NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Draft guidance consultation

Maralixibat for treating cholestatic pruritus in Alagille syndrome in people 2 months and over

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using maralixibat 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?

Issue date: July 2024

Note that this document is not NICE's final guidance on maralixibat. 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- After considering these comments, the committee will prepare the final draft guidance.
- Subject to any appeal by stakeholders, the final draft guidance may be used as the basis for NICE's guidance on using maralixibat in the NHS in England.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: 21 August 2024
- Second evaluation committee meeting: TBC
- Details of the evaluation committee are given in section 4

1 Recommendations

1.1 Maralixibat is not recommended, within its marketing authorisation, for treating cholestatic pruritus in Alagille syndrome in people 2 months and over.

1.2 This recommendation is not intended to affect treatment with maralixibat 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 healthcare professional consider it appropriate to stop. For children or young people, this decision should be made jointly by the healthcare professional, the child or young person, and their parents or carers.

Why the committee made these recommendations

Alagille syndrome is a rare genetic condition that often causes cholestatic pruritus (itching), which usually starts in early childhood. Most people with the syndrome have cholestasis. This causes a build-up of bile acids in the liver and skin, which can lead to cirrhosis of the liver and pruritus. There are no treatments that are licensed for cholestatic pruritus in Alagille syndrome available in the NHS. Standard care includes treatments for pruritus that are not licensed for use in this syndrome. Some people have a liver transplant because of constant severe pruritus. The company has positioned maralixibat as an add-on treatment to standard care.

Clinical trial evidence suggests that the pruritus of Alagille syndrome is more likely to respond to maralixibat plus standard care than to placebo plus standard care alone. But it is not clear how effective maralixibat is at reducing the pruritus in the longer term, and whether it would help people live longer compared with standard care alone. There is also no evidence on maralixibat use in adults.

There are also many uncertainties in the company's economic model, so the cost-effectiveness estimates are uncertain. Even when considering the condition's severity, and the effect of maralixibat on quality and length of life, the most likely

cost-effectiveness estimates are above the range NICE normally considers an acceptable use of NHS resources. So, maralixibat is not recommended.

2 Information about maralixibat

Marketing authorisation indication

2.1 Maralixibat (Livmarli, Mirum Pharmaceuticals) is indicated for the 'treatment of cholestatic pruritus in patients with Alagille syndrome (ALGS) 2 months of age and older'.

Dosage in the marketing authorisation

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

Price

2.3 The list price for maralixibat is £43,970 per 30 ml bottle of 9.5 mg/ml maralixibat oral solution (company submission).

2.4 The company had proposed a commercial arrangement, which was considered by committee as an exception. It was considered to be 'at risk', that is, it had not been provisionally agreed with NHS England. A final agreement would be needed for this arrangement to be the basis for a final recommendation.

3 Committee discussion

The evaluation committee considered evidence submitted by Mirum 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

Cholestatic pruritus in Alagille syndrome

3.1 Alagille syndrome is a rare condition caused by mutations in the genes encoding JAG1 or NOTCH2, which are involved in the NOTCH signalling pathway. It can affect different organs including the liver, heart, skeleton, eyes, kidneys and vascular system. It is thought to occur in 1 in 30,000 to 1 in 70,000 live births. About 85% of children with Alagille syndrome have cholestasis characterised by retained bile acids in liver cells because of impaired secretion or obstruction of bile flow. Some of these bile acids join the systemic circulation, leading to increased levels of serum bile acids (sBAs) and bilirubin, which are biomarkers of cholestasis. Cholestasis can lead to pruritus (itching) and a range of liver complications, including cirrhosis (in about 46% of people with the condition), ascites (57%) and portal hypertension (40%). Severe and debilitating pruritus affects about 74% to 88% of people with cholestasis, and commonly starts from between 6 and 14 months after birth. Other non-liver symptoms of cholestasis in Alagille syndrome include:

- hypercholesterolemia
- fat-soluble vitamin deficiency
- xanthomas (fatty deposits on the extensor surfaces)
- chronic fatigue and sleep disturbances
- neurocognitive deficits
- dental damage
- growth failure.

Complications unrelated to cholestasis include abnormalities of the heart, bony framework and eye, and renal dysplasia (kidneys not fully developed). Cardiovascular abnormalities may affect up to 94% of people with Alagille syndrome and are most common in people with cholestatic liver disease. The committee acknowledged that Alagille syndrome is a serious, rare condition that can affect a range of organs.

