**Pharmacoeconomics Review Article**

**Title page**

**A Review of Ruxolitinib for the Treatment of Myelofibrosis: A Critique of the Evidence**

Ros Wade,1 Robert Hodgson,1 Mousumi Biswas,1 Melissa Harden,1 Nerys Woolacott1

1Centre for Reviews and Dissemination (CRD), University of York, York, UK

**Corresponding author:** Ros Wade, Centre for Reviews and Dissemination (CRD), University of York, York YO10 5DD, UK. Telephone: +44 (0)1904 321051, E-mail: ros.wade@york.ac.uk.

**Running title:** Review of Ruxolitinib for the Treatment of Myelofibrosis

**Key words:** Ruxolitinib, myelofibrosis, splenomegaly

Word count: 5641

Contents

[Abstract 3](#_Toc458492642)

[1. Introduction 5](#_Toc458492643)

[2. The Decision Problem 6](#_Toc458492644)

[3. The Independent Evidence Review Group (ERG) Review 7](#_Toc458492645)

[3.1 Clinical Evidence 7](#_Toc458492646)

[3.1.1 Critique of the Clinical Evidence 9](#_Toc458492647)

[3.2 Cost Effectiveness Evidence 10](#_Toc458492648)

[3.2.1 Critique of the Cost Effectiveness Evidence 11](#_Toc458492649)

[3.3 Conclusions of the ERG Review 13](#_Toc458492650)

[4. NICE Guidance 14](#_Toc458492651)

[4.1 Preliminary Guidance 14](#_Toc458492652)

[4.1.1 Company’s Response to the Preliminary Guidance 14](#_Toc458492653)

[4.1.2 ERG Critique of the Response Submitted by the Company 15](#_Toc458492654)

[4.2 Final NICE Guidance 16](#_Toc458492655)

[5. ERG conclusion 16](#_Toc458492656)

# Abstract

As part of the National Institute for Health and Care Excellence’s (NICE) Single Technology Appraisal (STA) process, ruxolitinib was assessed to determine the clinical and cost effectiveness of its use in the treatment of disease-related splenomegaly or symptoms in adults with myelofibrosis. Ruxolitinib had previously been assessed as part of the STA process and was not recommended in NICE guidance issued in June 2013 (TA289). A review of TA289 was commissioned following the availability of new longer term survival data; a price discount patient access scheme (PAS) was also introduced. The Centre for Reviews and Dissemination (CRD) and Centre for Health Economics (CHE) Technology Appraisal Group at the University of York was commissioned to act as the independent Evidence Review Group (ERG). This article provides a summary of the company’s submission, the ERG review and the resulting NICE guidance issued in March 2016.

The main clinical effectiveness data were derived from two good quality multicentre randomised controlled trials (RCTs): Controlled myelofibrosis study with oral JAK inhibitor treatment (COMFORT)-II compared ruxolitinib with best available therapy (BAT) and COMFORT-I compared ruxolitinib with placebo. Both RCTs demonstrated a statistically significant reduction in splenomegaly and its associated symptoms in intermediate-2 and high risk myelofibrosis patients. Overall survival was statistically significantly improved with ruxolitinib compared with BAT at 3.5 years of follow-up in the COMFORT-II trial (HR 0.58, 95% CI 0.36 to 0.93). Grade 3-4 adverse events were more frequent in the ruxolitinib group than the BAT group; 42% compared with 25%. Evidence relating to patients with lower risk disease or low platelet counts (50 to 100 x 109/L) was less robust.

The company’s economicmodel was well presented and had an appropriate model structure. The base-case incremental cost-effectiveness ratio (ICER) was estimated to be around £45,000 per QALY gained (including the PAS discount). Extensive sensitivity and scenario analyses were presented, demonstrating that the estimated ICER was robust to a range of input values and assumptions made in the model. Alternative scenarios presented by the ERG showed only modest increases in the estimated ICER, primarily as a result of including an element of drug wastage within the model. Alternative scenarios resulted in estimated ICERs ranging from around £45,000 to £49,000 per QALY gained (including the PAS discount).

At the first appraisal meeting, the NICE Appraisal Committee concluded that ruxolitinib was clinically effective and was a cost effective use of National Health Service (NHS) resources for patients with high risk myelofibrosis, who meet NICE’s end of life criteria. Following the consultation, the company offered a revised PAS, resulting in a revised base case ICER of £31,229 per QALY gained. The company also presented new evidence on the cost-effectiveness of ruxolitinib in intermediate-2 and high risk subgroups and a revised version of the model. The NICE Appraisal Committee considered the new evidence and recommended ruxolitinib for the treatment of patients with intermediate-2 risk disease as well as patients with high risk disease, based on International Prognostic Scoring System (IPSS) prognostic factors.

**Key Points for Decision Makers**

* Good quality randomised controlled trials demonstrate the efficacy of ruxolitinib in splenomegaly and its associated symptoms in myelofibrosis patients. Overall survival was improved with ruxolitinib compared with best available therapy (BAT) at 3.5 years of follow-up, although the incidence of grade 3 or 4 adverse events was higher with ruxolitinib.
* The principal issues raised by the ERG relate to the company’s use of an individual patient discrete event simulation model which, while having a number of advantages over a Markov structure, placed additional demands on, and did not always fit well with, the available data from the COMFORT trials. Despite this, scenarios presented by the ERG resulted in only modest increases in the estimated ICER and ruxolitinib could be considered cost-effective according to NICE’s end of life criteria for patients with high risk disease.
* Accommodation by the company of the ERG’s concerns, together with a revised PAS, resulted in a revised base case ICER of £31,229 per QALY gained. In NICE’s final guidance, ruxolitinib was recommended for the treatment of patients with intermediate-2 risk disease as well as patients with high risk disease.

