NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Draft guidance consultation

Ruxolitinib for treating polycythaemia vera

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using ruxolitinib in the NHS in England. The evaluation committee has considered the evidence submitted by the company and the views of non-company stakeholders, clinical experts and patient experts.

This document has been prepared for consultation with the stakeholders. It summarises the evidence and views that have been considered, and sets out the recommendations made by the committee. NICE invites comments from the stakeholders for this evaluation and the public. This document should be read along with the evidence (see the committee papers).

The evaluation committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the recommendations sound and a suitable basis for guidance to the NHS?
- Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of age, disability, gender reassignment, pregnancy and maternity, race, religion or belief, sex or sexual orientation?

Note that this document is not NICE's final guidance on this technology. The recommendations in section 1 may change after consultation.

After consultation:

- The evaluation committee will meet again to consider the evidence, this evaluation consultation document and comments from the stakeholders.
- At that meeting, the committee will also consider comments made by people who are not stakeholders.
- After considering these comments, the committee will prepare the final draft guidance.
- Subject to any appeal by stakeholders, the final draft guidance may be used as the basis for NICE's guidance on using ruxolitinib in the NHS in England.

For further details, see NICE's manual on health technology evaluation.

The key dates for this evaluation are:

- Closing date for comments: 23 June 2023
- Second evaluation committee meeting: 12 July 2023
- Details of the evaluation committee are given in section 4

1 Recommendations

1.1 Ruxolitinib is not recommended, within its marketing authorisation, for treating polycythaemia vera in adults who cannot tolerate hydroxycarbamide (also called hydroxyurea) or when the condition is resistant to it.

1.2 This recommendation is not intended to affect treatment with ruxolitinib that was started in the NHS before this guidance was published. People having treatment outside this recommendation may continue without change to the funding arrangements in place for them before this guidance was published, until they and their NHS clinician consider it appropriate to stop.

Why the committee made these recommendations

Standard treatment to control blood cell count (cytoreductive therapy) in polycythaemia vera is hydroxycarbamide or interferon alfa. Ruxolitinib would be used for people who cannot tolerate hydroxycarbamide or when the condition is resistant to it.

Results from clinical trials suggest that ruxolitinib is more effective than standard treatment at controlling blood cell counts and reducing spleen size. But whether it increases how long people live is uncertain.

There are uncertainties in the company's cost-effectiveness model, particularly around its structure and how survival is modelled. This means that it is not possible to determine the most likely cost-effectiveness estimates. But all possible estimates are considerably above the range that NICE usually considers an acceptable use of NHS resources. So, ruxolitinib is not recommended.

2 Information about ruxolitinib

Marketing authorisation indication

2.1 Ruxolitinib (Jakavi, Novartis Pharmaceuticals) is indicated for 'the treatment of adult patients with polycythaemia vera who are resistant to or intolerant of hydroxyurea [hydroxycarbamide]'.

Dosage in the marketing authorisation

2.2 The dosage schedule is available in the summary of product characteristics for ruxolitinib.

Price

2.3 The list prices of ruxolitinib for 56-capsule packs are £1,428 (5 mg), £2,856 (10 mg), £2,856 (15 mg) and £2,856 (20 mg; all prices excluding VAT; BNF online accessed May 2023).

2.4 The company has a commercial arrangement. This makes ruxolitinib available to the NHS with a discount and it would have also applied to this indication if the technology had been recommended. The size of the discount is commercial in confidence. It is the company's responsibility to let relevant NHS organisations know details of the discount.

3 Committee discussion

The evaluation committee considered evidence submitted by Novartis Pharmaceuticals, a review of this submission by the external assessment group (EAG), and responses from stakeholders. See the committee papers for full details of the evidence.

The condition

Polycythaemia vera

3.1 Polycythaemia vera is a bone marrow condition that leads to an increase in the number of cells in the blood. It mostly affects the number of red blood cells. As more blood cells are made, the blood becomes thicker.

This can lead to complications such as gout, bleeding problems and blood clots. Blood clots can cause strokes, heart attacks, or blockage of an artery in the lungs (pulmonary embolism) or in a vein deep in a muscle (deep vein thrombosis). Polycythaemia vera may also cause an increase in white blood cells, which can lead to severe itching. In some cases, the extra cells collect in the spleen, which may then become enlarged (splenomegaly). Polycythaemia vera can also lead to other problems such as scarring of the bone marrow (myelofibrosis) and acute myeloid leukaemia. The patient experts highlighted that polycythaemia vera is a debilitating illness that significantly affects people living with the condition, and their families and carers. They said that the symptoms that affect people the most are severe fatigue, bone pain, itching and having an enlarged spleen. They also noted how highly disruptive frequent venesections are (when blood is removed from a person to reduce excess red blood cells). The patient experts also highlighted the extra psychological burden of being diagnosed with a rare condition. People with polycythaemia vera explained how the condition can worsen very quickly because they can feel good, but the next day they are in considerable pain and have to rest. They emphasised the significant disruption this has on their lives, and on families and carers. The patient experts also noted how 25% of people surveyed by MPN Voice and Leukaemia Care reported stopping first-line treatments because of side effects or declining treatment effectiveness. People with polycythaemia vera also explained how current treatment options can fail to have the desired effect and result in significant side effects. The clinical experts identified that current treatment options carry a high risk of developing leukaemia, which can be fatal within 3 to 6 months. They noted that, in people who cannot tolerate hydroxycarbamide or when their condition is resistant to it, there are few options other than busulfan. With busulfan treatment, there is a 20% risk of developing leukaemia. The clinical experts highlighted the unmet need for a treatment option that reduces symptoms and improves quality of life compared with current treatments. The committee concluded that polycythaemia vera is a debilitating