Impact of the condition

3.2 The patient and clinical experts explained that cholestatic pruritus is the symptom of Alagille syndrome that has the largest impact on a person with the condition's life. Unremitting and severe pruritus can have a devastating effect on all aspects of life for the person and their families. The patient experts explained that the physical distress of constant pruritus affects a person's sleep. This can often cause difficulty in concentrating, and delays in developmental and educational milestones because of tiredness, the physical need to itch and absence from school. The patient experts also explained the constant need to consider external factors that can exacerbate the pruritus such as the weather and type of clothing. Debilitating pruritus can lead to skin lesions and extensive scarring, which can have a profoundly negative impact on the person's quality of life. Also, constant scratching can be stigmatising, lead to bullying and have a social impact because of its association with having fleas and lice, and poor hygiene. This, along with the physical scarring can affect a person's self-confidence and social relationships, which can cause anxiety, and affect mental and psychological wellbeing. They explained that there can be feelings of hopelessness because of the lack of effective treatments, which can lead to suicidal ideation. The patient experts explained the wider impact of the condition on parents, carers and siblings. For example, pruritus usually starts when a person with Alagille syndrome is under 6 months, and the need for constant comfort means that parents' quality of sleep is also affected. Even though the primary carer is often the mother, the physical and mental burden of full-time caring can affect both parents. It can affect their ability to work, carry out daily activities and care for other siblings, whose mental and emotional wellbeing may also be affected. The impact on family life can be devastating, including relationship breakdown, separation and divorce, the need for psychotherapy or counselling, and an impact on finances. The committee acknowledged that cholestatic pruritus in Alagille syndrome

can have an extremely negative impact on the person with the condition, and on their family and carers.

Severity of cholestatic pruritus

3.3 The committee recalled feedback from the clinical and patient experts that highlighted the multidimensional and subjective nature of chronic pruritus (see section 3.2). It acknowledged the difficulty and complexity in assessing the severity and impact of pruritus, especially in a paediatric population. The committee noted that the eligibility criteria of the maralixibat clinical trials included a clinical outcome measure for the severity of pruritus. This was the Itch Reported Outcome (ItchRO), Patient (Pt) or Observer (Obs); see section 3.6). It noted that some of the trials also reported outcomes on pruritus severity using the Clinician Scratch Scale (CSS; see section 3.8). The ItchRO scores pruritus severity on a 5-point Likert scale from 0 to 4 (where 0 represents no itch and 4 represents the most severe itch). Similarly, the CSS use a 5-point Likert scale (where 0 represents no visible skin damage and 4 represents cutaneous mutilation, haemorrhage and scarring). The clinical expert explained that these scales are rarely used in clinical practice. Rather, healthcare professionals discuss the impact of the pruritus on the quality of life of the child and family, particularly on sleep quality and the ability to function during the day. They emphasised that the impact on quality of life was the main factor guiding management. The committee noted that the maralixibat trials had all reported health-related quality-of-life data using the Paediatric Quality of Life Scale (PedsQL). It acknowledged the impact that chronic pruritus can have on secondary conditions such as sleep disorders, anxiety, depression and quality of life. It concluded that the clinical assessment of pruritus severity should consider the intensity and the impact on health-related quality of life using validated tools.

Clinical management

Treatment pathway

3.4 There are no NICE guidelines or NICE technology appraisals guidance on managing cholestatic pruritus in Alagille syndrome. The current clinical management without maralixibat involves using off-label treatments, including:

- bile acid binding resins (such as colestyramine)
- ursodeoxycholic acid (a synthetic bile acid)
- rifampicin that makes bile acids less pruritic and more excretable by kidneys
- opioid antagonists (such as naltrexone).

Adjunctive therapies include antihistamines and sertraline. Surgical options include partial external biliary diversion and liver transplants. The clinical expert explained that these treatments are used to varying degrees for managing cholestatic pruritus in Alagille syndrome in the NHS. But colestyramine is rarely used because it is ineffective, unpleasant to take and has serious side effects because it binds to fat-soluble vitamins. They also explained that surgical biliary diversion is not used in the NHS. This is because there is a less than 50% chance of a good response and there are serious related complications such as the need for an abdominal stoma. One patient expert explained that their 2-year-old is on all of the off-label treatments, but these do not seem to have any real effect. They explained that feedback from other carers suggests that current treatments are ineffective at controlling the pruritus, and that the condition is very difficult to manage. The clinical expert concurred, citing historical data from a cohort of people with Alagille syndrome on standard care. This showed that more than 40% of children had liver transplants, and the main indication was for pruritus. They explained that this data suggests that standard care may likely not be effective for most people with cholestatic pruritus. They also highlighted

the long waiting list for liver transplants and that some people may die while waiting for an organ. It was noted that significant cardiovascular disease could also prevent people from having liver transplants. One patient expert highlighted that their treatment options were limited to rifampicin because of cardiovascular complications, which are common in people with Alagille syndrome (see section 3.1). The committee acknowledged that there were no licensed treatments for cholestatic pruritus for Alagille syndrome available on the NHS, and there is a high unmet need. It concluded that people with the condition and their families would welcome safe and effective treatments for cholestatic pruritus that could potentially delay or remove the need for liver transplants.