# 1. Introduction

The National Institute for Health and Care Excellence (NICE) is an independent organization responsible for providing national guidance to the NHS in England on the use of selected new health technologies. Single technology appraisals (STAs) evaluate a single product, device or other technology that has a single indication, for example, a new pharmaceutical product or licensed indication [[1](#_ENREF_1)]. The manufacturer or sponsor of the technology (hereafter referred to as the company) submits the principal evidence supporting the clinical and cost effectiveness of the product, and an external independent academic organization (the Evidence Review Group [ERG]) is commissioned to produce a review and critique of the evidence submitted [[2](#_ENREF_2)]. Clinical specialists, NHS commissioning experts and patient experts also provide evidence for consideration by the NICE Appraisal Committee in formulating their guidance [[1](#_ENREF_1)]. Once published, NICE technology guidance provides a legal obligation for NHS providers to reimburse technologies that have been recommended [[1](#_ENREF_1)].

Ruxolitinib has previously been assessed as part of the STA process and was not recommended for the treatment of disease-related splenomegaly or symptoms in adults with myelofibrosis in NICE guidance issued in June 2013 (TA289), although it was made available via the National Cancer Drugs Fund. Longer term survival data became available and a price discount patient access scheme (PAS) was introduced, therefore, a review of TA289 was commissioned.

This article presents a summary of the ERG’s independent critique of the company’s submission to NICE and its role in the subsequent development of NICE guidance for the use of ruxolitinib for the treatment of disease-related splenomegaly or symptoms in adults with myelofibrosis (review of TA289). The key issues that arose during the review process and subsequent committee decision making are summarised. Full details of the appraisal and the relevant documents can be found on the NICE website [[3](#_ENREF_3)].

# 2. The Decision Problem

Myelofibrosis is a myeloproliferative neoplasm which can develop *de novo* as primary myelofibrosis (PMF) or secondary to polycythaemia vera or essential thrombocythaemia, known as post-polycythaemia vera myelofibrosis (PPV-MF) and post-essential thrombocythaemia myelofibrosis (PET-MF). PMF, PPV-MF and PET-MF patients share a common transformation in the early haematopoietic stem cell of the Janus-associated kinase (JAK) 2 gene. The JAK/STAT (signal transducer and activator of transcription) pathway is essential for normal haematopoiesis, inflammatory responses and immune function. However, in patients with myelofibrosis, there is over-activation of the JAK/STAT signalling pathway, resulting in over-proliferation of blood cell precursors. Bone marrow becomes replaced with scar tissue, resulting in bone marrow failure and extramedullary haematopoiesis (blood cell production outside the bone marrow, in the liver and spleen), which results in swelling of the liver and spleen.

Myelofibrosis is a rare and debilitating disease associated with substantial morbidity and early mortality. The clinical features of myelofibrosis include splenomegaly (enlarged spleen), fatigue, pain, early satiety, dyspnoea, weight loss, night sweats, pruritis (itching) and progressive anaemia. Late stage myelofibrosis may transform to acute myeloid leukaemia. Survival varies considerably; median survival following diagnosis for patients with PMF is 4.0 to 5.7 years overall, while for patients with secondary myelofibrosis, median survival following diagnosis is 5.7 to 7.5 years [[4](#_ENREF_4)]. However, within each of these groups survival varies considerably according to a number of risk factors. Using the International Prognostic Scoring System (IPSS) developed by the International Working Group for Myelofibrosis research and Treatment (IWG-MRT), median overall survival is over 10 years for patients with low risk disease; approximately 8 years for patients with intermediate-1 risk disease; approximately 4 years for patients with intermediate-2 risk disease; and approximately 2 years for patients with high risk disease [[5](#_ENREF_5)]. A later study by the IWG-MRT developed a second risk score: the Dynamic International Prognostic Scoring System (DIPSS) [[6](#_ENREF_6)]. Using DIPSS, median survival was not reached in patients with low-risk disease; it was 14.2 years in intermediate-1 risk disease, 4 years in intermediate-2 risk disease, and 1.5 years in high risk disease [[6](#_ENREF_6)]. The prevalence of myelofibrosis is estimated to be 2.2 per 100,000 population [[4](#_ENREF_4)].

The only curative treatment for myelofibrosis is haematopoietic stem cell transplantation, however it is unsuitable for most patients and is associated with a high risk of complications and death. Therefore, clinical management focuses on symptom control. Treatments such as hydroxycarbamide, steroids and thalidomide are commonly used in the UK and are recommended for use by the British Committee for Standards in Haematology (BCSH) guidelines [[7](#_ENREF_7)]. The BCSH guideline for the diagnosis and management of myelofibrosis was revised in 2014 to include the recommendation of ruxolitinib as first line therapy for symptomatic splenomegaly and/or myelofibrosis-related constitutional symptoms regardless of JAK2 V617F mutation status [[8](#_ENREF_8)].

Ruxolitinib is a JAK2 inhibitor licensed for the treatment of disease-related splenomegaly or symptoms in adult patients with PMF, PPV-MF or PET-MF. Ruxolitinib has previously been assessed as part of the STA process and was not recommended for the treatment of disease-related splenomegaly or symptoms in adults with myelofibrosis in NICE guidance issued in June 2013 (TA289) [[9](#_ENREF_9)], although it was made available via the National Cancer Drugs Fund. Longer term survival data became available and a price discount PAS was introduced, therefore, a review of TA289 was commissioned. The NICE appraisal scope requested clinical and cost-effectiveness evidence for ruxolitinib with established clinical practice, compared with established clinical practice without ruxolitinib.

# 3. The Independent Evidence Review Group (ERG) Review

The company provided a submission to NICE on the use of ruxolitinib in the treatment of disease-related splenomegaly and symptoms in adults with myelofibrosis.