condition. It also concluded that there is high unmet need for effective treatments that improve survival and quality of life, and have manageable side effects.

Clinical management

Treatment pathway

3.2 The clinical and patient experts, and the company, identified the British Society for Haematology 2018 guidelines on treating polycythaemia vera as the most appropriate for the NHS. The guidelines recommend venesection and low-dose aspirin for everyone with polycythaemia vera. Cytoreductive therapy is recommended for people who are at high risk (65 years and over or with a history of thrombosis), have an uncontrolled haematocrit (percentage of red blood cells in the blood) or whose tolerability of venesections is poor. First-line cytoreductive therapy is hydroxycarbamide or interferon alfa. Second-line cytoreductive therapy is interferon alfa if hydroxycarbamide is used first line, or hydroxycarbamide if interferon alfa is used first line. Third-line cytoreductive therapies include anagrelide plus hydroxycarbamide, busulfan and radioactive phosphorous. Pipobroman is recommended by the British Society for Haematology for people with a limited life expectancy, but was not included in the NICE scope for this evaluation. The company explained that the clinical experts it consulted said that radioactive phosphorus is rarely used, so it was not included in the company's submission. The comparator presented in the company submission was called 'best available therapy'. It included hydroxycarbamide, interferon alfa, anagrelide and busulfan, with the use of each weighted by use in the MAJIC-PV clinical trial (see section 3.6). The EAG noted that the clinical experts it consulted agreed with the exclusion of radioactive phosphorous, and highlighted the limited use (if at all) of anagrelide and busulfan. The clinical experts also explained during the committee meeting that anagrelide, busulfan, radioactive phosphorous and pibobroman are very rarely used in clinical practice. They explained that this is because they are not licensed for, and have not been shown to be effective for treating,

polycythaemia vera. The EAG commented that the company's definition of best available therapy was appropriate, and that hydroxycarbamide and interferon alfa were the most used treatments. The committee concluded that hydroxycarbamide and interferon alfa were the most relevant treatment options for polycythaemia vera, and that the company had appropriately defined best available therapy.

Treatment positioning of ruxolitinib

3.3 The committee recalled the wording of the marketing authorisation for ruxolitinib. It noted that ruxolitinib is indicated for people with polycythaemia vera who are resistant or intolerant to hydroxycarbamide. The company explained that this meant ruxolitinib would be used as second- or third-line cytoreductive therapy depending on which line hydroxycarbamide was used and whether there was resistance or intolerance to it (see section 3.2). The clinical experts highlighted that ruxolitinib does not have to be used immediately after hydroxycarbamide. They said this is because ruxolitinib eligibility can be based on previous intolerance to hydroxycarbamide. They noted that the availability of ruxolitinib would give people another treatment option besides interferon alfa when there is resistance or intolerance to hydroxycarbamide. They explained that interferon alfa can exacerbate some of the symptoms of polycythaemia vera, such as itching. The patient experts described their first-hand experience of treatment with ruxolitinib. They noted that they had significant improvements in their condition and reduced side effects compared with hydroxycarbamide and interferon alfa. The patient and clinical experts highlighted that ruxolitinib can lead to improved control of blood cell counts and improve symptoms. For example, it can reduce fatigue, spleen size, pain and itchy skin. The clinical experts noted that ruxolitinib is already in widespread use for treating myelofibrosis. Ruxolitinib is recommended in NICE's technology appraisal guidance on ruxolitinib for treating disease-related splenomegaly or symptoms in adults with myelofibrosis. The clinical experts noted that there are potential risks with ruxolitinib, such as infections, skin cancer, weight gain, raised blood

pressure and high cholesterol levels. But they noted that these side effects can be mitigated against. The committee concluded that the company's proposed positioning of ruxolitinib in the treatment pathway was appropriate. It concluded that best available therapy, as defined in the company submission (see section 3.2), was an appropriate comparator at this point in the pathway.

