of median survival outcomes for the ruxolitinib arm (N=93) compared against the BAT arm (N=87) are reported for the MAJIC-PV trial in the unpublished trial manuscript:[16]

- Progression-free survival: HR 0.64 (95% CI 0.36 to 1.15; p=0.13
- Event-free survival: HR 0.58 (95% CI 0.35 to 0.94); p=0.03
- Major thrombosis event-free survival: HR 0.56 (95% CI 0.32 to 1.00; p=0.05
- Haemorrhagic event-free survival: HR 0.66 (95% CI 0.34 to 1.28; p=0.22.

3.2.6.9 HRQoL outcomes

The trials reported a range of HRQoL measures (with some evidence of selective reporting) (see Table 7 and section 3.2.4.3 above). However, the EAG's clinical experts commented that many of the HRQoL measures are not used in clinical practice nor widely in trials. Below we have prioritised those HRQoL measures that inform the economic analysis (EQ-5D, EORTC QLQ-C30, MPN-SAF), or are relevant to symptoms specified in the NICE scope (PSIS is an itching-specific instrument whilst MPN-SAF includes itching and fatigue among other symptoms). The EAG's clinical experts commented that the MPN-SAF and its derivatives such as MPN-10 are the HRQoL measures most used in clinical practice.

EQ-5D index score

The EQ-5D is specified as an outcome in the RESPONSE-2 and MAJIC-PV trials (Table 7 above) but is only reported for RESPONSE-2.

The company have presented EQ-5D scores from RESPONSE-2 in their submission (CS Figure 21 and the trial publication[23]) but these are difficult to interpret due to: (i) the scores are reported as percentage classes instead of their original scale; (ii) sample sizes are unclear since the numerators and denominators for the percentages are not provided; (iii) the use of percentages excludes any information on the variance of scores. The EAG have instead sourced the overall EQ-5D-5L scores from the RESPONSE-2 week 260 CSR, summarised in Table 15 below. These data suggest there is little difference in the change from baseline between the ruxolitinib and BAT arms, and within the ruxolitinib arm after crossover occurred.

The company note that a large proportion of patients reported no problems in all five EQ-5D domains at baseline, and they argue that EQ-5D is unsuitable for measuring HRQoL in PV (CS section B.3.4.1) (discussed below in section 4.2.7.2). However, point estimates of EQ-5D scores from RESPONSE-2 were used in a scenario analysis in the company's economic model (CS section B.3.4.1).

Table 15 Changes in EQ-5D-5L health index score in the RESPONSE-2 trial

Mean (SD) change from baseline [a]	Ruxolitinib (N=74)	BAT (N=75)	Difference [b]	Source
Week 28	[redacted]	[redacted]	[redacted]	Table 14.2-2.6 in week 260 CSR
Week 52	[redacted]	[redacted] [c]	[redacted]	
Week 80	[redacted]	NA	NA	
Week 104	[redacted]	NA	NA	
Week 156	[redacted]	NA	NA	
Week 247	[redacted]	NA	NA	
NA: not applicable. [a] Baseline mean varied with each assessment timepoint, presumably because not all patients had baseline measurements at all timepoints [b] calculated by reviewer. [c] patients who did not cross over to ruxolitinib				

MPN-SAF scores

MPN-SAF scores inform the company's economic analysis indirectly, via conversion to MF-8D scores (section 4.2.7.2). All three trials reported changes in MPN-SAF scores, although the reporting format is different for each trial, making comparisons across the trials difficult. This outcome also has some uncertainty relating to missing data.

- **RESPONSE** (CS section 2.7.1): At week 32, the proportion with ≥50% reduction in MPN-SAF total score (a clinically meaningful improvement) was 49% in the ruxolitinib arm (36/74) and 5% in the BAT arm (4/81) The reported sample sizes indicate that the ruxolitinib arm had 36/110 (33%) missing data and the BAT arm had 31/112 (28%) missing data compared to the full analysis set.

- **RESPONSE-2** (CS section B.2.7.2, CS Figure 20 and CS Appendix M.3.2): At week 28, the proportion with ≥50% reduction in MPN-SAF TSS was 45.3% in the ruxolitinib arm and 22.7% in the BAT arm. Sample sizes reported were 64 ruxolitinib patients and 22 BAT patients, meaning that the ruxolitinib arm had 46/110 (42%) missing data and the BAT arm had 90/112 (80%) missing data compared to the full analysis set.

- **MAJIC-PV** (unpublished manuscript[16]): Only the mean difference in the change from baseline in MPN-10* between the ruxolitinib and BAT arms is reported, for a range of timepoints from month 2 to month 60. The difference favoured ruxolitinib over BAT at all timepoints and was statistically significant up to around 24 months but statistical significance should be interpreted cautiously due to the large number of comparisons made (Table S8 in the draft trial manuscript[16]). The mean difference for ruxolitinib versus BAT at 60 months was −3.1 (−9.6 to 3.4); p=0.35. Sample sizes are not reported so the extent of missing data is unclear. (*NB the source table refers to "MPN-10" but the wording in the manuscript implies that this is synonymous with the MPN-SAF).

EORTC QLQ-C30

The EORTC QLQ-C30 measure was utilised only in the RESPONSE trial. Improvements from baseline occurred across all of the six subscales for the ruxolitinib arm, both at week 32 and (excluding crossovers) at week 256, whilst scores worsened slightly for five of the six subscales at week 32 in the BAT arm (Table 16). The threshold for a clinically meaningful change (10 points) was reached for the ruxolitinib arm at week 32; the largest improvement with ruxolitinib and the largest worsening with BAT were both for the Global health status subscale. Sample sizes and variance measures are not reported for this outcome.

Table 16 EORTC QLQ-C30 questionnaire functional and QoL scales in the RESPONSE trial

Mean change in score from baseline	Ruxolitinib		BAT Week 32	Difference at week 32 [a]	Source
Scale	**Week 32**	**Week 256**			
Global health status/QoL	10.86	9.49	−4.82	15.68	Vannucci et al. 2015;[22] Supplementary Figure 8 in Kiladjian et al. 2020[24]
Physical functioning	6.44	7.05	−1.51	7.95	
Role functioning	5.3	2.08	−0.41	5.71	
Emotional functioning	7.92	7.55	1.04	6.88	
Cognitive functioning	4.17	6.08	−3.33	7.50	
Social functioning	7.66	5.73	−0.42	8.08	
[a] calculated by reviewer; minimal clinically important difference is 10 points. NB variance estimates and sample sizes are not reported; results are for patients with both baseline and week 32 / 256 data					

Pruritis Impact Symptom Scale (PSIS)

The company report changes from baseline in PSIS scores for RESPONSE (CS Figure 11) and RESPONSE-2 (trial publication[23]). The severity of PV-related itching, the extent to which the patient was bothered by itching, and the extent to which the itching interfered with daily life were improved to a greater extent in the ruxolitinib arm than the BAT arm at 32 weeks, both for 24 hour and 7-day recall periods, in both trials. However, the sample size and variance estimates for this outcome are not reported in the CS or trial publications. It is also unclear whether this tool has been validated and what the minimum clinically important difference would be.

Overall, there is evidence that ruxolitinib improves patients' symptoms relating to itching, but with some uncertainty around how variable and clinically significant these findings are.

3.2.7 Subgroup analyses

The NICE scope specifies two subgroups: patients with and without splenomegaly. These subgroups are covered by the different trial populations: in the RESPONSE trial all patients had splenomegaly (based on imaging measurements), whilst the RESPONSE-2 trial excluded patients with splenomegaly (based on splenic palpation) (CS Table 6). Note that the MAJIC-PV trial included high-risk PV patients irrespective of splenomegaly and thus provides evidence from a further relevant population reflecting the mix of patients seen in clinical practice.

3.2.7.1 Pre-specified subgroups in the trials

The following subgroup analyses were conducted for the primary outcome in each trial. The subgroup analysis results reported in the CS and trial publications are consistent with those specified in the trial protocols for RESPONSE and RESPONSE-2. For MAJIC-PV the trial protocol specifies exploratory subgroups, but these are not reported in the CS or the trial draft manuscript.[16]

RESPONSE

Pre-specified subgroup comparisons (trial protocol section 9.4.4) were: baseline palpable splenomegaly (<10cm versus ≥10cm below the costal margin), sex (male versus female), age group (≤60 years versus >60 years), hydroxycarbamide intolerance or resistance, region (US versus non-US), race (White or Caucasian versus other) and ethnicity (Hispanic or Latino versus other).

A forest plot showing the odds of achieving the composite primary response outcome at week 32 for each subgroup is provided in CS Figure 24 but is missing odds ratios for one subgroup in each pair so the EAG are unable to interpret this (the week 32 CSR was not provided by the company).

RESPONSE-2

Pre-specified subgroup comparisons (trial protocol section 10.4.4) were: hydroxycarbamide intolerance or resistance, sex (male versus female), age group (≤60 years versus >60 years), risk category (0 risk factors versus 1-2 risk factors including age >60 and/or previous thromboembolism).

CS Figure 25 shows the odds of achieving complete haematological remission at week 28 for each of these subgroups. All odds ratios are greater than 5.0 and have overlapping

confidence intervals, suggesting that the odds of achieving the primary outcome did not differ between subgroups.

3.2.7.2 Post-hoc subgroup analyses in the trials

CS Appendix E reports subgroup analyses of patients who had received prior interferon-alfa, interferon-alfa as BAT, or ruxolitinib after crossover from receiving interferon-alfa as BAT. These subgroup, which are based on data pooled from RESPONSE and RESPONSE-2, have small sample sizes ranging from 13 to 30 patients and therefore their generalisability is uncertain.

3.2.8 Safety results

Adverse events in the RESPONSE and RESPONSE-2 trials are reported in CS sections B.2.10.1 and B.2.10.2 up to weeks 256 and 260 respectively and in Appendix F for earlier data cuts. Adverse events in the MAJIC-PV trial are reported up to 5 years in CS section B.2.11.3, CS Appendix Table 20 and the unpublished trial manuscript.[16]

NB as noted in section 3.2.4.4, in the CS adverse events are not reported consistently in the same format across the trials.

Most frequent adverse events

In RESPONSE and RESPONSE-2 the most frequent adverse events of any grade were seen in the BAT arms, especially for the disease symptom pruritus (BAT 32.6 and 31.9 per 100 patient years respectively; ruxolitinib 7.0 and 3.6 per 100 patient years respectively; crossover 6.1 and 3.4 per 100 patient years respectively). Thrombocytopaenia of any grade also had the highest rate in the BAT arms (BAT 16.3 and 15.0 per 100 patient years respectively; ruxolitinib 4.4 and 1.5 per 100 patient years respectively; crossover 1.2 and 1.5 per 100 patient years respectively). The most frequent adverse event of any grade that occurred more often in the ruxolitinib and crossover groups than in the BAT arms was anaemia (ruxolitinib 8.9 and 8.1 per 100 patient years respectively; crossover 8.8 and 9.2 per 100 patient years respectively).

In the MAJIC-PV trial, the most frequent adverse events were infections, gastrointestinal disorders, and vascular disorders.[16] The CS highlights Grade 3 anaemia which occurred in 7% of ruxolitinib patients compared to 1% of BAT patients.

Most frequent serious adverse events

In the RESPONSE trial the most frequent serious adverse event was pneumonia, but with similar rates across the trial arms (1.2 to 1.8 per exposure adjusted 100 patient years). Several serious adverse events were recorded only in the ruxolitinib and crossover groups but not the BAT group, notably squamous cell carcinoma, basal cell carcinoma, rectal haemorrhage, and herpes zoster infection. Adverse events which were classified as serious adverse events are not reported in the CS for the RESPONSE-2 trial, nor in the trial manuscript for the MAJIC-PV trial.[16]

Infections

In the RESPONSE trial, the total rate of infections per 100 patient years was highest in the BAT arm (BAT 59.8; ruxolitinib 18.9; crossover 19.1). The total rate of infections is not reported for the RESPONSE-2 trial, although individual infections are reported. For both RESPONSE and RESPONSE-2, the herpes zoster infection appears to have only occurred in the ruxolitinib arms and the crossover groups, although the adverse events in CS Tables 15 and 17 are not reported consistently across the trials and infrequent infections might not have been captured due to the reporting thresholds used in the tables (RESPONSE: ≥0.5 per 100 patient years; RESPONSE-2: for ≥3% of patients in any arm).

In the MAJIC-PV trial, infections were more common in the ruxolitinib arm (27 Grade 3/4 events) compared to the BAT arm (12 Grade 3/4 events). The most common infections for ruxolitinib patients were respiratory, genitourinary, and cutaneous herpes zoster. Herpes zoster infections at any Grade occurred in 9 ruxolitinib patients compared to 3 BAT patients. All infections are individually reported in Table S9B of the unpublished trial manuscript.[16]

Malignancies, including transformation to MF or AML

In the RESPONSE trial, second malignancies had a higher exposure-adjusted rate per 100 patient years in the ruxolitinib arm (7.0) and crossover group (4.5) than in the BAT arm (4.1); so too did rates of non-melanoma skin cancer: ruxolitinib arm (5.1), crossover group (2.7) and BAT arm (2.7). Exposure-adjusted rates per 100 patient years of transformation for both MF and AML were also higher in the ruxolitinib arm and crossover group although with slightly lower rates than reported for the malignancies.

In the RESPONSE-2 trial, second malignancies are reported in the CSR but not the CS: the week 260 CSR states that there may be some data overlap with this category.[32] Non-melanoma skin cancer had a slightly higher rate of occurrence in the ruxolitinib arm and crossover group than in the BAT arm, but rates of transformation to MF and AML were

slightly higher in the BAT arm. No patients transformed to AML in the ruxolitinib arm or crossover group.

In the MAJIC-PV trial, certain malignancies were more common in the ruxolitinib arm compared to the BAT arm: squamous cell (skin) carcinoma occurred in 11 versus 0 patients respectively, and transformation to AML occurred in 4 versus 0 patients respectively. Transformation to myelofibrosis was more common in the BAT arm: 5 ruxolitinib patients compared to 10 BAT patients. Further malignancies are fully reported in Table S9B of the unpublished manuscript.[16]

Thromboembolic events

In the RESPONSE and RESPONSE-2 trials, thromboembolic events had the highest rates (exposure-adjusted per 100 patient years) in the BAT arms (8.2 and 3.7 respectively), compared to the ruxolitinib arms (1.2 and 1.5 respectively) and crossover groups (2.7 and 2.9 respectively).

In the MAJIC-PV trial, Table S7 in the unpublished manuscript reports the number, toxicity, and CTCAE grade of minor and major thrombotic events but does not distinguish between the ruxolitinib and BAT arms.[16]

Deaths

One out of a total of six deaths in the RESPONSE trial was suspected to be related to the study drug (gastric adenocarcinoma) and none of the five deaths in the RESPONSE-2 trial were deemed to be related to the study drug. More deaths occurred in the MAJIC-PV population (n=32). The EAG speculate this may be due to the slightly older population and a greater proportion of patients (in the BAT arm) who had had a prior thromboembolic event (Appendix 9.2) indicating high-risk. However, only one death in each treatment arm in the MAJIC-PV trial was considered related to the study drug and none of the deaths were infection-related.

EAG conclusions on safety results

Adverse events are difficult to compare across the trials due to inconsistent reporting formats. Safety results appear to be broadly consistent across the trials, the biggest difference between them being the number of deaths occurring in the MAJIC-PV trial, reflecting high-risk population characteristics. The incidence rates of anaemia, specific infections including herpes zoster and non-melanoma skin cancers,were higher in the ruxolitinib arms and crossover groups. Overall rates of infections varied,

being highest in the BAT arm of RESPONSE and the ruxolitinib arm of MAJIC-PV (not reported in the CS for RESPONSE-2). Overall no new safety signals were observed.

3.2.9 Pairwise meta-analysis of intervention studies

No pairwise meta-analysis was conducted because the three trials included by the company each included a different population subgroup (people with splenomegaly in the RESPONSE trial, those without palpable splenomegaly in RESPONSE-2, and a high-risk subgroup with or without splenomegaly in MAJIC-PV). The trials also differed in other characteristics including the presence and timing of crossovers and timing of outcome assessments. The EAG agree that a pairwise meta-analysis was not appropriate.

3.3 Critique of studies included in the indirect treatment comparison (ITC)

The company conducted an indirect treatment comparison which they refer to as a MAIC (matched adjusted indirect comparison). MAIC is a misnomer since the company had individual patient data (IPD) available from both cohorts being compared and used these in a propensity score matching analysis (MAIC, in contrast, is applicable when IPD are available for only one of the cohorts being compared[33]). In this report we refer to the indirect comparison as an ITC.

3.3.1 Rationale for the ITC

The RESPONSE and RESPONSE-2 trials experienced early crossover of patients from the BAT arm to the ruxolitinib arm, from week 32 in RESPONSE and from week 28 in RESPONSE-2. Estimates of the effect of ruxolitinib on overall survival would therefore be confounded by crossover. Adjustment for crossover was not feasible due to the low frequency of deaths (only two on-treatment events at week 256 in RESPONSE; CS section 2.7.1). An ITC was conducted to estimate the effect of ruxolitinib on overall survival without confounding, by comparing long-term survival in the randomised ruxolitinib trial arm of RESPONSE against that in an external BAT cohort, using propensity score matching to balance the characteristics of the ruxolitinib and BAT cohorts.

As discussed below, the ITC is based only on the RESPONSE trial (plus the matching external BAT cohort). The ITC therefore provides an estimate of the effect of ruxolitinib on overall survival specifically for the splenomegaly subgroup, but not for the no palpable splenomegaly subgroup. The company consider the ITC to be a “supportive analysis and

presented for transparency and completeness" (clarification response A8). The ITC results for overall survival do not inform the company's economic analysis base case but do inform scenario analyses (section 5.2.2).

3.3.2 Identification, selection and feasibility assessment of studies for the ITC

The company did not include observational studies in their SLR (section 3.1 above), nor were other data sources for BAT considered (clarification response A8). An indirect comparison (referred to as a MAIC) containing a relevant PV registry (GEMFIN) is listed among the SLR results in CS Appendix D.1.3. The company acknowledge in their clarification response that a systematic search for other real-world registries was not performed, but they argue that a BAT cohort within the GEMFIN registry is likely to represent the most appropriate source of evidence at the time the analysis was conducted:

- The Spanish Registry of Polycythemia Vera set up in 2011 by GEMFIN (Grupo Español de Enfermedades Mieloproliferativas Filadelfia Negativas) referred to as the GEMFIN registry, is one of the largest registries of PV (N=[redacted] as at October 2016) (clarification response A8).
- Results have been published for a subgroup of GEMFIN patients with PV treated with BAT who are resistant to or intolerant of hydroxycarbamide (N=184).[34]
- IPD from GEMFIN were available to the company (clarification response A8).

GEMFIN is a Spanish registry but both the EAG's clinical experts agreed that there is a general lack of robust long-term BAT data for PV patients who are resistant or intolerant to hydroxycarbamide and they were not aware of any registries or other cohorts that would be more relevant than GEMFIN.

An ITC using data from the week 208 analysis of the RESPONSE trial with a subgroup of patients from GEMFIN as the comparator cohort had previously been published as a conference poster by Alvarez-Larrán et al. 2018.[35] The CS provides an update of the ITC using week 256 data from the RESPONSE trial but the GEMFIN data from 2016 (median follow up 3 years) was not updated. The ITC is reported in CS section B.2.9, CS Appendix sections D.1.4. to D.1.8, and in a confidential company slideshow.[36] The ITC used 110 patients from RESPONSE and 184 resistant or intolerant to hydroxycarbamide patients from GEMFIN who had at least one follow-up visit.

A later study by Alvarez-Larran 2022 compared BAT (N=272) and ruxolitinib (N=105) cohorts from GEMFIN using an April 2021 data cut. For OS, they reported a hazard ratio of 0.8 (95% CI 0.4, 1.7) which did not reach statistical significance.

3.3.3 Clinical heterogeneity assessment

RESPONSE and GEMFIN were compared in terms of baseline characteristics (clarification response Table 2). Eight of 10 covariates were considered most likely to be prognostic or treatment effect modifiers by company experts (clarification response Table 1). The EAG's experts also considered resistance to hydroxycarbamide, inadequate HCT, and high WBC as prognostic but these were not reported in GEMFIN.

There are notable imbalances in terms of age (61 vs 69 years), cytopaenia at lowest hydroxycarbamide dose (15% versus 7%), male sex (60% versus 47%), time since diagnosis of PV (8.9 versus [redacted] years), and diabetes ([redacted]% versus [redacted]%). JAK2 mutation status and leg ulcers also showed differences between studies (JAK2: 95% versus 89%; leg ulcers: [redacted]% versus [redacted]%). However, company experts did not rank either highly as a prognostic factor, and the EAG's experts concurred.

Hence, there are imbalances between RESPONSE and GEMFIN in terms of known prognostic factors between studies. Furthermore, other prognostic factors are not reported so differences between the cohorts are unknown.

3.3.4 Risk of bias assessment for studies included in the ITC

The company conducted a risk of bias assessment for the RESPONSE trial (CS Appendix D.3) but not for the GEMFIN cohort. We note that, in an ITC analysis, risks of bias can arise from within each included cohort (e.g. in selection of cases, management of patients, or assessment of outcomes) as well as from the matching method (e.g. inadequate control of confounding):

- In the RESPONSE trial the main risk of bias concern relevant to the ITC is that the trial was open-label, meaning that patient care in the ruxolitinib arm may have been influenced by investigators' knowledge of the treatment allocations (i.e. high risk of bias) (section 3.2.3).
- In the GEMFIN cohort, the retrospective ascertainment of cases could have led to selection bias (random selection from among the available cases could reduce this risk but would also reduce patient numbers)

- The propensity score matching analysis appears to have mitigated confounding to some extent but there is uncertainty as to whether residual confounding remains, due to the limited number of baseline characteristics that were included as covariates in the matching (section 3.4.1 below).

EAG comment on the studies included in the ITC

The EAG agree that GEMFIN is probably the best source of long-term BAT data available, although the availability of evidence has not been evaluated systematically. There are imbalances in prognostic factors between RESPONSE and GEMFIN and some prognostic factors were not reported.

3.4 Critique of the indirect treatment comparison

3.4.1 Data inputs to the ITC

Matching was conducted on OS only. Whilst the original propensity score matching used week 208 data for RESPONSE and week 728 data for GEMFIN [Alvarez-Larran et al. 2018[35] The analysis in the CS was updated using week 256 data from RESPONSE. The GEMFIN data for the matching were obtained in 2016 (median follow up 3 years). If a later cut of GEMFIN were used there would have been more patients and matching may have been more successful. However, as the data do not belong to the company, presumably this would not have been possible.

Evidence for selection of prognostic factors was based upon opinion of 2 clinicians, and those characteristics available and consistently reported in RESPONSE and GEMFIN. The top 8 prognostic factors were ranked by the experts (clarification response Table 1) but only 4 were included in the analysis. Experts were consulted for the Alvarez-Larran (2018) study [35] hence opinions are quite dated. Studies were matched on age, sex, history of thrombosis, and cytopaenia (CS Appendix D.1.6). Of the remaining 4 prognostic factors, uncontrolled myeloproliferation was excluded as there were no events in RESPONSE, duration of PV diagnosis was excluded as definitions differed by study. Diabetes was excluded as numbers were similar across studies and this factor was ranked low by experts (clarification response A11). No explanation is given for excluding failure to reduce massive splenomegaly but the variable is relatively balanced between studies ([redacted]% versus 1%).

No scenario analyses were conducted around variables selection as the company considered GEMFIN "insufficient to support further matching on lower ranked prognostic factors". The EAG disagree, as these rankings were based on the opinions of only two

experts. We would have preferred the company to conduct scenario analyses to explore the broader effect of variable selection on ITC results. However, such analyses are unlikely to be feasible as the company do not own the GEMFIN database (clarification response A11[f]).

Whilst the population matching adjusted for some prognostic factors, others were excluded or not reported, and no scenario analyses around inclusion of prognostic factors were conducted.

The feasibility of combining RESPONSE and RESPONSE-2 in the matching exercise was assessed. However, the company say that results of this "exploratory analysis" could not be located, and the explanation provided as to how this analysis resulted in a "poor fit" is unclear (clarification response A9).

3.4.2 Statistical methods for the ITC

Propensity score matching is an appropriate methodology when the company have access to individual participant data (IPD) for both groups. The company matched RESPONSE with the GEMFIN registry.

Only patients from RESPONSE randomised to ruxolitinib were included in the analysis (patients who crossed from BAT to ruxolitinib were not included). Seven patients included in the original Alvarez-Larran et al. 2018 ITC[35] were excluded from the company submission due to a lack of follow up data subsequent to being identified as resistant or intolerant to hydroxycarbamide (clarification response A10). It is unclear why these patients would have been included in the Alvarez-Larran et al. 2018 analysis.[35]

Multivariate regression was conducted using nearest neighbour matching with prognostic factors as predictors and treatment as the dependent variable. Sample size was reduced from [redacted] in GEMFIN and 110 in RESPONSE to [redacted] post-matching. Studies were reasonably well-matched following matching (Table 12, document B), although there was a [redacted]% difference in males. Two sensitivity analyses were conducted: (i) using a wider nearest neighbour threshold, and (ii) using an optimal matching approach. Results were consistent with the base case.

3.4.3 Summary of the EAG's critique of the ITC

- The chosen propensity scoring methodology is appropriate for the ITC

- GEMFIN appears to be the best choice of dataset for the BAT cohort, although the available evidence has not been evaluated systematically
- The analysis uses a historical data cut of GEMFIN, but as the company do not have access to the dataset, cannot be updated
- Only a limited set of prognostic factors were included in the analysis and these were based on solicited responses from two experts back in 2016
- No scenario analyses around inclusion of variables in the analysis were conducted
- There may have been missing prognostic factors including those identified by EAG experts (e.g. resistance to hydroxycarbamide, inadequate HCT and high WBC counts)
- No scenario analyses were conducted including patients from RESPONSE-2 or MAJIC-PV
- The company list a number of uncertainties in the ITC results including whether GEMFIN was representative of a UK population (they concede low use of IFN-alfa) the generalisability of the GEMFIN population, shorter follow up for GEMFIN (3 years versus 5 years for RESPONSE), a failure to use RESPONSE-2 in the matching, and being unable to include many covariates in the matching (CS section B.2.9.2)
- A published comparison of patients from GEMFIN reported no statistically significant difference in OS between those who received ruxolitinib and BAT

In conclusion, based on the above, in our opinion the OS estimates from the company ITC are highly uncertain

3.5 Overall survival results from the ITC

The overall survival results are shown in Table 17. However, as noted above, we believe these are highly uncertain.

Table 17 Overall survival results from the indirect treatment comparison

Analysis	Number of patients		Number of events		HR (95% CI)[a]
	BAT	Ruxolitinib	BAT	Ruxolitinib	
Pre-matching[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-matching[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-matching[c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

BAT: Best Available Therapy; CI: confidence interval; HR: hazard ratio; OS: overall survival. [a] Based on Cox proportional hazards model with a value less than 1 favouring ruxolitinib. [b] Treatment arm (BAT/ruxolitinib) was used to estimate HR. [c] Treatment arm (BAT/ruxolitinib) and covariates used in matching were used to estimate HR.

Source: Reproduction of CS Table 13

3.6 Additional work on clinical effectiveness undertaken by the EAG

No additional analyses have been conducted by the EAG, as no statistical code nor input data for the ITC were provided to validate the results.

3.7 Conclusions on the clinical effectiveness evidence

The EAG have not identified any key issues in the clinical efficacy evidence that could be resolved by acquiring any additional data or by using alternative analysis approaches. Limitations of the existing data and reporting mean that the clinical efficacy outcomes are subject to uncertainty that would be difficult to resolve unless new evidence (and clearer reporting of studies) becomes available. The three RCTs are all at high risk of bias meaning that variance estimates such as 95% confidence intervals will underestimate the uncertainty present. HRQoL outcomes are particularly at risk of bias due to lack of clarity around missing data, subjectivity of the outcomes in relation to the open-label nature of the RCTs, and selective reporting. Inclusion of the MAJIC-PV trial to compensate for confounding after early crossover in the RESPONSE and RESPONSE-2 company trials is appropriate and has additional advantages, e.g. consisting of a wholly UK population, but is limited by superficial and ambiguous reporting of some aspects of the trial.

Residual uncertainty in the clinical efficacy evidence is summarised in Table 18 below. Although safety outcomes are difficult to compare across trials due to inconsistent reporting, adverse events were generally as expected and do not raise any new concerns.

Table 18 Residual clinical efficacy uncertainties identified by the EAG

#	Source of uncertainty	Effect on certainty of evidence	EAG comment/resolution
1	Radioactive phosphorus is included as a comparator in the NICE scope but excluded from the company's decision problem (section 2.3.2).	Trial BAT arm evidence may not be entirely representative of the NHS PV population receiving BAT who are R/I to HC.	The EAG's clinical experts confirmed that radioactive phosphorus is hardly ever used in clinical practice.
2	Lack of standardisation of definition of R/I to HC in clinical practice (section 2.3.1).	The NHS PV population who are R/I to HC could be broader than in the trials and so influence the overall costs of introducing ruxolitinib.	The EAG's clinical experts confirmed baseline characteristics of the trials are generally reflective of the NHS PV population and subgroups who are R/I to HC.

3	High risk of bias in all three RCTs (section 3.2.3 and Appendix 9.3) • Open label • Selective reporting • Handling of missing data	Uncertainty around the outcomes is not fully captured in the variance measures such as 95% confidence intervals, where reported.	Open label aspect was not justified, however cannot be changed retrospectively. Clarification could be sought on the randomisation process, selective reporting, and missing data around HRQoL specifically.
4	Lack of data from MAJIC-PV: trial publication is unpublished • selective reporting • IPD not available	There are ambiguities around some aspects of the MAJIC-PV trial, e.g. relating to crossovers, missing data and why EQ-5D was not reported.	MAJIC-PV was an investigator-led trial and IPD could not be made available to the company. Final publication of the draft trial manuscript might improve some aspects of clarity.
5	Non-RCT evidence was not systematically searched for: • the SLR was structured to only identify RCTs • Clarification response A8 confirms no systematic search was done to identify real-world studies for the ITC. • Provenance of a study used for additional scenario analyses is not reported.	Uncertainty whether the GEMFIN registry cohort (Alvarez-Larran et al. 2018)[35] used in the ITC is the most appropriate (externally valid) BAT cohort. Uncertainty whether the GEMFIN registry cohort (Alvarez-Larran et al. 2022)[37] used in additional scenario analyses is the most appropriate (externally valid) source of evidence.	The EAG's clinical experts were not aware of any other long term BAT cohorts that would be more relevant and considered the GEMFIN BAT cohort broadly generalisable to the UK. Secondly, the ITC is considered by the company as supportive and not critical evidence. The EAG did not identify a need for the ITC or observational study results to inform the economic model as the included RCTs are sufficiently representative.
6	ITC methods: The results of the ITC are highly uncertain due to: • Limited adjustment for imbalances in prognostic factors between the treatment groups. • High risk of bias in the existing RESPONSE study and in case selection from the GEMFIN registry. • Used an old data cut from the GEMFIN registry • Scenario analyses were not conducted around selection of variables or around	The overall survival estimates from the ITC are uncertain.	The EAG are not aware of any other data that would provide for a more robust analysis. Selection bias in the GEMFIN cohort was partly resolved by propensity score matching. An updated data cut from the GEMFIN registry was not available as the company do not have access to the dataset. The results inform overall survival estimates (and no further outcomes, except that the published study also analysed thrombosis) in scenario analyses only, not in the base case. Results from a recent comparison of BAT and ruxolitinib patients from GEMFIN did not find a statistically

	including patients from RESPONSE-2 or MAJIC-PV		significant difference in overall survival.
BAT: best available therapy; HC: hydroxycarbamide; HRQoL: health-related quality of life; IPD: individual patient level data; ITC: indirect treatment comparison; PV: polycythaemia vera; RCTs: randomised controlled trials; R/I: resistant to or intolerant of; SLR: systematic literature review.			

4 COST EFFECTIVENESS

4.1 EAG comments on the company's review of cost-effectiveness evidence

The company conducted a systematic search for literature on economic evaluations, health state utilities and UK resource use and costs for adults with PV (CS Appendix G). The search strategy was appropriate and reasonably up to date (last updated June 2022). The EAG do not have any concerns about the design or conduct of the reviews. We discuss results for the reviews of utilities and costs/resource use, respectively, in sections 4.2.7.1 and 4.2.8.1 below.

The review of economic evaluations identified five studies, including assessments of the cost-effectiveness of ruxolitinib compared with BAT in populations with PV resistant or intolerant to HC in Ireland (NCPE 2016), the United States (Hong et al. 2020) and Scotland (SMC 2019).[38-40] The SMC have also reported an assessment for ropeginterferon alfa-2b compared with ruxolitinib in a high-risk PV population (SMC 2022).[41] See CS Appendix G Tables 31, 33 and 34 for further details.

4.2 Summary and critique of the company's submitted economic evaluation

4.2.1 NICE reference case checklist

The company summarise key features of their economic evaluation in CS Table 19. The EAG assessment of the company's economic analysis against the NICE reference case checklist is shown in Table 19 below.[42] The company's analysis meets all reference case criteria, except for use of NICE's preferred measure of health-related quality of life, the EQ-5D. Instead, the company use a condition-specific preference-based measure developed for myelofibrosis, the MF-8D, for their base case analyses.[43] See section 4.2.7.2 below for discussion and EAG critique of this decision.

Table 19 NICE reference case checklist

Element of health technology assessment	Reference case	EAG agrees submission meets reference case
Defining the decision problem	The scope developed by NICE	Yes
Comparator(s)	As listed in the scope developed by NICE	Yes
Perspective on outcomes	All health effects, whether for patients or, when relevant, carers	Yes
Perspective on costs	NHS and personal social services (PSS)	Yes

Element of health technology assessment	Reference case	EAG agrees submission meets reference case
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. Effectively lifetime (46 years from starting age at model entry)
Synthesis of evidence on health effects	Based on systematic review	Yes. Health effects from RESPONSE, RESPONSE-2 and MAJIC-PV trials. Scenario with OS HR from ITC. See 4.2.6 below for discussion.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	No. Base case analysis uses MF-8D measure (EQ-5D in scenario). See section 4.2.7.2 below for discussion.
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. MF-8D and EQ-5D valuations from UK general population sample.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit, except in specific circumstances	Yes. The criteria for use of QALY weighting for severity are not met, see Section 7 below.
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
Source: assessment by EAG. Criteria from NICE health technology evaluations: the manual, January 2022		

4.2.2 Model structure

4.2.2.1 State-transition model for RESPONSE and RESPONSE-2 ('primary analysis')

For their primary analysis, the company use a cohort state-transition model (STM) for the licensed population subgroups with and without splenomegaly, based on populations in the RESPONSE and RESPONSE-2 trials respectively (see CS section B.3.2). The same model is used to calculate separate results for each subgroup (pooled results for the whole licensed population are not presented). The model is implemented in Microsoft Excel and employs a 28-day cycle length with a maximum 46-year time horizon, which is effectively lifetime given the age of the cohort at model entry. No half-cycle correction was applied due to the short cycle length. The model was developed with input from an advisory board comprising five UK-based haematologists with PV experience, as well as published literature.

Overview of the model structure

A schematic of the STM structure is provided in CS Figure 35. The model includes three main health states, defined by therapy phases as opposed to disease stages (an approach used in TA386 and TA756 for the treatment of myeloid fibrosis with ruxolitinib and fedratinib).[15 44] Patients enter the model in either the ruxolitinib state or the BAT state, depending on the treatment arm. Patients remain in the ruxolitinib state until discontinuation of ruxolitinib or death. After discontinuation of ruxolitinib, patients move into the BAT state. Patients in the BAT state remain there until death.

In the base case analysis, the BAT state is partitioned into three sub-states, which represent different stages of treatment: first BAT; second or subsequent BAT; and no treatment (discontinuation of all BAT). The company use this BAT partition to model progressive decline in health-related quality of life as patients move through the BAT regimens: utility declines between first, second/subsequent and no further treatment substates, see section 4.2.7.3 below. The BAT partition is implemented using a series of tunnel states, which capture time since initiation of BAT. A scenario analysis with no BAT partition is also presented.

Key complications associated with PV (thromboembolic events (TE), progression to MF, progression to AML and myelodysplastic syndrome (MDS), and haemorrhage) are modelled as events rather than as health states. Incidence rates for these complications and for therapeutic phlebotomy are lower in the ruxolitinib state than in the BAT state, but ruxolitinib is associated with a higher incidence of NMSC. One-off costs and QALY losses are applied for incident cases of TE, MF, AML/MDS, NMSC, haemorrhage and therapeutic phlebotomy.

The company argue that inclusion of these events as health states would be particularly challenging, as it would require many assumptions and data that are not available for the population (CS section B.3.2.2).

