Single Technology Appraisal

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Contents:

The following documents are made available to stakeholders:

1. **Draft Guidance Document (DG)** as issued to consultees and commentators

2. **Comments on the Draft Guidance from** Gilead

3. **Consultee and commentator comments on the Draft Guidance** from:
   a. PBC Foundation
   b. British Association for the Study of the Liver
   c. British Hepatology Pharmacist Group
   d. Ipsen

4. **NHS England comments on the inclusion of fibrates as a comparator**

5. **Committee member comments on company response to the DG**

6. **External Assessment Group critique of company response to the DG**

7. **Post-ACM2 submission from Gilead**

8. **EAG critique of company post-ACM2 submission**

9. **EAG additional note on EAG suggested base case**

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

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Draft guidance consultation

Seladelpar for previously treated primary biliary cholangitis

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using seladelpar 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 2025

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 seladelpar 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: 20 August 2025
- Second evaluation committee meeting: 10 September 2025
- Details of membership of the evaluation committee are given in section 4

1 Recommendations

1.1 Seladelpar should not be used to treat primary biliary cholangitis, including pruritus, in adults:

- with ursodeoxycholic acid (UDCA), if the primary biliary cholangitis has not responded well enough to UDCA, or
- alone, if UDCA cannot be tolerated.

1.2 This recommendation is not intended to affect treatment with seladelpar 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.

What this means in practice

Seladelpar is not required to be funded and should not be used routinely in the NHS in England for the condition and population in the recommendations.

This is because the available evidence does not suggest that seladelpar offers value for money in this population.

Why the committee made these recommendations

Usual treatment in NHS practice for primary biliary cholangitis is UDCA. If UDCA does not work well enough, licensed add-on treatments include obeticholic acid or elafibranor. People who cannot tolerate UDCA may also have obeticholic acid or elafibranor. But there are uncertainties about the treatment pathway.

Clinical trial evidence shows that seladelpar reduces liver enzymes, which are raised in primary biliary cholangitis, more than placebo. This suggests that seladelpar could

delay the condition getting worse. Clinical trial evidence also shows that seladelpar reduces pruritus (itch) compared with placebo.

Seladelpar has not been compared in a clinical trial with obeticholic acid or elafibranor. The results of an indirect comparison are highly uncertain. It is unclear whether seladelpar reduces liver enzymes compared with obeticholic acid or elafibranor. But there may be a reduced itch with seladelpar compared with obeticholic acid.

It is not clear if the company's economic model fully reflects the effect of itching and other aspects of primary biliary cholangitis on quality of life. The cost-effectiveness estimates are uncertain and are above the range that NICE considers an acceptable use of NHS resources. So, seladelpar should not be used.

2 Information about seladelpar

Marketing authorisation indication

2.1 Seladelpar (Livdelzi, Gilead) is indicated for 'the treatment of primary biliary cholangitis (PBC), including pruritus, in adults in combination with ursodeoxycholic acid (UDCA) who have an inadequate response to UDCA alone, or as monotherapy in those unable to tolerate UDCA'.

Dosage in the marketing authorisation

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

Price

2.3 The list price of seladelpar is £3,155.00 per 30-tablet pack of 10 mg tablets (excluding VAT; company submission).

2.4 The company has a commercial arrangement, which would have applied if seladelpar had been recommended.

Carbon reduction plan

2.5 Information on the Carbon Reduction Plan for UK carbon emissions for Gilead will be included here when guidance is published.

3 Committee discussion

The evaluation committee considered evidence submitted by Gilead, 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

Primary biliary cholangitis

3.1 Primary biliary cholangitis (PBC) is a chronic, progressive autoimmune condition that leads to a build-up of bile in the liver. This happens because the body's immune system destroys bile ducts in the liver, causing cholestasis. This means that the flow of bile through the liver and biliary system is impaired or stalled. Over time, chronic cholestasis leads to scarring of the liver (fibrosis and cirrhosis) and liver failure, and can ultimately lead to death. The cause of PBC is not known, but it is thought to be a mix of environmental and genetic factors. PBC is typically diagnosed by testing for biochemical indicators of liver function (such as alkaline phosphatase [ALP]). Many people do not have symptoms until they have significant liver damage. Common symptoms include itchy skin (pruritus) and fatigue. Around 20,000 people in the UK have PBC, with an annual incidence of 2 to 3 per 100,000. PBC is more common in women (90%) and in people aged over 40 (75%). The patient group submissions described the challenges of living with PBC, such as severe fatigue and severe itching. The patient expert elaborated that there is both the physical fatigue felt in the body, affecting movement and the ability to do daily activities, and the mental and physical exhaustion that comes with itch. The patient expert and patient group submissions emphasised that the chronic symptoms greatly affect people with PBC and their families and carers. They said that symptoms affect quality of life, sleep and the

ability to work or manage daily activities. The patient experts reported that people with PBC often experience a delayed diagnosis, which can lead to feelings of isolation, confusion and frustration because of unexplained symptoms. While having a diagnosis can bring relief, the rarity of the condition means that many people have never heard of it, and this can reinforce feelings of isolation. PBC can progress unpredictably to cirrhosis or liver cancer, with some people eventually needing a liver transplant. This adds to feelings of anxiety and uncertainty. The committee concluded that PBC has a substantial effect on people's lives.

Clinical management

Treatment pathway and positioning of seladelpar

3.2 There are no NICE guidelines specifically for the treatment of PBC. The most relevant available guidelines were the British Society of Gastroenterology/UK-PBC primary biliary cholangitis treatment and management guidelines. These were developed before the appraisal of newer treatments such as NICE's technology appraisal guidance on elfibranor for previously treated primary biliary cholangitis (from now, TA1016). The clinical expert explained that first-line treatment for PBC is UDCA. People whose PBC has an inadequate response to UDCA (as defined as ALP above 1.67 times the upper level of normal [ULN]) have obeticholic acid (OCA), elafibranor or fibrates (off-label treatment), with or without UDCA as a second-line treatment. People who cannot tolerate UDCA have OCA, elafibranor or fibrates. OCA is recommended in NICE's technology appraisal guidance on OCA for treating primary biliary cholangitis (from now, TA443) and elafibranor is recommended in TA1016. The care pathway is structured in a way that to start licensed second-line treatments, people must first be seen by a multidisciplinary team. Patient organisations explained that there are frustrations with variation in care and difficulties accessing specialist teams, especially for the approximately 40% of people who need second-line treatment after UDCA. Access to these treatments can vary, and some, like OCA, have

more side effects. Treatments for symptoms, such as colestyramine for itching, can also be difficult to take. The company proposed that seladelpar should be primarily positioned as a second-line treatment for people who have an inadequate response to, or cannot tolerate, UDCA. It said that it may also be considered a third-line option for people who cannot tolerate or whose condition has not responded to OCA. The clinical expert noted that seladelpar and elafibranor have similar mechanisms of action. The reported benefit of seladelpar in normalising ALP levels meant that the clinical expert would be keen to use it as early in the treatment pathway as possible. The clinical expert noted that people needing OCA and seladelpar may be different: people with high transaminase levels may benefit from having OCA, and people with pruritus may benefit from having seladelpar. This is because OCA has an increased risk of worsening pruritus, and seladelpar may reduce pruritus. The committee noted that the clinical effectiveness of seladelpar in later lines of treatment is uncertain because data on third-line use is limited. The committee concluded that seladelpar would be positioned as a second-line treatment and that it may also be used at third line.

Excluding fibrates as comparators

3.3 Fibrates are anti-pruritic agents, but they are not licensed for treating PBC. They were not included as comparators in the NICE scope. In TA443 and TA1016, fibrates were not considered second-line treatments. Instead, they were viewed as add-on treatments for pruritus. In TA1016, bezafibrate was included in the model only for its role in treating pruritus. It was not considered a standalone second-line treatment in that context. The EAG considered that fibrates were a relevant comparator. The EAG explained that 1 of its 3 clinical experts had stated that fibrates were their preferred second-line treatment in clinical practice. It also highlighted a UK audit that found that 50% of people whose PBC had responded inadequately to UDCA had treatment with fibrates. The EAG also noted that fibrates were considered within NHS England's medicines repurposing programme, before the programme was suspended for

reasons including structural changes at NHS England. The EAG considered that fibrates were a relevant comparator. The clinical expert at the committee meeting said that fibrates were used as a second-line treatment. He noted that around 20% to 25% of people use fibrates as a second-line treatment, including a small proportion who used them as monotherapy at this position in the treatment pathway. The clinical expert noted that fibrates would only be used in the short term and around 1 in 10 people may need to stop because of liver injury. The patient expert expressed concern that some people with PBC are unable to access specialist treatment centres where OCA is available. Instead, these people are prescribed fibrates as an alternative. The patient expert was particularly concerned about people with PBC being prescribed fibrates. They said that fibrates have limited evidence of effectiveness and had not been through the regulatory process to assess their clinical effectiveness and safety for treating PBC and associated pruritus. The clinical expert agreed that although the preferred option would be to use the licensed second-line treatments, off-label fibrates were currently being used in clinical practice. The company disagreed that fibrates were a comparator, noting the inconsistency with previous appraisals and that fibrates are not licensed for treating PBC. The committee recalled that:

- A particular benefit of seladelpar was the reduction in pruritus: the marketing authorisation for seladelpar states that it is for the treatment of PBC, including pruritus, whereas the marketing authorisations for elafibranor and OCA state that these treatments are for treating PBC.
- There were people in clinical practice having fibrates as a second-line treatment as an alternative to OCA.
- NICE methods allow consideration of off-label treatments as comparators if there is evidence of their use in clinical practice.

The committee thought that overall, fibrates were potentially a comparator for people with PBC with pruritus in NHS clinical practice. It considered that it would be informative to see a comparison of benefits

and costs of seladelpar compared with fibrates. The committee concluded that it was appropriate to re-issue the scope, listing fibrates as a potential comparator to allow exploration of this. The committee also noted that input from NHS England, patients and clinicians would help the committee to define the comparators.

Clinical effectiveness

RESPONSE data source

3.4 The main source of clinical-effectiveness evidence for seladelpar was the RESPONSE trial. This was a phase 3, randomised, double-blind, placebo-controlled study that lasted for 12 months. It evaluated seladelpar in people with PBC that had an incomplete response or who could not tolerate UDCA. A total of 193 people were enrolled in the trial. Of these, 128 had seladelpar 10 mg, with a reduced 5 mg dose used in cases of intolerance. The remaining 65 people had placebo. Both the seladelpar and the placebo arms could include UDCA use, and around 94% of people in each arm had UDCA. There were 17% of people who had previously had OCA or fibrates. The primary outcome was the proportion of people achieving a composite biochemical response at month 12. This was defined as:

- ALP less than 1.67 times the ULN
- a reduction in ALP of at least 15%, and
- total bilirubin at or below 1.0 times the ULN.

Key secondary outcomes included the proportion of people with ALP normalisation (1.0 times at or below the ULN) at 12 months. The clinical and patient experts emphasised that an ALP level above normal, even if it was less than 1.67 times the ULN, was still associated with disease progression. So, they said that ALP normalisation was an important outcome. Another key secondary outcome was change from baseline in weekly averaged pruritus numerical rating scale score. The committee

noted that although ALP levels were an important outcome in terms of disease progression, they may not always reflect the symptoms people with PBC experience. The committee concluded that trial outcomes were informative for decision making.

RESPONSE results

3.5 In the RESPONSE trial, 61.7% of people in the seladelpar arm met the composite outcome at 12 months compared with 20% in the placebo arm. This composite outcome was primarily driven by improvements in ALP-related measures. The EAG noted that baseline bilirubin levels were already low in both the treatment and placebo arms. This suggests that people in RESPONSE were probably at an earlier stage of disease, when changes in bilirubin would be less pronounced. At 12 months, 25% of people in the seladelpar arm and no people in the placebo arm had normalised ALP levels. Among people with moderate to severe pruritus at baseline, having seladelpar statistically significantly reduced the pruritus numerical rating scale score compared with placebo. There was a least-squares mean change from baseline to -3.2 with seladelpar, compared with -1.7 with placebo. The committee concluded that seladelpar was clinically effective at improving ALP levels and reducing pruritus compared with placebo.

Positive treatment response for people having placebo in seladelpar trials

3.6 In RESPONSE, 20% of people in the placebo arm had meaningful clinical improvement in ALP response, despite having no active treatment other than background treatments. The committee noted that this placebo response was high. It noted that this positive response was not explained by changes in background treatment or UDCA dosing, which remained consistent with peoples' previous use. The EAG highlighted that such placebo responses are common in PBC trials. A clinical expert suggested that improved adherence to UDCA during clinical trials may contribute to these effects. Better adherence could lead to better outcomes, even in the

absence of new treatments. The committee suggested that it was possible that the positive treatment response was because of regression to the mean. The EAG noted that this could introduce uncertainty when comparing data from RESPONSE with data from trials for other treatments, especially if placebo effects differ across trials. There may also be uncertainty about real-world adherence over time and how well trial results reflect NHS clinical practice. The committee agreed that it was plausible that the placebo effect was caused by increased adherence to UDCA. But it thought that similar adherence to UDCA would also be expected in the seladelpar arm. The committee said that the relative treatment-effect estimates were not likely to be biased by adherence to UDCA, because the same effect would be seen in both trial arms. It concluded that the reason for the observed placebo response was uncertain.

Indirect comparison approach

Bayesian NMA and MAIC

3.7 There are no head-to-head trials directly comparing seladelpar with the comparators included by the company in the final scope (OCA and elafibranor). Instead, the company relied on indirect treatment comparisons (ITCs) using data from the following trials:

- RESPONSE for comparing seladelpar plus UDCA with UDCA plus placebo
- ELATIVE for comparing elafibranor plus UDCA with UDCA plus placebo
- POISE, COBALT and NCT03633227 for comparing OCA plus UDCA with UDCA plus placebo.

The company used different methodological approaches for comparing seladelpar with each comparator. To compare seladelpar with OCA, the company used a Bayesian network meta-analysis (NMA). For the

comparison with elafibranor, it used an anchored matching-adjusted indirect comparison (MAIC). The company thought that a Bayesian NMA was unsuitable for comparing seladelpar with elafibranor. This was because of differences in baseline bilirubin and cirrhosis rates between the RESPONSE and ELATIVE trials, which the company suggested violated the transitivity assumption. That is, the trials differed in ways beyond the treatments being compared, limiting the validity of indirect comparisons. The EAG did not think that differences in bilirubin or cirrhosis rates warranted this different approach. The company noted that there were differences in the definitions of ULN for the ALP measures and the specific ULN cut-offs by sex. The company recalculated the outcomes for seladelpar to adjust for differences in the sex-specific cut-offs between RESPONSE, POISE and ELATIVE. The MAIC method adjusted individual patient data from RESPONSE to match the baseline characteristics of the ELATIVE trial population. Four treatment-effect modifiers: age, baseline ALP, bilirubin, and cirrhosis were used for matching. These modifiers were consistent with those used in TA1016 and were supported by literature and expert opinion. The EAG did not consider it appropriate to use a separate approach for each comparator. It also noted that the MAIC analysis resulted in a small effective sample size (36% of the original sample), suggesting that differences between the populations in RESPONSE and ELATIVE were difficult to reconcile with matching. The EAG suggested that a Bayesian NMA is a more appropriate method overall. The committee concluded that it was not appropriate to use a separate method for each comparison and that a Bayesian NMA should be used for all indirect comparisons.

NMA uncertainty

3.8 The committee noted the extremely large credible intervals around the relative treatment-effect estimates for seladelpar compared with elafibranor and with OCA in the ALP outcomes. This indicated that there was a large degree of uncertainty around the results. The committee

thought that without further explanation, the credible interval raised concerns about the validity of the model. The company highlighted that in the appraisal of elafibranor there were wide credible intervals for the comparison of elafibranor with OCA. The committee noted that in the results reported in the individual trials (the naive results), seladelpar had a smaller estimated treatment effect than the estimate for OCA or elafibranor. But it noted that in the Bayesian NMA it had a larger treatment effect. The committee considered that this lacked face validity. The company raised some potential differences between the trials (such as key effect modifiers, including baseline bilirubin levels and the proportion of people with cirrhosis). It did this to explain why the results from the ITC were different to the original trial results. The committee was very concerned with the analysis and did not have confidence in the results from the company's indirect comparison. It asked for:

- the company to submit its Bayesian NMA model code with scenarios presented to explore uncertainty in the model
- updated reporting of the company's Bayesian NMA, including clear reporting of the original trial values and sources of numerical data for the comparators
- an explanation of how the trials and trial populations differed, with particular reference to effect modifiers
- an explanation of any difference between the reported trial results and the relative effects resulting from the Bayesian NMA
- consideration that Turner priors are suitable for odds ratios and would need careful justification to be used for an NMA of relative risks.

Adverse events and patient-reported outcomes

3.9 For adverse events and patient-reported outcome measures (PROMs) the company used the Bayesian NMA. Although there were potential issues with the NMA methodology and its reporting, the credible intervals around the adverse events and PROM results were narrower than around the ALP outcomes. The NMA suggested that seladelpar was associated with

less pruritus at 12 months compared with both the 5mg and the 10mg dose of OCA and the placebo. Seladelpar also had lower odds of upper respiratory tract infections compared with placebo and elafibranor. In terms of patient-reported outcomes, seladelpar showed a numerical reduction in pruritus at 12 months compared with placebo and OCA. But the 95% credible interval did not include no effect on the 5-D itch scale. The EAG noted that no minimum clinically important difference had been established for either the 5-D itch or PBC-40 scales. Although the company had presented results for elafibranor PROMs, these had come from a subset of people from the ELATIVE trial. The EAG commented that it had not seen the baseline characteristics for this group. The committee concluded that the results from the NMA appeared to support clinical opinion that seladelpar improves pruritus compared with OCA. But it would welcome clarity on the NMA as outlined in section 3.8 and comments on the clinical meaningfulness of these results in response to consultation.

Economic model

Company's modelling approach

3.10 The company used a cohort-level Markov state-transition model to evaluate the cost effectiveness of seladelpar, with or without UDCA, compared with OCA or elafibranor with or without UDCA. The model had 2 components, with health states defined by ALP levels and liver-disease progression. Transitions between health states were driven by ALP levels, which were used as a proxy for increased or decreased risk of disease progression. Disease progression was defined as the transition to the compensated cirrhosis or elevated bilirubin health state. After this point, the condition can no longer improve and will continue to worsen over time. The model incorporated both costs and the impact on health-related quality of life (utilities) for people in ALP and liver-disease states. It also accounted for the burden of pruritus at varying levels of severity within these health states. The model was run over a lifetime horizon of up to 50

years to capture long-term outcomes and costs. The committee noted that the model had a similar structure to the models used in the appraisals of OCA and elafibranor. But it included an additional health state for ALP normalisation, separate from the mild ALP normalisation health state (see section 3.4). The committee agreed that this was appropriate, and the model structure was appropriate for decision making.

Treatment discontinuation

3.11 For the first 12 months, the company's base case modelled treatment discontinuation rates directly from the clinical trials for seladelpar and each comparator. These included the RESPONSE trial for seladelpar, ELATIVE for elafibranor and POISE for OCA. But the EAG preferred using ITC-derived rates, which it said better accounted for differences in trial populations and aligned with the methodology used for other model parameters. These ITC-derived rates were notably higher for comparators, especially for OCA, which had a 26.95% discontinuation rate at 12 months compared with 9.59% in the company's base case. Seladelpar's discontinuation rate remained consistent at 6.73% across both approaches. For the period after 12 months, the company applied a discontinuation rate ratio of 0.28. This rate ratio was based on the ELATIVE and ELATIVE open-label extension study data for elafibranor and this ratio was applied to estimated discontinuation rates after 12 months for seladelpar and OCA. In contrast, the EAG favoured a lower ratio of 0.12 derived from RESPONSE and ASSURE (a long-term open-label trial of seladelpar). The clinical expert said that the EAG's rates seemed plausible. The committee concluded that it preferred the EAG's approach, with the caveat that the 12-month rates were derived from the ITC, which introduced uncertainty around these estimates (see section 3.7).

Utility values

Source of utility values

3.12 To derive utility values the company used disease-specific PBC-40 data collected in RESPONSE. It mapped this to EQ-5D-5L using real-world data from 90 people included in the ITCH-E study. It then also mapped from EQ-5D-5L to EQ-5D-3L using the Hernández-Alava algorithm. The company developed a mixed model for repeated measures (MMRM) to apply a disutility for each ALP health state. The MMRM model based on RESPONSE data also estimated disutility associated with pruritus. But the company instead used disutility values associated with pruritus from Smith et al. (2022). These values were based on EQ-5D-5L data from the GLIMMER study (a trial of linerixibat in people with moderate to severe pruritus over 16 weeks). The estimated disutilities from Smith et al. were larger compared with estimates from RESPONSE. The EAG favoured the MMRM-derived disutilities, arguing that they were more appropriate because of their internal consistency, alignment with UK cohort data, and direct use of RESPONSE data. The committee preferred the EAG approach. This was because it was based on trial data from people like those who would have treatment in clinical practice and it used data from a consistent source for all health states. But it acknowledged that this approach might underestimate the burden of pruritus because the disutility values were very small. The clinical expert said that it would expect a utility loss from pruritus. The committee said that the reason utility values decreased between the ALP health states was unclear. So, it was not possible to determine whether there had been double counting of disutility caused by pruritus. The committee noted that pruritus is caused by elevated bilirubin, which was only measured in the highest ALP elevation health state. The committee noted that the model did not explicitly model other symptoms of PBC (other than pruritus) that could also affect utility values. It recalled the large impact of fatigue reported by the patient expert. The committee also noted the EAG comment that the model included the costs of treatments for pruritus including fibrates, but did not model an improvement with treatment. The committee concluded that, although it preferred the EAG approach on utility values, there remained

uncertainty about whether the disutility associated with pruritus was underestimated using this approach. The committee said that additional evidence from the literature may be of value. The committee was also uncertain whether the model reflected the expected quality of life in the ALP health states. It said that a further explanation of why, and evidence of how, quality of life would differ between these health states was needed.

Cost-effectiveness estimates

Acceptable ICER

3.13 NICE's manual on health technology evaluations notes that, above a most plausible incremental cost-effectiveness ratio (ICER) of £20,000 per quality-adjusted life year (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 around the:

- methodology and results of the ITC, including whether seladelpar has a benefit in lowering ALP levels compared with OCA or elafibranor and the estimated treatment discontinuation rates from this (see section 3.7)
- relative benefits and costs compared with fibrates, which the committee said were potential comparators (see section 3.3)
- utility values, including:
  - the disutility associated with pruritus (see section 3.12)
  - differences in utility values across ALP health states, and the reason for any differences (see section 3.12).

Because the committee had requested analyses that may help to

resolve, reduce the magnitude or clarify the extent of the uncertainty, it did not identify an acceptable ICER range. The committee concluded that it would reconsider the acceptable ICER range at the second committee meeting. This would take into account any new analyses presented.

Company and EAG cost-effectiveness estimates

3.14 The ICERs cannot be presented because the comparators have confidential patient access schemes. But the company and EAG base-case ICERs and all presented scenarios were substantially above the range normally considered a cost-effective use of NHS resources. The committee's preferred model assumptions were mostly aligned with the EAG's, and included that:

- The baseline distribution of people across the modelled health states should reflect people who would have seladelpar in clinical practice, that is, people who have ALP elevation more than 1.67 times the ULN (see section 3.4).
- A Bayesian NMA should be done and validated for the ITC for all comparators (see section 3.8.
- The treatment discontinuation rates should be derived from the ITC and assumptions on the rate of discontinuation after 12 months should be derived from data from seladelpar trials, rather than from elafibranor trials (see section 3.11).
- The same data source (RESPONSE) should be used for utility values for ALP health states and pruritus (see section 3.12).

The committee was not satisfied that it had been presented with sufficient analyses to determine its preferred ICER range (see section 3.13). Because of uncertainties in the data it requested additional analyses, which were:

- A full description of the ITC and explanation of its uncertainty, with scenario analyses when appropriate to explore the impact of uncertainty (see section 3.8).
- A distribution of change from baseline on the pruritus scale in each treatment arm. This needs to include the proportion of people achieving a clinically meaningful improvement, to better inform the utility values and QALY gains associated with pruritus improvement (see section 3.9).
- Including fibrates as a comparator (see section 3.3).
- A commentary on how well the model reflects quality of life associated with PBC, with any supportive evidence from the literature (see section 3.12).

Other factors

Equality

3.15 The company noted that people with PBC may face long wait times for care, often between 3 and 4 months. They also have higher mortality rates while on liver transplant waiting lists compared with people with other liver diseases. Stakeholders for this appraisal noted that a recent UK-wide audit highlighted geographical disparities in access to specialist teams and second-line treatments, driven by differences in local resource availability. The committee was mindful of its duties under the Equality Act 2010. It identified that previous technology appraisals (for example, TA1016) identified other factors that should also be considered, such as:

- There is a particularly high prevalence of this condition in women, with around 90% of cases occurring in women globally.
- Men are more likely to present with advanced disease that responds poorly to treatment.
- Age also influences outcomes, with people diagnosed under 50 experiencing more severe and progressive disease and poorer treatment response than people diagnosed later in life.

- There are potentially poorer outcomes in men and younger people.
- Although the condition predominantly affects people over 40, younger women may have additional concerns about fertility.

The committee concluded that, although reducing differences in access and liver transplant waiting times were outside of its remit, it was mindful to consider current practice and its impact on patient experience. It said that it would also continue to consider potential differences in prevalence and outcomes within groups with protected characteristics under the Equality Act 2010.

Conclusion

Recommendation

3.16 The committee considered the clinical trial evidence for seladelpar in treating PBC, including its potential to improve liver biochemistry and pruritus compared with OCA and elafibranor, with or without UDCA. But it noted that seladelpar had not been directly compared with these treatments in clinical trials, and that the results of indirect comparisons were highly uncertain. The cost-effectiveness estimates were above the range NICE considers an acceptable use of NHS resources. So, seladelpar should not be used as an option for treating PBC.

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 Megan John

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, a project manager and an associate director.

Madiha Adam

Technical lead

Mary Hughes

Technical adviser

Kate Moore

Project manager

Elizabeth Bell

Principal technical adviser

ISBN: [to be added at publication]

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Gilead Response to Draft Guidance Document

August 2025

File name	Version	Contains confidential information	Date
ID6249_Seladelpar_ACD Response [REDACTED]_20.08.2025	1.0	Yes	20/08/2025

Contents

List of Tables

List of Figures

Executive Summary

Gilead are grateful for the opportunity to respond to feedback from the NICE appraisal committee summarised within the Appraisal Consultation Document (ACD). While we are disappointed that seladelpar did not receive an initial positive recommendation for previously treated primary biliary cholangitis (PBC), we are pleased that the committee recognised the substantial effect that PBC has on people's lives and appreciated the clinical value of seladelpar for improving alkaline phosphatase (ALP) levels and reducing pruritus in this patient population.

However, the company are particularly concerned about NICE's decision to re-issue the final appraisal scope for seladelpar, with the appraisal scope now listing fibrates as a potential comparator. This U-turn from the conclusion reached by NICE after the detailed consultation on scope raises questions of procedural fairness and creates an inconsistency with the appraisal scope of elafibranor for previously treated PBC (TA1016), which Gilead considers inequitable. We are encouraged by the position of the patient expert who expressed concern about the prescription of fibrates for the treatment of PBC, citing a lack of licence and limitations surrounding clinical effectiveness and safety.

We discuss this and other key committee requests in our response to the ACD in the following sections:

1. The inclusion of fibrates as a comparator in the re-issued appraisal scope (Topic 1)
    - There are no exceptional circumstances to justify the proposed change to the appraisal scope. The decision to include fibrates in the appraisal scope without sufficient justification or consultation with original stakeholders brings into question the fairness of this procedure.
    - The listing of fibrates as a comparator in the appraisal scope is inconsistent with earlier NICE appraisals in PBC. Without a valid justification, this appears unreasonable.
    - There is no clear evidence that fibrates are part of established clinical practice for the population in the NHS that is subject to this appraisal, a

criterion that must be met for an unlicensed product to be used as a comparator.

- The committee have not sufficiently considered the safety issues associated with fibrates use in PBC, which it must do before using an unlicensed product as a comparator.

- A comparison of the clinical benefits of seladelpar versus fibrates using conventional statistical methods, such as Bayesian network meta-analysis (NMA), is not feasible due to the violation of transitivity across inclusion criteria, baseline population characteristics, and composite response definitions.

2. Concerns with the company's approach to the indirect treatment comparison (ITC) (Topic 2)

    - Gilead wishes to reiterate that there is evidence for the presence of unbalanced effect modifiers between studies, which necessitates a matching-adjusted indirect comparison (MAIC) for a valid comparison (as per NICE TSD 18). This is supported by ITC experts and key opinion leaders (KOLs).

    - The three trials (RESPONSE, POISE, and ELATIVE) have differing definitions of the ULN for ALP. All ITCs submitted to NICE required redefining the definition of the upper limit of normal (ULN) in RESPONSE to match that of the comparator trial, thus NMA results consequently differed from the naive trial data and reversed the direction of the effect in the case of the comparison vs. seladelpar for the Toronto I (ALP response) outcome.