Positioning of maralixibat

3.5 For this evaluation, the company positioned maralixibat as an add-on treatment to standard care for the treatment of cholestatic pruritus in people with Alagille syndrome 2 months and over. The clinical expert explained that, in clinical practice, it is likely that maralixibat would be offered after people have had standard care. The committee agreed with the company's positioning of maralixibat as an add-on treatment to standard care. It concluded that the company's choice of comparator was appropriate.

Clinical evidence

Key clinical evidence

3.6 The key clinical-effectiveness evidence came from 3 phase 2 multicentre double-blind randomised controlled trials with open-label extensions. These compared maralixibat plus standard care (maralixibat from now) with placebo plus standard care (placebo from now) in children aged 1 year to 18 years. They were all diagnosed with Alagille syndrome and moderate to severe pruritus defined as an ItchRO(Obs) score of at least 2 for 2 consecutive weeks. The trials were:

- ICONIC: this was considered by the company to be the pivotal trial, and included 31 people from the UK, Europe and Australia. It comprised an 18-week open-label single-arm phase in which people had up to 380 micrograms/kg/day of maralixibat. This is the recommended dose in the summary of product characteristics for maralixibat. This was followed by a randomised withdrawal phase in which people were randomly allocated to either continue with maralixibat or to have placebo for 4 weeks. People were followed up for up to 204 weeks. The primary endpoint was the mean change from baseline in fasting sBA after 4 weeks of treatment in people whose condition had responded to treatment. Response was defined as at least a 50% decrease in sBA from baseline.
- ITCH: this included 37 people from Canada and the USA. People were randomised to have 13 weeks (5 weeks of dose escalation and 8 weeks on a stable dose) of maralixibat up to 280 micrograms/kg/day or placebo. They were followed up for up to 220 weeks in the extension study, IMAGINE 2. For ITCH, the primary endpoint was the mean change from baseline at 13 weeks in the ItchRO(Obs) score.
- IMAGO: this included 20 people from the UK. People were randomised to 13 weeks (5 weeks of dose escalation and 8 weeks on a stable dose) of maralixibat up to 280 micrograms/kg/day or placebo. They were followed up for 156 weeks or more in the extension study, IMAGINE. For IMAGO, the primary endpoint was the mean change from baseline at 13 weeks in fasting sBA levels.

To compare the longer-term effects of maralixibat with standard care, the company used data from everyone having maralixibat in these 3 trials ('aggregated maralixibat cohort'). In addition, it used data from an international registry of a natural history cohort of people with Alagille syndrome, the GALA (Global ALagille Alliance) cohort. This comprised 469 people from North America, Australia and Europe. They were matched in terms of baseline characteristics to the aggregated maralixibat cohort for age (1 year to 18 years), bilirubin, alanine aminotransferase and

gamma-glutamyl transferase levels. There were no criteria for pruritus severity or baseline sBA levels available in the GALA cohort. The primary outcome was event-free survival, a composite endpoint that included liver transplants, surgical biliary diversion, liver decompensation and death. The company highlighted that the aggregated maralixibat cohort included everyone in the trials who had maralixibat regardless of treatment response.

Generalisability of the maralixibat trials to the NHS

3.7 The EAG highlighted the differences in characteristics between the maralixibat trials and its marketing authorisation with respect to age, definition of pruritus and maralixibat doses. It highlighted that the company had not provided clinical evidence for adults, and had not investigated the differential impact of using licensed and unlicensed doses in the maralixibat trials. It also noted that the placebo arms excluded colestyramine, which the EAG thought is commonly used in clinical practice. The clinical expert explained that colestyramine was excluded from the maralixibat trials because of the potential confounding effect it would have on the outcomes. The committee noted the clinical expert's differing view to the EAG's, that colestyramine is rarely used in the NHS (see section 3.4). It was unclear whether maralixibat would be used as an add-on treatment to colestyramine (see section 3.5). The committee noted that if it is, this would impact the generalisability of the trials to NHS practice. The committee acknowledged the differences in the baseline characteristics of the people included in the maralixibat trials and its broader marketing authorisation. But it acknowledged the rarity of Alagille syndrome. It concluded that the trial populations are likely representative of people in the NHS who would have maralixibat.