Approach to estimation of transition probabilities

The STM structure requires probability estimates for transitions between the ruxolitinib, BAT and death states. These probabilities are estimated from OS and time to treatment discontinuation (TTD) data from the trials. This is challenging for two reasons. Firstly, OS is immature in all three trials due to the relatively good prognosis for people with PV. Secondly, although five-year OS is available for the ruxolitinib arms in the RESPONSE and RESPONSE-2 trials, data for the BAT arms is confounded by cross-over (no patients remained on BAT after 80 weeks). Five-year comparative data are available from the MAJIC-PV trial, as this is unlikely to have been affected by cross-over (Harrison et al. 2022 supplementary Figure S5D).[16]

The company describe their approach to estimating time to treatment discontinuation and overall survival in CS sections B.3.1.2 to B.3.3.4 (note there is an error in the numbering of these sections in the CS). The estimation process is complex; an overview of the EAG's understanding of the process is as follows:

- TTD for reasons other than death is estimated for the ruxolitinib arm using competing-risk analyses of individual patient data from the RESPONSE and RESPONSE-2 trials. These analyses are conducted separately for the two trials and provide separate estimates of ruxolitinib TTD (with deaths censored) for the populations with and without splenomegaly.
- As the numbers of deaths observed in the trials were low, pre- and post-discontinuation survival for the ruxolitinib arm are estimated from pooled data from the RESPONSE and RESPONSE-2 trials.
- Parametric distributions are fitted to the ruxolitinib arm TTD, pre-discontinuation survival and post-discontinuation survival for each trial population. The model combines these extrapolations to estimate OS for the ruxolitinib arm.
- OS for the BAT arm is derived from the modelled OS for ruxolitinib adjusted downwards using a time-varying hazard ratio estimated from MAJIC-PV. The treatment effect is not estimated from the RESPONSE and RESPONSE-2 trials because of the problem with cross-over.

- The rates of discontinuation for the first BAT regimen and for all BAT regimen are estimated from MAJIC-PV data.

Further details and EAG critique of the company's approach to estimation of TTD and OS extrapolations are provided in sections 4.2.6.1 and 4.2.6.2 below.

Other model parameters

In addition to TTD and pre- and post-discontinuation survival, the model uses input parameters to estimate incidence rates for key events and adverse reactions, utilities and resource use/costs. The company present a summary of input parameters for the base case model in CS Table 37. They made some corrections to the parameter values reported in the CS in response to clarification questions and noted that the values in the model were correct. We discuss and critique the clinical effectiveness, utility and resource use/cost parameters in sections 4.2.6, 4.2.7 and 4.2.8 below.

4.2.2.2 Partitioned survival model for MAJIC-PV population ('subgroup analysis')

Individual patient data from the MAJIC-PV trial was not available to the company, as the trial is investigator-led. Consequently, the company employed a partitioned survival model (PSM) to estimate cost-effectiveness for the MAJIC-PV population. In this approach, the proportion of patients in each health state at each time point is estimated based on conventional survival outcomes (usually PFS and OS), and explicit modelling of transitions between the health states, which requires individual patient data, is not needed.[45 46]

As in the primary analysis, the model for the MAJIC-PV population has three health states, based on treatment: 'on ruxolitinib, 'on BAT' and death. Hence, the survival data required is TTD for ruxolitinib and OS. In this model, the BAT health state is not partitioned as with the primary analysis model. Conversely to the primary analysis, the OS for BAT is extrapolated directly from reconstructed KM data reported in the unpublished MAJIC-PV trial paper, with the OS for ruxolitinib estimated indirectly using a time-varying treatment effect.

4.2.2.3 EAG critique of model structure

EAG comments on the modelling approaches: STM vs. PSM (Key issue 4)

- In methodological terms, the state-transition approach has the advantage that the OS extrapolation is structurally related to ruxolitinib discontinuation, unlike the partitioned survival approach in which these outcomes are modelled independently.[45 46] In the current appraisal, the company report scenario analysis with their primary STM

model exploring uncertainty over the extrapolations of both pre-discontinuation survival and post-discontinuation survival.

- NICE DSU TSD19 notes that empirical comparisons have shown that the STM and PSM approaches can produce markedly different results, and that "it is not clear which approach is more reliable".[45] Consequently, TSD19 recommends that STMs should be presented alongside PSMs to verify the plausibility of the PSM extrapolations and to explore key uncertainties in the OS extrapolations.[45]
- A further uncertainty in the current appraisal is whether differences in results from the company's primary and subgroup models relate to the modelling technique (STM versus PSM), or to the different trial populations and contexts of treatment. Exploration of alternative modelling approaches might help to clarify this point. It is not currently possible for the company to conduct an STM analysis for the MAJIC-PV trial population, as they do not have access to individual patient date. However, it would be possible for the company to compare STM and PSM approaches for analysis of the RESPONSE and RESPONSE-2 trial populations.

EAG comments on model structure, states and events (Key issue 5)

- The company's decision to use therapy phases as states, rather than stages of disease, means that their model structure does not reflect the natural history of PV. Although discontinuation of ruxolitinib is likely to be related to long-term survival, other intermediate outcomes such as progression to more aggressive forms of cancer and major thromboembolic or haemorrhagic events are likely to be more strongly prognostic.
- The company cite TA386 and TA756 appraisals as precedent for the use of therapy-based health states for MF. However, a 'supportive care' state after discontinuation of treatment for MF was used in TA386 and TA756. We suggest that the supportive care state may be more directly related to decline in quality of life than the post-ruxolitinib BAT state for PV in the current appraisal.
- We understand that modelling multiple PV-related complications as states rather than as events would add complexity and require additional assumptions and parameter estimates and add uncertainty. However, we note that there are large uncertainties associated with the current model structure. In particular, we are concerned that extrapolation of all-cause mortality from the trials may not reflect the full impact of PV due to time lags between the onset of major complications and related mortality, and the increasing incidence of PV complications with age.

- A more conventional structure for the MAJIC-PV PSM would have been to use a measure of disease progression to define the health states, in addition to treatment discontinuation. For example, the MAJIC-PV manuscript reports KM curves and relative treatment effects for progression-free survival and event-free survival (see section 3.2.6.8 above). One of these intermediate survival outcomes could be used to define pre and post- progression/event health states in a standard three-state PSM structure. We suggest that the company consider an alternative model structure, incorporating an intermediate survival outcome.

EAG comments on partitioning of the BAT state

- Clinical advice to the EAG is that there is not a clear sequence of lines of BAT treatment and long-term cessation of all BAT is considered to be rare. In the absence of alternatives, patients with PV who are resistant or intolerant to hydroxycarbamide continue to switch between currently available medical treatments, with dose adjustments and interruptions to manage symptoms and risks, although this often results in suboptimal control. There is uncertainty over the long-term rate of discontinuation of all BAT therapies and over the assumptions about disutilities for the BAT substates (see sections 4.2.6.1 and 4.2.7.3 respectively). We therefore do not use the BAT partition in the EAG preferred analyses, but we include it in scenario analysis. This is not considered to be a key issue, as the impact on the cost-effectiveness results is modest.

4.2.3 Population

The decision problem population is adults with PV who are resistant or intolerant to hydroxycarbamide, in line with the marketing authorisation for ruxolitinib and the current decision problem (CS B.3.2.1).

The company report three sets of cost-effectiveness results for different subgroups of this population. The primary analysis uses data from the RESPONSE and RESPONSE-2 trials to model subgroups with splenomegaly and without splenomegaly respectively. In addition, the company report results for a ‘high-risk subgroup’, based on the population in the MAJIC-PV RCT. The company argue that all three trial populations are generalisable to England and Wales (CS Table 38). See section 3.2.2 above for discussion of baseline characteristics for patients in the three trials.

The company argue that, collectively, the trial populations with and without splenomegaly in RESPONSE and RESPONSE-2 represent the entire licensed population; with a split of approximately 20% with splenomegaly and 80% without (CS B.3.2.1). Estimates of the prevalence of splenomegaly in practice vary depending on the assessment method and it is difficult to compare estimates from the different trials. In the MAJIC-PV trial, 25% of the population had palpable splenomegaly at baseline (Appendix 9.2 below).

EAG comments on model population (Key issue 1)

- The baseline characteristics of patients in the three clinical trials on which the company's economic analyses based are broadly similar, with the exception of splenomegaly. The EAG clinical advisers agree that all three populations are generally reflective of NHS patients with PV who are resistant to or intolerant of hydroxycarbamide, but that the slightly older population in MAJIC-PV was closer to the patients who they see (section 3.2.2 above). However, we note that estimated survival in the MAJIC-PV population appears noticeably worse than in the RESPONSE and RESPONSE-2 trial populations.
- The NICE scope requests subgroup analysis for patients with and without splenomegaly, which is currently only available from the RESPONSE and RESPONSE-2 trial populations. Expert advice to the EAG is that splenomegaly status would be known at the time patients of consideration for ruxolitinib treatment as patients are assessed by palpation, so this subgroup is identifiable. The EAG experts suggested that people with splenomegaly are more likely to benefit from treatment with ruxolitinib than patients without splenomegaly, although evidence of a difference in treatment effect is lacking. Further analysis to compare cost-effectiveness results for people with and without splenomegaly should be conducted as and when subgroup analysis by baseline splenomegaly status becomes available for the MAJIC-PV trial.

4.2.4 Interventions and comparators

The economic model compares the incremental cost-effectiveness of ruxolitinib to best available therapy (BAT). The intervention and comparator are consistent with the NICE scope. See section 4.2.8.2 below for comments on the dosing assumptions and mix of current treatments in UK practice.

4.2.5 Perspective, time horizon and discounting

The company analyses take the perspective of the NHS and Personal Social Services (PSS) in England, which aligns with the NICE manual for health technology assessments.[42] Costs and outcomes (life years and QALYs) are discounted at 3.5%. The company uses a lifetime horizon to reflect the chronic nature of PV, where lifetime is assumed to be 46 years from the start of the model. Given that the starting age of the patient population in the model is approximately 60-66 years, the company's scenario analysis with a shorter time horizon of 30 years may be more appropriate. We include this scenario in EAG additional analysis (section 6.2.2 below).

4.2.6 Treatment effectiveness and extrapolation

The clinical parameters used in the model consist of time to treatment discontinuation (TTD), parameters required to estimate overall survival (OS) and incidence rates for key complications, therapeutic phlebotomy and adverse events. These parameters were estimated from RESPONSE and RESPONSE-2 trial data, and from the unpublished manuscript for the MAJIC-PV trial, as summarised in CS Table 21. We summarise the clinical parameters used in the company's primary and subgroup models in Table 20 and Table 21 respectively. Description and EAG critique of the company's approach to estimating these parameters is provided in the following sections of this report.

Table 20 Summary of clinical parameters in the primary model (RESPONSE and RESPONSE-2 trial populations)

Parameter	Base case analysis	Source
Time to treatment discontinuation (TTD)		
Ruxolitinib TTD (excluding death)	Odds spline with 1 knot for both subgroups, CS Figure 39	Competing-risk analyses of RESPONSE and RESPONSE-2 IPD for the two subgroups
BAT TTD 1st BAT	KM and Gompertz tail	Extrapolation of reconstructed KM data for discontinuation of first BAT regimen in MAJIC-PV
BAT TTD all BAT	BAT OS / HR ([redacted])	HR approximated from numbers of deaths and discontinuations in the BAT arm of MAJIC-PV
Overall survival (OS)		
Ruxolitinib pre-discontinuation survival	Exponential for both subgroups (+ gen pop mortality constraint applied post- trial) CS Figure 41	Data from RESPONSE and RESPONSE-2 were pooled due to the small number of deaths observed within the trials (same extrapolations for both subgroups)
Ruxolitinib post-discontinuation survival	Exponential (+ gen pop mortality constraint over time horizon) CS Figure 44	
OS for ruxolitinib	Calculated indirectly by STM	-

OS for BAT	Ruxolitinib OS x time varying HR (1.10 before 3 years; 2.20 from year 3, waning from year 5 to HR=1 at year 20) CS Figures 47 and 48	HR estimated from piecewise Cox proportional hazards analysis of reconstructed MAJIC-PV KM data
Event rates		
Key complications and phlebotomy (ruxolitinib)	Exposure-adjusted incidence rates while on ruxolitinib CS Table 24	Incidence rates estimated from relevant trial for population when available
Key complications and phlebotomy (BAT)	Incidence for ruxolitinib adjusted for BAT with IRR CS Tables 25	Incidence-rate ratios calculated from pooled RESPONSE, RESPONSE-2 and MAJIC-PV
Adverse events	Incidence rates CS Table 23	Exposure-adjusted incidence rates (any grade) pooled for RESPONSE and RESPONSE-2
Source: summary produced by EAG BAT best available treatment; CS company submission; CQ clarification question response; gen pop, general population; HR hazard ratio; IPD individual patient data; IRR incidence-rate ratios; KM Kaplan-Meier; OS overall survival; TTD time to treatment discontinuation		

Table 21 Summary of clinical parameters in the subgroup model (MAJIC-PV population)

Parameter	Base case analysis	Source
Time to treatment discontinuation (TTD)		
Ruxolitinib TTD	Ruxolitinib OS x HR for TTD vs. OS ([redacted]) See CS Figure 42 and CQ response B5	Ruxolitinib OS adjusted with HR for TTD vs. OS. HR estimated from reconstructed KM for ruxolitinib arm of MAJIC-PV
TTD all BAT	BAT OS / HR ([redacted])	Estimated as above
Overall survival (OS)		
OS for BAT	Weibull extrapolation (+ gen pop mortality constraint over time horizon) CS Figure 46	Extrapolation fitted to MAJIC-PV reconstructed KM data for BAT arm
OS for ruxolitinib	BAT OS / time varying HR	BAT extrapolation adjusted by same HR as in primary analysis
Event rates		
Complications Phlebotomy Adverse events	Same as for primary analysis	
Source: summary produced by EAG BAT best available treatment; CS company submission; CQ clarification question; gen pop, general population; HR hazard ratio; KM Kaplan-Meier; OS overall survival; TTD time to treatment discontinuation		

4.2.6.1 Time to treatment discontinuation

4.2.6.1.1 Primary analysis (RESPONSE and RESPONSE-2 populations)

Ruxolitinib discontinuation

The TTD for ruxolitinib was modelled under a competing-risk framework, which is appropriate for the state-transition model. This allows the model to account for the increased likelihood of discontinuation due to death as patients age. The TTD for ruxolitinib due to reasons other than death and pre-discontinuation survival are initially modelled separately before being combined within the model 'trace' sheets.

The approach to fitting extrapolations for ruxolitinib discontinuation for reasons other than death is explained in CS section B.3.1.2. The analysis was conducted separately for people with and without splenomegaly, using individual patient data (with deaths censored) from the RESPONSE and RESPONSE-2 trials respectively (see CS Figure 38). The company followed recommended methods to fit and choose extrapolations in each population from NICE Decision Support Unit (TSD14).[47] See CS Appendix N.1 and N.2 for graphs and statistical measures of fit. For the base case, the company chose the odds spline model with one knot for both patients with and without splenomegaly (CS Figure 39). Other distributions were used in scenario analysis and the ICERs were moderately sensitive to the choice of distribution (CS Appendix P).

Figure 1 and Figure 2 below show the company's selected odds spline with one knot distribution and the EAG's preferred assumption of a Weibull distribution in comparison with KM data for TTD for ruxolitinib due to reasons other than death for the licensed population with and without splenomegaly, respectively. We prefer the Weibull distribution, because it has a better statistical fit for the RESPONSE trial and similar fit for RESPONSE-2.

Results with other selected distributions (lognormal, loglogistic, and the hazard spline with one knot) are shown in scenario analysis in Table 27.

Figure 1 TTD for ruxolitinib for the licensed population with splenomegaly

Abbreviations: TTD: time to treatment discontinuation; KM: Kaplan-Meier; OS: overall survival.

Source: Reproduced from CS Appendix N Figure 18 using selected distributions.

Figure 2 TTD for ruxolitinib for the licensed population without splenomegaly

Abbreviations: TTD: time to treatment discontinuation; KM: Kaplan-Meier; OS: overall survival.

Source: Reproduced from CS Appendix N Figure 19 using selected distributions.

BAT discontinuation

Time to discontinuation of the first BAT treatment was derived from reconstructed KM data from the MAJIC-PV trial (CS Figure 50). As with TTD for ruxolitinib, parametric extrapolations were fitted to the KM data and the fit and clinical plausibility assessed (see CS section B.3.3.5 and appendix N.6). As the data were mature, the company chose to use KM data directly for the 5-year follow up, with a Gompertz extrapolation for the remaining time horizon. The number of people remaining at risk in the KM at 4 and 5 years was 45 and 14 respectively.

The TTD for all BAT treatments is not reported in the MAJIC-PV manuscript; the company estimated a hazard ratio (HR) between OS and TTD of [redacted] using the number of reported deaths and discontinuations in the BAT arm from the unpublished manuscript.

See CS Figure 49 for the resulting distribution between the three BAT substates in the company's base case model. The TTD for second and subsequent BAT is estimated as the difference between the TTD for first BAT and TTD for all BAT. The time in no treatment is taken as the difference between OS for BAT and the TTD for all BAT.

The company assume that after discontinuation of ruxolitinib, patients are distributed to the three BAT substates in the same proportions as patients who were initiated on BAT at the same model cycle.

4.2.6.1.2 Subgroup analysis (MAJIC-PV population)

Time to discontinuation for ruxolitinib in the partitioned-survival model for the MAJIC-PV population was not modelled under a competing-risk framework, as the company did not have access to individual patient data from the MAJIC-PV trial. Instead, a HR of [redacted] was derived from reconstructed pseudo IPD for OS and TTD for ruxolitinib, which was then applied to the OS for ruxolitinib to obtain the TTD (note this HR was incorrectly reported in the CS, see correction in the company's response to clarification question B5). The company note that this approach follows clinical expert advice that TTD for ruxolitinib should be consistent with OS.

For discontinuation of BAT in the PSM, the same approach is used as described above for the primary STM model.

EAG comments on TTD extrapolations

- The company followed the recommended approach to fitting extrapolations for time to discontinuation of ruxolitinib and initial BAT treatment and provided clear reasons for their choice of distributions in the base case models.
- In the company's primary model, the distribution used for the extrapolation of TTD for ruxolitinib has a moderately large impact on the ICERs (CS Appendix P), because the STM structure means that TTD impacts on long-term survival as well as treatment-related utility and costs. (Note that this is not the case for the subgroup model (PSM) for the MAJIC-PV population, in which the TTD for ruxolitinib is linked via hazard ratio parameters to the OS extrapolation for BAT.)
- The company's choice of distribution (odds spline with one knot) for the extrapolation of ruxolitinib discontinuation in their primary is reasonable. We use a Weibull distribution in EAG preferred analysis, as this has a better statistical fit for the RESPONSE population. This results in a bigger difference in long-term continuation of ruxolitinib between the two subgroups, as shown in **Error! Reference source not found.** and Figure 2 above.
  Results with the Weibull and other selected distributions (lognormal, loglogistic, and the hazard spline with one knot) are shown in scenario analysis in Table 27.
- The model results are not sensitive to changes in the distributions used for extrapolation of time to discontinuation of BAT estimated from the MAJIC-PV trial (same distributions used in all three populations).

4.2.6.2 Overall survival

4.2.6.2.1 Treatment effect (OS HR for ruxolitinib versus BAT)

For the base case, the company used a time-varying HR estimated from reconstructed KM data from the MAJIC-PV trial. The company's clinical advisors noted that the KM curves appear to diverge after about 3.0 years (see CS Figure 47), which was in line with the experts' expectations based on intermediate outcomes (CS B.3.3.4). The company fitted a piecewise Cox proportional hazards model to reconstructed MAJIC-PV KM data to estimate hazard ratios before and after this cut point. CS Appendix O shows log-log and Schoenfeld residuals plots based on reconstructed KM data, which the company used to assess the timing of the change in HR.

The company reported scenarios with different cut-points (2.6 to 2.9 years) for their time-varying HR estimates. They also reported four other scenarios with fixed HR estimates applied throughout the time horizon: the HR from the unpublished report by the MAJIC-PV

investigators; the estimate from the company's ITC analysis (see section 3.5 above); a propensity score adjusted incidence rate ratio (IRR) of death from a retrospective analysis of Spanish registry data (Alvarez-Larrán et al. 2022)[37]; and an HR estimated from pooled RESPONSE and RESPONSE-2 trial data, without adjustment for crossover. As might be expected, the ICERs were highly sensitive to these very different HR estimates (CS Appendix P and company response to clarification question B2).

Table 22 Treatment effect estimates used in company analysis

Analysis	HR for OS (ruxolitinib vs. BAT)	Source
MAJIC-PV time-varying HR (base case)	0.91 (95% CI 0.38 to 2.18) 0 to 3 years 0.45 (95% CI 0.13 to 1.61) 3 to 5 years	CS section B.3.3.4
MAJIC-PV constant HR	0.73 (95% CI 0.36 to 1.50; p=0.39)	Harrison et al. 2022, Figure S5D [16]
Company ITC	[redacted]	CS Table 13
Spanish registry data	0.8 (95% CI 0.4 to 1.5; p=0.4)	Alvarez-Larrán et al. 2022 [37]
Pooled RESPONSE and RESPONSE-2 data	[redacted]	Company model
Source: EAG using data from company submission and model BAT best available treatment; CI confidence interval; HR hazard ratio; ITC indirect treatment comparison; OS overall survival		

Waning assumptions

In their base case, the company assume a gradual waning of the treatment effect after the trial period: with a linear increase in the HRs from the above estimates at year 5 to no effect (HR=1) at year 20 and beyond. This was based on clinical expert judgement that approximately twice the number of patients would be alive at 20 years with ruxolitinib compared with current treatment (see CS section B.3.3.4). The company tested various scenarios for the duration of the waning period, from 5 to 50 years. Results were sensitive to different waning assumptions.

EAG comments on the treatment effect for survival (Key issue 2)

- Evidence on the relative treatment effect on survival is highly uncertain. The confidence interval around the HR reported by the MAJIC-PV trial investigators is wide. The company's time-varying HR estimates are not unreasonable based on trends in the MAJIC-PV KM curves (CS Figure 47). The log-log and Schoenfeld residuals plots (CS Appendix Figures 26 and 27) provide support for the assumption of proportional hazards prior to 3.0 years and increasing divergence after this timepoint. However, these estimates are also highly uncertain. For the EAG analysis, we prefer to use the constant HR estimate as reported by the

MAJIC-PV trial investigators, but we report results with the company's time-varying HR estimates in scenario analysis.

- Other estimates of the treatment effect are used in the company's scenario analyses, including: estimates from pooled RESPONSE and RESPONSE-2 data, the ITC matched comparison with GEMFIN registry data, and the analysis of Spanish registry data (Alvarez-Larrán 2022)[37], see Table 22 above. We report EAG results with these scenarios for information but consider the MAJIC-PV trial to be the most robust source of evidence for relative treatment effects.

EAG comments on the waning of the OS treatment effect (Key issue 3)

- There is uncertainty over whether and how the treatment effect might change after the trial period. Given the uncertainties around the estimation of the treatment effect, we agree with the company's use of a waning assumption (linear increase in the HR from year 5 to HR=1 at year 20). We have not changed the waning period in EAG preferred analysis, but note a longer waning period, or the removal of waning, might be appropriate with the more conservative constant HR estimate that we use,

4.2.6.2.2 Ruxolitinib extrapolation for RESPONSE and REPONSE-2 populations

The OS for ruxolitinib was modelled indirectly using the extrapolations of TTD excluding death described above, and extrapolations of pooled data for pre-discontinuation survival and post-discontinuation survival(see CS sections B.3.1.2 and B3.3.2). Pooled data were used because of the small number of deaths observed in the trial, both pre- and post-ruxolitinib discontinuation.

The fitted extrapolations for pre-discontinuation survival are illustrated in CS Appendix N.3. The company choose an exponential distribution for their base case, which had the best statistical fit, with alternative distributions assessed in scenario analysis. They included a constraint to ensure that the hazard of death was no less than that for members of the general population of the same age and gender mix, but this was only applied after the trial period. In response to clarification question B4, the company added an option in the model to include the general population constraint throughout the time horizon (CQ response Figure 1).

Extrapolations for post-discontinuation survival are presented in CS B.3.3.2 and Appendix N.4. Again, the company chose an exponential distribution, which had the best fit to the trial

data and was considered clinically plausible by the company's experts. The general population mortality constraint was applied throughout the time horizon. The resulting extrapolation is illustrated in CS Figure 44.

The STM model combines the extrapolations for time to ruxolitinib discontinuation, pre-discontinuation survival and post-discontinuation survival to estimate OS for ruxolitinib.

EAG comments on the ruxolitinib OS extrapolation (primary analysis):

- The use of a competing-risk framework to estimate TTD, and subsequently OS for ruxolitinib is appropriate for the STM structure of the company's primary analyses. We agree with the pooling of data from the RESPONSE and RESPONSE-2 trials for estimation of pre- and post-discontinuation survival extrapolations, given the small numbers of deaths observed. However, this means that the comparative results for the patients with and without splenomegaly may not fully reflect survival differences between these subgroups.
- The company's base case extrapolation for pre-discontinuation survival is not adjusted for general population mortality during the trial period. This results in a lower mortality rate during the first five years of ruxolitinib treatment than for people of the same age and gender mix in the general population, which is not plausible. For the EAG preferred analysis we use the general population mortality constraint for pre-discontinuation survival throughout the time horizon. This results in mortality rates prior to discontinuation of ruxolitinib that are the same as for the general population, so he model is not sensitive to the distribution for extrapolation of pre-discontinuation survival.
- The model is somewhat sensitive to the distribution used for post-discontinuation survival. The company use an exponential extrapolation in their base case, which provides a reasonable fit to the trial data.

4.2.6.2.3 OS extrapolation for MAJIC-PV population

The OS for BAT was extrapolated directly from reconstructed OS KM data from the MAJIC-PV manuscript using a Weibull distribution (see CS B.3.3.3 and Appendix N.8).

The OS for BAT was derived by applying a relative treatment effect to the ruxolitinib OS extrapolation. In the base case analysis, the company used data from the MAJIC-PV trial, because comparative evidence from the RESPONSE and RESPONSE-2 trials was confounded by cross-over from the BAT arm to ruxolitinib.

The same estimates of the treatment effect were used in both STM and PSM models, and for all three trial populations (MAJIC-PV, RESPONSE and RESPONSE-2). See Table 22 below for the HR estimates used in the company's base case and scenario analyses.

The OS for ruxolitinib was derived from the OS for BAT by applying the time-varying treatment effect to the BAT OS, see discussion in section 4.2.6.2.1 above. Note that these HRs are the inverse of those used in the primary analysis, as the OS extrapolation for ruxolitinib in the MAJIC-PV population analysis was derived from the BAT OS extrapolation (in contrast with the primary analysis, where the OS extrapolation for BAT was estimated from the ruxolitinib OS extrapolation). The same gradual linear waning of the treatment effect from year 5 to year 20 employed in the primary analysis was also used in the MAJIC-PV population analysis.

4.2.6.3 Key complications (events)

The company incorporates five key complications as events in the economic model: TE, progression to AML or MDS, progression to MF, haemorrhage, and NMSC.

The incidence rates of key complications while on ruxolitinib were calculated based on the numbers of events reported in the RESPONSE, RESPONSE-2 and MAJIC-PV trials, adjusted by the duration of exposure to ruxolitinib or total follow-up time. CS Table 24 reports the exposure-adjusted incidence rates for patients on ruxolitinib for the three trials. Trial-specific data for the relevant population were used, where available.

The incidence of events whilst on BAT were estimated by applying a treatment effect in the form of incidence rate ratios (IRR) to the baseline incidence rate of events on ruxolitinib. To account for the small number of events and varying follow-up durations, the IRRs were estimated using the pooled number of events from the RESPONSE, RESPONSE-2 and MAJIC-PV trial. The IRRs used for each of the five events are reported in CS Table 25.

The company notes that none of the trials were powered to estimate the incidence of these key complications. They also note that assumptions were required for missing data, not reported for specific trials (see CS B.3.3.8).

EAG comment on estimated event rates for key complications (events)

- The incidence of key complications in the ruxolitinib arm was based on reported rates per patient year of exposure from the three trials. We note that these rates are fixed across the time horizon and are not adjusted for age.
- The incidence of the key events while patients were on BAT was estimated from relative rates (IRRs) from pooled trial data. This resulted in lower incidence of MF, TE and haemorrhage, and higher incidence of non-melanoma skin cancer while patients were on ruxolitinib than on BAT. There was very little difference between the treatments in estimated rates of conversion to AML/MDS.
- The company reported scenarios excluding the impact of the individual key events, and excluding all events in CS Appendix P. This showed limited impact on the ICERs.

4.2.6.4 Therapeutic phlebotomy

The rate of therapeutic phlebotomy for patients on ruxolitinib was derived from each of the three trials and applied to the respective analysis population: [redacted], [redacted], and [redacted] for RESPONSE, REPONSE-2, and MAJIC-PV, respectively (see CS section B.3.3.9). The unpublished MAJIC-PV manuscript did not report exposure time, and a total number of phlebotomy procedures was reported during the entire study period. as opposed to during ruxolitinib treatment only. Therefore, total follow-up time estimated from the pseudo-IPD for OS was used. As with complications, the number of phlebotomy procedures across all trials and the exposure time for ruxolitinib and BAT were pooled to calculate a treatment effect IRR of [redacted], which was applied to the rates for ruxolitinib to acquire the rate of phlebotomy for patients on BAT.

4.2.6.5 Adverse events

The model included adverse events occurring at a rate of ≥5 per 100 patient-years of exposure and at a rate of ≥3 per 100 patient-years of exposure in either arm of the RESPONSE and REPONSE-2 trials, respectively. CS Table 23 reports the pooled exposure-adjusted rates of 67 AEs. All grades of AEs were included in the model, with Grades 1 and 2 having a lesser impact than Grades 3 and 4. In the primary analysis, the rates of AEs from the RESPONSE and RESPONSE-2 trials were pooled for both patients with and without splenomegaly. The unpublished MAJIC-PV manuscript only reports AE categories experienced by ≥10% of patients and does not have data regarding Grade 1 or 2 AEs nor on the duration of exposure; the analysis for this population therefore used the same incidence of AEs used for the primary analysis.

4.2.7 Health related quality of life

4.2.7.1 Systematic literature review for utilities

The company identified two studies that reported utility estimates for people with PV from their systematic review of literature on health-related quality of life (CS Appendix H). The study by Lelonek et al. (2018) reported EQ-5D-3L values with a UK tariff for 102 people with PV and the JAK2V617F mutation.[48] Mean (SD) utility scores were the same for people with and without aquagenic pruritus: 0.8 (0.1) (see CS Appendix Table 40).

The second study was the Scottish Medicines Consortium (SMC) review for ropeginterferon (2022).[41] This included EQ-5D-3L utility scores for 1,142 adults with PV from the PROUD-PV and CONTINUATION-PV studies. Mean (SD) utility scores were cited of 0.881 (0.152) for 892 people with JAK2<50 and 0.876 (0.148) for 250 people with JAK2≥50 (CS Appendix Table 41). The company state that these data were collected from an international study which did not include UK patients, and that the value set was not reported. It is therefore not clear that these estimates would meet NICE reference case requirements.

Neither study was specific to the population of interest in this current appraisal. So, as utility data was available from the RESPONSE and RESPONSE-2 trials, the company did not use the above estimates in the economic model. The EAG agree with this judgement.

See CS Table 28 for a summary of utility values used in the economic model.

4.2.7.2 Study-based health related quality of life

Treatment specific utility values were derived from individual patient data from the trials, using regression analysis, with treatment and baseline values as covariates (see CS B.3.4.3 and company response to clarification question B9). For their base case, the company use utility estimates for condition-specific preference based utility instrument (the MF-8D),[43] derived from EORTC QLQ-30 and MPN-SAF data from the RESPONSE trial.

The MF-8D was developed for use in myelofibrosis and uses three items from the EORTC QLQ-30 and five from the MF Symptom Assessment Form (MF-SAF). The MF-SAF is similar to the MPN-SAF, but with differences in the wording of two items used in the MF-8D. The company therefore had to make the following assumptions to use the MF-8D for the PV population in the RESPONSE trial:

- That “pain under ribs on the left side” in the MF-SAF is equivalent to “abdominal pain” in the MPN-SAF
- And that “bone or muscle pain” in the MF-SAF is equivalent to “bone pain” in the MPN-SAF.

The company justify their preference for the MF-8D on the basis that the EQ-5D is not appropriate for capturing the impact of PV on health-related quality of life (CS section B.3.4.1). Their argument is based on:

- Published psychometric analysis which indicates that the EQ-5D and EORTC QLQ-C30 instruments to do capture the key symptoms of myelofibrosis.[43 49]
- Precedent from two NICE MF appraisals (TA386 and TA756), in which the NICE committees accepted use of the condition-specific MF-8D.[15 44]
- The similar nature of symptoms for PV and MF, including fatigue, early satiety, abdominal discomfort, inactivity, concentration, night sweats, itching and bone pain. EAG clinical expert advisors agreed that the symptoms of MF and PV are generally similar in nature but vary in severity.

The company also report results from an ‘exploratory’ psychometric analysis of EQ-5D-5L (mapped to UK EQ-5D-3L utility values using the NICE recommended method)[42 50 51]) and MPN-SAF data from RESPONSE-2 (CS B.3.4.1). Further information about the exploratory psychometric analysis was provided in a PowerPoint report in response to clarification question B8. This analysis included a comparison of ceiling effects, item correlation and standardised measures of change from baseline for the EQ-5D and MPN-SAF TSS.

The company also report a scenario analysis based on EQ-5D-5L data from the RESPONSE-2 trial (CS section B.3.4.1). For this analysis, UK 3L utility values were derived using the algorithm developed by Hernández Alava et al. 2020, as currently recommended by NICE.[42 50 51]

Health state utilities are appropriately adjusted in the model for aging of the population, using UK general population utility data (Hernandez et al. 2022).[50]

4.2.7.3 Disutility for BAT substates

In the primary analysis, reductions in utility values and disutilities are assigned for the BAT sub-health states as follows:

- From baseline to 1st BAT sub-health state: [redacted]

- From baseline to 2nd+ BAT sub-health state: [redacted]
- No treatment sub-health state: -0.05.

The higher disutility for the no treatment sub-health state is in line with the greater decline in health for patients with high-risk PV who are not on treatment.

4.2.7.4 QALY loss associated with key events

The QALY loss for reduced utility associated with key complications were calculated based on estimates of disutility and life expectancy derived from the literature (CS Table 27). In response to clarification question B10, the company states that sources used to calculate these QALY losses were not derived by systematic review.

The EAG noted in clarification B12 that although the QALY losses associated with key events include utility lost during expected survival following an event, the QALY losses do not include QALY loss for shortened life expectancy due to an event. The company stated that extrapolation of overall survival beyond the observed trial period implicitly accounts for the increase in death caused by a key complication; incorporating years of life lost due to an event could result in double counting. There is no possibility of determining the proportion of deaths due to a key event or due to other reasons, as overall survival is modelled directly for an average cohort and extrapolated over time, regardless of the cause of death.

For venesections, the company assume a QALY loss of -0.000103 per procedure, based on a decrement in utility of -0.037 procured from Matza et al. 2013 with the assumption that the decrement lasts for one day.[52] The company have confirmed in clarification response C3 the error in the company submission regarding the QALY loss associated with phlebotomy: the correct value of -0.000103 is implemented in the model.

4.2.7.5 QALY loss associated with adverse events

The impact of adverse events on HRQoL is not included in utility values but is captured in the model separately. The health disutility of an adverse event was based upon the health utility decrement and the duration of impact on quality of life of that particular adverse event. The company did not implement any health disutilities for Grade 1 or 2 adverse events, stating that this would simplify the model. CS Table 26 reports the disutilities and durations for the 36 categories of Grade 3 and 4 adverse events used in the model. Data for these adverse events were taken from values used in previous NICE appraisals and from the literature. For Grade 3 or 4 adverse events which no data could be sourced, the company

assumed a disutility of -0.075 for a duration of seven days, based on results used in NICE TA772.[53]

EAG comments on health-related quality of life

- This provides some evidence in favour of the MF-8D, including greater responsiveness and lower susceptibility to ceiling effects.
- However, the MF-8D was not developed for use in PV, and the company had to make assumptions to substitute the PV symptom score for the myelofibrosis symptom score used in the MF-8D. There is also a lack of direct evidence validating the EQ-5D and MF-8D in a PV population.
- EAG clinical experts advised that the MPN-SAF TSS is mostly used in MF as that is the most symptomatic myeloproliferative disorder, but as there is extensive symptom overlap between MF, essential thrombocythemia and PV, they consider that the instrument would capture PV symptoms.
- We use EQ-5D utilities in the EAG preferred analysis. This follows the NICE preference for use of the EQ-5D when available from relevant clinical trials, as this provides consistency across NICE appraisals. There is some evidence in favour of the MF-8D measure, but also uncertainty about its transferability from MF to PV.
- There is uncertainty regarding the accuracy of the QALY losses associated with key events, which do not consider the QALY loss associated with years of life lost. There is scope for further analyses regarding the QALY losses used, and whether more conservative QALY losses should be implemented to account for the lack of data regarding the potential decrease in life expectancy following a key event.