    - These differences in definitions of the ULN for ALP preclude synthesis of a single network of all three studies using a consistent definition of the ULN for ALP.

    - Despite this limitation, Gilead has provided an updated ITC report, including reporting of legacy requests from the EAG and a Bayesian NMA incorporating all three comparators, as data on file

3. ITC-derived discontinuation rates (Topic 3)

- Gilead proposes that the most realistic scenario would be to use the real-world evidence of discontinuation for obeticholic acid (OCA) and use the ITC to derive rates for seladelpar and elafibranor.
- As there is no clinical rationale for differing discontinuation rates for seladelpar vs. elafibranor, Gilead proposes using the average from the ITC for both.

4. Pruritus disutility (Topic 4)
    - Gilead agrees with the committee's sentiment that pruritus disutility values were significantly undervalued in the MMRM-derived values, owing to misaligned definitions of pruritus severity between the mapping study and the model.
    - A recent abstract by Hussain *et al*. (2023) in primary sclerosing cholangitis (PSC) is identified as an alternative source for pruritus disutilities.
    - Gilead also proposes an alternative approach to application of the Smith et al. values that uses an alternative 'no pruritus' anchor from which to calculate disutilities.

Updated company base case post-ACM1

Gilead would like to highlight the company has accepted committee's preference on baseline ALP distribution (i.e. no patients will start from ALP normalisation and mild state [ALP ≤ 1.67 x ULN]) and long-term discontinuation rate (ratio of 0.12 based on RESPONSE and ASSURE) as its base-case post appraisal committee meeting 1 (see ACD section 3.14).

Furthermore, Gilead is also in the process of revising its patient access scheme (PAS) discount, further reducing the discounted price of **£**[redacted] per pack to **£**[redacted] per pack.

The company base-case post factual accuracy check (FAC) and updated post-ACM1 (with previous and updated PAS prices) are presented in Table 1 below. All subsequent scenarios in this response document reflect a one-way change from the company's base-case post-ACM1.

Please note that these results reflect the corrections to the pruritus distribution provided by Gilead at clarification which did not appear to have been updated in the EAG's post factual accuracy model.

Table 1: Company base-case post FAC and post ACM1

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	iNHB at 20k
Company base-case post ACM1 (original PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]	-	-	-		[redacted]	-
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]
Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post FAC (original PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	384,050	15.458	9.416	[redacted]	[redacted]	[redacted]	Dominates	-9.79	[redacted]
Elafibranor + UDCA	451,437	15.703	10.033	[redacted]	[redacted]	[redacted]	Dominates	-12.54	[redacted]
Company base-case post ACM1 (updated PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]

Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post FAC (updated PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	384,050	15.458	9.416	[redacted]	[redacted]	[redacted]	Dominates	-9.79	[redacted]
Elafibranor + UDCA	451,437	15.703	10.033	[redacted]	[redacted]	[redacted]	Dominates	-12.54	[redacted]

Key: ACM1, appraisal committee meeting 1; FAC, factual accuracy check; ICER, incremental cost-effectiveness ratio; LY, life year; NHB, net health benefit; OCA, obeticholic acid; PAS, patient access scheme; QALY, quality-adjusted life year; UDCA, ursodeoxycholic acid.

Topic 1 The inclusion of fibrates in the re-issued appraisal scope

ACD section 3.3 *"The committee thought that overall, fibrates were potentially a comparator for people with PBC with pruritus in NHS clinical practice. It considered that it would be informative to see a comparison of benefits and costs of seladelpar compared with fibrates. The committee concluded that it was appropriate to re-issue the scope, listing fibrates as a potential comparator to allow exploration of this. The committee also noted that input from NHS England, patients and clinicians would help the committee to define the comparators."*

Company response:

- There are no exceptional circumstances to justify the proposed change to the appraisal scope. The decision to include fibrates in the appraisal scope without sufficient justification or consultation with original stakeholders brings into question the fairness of this procedure.
- The listing of fibrates as a comparator in the appraisal scope is inconsistent with earlier NICE appraisals in PBC. Without a valid justification, this appears unreasonable.
- There is no clear evidence that fibrates are part of established clinical practice for the population in the NHS that is subject to this appraisal, a criterion that must be met for an unlicensed product to be used as a comparator.
- The committee have not sufficiently considered the safety issues associated with fibrates use in PBC, which it must do before using an unlicensed product as a comparator.
- A comparison of the clinical benefits of seladelpar versus fibrates using conventional statistical methods, such as Bayesian NMA, is not feasible due to the violation of transitivity across inclusion criteria, baseline population characteristics, and composite response definitions.

1.1 There are no exceptional circumstances to justify the proposed change to the appraisal scope

The company has raised concerns with NICE in relation to the committee decision to amend the final published scope to include fibrates as a comparator, reversing the conclusion reached by NICE after previous detailed consultation on the scope.

NICE has confirmed that there is no explicit process detailed in the current health technologies evaluation manual to cover this circumstance.

However, PMG20 provides some guidance on what NICE considers a fair and transparent approach when similar circumstances arise in the development of other types of NICE guidance. PMG20 states that "*There can be exceptional circumstances when the final scope may need amending after it has been signed off and published on the NICE website. For example, amendments may be needed in the light of policy changes, the withdrawal of a medicine, or to include a NICE technology appraisal in development*."

During the scope consultation for seladelpar, NICE considered that "*fibrates are not a direct active comparator, but rather are part of background treatments that people with PBC will be offered*." In November 2024, NICE stated that "*NICE agrees that fibrates should be removed, so the scope has been updated to reflect this*." NICE subsequently removed fibrates as a comparator from the final scope (1). This aligned the final scope of seladelpar with that of elafibranor (TA1016) for the NICE appraisal (2).

When determining an appraisal scope, NICE is required to identify all potentially relevant comparators. At this stage of the evaluation, identifying comparators should be inclusive. (NICE Methods, 2.2.12). So NICE determined that fibrates were not potentially relevant comparators.

The committee has entirely reversed this position without sufficient justification or consultation with the original stakeholders, which brings into question the fairness of this procedure.

There are no exceptional circumstances that would justify this U-turn: since the issuance of the final scope, there has been no change to clinical practice in the NHS in relation to use of fibrates, and no clinical guidelines or clinical evidence published to support the inclusion of fibrates as a comparator.

NICE referred two other appraisals in which the scope was amended after the first committee meeting (TA1079 and TA724), however the circumstances in those appraisals can easily be distinguished from the current appraisal:

- **Fruquitinib for previously treated metastatic colorectal cancer (TA1079)** - Trifluridine-tipiracil with bevacizumab was added as a comparator because of the ongoing NICE appraisal for this new technology (TA1008) which if introduced would significantly affect the treatment pathway. The committee noted this technology would quickly become a relevant comparator. This is not the case in the current appraisal where the comparator being introduced and its place in the treatment pathway was fully known about at the time of initial scoping (3).

- **Nivolumab with ipilimumab and chemotherapy for untreated metastatic non-small-cell lung cancer (TA724)** - Pembrolizumab with pemetrexed and platinum chemotherapy was added as a comparator after the first appraisal committee meeting because of clinical expert input regarding a change to established clinical practice when the combination moved from the Cancer Drugs Fund to routine commissioning after initial scoping (4). As noted above, there has been no change, and there is no anticipated change, to clinical practice in the NHS in relation to use of fibrates to support their inclusion as a comparator since the initial scoping in this appraisal.

We note that PMG36 only expressly anticipates that the scope may need to be updated if there is a significant length of time between scoping and the evaluation (NICE Methods 2.9.3). This implies that scope may be updated if there has been a material change in circumstances. In this appraisal the timings have followed the expected timetable and there has been no change in circumstances.

1.2 The listing of fibrates as a comparator in the appraisal scope is inconsistent with earlier NICE appraisals in PBC and there is no justification for this

As noted in the company response to the EAG report, fibrates were not included as a comparator in the recent NICE appraisal of elafibranor (TA1016) (2), and whilst they were listed as a comparator for the final scope of NICE's appraisal of OCA (TA443), no evidence on fibrates as a comparator was presented to the committee (5). NICE is

committed to following a consistent approach and where products are indicated for a similar patient population, NICE must provide justification for the decision to deviate from an established precedent. The committee has not provided this justification.

The committee noted in the draft guidance that the EAG "*highlighted a UK audit that found that 50% of people whose PBC had responded inadequately to UDCA had treatment with fibrates*". This same audit was reviewed by the EAG in TA1016, however they concluded that whilst over half of people having second-line treatment for PBC had fibrates, "*these may have been used as an add-on treatment for itching, rather than to treat PBC*" (TA1016 Final Guidance 3.4). The clinical expert view in TA443 also supported the position that fibrates "*are not disease-modifying drugs and so … are not an appropriate comparator*" (TA443 FAD 4.3).

It would appear that the only reason the committee has for deviating from the position set in TA1016 and TA443 is the marketing authorisation for seladelpar states that it is for the treatment of PBC, including pruritus, whereas the marketing authorisations for elafibranor and OCA only state that these treatments are for treating PBC. It would be a perverse and unreasonable outcome if seladelpar was found not to be cost-effective because of an added benefit stated in its marketing authorisation over elafibranor and OCA.

1.3 There is no clear evidence that fibrates are considered to be part of established clinical practice for the population in the NHS that is subject to this appraisal

Fibrates do not have regulatory approval for use in PBC. Before introducing fibrates as a comparator, the committee must be satisfied that fibrates are part of established clinical practice for the population in the NHS (NICE Methods 6.2.4).

The draft guidance acknowledges that the EAG concluded that fibrates were a relevant comparator because "*1 of its 3 clinical experts had stated that fibrates were their preferred second line treatment in clinical practice*" and that a UK audit "*found that 50% of people whose PBC had responded inadequately to UDCA had treatment with fibrates*". The draft guidance also acknowledges that the clinical expert at the committee meeting "*noted that around 20% to 25% of people use fibrates as a second line treatment, including a small proportion who used them as monotherapy at this position in the treatment pathway*." The clinical expert also agreed that the preferred

option would be to use the licensed second-line treatments, although off-label fibrates were currently being used.

This information is not compelling evidence to support the argument that fibrates are considered to be part of established clinical practice for the population in the NHS that is subject to this appraisal. In particular, it is not clear that this information relates to patients:

- With an inadequate response to UDCA, specifically defined as ALP above 1.67 times the upper limit of normal. Without this specificity, other definitions of inadequate response may be applied, both on the level of ALP and what constitutes the upper limit of normal.
- In a specialist referral centre where licensed second line treatments are available following assessment by a multi-disciplinary team

In response to the draft guidance published by NICE, the company consulted the opinion of six clinical experts to gather feedback on fibrates and place in therapy for the treatment of patients with PBC in NHS clinical practice. Of the six clinical experts contacted by the company, five provided a response. Overall, fibrates were not considered as a second-line treatment for PBC. Clinical experts highlighted that their use is limited to specific scenarios, such as in patients who fail or do not tolerate approved therapies, for itch management in combination with NICE approved PBC treatments (prior to the approval of elafibranor), and in combination as a triple-therapy in patients unresponsive to NICE approved PBC treatments. Clinical experts also highlighted a preference to prescribe products with regulatory approval in the target indication. Verbatim responses can be found in Appendix A of the ACD response.

1.4 The committee have not sufficiently taken into account of the safety issues associated with fibrates use in PBC

Fibrates are not licensed for the population defined in the scope, and whilst NICE is entitled to consider off-label comparators, NICE must take into account "*the extent and quality of evidence, particularly for safety and efficacy, for the unregulated use*." (NICE Methods 6.2.4). The clinical experts in TA1016 stated that "*fibrates would not be widely used as a second-line treatment for PBC because of toxicity and limited evidence of efficacy*" (TA1016 Final Guidance 3.4), while in the draft guidance for seladelpar, the

patient expert highlighted that "*fibrates have limited evidence of effectiveness and had not been through the regulatory process to assess their clinical effectiveness and safety for treating PBC and associated pruritus*." This is supported by clinical guidelines, which highlight that the "*the evidence supporting their use remains limited to small groups of patients with limited follow-up* (6, 7)" However, it seems these concerns seem to have been overlooked by the current committee.

Clinical experts in consultation with Gilead have raised concerns over the safety profile of fibrates, particularly related to liver and renal toxicity. One clinical expert highlighted that the use of adjunctive bezafibrate was linked to a worsening eGFR in some older patients with PBC, while in PBC patients with cirrhosis, there have been documented incidences of worsening liver blood test results or bilirubin which have caused treatment cessation.

Based on published data, fibrates may lead to creatinine elevation, raising concerns about nephrotoxicity, and approximately 5–10% of patients, especially those on bezafibrate, experience musculoskeletal pain. Both fenofibrate and bezafibrate are contraindicated in those with renal impairment, as well as those with hepatic impairment for bezafibrate and those with active liver disease for fenofibrate. The efficacy of bezafibrate is limited in patients with portal hypertension, and those with cirrhosis have high discontinuation rates with fenofibrate (6).

1.5 An indirect comparison of benefits of seladelpar versus fibrates is not considered feasible due to violations of transitivity

At the request of the committee, Gilead conducted a feasibility assessment to assess the viability of an indirect treatment comparison (ITC) of the benefits of seladelpar versus fibrates for the treatment of patients with PBC.

A systematic literature review (SLR) was conducted to identify the clinical efficacy and safety of various pharmacological therapies used for the treatment of adult patients with PBC, followed by feasibility and ITC to draw quantitative comparisons between seladelpar and relevant comparators. For results and methods of the SLR, please refer to *Data on File – Fibrates Feasibility Assessment and ITC Report*.

A total of 24 studies from 183 publications were included in the clinical SLR, of which nine studies reported efficacy and safety outcomes for fibrates (bezafibrate, fenofibrate). However, only three studies assessing fibrates reported outcomes at 12

months. Analysable data for relevant outcomes (model-specific) was only reported in a randomised, controlled trial on fenofibrate in PBC patients with an incomplete response to UDCA by Li *et al*. (2022) (8).

However, the inclusion criteria for Li *et al* (2022) is not aligned with those from RESPONSE. In addition, significant heterogeneity is observed across the two studies based on various factors (as described in Table 2).

Table 2: Heterogeneity observed across methodological and clinical characteristics between RESPONSE and Li *et al*. (2022)

Factor	Description
Study characteristics	RESPONSE was a 12-month, double-blind, placebo-controlled study conducted in 193 PBC patients recruited across 90 sites in 24 countries (9). By contrast, Li *et al*. (2022) was a 12-month randomised, controlled, open-label study conducted in 48 non-cirrhotic PBC patients recruited from a single centre in China (8).
Population characteristics	The differences in age thresholds and baseline levels across RESPONSE and Li et al. (2022) presented significant challenges for comparing results and conducting a robust ITC. The mean age of patients enrolled in RESPONSE was higher (56.8 years) versus Li *et al*. (2022) (50.9 years). Furthermore, the proportion of females varied slightly between trials (94.2% in RESPONSE versus 89.6% in Li *et al*. [2022]) (8, 9).
ALP cut-off criteria for trial inclusion	Patients enrolled in RESPONSE required an ALP level ≥ 1.67 times the upper limit of normal (ULN) (9). By contrast, patients with ALP >ULN were recruited for the Li et al. (2022) study (8). Furthermore, the average ALP values at baseline (U/L) emphasised further differences across the studies (314.2 U/L in RESPONSE vs 197.5 U/L in Li et al. (2022) (8, 9). As ALP is a key effect modifier, variability in its baseline levels and inclusion criteria can lead to heterogeneity in treatment outcomes, which could violate the critical assumption of transitivity in an ITC. As the Li *et al.* (2022) ALP criteria were broader, it was not possible to match the RESPONSE data to the Li et al. (2022) study as it is not possible to 'add back' patients with lower ALP scores.
Cirrhosis (%)	The Li *et al.* (2022) study studied the efficacy and safety of fenofibrate in the treatment of non-cirrhotic PBC patients only, hence no patients with cirrhosis were enrolled on the study (8). By contrast, 14% of the patient population in RESPONSE had cirrhosis at baseline (9).
AMA positivity	AMA is considered to be a key biomarker for PBC. AMA positivity was highlighted as a key inclusion criterion for RESPONSE. AMA status was not provided in Li *et al*. (2022) (8).
Variability in composite response endpoint	In RESPONSE, composite response is defined as ALP <1.67×ULN. ≥15% ALP decrease from baseline, and total bilirubin ≤ 1.0× ULN. By contrast, Li *et al.* (2022) defined composite response as normalisation of ALP, GGT, and total bilirubin, without using standardized ULN-based thresholds (8). Due to differences in response definitions, a comparative analysis was not feasible for this outcome.

Inconsistent reporting of ALP change from baseline	Li *et al*. (2022) reported the change from baseline in ALP (U/L) using the median and IQR, which do not reflect the full data distribution (8). Additionally, most ITC/MAIC methods rely on SEs or 95% CIs to assess uncertainty. Since calculating the 95% CI from IQRs after digitization involves approximations, this method introduces bias and fails to account for the full distribution shape or the presence of outliers. Furthermore, biomarkers such as ALT, AST, bilirubin, GGT, and liver stiffness are often right-skewed, which violates the normality assumptions required by standard ITC/MAIC techniques.
Unclear definition of Toronto I	Li *et al.* (2022) did not clearly define Toronto I. Instead, they used the broader term Toronto. This lack of clarity raises concerns, as the reference cited to support their definition, Kumagi et al. (2010), describes two distinct thresholds for ALP response: ALP > 1.67 x ULN, and ALP > 1.76 x ULN (8). Hence, it remains unclear as to which of these ALP thresholds was used in the study.
Absence of ULN-based threshold for ALP	The Li *et al.* (2022) study did not provide a ULN-based threshold for ALP, amongst other biomarkers (e.g., total bilirubin) (8). In clinical trials involving biochemical markers, results are often normalised to the ULN to account for differences in laboratory reference ranges. When comparing response rates or treatment effects across studies (e.g., in meta-analyses or NMA), using ULN-based thresholds ensures the definitions of outcomes (such as "normalization" or "response") are uniform. The lack of standardization through ULN cut-offs introduces heterogeneity and violates transitivity in indirect comparisons.

Key: ALP: alkaline phosphatase; AMA: antimitochondrial antibody; CI: confidence interval; GGT: gamma-glutamyl transferase; ITC: indirect treatment comparison; MAIC: matching-adjusting indirect comparison; NMA: network meta-analysis; ULN: upper limit of normal.

Given the differences in the inclusion criteria, baseline population characteristics, and composite response definitions, an ITC of seladelpar versus fibrates may not be feasible without accounting for these discrepancies. While a mathematical comparison is technically possible for limited outcomes, it does not adhere to the statistical principle of transitivity, which is essential for valid and robust ITCs (as supported by NICE Technical Support Document 18).

A MAIC could be viewed as an alternative solution to compare the benefits of seladelpar versus fibrates for the treatment of patients with PBC. However, in this case, the extreme heterogeneity between the RESPONSE and Li *et al*. (2022) study populations render a MAIC unfeasible. After adjusting for key effect modifiers, such as ALP, cirrhosis, age, and sex, the RESPONSE effective sample size (ESS) decreases dramatically from 161 to three patients when matched to the Li *et al*. (2022) population. This substantial reduction in ESS reflects the fundamental differences between the trial populations and highlights the limitations of such adjustments in this context.

Topic 2 Concerns with the company's approach to ITC

ACD section 3.7 *"The company thought that a Bayesian NMA was unsuitable for comparing seladelpar with elafibranor. This was because of differences in baseline bilirubin and cirrhosis rates between the RESPONSE and ELATIVE trials, which the company suggested violated the transitivity assumption. That is, the trials differed in ways beyond the treatments being compared, limiting the validity of indirect comparisons. The EAG did not think that differences in bilirubin or cirrhosis rates warranted this different approach….The committee concluded that it was not appropriate to use a separate method for each comparison and that a Bayesian NMA should be used for all indirect comparisons."*

ACD section 3.8 *The committee noted that in the results reported in the individual trials (the naive results), seladelpar had a smaller estimated treatment effect than the estimate for OCA or elafibranor. But it noted that in the Bayesian NMA it had a larger treatment effect. The committee considered that this lacked face validity.*

The committee was very concerned with the analysis and did not have confidence in the results from the company's indirect comparison. It asked for:

- *the company to submit its Bayesian NMA model code with scenarios presented to explore uncertainty in the model*
- *updated reporting of the company's Bayesian NMA, including clear reporting of the original trial values and sources of numerical data for the comparators*
- *an explanation of how the trials and trial populations differed, with particular reference to effect modifiers*
- *an explanation of any difference between the reported trial results and the relative effects resulting from the Bayesian NMA*
- *consideration that Turner priors are suitable for odds ratios and would need careful justification to be used for an NMA of relative risks."*

Company response:

- Gilead wishes to reiterate that there is evidence for the presence of unbalanced effect modifiers between studies, which necessitates a MAIC for a valid comparison (as per NICE TSD 18). This is supported by ITC experts and KOLs.

- The three trials (RESPONSE, POISE, and ELATIVE) have differing definitions of the ULN for ALP. All ITCs submitted to NICE required redefining the definition of the ULN in RESPONSE to match that of the comparator trial, thus NMA results consequently differed from the naive trial data and reversed the direction of the effect in the case of the comparison vs. seladelpar for the Toronto I (ALP response) outcome.
- These differences in definitions of the ULN for ALP preclude synthesis of a single network of all three studies using a consistent definition of the ULN for ALP.
- Despite this limitation, Gilead has provided an updated ITC report, including reporting of legacy requests from the EAG and a Bayesian NMA incorporating all three comparators, as data on file

2.1 Evidence supporting effect modifiers

In the appraisal of elafibranor (TA1016), age at diagnosis, ALP levels at baseline, total bilirubin at baseline, cirrhosis, and ANA positive status were all considered to be effect modifiers based on discussion with clinicians. Clinical experts consulted by Gilead also validated these factors as being relevant effect modifiers in the present appraisal. This is supported by the subgroup analyses in RESPONSE which, while not powered to detect statistically significant differences, demonstrated the largest shifts in effect size point estimates in ALP, total bilirubin and ALP at baseline subgroups (Figure 1 and Figure 2).

The distribution of the effect modifiers included in the Gilead MAIC across the three relevant studies is presented in Table 3. In the elafibranor appraisal the homogeneity of the identified treatment effect modifiers in the ELATIVE and POISE trials was validated by clinical experts and a MAIC was not considered warranted. In contrast, external ITC expert statisticians consulted by Gilead concurred that the differences in baseline distribution of effect modifiers between RESPONSE and ELATIVE warranted a MAIC. Gilead considers that, while clinicians are best placed to identify clinical factors that are effect modifiers, statisticians are best placed to assess whether there is sufficient heterogeneity to warrant a MAIC.

Figure 1: Forest Plot of the Composite Endpoint Response Rate at Month 12 by Subgroup (ITT Analysis Set)

Figure 2: Forest Plot of the Response Rates of ALP Normalization at Month 12 by Subgroup

Table 3: Distribution of the effect modifiers in three key trials

Population characteristics	ELATIVE	RESPONSE	POISE
Baseline ALP mean U/L (SD)	321.9 (150.9)	314.3 (121.88)	323 (112.53)
Mean total bilirubin level-µmol/liter (SD)	9.6 (5.1)	12.9 (5.147)	11.1 (6.498)
Cirrhosis (%)	9.94 (8.3 in ELA and 13.2 in UDCA)	14	16
Age at diagnosis (SD) [95% CI]	--	49.23 (10.30)	47.32 (10.79)
Bilirubin >ULN at baseline (%)	3.7	13.0 (15.6 in SEL and 7.7 in UDCA)	8.3

Key: ALP, alkaline phosphatase; CI, confidence interval; ELA, elafibranor; SEL, seladelpar; SD, standard deviation; UDCA, ursodeoxycholic acid; ULN, upper limit of normal

2.2 Lack of face validity of the Bayesian NMA results vs. the naïve data

Gilead would also like to highlight that definitions of ULN in the three key trials (RESPONSE, ELATIVE, and POISE) are different (Table 4) and therefore it was necessary to adjust the ULN definition in the RESPONSE trial to the comparator trials. This adjustment led to the committee's observed discrepancies between the relative effects in the raw data vs, those in the NMA. This change in ULN definition alone leads to a reversal in direction of effect for seladelpar vs. elafibranor on the outcome of Toronto I (ALP response). To illustrate this, Table 5 presents the Bayesian NMA results before and after adjustment, alongside the original trial values. We hope this demonstrates the direction of the NMA results and naïve trial results for seladelpar vs. elafibranor before and after adjustment remain consistent, thereby addressing committee's concerns about the face validity of the NMA results.

Table 4: ALP and Bilirubin ULN cut-offs between RESPONSE, ELATIVE, and POISE trials

ULN cut-off	ALP ULN U/L		Bilirubin ULN micromoles/L	
	Female	Male	Female	Male
RESPONSE	116		18.8	
ELATIVE	104	129	20.5	

POISE	118	124	19.3	25.5
Average of ELATIVE and POISE	111	126	19.9	23.0

Key: ALP, alkaline phosphatase; ULN, upper limit of normal

Table 5: Bayesian NMA results and naïve trial comparison for Toronto I (ALP response) between seladelpar and elafibranor, before and after adjustment

RR (95%CrI)	Bayesian NMA (without outcome recalculation)	Bayesian NMA (with outcome recalculation)	Naïve comparison (without outcome recalculation)	Naïve comparison (with outcome recalculation)
Seladelpar vs. Elafibranor (Vague prior)	[redacted]	[redacted]	0.45	1.16
Seladelpar vs. Elafibranor (Informative prior: Turner et al.)	[redacted]	[redacted]		

Key: ALP, alkaline phosphatase; NMA, network meta-analysis

2.3 Model results with a Bayesian NMA incorporating all three comparators

As explained in section 2.2, the definitions of ULN in the three key trials (RESPONSE, ELATIVE, and POISE) are different (Table 4) and it is not possible to create a network of all 3 comparators with consistent definitions of ALP ULN. To reconcile the differing definitions of ULN between trials, Gilead has conducted an additional scenario analysis in the model using the average of two sets of Bayesian NMA results using ULN cut-offs matched to ELATIVE and POISE, respectively (both including all 3 comparators in a single network). Bayesian NMA results using ELATIVE and POISE matched ULN cut-offs (with adjustments for zero events and generic prior) and the average of the two are presented in Table 6.

The company base-case post ACM1 with the updated Bayesian NMA results as a single change is presented in Table 8 (original PAS) and Table 7 (updated PAS).

Table 6: NMA results using ELATIVE and POISE cut-off and the average of ELATIVE and POISE

	Elafibranor		OCA	
NMA options – ALP normalisation	**OR**	**RR**	**OR**	**RR**
NMA using ELATIVE ULN cut-off with adjustments for zero events [generic prior]	[redacted]	[redacted]	[redacted]	[redacted]
NMA using POISE ULN cut-off with adjustments for zero events [generic prior]	[redacted]	[redacted]	[redacted]	[redacted]
Average of ELATIVE and POISE [generic prior]	[redacted]	[redacted]	[redacted]	[redacted]
NMA options – Toronto I	**OR**	**RR**	**OR**	**RR**
NMA using ELATIVE cut-off [generic prior]	[redacted]	[redacted]	[redacted]	[redacted]
NMA using POISE cut-off [generic prior]	[redacted]	[redacted]	[redacted]	[redacted]
Average of ELATIVE and POISE [generic prior]	[redacted]	[redacted]	[redacted]	[redacted]

Key: ALP, alkaline phosphatase; NMA, network meta-analysis; OCA, obeticholic acid; OR, odds ratio; RR, relative risk; ULN, upper limit of normal

2.4 Consideration that Turner priors are suitable for odds ratios

NMA is conducted on the log-odds-ratio scale (binomial/logit), which allows application of Turner heterogeneity priors in their intended context. Although originally developed for log-odds ratios, Turner priors can also be applied to relative risks since they capture between-study heterogeneity rather than the effect measure itself. Empirical evidence indicates that heterogeneity is generally comparable across log-OR and log-RR scales, supporting this approach, particularly when relative risks are preferred for interpretability in settings with high event rates (10-12).