The ERG critically reviewed the evidence presented in the company’s submission by assessing: (i) whether the submission conformed to NICE methodological guidelines; (ii) whether the company’s interpretation and analysis of the evidence were appropriate; and (iii) the presence of other evidence or alternative interpretations of the evidence. In addition, the ERG identified areas requiring clarification, for which the company provided additional evidence [[10](#_ENREF_10)].

## 3.1 Clinical Evidence

The company’s submission incorporated a systematic review of studies evaluating the efficacy and safety of ruxolitinib in the treatment of patients with myelofibrosis. The main clinical effectiveness data were derived from two good quality multicentre randomised controlled trials (RCTs): Controlled myelofibrosis study with oral JAK inhibitor treatment (COMFORT)-II compared ruxolitinib with best available therapy (BAT) [[11](#_ENREF_11)] and COMFORT-I compared ruxolitinib with placebo [[12](#_ENREF_12)]. Both RCTs included patients with intermediate-2 or high risk myelofibrosis. The ERG did not identify any additional relevant RCTs.

Both RCTs demonstrated a statistically significant reduction in splenomegaly and its associated symptoms in intermediate-2 and high risk myelofibrosis patients. In COMFORT-II 28% of ruxolitinib patients achieved a 35% or greater reduction in spleen volume at week 48 (the primary outcome), compared with 0% of BAT patients (p<0.001), with a mean change in spleen volume of -30.1 versus +7.3% (p<0.001). The median time to initial response was 12.3 weeks in the ruxolitinib group, with 80% of responders maintaining a response at a median follow-up of 12 months. The median duration of maintenance of spleen response was 2.76 years in the ruxolitinib group [[11](#_ENREF_11)]. In COMFORT-I 42% of ruxolitinib patients achieved a 35% or greater reduction in spleen volume at week 24 (the primary outcome), compared with 0.7% of patients in the placebo group (p<0.001), with a mean change in spleen volume of -31.6 versus +8.1%. Most patients who achieved a 35% or greater reduction in spleen volume had achieved this by week 12, with 67% of responders maintaining a response at a median follow-up of 48 weeks [[12](#_ENREF_12)]. In both COMFORT trials, improvements in splenomegaly were observed in the control group when patients crossed over to receive ruxolitinib [[11](#_ENREF_11), [12](#_ENREF_12)].

Both COMFORT trials reported significant improvements in myelofibrosis-associated symptoms and health-related quality of life (HRQoL) compared with patients treated with BAT or placebo, measured using the disease specific European Organization for Research and Treatment Quality of Life Questionnaire (EORTC QLQ-C30), the Functional Assessment of Cancer Therapy-Lymphoma (FACT-Lym) scale and/or the modified Myelofibrosis Symptom Assessment Form (MFSAF) version 2 [[11](#_ENREF_11), [12](#_ENREF_12)]. In both COMFORT trials there were greater improvements in Global Health Status/Quality of Life in the ruxolitinib group than the control group; statistically significantly so in the COMFORT-I trial [[12](#_ENREF_12)]. In the COMFORT-I trial, significantly more ruxolitinib patients achieved a 50% or greater reduction in total symptom score using the MFSAF at week 24 than patients in the placebo group (45.9% versus 5.3%). Ruxolitinib patients had a 46.1% mean improvement in total symptom score at 24 weeks, compared with a 41.8% mean worsening in total symptom score in the placebo group [[12](#_ENREF_12)].

When compared with BAT in the COMFORT-II trial, ruxolitinib improved overall survival reaching borderline statistical significance at a median of 112 weeks follow-up (HR 0.52, 95% CI 0.27 to 1.00), and statistical significance at a median of 3 years (HR 0.48, 95% CI 0.28 to 0.85), and 3.5 years of follow-up (HR 0.58, 95% CI 0.36 to 0.93) [[13](#_ENREF_13), [14](#_ENREF_14)]. Compared with placebo ruxolitinib statistically significantly improved overall survival at a median follow-up of 51 weeks (HR 0.50, 95% CI 0.25 to 0.98) [[12](#_ENREF_12)] and 102 weeks (HR 0.58, 95% CI 0.36 to 0.95) in the COMFORT-I trial [[15](#_ENREF_15)]. However, the overall survival benefit did not reach statistical significance at three years (HR 0.69, 95% CI 0.46 to 1.03) [[16](#_ENREF_16)]. Neither trial was designed to be sufficiently powered to detect a statistically significant difference in overall survival. In addition, the majority of patients in the BAT and placebo groups crossed over to receive ruxolitinib during the course of the trial; therefore, the analyses are likely to underestimate the survival benefit of ruxolitinib. The company presented an overall survival analysis with adjustment for crossover, using a rank-preserving structural failure time model (RPSFTM). The RPSFTM analysis resulted in a larger reduction in the risk of death with ruxolitinib compared with BAT or placebo than that seen using an intention-to-treat (ITT) analysis; however, the results compared with placebo still did not reach statistical significance [[4](#_ENREF_4)].

Median overall survival was not reached in the ruxolitinib arm of either of the COMFORT trials; therefore, it was not possible to calculate the median (or mean) survival benefit associated with ruxolitinib compared with BAT or placebo. However, the company submission included a summary of an indirect comparison made between 100 PMF patients from the ruxolitinib arm of the COMFORT-II trial and 350 patients from the DIPSS cohort (a database of 519 PMF patients) [[4](#_ENREF_4)]. The number of observed deaths in the two cohorts were 30 (30%) on ruxolitinib and 256 (86%) on conventional care, generating estimates of median survival of 5 years (95% CI: 2.9-7.8) on ruxolitinib compared with 3.5 years (95% CI: 3.0- 3.9) for the DIPSS cohort [[17](#_ENREF_17)].