4.2.8 Resources and costs

4.2.8.1 Systematic literature review of costs and healthcare use

The company report the results of their review of cost and resource use data in CS Tables 45 and 46. They included three studies in their review, including the Scottish Medicines Consortium appraisal of ropeginterferon, but the company conclude that this data was not usable, because the population from which the data was sourced was not defined (SMC 2022).[38-40] The other two UK based studies were not used either, as one was considered too old and the other did not state the cost year.

4.2.8.2 Drug acquisition and administration

CS Table 29 reports the dosing schedule and costs of drugs used in the model. The 4-weekly treatment costs for ruxolitinib used in the model were [redacted] and [redacted] for the primary analysis with and without splenomegaly, respectively. As the unpublished MAJIC-PV trial manuscript does not contain data on dosage distribution, the company assumed a treatment cost for ruxolitinib in the MAJIC-PV population of [redacted], the same as the RESPONSE-2 trial. These prices for ruxolitinib are using the current Patient Access Scheme (PAS) discount for myelofibrosis of [redacted].

[redacted] Ruxolitinib is administered orally; there are no associated administration costs. The dosage for ruxolitinib in the model is based upon the RESPONSE and RESPONSE-2 trials for both primary and MAJIC-PV analyses.

For BAT, a 4-weekly treatment cost of £226.48 was used in the model for both primary and MAJIC-PV analyses. This was based on the distribution of treatments in the BAT arm of the MAJIC-PV trial (CS Table 20), but as pipobroman and radioactive phosphorus are no longer in use in England and Wales, they were excluded from the BAT composition in the model. Ruxolitinib, used in combination for a small number of patients in the MAJIC-PV BAT arm, was also excluded. Unit costs for the included BAT medications are shown in CS Table 29. We note that the company use the cost for a pegylated derivative of interferon-alfa, as this is now routinely used in NHS practice.

All patients on interferon-alfa require training on how to self-inject the drug, which involves one or two visits with a nurse or GP. However, according to clinical experts, approximately 5%-10% of patients with PV using interferon-alfa require continuous help from a nurse to administer the injection; the remaining patients on interferon-alfa are able to self-inject once trained and do not incur administration costs. Therefore, the model implements a one-off cost of £24.71 for patients on BAT to include the cost of training and district nurse visits.

EAG comments on drug acquisition and administration

- Clinical experts advising the EAG have noted that the majority of patients would continue to be treated with interferon-alfa or hydroxycarbamide (despite being resistant or intolerant to the treatment). Anagrelide and busulfan are seldom prescribed. Approximately 10-15% of patients resistant or intolerant to hydroxycarbamide would have no other suitable alternative.

- We have also been advised that the majority of patients with PV on interferon-alfa would self-administer the drug, but between 2-10% would require on-going nurse help for injection.

4.2.8.3 Patient management and monitoring

There were no UK cost studies or NICE appraisals for PV identified in the company's economic SLR. Therefore, resource utilisation data was obtained from questionnaires completed by five UK clinical experts with experience in PV. The clinicians provide estimates for the management and monitoring of PV over three time intervals: 0-6 months, 7-12 months, and 13+ months of treatment. CS Table 32 provides the estimated resource use and unit cost per cycle for the different resource categories; the same resource use and costs were used for both primary and MAJIC-PV analyses.

The management and monitoring costs used in the model per cycle for patients on ruxolitinib were estimated to be [redacted], [redacted], and [redacted] for 0-6 months, 7-12 months, and 13+ months of treatment, respectively. The corresponding costs used in the model for patients on BAT were [redacted], [redacted], and [redacted]. In the primary analysis where the BAT state is partitioned, patients in the "no treatment" sub-health state incurred an assumed cost of [redacted] per cycle, twice the cost of patients on BAT, and was fixed across all time intervals. This sub-health state was assumed to have a higher cost to represent the worsening of PV and a subsequent increase in management and monitoring when patients are no longer on treatment. The model also included a cost of £316 per therapeutic phlebotomy, and a cost of £6,774 for end of life care.

4.2.8.4 Adverse events and key events

The unit costs for Grade 3 and 4 adverse events are provided in CS Table 35, taken from the NHS reference costs 2020/21. Note that only 36 categories of the 67 adverse event categories were reported to have at least one Grade 3 or 4 event in either arm of the trials. The cost for the management of Grade 1 and 2 adverse events were assumed to be equivalent to the cost of two GP e-consultations at a total of £78.46.

CS Table 33 reports the management costs assumed for each of the five key events (TE, AML/MDS, MF, NMSC and haemorrhage). The company have noted in clarification response B13 and B14 the errors in costs in the table: the cost for the management of a TE event used in the model is £1,865, and the cost for a haemorrhage event is £2,023.

The cost for the management of a TE event, £1,865, was based upon the grade levels of events, unit costs, and the distribution of TE events in the ruxolitinib arms of RESPONSE and RESPONSE-2. CS Table 33 contains the unit costs from the NHS reference costs 2020/21 for Grade 3 and 4 TE events. The cost for an emergency department visit was assumed for the management of a Grade 1 or 2 TE event. The company have noted in clarification response B13 that the cost of an emergency department visit is stated incorrectly as £182 in the CS; the correct cost is £297.

The cost for the management of AML/MDS implemented in the model, £44,903, was also used in NICE TA386 and NICE TA756, and was taken from the results of a probabilistic decision model in AML by Wang et al. 2014.[15 44 54] The cost is the median value of the range of reported costs in Wang et al. 2014, who estimated 5-year medical costs for the management of AML in the UK.

The management cost for MF assumed in the model was £63,920. The costs for managing intermediate-2/high-risk MF which occurred in 57.3% of patients with MF was determined from TA386 using the total costs for ruxolitinib, £128,403, and BAT patients, £36,095.[15] The company were unable to find data on the management cost for the remaining 42.7% of patients with low/intermediate-1 MF, and so they assumed a cost of £72,190, double the cost of intermediate-2/high-risk MF in BAT patients. The company note that patients with low/intermediate-1 MF generally have a more favourable prognosis than patients with intermediate-2/high-risk MF, and will consequently have an increased duration of treatment, leading to higher overall resource use.

The management costs for NMSC and bleeding/haemorrhaging events used in the model were £1,058 and £2,023, respectively. The cost for an NMSC event was based on results in Vallejo-Torres et al. 2013, whilst the cost for a major haemorrhaging event was based on Crathorne et al. 2018; the management cost for a minor bleed was assumed to be equivalent to the cost of one emergency department visit, £297.[55 56]

EAG comments on resources and costs

- Clinical advice to the EAG was that in addition to an emergency department visit, patients with a grade 1 or 2 thromboembolic event would need a D-Dimer test and an ultrasound doppler scan. We include an additional cost for these tests in the EAG preferred analysis.

- In addition, interim or long-term treatment with warfarin or an oral anticoagulant would be initiated for some patients after a grade 1 or 2 thromboembolic event. We therefore include the cost of a single dose of an anticoagulant, as stated in the NICE guideline NG158.[57]. The effects of this cost change are discussed in section 5.3.3.2. We have not included the costs (or benefits) of thromboprophylaxis in our scenario analysis, as this would be difficult to estimate. However, we note that the impact of grade 1 or 2 thromboembolic events are likely to be underestimated in the model.
- Other estimated costs for adverse events were considered reasonable. It was noted that patients often consult with clinical nurse specialists for drug-related adverse effects, but the assumption of 1 or 2 GP online consultations was considered to be reasonable for the cost calculations. The company's use of a higher cost for low/intermediate-1 MF than for intermediate-2/high-risk MF was also considered reasonable due to the longer duration of treatment (median survival approximately 5-8 years and 1-3 years respectively).

5 COST EFFECTIVENESS RESULTS

5.1 Company's base case cost-effectiveness results for the primary analysis

The company report the deterministic base case results from their primary STM model in CS Table 39 for the licensed populations with and without splenomegaly (reproduced in Table 23 below). These and other results in this report use the current Patient Access Scheme (PAS) price for ruxolitinib ([redacted] price discount) agreed as part of the MF submission to NICE TA386,[15] with list prices used for all other drugs. Results with confidential discounts for comparator and concurrent medications are provided in a separate confidential addendum to this report.

Table 23 Company base case results: primary analysis

Treatment	Total			Incremental			ICER (£/QALY)
	Cost	LYG[a]	QALYs	Cost	LYG[a]	QALYs	
Licensed population with splenomegaly (RESPONSE trial population)							
BAT	£92,017	9.28	6.97	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	2.17	[redacted]	[redacted]
Licensed population without splenomegaly (RESPONSE-2 trial population)							
BAT	£86,809	10.46	7.80	-	-	-	-

Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	1.79	[redacted]	[redacted]

Source: Reproduced from CS Table 39.
BAT best available therapy; LYG life years gained; QALYs quality-adjusted life years; ICER incremental cost-effectiveness ratio.
[a] Note: life years gained are not discounted.

The base case results for the primary analysis show that for the licensed population without splenomegaly, ruxolitinib offers a mean QALY gain of [redacted] for an additional mean cost of [redacted] compared with BAT, producing an ICER of [redacted] per QALY gained. For the licensed population with splenomegaly, ruxolitinib provides a QALY gain of [redacted] for an additional cost of [redacted] against BAT, which results in an ICER of [redacted] per QALY gained.

5.1.1 Deterministic sensitivity analyses for the company's base case for the primary analysis

The company report deterministic sensitivity analysis results for the ten most influential parameters in CS Figure 53. The ranges of variation for the input parameters were based on 95% confidence intervals where available, or a range of +/- 20%. The company's results indicate that the assumptions regarding the treatment effect for OS are the main drivers of the model results for the primary analysis, increasing the ICER to [redacted] and [redacted] per QALY for the licensed population with and without splenomegaly, respectively. The discount rates for both cost and benefits and assumptions regarding utility values also have a notable impact on the ICER for the primary analysis.

5.1.2 Scenario analyses for the company's base case for the primary analysis

The company consider almost 100 scenarios for the primary analysis (see CS Appendix P) and report the top 20 most impactful scenarios in CS Figure 54.

Licensed population with splenomegaly (RESPONSE population)

Changing the source of the treatment effect (HR OS) from the MAJIC-PV trial to the retrospective analysis of Spanish registry data (Alvarez-Larrán et al. 2022) had the largest impact on the ICER, increasing the ICER to [redacted] per QALY, whilst limiting the treatment effect to 5 years has the second-largest effect, causing the ICER to rise to [redacted] per QALY.[37] Of the 20 scenarios provided in the CS, the top seven scenarios that increase the ICER the most involve the source of the treatment effect, treatment effect waning, and the time horizon. Using the treatment effect from ITC comparison with GEMFIN results in the lowest ICER per QALY, at [redacted]. We note that the CS did not report results for the scenario with a constant HR OS from the MAJIC-PV trial, but this was provided in response to

clarification question B2. This scenario increased the company's base case ICER for RESPONSE population to [redacted] per QALY.

Licensed population without splenomegaly (RESPONSE-2 population)

Limiting the treatment effect to 5 years resulted in the highest impact on the ICER, which increases to [redacted] per QALY; the second-largest effect arose by implementing Alvarez-Larrán et al. 2022 as the source of the treatment effect, giving an ICER of [redacted] per QALY.[37] As with the licensed population with splenomegaly, the top seven scenarios causing the highest increase in ICERs involved the source of treatment effect, treatment effect waning, and the time horizon. Also in line with the licensed population with splenomegaly, applying the ITC treatment effect from the comparison with GEMFIN rather than MAJIC-PV gives the greatest reduction in the ICER at [redacted] per QALY. The ICER for the scenario with the constant MAJIC-PV HR for the RESPONSE-2 population was [redacted] per QALY (company response to clarification question B2).

5.1.3 Probabilistic sensitivity analysis for the company's base case for the primary analysis

The company conducted a probabilistic sensitivity analysis (PSA) with input parameter distributions as presented in CS Table 37. They used appropriate probability distributions for the different parameters. An arbitrary SE of 10% was assumed where the SE was not reported, namely for the QALY loss for key events, management costs and end of life cost.

The results from 2,000 iterations are reported in CS Table 41, and CS Figure 52 illustrates the extent of uncertainty around the results with cost-effectiveness scatterplots and cost-effectiveness acceptability curves (CEACs). The EAG confirm that the probabilistic results for the licensed population either with or without splenomegaly are similar to the deterministic results. The estimated probability that ruxolitinib meets a cost-effectiveness threshold of £30,000 per QALY gained at the current PAS price for ruxolitinib [redacted] for both subgroups, with and without splenomegaly.

5.2 Company's base case cost-effectiveness results for the MAJIC-PV population

The company reports the results for the MAJIC-PV population in CS Table 42, reproduced in Table 24 below. This shows an estimated QALY gain of [redacted] and additional cost of [redacted] for ruxolitinib in comparison with current clinical management, resulting in an ICER of [redacted] per QALY gained.

Table 24 Company base case results: MAJIC-PV population

Treatment	Total			Incremental			ICER (£/QALY)
	Cost	LYG*	QALYs	Cost	LYG*	QALYs	
BAT	£83,317	8.02	6.11	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	1.63	[redacted]	[redacted]

Reproduced from CS Table 42.
Best available therapy; LYG life years gained; QALYs quality-adjusted life years; ICER incremental cost-effectiveness ratio.
*Note: life years gained are not discounted.

5.2.1 Deterministic sensitivity analyses for the company's base case for the MAJIC-PV population

The company illustrate the results of the ten most influential parameters from their deterministic sensitivity analyses in CS Figure 56. As with the primary analysis, the company's results show that the model is most sensitive to the treatment effect for OS, with the ICER increasing to [redacted] per QALY at the upper limit for the HR in the second time period (year 3-5). The discount rates for costs and benefits are also influential parameters for the MAJIC-PV population, as well as the hazard rate for the time to treatment discontinuation for ruxolitinib.

5.2.2 Scenario analysis for the company's base case for the MAJIC-PV population

The company report the results of the top 20 most impactful scenarios in CS Figure 56. Restricting the treatment effect to 5 years has the largest effect on the results, increasing the ICER to [redacted] per QALY, and implementing the treatment effect for OS reported by Alvarez-Larrán et al. 2022 produced the next-highest ICER of [redacted] per QALY.[37] Again, in line with the primary analysis, the most influential scenarios involve the treatment effect for OS and treatment effect waning, with the greatest reduction in the ICER obtained by from the ITC comparison with GEMFIN ([redacted] per QALY). The scenario with the constant HR estimated from the MAJIC-PV trial increase the ICER to [redacted] per QALY (company response to clarification question B2).

5.2.3 Probabilistic sensitivity analysis for the company's base case for the MAJIC-PV population

Probabilistic results for the MAJIC-PV population are provided in CS Table 44 and CS Figure 55. The EAG confirm that the probabilistic results for the MAJIC-PV population are similar to the deterministic results. As with the base case results, the probability that the ICER is below £30,000 per QALY gained is [redacted].

5.3 Model validation and face validity check

5.3.1 Company's model validation

The company state their approach to model validation in CS Section B.3.13. They report that two advisory board meetings were held with five clinical experts with experience in the management of patients with PV resistant or intolerant to hydroxycarbamide.[58 59]

The EAG note that the first advisory meeting, conducted on 24th June 2022, comprised only four clinical experts; however, the second cited advisory meeting took place over two dates (28th July 2022 and 8th August 2022) with five experts present.[58 59] Four of the five clinical experts who attended the advisory meetings are authors of the MAJIC-PV trial.

The model structure and appropriateness to the decision problem were discussed and validated with the clinical experts in these meetings, as well as the validity of model inputs such as costs and utilities. The company also report that a health economist, not involved in the development of the model, reviewed the model for coding errors, inconsistencies, and plausibility of inputs, and also subjected the model to stress testing of extreme scenarios to detect modelling errors.

The company note the following points:

- Long term predictions could not be compared against external data as long term data for the patient population are not available.
- Predicted life years for the licensed population without splenomegaly was higher compared to the licensed pop with splenomegaly, despite using different model structures and inputs. This is in line with clinical expectations.
- Predicted life years for the MAJIC-PV population were lower compared to estimates from the primary analysis for the RESPONSE and RESPONSE-2 trial populations. This reflects the poorer prognosis of the MAJIC-PV population.
- Prediction for the MAJIC-PV population also aligns with that observed in Alvarez-Larrán et al. 2022.[37]

5.3.2 EAG model validation

5.3.2.1 EAG verification procedures

The EAG conducted a series of quality checks on the company model, assessing its transparency and validity. A range of tests were performed to verify model inputs, calculations, and outputs:

- Cross-checking all parameter inputs against values reported in the CS, model, and cite sources
- Checking all model outputs against results stated in the CS, including the base case, PSA, DSA, and company scenarios for both the primary and MAJIC-PV population analyses
- Checking the individual formulae within the model
- Manually running scenarios and verifying model outputs against results reported in the CS and appendices for the DSA and scenario analyses
- Applying a range of extreme value and logic tests to check the plausibility of changes in results when parameters are changed
- Checking Visual Basic (VBA) code for errors, and re-running the code to ensure expected outputs were produced.

The model is well implemented and no coding errors were identified, however the EAG considers the failure to apply a general population mortality constraint to pre-discontinuation mortality within the 5-year trial period to be an error (see section 5.3.3.1 below).

The EAG identified several discrepancies between parameter values cited in the CS and the values used in the model (clarification questions B5, B11 to B18 and C3). The company confirmed that in all cases these related to errors in the description of model inputs in the CS, and that the correct values had been used in the model. Note also that the company confirmed that the columns in the table of scenario analyses in CS Appendix P are incorrectly labelled (clarification question C4).

5.3.2.2 Comparison of company extrapolations with trial and cohort data

Figure 3 and Figure 4 below present the model predictions for overall survival and time to treatment discontinuation for ruxolitinib and BAT for the licensed population with and without splenomegaly, respectively.

Figure 3 Predicted OS and TTD for ruxolitinib and BAT for the licensed population with splenomegaly

Abbreviations: OS: overall survival; TTD: time to treatment discontinuation; BAT: best available therapy; KM: Kaplan-Meier; Gen pop: general population mortality; ICER: incremental cost-effectiveness ratio; Rux: ruxolitinib.
Source: Reproduced from CS Appendix J Figure 14.

Figure 4 Predicted OS and TTD for ruxolitinib and BAT for the licensed population without splenomegaly

Abbreviations: OS: overall survival; TTD: time to treatment discontinuation; BAT: best available therapy; KM: Kaplan-Meier; Gen pop: general population mortality; ICER: incremental cost-effectiveness ratio; Rux: ruxolitinib.
Source: Reproduced from CS Appendix J Figure 13.

EAG comments on extrapolation distributions

- The company selected an odds spline model with one knot for the extrapolation of TTD for ruxolitinib due to reasons other than death in the primary analysis. The company make note of the potential for spline models with more than one knot to overfit the data. The EAG opt for a standard parametric distribution, the Weibull distribution, in our preferred assumptions to remove the uncertainty around spline models and utilise a more conservative approach.
- The remaining distributions chosen by the company are deemed appropriate by the EAG. Scenario analyses showing outcomes of selected distributions for OS and TTD for the primary analysis and the MAJIC-PV population analysis are provided in section 6.1.

5.3.3 Corrections to the company model

5.3.3.1 General population mortality constraint for pre-discontinuation survival

In the company's analyses for the RESPONSE and RESPONSE-2 populations, pre-discontinuation survival for ruxolitinib is only adjusted for general population mortality after the 5-year period of trial observation, which results in better predicted survival while patients remain on ruxolitinib than for people in the general population of the same age. The EAG raised this anomaly as a clarification question (B4), and the company provided an updated version of the model with an option to adjust pre-discontinuation survival for general population mortality over the entire time horizon. The ICERs for the RESPONSE and RESPONSE-2 populations with this adjustment were reported as a scenario analysis in Table 4 in the company's clarification response. We consider this a correction, as it is not plausible that people with PV would have better survival than the general population.

Full cost-effectiveness results for the company's primary base case analyses with the general population mortality correction applied are shown in Table 25 below. We use this correction in EAG additional in section 6.2. Note that as pre-discontinuation survival for ruxolitinib is only implemented in the primary analysis, the results for the MAJIC-PV population are unaffected.

Table 25 Company scenario analysis with the general population mortality constraint for pre-discontinuation survival: primary analysis

Treatment	Total			Incremental			ICER
	Cost	LYG[a]	QALYs	Cost	LYG[a]	QALYs	(£/QALY)
Licensed population with splenomegaly (RESPONSE trial population)							

BAT	£89,098	8.97	6.73	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	2.20	[redacted]	[redacted]
Licensed population without splenomegaly (RESPONSE-2 trial population)							
BAT	£82,203	9.88	7.37	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	1.87	[redacted]	[redacted]

Source: Company response to clarification question B4 and EAG analysis with company's model
BAT best available therapy; LYG life years gained; QALYs quality-adjusted life years; ICER incremental cost-effectiveness ratio.
[a] Note: life years gained are not discounted.

5.3.3.2 EAG scenario analysis for the cost of a grade 1-2 thromboembolic event

The company assumed a cost for management of all Grade 1-2 thromboembolic events of £297, equivalent to the cost of one emergency department visit. However, EAG clinical expert advisers noted that a D-dimer test and a vascular ultrasound would also be required to investigate a suspected thromboembolic event, as well as a single low-dose of an anti-coagulant (as per the NICE guideline NG158).[57] For the EAG analysis, we include the cost of a laboratory D-dimer test at £6.79 (NG158),[57] a single dose of enoxaparin sodium at £8.79, (BNF 2022)[8] and a vascular ultrasound costing £96.99 (NHS Reference costs 2020/21)[60]. This results in a small reduction in the ICERs (see Table 27 below).

Table 26 EAG scenario analysis for cost of grade 1-2 thromboembolic event

Treatment	Total			Incremental			ICER
	Cost	LYG[a]	QALYs	Cost	LYG[a]	QALYs	(£/QALY)
Licensed population with splenomegaly (RESPONSE trial population)							
BAT	£92,035	9.28	6.97	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	2.17	[redacted]	[redacted]
Licensed population without splenomegaly (RESPONSE-2 trial population)							
BAT	£86,849	10.46	7.80	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	1.79	[redacted]	[redacted]
MAJIC-PV population							
BAT	£83,339	8.02	6.11	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	1.63	[redacted]	[redacted]

Source: Company response to clarification question B4 and EAG analysis with company's model
BAT best available therapy; LYG life years gained; QALYs quality-adjusted life years; ICER incremental cost-effectiveness ratio.
[a] Note: life years gained are not discounted.

5.3.4 EAG summary of key issues and additional analyses

The company summarise and justify assumptions in their primary and subgroup (MAJIC-PV population) economic analyses in CS Table 38. We highlight key areas of uncertainty and

the rationale for additional EAG analyses in Appendix 9.5. Section 6.2 details the EAG's preferred assumptions and subsequent cost-effectiveness results. Additional scenario analyses are conducted on the EAG base case model in section 6.2.2.

6 EAG ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the EAG

Table 27 below shows cost-effectiveness results for selected company scenarios for the primary analysis for the licensed population with and without splenomegaly (RESPONSE and RESPONSE-2, respectively). As there are a large number of scenarios reported in CS Appendix P, we have selected scenarios that relate to key uncertainties and that have an impact on the ICERs.

Table 27 Selected scenarios applied to the company base case: primary analysis

Scenario	Treatment	RESPONSE			RESPONSE-2		
		Cost	QALYs	ICER	Cost	QALYs	ICER
Company base case	BAT	£92,017	6.97		£86,809	7.80	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HR OS: MAJIC-PV constant	BAT	£102,301	7.78		£94,479	8.52	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HR OS: pooled RESPONSEtrials	BAT	£103,377	7.86		£95,125	8.58	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HR OS: Alvarez-Larrán 2022	BAT	£105,234	8.01		£96,237	8.68	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HR OS: matched GEMFIN (ITC)	BAT	£75,644	5.66		£77,734	6.95	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
No BAT partition	BAT	£94,485	7.04		£89,043	7.87	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D utilities	BAT	£92,017	6.47		£86,809	7.22	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Faster waning: 5 to 10 years	BAT	£98,816	7.50		£92,756	8.35	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Slower waning: 5 to 50 years	BAT	£86,097	6.50		£81,321	7.29	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time horizon 30 years	BAT	£91,122	6.91	[redacted]	£86,368	7.77	
	Ruxolitinib	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
Ruxolitinib TTD lognormal	BAT	£94,803	7.18		£92,185	8.30	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ruxolitinib TTD loglogistic	BAT	£93,096	7.05		£90,099	8.11	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ruxolitinib TTD Weibull	BAT	£90,683	6.86		£88,983	8.00	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ruxolitinib TTD hazard spline 1	BAT	£90,118	6.82		£85,402	7.67	
	Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Ruxolitinib TTD Exponential	BAT Ruxolitinib	£85,860 [redacted]	6.48 [redacted]	[redacted]	£86,257 [redacted]	7.75 [redacted]	[redacted]
Remove impact of key events	BAT Ruxolitinib	£56,318 [redacted]	7.03 [redacted]	[redacted]	£63,023 [redacted]	7.90 [redacted]	[redacted]

Source: EAG analysis using company model and scenario analyses.
QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; HR: hazard ratio; OS: overall survival; ITC: indirect treatment comparison; TTD: time to treatment discontinuation.

Figure 5 and Figure 6 below show the KM data with the company's choice of distribution for TTD for ruxolitinib due to reasons other than death in comparison with the selected scenario distributions from Table 27 above for the licensed population with and without splenomegaly.

[redacted]

Figure 5 Comparison of selected scenario distributions for TTD for ruxolitinib for the licensed population with splenomegaly

Abbreviations: KM: Kaplan-Meier; TTD: time to treatment discontinuation; OS: overall survival.
Source: Reproduced from CS Appendix N Figure 18 using selected distributions.

Figure 6 Comparison of selected scenario distributions for TTD for ruxolitinib for the licensed population without splenomegaly

Abbreviations: KM: Kaplan-Meier; TTD: time to treatment discontinuation; OS: overall survival.
Source: Reproduced from CS Appendix N Figure 19 using selected distributions.

Table 28 below shows cost-effectiveness results for selected company scenarios for the MAJIC-PV population analysis. Again, from the many scenarios conducted by the company, we have selected scenarios that relate to key uncertainties and that have an impact on the ICERs.

Table 28 Selected scenarios applied to the company base case: MAJIC-PV population

Scenario	Treatment	Cost	QALYs	ICER
Company base case	BAT	£83,317	6.11	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
HR OS: MAJIC-PV constant	BAT	£83,317	6.11	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
HR OS: Pooled RESPONSE-trials	BAT	£83,317	6.11	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
HR OS: Alvarez-Larrán 2022	BAT	£83,317	6.11	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
HR OS: matched GEMFIN (ITC)	BAT	£83,317	6.11	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
EQ-5D utility values	BAT	£83,317	5.71	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
Faster waning: 5 to 10 years	BAT	£83,317	6.11	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
Slower waning: 5 to 50 years	BAT	£83,317	6.11	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
BAT OS: lognormal	BAT	£101,095	7.43	

	Ruxolitinib	[redacted]	[redacted]	[redacted]
BAT OS: loglogistic	BAT	£94,943	6.97	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
BAT OS: hazard spline 1	BAT	£98,348	7.23	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
BAT OS: Gompertz	BAT	£70,476	5.13	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
Time horizon: 30 years	BAT	£83,250	6.10	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
Remove impact of key events	BAT	£57,187	6.18	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
Source: EAG analysis using company model and scenario analyses. QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; HR: hazard ratio; OS: overall survival; ITC: indirect treatment comparison.				

Figure 7 shows the KM data for overall survival for BAT in the MAJIC-PV population analysis in comparison with the company's chosen Weibull distribution and selected scenario distributions from Table 28 above.

Figure 7 Comparison of KM with company base case distribution and selected scenario distributions for overall survival for BAT for the MAJIC-PV population analysis

Abbreviations: KM: Kaplan-Meier; BAT: best available therapy; OS: overall survival.
Source: Reproduced from CS Appendix N Figure 25 using selected distributions.

From the above tables, it is evident that the source of treatment effect for overall survival has a great impact on the ICER, with the exception of the hazard ratio derived from the ITC. As expected, reducing and increasing the treatment waning period also effects the ICER.

Although the company implemented an extended time horizon of 46 years for patients starting in the model at age 66, a 30-year time horizon has minimal effect on the ICER.

6.2 EAG's preferred assumptions

Based on the critique of the company's model, the EAG have identified the following preferred model assumptions:

- Correction for general population mortality for pre-discontinuation survival in the primary analysis
- Weibull distribution for extrapolation of TTD for ruxolitinib due to reasons other than death in the primary analysis
- A constant hazard ratio derived from the MAJIC-PV trial for overall survival
- No partitioning of the BAT health state in the primary analysis
- EQ-5D utility values
- EAG estimated cost assumed for the management of Grade 1-2 thromboembolic events.

6.2.1 Results using the EAG preferred model assumptions

The results for this analysis for the three trial populations are shown in Table 29 below. We also report cumulative analyses for the three populations in Table 30, Table 31, and Table 32 below, showing the progression from the company's base case model to the EAG base case model by applying EAG preferred assumptions one at a time.

Table 29 EAG preferred analysis results

Treatment	Total			Incremental			ICER
	Cost	LYG[a]	QALYs	Cost	LYG[a]	QALYs	(£/QALY)
RESPONSE trial population (with splenomegaly)							
BAT	£100,281	9.90	7.02	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	1.09	[redacted]	[redacted]
RESPONSE-2 trial population (without splenomegaly)							
BAT	£93,866	11.08	7.77	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	0.91	[redacted]	[redacted]
MAJIC-PV trial population							
BAT	£83,339	8.02	5.71	-	-	-	-
Ruxolitinib	[redacted]	[redacted]	[redacted]	[redacted]	0.92	[redacted]	[redacted]

Source: EAG analysis using the company's model
BAT best available therapy; LYG: life years gained; QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio.
[a] Note: life years gained are not discounted.

Table 30 Cumulative changes from the company base case model to the EAG preferred analysis: RESPONSE trial population (with splenomegaly)

Assumption	Treatment	RESPONSE		
		Cost	QALYs	ICER
Company base case	BAT	£92,017	6.97	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ General population mortality constraint	BAT	£89,098	6.73	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ Ruxolitinib TTD: Weibull	BAT	£87,837	6.64	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ HR OS: MAJIC-PV constant	BAT	£97,696	7.42	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ No BAT partition	BAT	£100,262	7.49	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ EQ-5D utilities	BAT	£100,262	7.02	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ Cost for Grade 1-2 TE events (EAG preferred analysis)	BAT	£100,281	7.02	
	Ruxolitinib	[redacted]	[redacted]	[redacted]

Source: EAG analysis using the company's model.
QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; TTD: time to treatment discontinuation; HR: hazard ratio; OS: overall survival; TE: thromboembolic.

Table 31 Cumulative changes from the company base case model to the EAG preferred analysis: RESPONSE-2 trial population (without splenomegaly)

Assumption	Treatment	RESPONSE-2		
		Cost	QALYs	ICER
Company base case	BAT	£86,809	7.80	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ General population mortality constraint	BAT	£82,203	7.37	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ Ruxolitinib TTD: Weibull	BAT	£84,052	7.54	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ HR OS: MAJIC-PV constant	BAT	£91,411	8.23	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ No BAT partition	BAT	£93,824	8.30	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ EQ-5D utilities	BAT	£93,824	7.77	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ Cost for Grade 1-2 TE events (EAG preferred analysis)	BAT	£93,866	7.77	
	Ruxolitinib	[redacted]	[redacted]	[redacted]

Source: EAG analysis using the company's model.
QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; TTD: time to treatment discontinuation; HR: hazard ratio; OS: overall survival; TE: thromboembolic.

Table 32 Cumulative changes from the company base case model to the EAG preferred analysis: MAJIC-PV trial population

Assumption	Treatment	Cost	QALYs	ICER
Company base case	BAT Ruxolitinib	£83,317 [redacted]	6.11 [redacted]	[redacted]
+ HR OS: MAJIC-PV constant	BAT Ruxolitinib	£83,317 [redacted]	6.11 [redacted]	[redacted]
+ EQ-5D utilities	BAT Ruxolitinib	£83,317 [redacted]	5.71 [redacted]	[redacted]
+ Cost for Grade 1-2 TE events (EAG preferred analysis)	BAT Ruxolitinib	£83,339 [redacted]	5.71 [redacted]	[redacted]

Source: EAG analysis using the company's model.
QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; HR: hazard ratio; OS: overall survival; TE: thromboembolic.

6.2.2 Scenario analyses conducted on the EAG base case model

Table 33 and Table 34 below show selected scenario analyses applied to the EAG preferred analysis for the primary analysis (RESPONSE and RESPONSE-2 populations) and for the MAJIC-PV population analysis respectively. The scenarios included in these tables include company base case assumptions, as well as scenarios chosen to illustrate key uncertainties.

Table 33 Scenario analyses on the EAG base case model: primary analysis

Scenario	Treatment	RESPONSE			RESPONSE-2		
		Cost	QALYs	ICER	Cost	QALYs	ICER
EAG base case	BAT Ruxolitinib	£100,281 [redacted]	7.02 [redacted]	[redacted]	£93,866 [redacted]	7.77 [redacted]	[redacted]
Ruxolitinib TTD odds spline 1	BAT Ruxolitinib	£101,830 [redacted]	7.13 [redacted]	[redacted]	£92,133 [redacted]	7.62 [redacted]	[redacted]
HR OS: MAJIC-PV time varying	BAT Ruxolitinib	£90,278 [redacted]	6.28 [redacted]	[redacted]	£86,499 [redacted]	7.13 [redacted]	[redacted]
BAT partition	BAT Ruxolitinib	£97,714 [redacted]	6.88 [redacted]	[redacted]	£91,454 [redacted]	7.61 [redacted]	[redacted]
MF-8D utilities	BAT Ruxolitinib	£100,281 [redacted]	7.49 [redacted]	[redacted]	£93,866 [redacted]	8.30 [redacted]	[redacted]
Company Grade 1-2 TE costs	BAT Ruxolitinib	£100,262 [redacted]	7.02 [redacted]	[redacted]	£93,824 [redacted]	7.77 [redacted]	[redacted]
Waning from year 5 to 10	BAT Ruxolitinib	£103,118 [redacted]	7.22 [redacted]	[redacted]	£96,080 [redacted]	7.96 [redacted]	[redacted]
Waning from year 5 to 30	BAT Ruxolitinib	£98,782 [redacted]	6.91 [redacted]	[redacted]	£92,542 [redacted]	7.66 [redacted]	[redacted]
Waning from year 5 to 50	BAT Ruxolitinib	£97,525 [redacted]	6.82 [redacted]	[redacted]	£91,424 [redacted]	7.56 [redacted]	[redacted]
Time horizon 30 years	BAT Ruxolitinib	£99,178 [redacted]	6.96 [redacted]	[redacted]	£93,194 [redacted]	7.73 [redacted]	[redacted]

Remove impact of key events	BAT Ruxolitinib	£62,184 [redacted]	7.09 [redacted]	[redacted]	£68,639 [redacted]	7.87 [redacted]	[redacted]

Source: EAG analysis using the company's model.
QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; TTD: time to treatment discontinuation; HR: hazard ratio; OS: overall survival; TE: thromboembolic.

Table 34 Scenario analyses on the EAG base case model: MAJIC-PV population analysis

Scenario	Treatment	Cost	QALYs	ICER
EAG base case	BAT Ruxolitinib	£83,339 [redacted]	5.71 [redacted]	[redacted]
HR OS: MAJIC-PV time-varying	BAT Ruxolitinib	£83,339 [redacted]	5.71 [redacted]	[redacted]
MF-8D utilities	BAT Ruxolitinib	£83,339 [redacted]	6.11 [redacted]	[redacted]
Company Grade 1-2 TE costs	BAT Ruxolitinib	£83,317 [redacted]	5.71 [redacted]	[redacted]
Waning from year 5 to 10	BAT Ruxolitinib	£83,339 [redacted]	5.71 [redacted]	[redacted]
Waning from year 5 to 30	BAT Ruxolitinib	£83,339 [redacted]	5.71 [redacted]	[redacted]
Waning from year 5 to 50	BAT Ruxolitinib	£83,339 [redacted]	5.71 [redacted]	[redacted]
BAT OS: lognormal	BAT Ruxolitinib	£101,122 [redacted]	6.96 [redacted]	[redacted]
BAT OS: loglogistic	BAT Ruxolitinib	£94,968 [redacted]	6.52 [redacted]	[redacted]
BAT OS: hazard spline 1	BAT Ruxolitinib	£98,374 [redacted]	6.77 [redacted]	[redacted]
BAT OS: Gompertz	BAT Ruxolitinib	£70,494 [redacted]	4.80 [redacted]	[redacted]
Time horizon: 30 years	BAT Ruxolitinib	£83,271 [redacted]	5.71 [redacted]	[redacted]
Remove impact of key events	BAT Ruxolitinib	£57,187 [redacted]	5.78 [redacted]	[redacted]

Source: EAG analysis using the company's model.
QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; HR: hazard ratio; OS: overall survival; TE: thromboembolic.