Table 7: Company base-case post ACM1* with updated Bayesian NMA results, updated PAS (£ [redacted]

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	iNHB at 20k
Company base-case post ACM1									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]
Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post ACM1 + average of Bayesian NMA results using ELATIVE and POISE matched ULN cut-offs									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	412,661	15.254	9.214	[redacted]	[redacted]	[redacted]	Dominates	-11.42	[redacted]
Elafibranor + UDCA	505,450	16.009	10.316	[redacted]	[redacted]	[redacted]	Dominates	-14.96	[redacted]
Company base-case post ACM1 + Bayesian NMA results using ELATIVE matched ULN cut-offs									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	407,109	15.039	9.032	[redacted]	[redacted]	[redacted]	Dominates	-11.32	[redacted]

Elafibranor + UDCA	496,137	15.753	10.095	-[redacted]	[redacted]	[redacted]	Dominates	-14.71	[redacted]
Company base-case post ACM1 + Bayesian NMA results using POISE matched ULN cut-offs									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					-[redacted]	
OCA + UDCA	420,180	15.547	9.463	[redacted]	[redacted]	[redacted]	Dominates	-11.55	[redacted]
Elafibranor + UDCA	517,035	16.346	10.615	[redacted]	[redacted]	[redacted]	1,095,856	-15.24	[redacted]

Key: ACM1, appraisal committee meeting 1; ICER, incremental cost-effectiveness ratio; LY, life year; NHB, net health benefit; NMA, network meta-analysis; OCA, obeticholic acid; PAS, patient access scheme; QALY, quality-adjusted life year; UDCA, ursodeoxycholic acid; ULN, upper limit of normal
Notes: *company base-case post ACM incorporated two changes preferred by the committee: baseline ALP level and long-term discontinuation (Table 1)

Table 8: Company base-case post-ACM1* with updated Bayesian NMA results, original PAS (£[redacted]

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	iNHB at 20k
Company base-case post ACM1									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]
Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post ACM1 + average of Bayesian NMA results using ELATIVE and POISE matched ULN cut-offs									

Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	412,661	15.254	9.214	[redacted]	[redacted]	[redacted]	Dominates	-11.42	[redacted]
Elafibranor + UDCA	505,450	16.009	10.316	[redacted]	[redacted]	[redacted]	Dominates	-14.96	[redacted]
Company base-case post ACM1 + Bayesian NMA results using ELATIVE matched ULN cut-offs									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	407,109	15.039	9.032	[redacted]	[redacted]	[redacted]	Dominates	-11.32	[redacted]
Elafibranor + UDCA	496,137	15.753	10.095	[redacted]	[redacted]	[redacted]	Dominates	-14.71	[redacted]
Company base-case post ACM1 + Bayesian NMA results using POISE matched ULN cut-offs									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	420,180	15.547	9.463	[redacted]	[redacted]	[redacted]	Dominates	-11.55	[redacted]
Elafibranor + UDCA	517,035	16.346	10.615	[redacted]	[redacted]	[redacted]	766,615	-15.24	[redacted]

Key: ACM1, appraisal committee meeting 1; ICER, incremental cost-effectiveness ratio; LY, life year; NHB, net health benefit; NMA, network meta-analysis; OCA, obeticholic acid; PAS, patient access scheme; QALY, quality-adjusted life year; UDCA, ursodeoxycholic acid; ULN, upper limit of normal
Notes: *company base-case post ACM incorporated two changes preferred by the committee: baseline ALP level and long-term discontinuation (Table 1)

Topic 3 ITC-derived discontinuation rates

ACD section 3.8 *"The EAG preferred using ITC-derived rates, which it said better accounted for differences in trial populations and aligned with the methodology used for other model parameters…These ITC-derived rates were notably higher for comparators, especially for OCA, which had a 26.95% discontinuation rate at 12 months compared with 9.59% in the company's base case. Seladelpar's discontinuation rate remained consistent at 6.73% across both approaches… The committee concluded that it preferred the EAG's approach, with the caveat that the 12-month rates were derived from the ITC, which introduced uncertainty around these estimates"*

Company response:

- Gilead proposes that the most realistic scenario would be to use the real-world evidence of discontinuation for OCA and use the ITC to derive rates for seladelpar and elafibranor.
- As there is no clinical rationale for differing discontinuation rates for seladelpar vs. elafibranor Gilead proposes using the average from the ITC for both.

3.1 Alternative method of generating discontinuation rates

Gilead would like to clarify that the rates quoted in the document: 26.95% (OCA ITC-derived), 9.59% (OCA company base-case), and 6.73% (seladelpar) represent cumulative discontinuation by month 12, not discontinuation rates as stated.

The EAG preferred to anchor the short-term discontinuation rates on the seladelpar RESPONSE trial and apply the ITC to obtain rates for elafibranor and OCA. Given that the EAG identified real-world evidence (Abbas et al., 2024) for OCA discontinuation rates in the UK (see EAG report addendum) (13), Gilead proposes that this is used as the anchor for applying discontinuation rates from the ITC. Specifically, the discontinuation rate reported in Abbas et al., 2024 is used to anchor the OCA rate in the ITC to derive corresponding discontinuation rates for seladelpar and elafibranor (13). Unlike OCA, which has been associated with poorer pruritus outcomes, there is no clinical rationale for the discontinuation rates of seladelpar and elafibranor to differ; small differences may simply be due to patient characteristics and/or timing of the

clinical studies and availability of alternative treatments. Therefore, Gilead proposes applying the average of the two resulting estimates for seladelpar and elafibranor. The updated discontinuation rates are presented in Table 9 below. The company base-case post ACM1 plus the updated discontinuation rates is presented in Table 10.

Table 9: Updated short-term (month 0-12) discontinuation rates

Treatment	Discontinuation rate	Source
OCA	22.1%	Abbas et al., 2024 (13)
Elafibranor	8.4%	ITC using OCA discontinuation (22.1%) as the anchor. Average of elafibranor (10.4%) and seladelpar (6.4%)
Seladelpar	8.4%	

Table 10: Results of company base-case post-ACM1* with updated discontinuation rates and PAS

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	iNHB at 20k
Company base-case post ACM1 (updated PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]
Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post ACM1 (updated PAS: £[redacted] + updated short-term (month 0-12) discontinuation rates									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	381,096	15.355	9.337	[redacted]	[redacted]	[redacted]	Dominates	-9.72	[redacted]
Elafibranor + UDCA	507,878	15.873	10.194	[redacted]	[redacted]	[redacted]	Dominates	-15.20	[redacted]
Company base-case post ACM1 (original PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]

Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post ACM1 (original PAS: £[redacted] + updated short-term (month 0-12) discontinuation rates									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	381,096	15.355	9.337	[redacted]	[redacted]	[redacted]	Dominates	-9.72	[redacted]
Elafibranor + UDCA	507,878	15.873	10.194	[redacted]	[redacted]	[redacted]	Dominates	-15.20	[redacted]

Key: ACM1, appraisal committee meeting 1; ICER, incremental cost-effectiveness ratio; LY, life year; NHB, net health benefit; OCA, obeticholic acid; PAS, patient access scheme; QALY, quality-adjusted life year; UDCA, ursodeoxycholic acid
Notes: *company base-case post ACM incorporated two changes preferred by the committee: baseline ALP level and long-term discontinuation (Table 1)

Topic 4 Pruritus disutility

ACD section 3.8 *"The committee preferred the EAG approach. This was because it was based on trial data from people like those who would have treatment in clinical practice and it used data from a consistent source for all health states. But it acknowledged that this approach might underestimate the burden of pruritus because the disutility values were very small. The clinical expert said that it would expect a utility loss from pruritus. […] The committee said that additional evidence from the literature may be of value."*

Company response:

- Gilead agrees with the committee's sentiment that pruritus disutility values were significantly undervalued in the MMRM-derived values, owing to misaligned definitions of pruritus severity between the mapping study and the model.
- A recent abstract by Hussain et al. (2023) in primary sclerosing cholangitis (PSC) is identified as an alternative source for pruritus disutilities.
- Gilead also proposes an alternative approach to application of the Smith et al. values that uses an alternative 'no pruritus' anchor from which to calculate disutilities.

4.1 Misalignment of pruritus severity definition between the mapping study and the model

In order to derive utility values for the pruritus NRS states of the model, Gilead conducted a mapping exercise using the ITCH-E data, which stratified patients according to their level of pruritus on the PBC-40 measure (unfortunately, no dataset was available that permitted generation of utility values by NRS category). Gilead would like to highlight that there is therefore a misalignment on how pruritus severity was defined in the mapping study compared to the model, as the mapping study defined mild pruritus and clinically severe pruritus on the PBC-40 itch domain score while the model defined mild, moderate and severe pruritus severity by NRS.

An analysis of the distribution of pruritus in the RESPONSE trial based on PBC-40 and NRS indicates that the utility values applied to the RESPONSE severe pruritus patients (derived from ITCH-E patients with PBC-40 ≥7) were likely generated from

patients with milder NRS scores from ITCH-E (21% mild, 41% moderate, and 37% severe based on mapping between the RESPONSE NRS and PBC-40 baseline data, see Table 11 and Table 12). This likely resulted in a 'dilution' of the disutility associated with severe pruritus. Gilead maintains that disutility from the mapping study is not the appropriate source for pruritus disutility due to this pruritus definition misalignment and inclusion of milder NRS scores in the disutility applied to severe patients.

Table 11: Comparison of pruritus severity based on PBC-40 itch score and NRS in the response trial at baseline

		NRS				
		0 None	**1-3 Mild**	**4-7 Moderate**	**≥7 Severe**	**Total**
PBC-40	**0 None**	[redacted]	[redacted]			100%
	1-7 Mild	[redacted]	[redacted]	[redacted]	2%	100%
	≥7 clinically severe		[redacted]	[redacted]	37%	100%

Key: NRS, numerical rating scale

Table 12:Comparison of pruritus severity based on PBC-40 itch score and NRS scale in the response trial at month 12

		NRS					
		0 None	**1-3 Mild**	**4-7 Moderate**	**≥7 Severe**	**Missing**	**Total**
PBC-40	**0 None**	[redacted]	[redacted]			[redacted]	100%
	1-7 Mild	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100%
	≥7 clinically severe		[redacted]	[redacted]	[redacted]	[redacted]	100%
	Missing	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100%

Key: NRS, numerical rating scale

4.2 Alternative literature source for pruritus disutility

In line with the committee's conclusion that additional evidence (on pruritus disutility) from the literature may be of value, Gilead has identified a study by Hussain et al

(2023) in primary sclerosing cholangitis (PSC) which reported mean EQ-5D utilities of 0.81, 0.77, 0.70, and 0.68 for no, mild, moderate, and severe itch (defined using the NRS), respectively. These health state utilities generate disutilities for mild, moderate, and severe itch of 0.04, 0.11, and 0.13 respectively (see Table 13) (14). Gilead considers this additional source of pruritus disutility based on Hussain et al (2023) to be more appropriate than the MMRM-derived values as the utility is based on the model-relevant outcome of pruritus defined by NRS score. Although it concerns a slightly different indication, it has the additional advantage that, in contrast with the Smith et al. literature values used in the model base case, the differences in utility between patients with different severities of pruritus (including no pruritus) have been captured in a single patient cohort. The company base-case post ACM1 with the updated pruritus disutilities is presented in Table 14.

Table 13: Pruritus disutilities based on Hussain et al (2023)

Utility of PSC patients with no itch	Utility of PSC patients by itch severity		Disutility by itch severity
0.81	Mild pruritus (on NRS scale)	0.77	**0.04**
	Moderate pruritus (on NRS scale)	0.70	**0.11**
	Severe pruritus (on NRS scale)	0.68	**0.13**

Key: NRS, numerical rating scale; PSC, primary sclerosing cholangitis

Table 14: Company base-case post-ACM1* with pruritus disutility based on Hussain et al and updated PAS

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	iNHB at 20k
Company base-case post ACM1 (updated PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]
Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post ACM1 (updated PAS: £[redacted] + pruritus disutility based on Hussain et al									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	10.670	[redacted]	[redacted]	[redacted]	Dominates	-10.42	[redacted]
Elafibranor + UDCA	499,347	15.826	11.033	[redacted]	[redacted]	[redacted]	Dominates	-13.93	[redacted]
Company base-case post ACM1 (original PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]
Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post ACM1 (original PAS: £[redacted] + pruritus disutility based on Hussain et al									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	

OCA + UDCA	421,717	15.606	10.670	[redacted]	[redacted]	[redacted]	Dominates	-10.42	[redacted]
Elafibranor + UDCA	499,347	15.826	11.033	[redacted]	[redacted]	[redacted]	Dominates	-13.93	[redacted]

Key: ACM1, appraisal committee meeting 1; ICER, incremental cost-effectiveness ratio; LY, life year; NHB, net health benefit; OCA, obeticholic acid; PAS, patient access scheme; QALY, quality-adjusted life year; UDCA, ursodeoxycholic acid
Notes: *company base-case post ACM incorporated two changes preferred by the committee: baseline ALP level and long-term discontinuation (Table 1)

4.3 Alternative application of the Smith et al. disutilities

In Gilead's submission disutilities by pruritus severity were generated by subtracting health state utility values generated by Smith et al. from the GLIMMER study (15, 16) from the utility value of an Italian study of PBC patients without pruritus (0.87). The EAG noted that the no pruritus utility value of 0.87 from the Italian study was higher than the general population (0.804), which lacked face validity. In response to this and to avoid relying on an external dataset while retaining Smith et al. as the preferred source, Gilead has derived a new baseline utility for the GLIMMER cohort as follows:

1. Age- and gender-matched general population utility for the GLIMMER cohort that was used to generate the Smith et al. utility values (mean age 55.8, 94% female, Levy et al, 2023 (16)): **0.83**

2. Utility multiplier vs. the general population for PBC without pruritus:

   - RESPONSE trial utility of patients without pruritus (RESPONSE PBC without pruritus calculated using MMRM from mapped EQ-5D data): **0.81**
   - Expected utility of the general population for RESPONSE participants: **0.83**
   - 0.81/0.83 = utility multiplier of **0.97**

3. Baseline utility for GLIMMER PBC without pruritus: **0.83 × 0.97 = 0.81**

Using this derived baseline utility (0.81), pruritus disutilities from Smith et al. have been recalculated (Table 15). Gilead believes that the updated values now represent more realistic estimates of pruritus disutilities based on the Smith et al study using an internally derived baseline PBC utility value, addressing the EAG's and the committee's concern of introducing an external dataset to the Smith et al study cohort. The company base-case post ACM1 with the updated pruritus disutilities based on the Smith et al study is presented in Table 16.

Table 15: Updated pruritus disutilities based on Smith et al

Utility of GLIMMER PBC patients without pruritus	Utility of PBC patients with (from the GLIMMER trial as reported in Smith et al)		Updated disutility by health state	Original disutility by health state*
0.81	Mild pruritus (on NRS scale)	0.75	-0.0550	-0.115

	Moderate pruritus (on NRS scale)	0.76	-0.0550	-0.115
	Severe pruritus (on NRS scale)	0.49	-0.3200	-0.38

Key: NRS, numerical rating scale; PBC, primary biliary cholangitis
Notes: *Calculated using baseline PBC utility value of 0.87 from the Italian study

Table 16: Company base-case post-ACM1* with pruritus disutility based on Smith et al (updated baseline utility) and updated PAS

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	iNHB at 20k
Company base-case post ACM1 (updated PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]
Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post ACM1 (updated PAS: £[redacted] + pruritus disutility based on Smith et al (updated baseline utility)									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	10.107	[redacted]	[redacted]	[redacted]	Dominates	-10.98	[redacted]
Elafibranor + UDCA	499,347	15.826	10.758	[redacted]	[redacted]	[redacted]	Dominates	-14.21	[redacted]
Company base-case post ACM1 (original PAS: £[redacted]									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]

Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]
Company base-case post ACM1 (original PAS: £[redacted] + pruritus disutility based on Smith et al (updated baseline utility)									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	
OCA + UDCA	421,717	15.606	10.107	[redacted]	[redacted]	[redacted]	Dominates	-11.57	[redacted]
Elafibranor + UDCA	499,347	15.826	10.758	[redacted]	[redacted]	[redacted]	Dominates	-14.81	[redacted]

Key: ACM1, appraisal committee meeting 1; ICER, incremental cost-effectiveness ratio; LY, life year; NHB, net health benefit; OCA, obeticholic acid; PAS, patient access scheme; QALY, quality-adjusted life year; UDCA, ursodeoxycholic acid

Notes:*company base-case post ACM incorporated two changes preferred by the committee: baseline ALP level and long-term discontinuation (Table 1)

References

1. National Institute for Health and Care Excellence. Seladelpar for previously treated primary biliary cholangitis ID6429 2025 [Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta11540.

2. National Institute for Health and Care Excellence. Elafibranor for treating primary biliary cholangitis (TA1016) 2024 [Available from: https://www.nice.org.uk/guidance/ta1016.

3. National Institute for Health and Care Excellence. Fruquintinib for previously treated metastatic colorectal cancer (TA1079) 2025 [Available from: https://www.nice.org.uk/guidance/ta1079.

4. National Institute for Health and Care Excellence. Nivolumab with ipilimumab and chemotherapy for untreated metastatic non-small-cell lung cancer (TA724) 2021 [Available from: https://www.nice.org.uk/guidance/ta724.

5. National Institute for Health and Care Excellence. Obeticholic acid for treating primary biliary cholangitis (TA443) 2017 [Available from: https://www.nice.org.uk/guidance/ta443.

6. Hirschfield GM, Beuers U, Corpechot C, Invernizzi P, Jones D, Marzioni M, et al. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. Journal of Hepatology. 2017;67(1):145-72.

7. Hirschfield GM, Dyson JK, Alexander GJM, Chapman MH, Collier J, Hübscher S, et al. The British Society of Gastroenterology/UK-PBC primary biliary cholangitis treatment and management guidelines. Gut. 2018;67(9):1568-94.

8. Li C, Zheng K, Chen Y, He C, Liu S, Yang Y, et al. A randomized, controlled trial on fenofibrate in primary biliary cholangitis patients with incomplete response to ursodeoxycholic acid. Ther Adv Chronic Dis. 2022;13:20406223221114198.

9. Hirschfield GM, Bowlus CL, Mayo MJ, Kremer AE, Vierling JM, Kowdley KV, et al. A Phase 3 Trial of Seladelpar in Primary Biliary Cholangitis. N Engl J Med. 2024;390(9):783-94.

10. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta- analysis of randomized controlled trials. Med Decis Making. 2013;33(5):607-17.

11. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. London: National Institute for Health and Care Excellence (NICE); 2014.

12. Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JP. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. Stat Med. 2015;34(6):984-98.

13. Abbas N, Smith R, Flack S, Bains V, Aspinall RJ, Jones RL, et al. Critical shortfalls in the management of PBC: Results of a UK-wide, population-based evaluation of care delivery. JHEP Rep. 2024;6(1):100931.

14. Hussain N, Hirschfield B, Ferguson J, Abbas N, Bhandal K, Rogers P, et al. P34 Burden, impact and variability of pruritus in primary sclerosing cholangitis (PSC) over time: a prospective observational study. Gut. 2023;72(Suppl 3):A34.

15. Smith H, Fettiplace J, Von Maltzahn R, Das S, McLaughlin MM, Jones D. S107 More Than Just an Itch: Impact of Cholestatic Pruritus in Primary Biliary Cholangitis (PBC) on Health-Related Quality of Life (HRQoL). Official journal of the American College of Gastroenterology | ACG. 2022;117(10S):e77-e8.

16. Levy C, Kendrick S, Bowlus CL, Tanaka A, Jones D, Kremer AE, et al. GLIMMER: A Randomized Phase 2b Dose-Ranging Trial of Linerixibat in Primary Biliary Cholangitis Patients With Pruritus. Clin Gastroenterol Hepatol. 2023;21(7):1902-12.e13.

Appendix A: Responses from clinicians in England regarding fibrates and place of therapy for the treatment of patients with PBC

Note:

The below questions have been sent out to six Therapeutic Experts in PBC. We have received responses from five. One clinician was unable to respond within the given timeframe and one clinician has not responded to the request.

- Gilead Sciences have not altered any of the responses.
- The responses are as written by the HCP to our questions.
- The responses are anonymised however further details can be shared upon request.
- We have only included a description of the clinician's place of work for reference to level of expertise in managing patients with PBC.

Clinician 1:

Place of work: Tertiary PBC treatment centre

1: What are the patient characteristics that would make you decide to use or recommend a fibrate, in place of an approved second line therapy such as Obeticholic acid or Elafibranor for a patient diagnosed with PBC who has been failed by UDCA?

"We would try fibrate if they had failed or not tolerated OCA and/or Elafibranor

In patients with an adequate biochemical response to UDCA +/- a second line agent above would consider adding a fibrate. This would mean switch from elafibranor if the patient was taking this as part of their therapy, but itch was not controlled"

2: If you prescribe for PBC patients where ALP>1.67xULN and you are choosing between OCA, Elafibranor and fibrates, out of every 100 patients, how many would you likely prescribe fibrates for, where the alternative treatment option could be OCA or elafibranor?

Our centres agreed approach for 2nd line treatment inexperienced non responders is:

1. OCA if no itch

2. Elafibranor if has itch

We would not consider BZF/Fibrate at this step

For second line experienced non responders we would only consider switching (from elafibranor or OCA) to or adding fibrates (to OCA) if they had failed therapy with the regulatory approved therapies

Clinician 2:

Place of work: Tertiary PBC treatment centre

1: What are the patient characteristics that would make you decide to use or recommend a fibrate, in place of an approved second line therapy such as obeticholic acid or elafibranor for a patient who has been failed by UDCA?

Before elafibranor was licensed, I would have used a fibrate for patients with pruritus or clinically significant portal hypertension. Now that elafibranor has been licensed, I would use elafibranor for such patients. I might feel more comfortable co-prescribing bezafibrate with a statin than elafibranor with a statin — but generally try to avoid either of these combinations.

2: If you prescribe for PBC patients where ALP>1.67xULN and you are choosing between OCA, elafibranor and fibrates, out of every 100 patients, how many would you likely prescribe fibrates for, where the alternative treatment option could be OCA or elafibranor?

I can't estimate the actual numbers. Now that elafibranor has been licensed, I tend to consider elafibranor when I would previously have considered fibrates. This is probably an extreme position though, based on a personal philosophy that we should be respectful of the drug regulatory and drug licensing process.

Clinician 3:

Place of work: Tertiary PBC treatment centre

1: What are the patient characteristics that would make you decide to use or recommend a fibrate, in place of an approved second line therapy such as obeticholic acid or elafibranor for a patient who has been failed by UDCA?

My starting point of principle is that we should not use an unlicensed therapy in place of a licensed therapy unless it is in the direct interests of that patient (and discussed and agreed with that patient). Cost to the NHS is not a justification. Prior to Elafibranor and Seladelpar there was an argument that it was in the best interests of a PBC patient needing SLT who had significant itch to use a fibrate rather than OCA and on that basis we used a lot of it. This has disappeared as a justification with the newer drugs. There are 2 settings in which I use fibrates

1. People who fall out of the right restrictions for 2nd line therapy (people who don't quite make the 1.67 cut-off but are young and high risk, people who I don't want to wait a full year to watch them fail therapy etc)

2. People who need or are very likely to need triple therapy.

2: If you prescribe for PBC patients where ALP>1.67xULN and you are choosing between OCA, elafibranor and fibrates, out of every 100 patients, how many would you likely prescribe fibrates for, where the alternative treatment option could be OCA or elafibranor?

5%. Only patients who are enroute to triple therapy and I want to use a PPAR first for itch reasons. Obviously Ela or Sela based triple therapy would be the ultimate therapy but we are precluded from that. I would never use fibrate as a sole 2nd line therapy in someone who has formally failed UDCA therapy

Clinician 4:

Place of work: Tertiary PBC treatment centre

1: What are the patient characteristics that would make you decide to use or recommend a fibrate, in place of an approved second line therapy such as obeticholic acid or elafibranor for a patient who has been failed by UDCA?

The main indication for Fibrates before Ela was approved was itch. I don't think small Trusts were prescribing Beza as opposed to referring patients for OCA though. Patients would still come through our PBC MDT.

The other main factor was the presence of cirrhosis. In cirrhotic patients, certainly CP B or C one would avoid OCA. I'm not saying we would then prescribe a Fibrate

as the alternative. At present, ELA is considered when patients have failed OCA, OCA + BEZA and meets 2LT targets.

2: If you prescribe for PBC patients where ALP>1.67xULN and you are choosing between OCA, elafibranor and fibrates, out of every 100 patients, how many would you likely prescribe fibrates for, where the alternative treatment option could be OCA or elafibranor?

Response not received to this question.

Clinician 5:

Place of work: Tertiary PBC treatment centre

1: What are the patient characteristics that would make you decide to use or recommend a fibrate, in place of an approved second line therapy such as obeticholic acid or elafibranor for a patient who has been failed by UDCA?

Currently I would always use licensed therapy i.e. OCA or ELA rather than a fibrate on initiating second line therapy for PBC. The only scenario where I can envisage using a fibrate here is if the patient with PBC on UDCA has low risk disease, does not have strong need for 2L therapy, but does has significant pruritus. I would use bezafibrate here primarily as anti-itch therapy.

2: If you prescribe for PBC patients where ALP>1.67xULN and you are choosing between OCA, elafibranor and fibrates, out of every 100 patients, how many would you likely prescribe fibrates for, where the alternative treatment option could be OCA or elafibranor?

5 or less

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. The Appraisal 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 provisional recommendations sound and a suitable basis for guidance to the NHS? NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations: • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced**.**
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	PBC Foundation

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months. [Relevant companies are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	£20,000 Bronze corporate sponsorship £4,000 Argentina Conference £40,000 Translation website £35,000 Summit 2025 £5,000 Online Drs Pulse module From Gilead 2024 Financials but received within last 12 months
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	
Name of commentator person completing form:	[redacted]

Comment number	**Comments** Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.
Example 1	We are concerned that this recommendation may imply that ……..
1	I have no questions, as such, but I would like to put on record my grave concerns about the conversations around fibrates, their use of and how appropriate they may or not be in patients with

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

	PBC. It also concerns me that whilst it has certain expertise, the AEG does not have expertise in PBC, nor the use of fibrates, and this came over very clearly in the information provided. There are a number of patterns around fibrate use, and these need to be explored to have a fuller understanding.
2	In the NICE territories, there are two types of premises that prescribe fibrates for PBC: specialist centres, and general hospitals (known as hubs and spokes in the ODN system). Many gastros in a general hospital will prescribe fibrates, as opposed to refer on to a specialist centre. Fibrate use in these general hospitals is obviously much more frequent than in the centres of excellence. I have a duty to point out that the OCA to fibrate ratio in hubs shows more faith in OCA, which has its own flaws, than in fibrates as a first second-line therapy. Fibrates were explored as a potentially licensed medicine under a repurposing programme and the initiative was scrapped.
3	Neither OCA or Elafibrinor were genuinely compared to fibrates and it concerns me as a patient advocate that when we have an opportunity to bring another viable treatment into the arena, the goal posts are moved. Bezafibrate, in many studies, is linked to both liver and kidney toxicity.
4	The bezurso trial had a mean ALP of approx 250, which data tells us is an ALP level more likely to respond to any of the major treatments, as opposed to the industry sponsored trials whose mean. FITCH is an interesting study which I am unsure about in terms of reaching genuine standards of changing care pathways. We, at the Foundation, have published data around patients on fibrates continuing to itch.
5	
6	

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about funding from the company and links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into one response. We cannot accept more than one set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.
- In line with the NICE Health Technology Evaluation Manual (sections 5.4.4 to 5.4.21), if a comment contains confidential information, it is the responsibility of the responder to provide two versions, one complete and one with the confidential information removed (to be published on NICE's website), together with a checklist of the confidential information. Please underline all confidential information, and separately highlight information that is submitted as 'confidential [CON]' in turquoise, and all information submitted as 'depersonalised data [DPD]' in pink. If confidential information is submitted, please submit a second version of your comments form with that information replaced with asterixis and highlighted in black.
- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the draft guidance document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. The Appraisal 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 provisional recommendations sound and a suitable basis for guidance to the NHS? NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations: • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced**.**
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	The British Association for the Study of the Liver (BASL)

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months. [Relevant companies are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	BASL Gilead (company) - BASL received £102,000 in sponsorship funding towards their Annual Conference and other educational events between August 2024 – July 2025. Ipsen Ltd (comparator) - BASL received £87,000 in sponsorship funding towards their Annual Conference and other educational events between August 2024 – July 2025. Advanz Pharma (comparator) – BASL received £5,700 in sponsorship funding towards their Annual Conference and other educational events between August 2024 – July 2025. Dr. Falk Pharma UK Ltd (comparator) - BASL received £20,201.62 in sponsorship funding towards their Annual Conference and other educational events between August 2024 – July 2025. Dr PalakTrivedi No direct funding (personal or through other means) to me as an individual. However, I have been awarded independent, peer-reviewed grants from Gilead sciences in the past (no funding yet – project to start Q4 2025/Q1 2026).
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None
Name of commentator person completing form:	[redacted]

Comment number	**Comments** Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.
Example 1	We are concerned that this recommendation may imply that ……..
1	Normalisation of ALP values (the key prognostic biomarker in primary biliary cholangitis, PBC) is achieved by one in every four patients treated with seladelpar. This metric is critical, as it confers survival advantage beyond a magnitude reduction below the 1.67x upper limit of normal threshold for those at highest risk of disease progression (i.e. PBC diagnosed below the age of 60 years,