Clinical effectiveness

RESPONSE and RESPONSE-2 trials

3.4 The main clinical evidence provided by the company for ruxolitinib was from the phase 3 RESPONSE and RESPONSE-2 trials. Both were multicentre open-label randomised trials funded by the company. They compared ruxolitinib with best available therapy and both trials lasted 5 years. Crossover from the best available therapy arm to the ruxolitinib arm was allowed (see section 3.5). Both trials included adults with polycythaemia vera who could not tolerate hydroxycarbamide or when the condition was resistant to it. RESPONSE included people with splenomegaly and RESPONSE-2 included people without splenomegaly. Everyone in the trials had to have an Eastern Cooperative Oncology Group performance status of 0, 1 or 2. RESPONSE recruited 222 people from 18 countries including 3 UK sites. RESPONSE-2 recruited 149 people from 12 countries not including the UK. The median age of people was about 61 years in RESPONSE and about 65 years in RESPONSE-2. The median time since diagnosis was about 8.8 years in RESPONSE and about 6.6 years in RESPONSE-2. The primary outcome of RESPONSE was primary response (controlled volume of red blood cells in the blood and a more than 35% reduction in spleen volume) at 32 weeks. This was statistically significantly improved for ruxolitinib compared with best available therapy (22.7% compared with 0.9%; $p<0.001$). The primary outcome of RESPONSE-2 was controlled volume of red blood cells in the blood at 28 weeks. This was statistically significantly improved for ruxolitinib compared with best available therapy (62.2% compared with 18.7%; $p<0.0001$). Overall survival for ruxolitinib at

5 years was 92% in RESPONSE and 96% in RESPONSE-2. Overall survival for best available therapy was not reported because crossover confounded results (see section 3.5). The committee concluded that RESPONSE and RESPONSE-2 show clinical advantages with ruxolitinib over best available therapy in controlling the volume of red blood cells in the blood and reducing spleen volume.

Crossover in RESPONSE and RESPONSE-2

3.5 Crossover was permitted in RESPONSE at 32 weeks and RESPONSE-2 at 28 weeks. Crossover from best available therapy to ruxolitinib was 88% in RESPONSE and 77% in RESPONSE-2. The company acknowledged the limitations of crossover in RESPONSE and RESPONSE-2. It explained that adjusting for crossover was not feasible because of the low number of deaths in RESPONSE and RESPONSE-2. So, it developed an indirect treatment comparison from overall survival data from RESPONSE for ruxolitinib and from real world GEMFIN registry data for best available therapy. Propensity score matching was done using individual patient level data from the respective sources. RESPONSE-2 data was not included because of considerable overlap with RESPONSE in the number of people in GEMFIN that could be matched. Using a combined population with RESPONSE and RESPONSE-2 would have resulted in a poor fit when estimating propensity scores for matching because these people could not be double counted. The indirect treatment comparison was not used to inform the company's base case. The overall survival results are academic in confidence so cannot be reported here. But the company noted that they showed statistically significantly improved survival for ruxolitinib compared with best available therapy. The company did identify limitations associated with the results because of:

- limited generalisability of the GEMFIN registry because of uncertainty about whether the Spanish population and treatments used reflect NHS clinical practice
- shorter follow-up time in GEMFIN than in RESPONSE

- the lack of inclusion of RESPONSE-2 data
- matching was only feasible for a limited number of covariates.

The EAG agreed with the limitations of the indirect treatment comparison and emphasised its limited scope because of only using RESPONSE data. It suggested that MAJIC-PV (see section 3.6) provided the best source of unconfounded evidence. The committee noted the efforts of the company to explore the effect of crossover on overall survival data in RESPONSE and RESPONSE-2. It concluded that overall survival data from the 2 trials was not suitable for decision making because of being confounded. It also concluded that the indirect treatment comparison was informative. But it did not think it was sufficient to be used in cost-effectiveness modelling as a source for overall survival data because of the limitations described by the company and because it only included RESPONSE data. The committee also agreed with the EAG that the best source of unconfounded evidence for overall survival was from MAJIC-PV (see section 3.6).

MAJIC-PV trial

3.6 Additional clinical evidence for ruxolitinib was from the phase 2 MAJIC-PV trial. This was a multicentre open-label randomised trial funded by Blood Cancer UK with an unrestricted funding grant from the company. It investigated ruxolitinib compared with best available therapy over 5 years. MAJIC-PV was a UK only trial, recruiting 190 people from 38 sites. Crossover was not specified in the trial protocol for MAJIC-PV, but the clinical experts noted that 10 people did crossover to ruxolitinib. MAJIC-PV recruited adults with high-risk polycythaemia vera who were intolerant of hydroxycarbamide or in whom the condition was resistant to it. High risk was defined as meeting at least 1 of these criteria:

- were 60 years or over
- had previously documented thrombosis deemed to be secondary to polycythaemia vera
- had significant or symptomatic splenomegaly

- had a platelet count of more than 1,000x10^9/litre
- had diabetes or hypertension needing pharmacological therapy for longer than 6 months.