6.3 Conclusions on the cost effectiveness evidence

The company's model generated base case ICERs of [redacted], [redacted], and [redacted] per QALY for the licensed populations with and without splenomegaly and the MAJIC-PV population analysis, respectively. In response to clarification question B4, the company performed

scenario analyses adjusting pre-discontinuation survival for general population mortality for the entire time horizon for the primary analysis. These scenarios produced ICERs of [redacted] and [redacted] for the licensed population with and without splenomegaly, respectively. The EAG considers this scenario as a correction (see section 5.3.3 above).

The EAG preferred model assumptions are the following:

- Correction to include the general population mortality constraint for pre-discontinuation survival throughout the time horizon (primary analysis)
- Weibull distribution for extrapolation of TTD for ruxolitinib due to reasons other than death, as we consider that this provides a better fit to the data than the odds spline model with one know that the company used (primary analysis)
- Treatment effect estimated using the constant HR estimate for OS, as reported by the MAJIC-PV trial investigators
- No partitioning of the BAT health state (primary analysis)
- EQ-5D utility values
- EAG estimated cost assumed for the management of Grade 1-2 thromboembolic events.

The EAG's correction and preferred assumptions increase the ICER to [redacted] per QALY for the licensed population with splenomegaly, [redacted] per QALY for the licensed population without splenomegaly, and [redacted] per QALY for the MAJIC-PV population analysis. These estimates are most sensitive to the assumptions regarding the source of treatment effect for overall survival and the source of utility values.

Alternative assumptions about the waning of the treatment effect also affect the ICER, and we note that EAG clinical advisors have suggested that they do not have reason to expect that the effectiveness of ruxolitinib would wane over time.

We also report a scenario removing the QALY loss and costs for major complications of PV to illustrate the impact of the way in which this has been modelled, not because we believe that it might be appropriate to exclude these impacts.

7 SEVERITY MODIFIERS

The company state that the QALY shortfall criteria for severity weighting, as defined in the 2022 NICE health technology evaluations manual,[42] are not met (CS B.3.6 and Table 36).

We show the absolute and proportional QALY shortfalls for the populations based on the company's base case analyses and EAG preferred assumptions in Table 35 below. The criteria for severity weighting are not met under the EAG's preferred assumptions.

Table 35 QALY shortfall analysis

Model (population)	Expected total QALYs [a]		QALY shortfall	
	General population [b]	Model	Absolute	Proportional
Company base case				
STM (RESPONSE population)	12.60	6.97	5.63	0.45
STM (RESPONSE-2 population)	11.13	7.80	3.32	0.30
PSM (MAJIC-PV population)	10.55	6.11	4.45	0.42
EAG preferred assumptions				
STM (RESPONSE population)	12.60	7.02	5.59	0.44
STM (RESPONSE-2 population)	11.13	7.77	3.36	0.30
PSM (MAJIC-PV population)	10.55	5.71	4.84	0.46

STM: state-transition model; PSM: partitioned survival model

[a] Discounted QALYs over the model time horizon (46 years from starting age)

[b] General population utilities by age and sex from Hernández Alava et al. 2022[51]

Source: Adapted from CS Table 36, with results for the EAG preferred analysis calculated from the company's model

8 REFERENCES

1. Stein BL, Moliterno AR, Tiu RV. Polycythemia vera disease burden: contributing factors, impact on quality of life, and emerging treatment options. *Annals of Hematology* 2014;93(12):1965-76. doi: 10.1007/s00277-014-2205-y [published Online First: 20141002]
2. MPN Voice. Polycythaemia vera 2020 [Available from: https://www.mpnvoice.org.uk/about-mpns/questions/polycythaemia-vera/ accessed 24/10/2022.
3. MPN Voice. About MPNs 2021 [Available from: https://www.mpnvoice.org.uk/about-mpns/ accessed 24/10/2022.
4. McMullin MF, Harrison CN, Ali S, et al. A guideline for the diagnosis and management of polycythaemia vera. A British Society for Haematology Guideline. *British Journal of Haematology* 2019;184(2):176-91.
5. The Haematological Malignancy Research Network (HMRN). Factsheets: Polycythaemia vera 2022 [Available from: www.hmrn.org/factsheets#polycythaemia_vera accessed 24/10/2022.
6. Harrison C, Sekhar M, Mitchell C, et al. Pan-London Haemato-Oncology Clinical Guidelines. Acute leukaemias and myeloid neoplasms. Part 4: Myeloproliferative neoplasms. London: RM Partners, South East London Cancer Alliance, North Central and East London Cancer Alliance, 2020.

7. Vannucchi AM, Barbui T, Cervantes F, et al. Philadelphia chromosome-negative chronic myeloproliferative neoplasms: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. *Annals of Oncology* 2015;26:v85-v99. doi: https://doi.org/10.1093/annonc/mdv203
8. BNF. British National Formulary: Joint Formulary Committee; 2022 [Available from: http://www.medicinescomplete.com accessed 05 August 2022.
9. Marchetti M, Vannucchi AM, Griesshammer M, et al. Appropriate management of polycythaemia vera with cytoreductive drug therapy: European LeukemiaNet 2021 recommendations. *The Lancet Haematology* 2022;9(4):e301-e11.
10. Barosi G, Birgegard G, Finazzi G, et al. A unified definition of clinical resistance and intolerance to hydroxycarbamide in polycythaemia vera and primary myelofibrosis: results of a European LeukemiaNet (ELN) consensus process. *British Journal of Haematology* 2010;148(6):961-63. doi: 10.1111/j.1365-2141.2009.08019.x [published Online First: 2009/11/26]
11. Medicines and Healthcare products Regulatory Agency (MHRA). Jakavi® 10 mg tablets: Summary of Product Characteristics. Available at: https://mhraproducts4853.blob.core.windows.net/docs/d61f7a73d3447fbfcf6d3936bed37a80946f480f [Last Accessed 08 September 2022] 2022 [
12. Medicines and Healthcare products Regulatory Agency (MHRA). Jakavi® 5 mg tablets: Summary of Product Characteristics. 2022 [Available from: https://mhraproducts4853.blob.core.windows.net/docs/9b08c7b88ef11cc23d00d10cd3f73959b27ef099 accessed 31/10/2022.
13. Medicines and Healthcare products Regulatory Agency (MHRA). Jakavi® 15 mg tablets: Summary of Product Characteristics. 2022 [Available from: https://mhraproducts4853.blob.core.windows.net/docs/8fab59846945bb488753c25883e9e9131e288eac accessed 31/10/2022.
14. Medicines and Healthcare products Regulatory Agency (MHRA). Jakavi® 20 mg tablets: Summary of Product Characteristics. 2022 [Available from: https://mhraproducts4853.blob.core.windows.net/docs/6667a689dcc1eaf7cd7e8a5d11c7b1029da2c80a accessed 31/10/2022.
15. National Institute for Health and Care Excellence. TA386: Ruxolitinib for disease-related splenomegaly or symptoms in adults with myelofibrosis (review of TA289). 2016 [Available from: https://www.nice.org.uk/guidance/TA386 accessed 05 August 2022.
16. Harrison C, Nangalia J, Boucher R, et al. Ruxolitinib versus Best Available Therapy for PV intolerant or resistant to Hydroxycarbamide in a Randomized trial (Manuscript in Development). *Confidential unpublished manuscript* 2022
17. Pardanani A, Harrison C, Vannucchi A, et al. MPN-184: Safety and Efficacy of Fedratinib, an Oral, Selective JAK2 Inhibitor, in Patients with Polycythemia Vera (PV) Resistant or Intolerant to Hydroxyurea: Results from a Phase II Dose-Ranging and Dose-Expansion Study. *Clinical Lymphoma, Myeloma and Leukemia* 2020;20:S332-S33.
18. NCT00928707. Phase II Study of GIVINOSTAT (ITF2357) in Combination With Hydroxyurea in Polycythemia Vera (PV): ClinicalTrials.gov; [Available from: Https://clinicaltrials.gov/ct2/show/NCT00928707 accessed 08/09/2022.
19. NCT00726232. Study to Determine the Safety and Efficacy of INCB018424 in Patients With Polycythemia Vera or Essential Thrombocythemia.: ClinicalTrials.gov; [Available from: https://clinicaltrials.gov/ct2/show/NCT00726232 accessed 08/09/2022.
20. NCT01632904. Randomized Switch Study From Hydroxyurea to Ruxolitinib for RELIEF of Polycythemia Vera Symptoms: the Relief Study (NCT01632904). ClinicalTrials.gov; [Available from: https://clinicaltrials.gov/ct2/show/NCT01632904 accessed 09/09/2022.
21. Mesa R, Vannucchi AM, Yacoub A, et al. The efficacy and safety of continued hydroxycarbamide therapy versus switching to ruxolitinib in patients with

polycythaemia vera: a randomized, double-blind, double-dummy, symptom study (RELIEF). *British Journal of Haematology* 2017;176(1):76-85.
22. Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera. *New England Journal of Medicine* 2015;372(5):426-35.
23. Passamonti F, Griesshammer M, Palandri F, et al. Ruxolitinib for the treatment of inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): a randomised, open-label, phase 3b study. *The Lancet Oncology* 2017;18(1):88-99.
24. Kiladjian J-J, Zachee P, Hino M, et al. Long-term efficacy and safety of ruxolitinib versus best available therapy in polycythaemia vera (RESPONSE): 5-year follow up of a phase 3 study. *The Lancet Haematology* 2020;7(3):e226-e37. doi: https://doi.org/10.1016/S2352-3026(19)30207-8
25. Barosi G, Birgegard G, Finazzi G, et al. Response criteria for essential thrombocythemia and polycythemia vera: result of a European LeukemiaNet consensus conference. *Blood* 2009;113(20):4829-33. doi: 10.1182/blood-2008-09-176818
26. Barosi G, Tefferi A, Besses C, et al. Clinical end points for drug treatment trials in BCR-ABL1-negative classic myeloproliferative neoplasms: consensus statements from European LeukemiaNET (ELN) and Internation Working Group-Myeloproliferative Neoplasms Research and Treatment (IWG-MRT). *Leukemia* 2015;29(1):20-26. doi: 10.1038/leu.2014.250
27. Scherber R, Dueck AC, Johansson P, et al. The Myeloproliferative Neoplasm Symptom Assessment Form (MPN-SAF): international prospective validation and reliability trial in 402 patients. *Blood* 2011;118(2):401-08. doi: 10.1182/blood-2011-01-328955 [published Online First: 20110502]
28. Emanuel RM, Dueck AC, Geyer HL, et al. Myeloproliferative neoplasm (MPN) symptom assessment form total symptom score: prospective international assessment of an abbreviated symptom burden scoring system among patients with MPNs. *Journal of Clinical Oncology* 2012;30(33):4098.
29. Novartis. Data on File. Protocol CINC424B2301: Randomized, open label, multicenter phase III study of efficacy and safety in polycythemia vera subjects who are resistant to or intolerant of hydroxyurea: JAK inhibitor INC424 tablets versus best available care (The RESPONSE Trial).
30. Novartis. Data on File. Protocol CINC424B2401: Randomized, open label, multicenter phase IIIb study evaluating the efficacy and safety of ruxolitinib versus best available therapy in patients with Polycythemia Vera who are hydroxyurea resistant or intolerant (RESPONSE-2).
31. Harrison C. MAJIC Clinical Study Protocol / Chief Investigator: Claire Harrison. Version 9.0 ed: University of Birmingham, 2018.
32. Novartis. Data on File. Final Clinical Study Report (Week 260): Randomized, open-label, multi-center phase IIIb study evaluating the efficacy and safety of ruxolitinib versus best available therapy in patients with polycythemia vera who are hydroxyurea resistant or intolerant (RESPONSE-2).
33. Phillippo D, Ades A, Dias S, et al. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Sheffield: Decision Support Unit, 2016.
34. Alvarez-Larrán A, Pereira A, Cervantes F, et al. Assessment and prognostic value of the European LeukemiaNet criteria for clinicohematologic response, resistance, and intolerance to hydroxyurea in polycythemia vera. *Blood* 2012;119(6):1363-9. doi: 10.1182/blood-2011-10-387787 [published Online First: 20111212]
35. Alvarez-Larran A, Verstovsek S, Perez-Encinas M, et al. PF628. Comparison of ruxolitinib and real-world best available therapy in terms of overall survival and thrombosis in patients with polycythemia vera who are resistant or intolerant to hydroxyurea. European Hematology Association Congress. Stockholm, Sweden, 2018.

36. Novartis. Data on File: Propensity Score Matching Analysis: Overall Survival at Week 256. Ruxolitinib (RESPONSE) versus BAT (GEMFIN).
37. Alvarez-Larrán A, Garrote M, Ferrer-Marín F, et al. Real-world analysis of main clinical outcomes in patients with polycythemia vera treated with ruxolitinib or best available therapy after developing resistance/intolerance to hydroxyurea. *Cancer* 2022;128(13):2441-48. doi: https://doi.org/10.1002/cncr.34195
38. Hong S, Veenstra DL. PCN153: Cost-Utility Analysis of Ruxolitinib versus Best Available Therapy for the Treatment of Hydroxyurea Resistant/Intolerant Polycythaemia Vera Without Splenomegaly in the United States. *Value in Health* 2020;23:S50. doi: 10.1016/j.jval.2020.04.1638
39. National Centre for P. Cost-effectiveness of ruxolitinib (Jakavi®) for polycythaemia vera. Available at: http://www.ncpe.ie/wp-content/uploads/2015/05/Web-summary-final.pdf. [Last Accessed 08 September 2022].
40. Scottish Medicines Consortium. Ruxolitinib (Jakavi) 2019 [Available from: https://www.scottishmedicines.org.uk/medicines-advice/ruxolitinib-jakavi-full-smc2213/ accessed 08/09/2022.
41. Scottish Medicines Consortium. Ropeginterferon alfa-2b 250 micrograms/0.5 mL solution for injection in pre-filled pen (Besremi®). 2022 [Available from: https://www.scottishmedicines.org.uk/medicines-advice/ropeginterferon-alfa-2b-besremi-full-smc2421/ accessed 08/09/2022.
42. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual 2022 [Available from: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation accessed 05/08/2022.
43. Mukuria C, Rowen D, Brazier JE, et al. Deriving a Preference-Based Measure for Myelofibrosis from the EORTC QLQ-C30 and the MF-SAF. *Value in Health* 2015;18(6):846-55. doi: https://doi.org/10.1016/j.jval.2015.07.004
44. National Institute for Health and Care Excellence. TA756: Fedratinib for treating disease-related splenomegaly or symptoms in myelofibrosis. Available at: https://www.nice.org.uk/guidance/ta756 [Last Accessed 05 August 2022] 2021 [Available from: https://www.nice.org.uk/guidance/ta756.
45. Woods B, Sideris E, Palmer S, et al. NICE DSU Technical Support Document 19: Partitioned Survival Analysis for Decision Modelling in. 2017
46. Woods BS, Sideris E, Palmer S, et al. Partitioned Survival and State Transition Models for Healthcare Decision Making in Oncology: Where Are We Now? *Value in Health* 2020;23(12):1613-21. doi: 10.1016/j.jval.2020.08.2094
47. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data, 2011.
48. Lelonek E, Matusiak L, Wrobel T, et al. Burden of aquagenic pruritus in polycythaemia vera. *Acta dermato-venereologica* 2018;98(2):185-90. doi: 10.2340/00015555-2812 [published Online First: 2017/10/04]
49. Rafia R, Mukuria CW. PCN127 - Exploratory Psychometric Analysis of The Eq-5d In A Myelofibrosis Population. *Value in Health* 2015;18(3):A211. doi: https://doi.org/10.1016/j.jval.2015.03.1224
50. Hernández Alava M, Pudney S, Wailoo A. Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from an English Population Study: Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, UK, 2020.
51. Wailoo MHA, Pudney S. Estimating EQ-5D by age and sex for the UK: The University of Sheffield, 2022.
52. Matza LS, Cong Z, Chung K, et al. Utilities associated with subcutaneous injections and intravenous infusions for treatment of patients with bone metastases. *Patient preference and adherence* 2013;7:855-65.

53. National Institute for Health and Care Excellence. TA722: Pemigatinib for treating relapsed or refractory advanced cholangiocarcinoma with FGFR2 fusion or rearrangement. Available at: https://www.nice.org.uk/guidance/ta722 [Last Accessed 05 August 2022] 2021 [Available from: https://www.nice.org.uk/guidance/ta722.
54. Wang H-I, Aas E, Howell D, et al. Long-Term Medical Costs and Life Expectancy of Acute Myeloid Leukemia: A Probabilistic Decision Model. *Value in Health* 2014;17(2):205-14. doi: https://doi.org/10.1016/j.jval.2013.12.007
55. Crathorne L, Reimer E, Harris S, et al. PSY79 - The Cost Of Major Bleeding Events Associated With Anticoagulation Treatment - An Estimation Of The Contribution Of The Reversal Agents To The Overall Cost Of Major Bleeding In A UK Setting. *Value in Health* 2018;21:S449. doi: 10.1016/j.jval.2018.09.2654
56. Vallejo-Torres L, Morris S, Kinge JM, et al. Measuring current and future cost of skin cancer in England. *Journal of Public Health (Oxford)* 2014;36(1):140-8. doi: 10.1093/pubmed/fdt032 [published Online First: 2013/04/05]
57. National Institute for Health and Care Excellence. Venous thromboembolic diseases: diagnosis, management and thrombophilia testing, NICE guideline NG158:
[A] Evidence reviews for D-dimer testing in the diagnosis of deep vein thrombosis and pulmonary embolism: National Institute for Health and Care Excellence, 2020.
58. Novartis. Data on File. Novartis UK Polycythemia Vera Advisory Board. 24th June 2022.
59. Novartis. Data on File. Novartis UK Polycythemia Vera Advisory Board. 08th August 2022.
60. Department of Health. National Schedule of Reference Costs 2020-2021. Available at: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/#ncc1819 [Last Accessed 05 August 2022] [
61. Passamonti F, Palandri F, Saydam G, et al. Ruxolitinib versus best available therapy in inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): 5-year follow up of a randomised, phase 3b study. *The Lancet Haematology* 2022;9(7):E480-E92. doi: 10.1016/S2352-3026(22)00102-8

9 APPENDICES

9.1 EAG critique of the methods of review

Systematic review components and processes	EAG response (Yes, No, Unclear)	EAG comments
Was an appropriate review question clearly defined using the PICOD framework or an alternative?	Partly	The review question was clearly defined as identifying RCTs on the clinical efficacy and safety of any treatment in PV patients who are resistant or intolerant of HC (CS Appendix D.1), supported by a PICOS table for eligibility criteria (CS Appendix Table 8). However, limiting the study design to RCTs, and not searching for observational studies, meant the SLR could not identify relevant studies to support the ITC.
Were appropriate sources of literature searched?	Yes	The core bibliographic medical databases MEDLINE (including MEDLINE In-Process, etc.), Embase, and the Cochrane Library for CDSR and CENTRAL were searched. Several relevant haematology and oncology conferences, ClinicalTrials.gov, and the bibliographies of relevant systematic reviews and meta-analyses were searched (CS Appendix D.1.1).
Did the searches span an appropriate time period?	Yes	The original and update searches covered from database inception to 8 June 2022 (CS Appendix D.1.1).
Were appropriate search terms used and combined correctly?	Yes	Disease terms for PV were combined with RCT terms that were closely based on a published and validated search filter. Both subject headings and free text terms were used. All search strings were reported (CS Appendix D.1.1).
Were inclusion and exclusion criteria specified? If so, were these criteria appropriate and relevant to the decision problem?	Yes, except criteria for the intervention/ comparators were broader than the decision problem	The eligibility criteria for the SLR are defined in CS Appendix Table 8. They are appropriate and relevant but broader than the decision problem because they include any pharmacological intervention for the treatment of PV. This explains why 4 out of the 8 studies identified in the SLR were excluded (discussed above in section 3.2).
Were study selection criteria applied by two or more reviewers independently?	Yes	Two independent reviewers applied the study eligibility criteria. Consensus was achieved by comparison and discussion, and a third independent reviewer made a final decision if necessary (CS Appendix D.1.2).
Was data extraction performed by two or more reviewers independently?	No, but the process is adequate	A single individual extracted information with a second individual verifying and checking for missed data. A third individual arbitrated a

		final decision if necessary (CS Appendix D.1.2).
Was a risk of bias assessment or a quality assessment of the included studies undertaken? If so, which tool was used?	Yes, except for the GEMFIN registry cohort	All RCTs identified in the SLR were quality assessed using the CRD checklist (CS Appendix D.1.3 and D.3). However, the GEMFIN registry cohort used in the ITC was not assessed.
Was risk of bias assessment (or other study quality assessment) conducted by two or more reviewers independently?	No, but the process is adequate	A single individual assessed risk of bias and a second individual confirmed the conclusions. A third individual arbitrated a final decision if necessary (CS Appendix D.1.2).
Is sufficient detail on the individual studies presented?	Yes	Study details of all the included studies are tabulated in CS Appendix D.1.3. Some missing documents were provided in response to clarification questions A2 to A6. The CSR for RESPONSE week 32 was not provided.
If statistical evidence synthesis (e.g. pairwise meta-analysis, ITC, NMA) was undertaken, were appropriate methods used?	Yes	The company conducted an ITC (CS section B.2.9) using appropriate propensity score matching methods in order to estimate OS that was not confounded by crossover. Discussed in sections 3.3 to 3.5 of this report.
CDSR: Cochrane Database of Systematic Reviews; CENTRAL: Cochrane Central Register of Controlled Trials; CRD: Centre for Reviews and Dissemination; CSR: clinical study report; HC: hydroxycarbamide; OS: overall survival; PICOS: population, intervention, comparator, outcome, study design; PV: polycythaemia vera; RCTs: randomised controlled trials; SLR: systematic literature review.		

9.2 Baseline characteristics of the included studies

Characteristic	RESPONSE		RESPONSE-2		MAJIC-PV		GEMFIN
	Ruxolitinib (n=110)	BAT (n=112)	Ruxolitinib (n=74)	BAT (n=75)	Ruxolitinib (n=93)	BAT (n=87)	BAT (n=[redacted])[e]
Age – years							
Median (range)	62.0 (34–90)	60.0 (33–84)	63 (NR)	67 (NR)	67 (34–88)	66 (28–85)	Mean ± SD [redacted]
IQR	-	-	54–71	61–74	-	-	-
>60 years – n (%)	-	-	46 (62)	57 (76)	-	-	-
Sex – n (%)							
Male	66 (60.0)	80 (71.4)	39 (53)	47 (63)	56 (60)	49 (56)	[redacted]
Female	44 (40.0)	32 (28.6)	35 (47)	28 (37)	37 (40)	38 (44)	-
Time since diagnosis – years							
Median (range)	8.2 (0.5–36)	9.3 (0.5–23)	6.5 (2.9–10.7)	6.7 (3.2–10.6)	-	-	-
Disease duration - months							
Median (range)	-	-	-	-	90 (0–365)	96 (4–388)	-
Previous lines of therapy							
Median (range)	-	-	-	-	1 (1–4)	2 (1–6)	-
Previous lines of antineoplastic therapy							
1	-	-	53 (72%)	52 (69%)	-	-	-
>1	-	-	21 (28%)	23 (31%)	-	-	-
Duration of prior HC/HU therapy – years							
Median (range)	3.1 (<0.1–20.9)	2.8 (<0.1–20.9)	2.83 (0.57–6.61) [a]	3.55 (0.57–7.03) [a]	-	-	
Resistance/intolerance (R/I) to hydroxycarbamide							
Both R/I – n (%)	-	-	-	-	19 (20)	27 (31)	-
Intolerant – n (%)	-	-	-	-	43 (46)	37 (43)	-
Resistant – n (%)	-	-	-	-	31 (33)	23 (26)	-

Characteristic	RESPONSE		RESPONSE-2		MAJIC-PV		GEMFIN
	Ruxolitinib (n=110)	BAT (n=112)	Ruxolitinib (n=74)	BAT (n=75)	Ruxolitinib (n=93)	BAT (n=87)	BAT (n=[redacted])[e]
Previous HC/HU treatment status – n (%)							
Unacceptable side effects	59 (53.6)	61 (54.5)	44 (59)	45 (60)	-	-	-
Inadequate response	51 (46.4)	51 (45.5)	30 (41)	30 (40)	-	-	-
ECOG performance status – n (%) [b]							
0	76 (69.1)	77 (68.8)	-	-	57 (61)	59 (68)	-
1	31 (28.2)	34 (30.4)	-	-	32 (34)	27 (31)	-
2	3 (2.7)	1 (0.9)	-	-	3 (3)	1 (1)	-
Prior thromboembolic event							
n (%)	39 (35.5)	33 (29.5)	21 (28)	18 (24)	26 (28)	38 (44)	[redacted][f]
Presence of *JAK2* V617F mutation							
n (%)	104 (94.5)	107 (95.5)	72 (97) [c]	69 (92)	-	-	[redacted]
Allele burden – % ± SD	76.2 ± 17.8	75.0 ± 22.6	-	-	-	-	
JAK2 mutation status							
Wild type – n (%)	-	-	-	-	3 (3)	1 (1)	-
JAK2V617F – n (%)	-	-	-	-	89 (96)	85 (98)	-
JAK2 exon 12 – n (%)	-	-	-	-	1 (1)	1 (1)	-
Spleen length							
Below costal margin – cm							
Median (range)	7.0 (0–24.0)	7.0 (0–25.0)	-	-	-	-	-
<10 cm – n (%)	71 (64.5)	67 (59.8)	-	-	-	-	-
>20 cm – n (%)	2 (1.8)	4 (3.6)	-	-	-	-	-
Overall length by ultrasound – cm							
Median (range) [g]	-	-	-	-	14 (9, to 26)	14 (9 to, 30)	-
Spleen volume – cm^3							

Characteristic	RESPONSE		RESPONSE-2		MAJIC-PV		GEMFIN
	Ruxolitinib (n=110)	BAT (n=112)	Ruxolitinib (n=74)	BAT (n=75)	Ruxolitinib (n=93)	BAT (n=87)	BAT (n=[redacted])[e]
Median (range)	1195 (396–4631)	1322 (254–5147)	-	-	-	-	
Palpable splenomegaly							
n (%)	-	-	-	-	23 (25)	22 (25)	-
Percentage HCT level – % [d]							
Mean ± SD	43.6 ± 2.2	43.9 ± 2.2	42.8 ± 1.46	42.7 ± 1.44	-	-	-
Median (range or IQR)	43.3 (range: 39.2–50.5)	44.0 (range: 37.6–50.5)	43.0 (IQR: 41.7–44.0)	42.7 (IQR: 41.7–44.0)	43 (range: 28–57)	43 (range: 34–52)	-
HCT category – n (%)							
40–45%	79 (71.8)	83 (74.1)	-	-	-	-	-
>45%	28 (25.5)	25 (22.3)	-	-	-	-	-
WBC count × 10^{-9}/L							
Mean ± SD	17.6 ± 9.6	19.0 ± 12.2	12.0 ± 8.19	13.0 ± 8.06	-	-	-
Median (range)	-	-	-	-	9 (2–73)	9 (2–37)	-
Platelet count × 10^{-9}/L							
Mean ± SD	484.5 ± 323.3	499.4 ± 318.6	469.5 ± 295.96	471.5 ± 350.38	-	-	-
Median (range)	-	-	-	-	401 (61–1546)	356 (99–1420)	-
Haemoglobin g/L							
Median (range)	-	-	-	-	136 (85-173)	136(65-163)	-
Phlebotomies within 24 weeks before screening							
≥2 – n (%)	-	-	58 (78)	57 (76)	-	-	-
Median (range)	2.0 (1–8)	2.0 (0–16)	-	-	-	-	
History of haemorrhage							
n (%)	-	-	-	-	3 (3)	6 (7)	-
Migraine or erythromelalgia							
n (%)	-	-	-	-	6 (6)	4 (5)	-

Characteristic	RESPONSE		RESPONSE-2		MAJIC-PV		GEMFIN
	Ruxolitinib (n=110)	BAT (n=112)	Ruxolitinib (n=74)	BAT (n=75)	Ruxolitinib (n=93)	BAT (n=87)	BAT (n=[redacted])[e]
Diabetes							
n (%)	-	-	-	-	7 (8)	3 (3)	-
Hypertension							
n (%)	-	-	-	-	33 (35)	25 (29)	-
Cytopenia at lowest hydroxycarbamide dose							
n (%)	17 (15)	-	-	-	-	-	[redacted]

Sources: CS Table 7; CS Table 12; CS Appendix M.2.1; Clarification response A11 Table 2.
[a] Manually converted duration in months from the source to duration in years for consistency. [b] ECOG performance status ranges from 0 to 5, with 0 indicating no symptoms and higher numbers indicating increasing disability. [c] For five patients (ruxolitinib, n=2; BAT, n=3) the *JAK2* V617F mutation was not confirmed by central laboratory assessment. These patients were not included as *JAK2* V617F mutation positive. [d] Value at the end of the HCT control period before randomisation. Patients who had an HCT of 40–45% within 14 days before their day 1 visit could proceed to randomisation; however, the HCT at baseline may have been higher or lower. [e] Excludes 7 patients without follow-up beyond the date of being identified as resistant or intolerant to hydroxycarbamide (clarification response A10). [f] At time of resistance/intolerance. [g] from clarification response C1.

9.3 Company and EAG risk of bias assessments for the RCTs

Question	Assessor	Trial		
		RESPONSE	**RESPONSE-2**	**MAJIC-PV**
Was randomisation carried out appropriately?	**Company**	Unclear risk of bias, randomisation methods not reported	Low risk of bias, random assignment of participants (1:1), using an interactive voice and web response system.	Unclear risk of bias, randomisation methods were not reported
	EAG	**Probably low risk of bias** The trial protocol states that an IRT system will assign a randomization number to the participant to link them to a treatment arm. However, the trial publication[22] does not confirm that this process was followed in practice.	**Agree, low risk of bias** An interactive voice and web response system was used to assign randomisation numbers to participants to link each participant to a trial arm. [61]	**Agree, unclear risk of bias** The trial protocol states that "randomisation will be based on a minimisation algorithm prepared by the trial statistician", but not reported whether or how this was conducted.
Was the concealment of treatment allocation adequate?	**Company**	Unclear risk of bias, concealment of treatment was not reported	Low risk of bias, an interactive voice and web response system was contacted by the investigator	Unclear risk of bias, concealment of treatment was not reported
	EAG	**High risk of bias** Due to being an open-label trial (NB the full allocation process is not explained and the trial publication[22] does not confirm that the stated process was followed in practice).	**High risk of bias** Due to being an open-label trial (NB the full allocation process is not explained and the trial publication[61] does not confirm that the stated process was followed in practice).	**High risk of bias** Due to being an open-label trial. Some patients "did not want to be in the BAT arm" (Figure S2 in the draft trial manuscript[16]).
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	**Company**	Low risk of bias, the authors of the primary publication reported that there were no significant differences between the two treatment groups with regard to baseline	Low risk of bias, baseline characteristics were generally similar between treatment groups. There were slight differences in median age and sex between the groups	Low risk of bias, authors reported that baseline characteristics at randomisation were balanced, however full patient characteristics were not reported

		characteristics and disease history		
	EAG	**Agree, low risk of bias** Baseline characteristics appear well balanced with minor exceptions (the ruxolitinib arm had 11% more females and 6% more people who had had a prior thromboembolic event than the BAT arm).	**Agree, low risk of bias** Baseline characteristics appear well balanced with minor exceptions (the ruxolitinib arm had 14% fewer people aged > 60 years and median age 4 years younger, 10% more females and a median 8.7 months less prior hydroxycarbamide therapy than the BAT arm).	**Unclear risk of bias** Most baseline characteristics appear balanced. However, 16% more BAT than ruxolitinib patients had prior thrombosis and the BAT arm also had a slightly longer disease duration and number of previous lines of therapy; whilst 11% more patients in the ruxolitinib arm were both intolerant and resistant to hydroxycarbamide.[16]
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blind to treatment allocation, what might be the likely impact on the risk of bias (for each outcome)?	**Company**	High risk of bias, open-label study. There was a potential for bias, particularly in PROs. Bias for ruxolitinib versus hydroxycarbamide may be particularly relevant as patients were already known to be hydroxycarbamide -resistant/ intolerant	High risk of bias, open-label study. There was a potential for bias in outcomes, particularly PROs. Bias for ruxolitinib versus HC/HU may be particularly relevant as patients were already known to be hydroxycarbamide-resistant/ intolerant The assessors were unaware of the treatment group assignments until database lock	High risk of bias, open-label study. Potential for bias, particularly in symptom and QoL scores.
	EAG	**Agree, high risk of bias** Note that being open label the trial has high risks of bias relating to: (i) elective patient crossover, (ii) patient care, and (iii) recording of outcomes, (iv) analysis of outcomes.	**Agree, high risk of bias** Note that being open label the trial has high risks of bias relating to: (i) elective patient crossover, (ii) patient care, and (iii) recording of outcomes, (iv) analysis of outcomes.	**Agree, high risk of bias** Note that being open label the trial has high risks of bias relating to: (i) patient care, (ii) recording of outcomes and (iii) analysis of outcomes.

Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	**Company**	High risk of bias, patients were able to cross over from BAT treatment arm to ruxolitinib at Week 32; 96 patients crossed over at or after Week 32 – this would have been affected by the open label nature of the study	High risk of bias, patients were able to cross over from BAT treatment arm to ruxolitinib at Week 28; 51 patients crossed over at or after Week 28 – this would have been affected by the open label nature of the study.	Unclear risk of bias, drop-outs were not reported
	EAG	**≤ week 32: Unclear risk of bias** Unclear whether patients were informed that they could cross over at week 32 and if so whether this would have affected their outcomes prior to week 32. CONSORT chart (CS Appendix Figure 4) does not identify dropout numbers or reasons prior to week 32. **> week 32: Agree, high risk of bias** Reasons as stated by the company	**≤ week 28: Unclear risk of bias** Unclear whether patients were informed that they could cross over at week 32 and if so whether this would have affected their outcomes prior to week 32. CONSORT chart (CS Appendix Figure 5) does not identify dropout numbers or reasons prior to week 32. **> week 28: Agree, high risk of bias** Reasons as stated by the company	**Probably low risk of bias** Table S4 of the unpublished manuscript[16] suggests numbers and reasons for dropout were broadly similar between trial arms.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	**Company**	Low risk of bias, the pre-defined outcome measures are all presented in the available records	High risk of bias, some outcomes measured are not reported, however analyses are promised in future publications but still not reported (e.g., changes ECOG status and spleen length)	High risk of bias, ISRCTN record lists outcome measures which are not reported in the available records
	EAG	**Efficacy outcomes: probably low risk of bias** Most of the pre-specified outcomes in the trial protocol	**Efficacy outcomes: Unclear risk of bias** The previously missing pre-specified outcomes (e.g. spleen	**Agree, high risk of bias** EQ-5D, MDASI and partial response rate are specified in the trial protocol, but results are not reported. Results

		have been reported, with some minor exceptions (MPN-PAF results not reported; overall clinico-haematologic response reported at 5 years but not at earlier timepoints). **HRQoL outcomes: high risk of bias** 32-week results for the MPN-SAF-TSS and PGIC are reported in the CS and publications only as % changes which have limited clinical interpretation, with no indication of the original scores, sample size or variance in scores. The week 32 CSR was not provided to the EAG.	length, ECOG performance status and WPAI score) are summarised in the week 260 CSR (NB individual patient ECOG PS scores are tabulated but not analysed). **HRQoL outcomes: high risk of bias** Changes in MPN-SAF TSS and PGIC are reported in the CS, publications and week 28 CSR only as % changes which have limited clinical interpretation, with no indication of the original scores, sample size or variance in scores. The week 28 CSR does report numbers achieving disease resolution, but only for a subgroup who had a baseline score of ≥20.	for the MPN-SAF are reported only as differences between arms, without the original scores for each arm.
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	**Company**	Low risk of bias, ITT analysis was used, with data from all patients who underwent randomisation. Patients with missing assessments that prevented the evaluation of the primary and secondary endpoints were considered non-responders	Low risk of bias, ITT analysis was applied for the primary and key secondary endpoints, including data from all patients randomly assigned to treatment Patients with missing assessments that prevented the study of the primary and secondary endpoints endpoint were considered non-responders	Unclear risk of bias, an mITT analysis was used (those who commenced study treatment and had at least one response assessment) but details of how missing data were accounted for were not given.