Clinical effectiveness

Assessment of treatment response

3.8 In its economic model, the company defined treatment response using the surrogate biomarker sBA in its base case rather than using clinical outcomes. Treatment response was defined as a 50% decrease in sBA levels from baseline at 12 weeks. The committee noted that the primary endpoint for clinical effectiveness in ICONIC and IMAGO was the absolute mean change from baseline in sBA levels (see section 3.7). It was not a dichotomous assessment of treatment response using the 50% cut off of sBA levels. The clinical expert explained that sBA is not a validated surrogate endpoint but is an important biomarker for treatment response in cholestatic pruritus. They confirmed that pruritus does not occur when sBA levels are within the normal or near-normal range. They explained that cholestatic pruritus stems from retained bile acids, and that a reduction in sBA level suggests a physiological reaction in response to treatment. The clinical expert thought that a 50% decrease in sBA from baseline indicates a large physiological change. They added that a smaller percentage decrease in sBA levels would also indicate a response to treatment. Also, in clinical practice, treatment response is usually assessed based on descriptive feedback of symptoms and their impact from the person with Alagille syndrome, or their family and carers. But the clinical expert explained that sBA levels may be used to confirm that a physiological response had occurred. The committee noted that the maralixibat trials also reported treatment response based on:

- a 70% threshold for sBA levels
- the clinical outcomes of ItchRO(Obs) using thresholds of -1, -1.5 and -2 change from baseline, and
- CSS using thresholds of -1 and -2 change from baseline.

The committee thought that the surrogate biomarker sBA and a threshold of 50% at 12 weeks likely indicates an adequate level of treatment response. But it would have preferred to have seen evidence of:

- the correlation between sBA levels and the clinical outcomes used in the trials, including assessment of health-related quality of life
- the level of change considered to be clinically meaningful in these clinical outcomes to inform decision making on the most appropriate criteria to assess treatment response.

Clinical-effectiveness results

3.9 In its pivotal trial, ICONIC, the company reported statistically significant reductions in sBA levels and ItchRO(Obs) scores after 4 weeks of treatment with maralixibat compared with placebo. When using different definitions of treatment response in ITCH (see section 3.8), there was a treatment response in more people having 13 weeks of maralixibat (below the recommended dose; see section 3.6) than having placebo. The company consider the absolute figures to be confidential, so they cannot be reported. The EAG thought that the maralixibat trials were exploratory, rather than confirmatory because:

- they were limited in terms of the short randomised periods (4 or 13 weeks)
- there was no wash-out period in ICONIC following initial maralixibat treatment in both arms
- of the lack of long-term comparative efficacy data because of the open-label single-arm extension follow-up phases.

It also noted that the statistical analyses were not adjusted for multiple testing, so they were subject to inflated type 1 error. In the GALA cohort comparison study, the company reported that, compared with people having standard care, people having maralixibat were less likely to have liver transplants, surgical biliary diversion and liver decompensation and to die. This was based on the composite-endpoint event-free survival (adjusted hazard ratio [HR] 0.305, 95% confidence interval [CI] 0.189 to 0.491; see section 3.6). The EAG thought that event-free survival may not be an appropriate proxy for mortality because people after a liver

transplant may be expected to have a near-normal life expectancy. It noted that the company had not adjusted for baseline sBA, a key prognostic factor. It also noted that results varied depending on the starting point for baseline (HR 0.305 when using the time maralixibat cohort entered study compared with HR 0.618 when using date of birth compared with HR 0.504 when using first eligible visits). It highlighted that the maralixibat doses used in the extension phases were much higher than the recommended dose, and that the 6-year data for maralixibat was immature. The committee acknowledged the limitations of the data in this rare condition, including the uncertainty in the long-term effectiveness of maralixibat and the lack of data in adults. It concluded that these areas of uncertainty would be taken into account in its decision making.

Economic model

Company model

3.10 In its submission, the company made the case that maralixibat plus standard care compared with standard care alone prolongs life and improves health-related quality of life. It assumed that, because treatment reduces the level of bile acids retained in the liver, it:

- slows down the progression of Alagille syndrome to severe stages of liver disease (cirrhosis, portal hypertension and ascites)
- delays the need for surgery (surgical biliary diversion or liver transplant).

The company's base case comprised a Markov model with 8 health states to estimate the cost effectiveness of maralixibat plus standard care compared with standard care alone for people with cholestatic pruritus in Alagille syndrome. The health states were:

- cholestasis and pruritus with response to treatment
- cholestasis and pruritus with loss of response to treatment
- cirrhosis

- portal hypertension
- ascites
- liver transplant (without cardiac or renal involvement, with a proportion of people remaining in the cycle to capture retransplant)
- after a liver transplant
- death (absorbing state).