Grade 3-4 adverse events were more frequent in the ruxolitinib group than the BAT group; 42% compared with 25% [[11](#_ENREF_11)]. The most frequently occurring grade 3-4 adverse events were anaemia and thrombocytopenia, which were generally managed by dose modifications and/or blood transfusions and rarely led to treatment discontinuation. Haemoglobin levels decrease rapidly following initiation of ruxolitinib treatment, but then increase over time, almost returning to the baseline level. Platelet levels decrease rapidly following initiation of ruxolitinib treatment, but then remain reasonably constant. Other adverse events affecting more than 20% of ruxolitinib patients were diarrhoea, peripheral oedema and fatigue, although fatigue and peripheral oedema were also frequently reported in the placebo and BAT groups [[11](#_ENREF_11), [12](#_ENREF_12)]. In the COMFORT-II trial 72% of patients in the ruxolitinib group and 18% of patients in the BAT group required dose reductions or interruptions due to adverse events; thrombocytopenia was the most common reason for dose modifications [[11](#_ENREF_11)].

In the COMFORT-II trial 38% ruxolitinib patients had discontinued treatment by 48-weeks [[11](#_ENREF_11)]. By a median of 3.5 years of follow-up 63% ruxolitinib patients had discontinued treatment, primarily because of adverse events (20%) and disease progression (18%) [[14](#_ENREF_14)]. Of those patients who had crossed-over to ruxolitinib from the BAT arm, 60% had discontinued treatment by 3.5 years of follow-up, primarily because of adverse events (18%) [[14](#_ENREF_14)]. In the COMFORT-I trial 50% of patients had discontinued treatment by 3 years of follow-up, primarily because of disease progression (23%) [[16](#_ENREF_16)].

The company submission also described four non-RCT studies that included patients who were not eligible for the COMFORT trials; ROBUST, JUMP, Study 258 and EXPAND. The ROBUST study was a small phase 2 study of patients from the UK who had intermediate-1, intermediate-2 or high risk disease [[18](#_ENREF_18)]. The JUMP study was a large phase 3b extended access study in 25 countries for patients who had intermediate-1, intermediate-2 or high risk disease [[19](#_ENREF_19)]. Study 258 was a small phase 2 dose-finding study of patients with low platelet counts (50 to 100 x 109/L) [[20](#_ENREF_20)]. The EXPAND study was a small phase 1b dose-finding study of patients with low platelet counts (50 to 99 x 109/L) [[21](#_ENREF_21)].

The results of the non-RCT studies were generally consistent with the COMFORT trials, suggesting that ruxolitinib may also be effective at reducing spleen size and symptoms in patients with intermediate-1 risk myelofibrosis and patients with a low platelet count (50 – 99x109/L), although adverse events were more frequently reported in patients with a low platelet count, particularly thrombocytopenia, as might be expected. Three of the non-RCT studies were very small and none had a non-ruxolitinib control group; therefore, evidence relating to patients with lower risk disease or low platelet counts was less robust [[18-21](#_ENREF_18)].

### 3.1.1 Critique of the Clinical Evidence

The company’s systematic review did not appear to have missed any relevant RCTs. The two included RCTs were good quality and the results are likely to be reliable. However, both were conducted in patients with intermediate-2 or high risk myelofibrosis and a platelet count of 100 x 109/L or higher, therefore, the results are only generalizable to this subgroup of myelofibrosis patients, rather than the full licensed population [[11](#_ENREF_11), [12](#_ENREF_12)]. Four non-RCT studies that included patients with intermediate-1 risk disease and platelet counts between 50-100 x 109/L were also described, although the results of these smaller studies without a non-ruxolitinib control group are less reliable [[18-21](#_ENREF_18)].

Significantly more patients in the ruxolitinib group than the control group in both COMFORT trials met the primary endpoint of achieving a 35% or greater reduction in spleen volume [[11](#_ENREF_11), [12](#_ENREF_12)]. However, 63% of patients had discontinued treatment at 3.5 year follow-up in the COMFORT-II trial [[11](#_ENREF_11)] and half of the patients in the COMFORT-I trial had discontinued treatment at 3 year follow-up, primarily because of disease progression or adverse events [[12](#_ENREF_12)]. Therefore, there is some uncertainty about the long-term effectiveness and tolerability of ruxolitinib.

The results relating to myelofibrosis-associated symptoms are likely to be reliable; ruxolitinib was associated with clinically meaningful improvements in myelofibrosis associated symptoms. Health related quality of life appeared to be improved with ruxolitinib, although data were missing for a large proportion of patients for some of the quality of life results, therefore, the reliability and generalisability of these results is unclear [[11](#_ENREF_11), [12](#_ENREF_12)].

Despite the fact that the COMFORT-II trial was not designed to be sufficiently powered to detect a statistically significant difference between treatment groups, and the cross-over nature of the trial, overall survival was statistically significantly improved with ruxolitinib compared with BAT at the later time points (3 years and 3.5 years) using both ITT and RPSFTM analyses. The RPSFTM was used to estimate the true effect of ruxolitinib on overall survival, adjusting for crossover. The company also considered the two stage approach, the iterative parameter estimation (IPE) and the inverse probability of censoring weights (IPCW) methods, presenting their reasoning for choosing the RPSFTM over other methods of adjusting for cross-over. Whilst all methods of adjusting survival estimates in the presence of treatment switching have limitations, the ERG considered the RPSFTM (which relies critically on the ‘common treatment effect’ assumption) to be appropriate in the present context. The overall survival benefit in the COMFORT-I trial did not reach statistical significance at the later time points, even after adjustment for crossover. Median survival was not reached, so the data were not mature and the duration of any survival benefit due to ruxolitinib could not be calculated from the trial data.

Evidence relating to patients with lower risk disease or low platelet counts was less robust and further data from randomised studies are required to confirm the efficacy and tolerability of ruxolitinib in these populations.