	EAG	**Primary and key secondary outcomes: Low risk of bias** ITT analysis: Missing response data were considered non-responders and missing phlebotomy ineligibility data were considered phlebotomy eligible (number of missing observations not reported). **HRQoL outcomes: High risk of bias** Missing data excluded; number and reasons for missing data not reported. (sources: CS and trial protocol)	**Primary and key secondary outcomes: Low risk of bias** ITT analysis: Missing response data were considered non-responders and missing data for remission outcomes were considered to represent no remission. **HRQoL outcomes: High risk of bias** Missing data excluded; number and reasons for missing data not reported (sources: CS and trial protocol)	**Primary and secondary outcomes: unclear risk of bias** The trial protocol states that for secondary outcomes "the amount of missing data will be reported but not imputed". However the amount of missing data is not reported. **HRQoL outcomes: High risk of bias** Missing data probably excluded; number and reasons for missing data not reported. Sample size is unclear for MPN-SAF. **All outcomes: unclear risk of bias** Lack of clarity around crossovers from ruxolitinib to BAT and receipt of ruxolitinib on the BAT arm (see section 3.2.3 for discussion).
Also consider whether the authors of the study publication declared any conflicts of interest/study funding.	**Company**	Unclear risk of bias, sponsor (Incyte and Novartis) involvement in study design and data analysis not reported, Author affiliations were disclosed	Low risk of bias, study funding and author conflicts of interest declared. The study was sponsored and designed by Novartis. Data were analysed and interpreted by Novartis in collaboration with all the authors. Novartis was unaware of treatment group assignments until database lock	Unclear risk of bias, nothing declared. Funder: Leukaemia & Lymphoma Research (UK)
	EAG	Conflicts of interest is not an independent domain of bias. Any risks of bias arising through conflicts of interest would be reflected in the bias assessments already reported above. For example, Novartis' (lack		

		of) awareness of treatment assignment should already be captured under the allocation concealment and blinding questions which indicate a high risk of bias.
Source: CS Appendix Table 14 with EAG additions. BAT: best available therapy; IRT: interactive response technology; ITT: intention to treat; MDASI: M.D. Anderson Symptom Inventory; mITT: modified intention to treat; MPN-PAF: Myeloproliferative Neoplasm Pruritis Assessment Form		

9.4 EAG summary of statistical methods in the RCTs

	RESPONSE	RESPONSE-2	MAJIC-PV
Analysis populations			
Summary	**Full analysis set: ITT analysis** (primary and two key secondary outcomes): all randomised patients included and analysed according to their hydroxycarbamide stratum and the treatment they were randomised to. **Safety set:** all randomised patients who received at least one dose of their allocated treatment, analysed according to the treatment they actually received. People randomized to the BAT arm who were intended to receive no therapy were included in the safety set. **Per protocol set:** A subset of the full analysis set patients who received at least one dose of study treatment and did not have a major protocol violation.	**Full analysis set: ITT analysis** (primary and key secondary outcome): all randomised patients included and analysed according to their hydroxycarbamide stratum and the treatment they were randomised to. **Safety set:** all randomised patients who received at least one dose of their allocated treatment, analysed according to the treatment they actually received. People randomized to the BAT arm who were intended to receive no therapy were included in the safety set. **Per protocol set:** A subset of the full analysis set patients who received at least one dose of study treatment and did not have a major protocol violation.	Modified ITT analysis: All patients who started treatment within one year of randomisation and had at least one response available. Safety population: Any patient starting treatment. The draft trial manuscript[16] states that as the modified ITT population included 10 patients switching to ruxolitinib, supporting analyses were performed censoring at the time they began ruxolitinib; these analyses did not affect the conclusions from the modified ITT analysis. However, results of these analyses are not reported.
EAG comment	The analysis populations for the primary, two key secondary, and safety outcomes are appropriate. Analysis populations are not specified for the remaining secondary outcomes and HRQoL measures. The per protocol population is not referred to in the CS which is reasonable given that the full analysis set is more robust.	The analysis populations for the primary, key secondary, and safety outcomes are appropriate. Analysis populations are not specified for the remaining secondary outcomes and HRQoL measures. The per protocol population is not referred to in the CS which is reasonable given that the full analysis set is more robust.	Limited details of the analysis populations are reported; analysis populations for HRQoL outcomes are unclear. Potential for bias due to unaccounted for missing data (see Appendix 9.3).

Sample size and power calculations			
Summary	**Primary outcome:** Assuming an HCT control rate of 10% in the BAT arm and 30% in the ruxolitinib arm, a sample size of 200 patients was deemed to be required to detect a significant difference with a two-sided test (0.05 significance level and 94% power) (CS Table 9). **Key secondary outcomes:** **Durable primary response:** According to the trial protocol, assuming 24% and 8% primary outcome responders in the ruxolitinib and BAT arms respectively at week 48, a large sample normal approximation would give 87% statistical power. An observed response rate as low as 17.1% in the ruxolitinib arm would achieve statistical significance relative to an observed response rate of 8% in the BAT arm. **CHR at week 32:** According to the trial protocol, the power for complete haematological remission at 32 weeks would be approximately 99% using a large sample normal approximation, meaning that an observed response rate as low as 40% in the ruxolitinib arm would achieve statistical significance relative to an observed response rate for the BAT arm of 27%.	**Primary outcome:** Sample size was calculated based on the results for the HCT control portion of the compound primary outcome, assuming HCT control rates of 50% in the ruxolitinib group and 20% in the BAT group (corresponding to an OR of 4·0). A total of 116 patients were needed to detect a significant difference between treatment groups with two-sided t-test at alpha=0·05 and 90% power. Planned enrolment was 130 patients (65 in each group) to allow for an estimated 10% attrition rate (CS Table 9 and trial publication[23]). **Key secondary outcome** According to the trial protocol, a total of 116 patients (58 patients in each treatment arm) would provide 90% power to detect a 30% increase in the rate of CHR at Week 28, between a BAT arm rate of 20% and a ruxolitinib arm rate of 50% (corresponding to an OR of 4.0) at a 5% significance level.	The complete response rate for the control group was estimated to be 35% and a clinically significant improvement would be 15%. Assuming complete response rates in the control and treatment group were 35% and 50% respectively, 90 patients would be required in each arm to detect a clinically significant difference of 15% with 78% statistical power at a 10% level of significance.[16] Apart from the primary outcome, additional hypotheses tests were unpowered, exploratory and not pre-specified [16]

EAG comment	The trial randomised 110 and 112 participants per arm so appears to be adequately powered for the primary outcome and probably also the two key secondary outcomes (the power calculation descriptions for the secondary outcomes do not specify the sample size). Adequacy of the sample size for detecting treatment effects in the remaining secondary outcomes is uncertain.	The trial randomised 74 and 75 participants per arm so appears to be adequately powered for the primary outcome and key secondary outcome. Adequacy of the sample size for detecting treatment effects in the remaining secondary outcomes is uncertain.	The complete response rates used for the power calculation in the ruxolitinib and BAT arms (50% and 35%) overestimate the observed rates reported in the trial (43% and 23%). The stated power calculation in the protocol uses a 10% error rate (relatively high) to achieve 78% power (relatively low). Nevertheless, a treatment effect on the primary outcome was detected (p=0.02). Adequacy of the sample size for detecting treatment effects in the remaining secondary outcomes is uncertain.
Methods to account for multiplicity			
Summary	A family wise α-level of 0.05 overall was applied for three pre-specified comparisons: the primary outcome and two key secondary outcomes. Conditional on significance of the primary outcome, treatment effects on the proportions of people achieving a CHR at week 32 and achieving a durable primary endpoint response at week 48 were tested at two-sided α = 0.05 for the two outcomes, controlling for multiplicity using the Hochberg procedure. [22] According to the trial protocol, no alpha adjustment was planned for the remaining secondary outcomes.	Not reported in the CS, week 28 CSR or study publication.[23] According to the trial protocol, the analysis of the key secondary outcome (proportion achieving CHR at week 28) was performed in a hierarchical manner (calculation method not specified). The key secondary outcome was tested at an α-level of 0.05 only if the primary outcome was significant at an α-level of 0.05. For all secondary efficacy outcomes, statistical tests were intended to be performed for descriptive purposes and not adjusted for multiple comparisons.	The CS, draft trial manuscript[16] and trial protocol do not mention whether any control for multiple outcome testing was applied.
EAG comment	The type I error control procedure is appropriate but only three outcomes are included. The likelihood of type I error in	The type I error control procedure is appropriate but only two outcomes are included. The likelihood of type I error in	No information available. The likelihood of nonsignificant treatment effects being declared significant is uncertain. Reliance

	testing the remaining secondary outcomes is uncertain.	testing the remaining secondary outcomes is uncertain.	on the statistical test results alone for inference is therefore inadvisable.
Analysis of outcomes			
Summary	**Primary outcome:** Responder rates were analysed using a Cochran-Mantel-Haenszel (CMH) test stratified by hydroxycarbamide tolerance status (resistant versus intolerant), 2-sided at the 5% significance level. The overall stratum-adjusted odds ratio was used as a measure of association between treatment and response. The adjusted proportion difference and its 95% CI were calculated using CMH weight and Wald-type CI or any other appropriate method (CS Table 9). The following is from the trial protocol (not reported in the CS): **Key secondary outcomes** (durable primary response and complete haematological response): Treatment groups were compared using a CMH test stratified on hydroxycarbamide tolerance as with the primary outcome. **All other secondary outcomes:** Are non-comparative in nature. These (except for durability of primary response and duration of primary response which can be evaluated in both treatment groups) will be evaluated only in the subjects originally randomized to	**Primary outcome:** A two-sided CMH test stratified by hydroxycarbamide tolerance status was conducted at the 5% level of significance. The odds ratio is presented with 95% Wald confidence limits (CS Table 9). The following is from the trial protocol (not reported in the CS or week 28 CSR): **Key secondary outcome** (complete hematological remission at week 28): Analysed using a two-sided stratified CMH test (stratification factors not reported in the CS, protocol or publications[23 61]). **Other secondary outcomes** (HCT control at weeks 52, and 80, complete hematological remission at weeks 52 and 80, and partial remission based on the ELN and IWG-MRT criteria at weeks 28, 52 and 80: A two-sided stratified CMH test at the 5% level of significance. **Other outcomes** (changes from baseline in HCT, summary of spleen length, number of phlebotomies from baseline to week 28, and HRQoL measures): Summarised with descriptive statistics.	**Primary outcome:** The trial protocol states that complete response was to be assessed using a normal test with continuity correction and unpooled variance and a $p<0.10$ considered statistically significant. Apart from the primary outcome, additional hypotheses tests were exploratory, unpowered, two-sided and considered $p<0.05$ statistically significant trial manuscript[16] and protocol). **HRQoL outcomes:** Changes from baseline and between-arm differences in change by timepoint were estimated using a linear mixed model which included covariates for categorical time point, treatment arm, and the interaction between time point and treatment arm. The difference between arms in proportion of patients with best post-baseline TSS response of 50% or greater was tested using a Chi-square test. [16] **Time-to-event outcomes:** Were predominantly analyzed using Kaplan-Meier methods, with differences in survival analyses determined using the Cox model, adjusting for the stratification factor (gender), and treatment (when not the primary variable of interest). [16]

	ruxolitinib and will be summarised descriptively.		
EAG comment	The statistical methods are reported in different sources with varying levels of detail but appear generally appropriate. The CS does not state whether the analyses were checked or validated.	The statistical methods are reported in different sources with varying levels of detail but appear generally appropriate. The CS does not state whether the analyses were checked or validated.	The statistical methods appear generally appropriate. However, no justification is provided for using a relatively high p-threshold for determining statistical significance of the primary outcome (but not secondary outcomes) which gives a 1 in 10 chance of nonsignificant effects being declared significant.
Handling of missing data			
Summary	**Primary and key secondary outcomes:** ITT analysis: Missing response data including patient withdrawals were considered non-responders and missing phlebotomy ineligibility data were considered phlebotomy eligible (number of missing observations not reported). **HRQoL outcomes:** Missing data excluded; number and reasons for missing data not reported. **Survival outcomes:** Censoring methods not reported (not specified in the CS, trial protocol or trial publication; the week 32 CSR was not provided to the EAG).	**Primary and key secondary outcomes:** ITT analysis: Missing response data including withdrawals were considered non-responders and missing data for remission outcomes were considered to represent no remission. **HRQoL outcomes:** Missing data excluded; number and reasons for missing data not reported **Survival outcomes:** (not stated in the CS; information from the trial protocol): For TFS, patients without an event by the analysis data cut-off were to be censored at the date of last adequate assessment. For OS, patients not known to have died before the data cut-off were to be censored at the date of the last assessment for patients who were on treatment or at the date of the last contact for patients in survival follow-up.	**Primary and secondary outcomes:** The trial protocol states that for secondary outcomes “the amount of missing data will be reported but not imputed”. However the amount of missing data is not reported in the CS or trial draft manuscript. [16] **HRQoL outcomes:** Missing data probably excluded; number and reasons for missing data not reported. Sample size is unclear for MPN-SAF. **Survival outcomes:** Censoring methods not reported.

EAG comment	Methods for handling missing data were appropriate for primary and secondary outcomes. Missing data were not accounted for in analyses of HRQoL and other exploratory outcomes. Number and reasons for missing data not fully reported.	Methods for handling missing data were appropriate for primary and secondary outcomes. Missing data were not accounted for in analyses of HRQoL and other exploratory outcomes. Number and reasons for missing data not fully reported.	Overall missing data were not accounted for, and the amount of missing data and reasons for data being missing were not reported.
Subgroup analyses			
Summary	Pre-specified subgroup comparisons (trial protocol section 9.4.4) were: baseline palpable splenomegaly (<10cm versus ≥10cm below the costal margin), sex (male versus female), age group (≤60 years versus >60 years), hydroxycarbamide intolerance or resistance, region (US versus non-US), race (White or Caucasian versus other) and ethnicity (Hispanic or Latino versus other). The odds of achieving the primary composite response outcome at week 32 were compared across subgroups by calculating odds ratios and their confidence intervals using logistic regression and displaying these in a forest plot. Post-hoc subgroup comparisons (not specified in the trial protocol) are reported in CS Appendix E for patients who had received prior IFN-alfa, IFN-alfa as BAT, or ruxolitinib after crossover from receiving IFN as BAT. These	Pre-specified subgroup comparisons (trial protocol section 10.4.4) were: hydroxycarbamide intolerance or resistance, sex (male versus female), age group (≤60 years versus >60 years), risk category (0 risk factors versus 1-2 risk factors including age >60 and/or previous thromboembolism). The odds of achieving HCT control at week 28 were compared across subgroups by calculating odds ratios and their confidence intervals using logistic regression and displaying these in a forest plot. Post-hoc subgroup comparisons (not specified in the trial protocol) are reported in CS Appendix E for patients who had received prior IFN-alfa, IFN-alfa as BAT, or ruxolitinib after crossover from receiving IFN as BAT. These subgroups pooled data from RESPONSE and RESPONSE-2.	Pre-specified subgroup comparisons (trial protocol section 13.3) were: hydroxycarbamide intolerance or resistance, blood count quartile at randomisation (3 classes), sex (male versus female), disease duration (5 classes), ruxolitinib starting dose (5mg or 10mg), number of prior treatments (4 classes), WBC count at trial entry (3 classes), haemoglobin at trial entry (4 classes), and splenomegaly (yes versus no). No analysis methods for subgroups were specified. The trial protocol states that due to the lack of statistical power for subgroup analyses, subgroup analysis results provided will be exploratory only. However, no subgroup analyses are reported in the CS or draft trial manuscript. [16]

	subgroups pooled data from RESPONSE and RESPONSE-2.		
EAG comment	The pre-specified subgroup analysis method is appropriate, but no justification is provided for the choice of subgroups analysed, which varied between the trials. The post-hoc IFN-alfa subgroups had small sample sizes ranging from 13 to 30 participants.	The pre-specified subgroup analysis method is appropriate, but no justification is provided for the choice of subgroups analysed, which varied between the trials. The post-hoc IFN-alfa subgroups had small sample sizes ranging from 13 to 30 participants.	No subgroup analysis method or results were reported.

BAT: best available therapy; CHR: complete haematological remission; CI: confidence interval; CMH test: Cochran-Mantel-Haenszel test; CSR: clinical study report; ELN: European LeukemiaNet; HRQoL: health-related quality of life; IFN: interferon; ITT: intention to treat; IWG-MRT: International Working Group - Myeloproliferative Neoplasms Research and Treatment; MPN-SAF: Myeloproliferative Neoplasm Symptom Assessment Form; OS: overall survival; TFS: transformation-free survival; US: United States; WBC: white blood cells.

9.5 EAG summary of key economic issues and additional analyses

Analysis	Company analysis	EAG comment	EAG additional analyses
Population and subgroups			
Primary analysis	Subgroup with splenomegaly (RESPONSE trial population)	All three trial populations represent subgroups of the population of interest The EAG considers that the MAJIC-PV analysis is likely to be more relevant as the trial was wholly UK-based and it included the majority of the licensed population	We report EAG analyses and scenarios for all three subgroups.
	Subgroup without splenomegaly (RESPONSE-2 trial population)		
MAJIC-PV analysis	'High risk' subgroup (MAJIC-PV trial population)		
Model structure			
Primary analysis	STM with three health states (On ruxolitinib, On BAT, death) Key PV complications modelled as events with one-off costs and QALY losses Partition of the BAT state: BAT 1, BAT 2+ and no further treatment	In theory, the STM has the advantage of modelling dependency between discontinuation of ruxolitinib and OS beyond the trial period. Whereas in the PSM, OS and ruxolitinib discontinuation are extrapolated independently However, neither model structure reflects post-trial dependencies between the onset of major complications and survival The BAT partition is subject to uncertainty over long-term trends in cessation of all therapy and disutilities	We do not include partitioning of the BAT state in the EAG preferred analysis. The BAT partition is included in EAG scenario analysis We also note uncertainty over the OS extrapolations as mortality due to complications is not explicitly modelled.
MAJIC-PV analysis	PSM with the same health states as the primary analysis and key PV complications modelled as events No partition of the BAT state		
OS extrapolations			
Primary analysis Survival pre- and post-discontinuation of ruxolitinib (competing risk	Extrapolations fitted to pooled IPD from RESPONSE and RESPONSE-2 Exponential distribution used in base case for pre- and post-discontinuation survival extrapolations. Scenarios with	The competing risk approach is appropriate for the STM, as is the pooling of trial data, given the low numbers of observed events Methods used to fit the survival extrapolations are appropriate and the	We apply the general population mortality constraint throughout the time horizon (company response to CQ B5)

Analysis	Company analysis	EAG comment	EAG additional analyses
analysis)	other distributions are reported in CS Appendix P General population mortality constraint applied after the trial period for pre-discontinuation survival (but throughout the time horizon for post-discontinuation survival).	exponential is a reasonable choice for the base case It is not plausible that mortality rates should be lower in the first five years of ruxolitinib treatment than for people of the same age in the general population	
Treatment effect HR for OS (ruxolitinib vs. BAT)	HR estimated from piecewise Cox proportional hazards analysis of reconstructed MAJIC-PV KM data Scenarios: constant HR from MAJIC-PV trial report; indirect comparison with GEMFIN; Alvarez-Larrán analysis of Spanish data; and pooled HR from RESPONSE and RESPONSE-2 (not corrected for crossover) Waning assumption: linear decline from year 5 to HR=1 at year 20	MAJIC-PV is the best available source for estimation of the relative effect on survival The company's piecewise HR estimates have some face validity, but they are highly uncertain, with wide and overlapping confidence intervals. There is no clear rationale for the company's waning assumptions, but they do potentially mitigate against uncertainty.	We opt for the constant HR reported by the MAJIC-PV investigators, which is more appropriate from a statistical perspective. We also report scenarios with more conservative waning scenarios.
Treatment to treatment discontinuation			
Primary analysis TTD for ruxolitinib due to reasons other than death from competing risk analysis	Odds spline with 1 knot for RESPONSE and RESPONSE-2 (separate competing risk analyses)	There is the potential for overfitting data using an odds spline model, and a parametric model is preferred.	The EAG selects a Weibull distribution for the extrapolation of data for both RESPONSE and RESPONSE-2.
Utilities			
Health state utilities	MF-8D from RESPONSE trial for base case (EQ-5D from RESPONSE-2 for scenario).	Although the company comments on the use of the MF-8D for myelofibrosis in previous appraisals, the MF-8D was not designed for patients with	The EAG uses the EQ-5D utility values in our preferred analysis. This is in accordance with NICE preferred methods and allows for

Analysis	Company analysis	EAG comment	EAG additional analyses
		polycythaemia vera. Assumptions were made in order to obtain PV symptom scores in place of myelofibrosis symptoms scores. There is a lack of direct evidence validating the EQ-5D and MF-8D in patients with PV.	consistency across NICE appraisals.
Resource use and costs			
Thromboembolic events	The company assume a cost equivalent with one emergency department visit, £297, for the management of all Grade 1 and 2 thromboembolic events.	EAG clinical experts suggested a higher cost associated with the management of Grade 1 and 2 thromboembolic events, taking into account the processes required to confirm and treat such an event.	The EAG applies additional costs in the base case for a D-dimer test, vascular ultrasound, and a single dose of an anticoagulant.

Single Technology Appraisal

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

EAG report – factual accuracy check and confidential information check

"Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release." (Section 5.4.9, NICE health technology evaluations: the manual).

You are asked to check the EAG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by **5pm on Wednesday 30 November** 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.

Please underline all confidential information, and separately highlight information that is submitted as '[redacted]' in turquoise, all information submitted as '[redacted]' in yellow, and all information submitted as '[redacted]' in pink.

Please note that page numbers referred to by the EAG are those in the EAG report version with tracked changes displayed

Issue 1 Aims of ruxolitinib treatment

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 25, first paragraph states: *'Ruxolitinib aims to reduce the occurrence of thromboembolic events that are caused by having too many red cells in the blood (but does not alter the underlying genetic basis of the disease)'*	Please amend the sentence to: ***'The aim of ruxolitinib is to improve symptoms and control haematocrit (HCT) levels in order to reduce the risk of thromboembolic events and the associated complications which can lead to death'***	The aim of ruxolitinib is to improve symptoms, control HCT to reduce the risk of thromboembolic events and reduce other PV complications.[1]	This is an incomplete statement rather than a factual inaccuracy (the aim of ruxolitinib is not explicitly stated in the cited reference). We have reworded the text on page 25 as suggested to ensure that we accurately reflect the company's interpretation of the aim of ruxolitinib.

Issue 2 Marketing authorisation details for ruxolitinib

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 25, second paragraph states: *'Ruxolitinib is licensed for the treatment of adult patients with PV who are resistant to or intolerant of hydroxycarbamide, and UK marketing authorisation was granted in January 2021.[11]'*	Please amend the sentence to: *'Ruxolitinib is licensed for the treatment of adult patients with PV who are resistant to or intolerant of hydroxycarbamide,* ***and EMA marketing authorisation was granted in January 2015,*** *with* ***a*** *UK marketing authorisation* ~~***was***~~ *granted in January 2021.[11]'*	It would be more appropriate to report the EMA marketing authorisation date here,[2] considering the marketing authorisation was originally granted by the EMA and subsequently adopted by the MHRA.	Not a factual inaccuracy. However, we have added the EMA approval date on page 25 as suggested in the interests of completeness. We have also added EMA to the table of abbreviations on page 8.

Issue 3 Identification of GEMFIN study

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 31, third paragraph states: *'The CS is not transparent about how the GEMFIN registry was identified and included in the SLR results. Nor is it transparent whether the GEMFIN registry is the only source of relevant comparator evidence suitable for use in the company's ITC analysis (see section 3.3.2 for the critique of studies included in the ITC).'*	Please amend the sentence to: *'**~~The CS is not transparent about how the GEMFIN registry was identified and included in the SLR results. Nor is it transparent~~** **A MAIC analysis based on RESPONSE and the GEMFIN registry was included in the SLR results. It is not transparent, however,** whether the GEMFIN registry was the only source of relevant comparator evidence suitable for use in the company's ITC analysis (see section 3.3.2 for the critique of studies included in the ITC).'*	GEMFIN was not included as part of the SLR results but as a study identified separately from the SLR. The published MAIC analysis that included GEMFIN was however included in the SLR. Further clarification is therefore required as to the results of the SLR in relation to GEMFIN.	We agree with the justification for amendment and have reworded the text on pages 31-32 accordingly.
Page 59, fourth paragraph states: *'A relevant PV registry (GEMFIN) is listed among the SLR results in CS Appendix D.1.3 but with no explanation of how the company became aware of it.'*	Please amend the sentence to: *'**A MAIC using data from a relevant PV registry (GEMFIN) is listed among the SLR results in CS Appendix D.1.3.**'*		We agree with the justification for amendment and have reworded the text on pages 59-60 accordingly.

Issue 4 Imbalances in MAJIC-PV baseline characteristics

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 36, second bullet point states: *'There is also an imbalance within the MAJIC-PV trial for one of the indicators of high-risk for PV where the BAT arm is more at risk than the ruxolitinib arm.'*	Please provide further clarification as to which factor this is regarding.	This statement is unspecific as to which factor is considered imbalanced by the EAG. Further clarification is therefore requested.	We have inserted text on page 36 as suggested to clarify that we are referring to prior thromboembolic events.

Issue 5 Timepoints that PSIS was measured in the RESPONSE trial

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 41, Table 7 states that PSIS was reported in the CS at Week 32 for RESPONSE. However, PSIS was reported for Week 32 and Week 256 in the RESPONSE trial and CS.	Please amend to: *'PSIS at Week 32* ***and Week 256****'*	This statement should be amended to accurately reflect the RESPONSE trial and CS.	Thank you for highlighting this error. We have amended the text as suggested in Table 7.

Issue 6 Statistical method MAJIC-PV

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 43, Table 7, under "Analysis of outcomes" states that "The statistical methods appear generally appropriate. NB alpha =0.1 and 80% confidence intervals are applied for the primary outcome giving a relatively high chance of nonsignificant findings being declared significant. No information on whether analyses were checked".	It is unclear where the value of 80% comes from.	We could not find the value cited.	Please note that on page 43 Table 8 was incorrectly cited as Table 7. We have corrected this cross-reference to the table. The cited value of 80% is stated in the trial protocol. We have added text in Table 8 (page 44) to clarify this.

Issue 7 Primary outcome in the MAJIC-PV trial

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 47, Table 11 report "Adjusted [a] OR 2.12 (95% CI 1.25 to 3.60); p=0.02".	Please amend to: "Adjusted [a] OR 2.12 (**90~~95~~**% CI 1.25 to 3.60); p=0.02".	The manuscript reports the 90% CI.	Thank you for highlighting this error. We have made the correction as suggested in Table 11.

Issue 8 Definition of HCT control in RESPONSE

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 45, Section 3.2.6.1 states: *'HCT control as defined by phlebotomy ineligibility and spleen volume at week 32 was the primary outcome in the RESPONSE trial and is referred to as the "primary response".'*	Please amend the sentence to: *'HCT control as defined by phlebotomy ineligibility and **a reduction of ≥35% in** spleen volume **from baseline** at week 32 was the primary outcome in the RESPONSE trial and is referred to as the "primary response".'*	The definition of "primary response" was a reduction in spleen volume, rather than the absolute spleen volume. This amendment will ensure accuracy to this definition.[1]	We have amended the text on page 45 as suggested to improve accuracy of the definition.

Issue 9 Inaccurate description of Alvarez et al (2022)

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 64, Section 3.4.3, last bullet point states: '*A published ITC of patients from GEMFIN reported no statistically significant difference in OS between ruxolitinib and BAT'*	Please amend the sentence to: *'A published **~~ITC~~ real-world comparison** of patients from GEMFIN **treated with ruxolitinib or BAT** reported no statistically significant difference in OS between ruxolitinib and BAT'.'*	This study is a real world comparison of ruxolitinib and BAT and not an ITC.	Thank you for highlighting this misinterpretation. We have reworded the text on page 64 to remove reference to ITC.

Issue 10 Outcomes from the clinical trials

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 48, final paragraph states: *'HCT is included as a component of the complete haematological remission reported in the MAJIC-PV trial (section 3.2.6.3). HCT levels are also reported separately in RESPONSE-2 and in MAJIC-PV.'*	Please amend the sentence to: *'HCT* ***control*** *is included as a component* ~~***of the complete haematological remission reported in the MAJIC-PV trial (section 3.2.6.3).***~~ ***in the primary outcomes of all three ruxolitinib clinical trials****. HCT levels are also reported separately in RESPONSE-2 and in MAJIC-PV.'*	The current wording is misleading. It should be clarified that HCT control was a primary outcome for MAJIC-PV, RESPONSE and RESPONSE-2 and that HCT levels specifically were reported for RESPONSE-2 and MAJIC-PV.	Not a factual inaccuracy. However, we have reworded the text as suggested on page 49 to reduce the possibility of misinterpretation. Please note that on page 49 we have also corrected a missing reference citation to the MAJIC-PV manuscript.

Issue 11 Clarification on the study included in the ITC

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 3.3.1, final paragraph states: *'*As discussed below, the company excluded the RESPONSE-2 and MAJIC-PV trials from the ITC and therefore the ITC is based only on the RESPONSE trial (plus the matching external BAT cohort)*.'*	Please amend the sentence to: *'*As discussed below, the company ~~**excluded the RESPONSE-2 and MAJIC-PV trials from the ITC and therefore the**~~ ITC is based only on the RESPONSE trial (plus the matching external BAT cohort)*.'*	The current wording is misleading. Analysis against RESPONSE-2 was not conducted at the time of the primary analysis due to the absence of events. Furthermore, ITC is not necessary for MAJIC as this is a RCT reporting OS directly comparing ruxolitinib vs. BAT.	Not a factual inaccuracy. However, we appreciate the text may be misleading and we have therefore amended the text on page 59 as suggested.

Issue 12 KM estimates description

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 3.2.6.8, Page 51 states: • "RESPONSE: Median OS 91.9% (95% CI 84.4% to 95.9%) (CS section B.2.7.1)" • "RESPONSE-2: Median OS not reached (CS section B.2.7.2)" • "MAJIC-PV: Median OS not reported; OS hazard ratio, ruxolitinib versus BAT 0.73 (95% CI 0.36 to 1.50; p=0.39 (CS section B.2.11.2)." • "RESPONSE: Median TFS [redacted] (95% CI [redacted]) (CS section B.2.7.1)" • "RESPONSE-2: Median TFS 94% (95% CI 85% to 98%) (CS section B.2.7.2)."	Please amend the sentences to: • "RESPONSE: **~~Median OS~~ KM estimates for OS at 5 years of** 91.9% (95% CI 84.4% to 95.9%) (CS section B.2.7.1). **Median OS not reached**" • "RESPONSE-2: **~~Median OS~~ KM estimates for OS at 5 years of 96% (95% CI: 87% to 99%).** Median OS not reached (CS section B.2.7.2)" • "MAJIC-PV: Median OS not **~~reported~~ reached**; OS hazard ratio, ruxolitinib versus BAT 0.73 (95% CI 0.36 to 1.50; p=0.39 (CS section B.2.11.2)." • "RESPONSE : **~~Median TFS~~ KM estimates for TFS at**	The KM estimates refer to estimate at 5 years. For clarity median OS/TFS typically refer to the time when 50% of patients had an event.	Thank you for highlighting our misinterpretation of the information. We have amended the text in section 3.2.6.8 (page 51) as suggested. (NB we could not find any reference in the submitted documents to the median OS not being reached in RESPONSE)

	5 years of [redacted] (95% CI [redacted]) (CS section B.2.7.1)" • "RESPONSE-2: ~~**Median TFS**~~ **KM estimates for TFS at 5 years of** 94% (95% CI 85% to 98%) (CS section B.2.7.2)."		

Issue 13 Typographical error in trial name

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 63, third paragraph states: *'Sample size was reduced from [redacted] in GEMFIN and 110 in response to [redacted] post-matching. Studies were reasonably well-matched following matching (Table 12, document B), although there was a [redacted]% difference in males.'*	Please amend the sentence to: *'Sample size was reduced from [redacted] in GEMFIN and 110 in **RESPONSE** to [redacted] post-matching. Studies were reasonably well-matched following matching (Table 12, document B), although there was a [redacted]% difference in males.'*	The trial name is missing capitalisation.	Thank you for highlighting this typographical error. We have corrected this on page 63.

Issue 14 Clarification on the line of treatment

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.2.1, Page 69, under heading: "Overview of the model structure", the paragraph states: *"In the base case analysis, the BAT state is partitioned into three sub-states, ...: first-line BAT; second-line or subsequent BAT;...utility declines between first-line, second/subsequent-line and no further treatment substates, see section* ***Error! Reference source not found.*** *below)."*	This should be corrected to: *'In the base case analysis, the BAT state is partitioned into three sub-states, which represent different stages of treatment: first-****~~line~~*** *BAT; second-****~~line~~*** *or subsequent BAT; and no treatment (discontinuation of all BAT). The company use this BAT partition to model progressive decline in health-related quality of life as patients move through the BAT regimens: utility declines between first-****~~line~~****, second/subsequent-****~~line~~*** *and no further treatment substates, see section* ***Error! Reference source not found.*** *below).'*	Patients enter the model following resistance to intolerance to HC/HU (2nd line+). The use of the term line of treatment may therefore be misinterpreted. We suggest removing the word "line" to avoid any misinterpretation.	We agree that this is potentially confusing and have made the requested changes on page 70.

Issue 15 Inaccurate description of the modelling approach for the primary analysis

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.1, Page 70 – under heading "Approach to estimation of transition probabilities", the paragraph states: "*The STM structure requires probability estimates for transitions between the ruxolitinib, BAT and death states, and for transitions between the three BAT substates. These probabilities are estimated from OS and time to treatment discontinuation (TTD) survival curves extrapolated from trial data*."	Please amend the sentence to: "*The STM structure requires probability estimates for transitions between the ruxolitinib, BAT and death states~~, and for transitions between the three BAT substates.~~ These probabilities are estimated from OS and time to treatment discontinuation (TTD) survival ~~curves extrapolated~~ data from* ***the*** *trial* ~~***data***~~*.* ***The BAT health state is partitioned onto three sub-health states***."	Transitions between the three BAT sub-health states are not explicitly modelled. Instead, the BAT health state is partitioned. For the primary analysis, the model does not use OS directly. Instead, OS for ruxolitinib is estimated from (1) TTD excluding death, (2) time to discontinuation from death only (otherwise referred as pre-discontinuation survival) and (3) the survival following ruxolitinib discontinuation (also referred as post-discontinuation survival).	We have removed the reference to 'transitions between the three BAT substates' from this paragraph as it confuses the main point, which relates to estimation of OS (page 70).
Section 4.2.1, Page 70 – under heading "Approach to estimation of transition probabilities", the paragraph states: "*The company describe their approach to fitting TTD and OS extrapolations in CS sections B.3.1.2 to B.3.3.4*"	Please amend the sentence to: "*The company describe their approach to fitting* ~~***TTD and OS extrapolations***~~ ***survival outcomes*** *in CS sections B.3.1.2 to B.3.3.4*"		We understand that TTD and OS extrapolations are not directly fitted in the primary model, so have reworded this sentence as follows: "The company describe their approach to estimating time to treatment discontinuation and overall survival…" (page 71).

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.1, Page 70 – under heading "Approach to estimation of transition probabilities", 2nd bullet point, the paragraph states: "*As the numbers of deaths observed in the trials were low, pre- and post-discontinuation survival for the ruxolitinib arm are estimated from a competing-risk analysis of pooled data from the RESPONSE and RESPONSE-2 trials*"	Please amend the sentence to: "*As the numbers of deaths observed in the trials were low, pre- and post-discontinuation survival for the ruxolitinib arm are estimated from* ***~~a competing-risk analysis of~~*** *pooled data from the RESPONSE and RESPONSE-2 trials*"	Only time to discontinuation due to death and discontinuation due to reasons other than death are competing (competing events). It is therefore not accurate to state that post-discontinuation survival is estimated from a competing risk analysis	We have removed the reference to the competing-risk analysis in this sentence (page 71).
Section 4.2.1, Page 71 – under heading "Other model parameters", the paragraph states: "*In addition to TTD and OS extrapolations, the model uses input parameters to estimate incidence rates for key events and adverse reactions, utilities and resource use/costs*"	Please amend the sentence to: *"In addition to TTD and* ***~~OS~~ extrapolations pre-and post-discontinuation survival****, the model uses input parameters to estimate incidence rates for key events and adverse reactions, utilities and resource use/costs*."		We have made the requested change (page 72).