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

	and those with a transient elastography [fibroscan] score above 10kPa). To date, **no currently licensed PBC therapy** has been able to achieve normalisation in ALP values in such a high proportion of patients. This clearly sets apart seladelpar from other medicines we use. *Ref. Corpechot et al.* Hepatology 2024
2	Additionally, much of the contemporary data from post-hoc analyses relating to the registrational RESPONSE phase 3 trial has not been covered (or not covered in enough depth. For instance, data showing ALP reductions amongst people with cirrhosis (LS mean -121.4 at month 12) is larger than seen with either obeticolic acid. Data in those with elafibranor thus far is not available specifically for this sub-grpup. *Ref.: Villamil et al.* Hepatology 2024
3	I am concerned that comparisons with off-label / unlicensed medicines such as fibric acid derivatives (fibrates) is being made here with regards effectiveness but not safety. No comparison with fibrates was made in the obeticholic acid or elafibranor panel reviews. Moreover, >10% of patients who receive fibrates in the UK population develop drug-induced elevations in liver transaminase values *Ref. Abbas et al.* Clin. Gastroenterol. Hepatol. 2023. This is not seen with Seladelpar. Moreover, the characteristics for participants in the BEZURSO trial were different (lower risk in terms of starting ALP values and pruritus severity for instance) than in the RESPONSE study.
4	Long-term efficacy and safety data regarding seladelpar in PBC (the ASSURE study – over) is extensive, and more than anything available in clinical trial data packages for obeticholic acid, bezafibrate or elafibranor. Ref. *Mayo et al.* Aliment. Pharmacol. 2024; Ther. *Levy et al.* Am. J. Gastroenterol. 2025; This is not well represented (or reflected) in the documents provided. These clearly demonstrate longevity of biochemical response and anti-pruritic effect, without adverse safety signals, with regards seladelpar. Thus, supporting its position as a key therapy to treat people with PBC. Moreover, I have presented data for up to five years of seladelpar use (*Trivedi et al.* Hepatology 2024) validating and reinforcing durability of treatment responses and safety. Again, this data is not available for fibrates, OCA or elafibranor and calls for re-evaluation.
5	
6	

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about funding from the company and links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into one response. We cannot accept more than one set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.
- In line with the NICE Health Technology Evaluation Manual (sections 5.4.4 to 5.4.21), if a comment contains confidential information, it is the responsibility of the responder to provide two versions, one complete and one with the confidential information removed (to be published on NICE's website), together with a checklist of the confidential information. Please underline all confidential information, and separately highlight information that is submitted as 'confidential [CON]' in turquoise, and all information submitted as 'depersonalised data [DPD]' in pink. If confidential information is submitted, please submit a second version of your comments form with that information replaced with asterixis and highlighted in black.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations.
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the draft guidance document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. The Appraisal 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 provisional recommendations sound and a suitable basis for guidance to the NHS? NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations: • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced**.**
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	British Hepatology Pharmacist Group

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months. [Relevant companies are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	**Yes:** • **Advanz £2000** • **Dr Falk £2000** • **Gilead £2000** • **All the above have provided sponsorship to run education study days for the British Hepatology Pharmacy Group via BASL. Not ongoing.**
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None
Name of commentator person completing form:	[redacted]

Comment number	**Comments** Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.
Example 1	We are concerned that this recommendation may imply that ……..
1	As fibrates will not go through the licensing process, we do not think it should be used as a comparator for PBC treatment. It was not considered as comparators in the recent appraisal of elafibranor for second-line PBC treatment and there is also no new real world evidence in terms of use of fibrates to warrant inclusion now. As fibrates are used in different ways across the country

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

	with variable access due to license status, including them will negatively impact patient access to seledelpar as a licensed treatment option reducing effective treatment for patients who have failed the previous two therapies which are only 50% effective.
2	In terms of uncertainties around treatment pathways, these existed with the appraisal of elafibranor, so again not sure we should be holding seledelpar to a different standard.
3	In terms of itch, bezafibrate is not licensed for itch and many other unlicensed therapies are used to manage itch so access is variable making it an inappropriate comparator.
4	There is uncertainty around the future availability of obeticholic acid in the UK since loosing its license in the EU which may result in even loss second line treatment options for patients.
5	
6	

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about funding from the company and links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into one response. We cannot accept more than one set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.
- In line with the NICE Health Technology Evaluation Manual (sections 5.4.4 to 5.4.21), if a comment contains confidential information, it is the responsibility of the responder to provide two versions, one complete and one with the confidential information removed (to be published on NICE's website), together with a checklist of the confidential information. Please underline all confidential information, and separately highlight information that is submitted as 'confidential [CON]' in turquoise, and all information submitted as 'depersonalised data [DPD]' in pink. If confidential information is submitted, please submit a second version of your comments form with that information replaced with asterixis and highlighted in black.
- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations.
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the draft guidance document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. The Appraisal 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 provisional recommendations sound and a suitable basis for guidance to the NHS? NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations: • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced**.**
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	Ipsen Ltd

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months. [Relevant companies are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	NA
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	NA
Name of commentator person completing form:	[redacted]

Comment number	**Comments** Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.
1	Typographical error: on page 6, “elafibranor” is misspelled. Please correct this.
2	During the committee meeting, some factual inaccuracies were noted. For example, it was stated that no positive data on pruritus were available for elafibranor. This is incorrect. Improvements in 5-D Itch scores and clinically meaningful improvements in PBC-40 Itch were observed, highlighting the potential of elafibranor to reduce both pruritus severity and its impact on quality of life. Published evidence clearly addresses this:

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

	• Kremer, Andreas E., et al. "P20 Effect of elafibranor on pruritus in primary biliary cholangitis: symptom severity and quality of life measurements from the phase III ELATIVE® trial." (2024): A22-A22. • Kowdley, K., C. Bowlus, and C. Levy. "Long-term efficacy and safety of elafibranor in primary biliary cholangitis: interim results from the open-label extension of the ELATIVE trial up to 3 years [AASLD abstract 5041]." *AASLD The Liver Meeting, Late Breaking Abstract Supplement*. 2024. • Kowdley, Kris V., et al. "Efficacy and safety of elafibranor in primary biliary cholangitis." New England Journal of Medicine 390.9 (2024): 795-805.
3	Another factual inaccuracy observed during the committee meeting was that the PBC-40 is not a validated tool. In fact, the PBC-40 was rigorously developed through patient interviews and has undergone extensive validation and psychometric testing in a large PBC population, making it a well-recognized and reliable measure of health-related quality of life in this setting. Reference: Jacoby A, Rannard A, Buck D, et al. Development, validation, and evaluation of the PBC-40, a disease specific health related quality of life measure for primary biliary cirrhosis. Gut 2005;54:1622-9
4	
5	
6	

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about funding from the company and links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into one response. We cannot accept more than one set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.
- In line with the NICE Health Technology Evaluation Manual (sections 5.4.4 to 5.4.21), if a comment contains confidential information, it is the responsibility of the responder to provide two versions, one complete and one with the confidential information removed (to be published on NICE's website), together with a checklist of the confidential information. Please underline all confidential information, and separately highlight information that is submitted as 'confidential [CON]' in turquoise, and all information submitted as 'depersonalised data [DPD]' in pink. If confidential information is submitted, please submit a second version of your comments form with that information replaced with asterixis and highlighted in black.
- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations.
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Draft guidance comments form

Consultation on the draft guidance document – deadline for comments 5pm on Wednesday 20 August 2025. Please submit via NICE Docs.

- If you have received agreement from NICE to submit additional evidence with your comments on the draft guidance document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Please return to: **NICE DOCS**

NICE National Institute for Health and Care Excellence

Seladelpar for treating primary biliary cholangitis [ID6429]

NHSE correspondence about the use of fibrates in clinical practice

Email from NHSE in response to whether fibrates should be considered a comparator to seladelpar for ID6429.

Our experts opinions are that fibrates will not go through the licencing process, and are used in different ways across the country, with variable utilisation owing to the license status. Therefore as they are not standard clinical practice, these should not be used as a comparator for PBC treatments.

Although fibrates are very cheap, they do not have the same objective data on long term safety that seladelpar has, and no-one is collecting the long-term safety and tolerability data on fibrates as is being collected for the licenced indications. Many clinicians are reluctant to initiate them for this reason. Also, once started, repeat prescribing becomes challenging as most GPs will not take on the repeat prescribing of fibrates, given their unlicenced status. It is felt that including them will negatively impact patient access to the licenced treatment options, thereby reducing effective treatment for patients who have not achieved effective management with the other therapies.

In relation to management of itch, bezafibrate Is not licenced for this indication, and there are many other unlicenced therapies also in use, so access is variable – and would make this an inappropriate comparator.

We note that the scope for elafibranor, didn't include fibrates as a comparator, so for consistency it would seem appropriate not to include fibrates as a comparator for seladelpar.

The elafibranor TA (1016) refers to:
'Off-label use of fibrates
3.4 The company base case did not include fibrates as a comparator for elafibranor. Bezafibrate was included in the company's model (see section 3.9 for more information about the model) for treating itching with UDCA and OCA, but not as a stand-alone second-line treatment. Submissions from professional organisations and NHS England identified fibrates as a potential comparator. The company explained that it did not include fibrates because they are used off-label and are not recommended by NICE. It added that fibrates have not been studied to regulatory standards, so there might be long-term safety concerns. The EAG referenced a UK audit that found that some people having second-line treatment for PBC had fibrates. But it noted these may have been used as an add-on treatment for itching, rather than to treat PBC. The clinical experts explained that fibrates are used in combination

with second-line treatments to treat itching. They added that fibrates would not be widely used as a second-line treatment for PBC because of toxicity and limited evidence of efficacy. The committee concluded that fibrates were not used primarily to treat PBC and were not an appropriate comparator for elafibranor.'

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Committee member comment on company's response to draft guidance regarding the ITC

1 BACKGROUND

Indirect treatment comparisons are required to compare seladelpar (SEL) to elafribranor (ELA) and OCA.

The company claims that separate approaches are required for each comparison:

- Bayesian NMA to compare SEL with OCA (informed by the RESPONSE and POISE RCTs)
- Anchored MAIC to compare SEL with ELA (informed by the RESPONSE and ELATIVE RCTs)

At ACM1 the committee queried the results and analyses methods for the NMA and why separate analyses were required. The company provided additional details of their analyses and now explain their rationale for separately comparisons.

The committee member has reviewed the additional documentation and provided comments below. Ultimately, the ITC results are very uncertain for all outcomes but particularly for "ALP normalisation at 12 months". Whilst some uncertainty is due to the methods selected by the company (see Sections 2.2.1 and 3), most of it is due to data availability for which there does not appear to be an immediate solution.

2 SEPARATE ANALYSES FOR EACH COMPARATOR

As a matter of principle, for a fully incremental comparison across more than 2 technologies, the relative effects used in the economic model should be calculated in a joint synthesis of all relevant studies. However, in this case, for the key efficacy outcomes the company argue that the outcome definitions across the 3 trials are not comparable and therefore outcomes observed in RESPONSE need to be adjusted to match the outcome definitions in ELATIVE and POISE, which also differ between them, so 2 different recalculations of the RESPONSE outcomes are required. The company also argue that a MAIC is required for the comparison to ELA due to the presence of effect modifiers, whilst a Bayesian NMA is suitable for the comparison to OCA.

2.1 Outcome recalculation

Is outcome recalculation really necessary? We could hypothesise that if the outcome definitions were sufficiently aligned across the 3 studies, recalculation might not be required as definitions within a trial would be the same for both arms and we might expect that relative effects would be unchanged. In the updated ITC report, now provided by the company, Tables 7 and 8 show the recalculated outcomes. The company notes that outcome recalculation leads to reversal of the direction of effect for the Toronto I (ALP response) outcome (company's response to DG, page 21 and section 2.2). This suggests that alternative outcome definitions have an impact on relative effects. However, it also calls into question results based on recalculated outcomes: which is the most relevant outcome definition? If it is the definition is RESPONSE then basing decisions on results based on other definitions may not be appropriate.

The company provide sensitivity analyses without outcome recalculation in the updated ITC report. Although point estimates are impacted, due to the uncertainty, 95% credible intervals (CrIs) for the Bayesian NMA with and without outcome recalculation largely overlap.

In the company response to draft guidance, Table 6 shows results of NMAs with different outcome recalculations and then with results for an average of the relative effects using both recalculations. No 95% intervals are presented, and I could not find equivalent results in the ITC report. However, this average is a simple mean across the ORs and RRs estimated from each analysis, they are not weighted by the precision of each estimate and therefore I don't think they are suitable for decision-making. However, Table 6 does show that the impact of different outcome definitions on the point estimates can be substantial.

Comments

I think the question of outcome recalculation is difficult, and one that does not have an obvious solution. In a way it is a clinical question – which is the most relevant outcome definition – but at the same time because the outcomes of the ELATIVE and POISE studies cannot be changed (as no IPD are available) we will always be left with different definitions in each of the comparator studies. As the recalculation appears to have some impact on the estimated relative effects, I am not sure that an analysis without recalculation is appropriate, particularly when putting all three studies into a single NMA. I can see the company's point for stating that a joint analysis is not appropriate (company's response to DG, page 21 and section 2.2) and that 2 separate analyses are required, one for each comparator.

2.2 MAIC vs NMA

The company argue that an anchored MAIC is required to compare SEL with ELA, due to differences in effect modifiers. I assume that this MAIC is carried out with the recalculated outcomes although this is not explicit in the analysis plan table in the ITC Report (Table 11).

As noted above, the differences in outcome definitions across the 3 trials is a strong argument for separate analyses but the fact that one comparison requires population adjustment (via MAIC) and the other does not, is not an argument for conducting separate analyses, as other methods such as ML-NMR are available to deal with this problem.

However, if separate analyses are being carried out, then the most appropriate analysis should be selected for each comparison. The problem with conducting a MAIC is that the relative effects obtained will be reflective of the ELA trial population (ELATIVE RCT) which differ in effect modifiers from the RESPONSE population, hence the need for adjustments. We therefore do not know what the comparative effectiveness would be in the RESPONSE population. So cost-effectiveness results derived for this comparison will apply to patients like those included in the ELATIVE RCT, and for outcomes defined as in the ELATIVE RCT.

Comments

I think the outcome recalculation issue leads to a requirement for separate analyses. Having accepted that, the most appropriate analysis for each comparison should be carried out. The issue of MAIC vs NMA for the SEL vs ELA comparison will depend on whether the committee is satisfied that adjustment for population differences is required and whether the results are suitable for decision making (given issues with ESS previously noted by the EAG and the estimates not necessarily applying to the SEL trial population). I have not reviewed the MAIC results in detail.

3 RELIABILITY OF NMA RESULTS (EFFICACY OUTCOMES)

The company has now provided an ITC report which includes details of the data and code used for each analysis. At ACM1 we questioned the validity of results due to the very wide 95% CrIs presented, particularly for the outcome “ALP normalisation at 12 months”. We also questioned the empirically informed prior distributions used (Turner priors) as they are specific to OR and the analyses appeared to be carried out on the RR or mean difference scales (Slide 25, ACM1 Part 1).

3.1 Wide intervals for “ALP normalisation at 12 months”

Table 16 of the updated ITC Report shows that no patients achieved ALP normalisation in the placebo arm of ELATIVE or RESPONSE. Table 27 of the updated ITC Report also

shows that no patients achieved ALP normalisation in the placebo arm of POISE. This means that the networks connecting SEL to ELA or OCA for this outcome are only linked by a zero event study, so the connection via placebo is very weak since there is not much information on the relative effects of SEL, ELA or OCA compared to placebo for this outcome. This may be because the RCTs were too small to detect ALP normalisation in the placebo arm (if it happens only to a small proportion of patients), or it may be that for the populations included in these trials, ALP normalisation is not possible on placebo and only a head to head study would be able to provide information of the relative effectiveness of the active therapies (see related comments on this in Section 5).

A similar issue was discussed at length in TA853 – see slides 11 and 12 from ACM2 of TA853 and slides 25-32 of ACM1.

The company note at the end of section 4.3.2 of the updated ITC report that "A continuity correction was applied to account for zero values by adding 0.5 to all treatment arms of studies included for both MAIC as well as Bayesian NMA". However, the values provided in the data tables do not include this correction and the specification of the correction in the quoted sentence is not enough to evaluate whether it was applied correctly. The usual continuity correction requires adding 0.5 to the results for all arms of the study, but it also requires adding 1 to the total number of patients in each arm of the study. It is unclear whether this was done. As noted in TA853 and in NICE DSU TSD2, when networks are only connected via a zero cell, a correction to the zero cell value is required. Different corrections can be used and applied in different ways and this was discussed at length in TA853, but in this appraisal the usual 0.5 continuity correction seems appropriate. (In TA853 issues relating to imbalance in trial sample sizes made other corrections more relevant but this does not apply here). However, even using this correction, comparisons with placebo will still be very uncertain, which explains the wide CrIs for all relative effects for this outcome.

Comments

Overall, this is a problem of trial design which cannot be resolved – I don't think the RCTs presented here can provide relative effectiveness estimates for this outcome (either they are too small, or ALP normalisation cannot be achieved on placebo and head-to-head studies are needed). This means that results for this outcome will be very uncertain regardless of analysis method (NMA or MAIC), type of continuity correction applied or prior distributions.

3.2 Prior distributions for the heterogeneity parameter

In the company's response to draft guidance a justification for the Turner priors is provided in Section 2.4. This explanation is not convincing. However, from the additional details

provided in the updated ITC Report, we can see that the Turner priors are justified. This is because, although the company presented the results of the NMAs as RR for dichotomous outcomes, the analyses are actually carried out on the OR scale, for which the quoted Turner priors are appropriate.

The updated ITC report also clarifies that the prior distributions used for continuous outcomes are derived from Rhodes et al., which is also appropriate.

The estimates of the heterogeneity are provided in the spreadsheets attached to the updated ITC Report (Appendix F). However, these do not show the 95% CrIs for the estimates, which I suspect are wide given the lack of data and reliance on the prior distributions for estimation of the between-study heterogeneity. Point estimates are sensitive to the choice of prior distribution.

Comments

The selected Turner/Rhodes prior distributions are appropriate.

3.2.1 Reporting of results on the RR scale

However, because the analysis is carried out on the OR scale, assumptions on the baseline risk (or probability of an event on the control/placebo arm) are required to then transform the relative effects from the log-OR scale to the log-RR scale. This baseline risk was calculated as the average probability of an event across the placebo arms of studies included in that NMA (as per NMA code provided). It is unclear why this was selected as a meaningful estimate of the probability of an event on placebo in the population of interest. NICE DSU TSD5 explains that evidence used to estimate this baseline risk should be “as specific as possible to the population of interest” (TSD5 Section 2.1) and specifically states that the evidence used for this should be justified. As this baseline risk is also used to estimate the absolute effects it can have an impact on the cost-effectiveness results (**NOTE:** I did not have time to check how results from the NMA are used in the model, are the ORs or RRs used directly, or the absolute probabilities calculated from the NMA model?). Depending on how results are used in the model, it may make more sense to use the probability of events from the RESPONSE trial, rather than an average across all trials. However, for ALP normalisation this would be zero, so an alternative source may be required. This also means that RRs and absolute probabilities from sensitivity analyses that use different studies (e.g. NMA will all 3 studies and NMA including additional OCA studies) will have used a different estimate of baseline risk, which makes comparisons across analyses confusing.

Similarly for continuous outcomes, the company use an average mean on the placebo arms of the included studies to calculate the mean on placebo. It is unclear whether this is appropriate.

Comments

This issue of the baseline probabilities/mean is probably a minor point, although the health economists may wish to consider how this affects absolute probabilities and how results are used in the model.

4 DISCONTINUATION RATES FROM ITC

In section 3.1 of the company's response to draft guidance the company propose that RWE of discontinuation for OCA (based on Abbas et al) should be used and ITC results should be used to derive the discontinuation rates for SEL and ELA. The EAG had proposed to use the discontinuation rates from the RESPONSE trial and to apply the ITC results for ELA and OCA to those instead.

I won't comment on which source of data is best to anchor the rates on, this should probably be discussed with clinical experts, or the EAG may have a view on the suitability of Abbas et al for this decision problem. In principle, either method is fine. The best evidence source should be used to anchor the discontinuation rates (for a given intervention), and the relative effects estimated from the ITC can be applied to obtain the discontinuation rates for the other interventions.

I will comment on the discontinuation ITC.

The company note that "there is no clinical rationale for the discontinuation rates of seladelpar and elafibranor to differ" (company response to draft guidance, section 3.1). They therefore propose to use a simple average of the estimated discontinuation rates for SEL and ELA in the model. This is inappropriate, as this simple average does not take into account the differences in precision of the two estimates. It is also unclear how the uncertainty around this average is calculated for use in the PSA (if it is indeed varied in the PSA). If the company believes, and can justify, that there are no differences in discontinuation rates between ELA and SEL, then the most appropriate synthesis of the evidence is one that includes this assumption up front, so a NMA that assumes the effect of SEL and ELA are the same (in practice this would mean coding ELA and SEL as the same treatment in an NMA also including placebo and OCA). Results from this simplified NMA would appropriately weight results by study size and number of events and would allow

appropriate estimation of the uncertainty in the discontinuation rates (under the assumption that they are the same for ELA and SEL). */

However, the company consider that a MAIC is required to compare the SEL to ELA for discontinuation rates due to the required adjustment for effect modifiers, and a separate NMA is required for the comparison to OCA. I am not sure if the arguments for important imbalances in effect modifiers for the effectiveness outcomes also apply to discontinuation, i.e. it is unclear whether the differences in discontinuation rates across studies would be affected by the same variables identified as effect modifiers for the effectiveness outcomes. However, if we are to accept that ELA and SEL have the same discontinuation rates (as argued by the company), I think it would still make sense to just hard-code that assumption into an NMA that also includes OCA, or to use the estimates from SEL and assume they also apply to ELA. Any of these options seems preferable to taking a simple average as proposed in the company's response to draft guidance (Table 9).

I note that I could not find where the values of 10.4% and 6.4% in Table 9 come from, so I cannot comment on whether they were calculated appropriately.

Comment

I am not sure if this is an important issue in terms of cost-effectiveness, but the company's proposition does not seem reasonable, as it fails to appropriately account for the available data and the uncertainty in the discontinuation rates.

5 FEASIBILITY ASSESSMENT AND ITC WITH FIBRATES

The company provided a document with details of the systematic literature review and feasibility assessment for indirect comparisons of SEL with fibrates: "Data on File – Fibrates Feasibility Assessment and ITC Report.docx."

Li 2022 was the only study identified as possibly being suitable for a comparison with SEL based on comparability and sufficient reporting of outcomes and study inclusion criteria. Li 2022 compared fenofibrate 200mg + UDCA to UDCA alone.

The company notes some differences in the populations included in Li 2022 compared to RESPONSE, but argue that a MAIC is not feasible as there is inadequate reporting of important covariates for population adjustment.

The company present results of a Bayesian NMA for the outcome of ALP normalisation which is the only key outcome for which comparable outcome definitions and sufficient detailed reporting were available (other outcomes are reported in appendices). Assuming the

code used for the NMAs presented here is the same as for the other NMAs, it was appropriate.

Data are presented in Table 6 of the fibrates feasibility assessment and ITC report document. Because there are no events observed in the placebo arm of RESPONSE, the same issues with uncertainty in results will apply as discussed in Section 2.1. However, I note that one event was observed in the UDCA alone arm of Li 2022, which suggests that there is a 4.2% probability of ALP normalisation on UDCA alone for the population in this study (Table 6). Table 6 also shows that the probability of ALP normalisation in Li 2022 is 54.2% which is much higher that what was observed in the other studies. This suggests that the included populations are very different across these studies and supports the company's assessment that an NMA may not be appropriate.

In their response to draft guidance the company argue that adverse events of fibrates are not being taken into account by the committee. An ITC with fibrates could take these into account if data were available for an NMA of e.g. discontinuation rates or adverse events to compare SEL with fibrates. This was not discussed by the company in the feasibility assessment for the ITC with fibrates.

Seladelpar for previously treated primary biliary cholangitis [ID6429]

EAG Review of company and other stakeholders' response to Draft Guidance

Produced by	Peninsula Technology Assessment Group (PenTAG) University of Exeter Medical School
Authors	Caroline Farmer[1] Jemma Perks[1] Alan Lovell[1] Maxwell S. Barnish[1] Valdemar Santo[1] Ahmed Abdelsabour[1] Dawn Lee[1] Taha Khan[2] Rebecca Jones[3] Gwilym Webb[4] Edward CF Wilson[1] [1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter [2] St Leonard's Practice, Exeter [3] Leeds Teaching Hospital NHS Trust [4] Cambridge University Hospitals NHS Foundation Trust
Correspondence to	Caroline Farmer 3.09 South Cloisters, St Luke's Campus, Heavitree Road, Exeter, EX1 2LU; c.farmer@exeter.ac.uk
Date completed	02/09/2025
Source of funding	This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR175575.

1. INTRODUCTION

On 9 July 2025, the NICE committee met to discuss the evidence submission for seladelpar for the treatment of primary biliary cholangitis (PBC), including pruritus, in adults. Seladelpar would either be used in combination with ursodeoxycholic acid (UDCA), if the primary biliary cholangitis had not responded well enough to UDCA, or alone, if UDCA could not be tolerated. The NICE committee made the decision not to recommend seladelpar. Within the NICE draft guidance, the committee identified the following key issues:

- That fibrates were potentially a comparator treatment for people with PBC with pruritus in NHS clinical practice and it would be informative to see a comparison of benefits and costs of seladelpar compared with fibrates.
- It was not appropriate for the company to produce separate methods for comparing seladelpar with its comparators and that a Bayesian NMA should be used for all indirect comparisons.
- The Bayesian NMA produced wide credible intervals and results that were considered to differ unexpectedly from the original trial results. The committee did not have confidence in the results and asked for:
    - The company to submit its NMA code with scenarios presented to explore uncertainty
    - Updated reporting of the NMA, including clear reporting of the original trial values and sources of numerical data for the comparators
    - An explanation of how the trials and trial populations differed, with particular reference to effect modifiers
    - An explanation of any difference between the reported trial results and the results of the NMA
    - Consideration that Turner priors are suitable for odds ratios (ORs) and would need careful justification to be used to calculate relative risks (RRs).
- The EAG's preferred approach should be used to model treatment discontinuation in the economic model, with the caveat that the 12-month rates derived from the indirect treatment comparisons were uncertain.
- The EAG's approach to modelling utility was preferred, but that there was an uncertainty about whether disutility associated with pruritus was underestimated and the committee was uncertain whether the model reflected the expect quality of life in the ALP health

states. Further evidence and rationale for the use of utility assumptions in the model was needed. This included the need for:

- A distribution of change from baseline on the pruritus scale in each treatment arm, including the proportion of people achieving a clinically meaningful improvement, to better inform the utility values and QALY gains associated with improvements in pruritus.
- A commentary on how well the model reflects quality of life associated with PBC, with any supportive evidence from literature.

The purpose of this document is to provide the EAG appraisal of the additional evidence provided by the company in response to the draft guidance. In its response, the company included a revised patient access scheme (PAS) discount for seladelpar. Updated economic analyses are presented based upon the company's revised PAS.

An additional cPAS appendix has been produced for Committee members to be viewed alongside this document. It contains updated economic analyses with all relevant price discounts applied.

2. EAG RESPONSE

2.1. Fibrates as a comparator to seladelpar

In its response to ACD, the company maintained its position that fibrates were not a comparator to seladelpar on the basis that:

- There was no clear evidence that fibrates were part of established clinical practice for the target population in the NHS
- That the committee had not considered the safety issues associated with fibrates

In addition, the company stated that an indirect treatment comparison (ITC) to compare seladelpar with fibrates was not feasible.

As part of its response, the company provided some qualitative data it had gathered from consultations with five clinicians to address the issue of whether fibrates were used as a second-line treatment for PBC. They also submitted a feasibility assessment it had conducted to determine that an ITC would not be feasible. The EAG provides a critique of this evidence in Sections 2.1.1 and 2.1.3 below. Consideration of the company's claim related to the safety considerations of fibrates is discussed in Section 2.1.2.

Finally, the company raised concerns that the revision to the scope of the appraisal to include fibrates was unfair and created inconsistency with previous appraisals of treatments for PBC (i.e. TA1016 NICE appraisal of elafibranor[1]). The EAG noted this consideration in its original report but considered that this issue related to the application of NICE processes and therefore was for the NICE committee to consider.

2.1.1. The role of fibrates for treating PBC in the NHS

In Appendix A of the company's response, the company provided quotes from clinicians' responses to two questions posed by the company:

- "What are the patient characteristics that would make you decide to use or recommend a fibrate, in place of an approved second line therapy such as Obeticholic acid or Elafibranor for a patient diagnosed with PBC who has been failed by UDCA?"
- "If you prescribe for PBC patients where ALP>1.67xULN and you are choosing between OCA, Elafibranor and fibrates, out of every 100 patients, how many would you likely

prescribe fibrates for, where the alternative treatment option could be OCA or elafibranor?"

It was not explicitly stated that clinicians' responses to the questions were produced in full. The company did not provide any information about how they recruited the clinical experts consulted and no baseline characteristics (e.g. geographical area, seniority, length of experience treating people with PBC) were provided, although all five were stated to be based within a 'tertiary PBC treatment centre'. The company stated that further details about the experts consulted could be shared upon request, but this was not possible within the time frame of the EAG's critique. This information may be useful for the NICE committee.