The median age of people was 66 years and the median time since diagnosis was 7.6 years. The primary outcome in MAJIC-PV was complete haematological remission in year 1. This was statistically significantly improved with ruxolitinib compared with best available therapy (43% compared with 26%; p=0.02). The committee noted that the choice of a 90% level of confidence for the primary outcome was not typical. It added that a 95% level of confidence was used in RESPONSE and RESPONSE-2. There was no statistically significant difference in overall survival at 3 years (88% for ruxolitinib and 87% for best available therapy). The hazard ratio for overall survival for ruxolitinib compared with best available therapy was 0.73 (95% confidence interval [CI] 0.36 to 1.50). The committee noted that the confidence interval for overall survival was very wide. It also noted that, because the confidence interval crossed 1, it is not known whether ruxolitinib improves or worsens survival. There was also no statistically significant difference in progression-free survival (84% for ruxolitinib compared with 75% for best available therapy; hazard ratio 0.64; 95% CI 0.36 to 1.15; p=0.13). The clinical experts noted that ruxolitinib statistically significantly improved event-free survival compared with best available therapy (hazard ratio 0.58; p=0.03). Event-free survival was defined as time to first occurrence of major thrombosis, haemorrhage, disease transformation or death. The committee noted that the mean dose of ruxolitinib in MAJIC-PV was 10 mg twice daily, with dose intensity increasing over time. The clinical experts noted that some people will have an increased dose for better control of their blood counts. They were unsure why dose intensity increased over time in MAJIC-PV but considered that this was likely because the number of people in the trial reduced over time. The committee concluded that MAJIC-PV showed clinical advantages with

ruxolitinib over best available therapy in inducing haematological remission and improving event-free survival.

Effect of ruxolitinib on overall survival

3.7 The committee considered whether ruxolitinib improved overall survival compared with best available therapy. It noted that none of the 3 clinical trials showed an overall survival benefit with ruxolitinib compared with best available therapy. This was because of confounded data in RESPONSE and RESPONSE-2 (see section 3.4 and section 3.5) and lack of statistical significance in MAJIC-PV (section 3.6). The committee noted the small number of deaths in the clinical trials over their 5-year follow-up durations:

- RESPONSE: 9 deaths in 112 people having best available therapy and 10 deaths in 110 people having ruxolitinib
- RESPONSE-2: 6 deaths in 75 people having best available therapy and 3 deaths in 74 people having ruxolitinib
- MAJIC-PV: 17 deaths in 87 people having best available therapy and 15 deaths in 93 people having ruxolitinib.

It considered that the small number of events causes considerable uncertainty in the estimated hazard ratios for overall survival. The clinical experts explained that the primary benefit of ruxolitinib was to improve quality of life for people with polycythaemia vera. They highlighted that ruxolitinib showed statistically significant improved event-free survival compared with best available therapy in MAJIC-PV (see section 3.6). They explained that this means people treated with ruxolitinib have fewer events that are known to be associated with increased risk of death, such as major thromboembolic events. The clinical experts considered that it is plausible that this would lead to improved overall survival. The committee concluded that it was plausible that ruxolitinib may improve overall survival compared with best available therapy, but that this and the size of any effect was unknown.

Generalisability

3.8 The committee considered the generalisability of RESPONSE, RESPONSE-2 and MAJIC-PV to NHS clinical practice. MAJIC-PV only included people from the UK. RESPONSE included people from 3 UK sites and RESPONSE-2 did not include anyone from the UK. The EAG highlighted that the clinical experts it consulted agreed that all 3 trial populations were reflective of NHS clinical practice. But it considered that MAJIC-PV was most generalisable to the NHS because of the age of those included (see section 3.4 and section 3.6). It also noted a concern expressed by the clinical experts that the definition of hydroxycarbamide intolerance in all 3 trials may not have reflected that used in NHS clinical practice. This was because there is no standard definition. The EAG also highlighted uncertainty on how much the MAJIC-PV population represented a high-risk subgroup. This was because baseline characteristics seemed similar to the other trials, but mortality was substantially higher. The company considered that all 3 trial populations represented people who would benefit from ruxolitinib and were relevant to decision making. The clinical experts consulted by NICE thought that all 3 trials were relevant to NHS clinical practice. One clinical expert expressed a preference for MAJIC-PV because of the very specific entry criteria and crossover present in RESPONSE and RESPONSE-2 (see section 3.4, section 3.5 and section 3.6). The clinical experts explained this specific entry criteria likely meant people recruited to RESPONSE and RESPONSE-2 were generally fitter than people in MAJIC-PV and in the NHS. The clinical experts noted that most people they saw in NHS clinical practice would have been eligible for MAJIC-PV. They also considered that most people in RESPONSE and RESPONSE-2 would have been eligible for MAJIC-PV. The committee noted that MAJIC-PV was considered to enrol a broader range of people than RESPONSE and RESPONSE-2. So, it considered that evidence from MAJIC-PV was likely to be most appropriate for assessing the use of ruxolitinib within its marketing authorisation. It also considered that the population recruited in MAJIC-PV best represented the polycythaemia vera population in NHS

clinical practice. The committee also recalled its previous conclusion that MAJIC-PV was the best source of unconfounded evidence because it had limited treatment arm crossover (see section 3.5). So, it concluded that MAJIC-PV was the most appropriate source of clinical-effectiveness evidence for its decision making.