## 3.2 Cost Effectiveness Evidence

The company presented a systematic review of the cost-effectiveness studies of ruxolitinib for the treatment of myelofibrosis. Two studies were identified; one compared ruxolitinib to BAT from a Canadian societal perspective [[22](#_ENREF_22)], the other was the model presented in the previous STA of ruxolitinib [[2](#_ENREF_2)]. Neither study was considered relevant, therefore a de novo model was developed.

The model submitted was an individual patient discrete event simulation model (Figure 1). The company justified the use of this approach on the basis that this type of model allows for increased flexibility and allows the progressive nature of myelofibrosis to be modelled in a more transparent way than a Markov model, which would require the excessive use of short-term ‘temporary’ states. The model compared ruxolitinib with BAT, which was assumed to consist of the basket of therapies used in the comparator arm of the COMFORT-II trial. The population modelled was intermediate-2 and high risk myelofibrosis patients, consistent with the COMFORT-II trial. The primary outputs of the model were quality-adjusted life-years (QALYs) and costs. The model time horizon was 35 years, which was designed to simulate a lifetime time horizon. Both costs and benefits were discounted at 3.5%, in line with NICE recommendations. A National Health Service (NHS) and Personal Social Services (PSS) perspective was taken. Costs were separated into drug acquisition costs, costs associated with the management of myelofibrosis, adverse event costs, costs associated with leukaemic transformation and end of life costs. The economic model tracks changes in HRQoL according to the different phases of treatment with utility values based on a time-trade-off (TTO) analysis of a newly developed condition-specific preference based measure for myelofibrosis [[23](#_ENREF_23)]. The new myelofibrosis specific measure was based on two existing measures, the MFSAF and the EORTC QLQ-C30, from which items were selected to capture the factors determining disease severity.

Figure Simplified schematic of the model structure



BAT, best available therapy

The model contained four mutually exclusive health states with ‘alive’ states being defined by therapy phase: on ruxolitinib (patients have some moderate improvement in symptoms and splenomegaly and HRQoL); on BAT (limited symptom control with minimal impact on HRQoL); on supportive care (associated with progressive worsening of symptoms until death); death. A 24 week stopping rule was applied, where ruxolitinib treatment is only continued if patients are considered to be responding to treatment. This stopping rule is based on definitions of response set out in the IWG-MRT/ELN guidelines [[24](#_ENREF_24)] and includes both symptom response and spleen size response. Time in each ‘health state’ is dependent upon initial treatment decision and treatment response and was modelled based on data on overall survival and time on treatment observed in the COMFORT-I and COMFORT-II trials.

The company’s submission presented both deterministic and probabilistic sensitivity analyses to demonstrate the robustness of the estimated ICER along with extensive scenario analysis examining the impact of a number of structural assumptions made in the base-case model. The base-case incremental cost effectiveness ratio (ICER) presented in the company’s submission, including a PAS discount, was estimated to be £44,905 per QALY gained in the deterministic analysis and £44,625 per QALY gained in the probabilistic sensitivity analysis. The probability of ruxolitinib being a cost-effective strategy at thresholds of £30,000, £40,000 and £50,000 per QALY gained was 0.33%, 4.32% and 95.02% respectively. The sensitivity and scenario analyses presented by the company showed that the ICER rarely exceeded £50,000 per QALY gained. Exceptions to this included reducing the time horizon to 5 years, using an alternative parametric function to estimate overall survival in BAT patients, and using ITT (rather than cross-over adjusted) analysis to estimate overall survival on BAT. None of these scenarios was considered particularly plausible.

### 3.2.1 Critique of the Cost Effectiveness Evidence

The ERG agreed that neither of the studies identified in the systematic review of cost-effectiveness studies of ruxolitinib was relevant to the appraisal. The Canadian study was not relevant to the UK and the UK study was part of the previous NICE STA submission for ruxolitinib and therefore based on more limited data.

The use of an individual patient discrete event simulation model appeared justified. Given the progressive nature of myelofibrosis, additional flexibility was needed to allow a number of structural assumptions of the model to be evaluated, which would not have been possible if a Markov structure had been adopted. The model structure was considered representative of how ruxolitinib is likely to be used in the NHS and the disease progression of myelofibrosis patients. The discrete event simulation structure adopted was, however, a more complex model structure than a Markov structure, placing additional demands on the available data. This required the company to make a number of assumptions, which while mostly plausible and justified by the available evidence were subject to a degree of uncertainty. These assumptions included:

* Post-discontinuation survival for patients initiating on ruxolitinib being the same for both early discontinuers and responders to treatment.
* Spleen responders and symptom responders were assumed to have spent the same amount of time on ruxolitinib and BAT, and assumed to have the same survival time.
* Time to death for patients on BAT was assumed to follow the same survival curve as used for discontinuation for reasons other than death.
* A 24 week survival benefit for non-responders was assumed for patients initiating on ruxolitinib.

The company’s submission presented empirical justification and/or extensive scenario analysis for all of these assumptions and there was generally limited impact on the ICER generated. The scenario analysis, however, did not address the joint uncertainty of these assumptions and as such a degree of uncertainty over the overall impact of these structural assumptions remains unexplored.

Further to the above the ERG also identified the following issues with the model:

* For patients who were considered non-responders to ruxolitinib after an initial 24 week treatment period it is assumed that overall survival would be increased by 24 weeks over patients initiating on BAT. No clinical evidence was provided to support this assumption, and while the ERG accepts that some OS benefit may be experienced by these patients, the evidence does not support the OS benefit assumed in the model.
* The comparator used in the model was BAT which comprised a basket of treatments used in the COMFORT-II trial. There were concerns regarding the composition of this basket of therapies and how well it represented UK practice. In particular the inclusion of lenalidomide was considered inappropriate as this drug is rarelyused in the UK. Furthermore, it was felt that this basket of therapies should have included allogeneic haematopoietic stem cell transplantation as, while not suitable for all myelofibrosis patients, it is the only curative therapy available for myelofibrosis and has been observed to result in significant survival benefit over other myelofibrosis treatments excluding ruxolitinib.
* The model presented in the company submission did not allow for any drug wastage. This was considered inappropriate due to the fact that adverse events are often managed with either dose reductions or interruptions.
* In the base-case model, the mortality rate for ruxolitinib responders while on treatment was based on mortality data from the COMFORT-II trial. The ERG considered this to be an inappropriate source given the limited sample size and infrequency of deaths in this population. The ERG therefore conducted an alternative analysis using an alternative mortality rate of 7.06% which was estimated using data on the mean time responders spent on ruxolitinib and mortality rates from UK life tables.