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.6, Page 75 – under heading "Treatment effectiveness and extrapolation", the paragraph states: "*The clinical parameters used in the model consist of time to treatment discontinuation (TTD), overall survival (OS) and incidence rates for key complications, therapeutic phlebotomy and adverse events*"	Please amend the sentence to: "*The clinical parameters used in the model consist of time to treatment discontinuation (TTD),* ***transitions between health states,*** *overall survival (OS)* ***when appropriate*** *and incidence rates for key complications, therapeutic phlebotomy and adverse events*."		The suggested wording is confusing. We have amended this sentence as follows: "The clinical parameters used in the model consist of time to treatment discontinuation (TTD), **parameters required to estimate** overall survival (OS) and incidence rates for key complications, therapeutic phlebotomy and adverse events." (page 76)
Section 4.2.6, Page 75 – Table 20 states: "*Competing-risk analysis of pooled RESPONSE and RESPONSE-2 IPD (same extrapolations for both subgroups). Data pooled due to small number of deaths observed within the trials*"	Please amend the sentence to: "~~***Competing-risk analysis of pooled RESPONSE and RESPONSE-2 IPD (same extrapolations for both subgroups).***~~ *Data* ***from RESPONSE and RESPONSE-2 were*** *pooled due to small number of deaths observed within the trials* ***(same extrapolations for both subgroups).***"		We have made the requested change in Table 20 (page 76).

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.6.2.2, Page 82, the paragraph states: "*The OS for ruxolitinib was modelled indirectly using the extrapolations of TTD described above, and extrapolations of pre-discontinuation survival and post-discontinuation survival estimated from a competing-risk analysis of pooled data from the RESPONSE and RESPONSE-2 trials (see CS B.3.1.2 and B3.3.2)*."	This should be corrected along the lines of: "*The OS for ruxolitinib was modelled indirectly using the extrapolations of TTD* ***excluding death*** *described above, and extrapolations of* ***pooled data for*** *pre-discontinuation survival and post-discontinuation survival* ***~~estimated from a competing-risk analysis of pooled data from the RESPONSE and RESPONSE-2 trials~~*** *(see CS B.3.1.2 and B3.3.2).*"		We have made the suggested change (page 83).
Section 4.2.6.2.2, page 83, under heading: "EAG comments on the ruxolitinib OS extrapolation (primary analysis):", the paragraph states: "*The competing-risk framework used for estimation of the OS extrapolation for ruxolitinib is appropriate for the STM structure of the company's primary analyses*."	This should be corrected to: "*The* ***use of a*** *competing-risk framework* ***~~used for estimation of the OS extrapolation~~ to estimate TTD, and subsequently OS*** *for ruxolitinib is appropriate for the STM structure of the company's primary analyses."*		We have made the suggested change (page 84).

Issue 16 Inaccurate description of the model in TA386 and TA756

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.2.3, Page 72 – under heading "EAG comments on model structure, states and events (Key issue 5)", the paragraph states: *"The company cite TA386 and TA756 appraisals as precedent for the use of therapy-based health states for MF. However, a 'supportive care' state after discontinuation of treatment for MF was used in TA386 and TA756. However, we suggest that the supportive care state would be more directly related to decline in quality of life than the post-ruxolitinib BAT state for PV in the current appraisal."*	Please consider removing this sentence or amending this sentence to be accurate.	In TA386 and TA756, patients initiated on ruxolitinib (TA386) or fedratinib (TA756) move to BAT following discontinuation of treatment (ruxolitinib/fedratinib), where they remain on BAT until close to end of life and move to the supportive care health state (last few months of life). In TA386/TA756, patients on active treatment (ruxolitinib/fedratinib) have a significantly greater quality of life compared with those that discontinue treatment and receive BAT, with the quality of life deteriorating further at the end of life (supportive care in MF). Data from the RESPONSE Trial show that quality of life for patients on ruxolitinib is significantly greater compared with that on BAT (not related to supportive care). We therefore believe that the statement is inaccurate as quality of life in TA386 and TA756 is	This is not a factual inaccuracy, but we have edited the following sentence to indicate greater uncertainty over this suggestion: "We suggest that the supportive care state **may** be more directly related to decline in quality of life than the post-ruxolitinib BAT state for PV in the current appraisal." (page 73) We do understand that, as in the TA386 and TA756 models for MF, quality of life in the current PV model differs between the 'On active treatment' and 'On BAT' health states. We agree that inclusion of this treatment-related difference in the current model is supported by the EQ-5D results from the RESPONSE trial. However, in this bullet point we are questioning whether it is the case that: "*Similar to the approach used in MF in TA386 and TA756, outcomes with respect to HRQoL and costs are largely defined by a patient's phase in the management of the condition*" (CS section B.3.2.2 page 93) We note that TA386 and TA756 models included a supportive care health state after BAT. The company comment that modelled time in this state was short ('a few months'), with the implication that it does not have a large impact on the results.

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
		different between patients on active treatment and those that discontinue and move to BAT. A supportive care health state was included in TA386 and TA756 as MF is considered to be a more progressive disease compared with PV.	We could not find the time spent in the supportive care state from the publicly available committee papers for TA386. However, Figure 1 in the Novartis PAS submission (28 July 2015) shows that change in HRQoL for supportive care was the third most influential parameter in the univariate sensitivity analysis, and the cost of red blood cell units on supportive care was the 7th most influential parameter. Thus the inclusion of a supportive care health state was clearly an important element of this model. It is more difficult to assess the impact of the supportive care state in the TA756 model, as uncertainties over the structure were not fully resolved. The TA386 and TA756 models also differed from the current model in other respects: both were individual-level simulations and included measures of response as well as the health states; and TA756 also included an AML health state. We recognise that PV is less progressive than MF, and that a different modelling solution may well be appropriate. We have included the question of model structure as a key issue in our report and will welcome further discussion on this point.

Issue 17 Appropriateness of using health states based on therapy phases

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.2.3, under heading "EAG comments on model structure, states and events (Key issue 5)", first bullet point states: *"The company's decision to use therapy phases as states, rather than stages of disease, means that their model structure does not reflect the natural history of PV"*	Please amend the sentence to: *'The company's decision to use therapy phases as states, rather than stages of disease, means that their model structure* ***~~does~~ may*** *not reflect the natural history of PV'*	The current wording is misleading. Clinical experts consulted agreed that the structure reflect the natural history of PV	This is not a factual inaccuracy. This statement is under the heading of "EAG comments", and is a conclusion presented by the EAG. The rationale for this conclusion is presented in the sentences that follow the statement.

Issue 18 Errors in HR reported to be used in the model for OS in the primary analysis

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.6, Table 20 – Page 76 Table 20 includes an error in the HR used for OS. The values are reported as "0.91" and "0.45".	This should be corrected to: "1.10" and "2.20".	The model uses the inverse of the HR reported in the MAJIC-PV manuscript as the HRs are applied to the ruxolitinib OS curve to derive BAT OS.	Thank you for highlighting this error. We have corrected this in Table 20 (page 77).

Issue 19 Absence of scenario pooling data for TTD

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.6.1.1 – Page 77, under heading "Ruxolitinib discontinuation", the paragraph states: "*The company also provided a scenario with ruxolitinib TTD estimated from pooled RESPONSE and RESPONSE-2 data (CS appendix N.3), but this had little impact on the ICERs*"	Please remove this sentence.	No scenario has been presented using the pooled RESPONSE/RESPONSE-2 TTD excluding death.	Thank you for pointing this out, we agree that there is no corresponding scenario. The error has occurred due to a scenario in Appendix P, page 161 labelled: "Do not pool Rux TTD excl death". The resulting ICERs are actually due to not pooling pre-discontinuation survival ("Pool prePS = No" in the model). We have removed this sentence from the report (page 78).

Issue 20 Incorrect reference in the EAG report

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.6.1.1 – Page 77, under heading "Ruxolitinib discontinuation" The text includes an error in the section referred to. The EAG report refers to "*Section B.1.2*".	This should be corrected to: "*section B.3.1.2*".	Incorrect reference to the section in the CS.	Thank you for highlighting this error. We have corrected this on page 78 as suggested.
Section 4.2.6.1.2 – Page 80, under heading "EAG comments on TTD extrapolations" 3rd bullet point The text includes an error in the Figure referred to. The EAG	This should be corrected to: "*Figure 1 and 2*".	Incorrect reference to the Figure in the EAG report.	Thank you for highlighting this error. We have corrected this on page 81 as suggested.

report refers to "*Figure 1 and Figure 1*".			
Section 4.2.8 – Page 89, under heading "Systematic literature review of costs and healthcare use" The text includes an error in the Tables referred to. The EAG report refers to "*CS Tables 45 and 45*".	This should be corrected to: "*CS appendix Table 45 and 46*".	Incorrect reference to the CS Table in the EAG report.	Thank you for highlighting this error. We have corrected this on page 90 as suggested.
Section 5.2.3 – page 96, under heading "Probabilistic sensitivity analysis for the company's base case for the MAJIC-PV population" The text includes an error in the Tables referred to. The EAG report refers to "*CS Figure 52*".	This should be corrected to: "*CS Figure 55*".	Incorrect reference to the CS Figure 55.	Thank you for highlighting this error. We have corrected this on page 98 as suggested.
Section 3.2.6.6 – page 50 The text includes an error. The 3rd bullet point states "*71% in MAJIC-PV (**Error! Reference source not found.** below*)".	This should be corrected to: "*Table 14 below"*.	Incorrect reference to the Table in EAG report	Thank you for highlighting this error. We have corrected this on page 50 as suggested.

Issue 21 Use of the EQ-5D 3L in scenario analysis (incorrectly reported in the CS)

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.7.2 – Page 86, the paragraph states: "*The company also report a scenario analysis based on EQ-5D-5L data from the RESPONSE-2 trial*"	This should be corrected to: "*The company also report a scenario analysis based on EQ-5D-**3L**~~**5L**~~ data from the RESPONSE-2 trial*"	The mapped EQ-5D 3L is used in scenario analysis. This was incorrectly reported in the CS as the EQ-5D 5L.	Not a factual inaccuracy. The study protocol and CSR for RESPONSE-2 specify that the EQ-5D-5L questionnaire was used. We also note in the following sentence that the NICE recommended Hernández Alava et al. 2020 algorithm was used to derive the UK utility values. For clarity we have added on page 88 that these are **3L** utility values.

Issue 22 Incorrect reporting of costs

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 4.2.8.4 – Page 92 The text includes an error in the cost for management of low/int1 risk MF used in the model. The value is reported as *"£71,190"*	This should be corrected to: "*£72,190*".	Error in transcription of costs used in the economic model.	We apologise for the error; the cost has been corrected (page 93).
The paragraph also states: "*double the cost of high-risk MF in BAT patients*"	This should be corrected to: "*double the cost of **int2**/high-risk MF in BAT patients*"		We have corrected the statement to include "intermediate-2" (page 93).

Issue 23 Clinical expert involvement in the MAJIC-PV trial

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 5.3.1 – Page 97, the paragraph states: "*Four of the five clinical experts who attended the latter advisory meeting were also involved in the MAJIC-PV trial*"	This should be corrected to: "*Four of the five clinical experts who attended the* ***~~latter~~*** *advisory meeting****s*** ***~~were also involved in the~~*** ***are authors of the*** *MAJIC-PV trial*"	4 out of the 5 clinical experts are listed as authors.	We have revised the sentence as requested (page 98).

Issue 24 Clarification regarding choice of extrapolation

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 5.3.2.2 – Page 100, under heading: "EAG comments on extrapolation distributions", the paragraph states: "*The company make note of the potential for spline models to overfit the data.*"	This should be corrected to: "*The company make note of the potential for spline models* ***with more than one knot*** *to overfit the data*"	The CS comment that splines models with more than one knots may lead to data over-fitting. Spline models with one knot are unlikely to lead to over-fitting. Data over-fitting is likely to happen as the number of knot increase.	The EAG have made the relevant changes on page 101.
Page 18, Issue 6, under heading "description of issue", the paragraph states: "*the company noted that the odds spline model implemented for the extrapolation of TTD for ruxolitinib due to reasons other than death has the potential to*	This should be corrected to: *"****The company used an odd spline odd model with one knot and note of the potential for spline models with more than one knot to overfit the data****"*		The EAG have made the relevant changes on page 18.

overfit the data in the primary analysis."			

Issue 25 Direction of the ICER

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 5.3.3.2 – Page 101, under heading: "EAG scenario analysis for the cost of a grade 1-2 thromboembolic event", the paragraph states: "*This results in a small increase in the ICERs*."	This should be corrected along the lines of: "*This results in a small* ***~~increase~~ improvement*** *in the ICERs*"	The EAG scenario improve the ICERs slightly.	We apologise for the error. The sentence has been corrected on page 102.

Issue 26 Error in baseline characteristics data

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 120–123, Appendix 9.2. Presence of JAK2 V617F mutation GEMFIN BAT data value is reported as [redacted]	Please amend these data to the updated sample value of [redacted] in line with Table 2 of the company's clarification questions response. Footnote g can subsequently also be removed.	The data currently reported for JAK2 mutation are for the full adjusted population published in Alvarrez-Larran et al. 2018, whereas the rest of the data in the column report for the N=184 adjusted population. This should be adjusted for consistency in this table and to align to the clarification question responses.	Thank you for highlighting the inconsistency of the data sources in Appendix 9.2. We have corrected this as suggested.

Confidentiality highlighting amendments

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG response
Page 52, Table 15	AIC highlighting is required for EQ-5D-5L health index scores in the RESPONSE-2 trial as these data are not anticipated to be published.	Please highlight all data in Table 15 as AIC.	We have added AIC highlighting to the data in Table 15 as suggested.
Page 60, first bullet point	AIC highlighting is required for the number of patients in the GEMFIN registry as of October 2016 as this is not anticipated to be published.	*'The Spanish Registry of Polycythemia Vera set up in 2011 by GEMFIN (Grupo Español de Enfermedades Mieloproliferativas Filadelfia Negativas) referred to as the GEMFIN registry, is one of the largest registries of PV (N=[redacted] as of October 2016).'*	We have added AIC highlighting to the number of patients on page 60 as suggested.
Page 61, second paragraph	AIC highlighting is required for: • Proportion of patients with diabetes (RESPONSE and GEMFIN) • Proportion of patients with diagnosis of PV (GEMFIN only) • Proportion of patients with leg ulcers (RESPONSE and GEMFIN) These data are not anticipated to be published.	*'There are notable imbalances in terms of age (61 vs 69 years), cytopaenia at lowest hydroxycarbamide dose (15% versus 7%), male sex (60% versus 47%), time since diagnosis of PV (8.9 versus [redacted] years), and diabetes ([redacted] versus [redacted]). JAK2 mutation status and leg ulcers also showed differences between studies.(JAK2: 95% versus 89%; leg ulcers: [redacted] versus [redacted]). However, company experts did not rank either highly as a prognostic factor, and the EAG's experts concurred.'*	We have added AIC highlighting on pages 60-61 as suggested.

Page 62, Section 3.4.1, first paragraph	Median GEMFIN follow up requires AIC highlighting as it is not anticipated to be published.	*'The GEMFIN data for the matching were obtained in 2016 (median follow up [redacted]).'*	
Page 62, Section 3.4.1, second paragraph	RESONSE data for failure to reduce massive splenomegaly requires AIC highlighting as these data are not anticipated to be published.	*'No explanation is given for excluding failure to reduce massive splenomegaly but the variable is relatively balanced between studies ([redacted] versus 1%).'*	We have added AIC highlighting on page 62 as suggested.
Page 63, Section 3.4.2, third paragraph	Data from the ITC analysis require AIC highlighting as these data are not anticipated to be published.	*'Sample size was reduced from [redacted] in GEMFIN and 110 in RESPONSE to [redacted] post-matching. Studies were reasonably well-matched following matching (Table 12, document B), although there was a [redacted]% difference in males.'*	We have added AIC highlighting on page 63 as suggested.
Page 93–106, Tables 23–26, Table 29	Life-years gained (LYG) data do not require CIC highlighting.	It has been agreed with NICE that the LYG data from the economic model may be unredacted. Please remove all CIC highlighting from LYG data in Tables 23–26, and Table 29.	We have removed the CIC highlighting from the LYG column in Tables 23-26 and Table 29.
Page 120–123, Appendix 9.2.	Baseline characteristics of GEMFIN require AIC highlighting as these data are not anticipated to be published. Baseline overall length of spleen by ultrasound for MAJIC-PV also requires AIC highlighting as these data are not yet published.	Please highlight all data in the GEMFIN column of the table presented in Appendix 9.2 as AIC. Please also highlight baseline overall length of spleen by ultrasound for both ruxolitinib and BAT groups in MAJIC-PV as AIC.	We have added AIC highlighting to all GEMFIN data and the MAJIC-PV overall spleen length data in Appendix 9.2 as suggested. We have also corrected the MAJIC-PV overall spleen length data in Appendix 9.2 in accordance with clarification response C1.

References

1. Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera. New England Journal of Medicine 2015;372:426-435.
2. European Medicines Agency (EMA). Jakavi® (ruxolitinib).

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

B.1 Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **20 January 2023**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

B.2 About you

Table 1: About you

Your name	XXXXXXXXXXX
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Novartis Pharmaceuticals UK Ltd
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	Since April 2005 Novartis has exclusively licensed glycopyrronium bromide and certain intellectual property relating to its use and formulation from Vectura and its co-development partner, Sosei Heptares. The following inhaled medications are comprised of, or contain glycopyrronium bromide: • Seebri® Breezhaler® (glycopyrronium bromide) (used as a maintenance treatment for Chronic Obstructive Pulmonary Disease [COPD]) • Ultibro® Breezhaler® (indacaterol/glycopyrronium bromide) is used as a maintenance treatment for COPD • Enerzair® Breezhaler® (indacaterol/glycopyrronium bromide/mometasone furoate) is used as a maintenance treatment for asthma uncontrolled with long-acting beta-agonists/inhaled corticosteroids (LABA/ICS) Phillip Morris International (a tobacco company) is currently in the process of acquiring Vectura Group plc.

B.3 Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2: Key issues

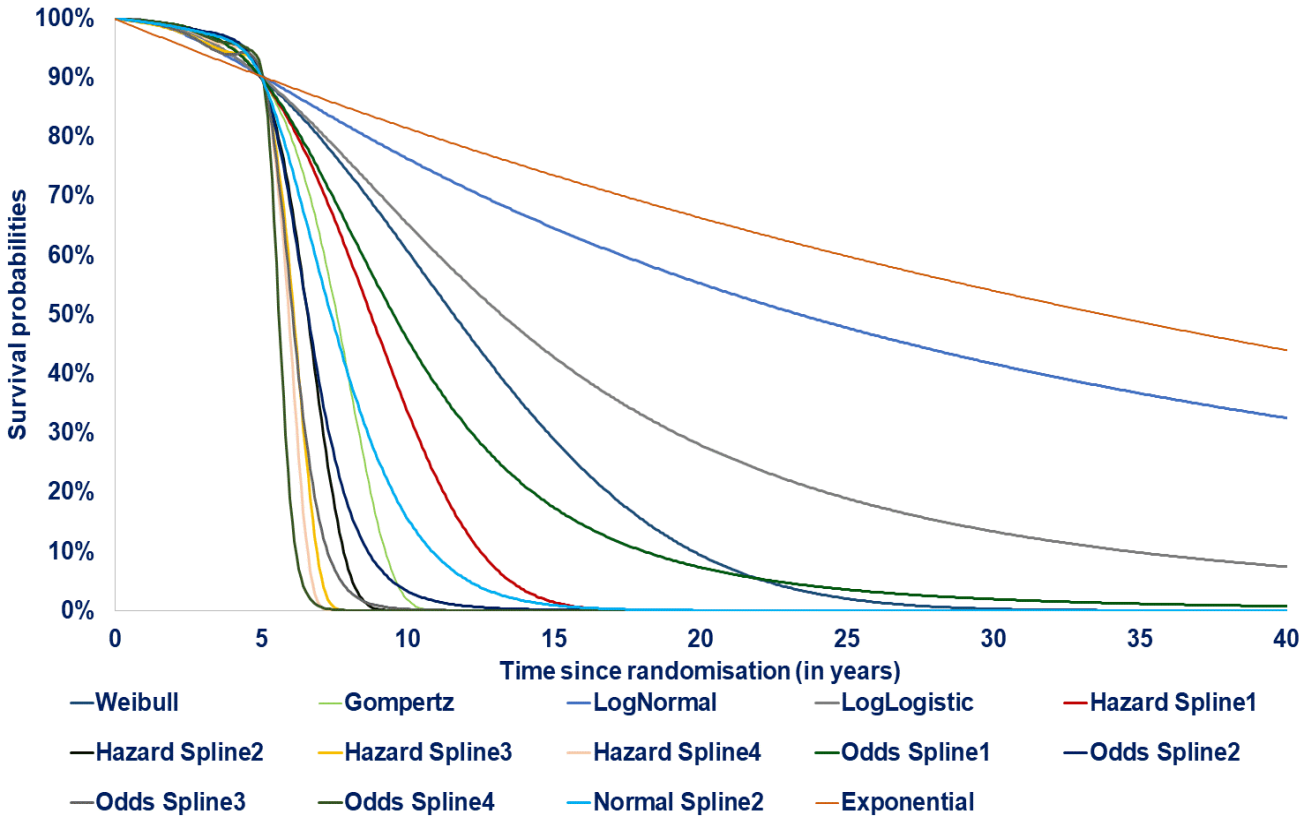
Key issue	Does this response contain new evidence, data or analyses?	Response
Issue 1: Relevance of the trial populations for modelling UK practice	*No*	The MAJIC-PV trial is a UK Phase 2 investigator-led study, which recruited high-risk patients with polycythaemia vera (PV) who are resistant to or intolerant (R/I) to hydroxycarbamide/hydroxyurea (HC/HU). In this trial, high-risk was defined as ANY of the following:[1] • Age >60 years; • Previously documented thrombosis (including transient ischemic attack [TIA]), erythromelalgia or migraine (severe, recurrent, requiring medications, and felt to be secondary to the myeloproliferative neoplasms [MPN]) either after diagnosis or within 10 years before diagnosis and considered to be disease related; • Significant splenomegaly (>5 cm below costal margin on palpation) or symptomatic (splenic infarcts or requiring analgesia); • Platelets >1000 x 109/L; • Diabetes or hypertension requiring pharmacological therapy for >6 months. Therefore, as highlighted by the external assessment group (EAG), the MAJIC-PV trial is directly relevant to the UK setting (EAG report, Section 4.2.3). It should be noted that the definition for high-risk used in the MAJIC-PV trial is also broader than that defined by the European LeukemiaNet (ELN).[1]

		The RESPONSE trials (RESPONSE in PV patients with splenomegaly and RESPONSE-2 in PV patients without splenomegaly) were Novartis sponsored multinational trials, with clinical experts consulted by the company and the EAG (EAG report, Section 3.2.2) indicating that the patients' baseline characteristics are generally reflective of patients with PV who are R/I to HC/HU in the UK.[2, 3] Following the EAG report, further clinical opinion was sought. Clinical experts reconfirmed that the population recruited in the MAJIC-PV trial is likely to represent most patients for whom ruxolitinib would be given in clinical practice. In line with earlier advice, they noted that there is no accepted definition of high-risk, and that any recommendation should not be limited and that clinicians should be able to decide patients to whom ruxolitinib should be offered. Novartis therefore believe that all three populations represent the entire adult patient population with PV who are R/I to HC/HU who would benefit from ruxolitinib and therefore the results of all three analyses are applicable and required for decision-making. There is therefore no change to the company submission (CS) base-case.
Issue 2: Modelling the relative treatment effect for overall survival	*No*	The company's base-case uses a time varying treatment effect for overall survival (OS) estimated from MAJIC-PV (hazard ratio [HR]: [redacted]).[1] The time intervals of 0 to 3 years and 3 to 5 years were decided upon following clinical opinion, visual inspection of the Kaplan–Meier (KM) data in MAJIC-PV and statistical analysis of change in the hazard over time. In contrast to this time varying-treatment effect, the EAG stated a preference to use a constant HR of 0.73 (95% CI 0.36 to 1.50; p=0.39) for the entire duration (0 to 5 years).[1] Novartis continue to believe that using a time-varying treatment effect better reflects the data and is more appropriate. As stated in the CS, clinical advisors noted the curves for OS started to diverge after approximately 3.0 years in the MAJIC-PV trial, which was in line with their expectation that a survival difference would not manifest immediately.[1] Further clinical input was sought as part of this technical engagement response and confirmed that using a time-varying treatment effect would be more reflective of the data and their clinical expectation. Clinical experts did not consider the use of a constant treatment effect to be appropriate or reflective of the data. Novartis further believe that should a constant treatment effect be used, as the case in the EAG base-case, treatment waning assumptions should be removed in line with the EAG's comment and its own clinical expert. The treatment effect estimated from the MAJIC-PV trial is used for all three analyses, in patients with and without splenomegaly in the absence of individual patient level data (IPD) from the MAJIC-PV trial. Further clinical opinion was

		sought and confirmed that there is no reason to believe the treatment effect in patients with and without splenomegaly to be different. There is therefore no change to the CS base-case.
Issue 3: Waning of the treatment effect	*No*	The company's base-case analysis used a conservative treatment waning assumption, where the treatment effect diminishes linearly from the end of trial follow-up (5 years) and stops at 20 years in which there would be no treatment effect (HR=1). This assumption was based on clinical expert opinion that approximately twice the number of patients would be alive at 20 years with ruxolitinib compared with current treatment. Novartis welcome the EAG's comment suggesting that it may be more appropriate to use a longer waning period, or to remove waning from the model entirely when used in combination with the more conservative fixed HR estimate. The EAG report further states that the "*EAG clinical advisors have suggested that they do not have reason to expect that the effectiveness of ruxolitinib would wane over time*". However, in their base-case, the EAG uses the company's waning assumption but at a constant treatment effect for OS (that does not align with the data and clinical expectation as discussed in Issue 2). While Novartis continue to believe that using a time-varying treatment effect is more appropriate, should a constant treatment effect be used (as demonstrated in the EAG base-case), waning should be removed in line with the EAG comments and their own clinical expert opinion. Further clinical opinion was also sought, which confirmed the EAG's clinical expert opinion that there is no reason to expect the effectiveness of ruxolitinib to wane over time as patients remain on treatment for an extended duration. Clinical experts further noted that ruxolitinib delivers stable response. There is therefore no change to the CS base-case.

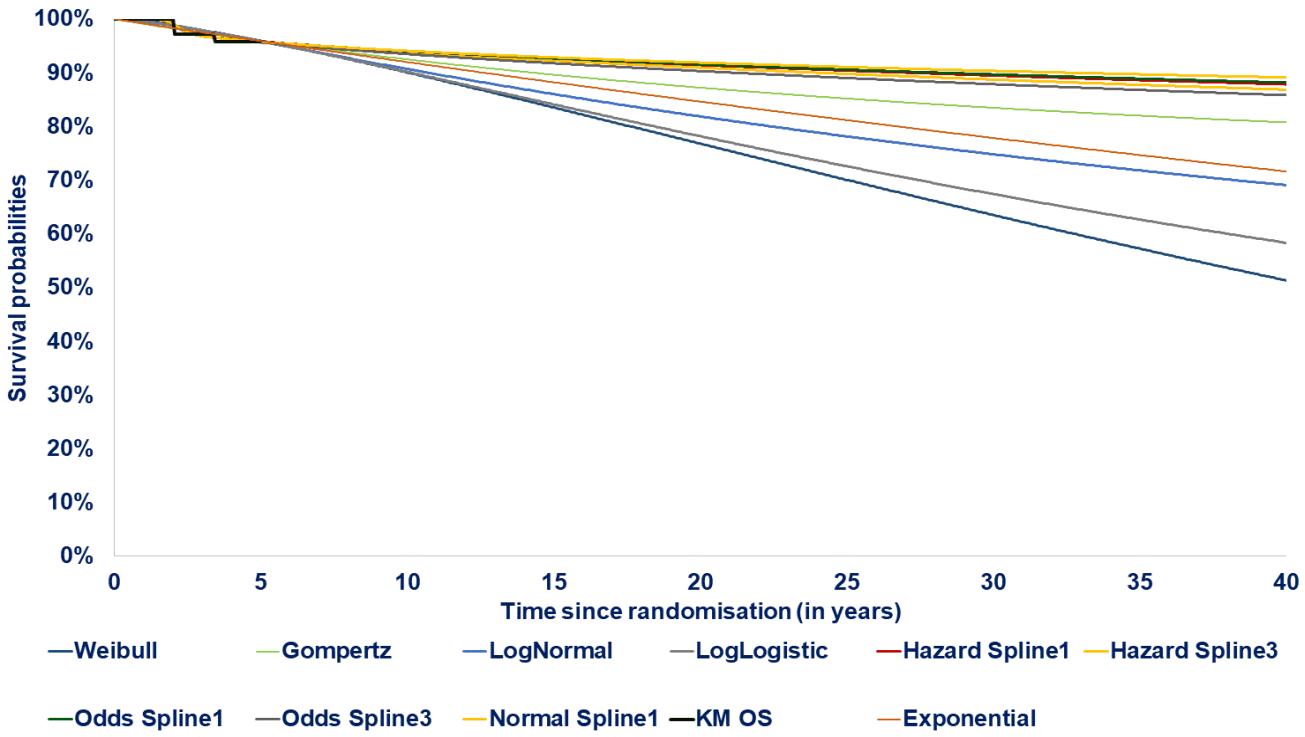
Issue 4: Modelling approach: state-transition or partitioned-survival	*Yes*	For the analyses based on the RESPONSE-trials, the company used a state-transition model (STM) approach for ruxolitinib whereby OS is estimated as a function of the time on treatment (ToT). This was primarily justified due to challenges in modelling survival directly (immaturity of the data with approximately 91% and 96% of patients still alive at the end of the follow-up periods for RESPONSE and RESPONSE-2, respectively) and to facilitate modelling of ToT (to account for the increasing likelihood of discontinuation due to death as patients age).[4, 5] A partitioned survival model (PSM) is used in the MAJIC-PV analysis in the absence of IPD. The EAG note that the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 19 reports that the STM and PSM models can give very different results and recommend the parallel development of both approaches when possible.[6] The EAG notes that results from the three analyses presented differ and it is unclear if the differences are due to the modelling approach or differences in population. To help identify and explore some of the uncertainty around the modelling approach, the EAG suggests that the company develop a PSM for the RESPONSE analysis as data are available to the company to perform such analysis. The company believes that the STM modelling approach presented in the CS for the RESPONSE-analyses is the most appropriate, however, we understand the EAGs request and need to separate out and identify sources of uncertainty. As such, for transparency the company attempted to conduct a PSM for the RESPONSE-analyses whereby OS is modelled directly (e.g., extrapolated beyond the clinical trial). IPD was obtained from the RESPONSE (n=110) and RESPONSE-2 (n=74) trials with a total of 10 (9.1%) and 3 (4.1%) deaths reported in the RESPONSE and RESPONSE-2 trials, respectively.[4, 5] In accordance with the NICE DSU TSD 14, a range of standard parametric distributions (exponential, Weibull, Gompertz, loglogistic, lognormal and generalised gamma) and more flexible models (hazard, normal and odds spline models with up to four knots) were explored in the extrapolation of the clinical trial data beyond the observed period.[7] The spline models (hazard, normal, odds) with one and four knots were estimated in R using the FlexSurv package. The KM and fit with the different distributions to OS are shown below in Figure 1 for RESPONSE and Figure 2 for RESPONSE-2.

Figure 1: Comparison of the KM and parametric distribution fits to OS for ruxolitinib in RESPONSE



Abbreviations: KM: Kaplan–Meier; OS: overall survival.

Figure 2: Comparison of the KM and parametric distribution fits to OS for ruxolitinib in RESPONSE-2



Abbreviations: KM: Kaplan–Meier; OS: overall survival.

		Due to the immaturity of the data resulting from the small number of events, there are significant variations in predictions between curves for the RESPONSE analysis, with a limited number of curves producing plausible extrapolations. For the RESPONSE-2 analysis, all extrapolations (prior adjustment for general population mortality) lead to a predicted survival above 63% at 30 years for ruxolitinib which is not plausible given the age of the cohort (mean age: 64.4 years [standard deviation {SD}]: 11.30). The company therefore does not believe that using a PSM for the RESPONSE-analyses is appropriate for decision-making. Nevertheless, in order to support the EAGs considerations, results are presented for information to provide further clarity to the committee and EAG. For the RESPONSE analysis, only the Weibull, log-logistic and odd spline model with one knot generated plausible predictions for OS, with the Weibull distribution in the middle and aligning the most with clinical expectation. For the RESPONSE-2 analysis, none of the distributions generated a plausible prediction for OS (before adjustment for general population mortality). The Weibull distribution is used for both analyses as this was deemed the most plausible extrapolation for RESPONSE and provided the lowest survival for RESPONSE-2. Other inputs remain identical with the original CS, with the exception of the addition of a constraint for time to treatment discontinuation (TTD) to ensure that TTD in the original CS is below OS for the PSM analysis. Incremental cost-effectiveness ratios (ICERs), quality-adjusted life years (QALYs) and costs for the company base-case using the STM approach and the alternative PSM are shown below in Table 1. Results are presented using the patient access scheme (PAS) discount agreed for myelofibrosis (MF) ([redacted]). Although results are similar between the PSM and STM for the RESPONSE-analyses and may provide some reassurance, Novartis urges caution in the interpretation of results from the PSM due to the immaturity of the data. This could be due to chance. While some of the differences in results between the MAJIC-PV and RESPONSE-analysis are likely to be inherently due to the modelling approach (recognition that different approaches generate different results within the same dataset), differences are more likely to be due to different inputs between analyses.

Table 1: ICERs from CS base-case and model using PSM approach (deterministic) – PAS discount agreed for MF

	Ruxolitinib			BAT			Inc cost	Inc LY	Inc QALY	ICER
	Cost	LY	QALYs	Cost	LY	QALYs				
RESPONSE										
CS base-case	[redacted]	11.45	[redacted]	[redacted]	9.28	[redacted]	[redacted]	2.17	[redacted]	[redacted]
Alternative structure	[redacted]	9.68	[redacted]	[redacted]	8.06	[redacted]	[redacted]	1.62	[redacted]	[redacted]
RESPONSE-2										
CS base-case	[redacted]	12.25	[redacted]	[redacted]	10.46	[redacted]	[redacted]	1.79	[redacted]	[redacted]
Alternative structure	[redacted]	14.18	[redacted]	[redacted]	12.96	[redacted]	[redacted]	1.22	[redacted]	[redacted]

Abbreviations: BAT: best available therapy; CS: company submission; ICER: incremental cost-effectiveness ratio; Inc: incremental; LY: life years; MF: myelofibrosis; PAS: patient access scheme; PSM: partitioned survival model; QALYs: quality-adjusted life years.

In conclusion, as justified in the CS (Section B.3.2.2), while the RESPONSE-trials provided 5-year data, the survival data for ruxolitinib are immature with approximately 91% and 96% of patients still alive at the end of the follow-up period for RESPONSE and RESPONSE-2, respectively, making it challenging to model OS directly (and therefore use a PSM approach). [4, 5] For transparency, and in response to the EAG request to help identify the source of uncertainty, the company provided an analysis using a PSM approach. Nevertheless, the company does not consider the PSM to be suitable for decision-making due to the large uncertainty and immaturity of the data, and the company believes that the STM modelling approach presented in the CS is the most appropriate. Furthermore, despite the absence of IPD, an alternative model structure using a STM approach was explored for MAJIC (Issue 5), albeit with some assumptions providing further reassurance of the validity of the CS base-case. While it is difficult to accurately identify the source of the differences in ICERs between the RESPONSE-analyses (which uses a STM) and the MAJIC-PV analysis (which uses a PSM), Novartis believes that differences are likely to be due to a combination of the modelling approach and population. There is therefore no change to the CS base-case.