Effect of splenomegaly on treatment choice

3.9 RESPONSE included people with splenomegaly whereas RESPONSE-2 included people without splenomegaly. The company's base-case economic model included separate cost-effectiveness estimates for people with and without splenomegaly (see section 3.10). The clinical experts explained that treatments offered do not vary by whether or not splenomegaly is present. But they said that identifying splenomegaly helps clinicians adopt more targeted disease monitoring. This is because splenomegaly may increase the chance of the condition being resistant to hydroxycarbamide or people being intolerant of it. Also, it may indicate that the condition is transforming into myelofibrosis. The clinical experts added that the presence of splenomegaly is not routinely checked or measured in clinical practice. Instead, splenomegaly investigations are prompted by people reporting symptoms, but the effect on quality of life can vary significantly. The clinical experts explained that, for some people, splenomegaly means difficulty in eating, which leads to weight loss, but others have very few symptoms. One clinical expert also highlighted that subgroup results from MAJIC-PV showed no evidence of a differential benefit for ruxolitinib in people with and without splenomegaly. The committee concluded that the presence of splenomegaly was not a treatment- or outcome-altering factor, so was not a subgroup-defining characteristic for decision making.

Economic model

Model based on RESPONSE and RESPONSE-2 data

3.10 The company developed a state-transition model to model the cost effectiveness of ruxolitinib compared with best available therapy. In its

base case, the company presented cost-effectiveness results separately for people with and without splenomegaly. The baseline characteristics of people in the model were aligned with RESPONSE and RESPONSE-2, respectively. The time to treatment discontinuation and overall survival in the ruxolitinib arm for each population was informed by individual patient data from RESPONSE and RESPONSE-2. The time to treatment discontinuation and overall survival in the best available therapy arm for each population was informed by data from MAJIC-PV (see section 3.13 and section 3.15). The company also developed a separate model based on MAJIC-PV data for the high-risk subgroup (see section 3.11). In the RESPONSE and RESPONSE-2 model, people were modelled to enter and have treatment in either a 'ruxolitinib' or 'best available therapy' state. People in the 'ruxolitinib' state could move to the 'best available therapy' state or 'death'. People in the 'best available therapy' state could move only to 'death'. In the 'best available therapy' state, people were also separated by treatment line (first, second and beyond, or no treatment). For each treatment state, the model captured:

- treatment-related adverse events
- key complications including thromboembolic events, bleeding or haemorrhage, progression to myelofibrosis and cancer
- venesections
- health-related quality of life
- resource use.

The EAG noted the company's model was appropriate and developed with suitable methods but expressed concern with the company's model structure. It outlined that a model based on disease stages rather than treatment stages would incorporate progression outcomes that are more prognostic of long-term survival than treatments. The committee concluded that the company's RESPONSE and RESPONSE-2 model was developed appropriately but shared the EAG's concerns about a model structure based on treatment rather than disease stages.

Model based on MAJIC-PV data

3.11 The company also developed a partitioned survival model for the MAJIC-PV high-risk subgroup population. A partitioned survival model was used because of the lack of individual patient data that is needed to estimate transition probabilities for a state-transition model. The MAJIC-PV model used the same model structure and modelled treatment stages as the RESPONSE and RESPONSE-2 model (see section 3.10). The EAG preferred the MAJIC-PV model, based on generalisability to the NHS (see section 3.8). But it noted the same concerns with the treatment stage-based model structure as for the RESPONSE and RESPONSE-2 model (see section 3.10). The committee recalled that it considered that MAJIC-PV data was the most appropriate source of unconfounded clinical-effectiveness evidence for assessing the cost effectiveness of ruxolitinib for polycythaemia vera (see section 3.8). It concluded that the company's MAJIC-PV model was also developed appropriately but subject to the same model structure concerns as with the RESPONSE and RESPONSE-2 model.

Updated progression-based model

3.12 During technical engagement, the company developed an updated model structure based on stages of disease progression. The updated structure was applied to the RESPONSE, RESPONSE-2 and MAJIC-PV models. In the updated model structure, people entered the model in 'progression-free on ruxolitinib' or 'progression-free on best available therapy' health states. People in the 'progression-free on ruxolitinib' health state could then move to 'progression-free on best available therapy', 'progressed disease' or 'death' health states. People in the 'progression-free on best available therapy' health state could then move to 'progressed disease' or 'death' health states. The progressed disease state was further divided into:

- low- or intermediate 1-risk myelofibrosis
- intermediate 2- or high-risk myelofibrosis

- acute myeloid leukaemia or myelodysplastic syndrome.