The ERG conducted a number of additional analyses considering a number of alternative assumptions and exploring their impact on the estimated ICER. Adjusting the survival rate for responders, removing lenalidomide from the BAT basket of therapies and assuming a 5% rate of drug wastage individually had only a modest impact on the estimated ICER, which remained at less than £50,000 per QALY gained (including the PAS discount) in all scenarios.

An alternative base-case was presented by the ERG, based on a combination of a number of scenario analyses presented by the company and assuming a 5% rate of drug wastage, excluding lenalidomide from the BAT basket of therapies, assuming that for non-responders time on BAT is adjusted for time spent on ruxolitinib, and assuming that the BAT discontinuation rate is underestimated by 20%. This alternative base-case was presented assuming both a 0% mortality rate for ruxolitinib responders and an alternative rate of 7.06%, based on a crude estimate of expected mortality made by the ERG. Again, this had only a modest impact on the estimated ICER, which remained at less than £50,000 per QALY gained (including the PAS discount).

## 3.3 Conclusions of the ERG Review

Evidence from two good quality RCTs demonstrated that ruxolitinib is effective at reducing splenomegaly and its associated symptoms and can increase overall survival in intermediate-2 and high risk myelofibrosis patients who can tolerate ruxolitinib and remain on therapy. However withdrawal rates were high with more than half of patients having discontinued therapy after three years. Evidence relating to patients with lower risk disease or low platelet counts (50 to 100 x 109/L) was less robust.

The de novo model was well presented and had an appropriate model structure, which showed good validation with the trial data from the COMFORT-II study. The extensive sensitivity and scenario analyses presented showed the estimated ICER to be largely robust to a range of input values and assumptions made in the model. The alternative scenarios presented by the ERG showed only a modest increase in the estimated ICER, which remained at less than £50,000 per QALY gained (including the PAS discount).

# 4. NICE Guidance

## 4.1 Preliminary Guidance

The NICE Appraisal Committee considered the evidence available on the clinical and cost effectiveness of ruxolitinib for the treatment of disease-related splenomegaly or symptoms in adults with myelofibrosis, alongside expert testimony from clinical specialists and patient representatives. It also took into account the effective use of NHS resources. The Committee deliberated over whether ruxolitinib for the treatment of myelofibrosis meets NICE’s end of life criteria. The preliminary NICE recommendation, recorded in the appraisal consultation document (ACD), was that ruxolitinib was recommended for the treatment of disease-related splenomegaly or symptoms in adults with high risk myelofibrosis that does meet the end of life criteria, with a life expectancy of less than 24 months. The Committee recognised that ruxolitinib was clinically effective and was a cost effective use of NHS resources in this subgroup of patients, with an ICER likely to fall between £45,000 and £49,000 per QALY gained (including the PAS discount). However, ruxolitinib was not a cost-effective use of NHS resources and was not recommended for the treatment of disease-related splenomegaly or symptoms in adults with intermediate risk myelofibrosis that does not meet the end of life criteria.

### 4.1.1 Company’s Response to the Preliminary Guidance

In response to the ACD, the company provided a revised base case. The base case was revised to account for (a) errors identified by the ERG on the inclusion of leukaemic transformation (LT) in the economic model, (b) adjusting the baseline utility by a factor of 10%, (c) change to the treatment pathway for responders to ruxolitinib assuming responders spend time on BAT after ruxolitinib discontinuation (Figure 2) and (d) the exclusion of lenolidomide. A revised PAS was also offered by the company.

Figure Updated simplified schematic of the model structure



Transition for non-responders

Transition for responders

BAT, best available therapy

These changes resulted in a revised base case ICER for ruxolitinib of £31,229 per QALY gained (including the revised PAS).

An exploratory analysis was conducted to illustrate the impact of myelofibrosis on caregivers’ quality of life. The company justified this analysis on the basis that the NICE methods guide states that “*For the reference case, the perspective on outcomes should be all direct health effects, whether for patients or other people*”. On this basis the company explained that the debilitating impact of myelofibrosis on patients means that many will require care from family and friends and that this burden should potentially be considered when evaluating the impact of ruxolitinib therapy. However, there was a lack of evidence regarding the extent to which myelofibrosis affects the quality of life of caregivers. Therefore, a series of assumptions was made to estimate the impact of myelofibrosis on caregiver quality of life in the analysis which was not observed in the COMFORT trials. The assumptions related to caregiver quality of life were that 57.48% of myelofibrosis patients required some help, quality of life of 1.76 caregivers was affected for each of the myelofibrosis patients that needed care, caregivers of myelofibrosis patients on BAT experienced a 0.1 decrement to their utility, and caregivers’ quality of life returned to that of the general population, whilst on ruxolitinib. However, costs associated with caregivers were not included in the analysis. The exploratory analysis reduced the revised base case ICER (including revised PAS) to £28,111 per QALY gained.

### 4.1.2 ERG Critique of the Response Submitted by the Company

The ERG reviewed the revised model. Two key drivers of the revised economic analysis were the inflation of the baseline utility by 10%, and responders on ruxolitinib spending 30% of their time alive on BAT instead of 45%, after discontinuation of ruxolitinib.