A summary of the responses from the clinicians consulted by the company is provided in Table 1. Overall, the EAG interpreted the responses to suggest that:

- Fibrates were previously used as a 2nd line treatment before the availability of elafibranor, which has now displaced their use at least to some extent
- Some clinicians in the NHS may be reluctant to use fibrates as a treatment for PBC due to concerns about using these off-label
- Fibrates may still be used in NHS practice to treat specific populations with PBC. This could include those who don't meet the threshold of >1.67xULN but experience pruritus or are likely to require treatment long-term, or those for whom OCA is contraindicated.

Two stakeholders to the appraisal (British Association for the Study of the Liver [BASL] and the British Hepatology Pharmacist Group) also raised concerns about the use of fibrates to treat PBC given that they were not licensed for this purpose.

On the basis of its consultation with its experts, the company raised several key uncertainties related to whether fibrates were a comparator treatment to seladelpar, including:

- whether fibrates were established practice in the NHS as a 2nd line treatment option
- whether this was in the target population (i.e. those who had experienced an inadequate response to UDCA [i.e. >1.67xULN])
- and whether when used, these were prescribed by multidisciplinary teams based in specialist treatment centres.

The EAG consulted its own clinical experts on the use of fibrates in clinical practice, responses are shown in Table 2. The EAG asked its experts the same two questions that the company

asked its own experts to understand prescribing patterns for fibrates and also asked additional questions to address the uncertainties raised by the company. Two experts provided treatment to people with PBC within multidisciplinary teams in specialist units and one expert treated people with PBC within primary care. Overall, there was a difference of opinion between experts who prescribe treatments for PBC in specialist units – one expert stated that fibrates were the preferred treatment option in the target population at 2nd line while the other expert stated that their centre would rarely use fibrates and prefer to use licensed treatments. The third expert who treats people with PBC in primary care stated that half of people in the target population may receive fibrates but stated that practice varies across centres.

Overall, based on the evidence available to the EAG at the time of appraisal, the EAG considered the following:

- **were fibrates used to treat PBC in the target population in the NHS?**
  - YES. At least in some centres this was the case (supported by 2 EAG experts, 1 company expert). The Abbas 2024[17] audit of NHS practice also supported this.
- **were fibrates a treatment for PBC rather than only for itch in people with PBC?**
  - YES. Fibrates are a PPAR agonist, similar to seladelpar and elafibranor. Studies evaluating fibrates showed that treatment with fibrates led to improvements in disease outcomes, such as ALP response.
- **were fibrates established practice for treating PBC?**
  - UNCLEAR/POSSIBLY. Fibrates were established practice for treating PBC in the target population in at least one specialist unit in the NHS (as stated by 1 EAG expert). The Abbas 2024 audit[17] suggested that 50% of people with PBC in the NHS received fibrates at 2nd line, which suggested that fibrates were established practice. The company stated that the population included in this 50% was not clear from the Abbas publication, which the EAG agreed with, though it considered it plausible that a meaningful number of the 50% might represent the target population of those who had an inadequate response to UDCA or who were in tolerant. Feedback from the company experts, EAG experts and appraisal stakeholders suggested that practice varies widely across areas of the NHS, which at least in part may be due to varying views on the use of unlicensed treatments. This meant that national audit data, such as the Abbas et al. 2024 study, may likely be a better representation of clinical practice than consultation

with a smaller group of experts/centres who may not be representative of all practices in the NHS.

- **were fibrates still used as a 2nd line treatment following the introduction of elafibranor in clinical practice?**
  - UNCLEAR. Experience with elafibranor was limited at the time of the appraisal and prescribing patterns were still evolving. Three company experts stated that the introduction of elafibranor had, to some extent, displaced the use of fibrates. One EAG expert suggested that fibrates may still be the preferred first line choice, though this may change in time. Given some hesitation from experts on the use of unlicensed treatments, the EAG considered it plausible that those uneasy with the use of using fibrates off-label may consider using elafibranor at 2nd line. However, at least at present, the EAG considered that this was uncertain and fibrates may still be being used as a 2nd line treatment option.

Table 1: Summary of company's clinical expert consultation

	Clinician #1	Clinician #2	Clinician #3	Clinician #4	Clinician #5
What are the patient characteristics that would make you decide to use or recommend a fibrate, in place of an approved second line therapy such as obeticholic acid or elafibranor for a patient who has been failed by UDCA?	Would not be used as a 2nd line therapy. Only used after inadequate response or intolerance to OCA and/or elafibranor. Possible use as a 3rd or 4th line treatment if inadequate response to UDCA +/- OCA or elafibranor	Prior to introduction of elafibranor, would have prescribed fibrates as a 2nd line treatment option. Now would use elafibranor in 2nd line except possibly if someone was receiving statins.	Prior to introduction of elafibranor, would have prescribed fibrates as a 2nd line treatment option. Now would use fibrates with people who don't quite make the ALP 1.67xULN cut-off but are young and high risk and those who are likely to need triple therapy.	Unclear. Primarily considered to treat itch. Not typically considered a 2nd line treatment but may prescribe fibrates where OCA not indicated for cirrhosis, although not universal. May be used in combination with OCA.	Primarily considered to treat itch. May be used 2nd line in people with low risk disease, no strong need for 2nd line therapy but does have significant pruritus.
If you prescribe for PBC patients where ALP>1.67xULN and you are choosing between OCA, elafibranor and fibrates, out of every 100 patients, how many would you likely prescribe fibrates for, where the alternative treatment option could be OCA or elafibranor?	Zero	Not estimable but few people due to concerns about fibrates being used off-label.	5%	No response	≤5%

Abbreviations: ALP, alkaline phosphatase; OCA, obeticholic acid; UDCA, ursodeoxycholic acid; ULN, upper limit of normal

Table 2: Summary of EAG's clinical expert consultation

	Clinician #1 (specialist unit)	Clinician #2 (specialist unit)	Clinician #3 (primary care)
Can you please confirm whether your centre will prescribe fibrates as a 2nd line treatment for people who have experienced an inadequate response to UDCA (i.e. ALP >1.67xULN)?	Yes	No, because they are not licensed for use in PBC and there are licensed alternatives so our practice is to use the licenced drugs. Only in selected cases – personalised medicine – under MDT discussion - might they be advised pre-OCA/elafibranor now.	Yes, potentially - as a GP, I have patients with PBC who have been started on fibrates as a second-line treatment.
What are the patient characteristics that would make you decide to use or recommend a fibrate, in place of an approved second line therapy such as obeticholic acid or elafibranor for a patient who has been failed by UDCA?	We use fibrate (usually bezafibrate, sometimes fenofibrate) first line for all patients who have failed UDCA therapy unless there is a contraindication to standard fibrate therapy.	There is no guideline to guide this so it's a very personalised approach, discussed carefully at our MDT meeting. As above, we would generally look at the licenced options first always. Being careful to consider whether patient has pruritus, with normal fibrosis assessment and low likelihood despite ALP of having a liver related event due to PBC before other comorbidities take effect first might lead us down a fibrate first route. The UK-PBC or global PBC scores can be helpful in that setting. In that setting a patient may benefit from trying	I will preface this by saying I do not really get involved with initiating second-line treatment of PBC. Nevertheless, a patient would be more likely to get a fibrate if they were scared about side effects or had polypharmacy or other co-morbidities because fibrates have been around a lot longer and we know they are quite safe whereas OCA is contraindicated in advanced fibrosis/cirrhosis a patient, if they already had dyslipidaemia again a fibrate would be preferred as getting 2 birds with one stone.

	Clinician #1 (specialist unit)	Clinician #2 (specialist unit)	Clinician #3 (primary care)
		bezafibrate. It's a small number of people.	
If you prescribe for PBC patients where ALP>1.67xULN and you are choosing between OCA, elafibranor and fibrates, out of every 100 patients, how many would you likely prescribe fibrates for, where the alternative treatment option could be OCA or elafibranor?	If all options are open, 100 out of 100 would receive fibrate first line at present (with different choices made when there are particular individual circumstances). Prior to about 2 years ago, we were giving some patients OCA first line but the MDT had a formal discussion and we agreed to move to fibrate first line for all unless there were additional considerations.	It's a small number of people – we generally wouldn't unless it was an adjunct to OCA or patient was itchy.	Not sure. If I had to guess, it would be 50:50 ish
Now that elafibranor is available, has this changed the way you would use fibrates to treat PBC?	Not for initial "second line" therapy but yes in relation to therapy after initial second line. We have moved a very small number of patients to elafibranor who have not tolerated standard fibrates or who have not had adequate response to standard fibrates. We have had discussions with patients about moving from 'unlicensed' to 'licensed' therapy even when they attained response on fibrate but	Because it has some data on pruritus, yes we are looking at this for patients with pruritus, or who haven't responded to OCA/OCA has caused itch so can't be used. Its only recently been approved so our experience is a little limited	I suspect fewer patients will now get fibrates as elafibranor doesn't cause pruritus as much and also improves ALP

	Clinician #1 (specialist unit)	**Clinician #2 (specialist unit)**	**Clinician #3 (primary care)**
	have not -- to my knowledge -- changed agent based on this.		

Abbreviations: ALP, alkaline phosphatase; MDT, multi-disciplinary team; OCA, obeticholic acid; UDCA, ursodeoxycholic acid; ULN, upper limit of normal

2.1.2. Safety considerations related to fibrates

The company stated that the NICE committee had not taken into consideration safety issues associated with fibrates in its request to include fibrates as a comparator to seladelpar. It described some concerns raised by its clinical experts and cited a clinical practice guideline from the European Association for the Study of the Liver (EASL)[2] as support for fibrates being associated with safety concerns.

In order to demonstrate the safety of fibrates for consideration by the NICE committee, the EAG considered that the company should have presented safety data for fibrates from its systematic literature review (SLR) that was presented in its original submission (and where studies evaluating fibrates in the target population were eligible for inclusion). In order to understand the relative safety of fibrates compared to seladelpar, the EAG considered that an indirect treatment comparison may be required.

2.1.3. Clinical effectiveness versus fibrates

The company stated that they conducted a feasibility assessment to assess the feasibility of conducting an indirect treatment comparison (ITC) to compare seladelpar and fibrates. The company's SLR identified nine studies evaluating fibrates for the treatment of PBC.

Consistent with the SLR reported in the company submission, the company limited the inclusion of studies to those reporting a 12-month follow-up. In its original report, the EAG noted that the choice of a 12-month limit on study inclusion was not justified and led to the exclusion of studies that were relevant to the decision problem for this appraisal. Notably, as the treatment effect of seladelpar was shown to occur within one month of treatment, after which there were no further improvements in treatment response, the EAG considered that an earlier follow-up duration would have been relevant for inclusion and so by limiting the inclusion criteria to a 12-month follow-up, the company had not fully considered the feasibility of an ITC comparing seladelpar and fibrates.

The company appraised three studies with 12-month follow-up for consideration in the ITC: Li et al. (2022),[3] BEZURSO (2018)[4] and Iwasaki et al. (2008).[5] Of these, two (BEZURSO and Iwasaki) were excluded on the basis that they did not report the relevant outcomes consistently with the RESPONSE trial.[6]

Li 2022[3] was subject to further consideration for feasibility for NMA, from which the company concluded the following:

- The studies differed in the participant eligibility criteria for ALP levels at baseline:
    - The RESPONSE trial[6] requires ALP levels ≥ 1.67× the upper limit of normal (ULN), indicating a population with higher ALP levels
    - The Li 2022[3] study used an inclusion threshold of ≥ 1× ULN, encompassing a wider spectrum of ALP levels
- The studies differed in average baseline ALP levels:
    - RESPONSE: mean ALP 314.2, representing participants with higher baseline ALP levels
    - Li 2022: mean ALP 197.5, suggesting a population with lower ALP levels compared to RESPONSE
- The studies varied in age, which was an effect modifier
    - RESPONSE: 56.8 years, capped at 75 years
    - Li 2022: 50.9 years, indicating a younger population with a narrower age range capped at 65 years
- The studies varied in the proportion of participants with cirrhosis, which was an effect modifier:
    - RESPONSE: Included 14% of participants with baseline cirrhosis, reflecting a mixed population with both cirrhotic and non-cirrhotic patients
    - Li 2022: Specifically excluded patients with cirrhosis, focusing solely on non-cirrhotic PBC patients, ensuring a more homogeneous population with less advanced liver disease
- The mean total bilirubin levels (effect modifier) were almost similar across the two trials

While the EAG considered that the company could have made some adjustments to the data from RESPONSE[6] to match the characteristics of the Li 2022[3] study, this would not have been possible for baseline ALP levels, as the Li study included a broader population. Overall, the EAG agreed with the company that differences in the design of studies would have made an ITC with 12-month follow-up data challenging.

The EAG conducted an appraisal of all studies that evaluated fibrates to determine whether an ITC comparing seladelpar and fibrates may have been feasible if other follow-up durations were included. Within the time available to the EAG, this appraisal focused only on the outcomes reported by the studies and did not take into consideration other study design characteristics (e.g. population characteristics, study quality). This appraisal identified inconsistencies between studies in the definition, follow-up and analysis of outcomes that would likely violate the transitivity assumption in an ITC if recalculation of outcomes from RESPONSE[6] was not feasible (Table 2). On that basis, even without consideration of other study design characteristics, the EAG concluded that it may not have been feasible to conduct an ITC to compare seladelpar and fibrates.

If an ITC was not feasible, then in order to present the committee with some evidence to compare the clinical effectiveness of fibrates, the EAG considered that the company could have presented the primary results from trials that had evaluated fibrates and conducted a narrative synthesis. This may have been able to identify (a) the best quality evidence for fibrates, (b) patterns in the results across timepoints, outcome definitions and different study populations, and (c) a plausible conclusion about the likely effectiveness of fibrates as compared with UDCA. While the comparability of these data with seladelpar would have been highly uncertain, it may have informed some exploratory analysis as to the cost effectiveness of seladelpar versus fibrates.

Table 3: Outcomes and follow-up duration in studies evaluating fibrates

	Comparison	Treatment duration (months)	Outcomes of interest	EAG comment
RESPONSE[6]	Seladelpar + UDCA vs UDCA	Up to 12	• Proportion of patients achieving composite response (ALP < 1.67× ULN; ≥ 15% decrease in ALP and total bilirubin ≤ 1.0× ULN) • ALP normalisation (≤ 1.0) • ALP response • Change in ALP • Pruritus intensity • Bilirubin change from baseline	NA
Li 2022[3]	Fenofibrate 200 mg + UDCA vs. UDCA	12	• Proportion of patients achieving normalisation of ALP, γ-GT, and TBil levels (<1.0xULN) • Change in ALP • Change in bilirubin	ALP normalisation appeared comparable to RESPONSE
BEZURSO[4]	Bezafibrate 400 mg + UDCA vs. UDCA	24; some outcomes available at 3 and 12 months	• Proportion of patients achieving composite response (normalisation of ALP, ALT, AST, bilirubin, albumin, normal prothrombin index) • ALP response (< 1.67× ULN using various criteria, including Toronto) • Change in ALP • Change in bilirubin • Pruritus intensity	Difficult to include in an ITC due to the reporting of medians (IQR) exclusively for continuous outcomes of interest. Dichotomous outcomes were reported at 24 months and not available at 12 months. The EAG was unable to find 12-month data for these outcomes, though it's possible these could be obtained from the study authors.

	Comparison	Treatment duration (months)	Outcomes of interest	EAG comment
Iwasaki 2008[5]	Bezafibrate 400 mg + UDCA vs. UDCA	12	• ALP normalisation (unclear, but possibly ≤1.5xULN) • Change in ALP • Change in bilirubin	Comparison with ALP normalisation would require recalculation of RESPONSE data to use the same criteria. Continuous outcomes were reported as median (IQR) only.
Hosonuma 2015[7]	Bezafibrate 400 mg + UDCA vs. UDCA	96	• Change in ALP • Change in total bilirubin	No shorter follow-up data reported. Data only reported in figures.
Itakura 2004[8]	Bezafibrate 400 mg + UDCA vs UDCA	6	• Change in ALP • Change in bilirubin	Data reported as mean and SD, 6-month timepoint only
Kanda 2003[9] (labelled Tatsuo 2003 in the company assessment)	Bezafibrate 400 mg + UDCA vs UDCA	6	• Change in ALP	Small study. Change in ALP reported as baseline and final value or % change.
Study 213 (Bonder 2024[10])	Bezafibrate 200 mg + OCA + UDCA vs. Bezafibrate 400 mg + OCA + UDCA vs. Bezafibrate 200	3	• Composite endpoint (ALP < 1.67× ULN; ≥ 15% decrease in ALP and total bilirubin ≤ 1.0× ULN)	Conference abstract

	Comparison	Treatment duration (months)	Outcomes of interest	EAG comment
	mg + UDCA vs. Bezafibrate 400 mg + UDCA			
Jones 2024 (labelled as 'Study 214 (Levy 2024)' in the company report)	Bezafibrate 100 mg + OCA + UDCA vs. Bezafibrate 400 mg + OCA + UDCA vs. Bezafibrate 100 mg + UDCA vs. Bezafibrate 400 mg + UDCA	3	• Percentage change in ALP • Proportion who achieved ALP ≤1.0xULN • ≥40% reduction in ALP • Percentage change in bilirubin • Proportion who achieved total bilirubin ≤0.6xULN	ALP normalisation could be comparable at 3 months. Conference report. Further follow-up data expected.
Liberopoulos 2010[11]	Fenofibrate 200 mg + UDCA vs. UDCA	2	• Change in ALP	Data reported as mean and SD, 6-month timepoint only

Abbreviations: ALP, alkaline phosphatase; γ-GT, gamma-glutamyl transferase; TBil, total bilirubin; UDCA, ursodeoxycholic acid; ULN, upper limit of normal

Note: studies listed were those identified by the company in their SLR

2.1.4. Cost-effectiveness versus fibrates

The EAG considered producing threshold analysis to inform how much more effective seladelpar would need to be to be cost-effective versus fibrates, however, this was not possible within the Excel model supplied by the company for a number of technical reasons: firstly there were multiple endpoints informing effectiveness and secondly the model requires a macro to be run for calibration to produce the results.

As noted in the EAG report, however, if fibrates were assumed of at least equal effectiveness to OCA, then they were likely to represent a much more cost-effective treatment for PBC than either OCA, elafibranor or seladelpar, given the list price for 100x200mg bezafibrate is approximately £8.63 (BNF 2025) and fenofibrate 28x160mg is £2.87 (Drug Tariff), whereas list prices for the OCA (30x5mg), elafibranor (30x80mg) and seladelpar (30x10mg) are £2,384, £2,867 and £3,155 respectively (BNF 2025); a 600 to 1,000 fold difference in list price.

2.1.5. Conclusions

Overall, while the views of clinical experts and stakeholders to the appraisal suggested that the use of fibrates varies widely across the NHS, on balance the EAG maintained its view that fibrates were a relevant comparator to seladelpar. This was based on data suggesting that up to 50% of people with PBC receive fibrates as 2nd line treatment from the Abbas et al [17] audit, which the EAG expected to include a reasonable sized population of people in the target population who would be eligible for seladelpar. The EAG also noted that fibrates are also PPAR agonists, like seladelpar and elafibranor, and that studies had reported positive effects of fibrates on disease outcomes.

Overall, it was not possible to reach a conclusion about the clinical and cost effectiveness of seladelpar as compared with fibrates based on the information provided by the company in its response. To some extent, the feasibility of doing this was limited by the evidence base for fibrates and seladelpar, which was not homogeneous in study design characteristics and so would make any indirect comparison challenging. However, the EAG considered that further evidence and analysis could have been presented by the company to explore this, though this evidence would have almost certainly been highly uncertain. Given the significantly reduced costs of fibrates, the EAG expected that seladelpar would likely not be cost effective in comparison with fibrates, though was unable to explore this through scenario analysis due to constraints in the company model.

2.2. ITC including all comparators

Overall, the NICE committee stated in the draft guidance that they preferred the Bayesian NMA over the anchored MAIC. The committee had a number of concerns with the validity of the analysis produced including that the naïve trial results showed that seladelpar had a smaller estimated treatment effect than the estimate for OCA or elafibranor but in the Bayesian NMA it had a larger treatment effect. The Committee asked for:

- the company to submit its Bayesian NMA model code with scenarios presented in order to explore uncertainty in the model
- updated reporting of the company's Bayesian NMA, including clear reporting of the original trial values and sources of numerical data for the comparators
- an explanation of how the trials and trial populations differed, with particular reference to effect modifiers
- an explanation of any difference between the reported trial results and the relative effects resulting from the Bayesian NMA
- consideration that Turner priors are suitable for odds ratios and would need careful justification to be used for an NMA of relative risks.

2.2.1. NMA Code

The company submitted their Bayesian model NMA code (which was written originally in R) in a word file stripped of the information of which libraries were used and without the input files used to run the NMA (input data were also submitted in a word file). This limited the EAG's ability to review the code and meant that analyses could not be reproduced as part of quality checking. No obvious errors were identified in cursory code review or in review of the input data in the word file.

2.2.2. Updated reporting

The company supplied an updated NMA report and a feasibility assessment report for the comparison to fibrates. Table 3 and Table 4 show the input data used for ALP normalisation and ALP response using the Toronto I for each of the treatments. Data for RESPONSE was recalculated using the ELATIVE and POISE ALP ULN cut-offs for comparison to elafibranor +

UDCA and OCA + UCDA, respectively. Calculations were not supplied to the EAG to allow these analyses to be checked. The EAG was able to match the other information provided for ALP normalisation and response according to the Toronto I criteria back to the source papers.

Table 4: Input data for ALP normalisation (ALP ≤ 1.0× ULN) at 12 months (used in model to inform transition to ALP normalisation)

Study Name	Treatment	N	n	%	RR (95%CI)*
RESPONSE[6]	Seladelpar 10 mg + UDCA	128	32	25	16.500 (2.305, 118.088)
	Placebo + UDCA	65	0	0	
RESPONSE (ELATIVE matched ALP ULN cut-off)	Seladelpar 10 mg + UDCA	[redacted]	[redacted]	[redacted]	[redacted]
	Placebo + UDCA	[redacted]	[redacted]	[redacted]	[redacted]
RESPONSE (POISE matched ALP ULN cut-off)	Seladelpar 10 mg + UDCA	[redacted]	[redacted]	[redacted]	[redacted]
	Placebo + UDCA	[redacted]	[redacted]	[redacted]	
POISE[12] (Table 7 CADTH report)	OCA 5-10 mg + UDCA	70	1	1.4	1.057 (0.067, 16.578)
	OCA 10 mg + UDCA	73	5	6.9	5.068 (0.607, 42.334)
	Placebo + UDCA	73	0	0	
Li 2022[3]	200 mg/day fenofibrate + UDCA	24	13	54.2	13.542 (1.916, 95.687)
	UDCA	24	1	4.2	
ELATIVE[13] (Table 6 FDA label)	Elafibranor 80 mg + UDCA	108	16	15	8.000 (1.090, 58.741)
	Placebo + UDCA	53	0	0	

Abbreviations: ALP: Alkaline phosphatase; CADTH: Canadian Agency for Drugs and Technologies in Health; FDA: Federal Drugs Administration; N: total number in the arm in the study; n: number of events observed; OCA: Obeticholic acid; UDCA: Ursodeoxycholic Acid; ULN: Upper limit of normal

Note: *crude risk ratio calculated based on event rates. Calculation was performed using https://www.gigacalculator.com/calculators/relative-risk-calculator.php. It required a zero cell correction of >=1 unit, therefore a 1 correction was added to zero event arms

Table 5: Input data for ALP response (Toronto I: ALP ≤1.67 × ULN) at 12 months (used in model to inform transition to mild ALP elevation)

Study Name	Treatment	N	n	%	RR (95%CI)*
RESPONSE[6]	Seladelpar 10 mg	128	84	65.6	2.509 (1.637, 3.847)
	Placebo	65	17	26.2	
RESPONSE (ELATIVE	Seladelpar 10 mg	[redacted]	[redacted]	[redacted]	[redacted]

matched ALP ULN cut-off)					[redacted]
	Placebo	[redacted]	[redacted]	[redacted]	
RESPONSE (POISE matched ALP ULN cut-off)	Seladelpar 10 mg	[redacted]	[redacted]	[redacted]	[redacted]
	Placebo	[redacted]	[redacted]	[redacted]	
POISE[12] (Table 3 FDA label)	OCA 5-10 mg	70	33	47.1	2.868 (1.616, 5.089)
	OCA 10 mg	73	40	54.8	3.333 (1.909, 5.822)
	Placebo	73	12	16.4	
Li 2022[3]	200 mg/day fenofibrate + UDCA	Not included in NMA – the company argued that the definition used in the paper for Toronto criteria was not clear which the EAG agreed was the case			N/A
	UDCA				
ELATIVE[13] (Table 6 FDA label)	Elafibranor 80 mg	108	56	51.9	5.393 (2.298, 12.653)
	Placebo	53	5	9.4	

Abbreviations: ALP: Alkaline phosphatase; EAG: external assessment group; FDA: Federal Drugs Administration; N: total number in the arm in the study; n: number of events observed; NMA, network-meta analysis; OCA: Obeticholic acid; RR, risk ratio; UDCA: Ursodeoxycholic Acid; ULN: Upper limit of normal

Note: *crude risk ratio calculated based on event rates. Calculation was done using https://www.gigacalculator.com/calculators/relative-risk-calculator.php. It required a zero cell correction of >=1 unit, therefore a 1 correction was added to zero event arms

Input data sources were not provided in the NMA report for the other two outcomes which inform the economic analysis (pruritus and all cause discontinuation). The company provided a separate file that contained numerical information for the data input but not the source of the data.

The EAG were able to match the numbers for pruritus in original source publications (RESPONSE CSR[14] Table 48, Kowdley 2024[13] Table 3, Nevens 2016[12] Table 2, NCT03633227 clinicaltrials.gov). The EAG were able to match the numbers for treatment discontinuation for RESPONSE (Hirschfield 2024[6]) and POISE (Nevens, 2016[12] supplementary appendix). For ELATIVE it appeared that the company incorrectly used the data for discontinuation of participants from the trial rather than discontinuation from treatment in the NMA (Kowdley 2024 Supplementary Information Appendix[15] Figure S1: the NMA input file assumed 9 vs 4 discontinued treatment whereas the data indicates that 12 vs 6 discontinued treatment). However, the EAG considered that this would only have had a small effect on the results, which would have meant a small bias in in favour of elafibranor.

2.2.3. Comparison of trial results with NMA results

Recalculation of RESPONSE[6] data to correspond with the POISE[12] ULN threshold had a limited impact on results (Table and Table 4). However, recalculation using the ELATIVE[13] ULN threshold had a relatively large impact on results. For ALP normalisation (ALP ≤ 1.0× ULN), the response rate for seladelpar + UDCA was reduced from 25 to 18%. For ALP response using the Toronto I ULN threshold, the response rate for seladelpar + UDCA was reduced slightly from 65.6% to 60.2% but the placebo response rate was greatly reduced (from 26.2% to 12.3%) meaning that the relative effect for seladelpar + UDCA was nearly doubled in size.

The EAG identified an inconsistency between the naïve results from the trial and the NMA data in the comparison between seladelpar and elafibranor. In analyses both with and without adjustment for differences in ALP ULN cut-off, the effect size for ALP response for elafibranor + UDCA (relative to placebo + UDCA) was greater than seladelpar + UDCA (relative to placebo + UDCA; see Table 4). This did not align with the company's NMA results following outcome recalculation (Table 5 of company's ACD response), which indicated that seladelpar + UDCA was associated with a numerical benefit for seladelpar + UDCA versus elafibranor + UDCA (i.e. RR of >1). The EAG considered it plausible that the direction of the relative effect had been reversed by mistake for the recalculated comparison (i.e. are presented for elafibranor + UDCA vs seladelpar + UDCA, rather than as labelled). The EAG could not, however, identify any error in the code and input word documentation supplied that would indicate that this was the case. The results without outcome recalculation aligned more closely with the input data and the trial naïve data.

The EAG also noted a possible error in the company's reporting of the NMA results in Table 6 of the company response, in that the odds ratios and relative risk ratios appear to be reported the wrong way round when compared to Table 60 of their ITC report (Table 5). Values used in the model were correctly specified, however.

2.2.4. Consideration of Turner priors for Odds Ratios

The company explained that the log-odds-ratio scale (binomial/logit) on which the NMA was conducted allows application of Turner heterogeneity priors and that empirical evidence indicates that heterogeneity is generally comparable across log-OR and log-RR scales. The EAG considered this reasonable.

2.2.5. Adjustment for zero events

In its report, the EAG expressed a preference to see the NMA results with and without adjustments for zero events, which the company provided for ALP normalisation in its updated ITC report (reproduced below in Table 5). The EAG noted that the adjustment for zero events had a substantial impact both on the confidence intervals and the point estimates within the NMA, particularly in the comparison between seladelpar and OCA. Adjusting for zero events appeared to favour elafibranor and OCA over seladelpar in the reported point estimates.