The company used 5-year risk of progression-free survival events data from the trials to derive transitions between different health. It also used assumptions or external literature when data was not available from the trials. The company did not use the progression-based model in its base case. This was because that model relied on more assumptions and was associated with more uncertainty than the original model structure. It noted that the cost-effectiveness results of the progression-based model were more favourable to ruxolitinib than the original model structure, so suggested that the original model structure was conservative. The EAG noted that the progression-based model used progression-free survival, with overall survival modelled as a surrogate for disease progression or transformation. This was to capture the prognostic value of progression on survival. The company stated that it was not possible to construct a model based on event-free survival. This was because of the lack of information in the MAJIC-PV published data on the number of event-free survival events and lack of individual patient data. The EAG preferred the updated model structure in principle because it modelled progression directly. But it did not use this model structure in its base case because it did not have sufficient opportunity to review and validate the model inputs. The EAG noted it would have greater confidence in the updated model structure if inputs, assumptions and outcomes could be validated by clinical experts or external registry data. The clinical experts highlighted that they would prefer a model based on clinical events, so favoured the updated model structure. But they also noted that treatment changes usually follow changes in clinical events anyway. The committee noted that the clinical trials did not show a statistically significant overall survival benefit with ruxolitinib compared with best available therapy. But it noted that ruxolitinib did statistically significantly improve event-free survival compared with best available therapy in MAJIC-PV (see section 3.7). It also noted that it was plausible that ruxolitinib may improve overall survival by:

- reducing the occurrence of events that are associated with an increased risk of death and
- delaying disease progression.

The committee concluded that it preferred the company's updated progression-based model structure in principle. This was because it captured the prognostic value of preventing progression on survival, rather than modelling overall survival directly. But it highlighted considerable uncertainty in the updated model structure. This was because the EAG had not had a chance to fully review the model inputs and assumptions, and outcomes had not been validated. The committee noted that, to accurately evaluate the validity of the updated progression-based model structure, it would want the EAG to provide a review of the appropriateness of the inputs and assumptions used. The committee concluded that it preferred the updated progression-based model structure in principle, but that further information would help for it to fully evaluate the appropriateness of this model. This would include:

- probabilistic results for the updated progression-based model with committee preferred assumptions (see section 3.18)
- full independent clinical assessment to assess plausibility of the model results
- validation of the model results for the relative effects on overall survival compared to MAJIC-PV results and longer-term registry data.

Long-term treatment effect on overall survival

3.13 Overall survival for best available therapy in the RESPONSE and RESPONSE-2 model was estimated by applying the overall survival hazard ratio for best available therapy compared with ruxolitinib from the MAJIC-PV trial to the estimated overall survival for ruxolitinib from RESPONSE and RESPONSE-2. This was because crossover in RESPONSE and RESPONSE-2 meant that overall survival data for ruxolitinib was confounded (see section 3.10). Overall survival for ruxolitinib in the MAJIC-PV model was estimated by applying the overall

survival hazard ratio for ruxolitinib compared with best available therapy from the MAJIC-PV trial to the estimated overall survival for best available therapy from MAJIC-PV. This was because of the lack of individual patient data for MAJIC-PV (see section 3.11). This meant that the same overall survival treatment effect was assumed across the RESPONSE, RESPONSE-2 and MAJIC-PV models. The company estimated a time-varying hazard ratio for overall survival using a piecewise Cox proportional hazards model. The company explained that it assumed a larger treatment effect after 3 years based on expert advice and visual inspection of the Kaplan–Meier curve from MAJIC-PV. It estimated an overall survival hazard ratio for ruxolitinib compared with best available therapy of 0.91 (95% CI 0.38 to 2.18) for years 0 to 3 and 0.45 (95% CI 0.13 to 1.61) from years 3 to 5. The EAG explained that the company's methods for applying the piecewise hazard ratio were appropriate and that the estimates may be clinically plausible. But it noted high uncertainty in the treatment effect because of wide confidence intervals. So, the EAG highlighted a preference for a constant hazard ratio because of less wide confidence intervals than the piecewise hazard ratio. The EAG also noted that the use of time-varying hazard ratios was more favourable to ruxolitinib. It considered that a more conservative approach was more appropriate because of the high level of uncertainty. The constant overall survival hazard ratio for ruxolitinib compared with best available therapy was estimated as 0.73 (95% CI 0.36 to 1.50). The company emphasised how challenging it was to show a survival gain in polycythaemia vera. This was because of the relatively low risk of death in polycythaemia vera, and low number of deaths in trials, which reduced statistical power. The clinical experts were presented with different modelled overall survival curves for the RESPONSE and RESPONSE-2 model during the committee meeting. The curves varied by a constant or time-varying hazard ratio and by including or excluding treatment waning (see section 3.14). The clinical experts explained that it is plausible that ruxolitinib improves overall survival compared with best available therapy (see section 3.7). The committee considered the extent to which the

company's model accurately predicted long-term treatment effects, specifically an estimated survival gain. It recalled the small number of deaths in the clinical trials and the very similar mortality rate across treatment arms (see section 3.7). It also noted that the confidence intervals surrounding both the constant and time-varying hazard ratios were very wide. So, it concluded that both the constant and time-varying hazard ratios used by the company and EAG were highly uncertain. The committee also concluded that it would like to see:

- full probabilistic sensitivity analyses results exploring the estimated hazard ratio for overall survival and
- scenario analyses results presenting more conservative assumptions for survival gain, including an overall survival hazard ratio equal to 1 in the original model structure.