The adjustment of the baseline utility was supported by clinical opinion instead of any empirical data. The ERG recognised that some inflation may be appropriate as this could appropriately represent the baseline utility of patients enrolled in COMFORT-II, although there was limited evidence to suggest an appropriate inflation factor. Therefore, the ERG was uncertain about the plausibility of the magnitude of the presented inflation factor, which was not considered a plausible scenario in the original analyses. The adjustment of the baseline utility had a large impact on the ICER.

The ERG noted that the original assumption that patients discontinuing ruxolitinib therapy moved directly to supportive care was potentially a conservative assumption. However, the ERG was concerned regarding the changes to this assumption implemented in the company’s revised model. Specifically, the ERG noted that the revised 30% rate was almost entirely arbitrary and based on minimal data. Furthermore, it is far from certain that all patients discontinuing ruxolitinib will transit to BAT and it is far more likely that only a proportion will. This was not reflected in the new assumption. This change to the model also made the optimistic assumption that patients on BAT following ruxolitinib therapy would have the same quality of life as those initiating on BAT. The assumption seems far from obvious as there is clear potential for patients on BAT following ruxolitinib therapy to have lower quality of life than those who initiate on BAT. There was therefore significant potential that this new assumption was overly optimistic and underestimated the ICER.

Therefore, the ERG carried out two scenario analyses. The first assumed no inflation to base line utility and that patients moved directly to supportive care after discontinuation of ruxolitinib. The second assumed a 10% inflation in base line utility and that 30% of patients’ time after discontinuing ruxolitinib was spent on BAT. The ICERs for the analyses were £35,632 and £31,676 respectively per QALY gained (with the revised PAS). The ERG considered the ICERs from these two scenarios to be at least as plausible as the ICER presented in the company’s resubmission of £31,229 (with the revised PAS).

The exploratory analysis incorporating care givers’ quality of life within the economic analysis had a large impact on the ICER. However, the ERG was uncertain about the quality of the evidence used to justify this analysis, particularly due to the fact that the quality of life losses were observed in a different disease area. The ERG therefore considered this analysis to be subject to significant uncertainty.

## 4.2 Final NICE Guidance

Following the consultation on the preliminary guidance, the NICE Appraisal Committee released the following final guidance to the NHS (review of TA289) [[25](#_ENREF_25)]:

Ruxolitinib is recommended as an option for treating disease-related splenomegaly or symptoms in adults with primary myelofibrosis (also known as chronic idiopathic myelofibrosis), post polycythaemia vera myelofibrosis or post essential thrombocythaemia myelofibrosis, only:

* in people with intermediate-2 or high-risk disease, and
* if the company provides ruxolitinib with the discount agreed in the patient access scheme.

People whose treatment with ruxolitinib is not recommended in this NICE guidance, but was started within the NHS before this guidance was published, should be able to continue treatment until they and their clinician consider it appropriate to stop.

# 5. ERG conclusion

This was a typical example of the NICE STA process. This case study, however, highlights that even where the company submission is relatively strong, the ERG still has a role in interpreting the clinical and cost-effectiveness evidence submitted by companies.

The principal issues raised in this STA relate to the model approach taken by the company. The use of an individual patient discrete event simulation model can be considered relatively novel and it has a number of advantages over a Markov structure, allowing the progressive nature of myelofibrosis to be modelled in a more transparent way than would have been possible using a Markov model. It was, however, also clear that the use of a discrete event simulation model placed additional demands on the available data and the structure of the model did not always fit well with the data available in the COMFORT trials that were used to populate the model. As a consequence, the issues raised by the ERG focused largely on exploring the uncertainty around a number of inputs for which evidence from the COMFORT-I and COMFORT-II trials was least robust, and on how these data were incorporated into the economic model. Sensitivity and scenario analyses presented by both the company and the ERG, however, demonstrated that the results of the cost-effectiveness analysis were robust to a variety of input values and as such the estimated ICER did not exceed the cost-effectiveness threshold in any plausible scenarios.

**Acknowledgements:** We would like to thank Professor Mary Frances McMullin, Professor of Clinical Haematology at Queens University Belfast and Consultant Haematologist at Belfast City Hospital, for clinical advice throughout the project.

**Compliance with Ethical Standards:** This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment Programme (project number 14/206/01) and will be published as part of a compendium of ERG articles in *Health Technology Assessment*. See the HTA programme website (<http://www.hta.ac.uk>) for further project information. This summary of the ERG report was compiled after the Appraisal Committee’s review and incorporates additional information and comment from the authors on the STA process and iterations of the NICE guidance not covered by the HTA report.

The views and opinions expressed herein are those of the authors and do not necessarily reflect those of NICE or the Department of Health.

This work is Crown copyright (UK).

**Author contributions:** Ros Wade, Robert Hodgson, Mousumi Biswas, Melissa Harden and Nerys Woolacott all formed part of the ERG that produced the ERG report that this paper describes. Ros Wade wrote the first draft of the manuscript. All authors commented on the manuscript and approved the final version.

All authors (Ros Wade, Robert Hodgson, Mousumi Biswas, Melissa Harden and Nerys Woolacott) do not have any conflicts of interest to declare.

**References**

1. National Institute for Health and Clinical Excellence (NICE). Guide to the single technology appraisal (STA) process. London: NICE; 2006.

2. Wade R, Rose M, Neilson AR, Stirk L, Rodriguez-Lopez R, Bowen D, et al. Ruxolitinib for the treatment of myelofibrosis: a NICE single technology appraisal. Pharmacoeconomics. 2013;31 (10):841–52.

3. National Institute for Health and Care Excellence (NICE). Ruxolitinib for treating disease-related splenomegaly or symptoms in adults with myelofibrosis (TA386). 2016 [cited 2016 5 April]; Available from: <https://www.nice.org.uk/guidance/ta386>

4. Novartis. Single technology appraisal. Ruxolitinib for treating disease-related splenomegaly or symptoms in adults with myelofibrosis (review of TA289). Company evidence submission: Novartis; 2015.