Issue 5: Model structure: health states and events	*Yes*	The EAG raised a number of concerns with the model structure and proposed health states in the original CS. In particular, the EAG had the following concerns: • (1) Health states were based on therapy phases (on ruxolitinib & on best available therapy [BAT]), rather than a measure of disease progression (progression-free survival [PFS] or event-free survival [EFS]); • (2) PFS or EFS are more likely to be prognostic to survival compared with treatment discontinuation; • (3) The lack of data informing the partitioning of the BAT health state; • (4) Not accounting for the increased incidence of events over time (use of a fixed incidence rate of events). To address some of these uncertainties, the EAG suggested that the company should consider an alternative model structure based on a measure of disease progression such as PFS and EFS, both of which are reported in the MAJIC-PV manuscript,[1] in addition to simplifying the approach to modelling the subsequent type of events. As justified in the CS (CS, Section B.3.2.2), the structure and health states were selected and validated following discussion with clinical experts with the aim to: • (1) Capture the key contributors of quality of life (based on therapy phases as in TA386 and TA756 for MF) and difference in quality of life between patients on ruxolitinib and BAT in the RESPONSE-trials;[5, 8-10] • (2) Reflect the natural history of PV (inclusion of PV-complications as events), while also taking into account the limitations of the data available (the sample size in the trials and the number of events are too small to robustly construct a model based on events and there is high and early cross-over in the RESPONSE-trials) and absence of IPD from the MAJIC-PV trial.[5, 8] Additional uncertainty and complexity associated with modelling PV-related complications as health states is acknowledged in the EAG report (EAG report, Section 4.2.2.3). In response to the EAG report and to help resolve some of the uncertainties raised, an alternative model structure has been explored for all three analyses (RESPONSE, RESPONSE-2, MAJIC-PV) and presented whereby health states are based on a measure of disease progression, as suggested by the EAG. The alternative model structure uses PFS with OS modelled as a surrogate for progression/transformation (to capture the prognostic value of progression on survival). It should be noted that compared with solid tumours, where progression is defined as an increase in tumour size, progression in PV is defined as transformation into acute myeloid leukaemia/myelodysplastic syndrome (AML/MDS), MF or death in the MAJIC-PV trial.[1] Further details on the methods and key assumptions are provided in Appendix 1 of this

		response. It should be recognised that it is not possible to construct a model based on EFS (that would include major thrombo-embolic and haemorrhagic events in addition to MF and AML/MDS) due to a lack of information in the MAJIC-PV manuscript on the number of each type of EFS events and the absence of IPD. In brief, the alternative model structure uses an STM approach (to capture the prognostic value of progression on survival) and is composed of four key health states ("progression-free on ruxolitinib", "progression-free on BAT", "progressed disease" and "death"). Analyses are presented for all three populations (RESPONSE, RESPONSE-2 and MAJIC-PV) to cover the entire PV population eligible for ruxolitinib (see response to Issue 1). Patients on ruxolitinib enter the model progression-free and remain on treatment until discontinuation moving to the BAT progression-free health-state in the absence of progression/transformation. Following transformation into MF or AML/MDS, patients enter the progressed disease health state and remain in this health state until death. Patients move directly to the death health state when progression is attributable to death. Patients entering the model on BAT remain in this health state until transformation/progression and can either move to the progressed disease health state (in case of transformation) or death health state directly. In line with the EAG preferred assumption, no partition of the BAT health state is assumed. PFS is modelled under a competing-risk framework (with each competing event modelled separately) to capture the effect of age on the incidence of the key events (MF, AML/MDS) and death. Patients on ruxolitinib are assumed to discontinue treatment following transformation in line with the trials and clinical practice (unless retreatment for MF), with the progression-free health state split (using health state occupancy – area under the curve [AUC]) into time "On ruxolitinib" and time "Off ruxolitinib" (e.g., on BAT). Transitions from the progression-free health state used to derive PFS (progression-free to pre-progression death, MF and AML/MDS) are derived from the 5-year risk of events reported in the respective trials when possible or using a set of assumptions when data were inconsistent or not available (Appendix 1), supplemented by data from the literature to inform how the risk of events may vary over time.[5, 8, 11] Transitions from the progressed disease (transformation into MF [low/intermediate-1 risk vs. intermediate-2/high risk MF] or AML/MDS) to death are informed by external data (Appendix 1).[9, 12, 13] Model predictions are shown below in Figure 3–Figure 5 and show that despite assumptions and use of external data, model predictions for PFS and OS are plausible and are generally aligned with observed data from the trials at 5 years (with the exception of the RESPONSE-2 analysis in which PFS and OS were underpredicted compared with the

observed data due to the conservative approach used in the model where patients were allowed to transform into AML/MDS within the first 5 years, despite no AML/MDS observed in RESPONSE-2 for patients treated with ruxolitinib).[5]

ICERs, QALYs and costs using the company base-case and alternative structure are presented below in Table 2. Results are presented using the PAS discount agreed for MF ([redacted]). For completeness, results from sensitivity (deterministic and probabilistic sensitivity analysis) and scenario analysis are presented in Appendix 1.

Table 2: ICERs from CS base-case and alternative model structure based on a measure of progression (deterministic results) – PAS discount agreed for MF

	Ruxolitinib			BAT			Inc cost	Inc LY	Inc QALY	ICER
	Cost	LY	QALYs	Cost	LY	QALYs				
RESPONSE										
CS base-case	[redacted]	11.45	[redacted]	[redacted]	9.28	[redacted]	[redacted]	2.17	[redacted]	[redacted]
Alternative structure	[redacted]	13.28	[redacted]	[redacted]	10.68	[redacted]	[redacted]	2.60	[redacted]	[redacted]
RESPONSE-2										
CS base-case	[redacted]	12.25	[redacted]	[redacted]	10.46	[redacted]	[redacted]	1.79	[redacted]	[redacted]
Alternative structure	[redacted]	13.17	[redacted]	[redacted]	11.83	[redacted]	[redacted]	1.34	[redacted]	[redacted]
MAJIC-PV										
CS base-case	[redacted]	9.65	[redacted]	[redacted]	8.02	[redacted]	[redacted]	1.63	[redacted]	[redacted]
Alternative structure	[redacted]	10.08	[redacted]	[redacted]	8.18	[redacted]	[redacted]	1.90	[redacted]	[redacted]

Abbreviations: BAT: best available therapy; CS: company submission; ICER: incremental cost-effectiveness ratio; Inc: incremental; LY: life years; MF: myelofibrosis; PAS: patient access scheme; QALYs: quality-adjusted life years.

		Figure 3: Comparison of model predictions and KM using alternative model structure for RESPONSE  **Abbreviations**: BAT: best available therapy; KM: Kaplan–Meier; OS: overall survival; PFS: progression-free survival; Rux: ruxolitinib.

Figure 4: Comparison of model predictions and KM using alternative model structure for RESPONSE-2

Abbreviations: BAT: best available therapy; KM: Kaplan–Meier; OS: overall survival; PFS: progression-free survival; Rux: ruxolitinib.

		Figure 5: Comparison of model predictions and KM using alternative model structure for MAJIC-PV  **Abbreviations**: BAT: best available therapy; KM: Kaplan–Meier; OS: overall survival; PFS: progression-free survival; Rux: ruxolitinib. In summary, this alternative model structure provides evidence supporting the estimates in the original CS. Furthermore, it suggests that the results in the original CS are potentially conservative. Despite the limitations associated with both analyses and recognising that different model structures will provide non-identical estimates, this alternative model structure provides reassurance and indicates that the original model structure and assumptions used in the company submission are suitable for decision-making and generating plausible ICERs. Finally, it should be recognised that several assumptions and adjustments were necessary for the development of the alternative model structure based on progression and the uncertainty associated with these should be acknowledged. In conclusion, the company believes that the CS original model structure and assumptions (around the treatment effect for OS) are the most appropriate for decision-making as this relies on less assumptions and therefore is less uncertain. There is therefore no change to the CS base-case.

Issue 6: Extrapolation of time to ruxolitinib discontinuation	*Yes*	For the analyses based on the RESPONSE-trials, the company used data from each trial for TTD due to reasons other than death. This approach was considered favourable over pooling the data from both trials due to the visual difference in the KM (CS, Figure 37). A total of [redacted] and [redacted] discontinuations due to reasons other than death were reported in the RESPONSE and RESPONSE-2 trials respectively.[4, 5] In contrast to this, the EAG noted that in the CS pre- and post-discontinuation survival data were pooled. This was because of the small number of events for discontinuations due to death ([redacted] in RESPONSE and [redacted] in RESPONSE-2) and death following discontinuation ([redacted] in RESPONSE and [redacted] in RESPONSE-2). The EAG suggested that the company present an additional scenario using the pooled data from RESPONSE and RESPONSE-2 for TTD due to reasons other than death. Additionally, the EAG stated a preference for the Weibull distribution in place of the odds spline model selected by the company. Novartis does not believe that pooling the data for TTD due to reasons other than death is appropriate, given the differences between the two populations in their rate of discontinuation. In response to the EAGs concerns, and in order to help resolve this uncertainty, a scenario is presented pooling the data for TTD for reasons other than death, using a Weibull distribution as requested by the EAG. ICERs, QALYs and costs for the company base-case using separate TTD and odds spline model and the pooled analysis using the Weibull distribution are shown below in Table 3. Results are presented using the PAS discount agreed for MF ([redacted]%). There is therefore no change to the CS base-case.

Table 3: ICERs from CS base-case and using pooled TTD (deterministic) – PAS discount agreed for MF

	Ruxolitinib			BAT			Inc cost	Inc LY	Inc QALY	ICER
	Cost	LY	QALYs	Cost	LY	QALYs				
RESPONSE										
CS base-case	[redacted]	11.45	[redacted]	[redacted]	9.28	[redacted]	[redacted]	2.17	[redacted]	[redacted]
Pooled TTD	[redacted]	12.10	[redacted]	[redacted]	10.05	[redacted]	[redacted]	2.05	[redacted]	[redacted]
RESPONSE-2										
CS base-case	[redacted]	12.25	[redacted]	[redacted]	10.46	[redacted]	[redacted]	1.79	[redacted]	[redacted]
Pooled TTD	[redacted]	11.48	[redacted]	[redacted]	9.52	[redacted]	[redacted]	1.96	[redacted]	[redacted]

Abbreviations: BAT: best available therapy; CS: company submission; ICER: incremental cost-effectiveness ratio; Inc: incremental; LY: life year; MF: myelofibrosis; PAS: patient access scheme; QALYs: quality-adjusted life years; TTD: time to treatment discontinuation.

Issue 7: Source for utility estimates: MF-8D or EQ-5D	No	The company's base-case uses the utility values estimated from the Myelofibrosis- 8 Dimensions (MF-8D). This is justified in the CS due to the inappropriateness of the EuroQol-5 Dimensions (EQ-5D), as demonstrated using psychometric analyses submitted as part of the CS and clinical advice that symptoms in PV and MF are similar: the EQ-5D was also not considered appropriate in MF in TA386 and TA756.[9, 10] While the EAG acknowledges that there is evidence in favour of the MF-8D, the EAG preferred base-case uses the EQ-5D. This is justified in the EAG report by the MF-8D being developed in MF rather than PV, assumptions which are required to map questions from the Myeloproliferative Neoplasm Symptom Assessment Form (MPN-SAF) to the Myelofibrosis Symptom Assessment Form (MF-SAF), and to ensure consistency between appraisals. To help resolve this uncertainty, the EAG suggests the company: • (1) Provide direct evidence that the EQ-5D is not appropriate; • (2) Provide comparative evidence of the psychometric properties of the EQ-5D and MF-8D;

		• (3) Obtain further clinical opinion on the validity of substituting questions from the MPN-SAF to the MF-SAF. While Novartis acknowledges some of the uncertainties raised by the EAG, Novartis firmly believes that the MF-8D should be used to ensure consistency with appraisals in similar conditions. In both TA386 and TA756, the appraisal committee and EAGs recognised limitations of the EQ-5D in MF (a condition with symptoms resembling those for PV notably itching), with decision-making ultimately based on the MF-8D.[9, 10] As part of the CS, Novartis submitted direct empirical evidence on the lack of appropriateness of the EQ-5D in terms of construct validity and responsiveness using psychometric analyses as recommended in the NICE method guide.[14] The Guide states *"In some circumstances the EQ-5D may not be the most appropriate measure. To make a case that the EQ-5D is inappropriate, provide qualitative empirical evidence on the lack of content validity for the EQ-5D, showing that key dimensions of health are missing. This should be supported by evidence that shows that EQ-5D performs poorly on tests of construct validity (that is, it does not perform as would be expected) and responsiveness in a particular patient population."* To determine whether the EQ-5D is appropriate in PV, psychometric analyses were presented in the CS using RESPONSE-2 where both the EQ-5D and MPN-SAF were collected (the EQ-5D was not collected in RESPONSE). As recommended in the NICE method guide,[14] the appropriateness of the EQ-5D was examined in terms of psychometric criteria of convergent validity, ceiling thresholds and responsiveness relative to the MPN-SAF. The specific tests examined whether the EQ-5D was related to PV-specific symptoms (convergent validity) and reflected changes. As described in the CS (Section B.3.4.1): • **The EQ-5D was associated with high ceiling effect compared with the MPN-SAF indicating that the EQ-5D does not reflect the symptom burden in PV.** A large proportion ([redacted]%) of patients reported no problems in all 5 EQ-5D dimensions at baseline. The MPN-SAF total score did not show a comparable ceiling effect [redacted]%).[15] • **The EQ-5D lacks construct validity.** The EQ-5D construct validity as measured by convergence is inconsistent across MPN-SAF domains at baseline. • **Lack of responsiveness of the EQ-5D.** The standardised response mean (SRM) for the MPN-SAF total score was medium to large (>|0.5|) over time indicating that participants had large to medium improvement in PV key symptoms. In contrast, the SRM remained small to very small for the EQ-5D indicating that the EQ-5D lack responsiveness in PV.

Novartis therefore firmly believe that direct empirical evidence has already been presented to support the lack of appropriateness of the EQ-5D. It should be noted that the same type and level of evidence was presented as the basis for decision-making by the NICE committee in TA386 for the assessment of ruxolitinib in MF, and subsequently accepted in TA756 for fedratinib.[9, 10]

In addition to direct empirical evidence, further clinical opinion was obtained on whether the EQ-5D was likely to be appropriate in PV. Clinical experts noted evidence in MF about the lack of appropriateness of the EQ-5D and considered that symptoms were broadly similar between MF and PV. Clinical experts therefore did not believe the EQ-5D to be appropriate and noted that the detrimental effect on quality of life of itching, a key symptom in PV and MF, is unlikely to be captured by the EQ-5D.

The MPN-SAF is a condition specific questionnaire and therefore reflects symptoms in PV. However, to be used in a cost-effectiveness analysis, a preference-based measure is required which is not available in PV. In contrast, the MF-8D (preference-based measure) was developed for MF based on questions from the MF-SAF and European Organisation for Research and Treatment of Cancer Core Quality of Life Questionnaire (EORTC QLC-30), a condition exhibiting similar symptoms to PV. Novartis therefore believe that utility estimates derived from the MF-8D are more appropriate than those derived from the EQ-5D which was shown to lack construct validity and responsiveness in both PV and MF.

As acknowledged in the CS, and raised by the EAG, assumptions are however required for 2 of the 8 dimensions used to derive the MF-8D due to difference in wordings between the MF-SAF (specific to MF) and the MPN-SAF (broader to MPN). As shown below in Table 4 substituted dimensions were very similar between the MPN-SAF and MF-SAF.

Table 4: Comparison in dimensions with non-identical wording between MF-SAF and MPN-SAF

MF-SAF	MPN-SAF
Pain under ribs on the left side	Abdominal pain
Bone or muscle pain	Bone pain

Abbreviations: MF-SAF: Myelofibrosis Symptom Assessment Form; MPN-SAF: Myeloproliferative Neoplasm Symptom Assessment Form

In response to technical engagement, further clinical validation was obtained confirming that both questionnaires were similar. Clinical experts further noted that while assumptions are required, this is preferable to using the EQ-5D.

		The EAG further suggests that comparative evidence for the psychometric performance of the MF-8D and EQ-5D utilities for a population with PV could help resolve this issue. To be able to do this, a dataset is required containing both the EQ-5D and the MPN-SAF and EORTC QLC-30 (to derive the MF-8D). Such dataset does not exist unfortunately. Perhaps more importantly, the MF-8D would be derived from the MPN-SAF, therefore evaluating the psychometric properties of the MF-8D against the MPN-SAF would be misleading and counter-intuitive as one is estimated from the other. As the MF-8D would be derived from the MPN-SAF, the MF-8D is expected to have good psychometric properties (as this is assessed against the MPN-SAF). In summary, in line with the NICE method guide[14], direct empirical evidence was provided in the CS to support the lack of appropriateness of the EQ-5D and use of a condition specific measure (these findings were in line with those in MF that supported the use of a condition preference-based measure). While Novartis acknowledges the limitations raised by the EAG, clinical experts strongly believe that the MF-8D is more appropriate than the EQ-5D. Novartis further believe that using the MF-8D would be consistent with previous NICE appraisals other MPNs conditions, such as MF. There is therefore no change to the CS base-case as the MF-8D is likely to be more appropriate than the EQ-5D.

B.4 Additional issues

Novartis have no further comments or additional issues, aside from Issue 1–7 discussed above.

B.5 Summary of changes to the company's cost-effectiveness estimate(s)

As discussed in Issue 1–7 above, no changes are made to the company's base-case cost-effectiveness estimates following technical engagement. While Novartis provided additional analyses in response to technical engagement, these analyses were provided for transparency and provide reassurance to the EAG and NICE committee only.

References

1. Harrison C NJ, Boucher R, et al. Ruxolitinib versus Best Available Therapy for PV intolerant or resistant to Hydroxycarbamide in a Randomized trial (Manuscript in Development), 2022.
2. Passamonti F, Griesshammer M, Palandri F, et al. Ruxolitinib for the treatment of inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): a randomised, open-label, phase 3b study. The Lancet Oncology 2017;18:88-99.
3. Vannucchi AM, Kiladjian JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera. N Engl J Med 2015;372:426-35.
4. Kiladjian J-J, Zachee P, Hino M, et al. Long-term efficacy and safety of ruxolitinib versus best available therapy in polycythaemia vera (RESPONSE): 5-year follow up of a phase 3 study. The Lancet Haematology 2020;7:e226-e237.
5. Passamonti F, Palandri F, Saydam G, et al. Ruxolitinib versus best available therapy in inadequately controlled polycythaemia vera without splenomegaly (RESPONSE-2): 5-year follow up of a randomised, phase 3b study. The Lancet Haematology 2022;9:E480-E492.
6. National Institute for Health and Care Excellence (NICE). NICE DSU Technical Support Document 19: Partitioned survival analysis as a decision modelling tool. Available at: https://www.sheffield.ac.uk/nice-dsu/tsds/full-list (accessed 12 January 2023).
7. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. Available at: https://www.sheffield.ac.uk/nice-dsu/tsds/full-list (accessed 25 October 2022), 2011.
8. Kiladjian J-J, Zachee P, Hino M, et al. Long-term efficacy and safety of ruxolitinib versus best available therapy in polycythaemia vera (RESPONSE): 5-year follow up of a phase 3 study. The Lancet Haematology 2020;7:e226-e237.
9. National Institute for Health and Care Excellence (NICE). Ruxolitinib for treating disease-related splenomegaly or symptoms in adults with myelofibrosis [TA386]. Available at: https://www.nice.org.uk/guidance/ta386 (accessed 28th June 2022).
10. NICE. TA756: Fedratinib for treating disease-related splenomegaly or symptoms in myelofibrosis. Available at: https://www.nice.org.uk/guidance/ta756 [Last Accessed 05 August 2022].
11. Szuber N, Mudireddy M, Nicolosi M, et al. 3023 Mayo Clinic Patients With Myeloproliferative Neoplasms: Risk-Stratified Comparison of Survival and Outcomes Data Among Disease Subgroups. Mayo Clin Proc 2019;94:599-610.
12. Tang G, Lopez JEH, Wang SA, et al. Characteristics and clinical significance of cytogenetic abnormalities in polycythemia vera. Haematologica 2017;102:1511.
13. Tefferi A, Lasho TL, Jimma T, et al. One thousand patients with primary myelofibrosis: the mayo clinic experience. Mayo Clin Proc 2012;87:25-33.
14. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual (PMG36). January 2022. Available at: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation (accessed 25 October 2022).
15. Novartis. Data on File. Final Clinical Study Report (Week 260): Randomized, open-label, multi-center phase IIIb study evaluating the efficacy and safety of ruxolitinib versus best available therapy in patients with polycythemia vera who are hydroxyurea resistant or intolerant (RESPONSE-2).

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

Patient expert statement and technical engagement response form

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources. The external assessment report (EAR) and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking you about living with polycythaemia vera or caring for a patient with polycythaemia vera. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

A patient perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

Patient expert statement

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise. We have given guidance on the issues in which we expect this to be the case and advice on what you could consider when giving your response.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Help with completing this form

If you have any questions or need help with completing this form please email the public involvement (PIP) team at pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Your response should not be longer than 15 pages.

Please note, **part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

Patient expert statement

NICE National Institute for Health and Care Excellence

The deadline for your response is **5pm** on **20 January 2023.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Living with this condition or caring for a patient with polycythaemia vera

Table 1 About you, polycythaemia vera, current treatments and equality

1. Your name	
2. Are you (please tick all that apply)	☒ A patient with polycythaemia vera? ☐ A patient with experience of the treatment being evaluated? ☐ A carer of a patient with polycythaemia vera? ☐ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	MPD Voice
4. Has your nominating organisation provided a submission? (please tick all options that apply)	☐ No (please review all the questions and provide answers when possible) ☒ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement ☐ Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement ☐ I agree with it and **will be** completing
5. How did you gather the information included in your statement? (please tick all that apply)	☐ I am drawing from personal experience ☐ I have other relevant knowledge or experience (for example, I am drawing on others' experiences). Please specify what other experience: ☐ I have completed part 2 of the statement **after attending** the expert engagement teleconference

	☐ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference ☐ I have not completed part 2 of the statement
6. What is your experience of living with polycythaemia vera ? **If you are a carer (for someone with polycythaemia vera) please share your experience of caring for them**	I was originally diagnosed with PV in 2004 following a chest infection and a consequent blood test. I was referred to Guys and St Thomas Hospital Haematology team. Initially my condition was controlled with venesections, daily aspirin and regular monitoring consultations. Over this time I developed symptoms such as itching skin on my legs every time I came into contact with water for which I was prescribed drugs that failed to work. I also developed Rosacea on my face, an enlargement of the spleen and increasing fatigue going about my daily business at work and at home. By 2010 my PV had advanced and following the deterioration in my condition I was asked to consider joining a new trial of Vorinostat. At the start the itching skin improved and my spleen reduced considerably in size. However my fatigue became worse and after a short time my hair thinned and started to fall out, I lost weight and eventually began to dread every dose. The lead consultant and I agreed that I discontinue the trial. I was then prescribed Hydroxycarbomide (2011 I think) which failed to have the desired impact. My bloods again became difficult to control, I still had itching, fibre in my marrow, rosacea and my spleen enlarged again. I had regular venesections to control blood counts, but worst of all my fatigue increased to the point that it was severely impacting on my work. On top of the experience with Vorinostat this was become more and more difficult to manage. In 2012 I discussed the MAJIC trial with my consultants at Guys and St Thomas's. I was given the opportunity to join the trial and I was randomised to use Ruxolitinib. I was therefore 001. The first person on the trial to gain access to this drug for PV. The doctors were wonderful, supportive and informative. The improvement in my condition was immediate and transformational, dramatically so I would say! All previous symptoms subsided, no more itching, rosacea, my spleen went back to practically normal and I got my life back.

	PV and the associated previous therapies were negatively impacting on my career (I was a Vice Principal of a Further Education College) and personal life to the degree that I was failing after a long successful career and I was looking for ways to stop working. My family were constantly worried that my deterioration would lead to premature infirmity etc. Following the end of the trial the consultant at Guys and St Thomas's Hospital managed to negotiate a compassionate, life time supply of Ruxolitinib with Novartis and I have continued to take 45 mg per day since August 2012 (12 years now) and hope to continue to do so. The only concern I have had has been with the development of a Basal Cell Carconoma on my nose and other worries with regard to another skin patch on my forehead which is being looked into currently. Otherwise I have been living a normal, lively and fulfilling life since I started taking Ruxolitnib in 2012. I am now retired (at the age of 60) and spend my time restoring an old run down house, traveling, walking, eating and drinking and enjoying the company of my family and friends.
7a. What do you think of the current treatments and care available for polycythaemia vera on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	I think that all of the doctors and nurses that I have encountered over the last 20 years of living with Polycythemia Vera have been nothing short of amazing. They have been committed to providing the absolute best care possible, well organised, knowledgeable, excellent communicators, diligent in the extreme, excellent team contributors, and fighters for the best of professional and clinical standards. I simply cannot praise them enough. With regard to treatments my experience has been that they have explored every avenue available to them to ensure the best of care and solutions for me and for others that I have encountered. They have explored every avenue for me and have found the therapy and drug that works for me and have strived to find ways for me to maintain access to it. I have spoken to other sufferers of the same condition , mainly through events organised at Guys and St. Thomas's, and when talking to them about the success of Ruxolitinib for me have encountered people who's condition was worse than mine who had not had the 'good fortune' to gain access to the drug for whatever reason,

	who were saddened if not (in one case I can remember) envious of me for being able to have this drug. I believe that this drug is a game changer for people with PV. At least it has been just that for me.
8. If there are disadvantages for patients of current NHS treatments for polycythaemia vera (for example, how they are given or taken, side effects of treatment, and any others) please describe these	I don't see any myself. I take one pill in the morning and two in the evening before bed. They sit on my bedside table. I also take 75mg of aspiring daily. I have a small worry about skin cancer but, in my group of friends I know several who have similar complaints that don't have PV or anything similar. I have to regularly visit Haematology clinics, but that actually has been a pleasure and always reassuring.
9a. If there are advantages of ruxolitinib over current treatments on the NHS please describe these. For example, the effect on your quality of life, your ability to continue work, education, self-care, and care for others? **9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why?** **9c. Does ruxolitinib help to overcome or address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these**	9a I think that I have answered this in question 6 above. Ruxolinib has simply given me my life back (sounds dramatic but it has!) I live a full life now, where I didn't before, I am retired now, though I am very active physically and mentally, I have learned French recently and have developed loads of practical skills in recent years, including building etc. I have a growing family and I have a very active family life, along with my sisters I care for my aging mother who has Parkinsons Disease and so on and so forth. 9b. The most important of these is that I can engage fully in my family. I have recently become a grandfather which I love. I have a great time enjoying retirement with my lovely wife and aim to continue to do so for some years to come. 9c. All positive really.
10. If there are disadvantages of ruxolitinib over current treatments on the NHS please describe these. For example, are there any risks with ruxolitinib? If you are concerned about any potential side effects you have heard about, please describe them and explain why	I have talked about skin cancers earlier. I have not experienced any other side effects. I did have a period of feeling fatigued a couple of years ago. My consultant increased my dose by 5mg per day and that seemed to improve. I feel well now.
11. Are there any groups of patients who might benefit more from ruxolitinib or any who may benefit less? If so, please describe them and explain why	If there are people who suffer from PCV who have had similar responses to other treatments as I have done I feel it is a great shame if Ruxolitinib is not available to them to try from a range of alternative therapies. It has worked for me where others didn't. I don't know how it would react with other conditions or other drug regimes

Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	that people are using for other conditions but it has been very suitable for me. It has provided a route back to normal health for me.
12. Are there any potential equality issues that should be taken into account when considering polycythaemia vera and ruxolitinib? Please explain if you think any groups of people with this condition are particularly disadvantaged Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here.	Not that I can imagine at the moment. It has been positive to me and I have never come across any situations where I have witnessed any negative impact on disadvantaged groups in this respect.
13. Are there any other issues that you would like the committee to consider?	No thank you. I think I have related my experience here.

Part 2: Technical engagement questions for patient experts

Issues arising from technical engagement

The issues raised in the EAR are listed in table 2. We welcome your comments on the issues, but you do not have to provide a response to every issue, such as the ones that are technical, that is, cost effectiveness-related issues. We have added a comment to the issues where we consider a patient perspective would be most relevant and valuable. If you think an issue that is important to patients has been missed in the EAR, please let us know in the space provided at the end of this section.

For information: the patient organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR, the patient organisation responses will also be considered by the committee.

Table 2 Issues arising from technical engagement

Relevance of the trial populations for modelling UK practice	
Modelling the relative treatment effect for overall survival	
Waning of the treatment effect	
Modelling approach: state-transition or partitioned-survival	
Model structure: health states and events	

Patient expert statement

Are there any important issues that have been missed in EAR?	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Ruxolitinib worked for me where all other drugs and therapies thought to be of use to me did not have the desired impact
- I have had minimal side effects to using Ruxolitinib for 12 years now apart from some concern/doubts about a link to a basal cell carcinoma several years ago.
- The health professionals working in this field are second to none and are totally committed to achieving the best possible solutions for sufferers of this and other related diseases and conditions.
- Click or tap here to enter text.
- Click or tap here to enter text.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☒ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see NICE's privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

Clinical expert statement and technical engagement response form

Thank you for agreeing to comment on the external assessment report (EAR) for this evaluation, and for providing your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature. The EAR and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR. You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **20 January 2023.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

Part 1: Treating polycythaemia vera and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Claire Harrison
2. Name of organisation	Guys and St Thomas' NHS Foundation Trust
3. Job title or position	Consultant Haematologist
4. Are you (please tick all that apply)	☐ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with polycythaemia vera? ☒ A specialist in the clinical evidence base for polycythaemia vera or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None
8. What is the main aim of treatment for polycythaemia vera? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	Reduce the risk of thrombotic and haemorrhagic events by controlling the haematocrit, WBC and platelet count Improve quality of life for PV patients Reduce splenomegaly

	Reduce risk of disease transformation
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	Normalised WBC Platelets and controlled HCT Reduction (50%) of frequency of venesection Normal or reduced spleen size Reduction in symptom score using validated tool such as MPN 10 or MPN SAF
10. In your view, is there an unmet need for patients and healthcare professionals in polycythaemia vera?	Availability of treatment options for patients not responding to or intolerant of front line therapies. Lack of clarity of impact of therapies on long term events Burden of symptoms not controlled with standard therapies
11. How is polycythaemia vera currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	As described in the current submissions Risk assessment Venesection and aspirin Cytoreductive therapy (either HC or IFN) for high risk patients Symptom control For patients who have unresponsive symptoms Nand who fail or are intolerant of HC and then IFN (if able to use this not all patients are to eg depression autoimmune disease) the technology would have a significant impact upon care
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic)	Ruxolitinib is an oral tablet therapy given as an out patient treatment in much the same way as HC. Specialists in hospitals prescribe and monitor both therapies. Ruxolitinib is already used in patients with myelofibrosis (a related blood condition)

• What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	I expect ruxolitinib to reduce requirements for health care support both for venesctions and symptom management. I expect this drug will require less monitoring and patients will have more stable disease but most importantly those patients with symptoms (commonly pruritus or itching, which is often intractable and life changing; or fatigue which frequently impacts the ability of patients to work and engage in normal family/social life) This is totally in line with the carefully curated evidence from patient advocacy and is well demonstrated in the clinical trials.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	No subgroup of patients is more likely to benefit but in particular patients with significant pruritus and fatigue benefit Those patients who have no additional therapy options (options other than HC or IFN are leukaemogenic)
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	No different from current care
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	No testing beyond standard of care is required
17. Do you consider that the use of the technology will result in any substantial health-related benefits that	Instruments that commonly are used to measure quality of life such as EQ5D or EORTIC QLQC30 do not reflect the difficult symptoms that PV patients

are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	experience in particular fatigue and itch. These symptoms are best measured with the validated symptom tool MPN10/MFSAF/MPNSAF
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	This technology represents a step change for patients in terms of quality of life and control of diease for those patients who have failed HC and have 1. Uncontrolled symptoms 2. Uncontrolled myeloprolfieration I wish to emphasize that ALL randomised clinical trials conducted with this agent have demonstrated this
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Side effects of this agent which require review are almost identical to HC – lowered blood counts and risk of skin cancer. An additional side effect is weight gain.
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Patients in the UK participated in all the clinical trials mentioned in the submission and in particular the MAJIC PV study was solely a UK study

21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
22. How do data on real-world experience compare with the trial data?	RWE reflects trial data
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why.	No inequalities but I wish to point out that if ruxolitinib is NOT approved there will be patients with intractable symptoms who will be profoundly disadvantaged. Furthermore patients who have failed both HC and IFN only have other treatment options such as busulfan which substantially increase the risk of leukaemia which is fatal within 3-6 months. Finally since ruxolitinib is approved in Scotland and as far as I am aware all other countries in the EU such as France, Germany, Italy, Spain if this drug is NOT approved English patients will be substantially disadvantaged.

More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Relevance of the trial populations for modelling UK practice	I agree with the view that the MAJIC PV patient population more closely resembles the appropriate population for review. The RESPONSE and RESPONSE 2 trials entry criteria were very specific eg number of venesections, requirement for splenomegaly
Modelling the relative treatment effect for overall survival	I agree with the modelling of benefit for overall survival on the MAJIC PV data this is the study with the UK population with no cross over
Waning of the treatment effect	It is difficult to judge waning of treatment effect when studies (with the exception of MAJIC PV) employed cross over. In my experience patient do very well on this agent and have a very well sustained benefit this is illustrated in the MAJIC PV study in terms of symptoms and control of blood count as well as time to treatment discontinuation (ie consistent). I note the comments of the EAG and discussion of longer waning.

	I would wish to make 3 additional comments. 1. We have not in over 10 years of using ruxolitinib been able to identy specific evidence of resistance nor biological mechanisms of resistance and in my experience in PV this is very rare 2. The mean age at diagnosis for patients with PV is approximately 67years this may impact upon considerations of waning of effect 3. Please remember the impact of symptoms on quality of life and lack of data for this
Modelling approach: state-transition or partitioned-survival	I note the EAR comments as follows "The incidence of the key events while patients were on BAT was estimated from relative rates (IRRs) from pooled trial data. This resulted in lower incidence of MF, TE and haemorrhage, and higher incidence of non-melanoma skin cancer while patients were on ruxolitinib than on BAT. There was very little difference between the treatments in estimated rates of conversion to AML/MDS." These are the findings I would expect: TE and haemorrhage relate to control of blood parameters I would thus expect these to be reduced by ruxolitinib. MF is characterised by splenomegaly (reduced by ruxolitinib) and risk of MF relates to factors such as leucocytosis Skin cancer does occur in patients treated with ruxolitinib this is a consistent finding possibly related in addition to prior exposure to HC. Concerning rates of AML/MDS these are rare events occurring after many years and hard to model Modelling does not include the major impact of symptom control.

Model structure: health states and events	Modelling a disease based on clinical events is in my opinion most clinically relevant. Patients movement between treatments unless triggered by intolerance is however often related to a change in disease status. Finally I would comment that the strongest benefit of ruxolitinib – symptom, quality of life, control of disease is not mentioned in this discussion
Extrapolation of time to ruxolitinib discontinuation	I would suggest use of the MAJIC PV data for this purpose.
Are there any important issues that have been missed in EAR?	Symptom benefit for patients is understated and underestimated. Patients with PV can have a very heavy symptom burden which has a massive impact on their quality of life and are intractable. Comments related to symptom scoring have in my opinion been understated. I would urge the EAR to closely examine MF8D and MPNSAF the questions are identical (though fewer in MF8D) the tools have been extensively validated in the entire MPN (PV, ET and MF) population and have been mapped to EQ5D. EQ5D has not been validated in PV patients.

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

PV is a chronic disease associated with major impact upon patient quality of life, risks of thrombosis, and disease progression, these facets of the disease are more pronounced for patients resistant or intolerant to HC.
Second or third line treatments are not as effective as ruxolitinib in controlling blood counts but also very importantly quality of life relating to difficult symptoms a fact that has consistently been shown in randomised clinical trials.
Ruxolitinib is a very well tolerated oral therapy already available for patients with myelofibrosis (a related condition) in England.
Beyond HC or IFN other agents available for PV patients substantially increase the risk of development of leukaemia a devastating incurable complication – for example those patients moving from HC to busulfan have a 20-30% risk of leukaemia this is not sufficiently mentioned in my opinion.
Those patients with intractable pruritus or fatigue deserve specific mention these disease related symptoms are not responsive to standard therapies and significantly impact upon patients quality of life.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

Clinical expert statement and technical engagement response form

Thank you for agreeing to comment on the external assessment report (EAR) for this evaluation, and for providing your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature. The EAR and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR. You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **20 January 2023.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

Part 1: Treating polycythaemia vera and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Tim Somervaille
2. Name of organisation	Cancer Research UK Manchester Institute, The University of Manchester & The Christie NHS Foundation Trust
3. Job title or position	Professor of Haematological Oncology
4. Are you (please tick all that apply)	☐ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with polycythaemia vera? ☒ A specialist in the clinical evidence base for polycythaemia vera or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None.
8. What is the main aim of treatment for polycythaemia vera?	Through maintaining a haematocrit of <45% by venesection and/or cytoreductive therapies such as hydroxycarbamide, the goal is to reduce the risk of adverse

(For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	events arising from the presence of disease in particular including thrombosis and to improve patient quality of life through amelioration of symptoms.
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	No thrombotic events in the long term and a good quality of life (e.g. reduced frequency of hospital visits for monitoring of blood counts +/- venesection, improvement of disabling symptoms)
10. In your view, is there an unmet need for patients and healthcare professionals in polycythaemia vera?	Yes. Some patients are intolerant of standard of care cytoreductive therapy hydroxycarbamide and benefit from ruxolitinib which is effective therapy in this disease through its ability to inhibit production of early erythroid cells by inhibiting the JAK2 tyrosine kinase; and to improve symptoms.
11. How is polycythaemia vera currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	PV is managed in the UK as per the BCSH guidelines found here: https://onlinelibrary.wiley.com/doi/full/10.1111/bjh.15648 and with a US and EU view here: https://onlinelibrary.wiley.com/doi/full/10.1002/ajh.26008 The pathway is in general well defined however there is a clear unmet need for patients who are intolerant of hydroxycarbamide and ruxolitinib fills this gap.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic)	No, ruxolitinib would represent an improvement in care for a subset of polycythaemia patients who are hydroxycarbamide intolerant. Ruxolitinib is however much more expensive than hydroxycarbamide. Ruxolitinib should be prescribed in a specialist clinic by an expert experienced in chemotherapy. No additional investment is needed to introduce the technology because it is already in widespread use in a related myeloproliferative neoplasm called myelofibrosis.