Treatment waning

3.14 The company's base case assumed that the overall survival treatment effect for ruxolitinib compared with best available therapy would diminish linearly and entirely from 5 years to 20 years. The EAG also maintained treatment waning in its base case (along with a constant hazard ratio up to 5 years; see section 3.13). It noted that this was a conservative assumption and that treatment waning had a large impact on cost-effectiveness results. But it highlighted that clinical experts it consulted identified no reason to assume a loss of long-term treatment effect. The company challenged the EAG's use of a constant hazard ratio and treatment waning. It thought that this combination was overly conservative because the constant hazard ratio was already a conservative assumption. The clinical experts consulted by NICE noted it was difficult to judge treatment waning because of crossover in RESPONSE and RESPONSE-2. They recalled personal experience of polycythaemia vera responding well to ruxolitinib and noted the median age of people diagnosed is 67 years. They noted age was an important factor. This was because they expected that treatment waning over 20 years or longer

may not have a large effect because of general life expectancy coming into effect. The clinical experts also highlighted that they knew of no evidence for treatment resistance and that ruxolitinib is known to have a sustained benefit. But they did note that treatment waning may occur because of the condition changing and disease progression. The committee concluded that treatment waning should be included in overall survival modelling, even when combined with a constant hazard ratio. It noted that this was to adopt a conservative approach because of uncertainty with the overall survival treatment effect, as outlined in section 3.13.

Treatment discontinuation

3.15 The company used time to treatment discontinuation data from the MAJIC-PV trial for the best available therapy arm for all 3 trial population models. For the RESPONSE and RESPONSE-2 models, the company used time to treatment discontinuation data for ruxolitinib from the corresponding trials. For the MAJIC-PV model, the company adjusted overall survival data for ruxolitinib based on the hazard ratio for time to treatment discontinuation compared with overall survival from MAJIC-PV. This was because of the lack of patient level data for this trial (see section 3.11). The company explained that, for RESPONSE and RESPONSE-2, time to treatment discontinuation data was extrapolated using an odds spline with 1-knot distribution based on visual and statistical fit. It highlighted that the clinical experts it consulted identified that treatment would more likely be stopped early on, before the condition stabilised, which informed its preference for the chosen distribution. The EAG explained that it preferred a Weibull distribution. This was because it had the best statistical fit for RESPONSE data and a suitable fit to the RESPONSE-2 data. The company noted statistical fit was similar across other time to treatment discontinuation distributions for ruxolitinib. The committee noted that, in the company's original model structure, the choice of time to treatment discontinuation extrapolation had a limited effect on cost-effectiveness results. The committee concluded that it could

not determine a preferred extrapolation distribution because this was likely to be affected by its requests to change the model structure (see section 3.6).

Utility values

Source of utility values

3.16 The NICE reference case stipulates that EQ-5D utility values should be used in company submissions unless there is empirical evidence to deviate from this measure. Data for the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) and Myeloproliferative Neoplasm Symptom Assessment Form (MPN-SAF) measures were collected in RESPONSE. Data for EQ-5D-5L and MPN-SAF measures were collected in RESPONSE-2 and MAJIC-PV. The EQ-5D and EORTC measures are generic measures of quality of life, whereas MPN-SAF is a disease-specific measure for myeloproliferative neoplasms. The company used Myelofibrosis 8 dimensions (MF-8D; a myelofibrosis disease-specific measure) utility values in its economic model. It did this by incorporating 3 items from EORTC QLQ-30 and 5 items from MPN-SAF data from the RESPONSE trial in its base case. Only RESPONSE data was used because it was the only trial to collect EORTC QLQ-30 data. The company explained the decision based on this evidence from RESPONSE-2:

- EQ-5D has a ceiling effect. This limited the maximum score that could be recorded because a higher percentage of people reported no problems in all 5 EQ-5D measures at baseline compared with items from MPN-SAF.
- EQ-5D lacks construct validity (or how accurate it can assess its intended measure) because convergence was inconsistent across MPN-SAF domains at baseline.
- EQ-5D lacks responsiveness because medium to large changes in scores for MPN-SAF were small to very small for EQ-5D.