Table 6: ALP normalisation results for Bayesian NMA using generic Turner prior

Seladelpar vs.	MAIC	Bayesian NMA with adjustment of zero events		Bayesian NMA without adjustment of zero events	
	RR (95% CrI)	RR (95% CrI)	OR (95% CrI)	RR (95% CrI)	OR (95% CrI)
ELATIVE matched ALP and total bilirubin ULN cut-off					
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Elafibranor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OCA 5-10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OCA 10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
POISE matched ALP ULN cut-off					
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Elafibranor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OCA 5-10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OCA 10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ALP: Alkaline phosphatase; CrI: Credible interval; EAG: Evidence assessment group; MAIC: Matching-adjusted indirect comparison; NMA: Network meta-analysis; OCA: Obeticholic acid; OR: Odds ratio; RR: Risk ratio; ULN: Upper limit of normal

Notes: Replication of Table 60 of the Data on File Updated ITC Report 20.08.2025

2.2.6. Impact of ULN cut-off on absolute treatment effect

The results in Table 6 demonstrate that varying the ULN cut-off used in the analysis had a limited impact on the absolute treatment effects within the NMA for ALP normalisation, however it had considerably more impact for ALP response (which was used to drive the transition to

mild ALP elevation in the economic model). When using Toronto, seladelpar was more effective than elafibranor for both ALP normalisation and ALP response. This did not align with the direction of effect in the naïve trial data following adjustment for differences in cut-offs across trials.

Table 7: Absolute treatment effect (normalisation rate at 12 months) for Bayesian NMA with adjustment of zero events using generic Turner prior

Absolute treatment effect	ALP normalisation (ALP ≤1.0 × ULN)		ALP response (Toronto I: ALP ≤1.67 × ULN)	
	ELATIVE matched ALP and total bilirubin ULN cut-off	POISE matched ALP ULN cut-off	ELATIVE matched ALP and total bilirubin ULN cut-off	POISE matched ALP ULN cut-off
Placebo	[redacted]	[redacted]	[redacted]	[redacted]
Seladelpar	[redacted]	[redacted]	[redacted]	[redacted]
Elafibranor	[redacted]	[redacted]	[redacted]	[redacted]
OCA 5-10 mg	[redacted]	[redacted]	[redacted]	[redacted]
OCA 10 mg	[redacted]	[redacted]	[redacted]	[redacted]

Notes: Data taken from Table 62 and Table 64 of the Data on File Updated ITC Report 20.08.2025

2.2.7. Ability to produce one NMA

The company stated that the differences in definitions of the ULN for ALP precluded synthesis of a single network of all three studies using a consistent definition of the ULN for ALP. The company instead conducted an additional scenario analysis in the model using the average of two sets of Bayesian NMA results using ULN cut-offs matched to ELATIVE[13] and POISE,[12] respectively (both including all 3 comparators in a single network). This was presented in Table 6 of the company response using a generic Turner prior. The EAG noted that the OR / RR in the comparison to elafibranor appeared to be presented the wrong way round for the TORONTO I NMA using the ELATIVE cut-off. The direction of the other results appeared consistent with the input data. The EAG did not consider this analysis particularly informative. Instead, the EAG considered that the results from the NMAs using consistent cut-offs against with the comparator of interest should be prioritised (i.e. OR from POISE matched for OCA and OR from ELATIVE matched for elafibranor), as the least biased of the available results.

As stated in its original report, the EAG considered that a single analysis to support decision-making framework was generally preferable, requiring a single indirect comparison and single fully incremental analysis across all treatment options for a decision problem. However, based

on the information provided by the company in its response, the EAG agreed with the company that producing one consistent NMA was not possible in this case due to differences in definitions of the ULN across trials. As the treatment effect for seladelpar was different using the definitions from POISE[12] and ELATIVE,[13] assuming consistent relative effects across the cut-offs would not be appropriate.

In the circumstances, the EAG considered that the best analysis option given the data available was to conduct two separate NMAs predicting the relative effectiveness of seladelpar vs its comparators in the POISE[12] and ELATIVE[13] populations individually. The implications of this for interpretation of the economic analysis are described in Section 3. The EAG maintained its view that a MAIC comparing seladelpar and elafibranor was not useful for decision-making, which is discussed further in Section 2.2.9.

2.2.8. EAG NMA

Given that the EAG was unable to run the code provided by the company without the input files, the EAG conducted a cross-check of the NMA results using the multinma package for ALP normalisation and ALP response. Both a Bayesian NMA using the Turner prior and a frequentist NMA were conducted. The multinma package only allowed the input of integer values and therefore a continuity correction of 1 rather than 0.5 was used in the EAG's Bayesian analysis. This did not appear to have a large impact on the point estimates, which were consistent with the frequentist results.

The results of the EAG's analysis were broadly consistent with the company's NMA in terms of point estimates for ALP normalisation; particularly for elafibranor. In comparison to OCA, the EAG NMA was somewhat less favourable. The credible intervals were narrower in the EAG analyses, this was thought to be due to the larger continuity correction used. As with the sensitivity analysis presented by the company without adjustment for zero values, this demonstrated that the credible intervals were highly dependent on the adjustments made to the placebo arms of the trials and that the results of the NMA were highly uncertain.

For ALP response using Toronto I, the results were again broadly consistent except that the comparison of seladelpar and elafibranor flipped direction (as would be expected from the raw data as highlighted in Section 2.1.3).

Table 8: EAG NMA analyses

	NMA type	OR (95% CrI)
ELATIVE matched ALP and total bilirubin ULN cut-off **ALP normalisation vs elafibranor**	Bayesian with +1 continuity correction	[redacted]
	Frequentist with +0.5 continuity correction	[redacted]
	Company result +0.5 continuity correction	[redacted]
POISE matched ALP ULN cut-off **ALP normalisation vs OCA 5-10mg**	Bayesian with +1 continuity correction	[redacted]
	Frequentist with +0.5 continuity correction	[redacted]
	Company result +0.5 continuity correction	[redacted]
ELATIVE matched ALP and total bilirubin ULN cut-off **ALP response (Toronto I: ALP ≤1.67 × ULN) vs elafibranor**	Bayesian	[redacted]
	Company result	[redacted]
POISE matched ALP ULN cut-off **ALP response (Toronto I: ALP ≤1.67 × ULN) vs OCA 5-10mg**	Bayesian	[redacted]
	Company result	[redacted]

Abbreviations: ALP: Alkaline phosphatase; CrI: Credible interval; EAG: Evidence assessment group; NMA: Network meta-analysis; OCA: Obeticholic acid; OR: Odds ratio; ULN: Upper limit of normal

Notes: company results taken from ITC report as draft guidance response had flipped odds ratios and relative risks

2.2.9. Appropriateness of the MAIC

The EAG maintained its view that the company MAIC for the comparison to elafibranor was not appropriate for decision-making. The company noted in their response that the presence of unbalanced effect modifiers between studies necessitates a MAIC for a valid comparison (as per NICE TSD 18) and that "Gilead considers that, while clinicians are best placed to identify clinical factors that are effect modifiers, statisticians are best placed to assess whether there is sufficient heterogeneity to warrant a MAIC."

TSD18 states:

- "Submissions using population-adjusted analyses in a connected network need to provide evidence that they are likely to produce less biased estimates of treatment differences than could be achieved through standard methods."
- "Quantitative evidence must be presented that population adjustment would have a material impact on relative effect estimates due to the removal of substantial bias." Comparison to minimum clinically important differences is noted as one way to look at this and the TSD also notes: "the qualification of "substantial" bias should be considered in both a clinical … and statistical context"

TSD18 also does not recommend MAIC over other methods (e.g. STC).

The EAG noted that the forest plot supplied for ALP normalisation indicated that the effect modifier of most significance was cirrhosis; followed by total bilirubin. Neither of these would be a statistically significant treatment effect modifier as confidence intervals overlapped (although of course the study was not powered to look for this). There was no clear evidence of effect modification for either age or baseline ALP presented in the forest plot. A forest plot was not presented for ALP response (Toronto I: ALP ≤1.67 × ULN).

The EAG also noted that in the elafibranor appraisal, the company stated that the importance of any differences in trial design, patient populations and outcomes assessed was "validated with clinicians", not determined only by clinicians. This was considered appropriate as evidence for effect modification for OCA was relatively limited. A forest plot for POISE[12] could not be identified. Forest plots for the ELATIVE[13] study showed evidence of effect modification for baseline bilirubin and ALP (non-overlapping confidence intervals for the primary endpoint) with no clear effect modification for age or cirrhosis (Figure 27 and Figure 28, CS TA1016).

Statistical tests for difference between ELATIVE[13] and RESPONSE[6] indicated that the only effect modifier with a significant difference between the trials was bilirubin (Table 12 Updated ITC Report).

Other than statistical differences in trial populations, it was also important to consider whether such differences were clinically meaningful as noted in TSD18. The company did not present any evidence to address this point. The EAG was also not aware of any information on MCIDs for the effect modifiers in this population. Given this, the EAG consulted its clinical experts to understand the meaningfulness of differences across trials using the table provided by the

company in their response which is reproduced below (Table 3). Two clinical experts were able to respond within the timeframe. They considered that:

- ALP was the best marker, but the differences in baseline ALP were not meaningful and did not favour any particular treatment
- The next most important characteristic was cirrhosis. The lower proportion with cirrhosis in ELATIVE[13] at baseline might be anticipated to favour elafibranor (vs. both seladelpar and OCA)
- The somewhat lower bilirubin in ELATIVE[13] might be anticipated to favour elafibranor in terms of assessing the proportion of patients responding after treatment, but was less important than differences in cirrhosis or ALP

Given the above, the EAG did consider that there may be some level of bias towards elafibranor in the NMA. However, the MAIC submitted was not considered to meet the criteria of "likely to produce less biased estimates" compared to the NMA. The large drop in sample size associated with the company's MAIC (to 36% of the original sample) and large proportion of the effective sample size (ESS) concentrated in patients with high weights[a] indicated clear evidence that the matching process had struggled to account for baseline differences and had the potential to introduce an additional source of bias, as well as additional uncertainty in the estimates.

The company presented a sensitivity analysis adjusting only for bilirubin and cirrhosis, which the EAG considered to be highly relevant given that these were the only two factors identified as having a potentially clinically meaningful difference between the trials. According to the company's ITC report, this resulted in an identical RR to the main MAIC for ALP normalisation [redacted] for seladelpar versus elafibranor [redacted]. The EAG was unclear whether this was a reporting error, on the basis that the chance of identical results would be relatively low. The EAG also noted that the numerical estimate produced by the MAIC, when the width of the confidence intervals was considered, was relatively similar to the NMA [RR: [redacted]. For ALP response using Toronto I, the sensitivity analysis adjusting for

[a] [redacted]

only bilirubin and cirrhosis resulted in a similar (not identical) RR to the main MAIC [redacted] vs [redacted] and a similar result to the Bayesian NMA with outcome recalculation RR [redacted]. The EAG considered that these results were more robust than those for ALP normalisation as the issues with adjustment for zero events in the placebo arm did not apply. Similarly, the MAIC and Bayesian NMA with the Turner prior produce similar results for pruritus [redacted]. These results indicated that the MAIC methodology was not producing any meaningful difference in the outcomes for which the most robust data were available.

Table 9: Distribution of the effect modifiers in three key trials

Population characteristics	ELATIVE[13]	RESPONSE[6]	POISE[12]
Baseline ALP mean U/L (SD)	321.9 (150.9)	314.3 (121.88)	323 (112.53)
Mean total bilirubin level-µmol/liter (SD)	9.6 (5.1)	12.9 (5.147)	11.1 (6.498)
Cirrhosis (%)	9.94 (8.3 in ELA and 13.2 in UDCA)	14	16
Age at diagnosis (SD) [95% CI]	--	49.23 (10.30)	47.32 (10.79)
Age at screening (mean ± SD)	[redacted]	[redacted]	
Bilirubin >ULN at baseline (%)	3.7	13.0 (15.6 in SEL and 7.7 in UDCA)	8.3

Abbreviations: ALP, alkaline phosphatase; CI, confidence interval; ELA, elafibranor; SEL, seladelpar; SD, standard deviation; UDCA, ursodeoxycholic acid; ULN, upper limit of normal

Note: Reproduced from Table 3 of company response and Table 12 of the Updated ITC Report

2.2.10. Conclusions

Whilst the EAG welcomed the additional information and analyses provided by the company, these were unable to resolve the fundamental issues presented within the indirect treatment comparisons, which were:

- It was not possible for the EAG to check the company's re-analysis using the alternative ALP cut-offs

- Differences in ALP cut-offs across trials which precluded the use of all Phase 3 trials in a single analysis
- Considerable uncertainty in relative treatment effects due to low sample sizes and event numbers
- For ALP normalisation, there was considerable uncertainty in the NMAs in large part due to there not being any events in the placebo arms of RESPONSE,[6] ELATIVE[13] and POISE.[12] This resulted in extremely wide credible intervals, which depended on the continuity correction used
- A potential issue with the company NMA for ALP response using Toronto I in the comparison of seladelpar and elafibranor, where naïve trial data (and EAG analysis) were contrary to the results of the company analysis
- Inconsistency in placebo response rates across trials, the cause of which remains uncertain

Overall, the EAG considered that the best analysis option given the data available was the use of the two separate NMAs predicting the relative effectiveness of seladelpar versus its comparators in the POISE[12] and ELATIVE[13] populations individually. In the comparison to elafibranor, this produced near identical results to the MAIC for the outcomes for which the most information is available (ALP response using Toronto I and pruritus).

2.3. NMA used in the economic model and implications for decision making

As discussed in Section 2.2, the EAG considered that the most appropriate approach (given the data available) was to use the results from two NMAs to predict the relative effectiveness of seladelpar versus its comparators in the POISE and ELATIVE populations, individually. To maintain internal consistency in the model, the EAG calibrated the hazard ratios for ALP normalisation and Toronto I criteria using the odds ratios from the NMA. The ELATIVE cut-off was used for elafibranor, while POISE cut-off was used for OCA.

This effectively divided the decision problem into two, prohibiting a robust fully incremental analysis. In this case, in theory, there should be one comparison relating to the ELATIVE population and one related to the POISE population. However, this was not possible for the EAG to implement with the available data as the model took its base transitions for both comparisons from RESPONSE and the EAG did not have access to transition probabilities for the timepoints included in the model for the other two trials (0-1,1-3, 3-6, 6-9 and 9-12 months). The model therefore assumed that relative effects from a comparison conducted in a different population could be applied to baseline transitions from RESPONSE. This would appear reasonable for the POISE population, as application of alternative cut-offs made little difference to the trial outcomes, but less reasonable in the comparison to elafibranor using the ELATIVE study cut-offs. Given the limitations with the available data, the EAG therefore presents results as per the original company analysis (including fully incremental analysis) using the baseline transitions from RESPONSE but noted the additional uncertainty of this for the comparison to elafibranor.

2.4. Treatment discontinuation

The company proposed anchoring 0-12 months discontinuation on OCA discontinuation from Abbas et al. (2023),[16] a study that had previously been used by the EAG in a scenario analysis to model OCA cumulative discontinuation. The EAG noted that, in their ACD response, the company cited Abbas et al. (2024),[17] which was a different study from the earlier Abbas 2023 source used by the EAG. Using this as the anchor, the company applied the ITC to derive discontinuation rates for seladelpar and elafibranor. The company claimed that there was no clinical rationale to assume differences between seladelpar and elafibranor, and therefore it assumed equal rates of 8.4% at 12 months.

The EAG disagreed with this approach. In the EAG's view, discontinuation should have remained anchored to the RESPONSE trial, consistent with how the model estimated treatment

effectiveness. Individual patient data from RESPONSE informed the transition probabilities for effectiveness in the seladelpar arm for the first 12 months, with comparator effectiveness derived from the ITC. Anchoring discontinuation to a different external source (e.g. Abbas 2023) would break this internal consistency, effectively breaking the link between discontinuation rates and effectiveness. The EAG therefore considered that RESPONSE should remain the anchor for discontinuation, with comparators' discontinuation estimated relative to seladelpar via the ITC.

The EAG also rejected the company's assumption of identical discontinuation for seladelpar and elafibranor, noting it was reasonable to expect patients and clinicians to continue longer on a treatment that shows fewer incidents of severe pruritus, consistent with the results for pruritus in the ITC.

2.5. Utility values for health states and pruritus

The company explored two alternative scenarios for modelling pruritus disutility. The first used re-anchored values from Smith et al. (GLIMMER trial; N=36),[18] which adjusted for the high utility observed among patients without itching that exceeded that of the general population. Despite this adjustment, the EAG remained concerned about face validity, as the re-anchored disutility for severe pruritus (-0.32) was of the same magnitude as the disutility applied in a previous NICE appraisal (TA679) for hospitalisation due to heart failure. The EAG noted that the Smith et al. authors concluded that patients with severe pruritus had similar quality of life to those with severe Parkinson's disease. The EAG presented these data to its clinical experts, one of whom was able to respond in the timeframe. They indicated that the estimates lacked face validity on the basis that some patients with an NRS of 7 would still be able to work, and therefore disutility would not be as low as for hospitalisation due to heart failure or severe Parkinson's disease. The EAG therefore considered that the Smith et al. data was, at least to some extent, implausible and lacking in face validity.

The second scenario used estimates from Hussain et al. (2023),[19] which the EAG considered provided values that were more clinically plausible than Smith et al. However, the study population had primary sclerosing cholangitis (rather than PBC), and results were available only as a conference abstract, which limited detailed appraisal of the study methods.

The EAG still considered that the most appropriate values were those derived from the MMRM2 mapping model based on EQ-5D-3L data from RESPONSE.[6] These estimates were internally

consistent, based on trial data, and aligned with findings from Rice et al.[20] (a large UK study of 2,240 participants with PBC that reported only a small disutility of -0.018 for itch, which was not statistically significantly associated with HRQoL impairment). Rice et al. attributed the lack of impact of pruritus to the mitigating effect of treatment. As the current model already included the cost of anti-pruritic drugs, the EAG considered that including high disutility values from less credible sources would result in overestimating the impact of pruritus within the model.

3. EAG RESPONSE TO CHANGES TO THE COMPANY'S COST EFFECTIVENESS ESTIMATES

The analyses presented in this document used the confidential price for seladelpar [redacted] and list prices for obeticholic acid, elafibranor and other drugs.

Table 10: Assumptions in the EAG and company base cases

	EAG base case	Company (Post ACM1)	Company (suggested revisions)
Relative effect	Two NMAs: ELATIVE-matched ULN for elafibranor and POISE-matched ULN for OCA; ORs applied to calibrate hazards	Anchored MAIC for elafibranor (RR) NMA with outcome recalculation (generic Turner prior) for OCA (OR)	Same as Post ACM1
Discontinuation (0-12 months)	ITC anchored to RESPONSE (seladelpar as anchor; comparators relative to seladelpar).	from the trials	ITC anchored to Abbas 2023 (OCA discontinuation as anchor).
Pruritus disutilities	MMRM2 mapping model from EQ-5D-3L (pivotal seladelpar trial).	Smith et al.	Adjusted Smith et al. (re-anchored) or Hussain 2023 disutilities.

Abbreviations: ACM, appraisal committee meeting; ITC, indirect treatment comparison; MAIC, matched adjusted indirect comparison; MMRM, mixed models for repeated measures; NMA, network meta-analysis; OCA, obeticholic acid; OR, odds ratio; RR, risk ratio; ULN, upper limit of normal

Table 11: Results of the company analyses

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	Change in NHB from the company base case
Company base-case post ACM1 (short term discontinuation unadjusted from trials and pruritus disutilities from Smith et al)									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	-
OCA + UDCA	421,717	15.606	9.512	[redacted]	[redacted]	[redacted]	Dominates	-11.57	-
Elafibranor + UDCA	499,347	15.826	10.153	[redacted]	[redacted]	[redacted]	Dominates	-14.81	-
Company base-case post ACM1 + Company suggestion for short term discontinuation ITC anchored to Abbas et al (0-12 months) + Company suggestion pruritus disutility based on adjusted Smith et al (decrement -0.32 vs previous -0.38)									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	[redacted]
OCA + UDCA	380,668	15.334	9.909	[redacted]	[redacted]	[redacted]	Dominates	-9.12	2.45
Elafibranor + UDCA	508,104	15.882	10.809	[redacted]	[redacted]	[redacted]	Dominates	-14.60	0.22
Company base-case post ACM1 + Company suggestion for short term discontinuation ITC anchored to Abbas et al (0-12 months) + Company suggestion pruritus disutility based on Hussain et al.									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	[redacted]

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	Change in NHB from the company base case
OCA + UDCA	380,668	15.334	10.413	[redacted]	[redacted]	[redacted]	Dominates	-8.62	2.95
Elafibranor + UDCA	508,104	15.882	11.088	[redacted]	[redacted]	[redacted]	Dominates	-14.32	0.50

Abbreviations: ACM, appraisal committee meeting; ICER, incremental cost effectiveness ratio; inc, incremental; m, months; MMRM, Mixed model for repeated measures; NHB, net health benefit; NMA, network meta-analysis; OCA, Obeticholic acid; QALYs, quality adjusted life years; RR, risk ratio; SE, standard error; UDCA, Ursodeoxycholic acid.

Note: NHB calculated at £20,000 per QALY.

Table 12: Results of the EAG base case

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	Change in NHB from the company base case
EAG updated base case (NMA with ELATIVE cut-off used for elafibranor, and NMA with POISE cut-off used for OCA, short term discontinuation unadjusted from trials and pruritus disutilities from RESPONSE)									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	[redacted]
OCA + UDCA	368,928	15.277	10.960	[redacted]	[redacted]	[redacted]	Dominates	-7.49	4.09

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	Change in NHB from the company base case
Elafibranor + UDCA	488,924	15.718	11.468	[redacted]	[redacted]	[redacted]	Dominates	-12.98	1.84

Abbreviations: ACM, appraisal committee meeting; ICER, incremental cost effectiveness ratio; inc, incremental; m, months; MMRM, Mixed model for repeated measures; NHB, net health benefit; NMA, network meta-analysis; OCA, Obeticholic acid; QALYs, quality adjusted life years; RR, risk ratio; SE, standard error; UDCA, Ursodeoxycholic acid.

Note: NHB calculated at £20,000 per QALY.

Table 13: Results of the EAG scenario analyses

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	Change in NHB from the company base case
EAG-calculated OR for Toronto I (ELATIVE-matched cut-off for elafibranor)									
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	[redacted]
OCA + UDCA	368,928	15.277	10.960	[redacted]	[redacted]	[redacted]	Dominates	-7.49	4.09
Elafibranor + UDCA	489,666	15.740	11.491	[redacted]	[redacted]	[redacted]	Dominates	-12.99	1.82
EAG base case + disutilities from Hussain et al.									

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER	NHB at 20k	Change in NHB from the company base case
Seladelpar + UDCA	[redacted]	[redacted]	[redacted]					[redacted]	[redacted]
OCA + UDCA	368,928	15.277	10.361	[redacted]	[redacted]	[redacted]	Dominates	-8.09	3.49
Elafibranor + UDCA	488,924	15.718	10.934	[redacted]	[redacted]	[redacted]	Dominates	-13.51	1.30

Abbreviations: ACM, appraisal committee meeting; ICER, incremental cost effectiveness ratio; inc, incremental; m, months; MMRM, Mixed model for repeated measures; NHB, net health benefit; NMA, network meta-analysis; OCA, Obeticholic acid; QALYs, quality adjusted life years; RR, risk ratio; SE, standard error; UDCA, Ursodeoxycholic acid.

Note: NHB calculated at £20,000 per QALY.

4. REFERENCES

1. National Institute for Health and Care Excellence (NICE). Elafibranor for treating primary biliary cholangitis (TA1016). London: NICE; 2024. Available from: https://www.nice.org.uk/guidance/ta1016.

2. Hirschfield GM, Beuers U, Corpechot C, Invernizzi P, Jones D, Marzioni M, et al. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. J Hepatol. 2017;67(1):145-72.

3. Li C, Zheng K, Chen Y, He C, Liu S, Yang Y, et al. A randomized, controlled trial on fenofibrate in primary biliary cholangitis patients with incomplete response to ursodeoxycholic acid. Ther Adv Chronic Dis. 2022;13:20406223221114198.

4. Corpechot C, Chazouillères O, Rousseau A, Le Gruyer A, Habersetzer F, Mathurin P, et al. A Placebo-Controlled Trial of Bezafibrate in Primary Biliary Cholangitis. NEJM. 2018;378(23):2171-81.

5. Iwasaki S, Ohira H, Nishiguchi S, Zeniya M, Kaneko S, Onji M, et al. The efficacy of ursodeoxycholic acid and bezafibrate combination therapy for primary biliary cirrhosis: A prospective, multicenter study. Hepatol Res. 2008;38(6):557-64.

6. Hirschfield GM, Bowlus CL, Mayo MJ, Kremer AE, Vierling JM, Kowdley KV, et al. A Phase 3 Trial of Seladelpar in Primary Biliary Cholangitis. NEJM. 2024;390(9):783-94.

7. Hosonuma K, Sato K, Yamazaki Y, Yanagisawa M, Hashizume H, Horiguchi N, et al. A prospective randomized controlled study of long-term combination therapy using ursodeoxycholic acid and bezafibrate in patients with primary biliary cirrhosis and dyslipidemia. Am J Gastroenterol. 2015;110(3):423-31.

8. Itakura J, Izumi N, Nishimura Y, Inoue K, Ueda K, Nakanishi H, et al. Prospective randomized crossover trial of combination therapy with bezafibrate and UDCA for primary biliary cirrhosis. Hepatol Res. 2004;29(4):216-22.

9. Kanda T, Yokosuka O, Imazeki F, Saisho H. Bezafibrate treatment: a new medical approach for PBC patients? J Gastroenterol. 2003;38(6):573-8.

10. Bonder A, Jones D, Levy C, Zou H, Civitarese A, Nevens F, editors. Metabolic outcomes and key efficacy data from a phase 2 trial evaluating combination obeticholic acid and bezafibrate in patients with primary biliary cholangitis. DDW 2024; 2024; Washington D.C.

11. Liberopoulos EN, Florentin M, Elisaf MS, Mikhailidis DP, Tsianos E. Fenofibrate in primary biliary cirrhosis: a pilot study. Open Cardiovasc Med J. 2010;4:120-6.

12. Nevens F, Andreone P, Mazzella G, Strasser SI, Bowlus C, Invernizzi P, et al. A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis. NEJM. 2016;375(7):631-43.

13. Kowdley KV, Bowlus CL, Levy C, Akarca US, Alvares-da-Silva MR, Andreone P, et al. Efficacy and Safety of Elafibranor in Primary Biliary Cholangitis. NEJM. 2024;390(9):795-805.

14. Gilead. Data on File - RESPONSE Clinical Study Report. 2023 2023.

15. Kowdley KV, Bowlus CL, Levy C, Akarca US, Alvares-da-Silva MR, Andreone P, et al. Efficacy and Safety of Elafibranor in Primary Biliary Cholangitis. Supplementary Appendix. NEJM. 2024;390(9):795-805.

16. Abbas N, Culver EL, Thorburn D, Halliday N, Crothers H, Dyson JK, et al. UK-Wide Multicenter Evaluation of Second-line Therapies in Primary Biliary Cholangitis. Clin Gastroenterol Hepatol. 2023;21(6):1561-70.e13.

17. Abbas N, Smith R, Flack S, Bains V, Aspinall RJ, Jones RL, et al. Critical shortfalls in the management of PBC: Results of a UK-wide, population-based evaluation of care delivery. JHEP Rep. 2024;6(1):100931.

18. Smith H, Fettiplace J, Von Maltzahn R, Das S, McLaughlin MM, Jones D. More Than Just an Itch: Impact of Cholestatic Pruritus in Primary Biliary Cholangitis (PBC) on Health-Related Quality of Life (HRQoL). Am J Gastroenterol. 2022;117(10):S78.

19. Hussain N, Hirschfield B, Ferguson J, Abbas N, Bhandal K, Rogers P, et al. P34 Burden, impact and variability of pruritus in primary sclerosing cholangitis (PSC) over time: a prospective observational study. Gut. 2023;72:A34.

20. Rice S, Albani V, Minos D, Fattakhova G, Mells GF, Carbone M, et al. Effects of Primary Biliary Cholangitis on Quality of Life and Health Care Costs in the United Kingdom. Clin Gastroenterol Hepatol. 2021;19(4):768-76.e10.

Seladelpar for Previously Treated Primary Biliary Cholangitis [ID6429]

Additional Evidence Prior to Appraisal Committee Meeting 3

Executive Summary

Following ACM 2, the appraisal committee preferred

(1) the biomarker response rate indirect treatment comparison (ITC) unadjusted for differing alkaline phosphatase (ALP) normalisation thresholds between trials

(2) discontinuation estimated from the ITC, using the seladelpar (SEL) RESPONSE discontinuation rate as the reference

(3) pruritus disutility values estimated via mapping from the ITCH-E study.

Gilead strongly disagree with the committee's conclusions regarding the preferred source of discontinuation rates and pruritus disutility values.

Regarding discontinuation rates, Gilead considers that methodological differences between the seladelpar models and precedent models, notably that of obeticholic acid (OCA), generate counter-intuitive results for decision-making. Specifically, even when SEL is substantially discounted and parity priced to OCA, reducing the seladelpar discontinuation rate reduces cost-effectiveness. Since the committee has agreed in principle that SEL has the better clinical profile, this implies that NICE is willing to reject a product with a better clinical profile given a like-for-like price to OCA, a product which has already been considered value for money for the NHS.