5. Cervantes F, Dupriez B, Pereira A, Passamonti F, Reilly JT, Morra E, et al. New prognostic scoring system for primary myelofibrosis based on a study of the International Working Group for Myelofibrosis Research and Treatment. Blood. 2009;113(13):2895-901.

6. Passamonti F, Cervantes F, Vannucchi AM, Morra E, Rumi E, Pereira A, et al. A dynamic prognostic model to predict survival in primary myelofibrosis: a study by the IWG-MRT (International Working Group for Myeloproliferative Neoplasms Research and Treatment). Blood. 2010;115(9):1703–8.

7. Reilly JT, McMullin MF, Beer PA, Butt N, Conneally E, Duncombe A, et al. Guideline for the diagnosis and management of myelofibrosis. Br J Haematol. 2012;158(4):453-71.

8. Reilly JT, McMullin MF, Beer PA, Butt N, Conneally E, Duncombe AS, et al. Use of JAK inhibitors in the management of myelofibrosis: a revision of the BCSH guidelines for investigation and management of myelofibrosis 2012. Br J Haematol. 2014;167(3):418-20.

9. National Institute for Health and Care Excellence. NICE TA289 and ruxolitinib final appraisal determination document. 2013. Ruxolitinib for disease-related splenomegaly or symptoms in adults with myelofibrosis. Available from: <http://www.nice.org.uk/nicemedia/live/14194/64265/64265.pdf>. (Accessed 20 February 2014). 2013.

10. Hodgson R, Wade R, Biswas M, Harden M, Woolacott N. Ruxolitinib for disease-related splenomegaly or symptoms in adults with myelofibrosis (review of TA289): A Single Technology Appraisal: CRD and CHE Technology Assessment Group; 2015.

11. Harrison C, Kiladjian JJ, Al-Ali HK, Gisslinger H, Waltzman R, Stalbovskaya V, et al. JAK inhibition with ruxolitinib versus best available therapy for myelofibrosis. N Engl J Med. 2012;366(9):787–98.

12. Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. N Engl J Med. 2012;366(9):799–807.

13. Cervantes F, Vannucchi AM, Kiladjian J-J, Al-Ali HK, Sirulnik A, Stalbovskaya V, et al. Three-year efficacy, safety, and survival findings from COMFORT-II, a phase 3 study comparing ruxolitinib with best available therapy for myelofibrosis. Blood. 2013;122:4047-53.

14. Harrison C, Niederwieser D, Vannucchi A, Kiladjian JJ, Tiziano Barbui T, Gisslinger H, et al. Results from a 3.5-year update of COMFORT-II, a phase 3 study comparing ruxolitinib (RUX) with best available therapy (BAT) for the treatment of myelofibrosis. Haematologica. 2014;99 (suppl 1):126.

15. Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. Efficacy, safety and survival with ruxolitinib in patients with myelofibrosis: results of a median 2-year follow-up of COMFORT-I. Haematologica. 2013;98:1865-71.

16. Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. Efficacy, safety, and survival with ruxolitinib in patients with myelofibrosis: results of a median 3-year follow-up of COMFORT-I. Haematologica. 2015;100:479-88.

17. Passamonti F, Maffioli M, Cervantes F, Vannucchi AM, Morra E, Barbui T, et al. Impact of ruxolitinib on the natural history of primary myelofibrosis: a comparison of the DIPSS and the COMFORT-2 cohorts. Blood. 2014;123:1833-5.

18. Mead AJ, Milojkovic D, Knapper S, Garg M, Chacko J, Farquharson M, et al. Response to ruxolitinib in patients with intermediate-1-, intermediate-2-, and high-risk myelofibrosis: results of the UK ROBUST Trial. Br J Haematol. 2015;170(1):29-39.

19. Martino B, le Coutre P, Griesshammer M, Illmer T, Schlag R, Waller CF, et al. Safety and efficacy of Ruxolitinib in an open-label, multicenter, single-arm, expanded-access study in patients with myelofibrosis (MF): an 1144-patient update Blood. 2014;124(21):3197.

20. Talpaz M, Paquette R, Afrin L, Hamburg SI, Prchal JT, Jamieson K, et al. Interim analysis of safety and efficacy of ruxolitinib in patients with myelofibrosis and low platelet counts. J Hematol Oncol. 2013;6(1):81.

21. Harrison CN, Gisslinger H, Miller CB, Kiladjian J, Atienza E, Stalbovskaya V, et al. Expand: a phase 1b, open-label, dose-finding study of ruxolitinib in patients with myelofibrosis and baseline platelet counts between 50 x 109/L and 99 x 109/L. Blood. 2012;120(21):177.

22. El Ouagari K, Knight CJ, Mendelson ET. Cost-effectiveness of ruxolitinib versus best-available therapy for medical treatment of myelofibrosis: Canadian societal perspective. Blood. 2012;120(21):4255.

23. Mukuria C, Rowen D, Brazier JE, Young TA, Nafees B. Deriving a Preference-Based Measure for Myelofibrosis from the EORTC QLQ-C30 and the MF-SAF. Value Health. 2015;18(6):846-55.

24. Tefferi A, Cervantes F, Mesa R, Passamonti F, Verstovsek S, Vannucchi AM, et al. Revised response criteria for myelofibrosis: International Working Group-Myeloproliferative Neoplasms Research and Treatment (IWG-MRT) and European LeukemiaNet (ELN) consensus report. Blood. 2013;122(8):1395-8.

25. National Institute for Health and Care Excellence (NICE). Ruxolitinib for treating disease-related splenomegaly or symptoms in adults with myelofibrosis (TA386). Final Appraisal Determination (FAD). 2016 [cited 2016 5 April]; Available from: <https://www.nice.org.uk/guidance/TA386/documents/final-appraisal-determination-document>