The company also noted that NICE's technology appraisal guidance on ruxolitinib and on fedratinib for treating disease-related splenomegaly or symptoms in adults with myelofibrosis accepted the use of MF-8D over EQ-5D. The company highlighted that the symptoms of polycythaemia vera and myelofibrosis are very similar. The EAG considered that the company did not provide sufficient evidence to reject the use of EQ-5D and used it in its base case. It noted that the test of convergent validity showed a strong correlation between EQ-5D and MPN-SAF total symptom score. This suggested that, even if some polycythaemia vera symptoms are not explicitly included in the EQ-5D descriptive system, the symptoms may still be reflected in one or more of the EQ-5D dimensions, and so in overall utility value. The EAG also noted that the estimated utility differences in the treatment arms of the clinical trials were similar whether EQ-5D or MF-8D measures were used. The clinical and patient experts explained that symptom improvements for people with polycythaemia vera are highly underestimated in EQ-5D measures. This is because key symptoms such as itching and fatigue are not well captured. This is because, in this context, itching is severe and highly debilitating. They added that EQ-5D is not validated in polycythaemia vera and MF-8D best reflects the lived experience of people with the condition. The EAG noted that itching should be captured within EQ-5D measures because it captures pain. But it suggested that there was uncertainty in how well it captures fatigue. The committee acknowledged the substantial burden on quality of life of polycythaemia vera, including the substantial burden of symptoms such as itching and fatigue. It recalled that the EAG explained about the strong correlation between EQ-5D and MF-8D scores. This suggested that the effect of symptoms on quality of life should still have been reflected in EQ-5D scores and overall utility. The committee concluded that EQ-5D was the most appropriate utility measure to use in the economic model, but that MF-8D should be used in scenario analyses.

Cost-effectiveness estimates

Committee's preferred assumptions

3.17 The committee concluded that its preferred assumptions for the cost-effectiveness modelling of ruxolitinib compared with best available therapy were:

- best available therapy as defined in the company submission was an appropriate comparator (see section 3.3).
- MAJIC-PV was the most appropriate trial for decision making for the full marketing authorisation (see section 3.8).
- the updated progression-based model structure was preferred in principle (see section 3.12), although further validation of this structure, the evidence used to parametrise the model and its final outcomes was needed (see section 3.19).
- EQ-5D was the most appropriate utility measure (see section 3.16).

The committee considered that treatment waning was appropriate, as per the company's and EAG's bases cases (see section 3.14) as a conservative assumption. But it may not be appropriate in the updated progression-based model once critiqued by the EAG. The committee considered that a preferred extrapolation distribution for time to treatment discontinuation could not be determined. This was because would likely be affected by requests to change the model structure (see section 3.15).

Uncertainty in the cost-effectiveness estimates

3.18 The committee recalled the uncertainties in the evidence base and in the company's modelling assumptions. The committee considered that there remained substantial uncertainty in the cost-effectiveness estimates generated using its preferred assumptions because of:

- uncertainty in the size of the overall survival treatment effect estimated for ruxolitinib compared with best available therapy (see section 3.13)
- uncertainty in the updated model structure because the EAG had not had chance to fully review the model inputs and assumptions, and the outcomes had not been validated (see section 3.12).

The committee considered that it would like to see the following analyses and further evidence to enable it to decide on the cost effectiveness of ruxolitinib compared with best available therapy:

- probabilistic sensitivity analyses results exploring the estimated hazard ratio for overall survival (see section 3.13)
- scenario analyses results presenting more conservative assumptions for survival gain, including an overall survival hazard ratio equal to 1 in the original model structure (see section 3.13)
- a review of the appropriateness of the inputs and assumptions used in the updated progression-based model structure by the EAG
- probabilistic results for the updated progression-based model with committee preferred assumptions (see section 3.18)
- full independent clinical assessment to assess plausibility of the results
- validation of the model results for the relative effects on overall survival compared to MAJIC-PV results and longer-term registry data.

Company and EAG cost-effectiveness estimates

3.19 The deterministic cost-effectiveness results included confidential prices for ruxolitinib and other treatments. So, the exact results cannot be reported here. The company's deterministic base-case incremental cost-effectiveness ratio (ICER) for ruxolitinib against best available therapy was considerably above the range normally considered cost effective. Also, the EAG's corresponding base-case ICER was also vastly higher than the typical cost-effectiveness threshold. The committee concluded that it could not recommend ruxolitinib for routine use. This was because the most plausible ICER was likely considerably above the range normally considered cost effective and because of the issues with the company's model and uncertainty in all the cost-effectiveness estimates.

Other factors

Equality

3.20 The committee did not identify any equality issues.

Conclusion

Recommendation

3.21 The committee recalled the high uncertainty associated with the company's model and long-term treatment-effect estimations for overall survival. It considered that more evidence was needed to generate robust cost-effectiveness estimates by assessing the appropriateness of the model structure and exploring the uncertainty surrounding the overall survival estimates. It recalled that both the EAG's and company's base cases were associated with uncertainty and that the cost-effectiveness estimates were considerably above the range normally considered a cost-effective use of NHS resources. So, it did not recommend ruxolitinib for treating polycythaemia vera in adults who cannot tolerate hydroxycarbamide or when the condition is resistant to it.

4 Evaluation committee members and NICE project team

Evaluation committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by committee D.

Committee members are asked to declare any interests in the technology being evaluated. If it is considered there is a conflict of interest, the member is excluded from participating further in that evaluation.

The minutes of each evaluation committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Chair

Dr Stephen Smith

Vice chair, technology appraisal committee D

NICE project team

Each evaluation is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the evaluation), a technical adviser and a project manager.

Owen Swales and Alice Pritchard
Technical leads

Lizzie Walker
Technical adviser

Celia Mayers
Project manager

ISBN: [to be added at publication]