Furthermore, real-world evidence suggests that treatment discontinuation rates in clinical practice vary substantially from those in the clinical studies and predicted by the ITC, making both trial-derived and ITC-derived estimates unreliable inputs for the model. Parameterising the model with an input that is both highly uncertain and to which results are highly sensitive would undermine the credibility of the analysis. Gilead therefore considers that the fairest and most reasonable approach is to apply the same discontinuation rate to all three comparators. This permits a cleaner comparison underpinned by treatment efficacy on liver biomarkers and pruritus and unbiased by highly uncertain discontinuation rates.

Regarding source of pruritus disutility, important context for preferring disutility values from the literature was omitted from the ACM 2 committee slides. Specifically, the committee's preferred mapped disutility values were generated from the ITCH-E study using a different clinical measure of pruritus severity (PBC-40) to that used in the model (numerical rating scale [NRS]). Analysis of the RESPONSE data, which collected both measures, demonstrates little overlap between responses to the two measures: a large proportion of patients reporting Clinically Severe pruritus on the PBC-40 measure (the measure used for mapping) reported mild or moderate pruritus on the NRS scale (the pruritus measure used in the model). This misalignment in severity generated disutility values for Severe pruritis from ITCH-E that underestimated the disutility of Severe pruritus on the NRS scale used in the model. It also renders impossible the derivation of utility values for both liver biomarker health states and pruritus from RESPONSE.

Consequently, to confirm the higher disutility values for severe pruritus previously identified by Gilead in the literature, Gilead has conducted further qualitative and quantitative pruritus quality of life (QoL) studies in external PBC datasets to capture severity on the NRS scale, which aligns with the measure used in the economic model. Specifically, Gilead has conducted quantitative analysis of an Adelphi-owned Disease-Specific Programmes (DSP) study data set.

Gilead considers the results from the DSP study to be the most relevant to decision-making because they address the key limitations identified with the ITCH-E mapped values. The ITCH-E mapping approach significantly underestimated the disutility of severe pruritus, since the utility values would have been derived from patients with mild and moderate pruritus on the NRS scale, as previously explained during consultation. Further stratification of PBC-40 pruritus scores to generate separate moderate and severe disutilities was not feasible given the poor correlation observed in RESPONSE, making it impossible to use utility data for both liver biomarker and pruritus health states from a single consistent source — the committee's stated preference. The DSP study overcomes these limitations. The disutility values were elicited directly using the EQ-5D, placing them higher within NICE's hierarchy of preferred utility sources than those mapped from the PBC-40. The underlying model adjusted for all available demographic and clinical covariates, including ALP levels, directly addressing the committee's concern regarding potential double-counting of the effects of ALP and pruritus. By adjusting for all known covariates reported to influence quality of life in pruritus patients, where data permitted, these values provide a more methodologically robust and clinically reflective estimate of the true burden of pruritus.

In summary, Gilead's preferred base case for consideration during ACM 3 changes two of committee's preferred assumptions post-ACM 2:

1) To assume the same short and long-term discontinuation rates for all three comparators (16% short term, the average of the results from the ITC; 0.9% long-term, based on the ratio of the long vs. short-term SEL discontinuation rates in RESPONSE).

2) To apply the pruritus disutility values from the DSP study.

Gilead's updated base case results are presented in Table 1 using the current discounted price of [redacted] per pack for SEL and list price for OCA and ELA. All other assumptions are as per the committee's preferred base case post-ACM 2.

Table 1: Company's updated base case cost-effectiveness results

Intervention	Costs (£)	QALYs	Inc. costs (£)	Inc. QALYs	ICER
SEL + UDCA	[redacted]	[redacted]	-	-	-
OCA + UDCA	[redacted]	[redacted]	[redacted]	[redacted]	Dominates
ELA + UDCA	[redacted]	[redacted]	[redacted]	[redacted]	Dominates

Discontinuation

Low UDCA Discontinuation in POISE

Discontinuation assumptions used in the seladelpar economic model significantly drive cost-effectiveness results.

As can be seen from the descriptive data presented in Table 2, OCA has the highest relative discontinuation compared to placebo at 12 months (RR=2.7)[1], driven by substantially lower discontinuation in the ursodeoxycholic acid (UDCA) /placebo arm of POISE relative to the other trials in the ITC. While patients receiving elafibranor (ELA) discontinued roughly at the same rate as patients on UDCA (RR=0.98), fewer patients discontinued treatment with SEL compared to UDCA as observed within the RESPONSE trial (RR=0.63).[2, 3]

Table 2: Discontinuation at 12 Months (Descriptive Statistics)

Intervention	Active			Control (UDCA / Placebo)			RR
	n	N	Rate	n	N	Rate	
SEL	10	128	7.81%	8	65	12.31%	0.63
ELA	12	108	11.11%	6	53	11.31%	0.98
OCA	16	144	11.11%	3	73	4.11%	2.70
ELA: Elafibranor; OCA: Obeticholic Acid; SEL: Seladelpar; UDCA: Ursodeoxycholic Acid							

Gilead conducted an indirect treatment comparison (ITC) on the 12-month discontinuation data, and derived relative effects, as measured by odds ratios, which are shown below [4].

As presented in Table 3, SEL was associated with numerically lower odds of all-cause discontinuation compared to ELA 80 mg, OCA 5-10 mg, and OCA 10 mg.

Table 3: Discontinuation at 12 Months (ITC Results)[4]

SEL vs.	OR (95% CI)	Discontinuation rate
SEL	-	7.8%
ELA	0.58 (0.11, 2.92)	12.7%
OCA 5-10 mg	0.24 (0.03, 1.38)	26.1%
OCA 10 mg	0.19 (0.03, 1.04)	30.8%
OR: Odds Ratio		

During the appraisal process the Evidence Assessment Group (EAG) suggested applying the discontinuation results from the Gilead ITC to the SEL discontinuation probability to estimate the discontinuation rates of ELA and OCA.

Table 3 shows that the discontinuation rate applied for OCA in the economic model as per the committee's preferred ITC inputs is much higher (26.1%) than the rates applied for both SEL (7.8%) and ELA (12.7%).

Gilead would like to highlight that the OCA discontinuation rates based on the ITC results are driven by the low UDCA discontinuation in the POISE trial (~4.11%, [1, 5]) potentially due to exclusion of patients with severe pruritus from the trial [6], which artificially inflated the relative discontinuation rate of OCA in the ITC. This in turn produces counterintuitive results, as outlined in the following section.

Counterintuitive Outcomes Arising from Comparator Discontinuation Dynamics

In Table 4 and Table 5 below, we illustrate how the discontinuation rates in the model generate counterintuitive results, by exploring the impact on cost-effectiveness when SEL and OCA are parity-priced and assuming a range of hypothetical discounts from 40% to 70% off the list price per patient for OCA. Firstly, we set SEL and OCA to have the same annual drug costs (£1,451.28) anchored to a hypothetical 40.0% discount off the OCA list price. This corresponds to a 54.7% discount off SEL list price. In the first half of Table 4, we leave all other assumptions as per the committee's preferred post-ACM 2 base case. Under these conditions SEL generates incremental QALYs, but the incremental costs generate an ICER of £48,311, above NICE's cost-effectiveness threshold.

In the second half of Table 4, we keep all assumptions as per the first half, but we double the SEL discontinuation rate to 16.0%. As expected, both incremental costs and QALYs decrease, but the former clearly increases more such that the ICER reduces to £35,211, above NICE's acceptable cost-effectiveness range but closer to it than when discontinuation was lower. Table 5 shows the same scenario but with a hypothetical discount of 70.0% for OCA (£725.64 annual drug costs) and SEL priced at parity. Although both scenarios now fall within NICE's acceptable ICER range, the same effect is observed, with higher SEL discontinuation generating better cost-effectiveness.

This observation is clearly counterintuitive within the context of NICE decision-making; the committee has accepted that SEL has a better clinical profile than OCA, with better control of liver inflammation based on the ITC of liver biomarker tests and better control of pruritus. Even with a hypothetical discount of 70.0% for OCA, SEL (if priced at parity) would be deemed less cost-effective with better adherence. In other words, the model penalises SEL for keeping patients on a clinically superior treatment for longer. The committee's position therefore implies that NICE would reject a clinically superior product at a comparable price to one that it has already deemed value for money for the NHS [7].

The reason for this lies within differences in model structures and assumptions. For example, OCA + UCDA generates [redacted] QALYs in our model vs. ~12.5 – 13.5 QALYs (depending on whether EAG or company estimates) generated in the model submitted in the OCA technology appraisal [5]. For reasons unknown, no comparator, even OCA, is cost-effective vs. UDCA within the present model structure. This means that products with higher discontinuation rates (reflecting a poorer clinical profile) will be more cost-effective, regardless of their relative efficacy.

Table 4: Equal Discontinuation for SEL & OCA - Cost-Effectiveness Results assuming 40% discount off OCA list price

Intervention	Costs (£)	QALYs	Inc. costs (£)	Inc. QALYs	ICER
Committee base case post-ACM 2, but with equal drug costs (annual cost: £ 1,451.28) at an assumed 40% discount of OCA list price					
SEL + UDCA	[redacted]	[redacted]	[redacted]	[redacted]	48,311
OCA + UDCA	[redacted]	[redacted]			
As above, but doubling the SEL discontinuation rate from 8% to 16%					
SEL + UDCA	[redacted]	[redacted]	[redacted]	[redacted]	35,211
OCA + UDCA	[redacted]	[redacted]			

Table 5: Equal Discontinuation for SEL & OCA - Cost-Effectiveness Results assuming 70% discount off OCA list price

Intervention	Costs (£)	QALYs	Inc. costs (£)	Inc. QALYs	ICER
Committee base case post-ACM 2, but with equal drug costs (annual cost: £ 725.64) at an assumed 70% discount of OCA list price					
SEL + UDCA	[redacted]	[redacted]	[redacted]	[redacted]	15,031
OCA + UDCA	[redacted]	[redacted]			
As above, but doubling the SEL discontinuation rate from 8% to 16%					
SEL + UDCA	[redacted]	[redacted]	[redacted]	[redacted]	8,115
OCA + UDCA	[redacted]	[redacted]			

Uncertainty due to discontinuation and proposed base case for discontinuation rates

Gilead has actively sought to obtain real-world data on both OCA and ELA discontinuation rates through additional literature searches and internal databases. Thus far, estimates differ widely between sources and from the clinical studies, demonstrating how using discontinuation derived from an ITC of the clinical studies is unlikely to reflect real-world adherence (see Table 6). Noticeably, in a recently published conference abstract, a 42% discontinuation rate was reported from data collected across five UK centres by hepatology pharmacists on all patients treated with ELA since it became licensed [8]. Gilead's internal data from IQVIA Longitudinal Access and Adjudication Data (LAAD) showed a discontinuation rate of 34% for ELA and 46% for SEL (6-month follow-up for both).

Given how different discontinuation is likely to be in clinical practice vs. the clinical studies, the counterintuitive results generated by different assumptions of treatment discontinuation and the high sensitivity of the model to discontinuation, we would propose assuming that all three comparators have the same discontinuation rates. This generates cost-effectiveness results underpinned by the key relative clinical effectiveness outcomes of liver biomarkers and pruritus, instead of cost-effectiveness results overly driven by discontinuation rate assumptions that have counterintuitive implications for the NHS.

Table 6 Additional evidence of real-world discontinuation rates

Source	Treatment	Discontinuation rate
Ronca et al., 2025 [9]	OCA	33% (retrospective study; no period specified)
Roberts et al., 2020 [10]	OCA	17% (retrospective study; no period specified)
Abbas et al. 2023 [11]	OCA	22.1% (annual)
Jones [8]	Elafibranor	42% (no period specified)
IQVIA Longitudinal Access and Adjudication Data (LAAD)	Elafibranor	34% (6-month)
	Seladelpar	46% (6-month)

Table 7 presents cost-effectiveness results when all three comparators assume the same short and long-term discontinuation rates (15.6% short term, the average of the results from the ITC; 0.9% long-term, based on the ratio of the long vs. short-term SEL discontinuation rates in

RESPONSE [0.12*7.8%=0.9%]). Results are presented using the current discounted price of [redacted] per pack for SEL and list price for OCA and ELA. All other assumptions are as per the committee's preferred base case post-ACM 2.

Table 7: cost-effectiveness results assuming equal short-term discontinuation rates of 15.6% and long-term rate of 0.9%

Intervention	Costs (£)	QALYs	Inc. costs (£)	Inc. QALYs	ICER
SEL + UDCA	[redacted]	[redacted]			
OCA + UDCA			[redacted]	[redacted]	Dominates
ELA + UDCA	[redacted]	[redacted]	[redacted]	[redacted]	Dominates

Note: results generated in committee's preferred base case EAG model, changing only discontinuation

Pruritus

For patients with PBC, chronic pruritus is often described as an unrelenting and tortuous symptom that profoundly disrupts daily quality of life. Many patients recount the pruritus as invasive, pervasive, and mentally consuming. One patient described how pruritus was taking over their thought and focus, causing them to "struggle to do ordinary things because I couldn't stop itching" [4], while another described the pruritus "as being like a car alarm. Like, you can't stop thinking about it" [11]. The severity of itching frequently intensifies at night, causing sleep deprivation and exhaustion, irritability, and emotional fragility during the day. Physical scratching, while momentarily relieving, often results in painful skin damage. One patient highlighted that "I would sometimes scratch them [my arms] so much, particularly during the night, I would draw blood", while another highlighted that "I used to go to work and go to the toilet to scratch them [my feet]. I used to be driving and wanted to pull over to scratch them – I even took a knife to them a few times" [12]. Patients also report social withdrawal, anxiety, and profound frustration, noting that pruritus is often the most debilitating and psychologically tormenting symptom of PBC [13, 14]. Clinical analyses highlight that the relentless combination of physical torment, sleep disturbance, and emotional burden underscores the unmet need for a holistic treatment option that addresses the symptoms of PBC, including pruritus, and its severe impact on quality of life [15, 16].

These lived experiences are reflected in our patient survey (n = 152), commissioned by Gilead through a third-party consultancy with UK PBC foundation. By the data cut-off date (March 12th), a total of 152 responses were received. Over half of respondents (55%) described feeling frustrated by their itch, with many also reporting embarrassment, low mood, and anxiety (Figure 1). Yet itch remains deprioritised in routine care. Only 36% reported that itch was discussed at clinic appointments, compared to 86% reporting discussion of test results (Figure 2). The impact on daily life is considerable: in the preceding four weeks, 47% reported itch-disturbed sleep and a third had scratched their skin raw (Figure 3).

Figure 1 PBC foundation survey: how itch makes patients feel

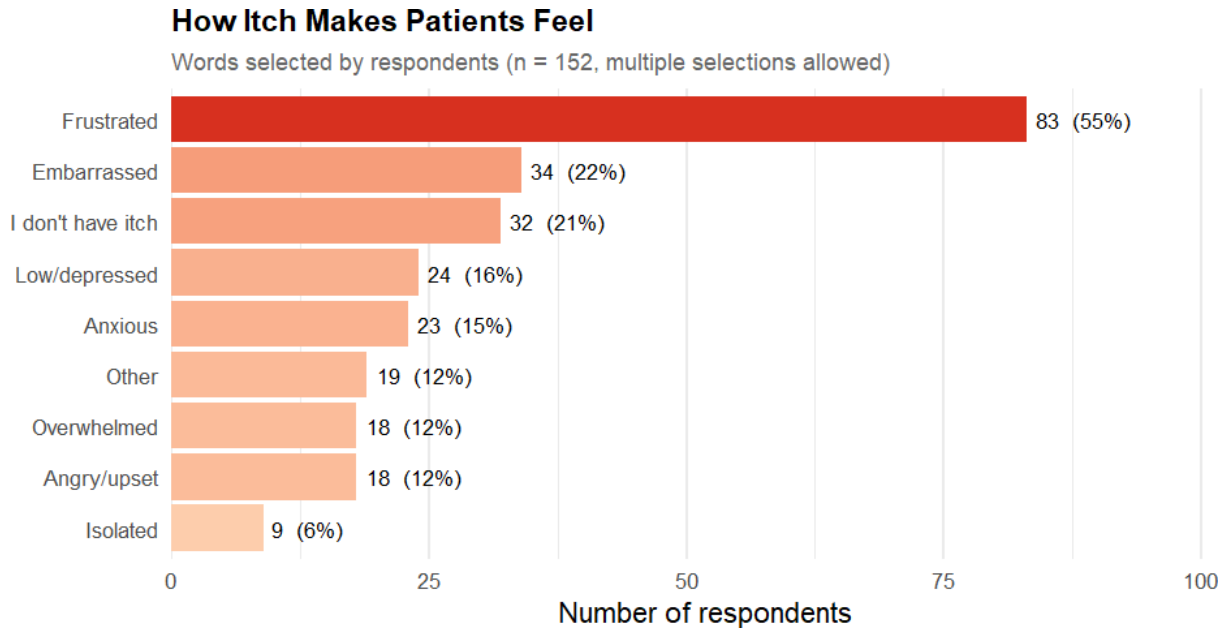


Figure 2 PBC foundation survey: what is discussed at liver clinic appointments

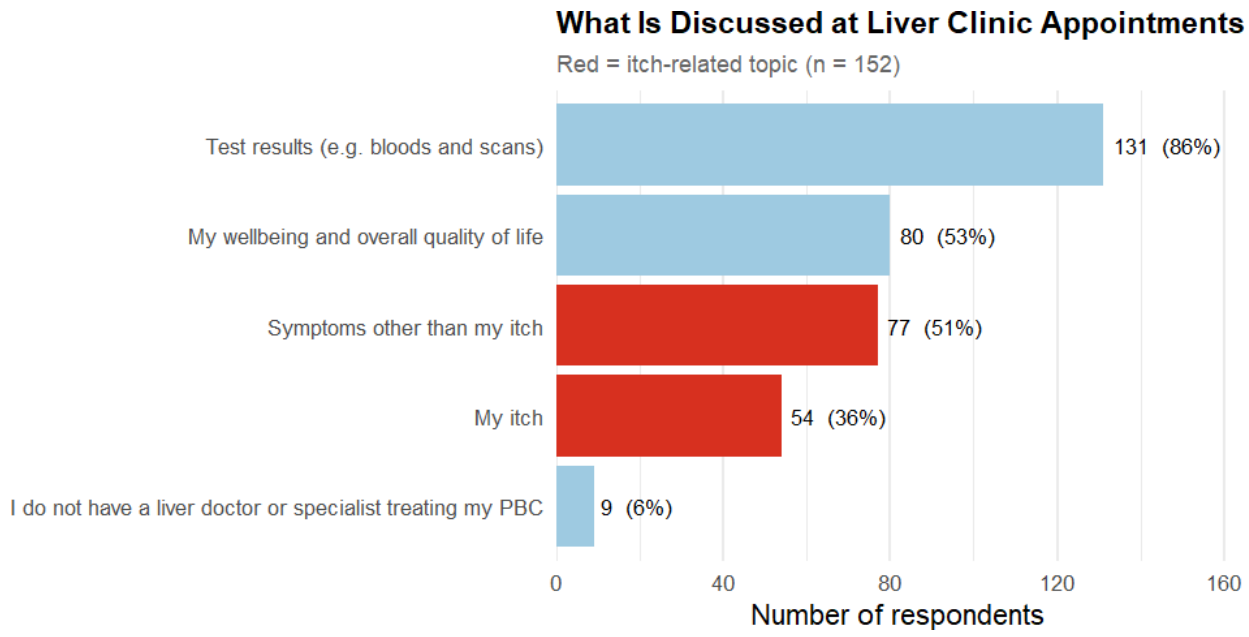
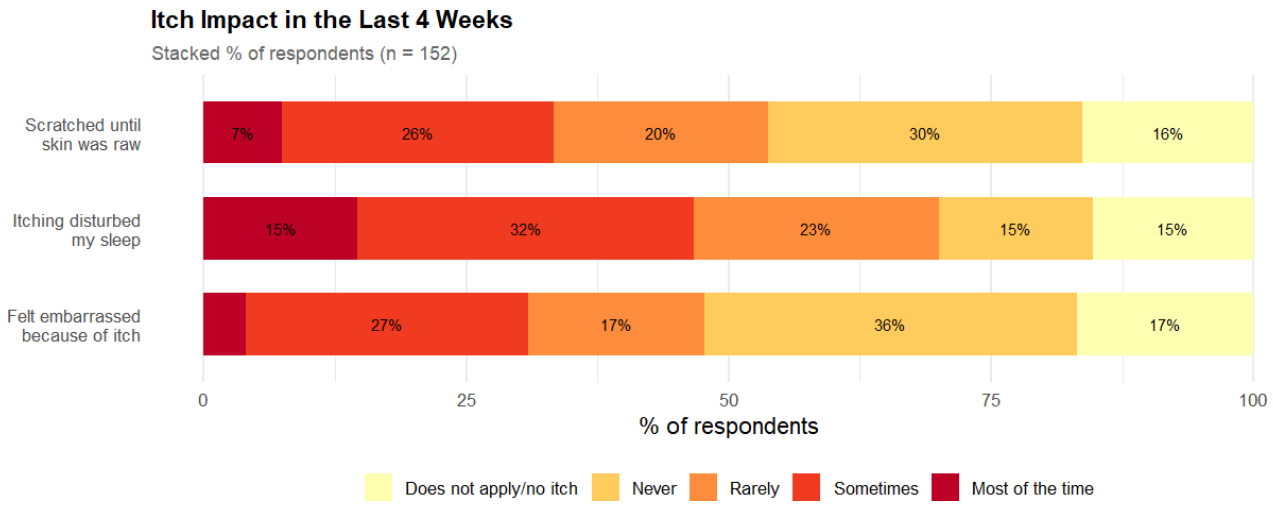


Figure 3 PBC foundation survey: itch impact in the last 4 weeks



It is against this backdrop that Gilead raises serious concerns regarding the disutility values for pruritus in the committee's preferred base-case, which is based on the mapping from the ITCH-E study. The scale of the impact described above is simply not reflected in the values currently preferred by the committee. As explained during consultation, this disconnect stems from the fundamental methodological limitations in the mapping study, which Gilead considers render submitted mapped disutilities for pruritus from the ITCH-E study invalid. This important context for preferring disutility values from the literature, which was omitted from the ACM 2 committee slides, is repeated below:

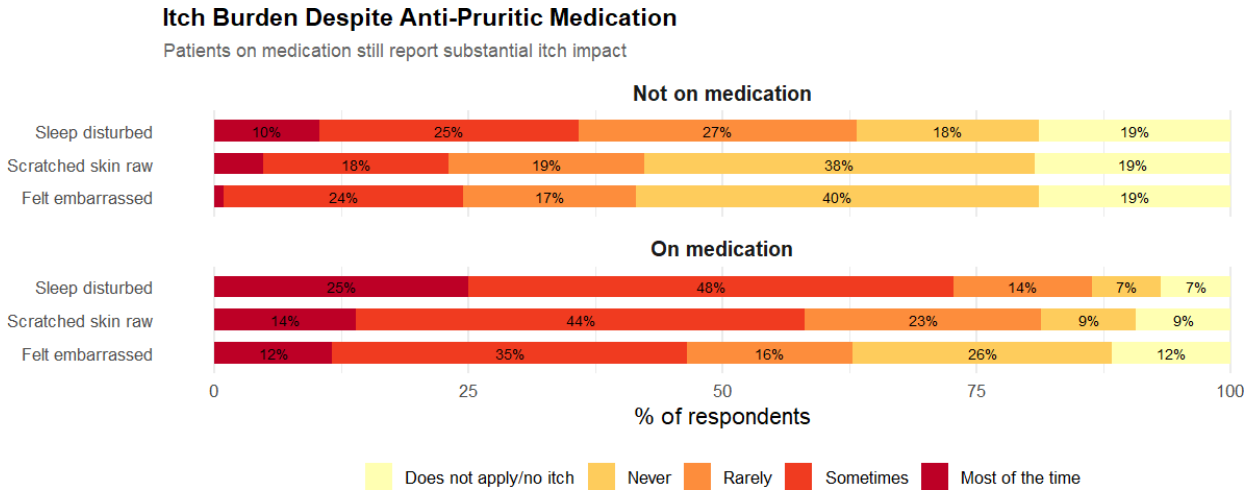
- The economic model captures severity of pruritus on the numerical rating scale (NRS). The severity is stratified into four categories, None (zero), Mild (1-3), Moderate (4-7) and Severe (≥7)
- In contrast, the disutilities generated from the ITCH-E mapping study stratified pruritus severity on the PBC-40 itch scale, into None (0), Mild (1-7) and Clinically Severe (≥7)
- In the economic model, the disutilties mapped from patients in ITCH-E with Clinically Severe (≥7) scores on the PBC-40 instrument were allocated to patients in the model with Severe (≥7) NRS scores
- Analysis of the RESPONSE data (which measured pruritus on both the PBC-40 instrument and NRS scale) shows that there is little overlap between PBC-40 Clinically Severe scores and NRS Severe scores:
    - At baseline, of the patients who were classed as Clinically Severe on the PBC-40, 21% and 41% were classed as mild or moderate on the NRS scale, respectively
    - At month 12, of the patients who were classed as Clinically Severe on the PBC-40, 50% and 37% were classed as mild or moderate on the NRS scale, respectively
- This means that the disutilities for Severe pruritus from the ITCH-E mapping study almost certainly significantly under-represented the magnitude of the disutility of Severe pruritus on the NRS scale, since the utility values would have been sampled from patients with mild and moderate pruritus on the NRS scale.
- Further stratification of the PBC-40 pruritis scores to generate separate Moderate and Severe scores from the aggregate Clinically Severe score would not be appropriate, since the two did not correlate well with NRS severity in RESPONSE.

- It is therefore not possible to use utility data for the liver biomarker and pruritus health states from a single consistent source, the stated preference of the committee.

For the above reasons, Gilead strongly advocated for use of values from literature in preference to mapped disutilities from ITCH-E. Gilead would like to counter the Rice et al. (2021) [17] paper cited in the ACM2 document by the EAG to support small disutility values for pruritus. In this study, pruritus was not captured using a validated severity scale such as the NRS or PBC-40 pruritus domain; therefore, the severity distribution among the 48.2% of pre-transplant patients reporting itching is unknown. The resulting disutility of −0.015 may be heavily diluted by milder cases and should not be considered reflective of the impact on patients with moderate-to-severe pruritus relevant to this appraisal.

The authors suggest the low impact may reflect the effect of treatment; however, this does not alter the interpretation, as patients whose pruritus is well controlled would simply present with lower severity at the time of measurement. Gilead would also highlight extensive publications from the PBC Foundation reporting that current pruritus medications offer little to no relief for PBC patients [18, 19], contradicting the assumption that treatment effects adequately account for the low disutility observed in this study. This is consistent with the survey findings with patients currently taking anti-pruritic agents reporting comparable or worse outcomes across all measures (Figure 4).

Figure 4 PBC foundation survey: itch burden despite anti-pruritic medication



Further, Gilead acknowledges the committee's concerns regarding the potential double-counting of pruritus disutilities and ALP levels. We believe these concerns are sufficiently addressed by the additional evidence detailed below, since we adjust for factors such as ALP level and fatigue when generating the disutility values.

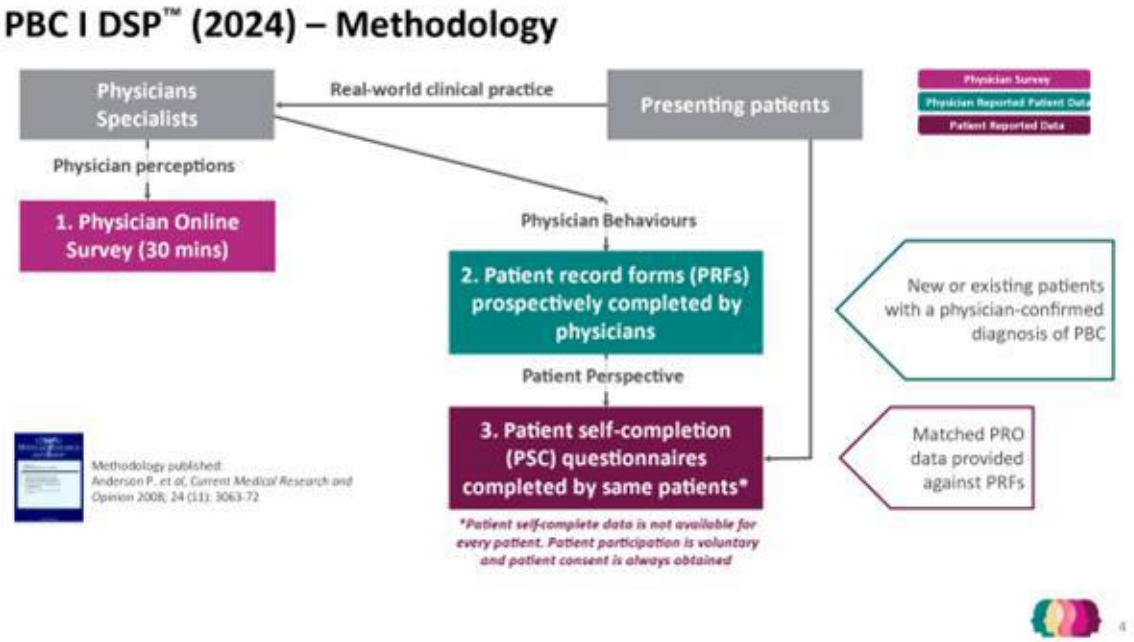
To further support the invalidity of the mapped values from ITCH-E used in the submission Gilead has conducted a quantitative analysis of an [Adelphi] Disease Specific Programmes (DSP) study in PBC to capture severity on the NRS scale, which aligns with the measure used in the economic model.