Progression in the model depended on treatment response, defined as a 50% decrease in sBA from baseline at 12 weeks (see section 3.8). Once in the treatment non-response health state, people progressed through the different stages of liver disease until liver transplant was indicated based on transition probabilities. Some people may need a retransplant. The company assumed this occurred once in the cycle after the initial liver transplant. The company assumed that 30% of people with severe cardiac or renal complications would not be eligible for liver transplants. It included a cycle length of 12 weeks to align with the stopping guidance for treatment non-response in the summary of product characteristics for maralixibat. The company did a scenario analysis that included 'surgical biliary diversion' health states. The committee recalled the clinical expert's feedback that surgical biliary diversions are not often used in the NHS (see section 3.4). It concluded that the company's model structure for its base case was suitable for decision making.

Modelling treatment response in maralixibat and standard care arms

3.11 In its base case, the company used sBA data from the initial 12-week single-arm open-label prerandomised phase of ICONIC to inform the treatment response of the maralixibat arm. It assumed that, in the standard care arm, there would be no response to treatment in anyone. The EAG highlighted that data from the randomised phase of ITCH showed that Alagille syndrome would respond to treatment in some people in the placebo plus standard care arm using the 50% sBA response threshold applied in the company's base case. The clinical expert noted the small North American sample size and the uncertainty in

the data from ITCH. But they thought that there would be some people in NHS clinical practice whose condition would likely respond to standard care. The committee acknowledged the methodological limitations, including the small population size in ICONIC and ITCH because of:

- the rarity of the condition (see section 3.6)
- the potential for a placebo effect
- the variability of the condition and regression to the mean.

But it recalled the clinical expert's testimony that sBA is a useful surrogate biomarker because it can confirm a physiological response to treatment (see section 3.8). So, the committee thought that the company's assumption that the condition would not respond to standard care in anyone did not likely reflect clinical practice. But it noted the company's positioning of maralixibat as an add-on treatment to standard care (see section 3.5). The committee also considered whether maralixibat would be added only when the range of off-label treatment options available in standard care had been exhausted. It acknowledged that there was high uncertainty around the assumptions of treatment response for maralixibat and standard care, including the data source. The committee concluded that it would have preferred to have seen scenario analyses using different outcome measures and thresholds to define treatment response for both maralixibat and standard care. These could include outcome measures and thresholds used in ITCH for sBA, ItchRO(Obs) and CSS.

Modelling overall survival in non-response health states

3.12 The company used overall-survival Kaplan–Meier data from the treatment-naive population of the GALA study to extrapolate long-term mortality risk. It did this using the log-logistic distribution in the non-response health states. It applied the mortality risk from age 2 months in line with maralixibat's marketing authorisation. The company explained that it had applied general population life expectancy such that the model predicted lower survival in people with Alagille syndrome than in the general population. The EAG thought that all the extrapolations were

unrealistic because of the immature survival data. It preferred using the exponential distribution. This was because the median survival was 77 years compared with 217 years using the company's preferred log-logistic distribution. The EAG also preferred applying the mortality risk from age 2 years, which represented the lower quartile of the GALA population age. The clinical expert confirmed that the mortality risk is greater in people with Alagille syndrome at all ages than in the age-matched general population. They explained that, typically, people with Alagille syndrome have a higher risk of dying when they are younger but the risk decreases as the person grows older. The committee acknowledged the unrealistically high median survival of 217 years using the company's preferred log-logistic distribution. It noted that background mortality for the general population was applied such that the long-term extrapolations made little difference to which distribution was chosen. It thought that the company's model best reflected the early mortality of the condition. It concluded that the company's modelling of overall survival for the non-response health states was acceptable for decision making.

Modelling overall survival in people having maralixibat

3.13 The company modelled a greater probability of surviving in people whose condition had responded to maralixibat. It did this using data for a surrogate outcome for overall survival, the composite endpoint of event-free survival used in the GALA cohort comparison study (see section 3.9). It used the hazard ratio of event-free survival data (adjusted HR 0.305; see section 3.9) to adjust the overall survival of the standard care arm to derive the overall survival of people having maralixibat whose condition had responded to treatment. The EAG thought that there was high uncertainty about the hazard ratio for event-free survival because of issues it had previously outlined (see section 3.9). It also highlighted the uncertainty in the crude estimates of death rates in the aggregated maralixibat and GALA standard care cohorts. So, in its base case, the EAG preferred to assume that there was no difference in mortality risk for people whose condition had responded or not to maralixibat (HR 1). The