• What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	Yes – ruxolitinib can be immensely beneficial for a subgroup of hydroxycarbamide intolerant patients because it not only suppresses production of early erythroid progenitor cells thereby assisting in maintaining the haematocrit at <45% and reducing the requirement for intermittent hospital-based venesections, it also often dramatically improves symptoms such as sweats, itching and fatigue which in patients with PV contribute to conferring a significantly inferior quality of life. The jury is out as to whether this will prolong life given the long life expectancy of patients with polycythaemia vera – follow up clinical trials are ongoing and in time they may address this question. However I am in no doubt that for a well selected group of patients ruxolitinib will significantly increase quality of life through improving really deleterious symptoms. For example, a small daily dose of ruxolitinib can be sufficient to completely abolish the severe aquagenic pruritus that some PV patients experience which can be so bad that patients are unable to shower or bathe in view of the skin pain it induces. Patients with PV may also in addition experience headaches, sweats, spleen pain (from a disease-related enlarged spleen) and fatigue and all of these symptoms are typically improved with ruxolitinib. In my personal experience, a number of my PV patients who have received compassionate access ruxolitinib have described the treatment as positively life changing.

14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	This is not being considered for the general population. The proposal is to have this available for hydroxycarbamide-intolerant or resistant PV patients.
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	It will make things easier for both patients and professionals. Both will benefit from having an effective therapy which improves symptoms and QoL, controls the haematocrit, and reduces the number of clinic appointments and venesections required.
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	PV patients should be able to access ruxolitinib if they are intolerant or resistant to hydroxycarbamide. If there is no therapeutic benefit in terms of symptom or disease control after six months, consideration should be given to finding an alternative therapy (e.g. busulphan, interferon).
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	I worry that the appraisal may have under-estimated the QoL improvement that some PV patients derive from ruxolitinib. Symptom scoring systems have been developed which are effective and particular to myeloproliferative disorders because they include specific annotation of symptoms more frequently seen in patients with these conditions (e.g. MPN-SAF). These seem more likely to be an accurate reflection of benefit than more generic QoL scales.
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes, as already mentioned for some patients with HC-intolerant or resistant PV, ruxolitinib can be life changing and so for those it does represent a step change improvement. It improves the way current need is met by offering a therapeutic

• Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	approach that actually works in delivering enhanced QoL, rather than the currently available options.
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	In general ruxolitinib is very well tolerated and in my experience most patients with a myeloproliferative disorder feel better on it. From my viewpoint the important issues to discuss with the patients up front before starting are a longer term increased risk of skin cancers and also a risk of latent virus reactivation (e.g. shingles) & in some patients weight gain.
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	In large part yes, although I agree that the UK-based MAJIC trial is most likely to be reflective of UK practice. The control of HCT in RESPONSE2 is important. Reduction of spleen volume by 35% or more in RESPONSE (primary outcome measure) is much less of an important goal of therapy in the real world. Likewise achieving a CHR (POM in MAJIC) is less important in real world than venesection minimisation and symptom improvement: in reality you want the patients to have fewer venesections and to feel better and to have more stable blood counts. All this can be achieved without necessarily having a CHR or a spleen volume reduction of >35%. RESPONSE2 shows the venesection requirement reduction handily; the other differences in POM in MAJIC and RESPONSE are surrogates and likely underestimate the utility of the technology (as shown in Table 14). The improvements in symptoms noted in RESPONSE and RESPONSE2 (p54 of EAG report) chime with my personal experience of treating this patient population myself, as per the MPN-SAF scores (which I think are much more appropriate than the EQ-5D-5L analysis shown in Table 15).

	I think all significant adverse events with the technology are mentioned in the EAG document. Note that all three studies enrolled patients who were hydroxycarbamide intolerant/resistant whereas, particularly for younger patients, if ruxolitinib were available as an option for therapy I would nevertheless consider a trial of pegylated interferon ahead of ruxolitinib given the longer experience we have with IFN in haematological malignancies, its well established safety and efficacy profile, and the absence of issues pertaining to skin cancer and viral reactivation. That said, there are many patients where IFN is not an option (e.g. prior intolerance, past history of psychiatric disorder and so on) so a trial of IFN should not be mandated ahead of ruxolitinib.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
22. How do data on real-world experience compare with the trial data?	They marry up well.
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could	There is the issue that ruxolitinib is available for this disease indication in Scotland, as well as in a number of EU countries.

• exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Relevance of the trial populations for modelling UK practice	I think all three trials are relevant for the UK population.
Modelling the relative treatment effect for overall survival	No views on this which diverge from those in the EAG report.
Waning of the treatment effect	No views on this which diverge substantially from those in the EAG report. I would note how difficult it is for any of us to be confident that the assumptions made in generating the models are a close representation of future reality.

Modelling approach: state-transition or partitioned-survival	No views on this which diverge substantially from those in the EAG report. I would note how difficult it is for any of us to be confident that the assumptions made in generating the models are a close representation of future reality.
Model structure: health states and events	No views on this which diverge substantially from those in the EAG report. I would note how difficult it is for any of us to be confident that the assumptions made in generating the models are a close representation of future reality.
Extrapolation of time to ruxolitinib discontinuation	No views on this which diverge substantially from those in the EAG report. I would note how difficult it is for any of us to be confident that the assumptions made in generating the models are a close representation of future reality.
Are there any important issues that have been missed in EAR?	In general the document appears thorough, fair and balanced. I would once more highlight my answer to Q17 above: " I worry that the appraisal may have under-estimated the QoL improvement that some PV patients derive from ruxolitinib….."

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

There is an unmet need in the UK for better treatments for patients with polycythaemia vera who are intolerant of or resistant to hydroxycarbamide

The JAK2 inhibitor ruxolitinib is very good therapy for many patients with polycythaemia vera who are intolerant of or resistant to hydroxycarbamide because it is effective in suppressing the exuberant erythropoiesis associated with PV, thereby minimising a requirement for venesection in this population of patients with few other effective options for treatment.

In addition, ruxolitinib can dramatically improve patient QoL where there is a significant symptom burden associated with hydroxycarbamide-intolerant PV; in some patients plagued by symptoms such as aquagenic pruritus, sweats, fatigue, spleen discomfort and so on, it can be positively life-changing.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ruxolitinib for treating polycythaemia vera (review of TA356) [ID5106]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under [redacted], all information submitted under [redacted], and all information submitted under [redacted] in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **20 January 2023**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

About you

Table 1 About you

Your name	[redacted], MPN Voice and [redacted], Leukaemia Care
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	MPN Voice and Leukaemia Care
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

NICE National Institute for Health and Care Excellence

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Relevance of the trial populations for modelling UK practice	Yes/No	no comment
Modelling the relative treatment effect for overall survival	Yes/No	no comment
Waning of the treatment effect	Yes/No	no comment
Modelling approach: state-transition or partitioned-survival	Yes/No	no comment
Model structure: health states and events	Yes/No	no comment
Extrapolation of time to ruxolitinib discontinuation	Yes/No	no comment
Source for utility estimates: MF-8D or EQ-5D	Yes/No	As representatives of the patient communities, we strongly believe that the use of the MF-8D framework is the most accurate and precise way of performing quantitative assessment of quality of life for PV patient populations.

		The QoL measures used in MF-8D truly reflect the lived experience of MPN patients in a way that a generic instrument such as EQ-5D does not. MF-8D does not over-emphasise typical MPN symptoms, but it measures them more precisely and allows a more accurate comparison of treatments to be achieved. MF-8D was originally designed to assess the QoL of Myelofibrosis patients but, in our opinion, is also valid for PV patients. The Landmark studies (cited in our main submission), which used the related MPN-SAF, showed a very similar burden of illness for MF and PV patients (TSS scores of 21.2 and 17.4 respectively), indicating that it is valid to measure the QoL of the two patient groups using the same framework. From the perspective of the patient community, we therefore recommend that MF-8D is the instrument used to assess utility in this assessment.

CONFIDENTIAL UNTIL PUBLISHED

Evidence Assessment Group Report commissioned by the NIHR Evidence Synthesis Programme on behalf of NICE

Ruxolitinib for treating polycythaemia vera (ID5106)

Evidence Review Group's summary and critique of the company's response to technical engagement

Produced by	Southampton Health Technology Assessments Centre (SHTAC)
Authors	Joanne Lord Asyl Liyakat Hawa Geoff Frampton
Correspondence to	Prof Joanne Lord Southampton Health Technology Assessments Centre (SHTAC) Wessex Institute Alpha House Enterprise Road, University of Southampton Science Park Southampton SO16 7NS www.southampton.ac.uk/shtac
Date completed	3 February 2023

[redacted]

[redacted]

Contents

List of tables

1. Introduction

This document is the Evidence Assessment Group's (EAG) summary and critique of the response by the company, Novartis Pharmaceuticals UK Ltd, to the key issues for technical engagement (TE) proposed in the EAG report for this appraisal. The EAG received the company's response on 24th January 2023.

The company's TE response form contains the following information:

- A written response to each of the 7 key issues, 3 of which include new evidence and/or analyses (see Table 1 below).
- A set of updated cost-effectiveness results for additional scenario analyses provided by the company in response to key issues numbers 4, 5 and 6.
- A separate appendix to the company's response form, providing additional detail on the methods and results of the alternative model structure for key issue 5.
- An updated version of the company's economic model accompanies the response form, including alternative model structures for key issues 4 and 5.

In this report we present the following:

- Our critique of the company's response to each of the 7 issues for technical engagement (Section 2)
- A validation of the results of the company's updated cost-effectiveness analysis, and the results of additional EAG scenario analyses (Section 3)

Table 1 Summary of key issues for technical engagement

Issue number	Summary of issue	Does this response contain new evidence, data or analyses?
1	Relevance of the trial populations for modelling UK practice	No
2	Modelling the relative treatment effect for overall survival	No
3	Waning of the treatment effect	No
4	Modelling approach: state-transition or partitioned-survival	Yes
5	Model structure: health states and events	Yes
6	Extrapolation of time to ruxolitinib discontinuation	Yes
7	Source for utility estimates: MF-8D or EQ-5D	No

2. Critique of the company's response to key issues for technical engagement

2.1 Issue 1 – Relevance of the trial populations for modelling UK practice

The company's response to TE does not include any new information on this issue. The relevance of the three clinical trial populations to the patient population in UK practice is discussed in sections 3.2.1, 3.2.2 and 4.2.3 of the EAG report.

The EAG maintain the view that the MAJIC-PV trial is most likely to reflect UK clinical practice, although the data available from the trial are limited. The RESPONSE and RESPONSE-2 trials are also relevant for the model as they included subgroups with and without splenomegaly, respectively. Clinical experts advising the EAG agreed that baseline characteristics for the three trial populations are generally reflective of the licensed population, although there is some continuing uncertainty due to:

- A lack of standardisation of definitions of hydroxycarbamide resistance and intolerance in clinical practice and differences in the criteria used in the trials.
- Differences in inclusion criteria. It is not clear to what extent the MAJIC-PV population represents a high-risk subgroup, although baseline characteristics appear similar to the other trials, the mortality rate was substantially higher.

2.2 Issue 2 – Modelling the relative treatment effect for overall survival

The company do not present any new information on this issue and reiterate their view that time varying estimates of the hazard ratio (HR) for overall survival (OS) are more appropriate and better reflect the MAJIC-PV results than a constant HR across the five-year follow up. Discussion of this issue is provided in section 4.2.6.2.1 of the EAG report.

The EAG maintain a preference for the constant HR estimate, due to high uncertainty over the treatment effect and the post hoc nature of the piecewise analysis. The confidence interval around the constant HR is wide, and those around the company's piecewise HR estimates are even wider. Other sources for estimates of the treatment effects do not help to resolve this uncertainty (see EAG Table 22).

However, we consider that the company's methods used to select the timepoint and estimate the piecewise HRs from the MAJIC-PV KM results are appropriate (CS section B.3.3.4 and Appendix O). We also acknowledge that the assumption of a delayed impact on survival may be clinically plausible, due to lags in onset of PV-related complications and disease progression. For this reason, we reported results using the company's piecewise HR

estimates for overall survival as a scenario analysis with other EAG preferred assumptions (EAG report Tables 33 and 34).

2.3 Issue 3 – Waning of the treatment effect

The company state that their base-case assumption of linear waning of the HR for overall survival from year 5 to HR=1 at year 20 was conservative, based on clinical expert opinion that approximately twice as many patients would be alive at 20 years with ruxolitinib than with current treatment. The company have not changed their base case in response to TE, but they argue that waning should be removed from the EAG preferred analysis because this includes the more conservative constant HR for overall survival (Issue 2 above).

As noted in section 4.2.6.1 of the EAG report (pages 82 to 83), there is uncertainty over whether and how the treatment effect might change after the trial period. Feedback from clinical experts consulted by the EAG was that there is no particular reason to assume a loss of effect with long-term ruxolitinib treatment, based on its mechanism of action. But given the uncertainties around estimation of the treatment effect as discussed above, we decided to retain the company's base case assumption of waning between years 5 and 20 in our preferred analysis, and to report scenario analysis around this assumption (EAG report Tables 33 and 34). For this scenario analysis we increased the assumed waning period to the maximum coded in the company's model of 5 - 50 years (the company have not coded removal of waning as an option). We have not changed this position but suggest that waning remains an area of uncertainty for consideration by the NICE committee.

2.4 Issue 4 – Modelling approach: state-transition or partitioned-survival

The company implemented a state-transition model (STM) for their 'primary' analysis, which is based on RESPONSE and RESPONSE-2 data, with overall survival (OS) for the ruxolitinib arm estimated indirectly from time to treatment discontinuation (TTD) and pre- and post-discontinuation survival. The company opted for this approach due to data immaturity, as few deaths were observed in the RESPONSE and RESPONSE-2 trials.

As requested by the EAG, the company have developed an alternative partitioned survival model (PSM) for the primary analysis to allow comparison with results from the MAJIC-PV analysis, which used a PSM due to the absence of IPD. Figures 1 and 2 in the company's TE response show KM data and extrapolations of OS from the RESPONSE and RESPONSE-2 trials respectively. The company chose the Weibull distribution for the extrapolation of ruxolitinib OS for both RESPONSE and RESPONSE-2. Table 1 in the TE

response compares cost-effectiveness from the CS base case STM model and the alternative PSM approach.

The company conclude that the PSM results based on direct extrapolation of OS data from RESPONSE and RESPONSE-2 are not suitable for decision making due to data immaturity and that the CS base case model remains the most appropriate approach. They also note that it is difficult to identify the source of the ICER differences between the analyses based on RESPONSE / RESPONSE-2 data and those based on MAJIC-PV, but that these are likely to be due to a combination of the modelling approach and population.

The EAG agrees that the alternative PSM reported in response to key issue 4 is subject to high uncertainty. However, we note that the CS base case STM model is subject to the same fundamental uncertainty due to immaturity of the OS data in the RESPONSE and RESPONSE-2 trials. Table 1 of the company's TE response shows that the model results are sensitive to the structural uncertainty over the use of an STM or PSM approach.

Comparison of the STM and PSM results in Table 1 of the company's TE response leads us to question whether either model accurately captures differences in survival between the splenomegaly subgroups. The difference between the ICERs for the populations with and without splenomegaly is larger with the PSM ([redacted] versus [redacted] per QALY) than with the STM ([redacted] versus [redacted] per QALY). We suggest that this may be at least partly a consequence of the use of pooled data from RESPONSE and RESPONSE-2 to estimate pre- and post-discontinuation survival in the STM, but the use of separate data for these trials to estimate OS in the PSM. If there are real differences in survival for people with and without splenomegaly during treatment with ruxolitinib, or during standard treatment, then the STM will underestimate the ICER difference between these subgroups.

Regarding the comparison with the ICER from the PSM analysis of MAJIC-PV ([redacted] per QALY, CS Table 39), we note that this is lower than both STM and PSM estimates using RESPONSE and RESPONSE-2 data. One might expect the ICER for the MAJIC-PV trial, which included a mixed population of patients with and without splenomegaly, to lie between the RESPONSE and RESPONSE-2 estimates. The reason for this difference is not clear, although it might be explained by other differences in the trial populations (see Issue 1 above).

2.5 Issue 5 – Model structure: health states and events

The company report cost-effectiveness results with an alternative model structure in response to EAG concerns that the company's model did not reflect the natural history of PV, including: the use of health states based on treatment phases rather than stages of disease; a lack of evidence underpinning the best available therapy (BAT) substates; and use of constant risks for key complications that did not increase with age. An outline of the alternative model structure and results is provided in the company's TE response, with more detail in an appendix.

The alternative model structure is illustrated in Figure 1 below, reproduced from the company's TE response appendix for convenience.

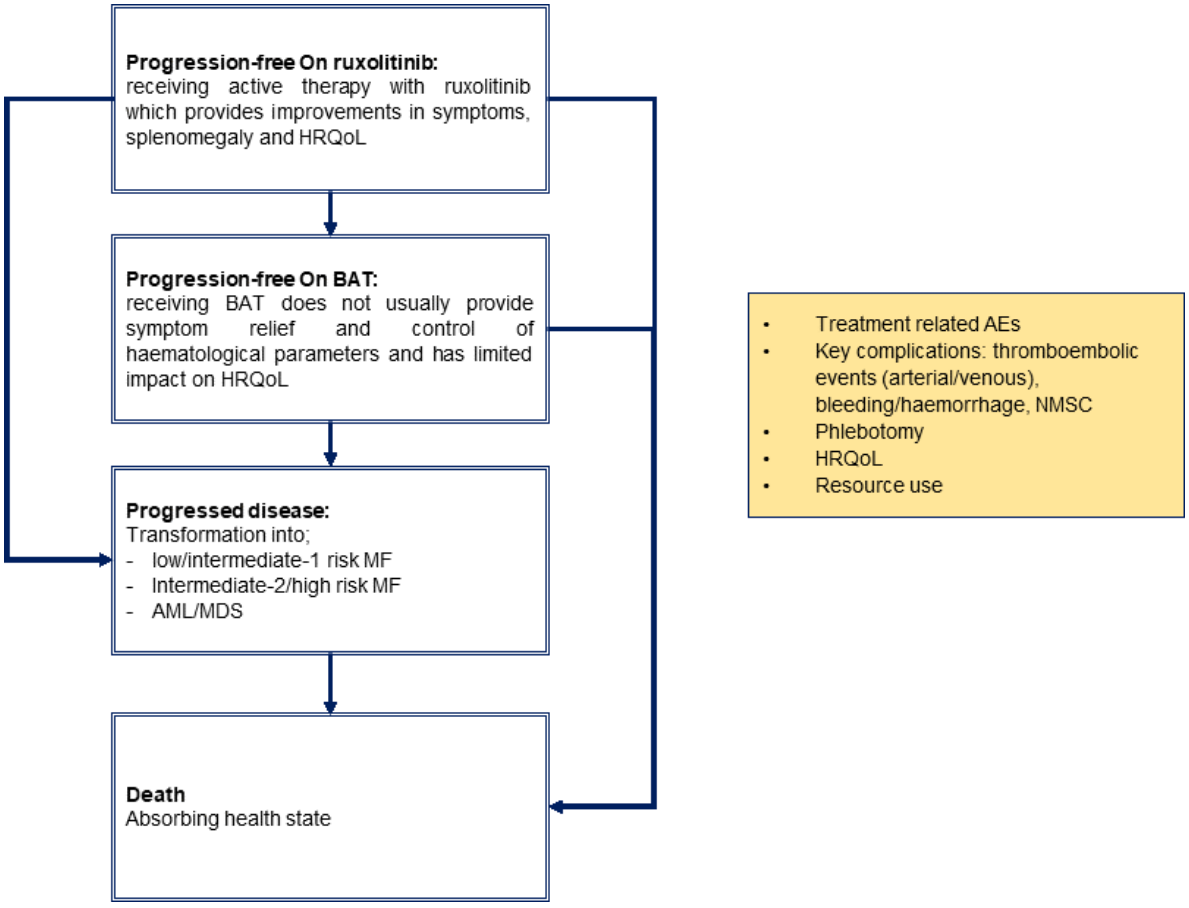


Figure 1 Simplified model structure schematic

Abbreviations: AE: adverse event; AML: acute myeloid leukaemia; BAT: best available therapy; HRQoL: health-related quality of life; MDS: myelodysplastic syndrome; MF: myelofibrosis; NMSC: non-melanoma skin cancer.
Source: Reproduced from Figure 1 in the company's response to technical engagement

The revised model includes a health state for progressed disease in addition to health states for 'progression-free' on ruxolitinib and on BAT, and death. For the purpose of this analysis, progression is defined as in the MAJIC-PV manuscript as transformation to myelofibrosis

(MF), acute myeloid leukaemia (AML) or myelodysplastic syndrome (MDS). An STM approach is used, with OS dependent on progression and treatment status as well as age.

In this alternative model structure, the company did not implement a partitioning of the BAT health state, which is in line with the EAG preferred assumption. Where possible, the company used the 5-year risk of events data from each trial to derive transitions from the progression-free health state to pre-progression death, MF, and AML/MDS; these transitions are in turn used to derive PFS. Where data were not available, the company made use of a set of assumptions supported by data from the literature. External data were used to inform transitions from progressed disease to death.

Table 2 below presents the results for the CS base case and EAG preferred assumptions applied to the original model provided with the company submission, along with results for the CS base case and EAG preferred assumptions applied to the alternative model structure including PFS. The EAG have checked through the model provided by the company and confirm that the correct modifications have been implemented to create the alternative model structure and produce the below results. Section 3.2 reports cumulative results changed from the company alternative structure model to the corresponding EAG model.

EAG comments on Issue 5

- Given the time constraints and insufficient detail, the EAG have not been able to conduct a critique of the external sources used to inform the transitions from the progressed disease state to the death state. However, the EAG favour the alternative model structure, as it reflects disease progression more directly.
- Results for the three trial populations in the alternative PFS model structure appear to have better face validity than in the original CS base case results (Table 2). Considering the company's primary analysis for the licensed population, the estimated ICER using the alternative model structure for the RESPONSE population ([redacted] per QALY) is lower than the ICER for the RESPONSE-2 population ([redacted] per QALY). The corresponding ICER for the MAJIC-PV population, which contains patients both with and without splenomegaly, is between the former two ICERs, at [redacted] per QALY.
- The EAG suggest that an independent clinical assessment should be conducted to assess the plausibility of the results. Further validation to demonstrate that the relative effects on overall survival from the model reflect the MAJIC-PV results, and that the results are consistent with longer-term registry data would strengthen confidence in the alternative model structure.

Table 2 Company and EAG results with the CS models and alternative PFS model structure

	Ruxolitinib			BAT			Incremental			
	Cost	LY	QALYs	Cost	LY	QALYs	Cost	LY	QALYs	ICER
RESPONSE										
CS base-case	[redacted]	11.45	[redacted]	[redacted]	9.28	[redacted]	[redacted]	2.17	[redacted]	[redacted]
EAG base-case	[redacted]	10.99	[redacted]	[redacted]	9.90	[redacted]	[redacted]	1.09	[redacted]	[redacted]
CS alternative PFS structure	[redacted]	13.28	[redacted]	[redacted]	10.68	[redacted]	[redacted]	2.60	[redacted]	[redacted]
EAG alternative PFS structure	[redacted]	13.28	[redacted]	[redacted]	10.68	[redacted]	[redacted]	2.60	[redacted]	[redacted]
RESPONSE-2										
CS base-case	[redacted]	12.25	[redacted]	[redacted]	10.46	[redacted]	[redacted]	1.79	[redacted]	[redacted]
EAG base-case	[redacted]	11.99	[redacted]	[redacted]	11.08	[redacted]	[redacted]	0.91	[redacted]	[redacted]
CS alternative PFS structure	[redacted]	13.17	[redacted]	[redacted]	11.83	[redacted]	[redacted]	1.34	[redacted]	[redacted]
EAG alternative PFS structure	[redacted]	13.17	[redacted]	[redacted]	11.83	[redacted]	[redacted]	1.34	[redacted]	[redacted]
MAJIC-PV										
CS base-case	[redacted]	9.65	[redacted]	[redacted]	8.02	[redacted]	[redacted]	1.63	[redacted]	[redacted]
EAG base-case	[redacted]	8.94	[redacted]	[redacted]	8.02	[redacted]	[redacted]	0.92	[redacted]	[redacted]
CS alternative PFS structure	[redacted]	10.08	[redacted]	[redacted]	8.18	[redacted]	[redacted]	1.90	[redacted]	[redacted]
EAG alternative PFS structure	[redacted]	10.08	[redacted]	[redacted]	8.18	[redacted]	[redacted]	1.90	[redacted]	[redacted]

Reproduced from company TE response Table 2 and EAG analyses.

CS: company submission; EAG: evidence assessment group; LYs: life years; QALY: quality-adjusted life year; BAT: best available therapy; ICER: incremental cost-effectiveness ratio.

2.6 Issue 6 – Extrapolation of time to ruxolitinib discontinuation

For time to treatment discontinuation (TTD) due to reasons other than death, the company base case uses data from the RESPONSE and RESPONSE-2 trials, separately. In contrast, data for pre-discontinuation survival (TTD due to death) and post-discontinuation survival are pooled because of the few events that occurred.

As requested by the EAG, the company have presented results from a scenario in which the data for TTD due to reasons other than death are pooled, which was referenced in the text but not provided in the initial company submission. A Weibull distribution is used in place of an odds spline model with one knot, the latter of which is implemented in the CS base case model; the Weibull distribution for TTD due to reasons other than death is an EAG preferred assumption. The EAG have checked through the updated model provided by the company and confirm that the correct adaptations have been made to produce results for the scenario.

Table 3 below compared the company and EAG base case results with corresponding results using pooled data for TTD due to reasons other than death. For the RESPONSE population, the ICER increases to [redacted] and [redacted] per QALY for the company and EAG models respectively, whilst the ICERs are reduced for the RESPONSE-2 population, to [redacted] and [redacted] per QALY for the company and EAG models, respectively.

Table 3 Base case results and scenario with pooled TTD for reasons other than death

	Ruxolitinib			BAT			ICER
	Cost	LY	QALYs	Cost	LY	QALYs	
RESPONSE							
CS base-case	[redacted]	11.45	[redacted]	[redacted]	9.28	[redacted]	[redacted]
EAG base-case	[redacted]	10.99	[redacted]	[redacted]	9.90	[redacted]	[redacted]
CS pooled TTD	[redacted]	12.10	[redacted]	[redacted]	10.05	[redacted]	[redacted]
EAG pooled TTD	[redacted]	11.77	[redacted]	[redacted]	10.73	[redacted]	[redacted]
RESPONSE-2							
CS base-case	[redacted]	12.25	[redacted]	[redacted]	10.46	[redacted]	[redacted]
EAG base-case	[redacted]	11.99	[redacted]	[redacted]	11.08	[redacted]	[redacted]
CS pooled TTD	[redacted]	11.48	[redacted]	[redacted]	9.52	[redacted]	[redacted]
EAG pooled TTD	[redacted]	11.02	[redacted]	[redacted]	10.01	[redacted]	[redacted]

Reproduced from company TE response Table 3 and EAG analyses.
CS: company submission; EAG: evidence assessment group; LYs: life years; QALY: quality-adjusted life year; BAT: best available therapy; ICER: incremental cost-effectiveness ratio; TTD: time to treatment discontinuation.

The EAG agree that, due to the data immaturity, the pooling of trial data for TTD due to reasons other than death is not appropriate. We note that additional long-term trial data is required to reduce uncertainty over the difference in cost-effectiveness between patients with and without splenomegaly.

2.7 Issue 7 – Source for utility estimates: MF-8D or EQ-5D

The company's response to TE does not include new information on this issue, although they state that further clinical validation was obtained in response to TE, which reiterated previous conclusions. The company reported MF-8D and EQ-5D utility estimates derived from trial data in CS section B.3.4.1 and in response to clarification questions B7, B8 and B9. EAG commentary and conclusions on these study-based utility estimates was provided in section 4.2.7 (pages 87-88 and 90) of the EAG report.

The NICE methods guide states that to make a case that the EQ-5D is not appropriate for a particular patient group, empirical evidence of a lack of content validity, supported by evidence of construct validity and responsiveness should be provided.[1] These concepts are defined, and appropriate tests suggested in NICE Decision Support Unit Technical Support Document 8 (Brazier and Longworth 2011).[2]

The company has previously reported results from 'exploratory' psychometric analyses conducted with EQ-5D and MPN-SAF data for patients randomised to ruxolitinib (N=[redacted]) in the RESPONSE-2 trial (CS section B.3.4.1 and a PowerPoint presentation provided in response to clarification question B8). In these analyses EQ-5D-5L data were appropriately mapped to 3L UK general population values, using NICE-preferred methods. We summarise available evidence to assess whether the NICE criteria for inappropriateness of the EQ-5D are met below:

Content validity *– evidence that the EQ-5D does not reflect all dimensions of health that are important to patients*

- The company cite expert and patient opinion about the symptoms of PV (fatigue, early satiety, abdominal discomfort, inactivity, concentration, night sweats, itching and bone pain), the similarity of symptoms for PV and MF, and the absence of such symptoms from the EQ-5D health state descriptions.
- The MPN Voice and Leukaemia Care TE response also states that the MF-8D reflects the lived experience of MPN patients in a way that a generic instrument such as the EQ-5D does not.

- The company report evidence of ceiling effects for the EQ-5D: [redacted] of patients reported no problems in all 5 EQ-5D dimensions at baseline (compared with [redacted] of patients within [redacted] of the maximum MPN-SAF total score).

Construct validity *– evidence that the EQ-5D does not reflect known differences between groups or does not converge with other relevant measures*

- Convergent validity was assessed with pairwise correlations between EQ-5D and MPN-SAF total scores and dimensions at baseline in the RESPONSE-2 ruxolitinib arm (n=[redacted]). The correlation was classified as 'very strong' (Pearson correlation coefficient [redacted]) between the EQ-5D utility and the MPN-SAF total symptom score.
- Correlations between the EQ-5D utility and some MPN-SAF symptoms were classified as weak (<0.3) for fever and unintentional weight loss, and 'moderate' (≥0.3 to <0.5) for abdominal discomfort, early satiety and night sweats.

Responsiveness *– evidence that the EQ-5D cannot detect change over time*

- Standardised response means for EQ-5D utility and MPN-SAF total scores at week 4, 8, 16, 28, 52 and 80 are shown in CS Figure 51. At all timepoints, the standardised response was lower for the EQ-5D ([redacted]) than for the MPN-SAF ([redacted]).

The EAG consider that the above evidence is weak and does not constitute sufficient grounds to reject use of the EQ-5D in our preferred analyses. In particular, we highlight that the test of convergent validity showed a strong correlation between the EQ-5D utility and the MPN-SAF total symptom score. This suggests that although some individual PV symptoms are not explicitly included in the EQ-5D descriptive system, they may still be reflected in one or more of the EQ-5D dimensions, and hence in the overall utility value.[2]

Furthermore, the EQ-5D did succeed in detecting a difference between the treatment arms in the RESPONSE-2 trial, see CS Figure 21 and Tables 5 and 6 in the company's clarification response. The company's estimated utility gain for ruxolitinib compared with BAT from RESPONSE-2 data was [redacted] (95% CI: [redacted]) (clarification question response Table 7). For comparison, the estimated utility gain from RESPONSE MF-8D data was [redacted] (95% CI: [redacted]). Although one might expect a disease-specific instrument such as the MF-8D to be more responsive than a generic instrument and differences between two trial populations, these estimates are similar.

3. Summary of additional EAG analysis

3.1 EAG model validity

The EAG conducted checks on the updated model provided in the company's TE response:

- The company base case analysis in the updated TE model does not differ from the company base case analysis in the model provided with the company submission
- The company implemented the partial survival model (PSM) discussed in Issue 4 correctly with no other changes to the original base case model.
- The company implemented the alternative model structure reported in Issue 5 correctly, with no other changes to the original base case model.
- All results provided by the company in their TE response have been checked against the model and are correct.
- The EAG successfully recreated the results with EAG preferred assumptions (Table 29 in the EAG report) in the updated model provided by the company

3.2 Results and conclusions

In their TE response, the company have stated that there is no change to their base case model. After reviewing the TE response and results, we do not wish to make any changes to the EAG preferred analyses, reported in Section 6 of the EAG report.

3.3 Additional results for Issue 5

Table 4, Table 5 and Table 6 below show the cumulative changes in results from the company base case and alternative PFS model structure to the corresponding model using the EAG preferred assumptions, as discussed in Section 2.5 above. We note that three of the changes that we applied to the company's original base case model are redundant in the alternative PFS model structure:

- A general population mortality constraint is applied throughout the time horizon in the alternative model structure;
- The alternative model structure does not make use of the relative treatment effect (HR) on OS from the MAJIC-PV trial
- The alternative model structure does not include a partition of the BAT state

Table 4 Cumulative changes from the company base case model to the EAG preferred analysis with alternative PFS structure: RESPONSE trial population

Assumption	Treatment	RESPONSE		
		Cost	QALYs	ICER
Company base case	BAT	£92,017	6.97	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
Company alternative structure	BAT	£123,685	8.10	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ General population mortality constraint	BAT	£123,685	8.10	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ Ruxolitinib TTD: Weibull	BAT	£123,685	8.10	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ HR OS: MAJIC-PV constant	BAT	£123,685	8.10	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ No BAT partition	BAT	£123,685	8.10	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ EQ-5D utilities	BAT	£123,685	7.67	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ Cost for Grade 1-2 TE events *(EAG alternative structure)*	BAT	£123,716	7.67	
	Ruxolitinib	[redacted]	[redacted]	[redacted]

Source: EAG analysis using the company's model.
QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; TTD: time to treatment discontinuation; HR: hazard ratio; OS: overall survival; TE: thromboembolic.

Table 5 Cumulative changes from the company base case model to the EAG preferred analysis with alternative PFS structure: RESPONSE-2 trial population

Assumption	Treatment	RESPONSE-2		
		Cost	QALYs	ICER
Company base case	BAT	£86,809	7.80	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
Company alternative structure	BAT	£123,336	8.83	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ General population mortality constraint	BAT	£123,336	8.83	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ Ruxolitinib TTD: Weibull	BAT	£123,336	8.83	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ HR OS: MAJIC-PV constant	BAT	£123,336	8.83	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ No BAT partition	BAT	£123,336	8.83	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ EQ-5D utilities	BAT	£123,336	8.32	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ Cost for Grade 1-2 TE events *(EAG alternative structure)*	BAT	£123,429	8.32	
	Ruxolitinib	[redacted]	[redacted]	[redacted]

Source: EAG analysis using the company's model.
QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; TTD: time to treatment discontinuation; HR: hazard ratio; OS: overall survival; TE: thromboembolic.

Table 6 Cumulative changes from the company base case model to the EAG preferred analysis with alternative PFS structure: MAJIC-PV trial population

Assumption	Treatment	Cost	QALYs	ICER
Company base case	BAT	£83,317	6.11	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
Company alternative structure	BAT	£91,002	6.22	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ HR OS: MAJIC-PV constant	BAT	£91,002	6.22	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ EQ-5D utilities	BAT	£91,002	5.88	
	Ruxolitinib	[redacted]	[redacted]	[redacted]
+ Cost for Grade 1-2 TE events *(EAG alternative structure)*	BAT	£91,033	5.88	
	Ruxolitinib	[redacted]	[redacted]	[redacted]

Source: EAG analysis using the company's model.
QALYs: quality-adjusted life years; ICER: incremental cost-effectiveness ratio; BAT: best available therapy; HR: hazard ratio; OS: overall survival; TE: thromboembolic.

References

1. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual 2022 [Available from: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation accessed 05/08/2022.
2. Brazier J, Longworth L. TSD 8 An introduction to the measurement and valuation of health for NICE submissions. NICE DSU Technical Support Document. Decision Support Unit, 2011.