company's submission stated that the long-term reduction in hazard for event-free survival was used to estimate reductions in individual components of event-free survival related to maralixibat treatment. The EAG clarified that event-free survival data from the GALA cohort comparison study was only used as a proxy for overall survival as outlined above. The committee recalled that the GALA cohort comparison study included everyone who had maralixibat regardless of response (see section 3.6). This may have led to an underestimate of the potential benefit of maralixibat. The company explained that it did not have sBA data for the GALA cohort, so could not dichotomise outcomes based on sBA response thresholds. The committee thought that it would have preferred to have seen the GALA cohort comparison study analysed by sBA response thresholds. But it acknowledged the limitations of the available data. The committee had concerns about the lack of clarity about how treatment response related to event-free survival and mortality in the model. The committee agreed it wanted to ensure that the model was producing credible estimates. So, it would have preferred to have seen time-to-event data reported for each event (liver transplants, surgical biliary diversion, liver decompensation and death) in the endpoint event-free survival from the GALA study for standard care compared with the predictions from the model.

Stopping rule

3.14 The summary of product characteristics for maralixibat states that 'alternative treatment should be considered in patients for whom no treatment benefit can be established following 3 months of continuous daily treatment with maralixibat'. The committee recalled that, in the company's base case, treatment response was defined as a 50% decrease from baseline in sBA levels at 12 weeks (see section 3.8). It recalled the clinical expert testimony that treatment response is normally assessed clinically based on feedback on quality of life from people and their carers. Then, sBA levels are used to confirm a physiological response (see section 3.8). The EAG highlighted that fewer people having

maralixibat met the company's treatment sBA response threshold at week 18 than week 12 in the initial prerandomised phase of ICONIC. The clinical experts explained that sBA levels may fluctuate but that any reduction in sBA level would show a physiological response to treatment (see section 3.8). So, it is unlikely that a defined sBA level would be used to indicate when to stop maralixibat in clinical practice. The committee noted that the company had not included treatment effect waning in its model. Instead, the company had modelled the probability of stopping treatment with maralixibat. It assumed this was equivalent to the proportion of people who had stopped maralixibat because of adverse events in the initial 18-week prerandomised phase of ICONIC. The clinical expert explained that many people with Alagille syndrome and their families also withdraw consent to continue treatment because of inadequate benefit. The committee thought that clarification was needed on the stopping rule and its implementation in clinical practice.

Transition probability from non-response to liver transplant state

3.15 In its base case, the company assumed that 47% of people whose condition does not respond to treatment would have a liver transplant by age 4 years. The company explained that there is a range of estimates for the proportion of people who would be expected to have a liver transplant in the literature. It used a proportion in the model from the higher end of the range, which it said was a conservative estimate. The clinical expert thought that the figures used in the company's model were higher than what they had seen. They suggested that about 25% of people would likely progress to having a liver transplant by age 4 years and 40% by age 11 years. The clinical expert highlighted that there is a shortage of donor organs and that children may die while on the waiting list for a liver transplant. The committee concluded that a transition probability of 25% of people whose condition does not respond to treatment would have a liver transplant by age 4 years should be used in the economic model.

Health-related quality of life

Health-state utility values

3.16 To determine patient and carer health-state utilities in Alagille syndrome, the company applied utilities from its vignette study that used:

- EQ-5D-5L
- time trade-off
- visual analogue scale valuation methods.

This study included 200 people from the UK general population. There were vignettes for 4 patient health states (progressive cholestasis, non-progressive cholestasis, successful liver transplant, chronic liver transplant rejection) and 3 carer health states. These were developed from a systematic literature review, clinical trial data, and healthcare professional and carer interviews. The company explained that the vignettes were derived from experts and then validated by experts and people with the condition and carers. In its base case, the company used utilities derived from the EQ-5D-5L. It considers the utilities to be confidential, so they cannot be reported. It explained that the vignette study approach was selected because PedsQL data from ICONIC could not be mapped to EQ-5D. This is because many people in ICONIC were not in school. This meant that not all the domains in the PedsQL were collected, so made mapping unfeasible. The EAG thought that the health states used in the company's vignette study did not directly reflect those used in economic model. It thought that the utilities lacked external validity. So, in its base case, the EAG preferred to use the values from Kamath et al. (2015) in NICE's highly specialised technology guidance on odevixibat for treating progressive familial intrahepatic cholestasis. Kamath et al. did a mapping study in a mixed population. This included 70 children with Alagille syndrome (aged 5 years to 18 years) and parent proxy PedsQL scores on 98 children with Alagille syndrome (aged 2 years to 18 years). The committee noted that the utilities used by the company

in the non-response and liver-related complications health states were much lower than those from Kamath et al. It recalled the feedback from the patient and clinical experts (see section 3.2) and thought that the utilities from Kamath et al. were likely too high. The committee had concerns about the methodology used in the company's vignette study because no quantitative data seemingly informed them. It would have preferred to have seen further detailed information about the methods of validation and methodology used in the time trade-off study, including full descriptions of the health states. The committee concluded that it would use both the EAG's and company's utilities in its decision making.

Carer disutility

3.17 In its base case, the company derived carer disutility using its vignette study. Carer disutility was applied in the treatment response health state and all the non-response health states. The company considers the utilities to be confidential, so they cannot be reported. Carer disutility was applied to 1.7 carers per person with Alagille syndrome in all health states until age 18 years. The company also commissioned an online survey of the impact of caring for 105 carers of people with Alagille syndrome and cholestasis. It then compared the data with 95 UK-matched parents from the general population. The committee noted that the mean EQ-5D score for carers (0.82) was similar to the UK-matched sample (0.79). The EAG thought that the utility values lacked external validity. This was because the disutility for the treatment non-response health state was 6 times that of the disutility in the treatment response health state. In its base case, the EAG preferred to assume that there was no carer disutility. But it provided a scenario analysis using assumptions on carer disutility from NICE's highly specialised technology guidance on odevixibat. This was a disutility of 0.05 for treatment response health state and 0.1 for treatment non-response health states applied to 1 carer per person with Alagille syndrome until age 18 years. The committee thought that the disutility for carers used by the company was much higher than disutilities used in other NICE technology appraisals and NICE highly specialised technology

evaluations. It had concerns about the methodology used in the vignette study (see section 3.16), and was unclear about how the carer disutility was applied in the model. The committee recalled the patient expert feedback on the impact of caring for a person with Alagille syndrome and cholestatic pruritus (see section 3.2), and that both parents are affected. It concluded that it was appropriate to model carer disutility. But it thought that the likely estimates were between the company's base case and EAG's scenario analysis using assumptions from NICE's highly specialised technology guidance on odevixibat.

Costs

Bodyweight assumptions that affect maralixibat cost

3.18 In its submission, the company highlighted that people with Alagille syndrome are expected to be shorter and weigh less than the general population. This is because of difficulty digesting fats and absorbing fat-soluble vitamins that leads to malnutrition. In its base case, the company used the fifth percentile bodyweight band from the general population, in line with data observed from ICONIC. Normal distributions based on bodyweight and standard deviation in each cycle were used to estimate the average cohort doses for maralixibat at each timepoint. The EAG thought that the impact of effective treatment in boosting growth should be considered. So, it provided a scenario analysis using the bodyweight band of the twenty-fifth percentile of the general population after 2 years in the model. The company also provided a scenario analysis using the bodyweight band of the seventy-fifth percentile of the general population from the start of the model. The committee thought that it had not seen evidence of the natural growth history of people with Alagille syndrome and cholestatic pruritus, and the impact of effective treatment. It concluded that this was an area of uncertainty and that it would consider a range of bodyweight bands in its decision making.

Severity

Severity of the condition

3.19 The committee considered the severity of the condition (the future health lost by people living with the condition and having standard care in the NHS). The committee may apply a greater weight to quality-adjusted life years (QALYs; a severity modifier) if technologies are indicated for conditions with a high degree of severity. The company provided absolute and proportional QALY shortfall estimates in line with NICE's health technology evaluations manual. The company consider the QALY shortfall estimates to be confidential, so they cannot be reported. Using the company's utilities, the committee thought that the severity weight of 1.2 applied to the QALYs was appropriate. But the committee noted that the estimates of the QALY shortfall depended on whether the utilities were derived from the company-preferred vignette study or the EAG-preferred Kamath et al. (2015) mapping study. It concluded that it would consider both cases in its decision making.

Cost-effectiveness estimates

Areas needing clarification

3.20 The committee thought that neither the company's nor EAG's base cases included all its preferred assumptions. It thought that there were many areas of uncertainty (see section 3.21). It also noted that it had only preferred the company's modelling of overall survival for the non-response health states which used the log-logistic distribution applied from age 2 months (see section 3.12). In addition to the areas of uncertainty highlighted in section 3.21, the committee said that it would like clarification on:

- the criteria that would be used to start maralixibat in NHS clinical practice, for example, using ItchRO(Obs) as in the maralixibat trials or a health-related quality-of-life scale or both (see section 3.3)

- using bile acid binding resins such as colestyramine in the NHS (see section 3.4 and section 3.7)
- the likely use of maralixibat as an add-on treatment to standard care in NHS clinical practice (see section 3.5)
- the evidence of the correlation between sBA levels and the clinical outcomes used in the maralixibat trials, including assessment of health-related quality of life (see section 3.8)
- the evidence of the level of change considered to be clinically meaningful in the clinical outcomes (see section 3.8)
- the details of the stopping rule and how it will be implemented in NHS clinical practice (see section 3.14)
- the methodology used in the company's vignette studies for the health-state utility values and carer disutility (see section 3.16)
- detailed information about the methods of validation and the methods used in the time trade-off study including full descriptions of the health states (see section 3.16 and section 3.17)
- the approach of applying carer disutility in the model (see section 3.17).

Company and EAG cost-effectiveness estimates

Acceptable ICER

3.21 NICE's manual on health technology evaluations notes that, above a most plausible incremental cost-effectiveness ratio (ICER) of £20,000 per QALY gained, judgements about the acceptability of a technology as an effective use of NHS resources will take into account the degree of certainty around the ICER. The committee will be more cautious about recommending a technology if it is less certain about the ICERs presented. But it will also take into account other aspects including uncaptured health benefits. The committee noted the high level of uncertainty, specifically related to the:

- definition of severity of cholestatic pruritus and eligible population in clinical practice (see section 3.3)

- positioning of maralixibat as an add-on treatment to standard care after standard care options had been exhausted (see section 3.5)
- limited evidence on the longer-term comparative effectiveness of maralixibat and lack of data in adults (see section 3.6 and section 3.9)
- treatments used in standard care, in particular, colestyramine (see section 3.7)
- assessment of treatment response (see section 3.8)
- assumptions of the treatment response in maralixibat and standard care in the model (see section 3.9)
- modelling of overall survival in people whose condition responds to maralixibat (see section 3.13)
- the stopping rule for maralixibat and its implementation in clinical practice (see section 3.14)
- the transition probability used for moving from the treatment non-response health state to the liver transplant health state (see section 3.15)
- methods used to derive the health-state utility values, and their plausibility and impact on the severity modifier (see section 3.16 and section 3.19)
- methods used to derive carer disutility, associated assumptions and implementation in the model, and plausible values (see section 3.17)
- assumption of bodyweight of population throughout the time horizon (see section 3.18).

The committee considered:

- the rarity of Alagille syndrome and uncertainty of the evidence (see section 3.1 and section 3.9)
- the high unmet need because of a lack of licensed treatment options (see section 3.4)
- the uncaptured benefits on siblings in the model (see section 3.2 and section 3.24).

It noted that NICE's manual on health technology evaluations states that, in some cases, evidence generation is particularly difficult. This includes rare diseases, treatments for a population that mainly comprises children, and innovative and complex technologies. In these cases, the committee may be able to make recommendations accepting a higher degree of uncertainty. So, taking into account these factors and the high level of uncertainty, the committee concluded that an acceptable ICER would be around £30,000 per QALY gained.

Company and EAG cost-effectiveness estimates

3.22 The committee considered the cost effectiveness of maralixibat plus standard care compared with standard care alone. It noted that none of the company's and EAG's base cases and scenarios provided its preferred ICER (see section 3.20). The committee thought that the range of ICERs addressing the areas of uncertainty listed in section 3.21 would likely be above the range that NICE normally considers an acceptable use of NHS resources. The company consider the ICERs to be confidential, so they cannot be reported here.

Other factors

Equality

3.23 The company suggested that most carers are mothers who may experience a greater carer burden compared with the other parent. The committee recalled feedback from the patient experts that the primary carer is often the mother, but that both parents are affected by the condition (see section 3.2). The committee concluded that this issue could not be addressed within a technology appraisal.

Uncaptured benefits

3.24 The committee considered whether maralixibat was innovative. It noted that maralixibat is the first UK licensed treatment option for cholestatic pruritus in Alagille syndrome, a rare multi-organ condition. Maralixibat represents a 'step-change' in a treatment pathway addressing a high

unmet need for people with no licensed options available on the NHS. The committee noted that it had identified additional benefits of maralixibat not captured in the economic modelling. But it had already considered this in its decision making (see section 3.21).

Conclusion

Recommendation

3.25 The committee concluded that the range of cost-effectiveness estimates for maralixibat plus standard care compared with standard care alone are likely to be above the range that NICE normally considers an acceptable use of NHS resources. So, maralixibat is not recommended for treating cholestatic pruritus in Alagille syndrome.

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 A.

Committee members are asked to declare any interests in maralixibat 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

James Fotheringham

Vice-Chair, technology appraisal committee A

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.

Sharlene Ting

Technical lead

Albany Chandler

Technical adviser

Jennifer Upton and Thomas Feist

Project managers

ISBN: [to be added at publication]