A summary of the DSP study and the disutilities generated from it is provided below (all results presented in this section are run with the updated PAS price [redacted] for SEL:

Adelphi DSP QoL Data

Adelphi DSPs are large, multinational, observational studies of clinical practice for a range of common chronic diseases. Its methodology has been published and validated [20, 21]. Patients with PBC were recruited as part of routine clinical practice, and patient-reported outcomes (PROs) were collected, including pruritus on the NRS scale, the PBC-40 and EQ-5D questionnaires, to assess disease-specific symptoms and generic health status. The DSP included a combination of abstracted physician- reported medical record data and patient-reported survey data. Figure 5 below presents the PBC I DSP methodology.

Figure 5 PBC DSP methodology



Adjusted differences in EQ-5D utility were estimated using a multivariable linear regression model, with EQ-5D utility as the dependent variable and pruritus severity category (on NRS scale) as the primary explanatory variable.

The model adjusted for available demographic and clinical covariates in the Adelphi DSP dataset (age, gender, ethnicity and insurance, age, gender, ethnicity, insurance status, BMI, fatigue, use of anti-pruritus medications (currently), and ALP levels). Notably, the inclusion of ALP levels as a covariate helps address the committee's concern regarding the potential double-counting of the effects of ALP levels and pruritus. Additionally, Gilead has adjusted for all known covariates reported to have an impact on the quality of life of pruritus patients [22], where data were available within the datasets. As these adjusted disutility values were elicited directly on the EQ-5D, they sit higher within NICE's hierarchy of preferred source of utility values than those mapped from the PBC-40 in RESPONSE, hence are included in our updated base case. Descriptive data of the baseline characteristics of the DSP dataset are presented in Table 11.

The results of the re-analysis are presented Table 8 and the committee base-case post ACM 2 with pruritus disutility from the re-analysis is presented in Table 9.

Table 8 Analysis of Adelphi DSP data, with stratification of pruritus severity on NRS scale

Pruritus category	Mean	Adjusted disutility	95% CI	P - Value

			Lower	Upper	
No Pruritus	0.86				
Mild Pruritus (NRS ≥1-<4)	0.796	[redacted]	[redacted]	[redacted]	0.8538
Moderate Pruritus (NRS ≥4–<7)	0.778	[redacted]	[redacted]	[redacted]	0.0573
Severe Pruritus (NRS ≥7)	0.707	[redacted]	[redacted]	[redacted]	<0.0001*

**statistically significant*

Table 9 Committee base case post-ACM 2, updated with pruritus disutilities from analysis of Adelphi DPS data

Intervention	Costs (£)	QALYs	Inc. costs (£)	Inc. QALYs	ICER
Committee base case post-ACM 2, but with pruritus disutilities from analysis of Adelphi DPS data					
SEL + UDCA	[redacted]	[redacted]			
OCA + UDCA	[redacted]		[redacted]	[redacted]	Dominates
ELA + UDCA	[redacted]				Dominates

Note: results generated in committee's preferred base case EAG model, changing only pruritus disutilities

Cost-Effectiveness Results

Table 10 presents cost-effectiveness results generated from pruritus disutility values from Adelphi DSP dataset. We consider these to be the most relevant to decision-making because they address the key limitations identified with the ITCH-E mapped values. As set out above, the ITCH-E mapping approach almost certainly significantly underestimated the disutility of severe pruritus, since the utility values would have been derived from patients with mild and moderate pruritus on the NRS scale. Further stratification of PBC-40 pruritus scores to generate separate moderate and severe disutilities was not feasible given the poor correlation observed in RESPONSE, making it impossible to use utility data for both liver biomarker and pruritus health states from a single consistent source — the committee's stated preference. The DSP study overcomes these limitations. The disutility values were elicited directly using the EQ-5D, placing them higher within NICE's hierarchy of preferred utility sources than those mapped from the PBC-40. The underlying model adjusted for all available demographic and clinical covariates, including ALP levels, directly addressing the committee's concern regarding potential double-counting of the effects of ALP and pruritus. By adjusting for all known covariates reported to influence quality of life in pruritus patients, where data permitted, these values provide a more methodologically robust and clinically reflective estimate of the true burden of pruritus.

Results are presented using the committee's preferred ITC-based discontinuation rates post ACM 2 (top half of Table 10) and with Gilead's preferred base case for ACM 3 (the same discontinuation rate assumed for all three comparators). Results are run using updated PAS price of [redacted] for SEL.

Table 10: cost-effectiveness results using pruritus disutilities from DSP

Intervention	Costs (£)	QALYs	Inc. costs (£)	Inc. QALYs	ICER
Committee base case post-ACM 2, but with pruritus disutilities from DSP					
SEL + UDCA	[redacted]	[redacted]			
OCA + UDCA	[redacted]		[redacted]	[redacted]	Dominates
ELA + UDCA	[redacted]				Dominates

As above, but assuming equal discontinuation rates across all three comparators (Base case)												
SEL + UDCA	[redacted]											
OCA + UDCA	[redacted]											Dominates
ELA + UDCA	[redacted]											Dominates

Note: results generated in committee's preferred base case EAG model, changing only pruritus disutilities and discontinuation

References

1. Nevens, F., et al., *A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis.* New England Journal of Medicine, 2016. **375**(7): p. 631-643.
2. Hirschfield, G.M., et al., *A Phase 3 Trial of Seladelpar in Primary Biliary Cholangitis.* N Engl J Med, 2024. **390**(9): p. 783-794.
3. Kowdley, K.V., et al., *Efficacy and Safety of Elafibranor in Primary Biliary Cholangitis.* N Engl J Med, 2024. **390**(9): p. 795-805.
4. Patient Worthy. *An Itch That Was Taking Over My Thoughts: A Story About Primary Biliary Cholangitis*. 2026; Available from: https://patientworthy.com/2026/02/25/an-itch-that-was-taking-over-my-thoughts-a-story-about-primary-biliary-cholangitis.
5. Gilead, *Data on File - Clinical Efficacy, Safety and Tolerability of Approved Medications for the Treatment of Primary Biliary Cholangitis*. 2025.
6. National Institute for Health and Care Excellence, *Obeticholic acid for treating primary biliary cholangitis.* 2017.
7. Jones, A., et al., *P34 Initial clinical response and safety of elafibranor in primary biliary cholangitis: a multi-centre pharmacist-led real-world evaluation.* Gut, 2025. **74**(Suppl 4): p. A35-A35.
8. Ronca, V., et al., *Fibrate-Oca (FI-OCA) Study - a Global snapshot of PBC practice around the Globe.* Digestive and Liver Disease, 2025. **57**: p. S35-S36.
9. Roberts, S.B., et al., *Real-World Effectiveness of Obeticholic Acid in Patients with Primary Biliary Cholangitis.* Hepatol Commun, 2020. **4**(9): p. 1332-1345.
10. Abbas, N., et al., *UK-Wide Multicenter Evaluation of Second-line Therapies in Primary Biliary Cholangitis.* Clin Gastroenterol Hepatol, 2023. **21**(6): p. 1561-1570.e13.
11. NPR. *Chronic itch is miserable. Scientists are just scratching the surface*. 2024; Available from: https://www.npr.org/2024/11/12/nx-s1-5186946/chronic-itch-primary-biliary-cholangitis.
12. BBC News. *'Doctors mistook my liver disease for menopause'*. 2025; Available from: https://www.bbc.co.uk/news/articles/ced99wx4y0eo.
13. Dutch Liver Patient Association. *Primary biliary cholangitis - 12 patient stories*. 2022; Available from: https://rare-liver.eu/media/pbc_-_12_patient_stories.eng.pdf.
14. PBC Foundation. *Primary biliary cholangitis: A snapshot of the experiences of patients and clinicians in Scotland*. 2025; Available from: https://www.pbcfoundation.international/wp-content/uploads/2025/08/PBC-in-Scotland-Report.pdf.
15. Corrigan, M., et al., *Primary biliary cholangitis: the lived experience and its impact on healthcare preferences – a qualitative study.* Frontline Gastroenterology, 2026: p. flgastro-2025-103288.
16. Hirschfield, G., et al., *Holistic Patient Care in Primary Biliary Cholangitis: Managing Both the Disease and the Symptoms.* EMJ Hepatology, 2022: p. 22-31.
17. Consortium, U.-P., et al., *Effects of Primary Biliary Cholangitis on Quality of Life and Health Care Costs in the United Kingdom.* Clinical Gastroenterology and Hepatology, 2021. **19**(4): p. 768-776.e10.
18. Mitchell-Thain, R., et al., *Unmet Need in Cholestatic Pruritus Emphasized by Patients with PBC: Results from PBC Foundation App Survey.* 2021.
19. Mitchell-Thain, R., et al. *Unmet Need in Cholestatic Pruritus Emphasized by Patients with PBC: Results from PBC Foundation App Survey*. 2021.
20. Anderson, P., et al., *Real-world evidence generation from patients, their caregivers and physicians supporting clinical, regulatory and guideline decisions: an update on Disease Specific Programmes.* Curr Med Res Opin, 2023. **39**(12): p. 1707-1715.

21. Anderson, P., et al., *Real-world physician and patient behaviour across countries: Disease-Specific Programmes - a means to understand.* Curr Med Res Opin, 2008. **24**(11): p. 3063-72.
22. Sierra, L., et al., *Primary biliary cholangitis has the worst quality of life indicators among the autoimmune liver diseases: A United States cohort.* Annals of Hepatology, 2025. **30**(2): p. 101945.

Appendix

Table 11 Adelphi DSP Descriptive analysis

Baseline Characteristics	Overall N = 207[1]	No Pruritus N = 77[1]	Mild Pruritus (NRS ≥1-<4) N = 61[1]	Moderate Pruritus (NRS ≥4–<7) N = 51[1]	Severe Pruritus (NRS ≥7) N = 18[1]
Age					
N	207.0	77.0	61.0	51.0	18.0
Mean (SD)	53.7 (10.6)	55.2 (10.8)	52.5 (11.7)	54.3 (9.3)	49.1 (7.4)
Median	55.0	56.0	54.0	55.0	49.0
Q1, Q3	46.0, 60.0	46.0, 62.0	43.0, 61.0	48.0, 60.0	45.0, 55.0
Min, Max	21.0, 85.0	21.0, 82.0	27.0, 76.0	30.0, 85.0	31.0, 60.0
Sex					
Female	181 (87%)	67 (87%)	54 (89%)	42 (82%)	18 (100%)
Male	26 (13%)	10 (13%)	7 (11%)	9 (18%)	0 (0%)
Ethnicity					
White	159 (77%)	59 (77%)	46 (75%)	41 (80%)	13 (72%)
Other	1 (0.5%)	0 (0%)	0 (0%)	0 (0%)	1 (5.6%)
Not Available	47 (23%)	18 (23%)	15 (25%)	10 (20%)	4 (22%)
Insurance					
Don't know	2 (1.0%)	1 (1.3%)	0 (0%)	0 (0%)	1 (5.6%)
Gesetzliche Krankenversicherung	67 (32%)	17 (22%)	23 (38%)	21 (41%)	6 (33%)
Gesetzliche mit privater Zusatzversicherung	5 (2.4%)	3 (3.9%)	0 (0%)	2 (3.9%)	0 (0%)
NHS	6 (2.9%)	1 (1.3%)	0 (0%)	4 (7.8%)	1 (5.6%)
Private Krankenversicherung	8 (3.9%)	1 (1.3%)	5 (8.2%)	1 (2.0%)	1 (5.6%)
Protection Universelle Maladie (PUMa)	19 (9.2%)	6 (7.8%)	8 (13%)	4 (7.8%)	1 (5.6%)
PUMa + CMU-C, mutuelle ou assurance privÃ©e	26 (13%)	11 (14%)	7 (11%)	6 (12%)	2 (11%)
Servizio Sanitario Nazionale	27 (13%)	13 (17%)	11 (18%)	3 (5.9%)	0 (0%)

Servizio Sanitario Nazionale e Assicurazione sanitaria privata	1 (0.5%)	1 (1.3%)	0 (0%)	0 (0%)	0 (0%)
Sistema Nacional de Salud (SNS)	42 (20%)	20 (26%)	6 (9.8%)	10 (20%)	6 (33%)
Sistema Nacional de Salud (SNS), y seguro mÃ©dico privado	4 (1.9%)	3 (3.9%)	1 (1.6%)	0 (0%)	0 (0%)
BMI					
N	207.0	77.0	61.0	51.0	18.0
Mean (SD)	24.7 (3.8)	24.8 (4.6)	24.3 (2.9)	25.1 (3.7)	24.7 (3.6)
Median	24.2	24.5	24.1	24.6	24.3
Q1, Q3	22.1, 26.3	21.8, 26.8	22.5, 26.2	23.0, 26.9	22.0, 26.1
Min, Max	17.6, 45.0	18.7, 45.0	17.6, 33.5	19.6, 37.2	20.0, 32.4
Fatigue[2]					
Yes	131 (63%)	34 (44%)	43 (70%)	40 (78%)	14 (78%)
No	76 (37%)	43 (56%)	18 (30%)	11 (22%)	4 (22%)
Anti pruritus medication[3]					
Yes	71 (34%)	13 (17%)	19 (31%)	34 (67%)	5 (28%)
No	136 (66%)	64 (83%)	42 (69%)	17 (33%)	13 (72%)
ALP[4]					
N	173.0	66.0	58.0	37.0	12.0
Mean (SD)	173.3 (193.6)	131.7 (72.7)	213.9 (306.9)	170.7 (88.8)	213.9 (142.5)
Median	127.0	110.5	146.5	150.0	147.0
Q1, Q3	100.0, 186.0	90.0, 140.0	105.0, 250.0	122.0, 190.0	120.0, 310.0
Min, Max	20.0, 2,399.5	20.0, 400.0	30.0, 2,399.5	30.0, 420.0	84.0, 474.0

1 N : Number of observations with non-missing values.

2 Fatigue: binary (Yes/No).

3 Anti pruritus medication: Patient currently receiving prescribed treatment to treat pruritus/itch caused by their PBC/PSC.

4 ALP: Alkaline phosphatase (U/L) (Most recent test result).

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Addendum: EAG response to additional company submission prior to ACM3

April, 2026

Produced by	Peninsula Technology Assessment Group (PenTAG) University of Exeter Medical School
Authors	Caroline Farmer[1] Jemma Perks[1] Alan Lovell[1] Maxwell S. Barnish[1] Valdemar Santo[1] Ahmed Abdelsabour[1] Saul Stevens[1] Dawn Lee[1] Taha Khan[2] Rebecca Jones[3] Gwilym Webb[4] Edward CF Wilson[1] [1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter [2] St Leonard's Practice, Exeter [3] Leeds Teaching Hospital NHS Trust [4] Cambridge University Hospitals NHS Foundation Trust
Correspondence to	Caroline Farmer; c.farmer@exeter.ac.uk

Source of funding	This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR175575.
Declared competing interests of the authors	None

This report should be referenced as follows: Farmer, Perks, Lovell, Barnish, Santo, Abdelsabour, Lee, Khan, Jones, Webb, Wilson. Seladelpar for previously treated primary biliary cholangitis [ID6429]: EAG response to additional company submission prior to ACM3. Peninsula Technology Assessment Group (PenTAG), 2026.

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

Abbreviations

Term	Definition
ACM	Appraisal committee meeting
ALP	Alkaline phosphatase
BMI	Body Mass Index
CI	Confidence interval
DE	Germany
DSP	Disease Specific Programmes
EAG	External Assessment Group
ELA	Elafibranor
ES	Spain
HRQoL	Health Related Quality of Life
ICER	Incremental cost-effectiveness ratio
IT	Italy
ITC	Indirect treatment comparison
LAAD	Longitudinal Access and Adjudication Data
MMRM	Mixed model for repeated measures
NHB	Net health benefit
NICE	National Institute for Health and Care Excellence
NRS	Numerical rating scale
OCA	Obeticholic acid
OLE	Open label extension
OLS	Ordinary Least Squares
PBC	Primary biliary cholangitis
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
SNS	Sistema Nacional de Salud
UDCA	Ursodeoxycholic acid
UK	United Kingdom
U/L	Units per litre

1. INTRODUCTION

The purpose of this addendum is to summarise the appraisal by the External Assessment Group (EAG) of evidence submitted by the company prior to a third NICE appraisal committee meeting (ACM) for seladelpar for the treatment of primary biliary cholangitis (PBC) following intolerance or inadequate response to ursodeoxycholic acid (UDCA).

Following ACM2, the Committee preferred:

(1) the biomarker response rate indirect treatment comparison (ITC) unadjusted for differing alkaline phosphatase (ALP) normalisation thresholds between trials

(2) discontinuation estimated from the ITC, using the seladelpar (SEL) RESPONSE discontinuation rate as the reference

(3) pruritus disutility values estimated via mapping from the ITCH-E study.

In its submission, the company accepted (1) that the biomarker response rate should be unadjusted for ALP normalisation thresholds but presented evidence to challenge the NICE committee's preferences related to (2) the source of discontinuation rates and (3) pruritus disutility values used in the economic model.

The Company also offered an additional reduction in price (Table 1).

Table 1 Discounted price of seladelpar

Previous discounted price	[redacted]
New discounted price	[redacted]

The EAG appraisal of the Company's evidence is presented in Section 2 for discontinuation rates and Section 3 for pruritus disutility value. Section 4 summarises the EAG's suggested base case in the light of the evidence and Section 5 presents results of the scenario analyses. A confidential appendix accompanies this report showing recalculated results tables with confidential discounts for comparator drugs included.

2. SOURCE OF DISCONTINUATION RATES

The company's post ACM2 submission notes that discontinuation rates are a key driver of cost-effectiveness results.

2.1. Background

In the company's model, discontinuation from treatment was modelled in two phases: 0-12m, and 12m+. Discontinuation in the first 12m was based on trial data, and post 12m was based on a relative discontinuation rate observed within short term (12m trials) and longer-term open label extensions.

2.2. Post-12m discontinuation: EAG coding misinterpretation

At time of writing, there is agreement between the company and EAG as to the source for the longer-term discontinuation rate (based on relative annual discontinuation rate observed between RESPONSE and ASSURE, the pivotal RCT and OLE for seladelpar). However, the EAG's intention was for this to be the source of the *relative* rate of discontinuation across all three comparators, rather than the *absolute* discontinuation rate post 12m to be equal across all comparators. The EAG had inadvertently coded this as equal absolute discontinuation rates rather than equal relative rates. The EAG has added a correction for this in its analyses below (scenarios 7 onwards, section 5), but notes that this point is moot if equal discontinuation rates within the first 12m are accepted.

2.3. Year 1 discontinuation

In its original (pre-AC1) submission, the company obtained all cause discontinuation rates for all treatments from their respective pivotal RCTs (RESPONSE for seladelpar, ELATIVE for elafibranor and POISE for OCA). The EAG critiqued this on the basis that these were naïve comparisons (i.e. raw results from the trials without adjusting for differences in populations and/or study designs) and noted that the company had conducted an ITC for discontinuation rates but not used the results from this in its analysis. As the purpose of the ITC is to correct for differences in populations, the EAG preferred use of these as a source for discontinuation rates.

2.4. Company submission Post ACM2

Following ACM2, the committee requested the company to use the ITC as its primary source of discontinuation. The company responded arguing that the ITC results were implausible due to

the substantially lower discontinuation rate in the placebo arm of the POISE study, possibly caused by the exclusion of patients with severe pruritus from that study (Table 2). On reviewing the baseline characteristics for the control arms for the three trials, the EAG was unable to determine any reason for the discontinuation rate to be systematically lower in the POISE study compared with RESPONSE or ELATIVE. Furthermore, the EAG noted the low numbers of events and the corresponding high degree of uncertainty.

Table 2: Discontinuation at 12 Months (Descriptive Statistics)

Intervention	Active			Control (UDCA / Placebo)			RR
	n	N	Rate	n	N	Rate	
SEL	10	128	7.81%	8	65	12.31%	0.63
ELA	12	108	11.11%	6	53	11.31%	0.98
OCA	16	144	11.11%	3	73	4.11%	2.70

Abbreviations: ELA: Elafibranor; OCA: Obeticholic Acid; SEL: Seladelpar; UDCA: Ursodeoxycholic Acid

Source: Reproduced from company post ACM2 additional evidence 16032026, Table 2. Data sources: SEL – RESPONSE (Hirschfield et al. 2024[1]); ELA – ELATIVE (Kowdley et al. 2024[2]); OCA – POISE (Nevens et al. 2016[3]).

In its response, the Company also provided an example of counterintuitive results with a scenario from its model where cost-effectiveness deteriorates with higher adherence (or rather, higher discontinuation rates for SEL lead to lower ICERs). The EAG notes that this is not necessarily a counterintuitive result as for very high-cost drugs the dynamics of the model can be such that shorter treatment periods can indeed be more cost-effective than longer ones (specifically, if incremental costs accrue faster than incremental benefits, the ICER will deteriorate with a longer time on treatment).

The Company then presented a review showing the variance in discontinuation rates across the literature and real world evidence sources, concluding that discontinuation is highly variable, and in particular noting higher discontinuation rates observed in retrospective and other real world data studies (Table 3).

The EAG noted the higher discontinuation rates from real world data, and further noted that OCA discontinuation appeared to be lower than either ELA or SEL. The IQVIA data in particular suggested very high discontinuation rates for ELA and SEL (34% and 46% at 6 months respectively, whereas trial data (Table 2) suggested ~11% at 12 months).

Table 3 Additional evidence of real-world discontinuation rates

Source	Treatment	Discontinuation rate
Ronca et al., 2025[4]	OCA	33% (retrospective study; no period specified)
Roberts et al., 2020[5]	OCA	17% (retrospective study; no period specified)
Abbas et al. 2023[6]	OCA	22.1% (annual)
Jones et al., 2025[7]	Elafibranor	42% (no period specified)
IQVIA Longitudinal Access and Adjudication Data (LAAD)*	Elafibranor	34% (6-month)
	Seladelpar	46% (6-month)

Abbreviation: OCA, Obeticholic acid

Source: Reproduced from company post ACM2 additional evidence 16032026, Table 6.

Note: *IQVIA LAAD data, internal to Gilead.

Given the data, the Company proposed assuming discontinuation rates to be equal across all three treatments, using the mean of the ITC rate of 15.6% over the first 12m (followed by a reduction to 0.9% for the remainder of the model time horizon, based on the relative discontinuation rate seen between RESPONSE and ASSURE).

2.5. EAG opinion

The EAG noted that the real-world evidence on discontinuation suggests much higher rates than observed in clinical trials. This is a common observation across all disease areas and treatments as trials are conducted under strict protocols; knowledge of taking part in a study may motivate patients to adhere to their treatments more than in 'real life' settings. The EAG considered that it is reasonable to assume an equal rate of discontinuation across all treatments, but that given the real-world data, levels higher than an annual rate of 15.6% should be explored. It therefore presented scenario analyses with annual rates of 20%, 30% and 40%.

3. SOURCE OF PRURITUS DISUTILITY VALUES

3.1. Background

EQ-5D data were not collected in the pivotal phase 3 study for seladelpar (RESPONSE).[1] Instead, the disease-specific PBC-40 was used. The Company converted these to EQ-5D-3L-based utilities by ALP level using a mapping algorithm via the ITCH-E study (algorithm "MMRM2" in the company submission),[8] and the model included covariates for pruritus. In summary, RESPONSE measured PBC-40. The ITCH-E study collected PBC-40 and EQ-5D data from a cohort of individuals with PBC currently receiving UDCA or OCA. The mapping study mapped EQ-5D data onto PBC-40, allowing conversion from PBC-40 data collected in RESPONSE to EQ-5D.

Whilst the Company chose to use the MMRM2 utilities for ALP levels, it rejected their use to estimate the impact of pruritus and subsequently conducted a systematic review of HRQoL in PBC. From this it identified estimates from Smith et al.,[9] reported as a conference abstract estimating EQ-5D-5L-based HRQoL as a function of itch severity from the baseline results of the GLIMMER study, a phase 2b dose finding study of linerixibat in patients with PBC with pruritus (Levy et al. 2022).[10] The EAG was unclear as to why the trial-based estimate and ITCH-E study were rejected by the Company and considered Smith et al. to lack face validity as the utility of the PBC population was higher than the general population. Furthermore, including extra external data adds additional uncertainties in terms of transferability.

Following ACM2, the Committee requested the Company use the pruritus disutilities from the ITCH-E study (i.e. MMRM2 algorithm).

3.2. Company submission Post ACM2

The company's submission comprised a survey conducted by the PBC foundation and a new analysis of quality of life data from the Adelphi Disease Specific Programmes (DSP) study in PBC.

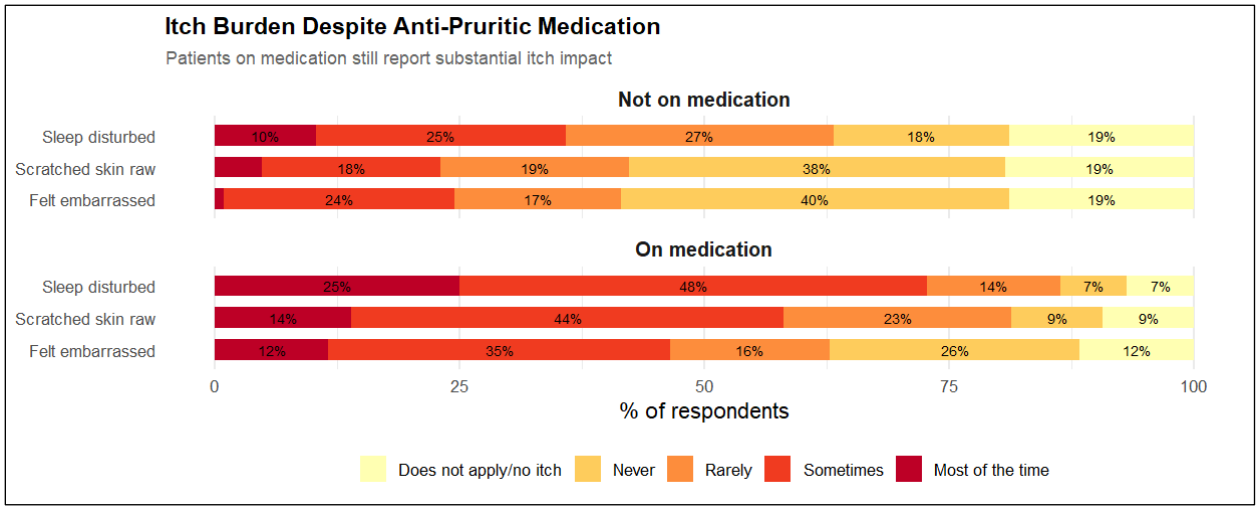
3.2.1. PBC foundation survey

The Company reported the results of a patient survey (n=152 as at data cut 12th March 2026). This showed that over half of respondents (55%) were frustrated by itch, with other terms highlighted by respondents including embarrassment (22%), feeling low/depressed (16%) and

anxious (15%), inter alia. Twenty one percent reported that they had no itch. Around 36% of respondents discussed itch at liver clinic appointments.

When asked about the impact of itch over the last 4 weeks on a 5 point Likert scale (the itch domain questions of the PBC-40), across all respondents (including those who reported no itch) 7% reported scratching until skin was raw most of the time, 15% reported sleep disturbance most of the time and 3% reported being embarrassed by their itch most of the time. Around one third of respondents reported these impacts most of the time or sometimes. Furthermore, the survey observed that the reports of disturbed sleep, raw skin and embarrassment were higher amongst those receiving medication for pruritus, suggesting a lack of efficacy of treatments.

Figure 1: PBC foundation survey: itch burden by anti-pruritic medication



Source: Reproduced from company post ACM2 additional evidence 16032026, Figure 4.

The Company therefore rejected the analysis based on the MMRM2 algorithm and ITCH-E study, considering the ITCH-E study to be invalid:

The Company noted that its model captured severity of pruritus using the numerical rating scale (NRS, a subjective score from 0-10) from the RESPONSE study data, stratified into none (zero), mild (1-3), moderate (4-7) and severe (7+) rather than the PBC-40 (which was also collected in the RESPONSE study). The ITCH-E study stratified by PBC-40 score from none (zero), mild (1-7) and clinically significant (7+). The company analysed the individual patient data from RESPONSE, comparing PBC-40 responses with NRS, finding that at baseline, 61% of those classed as clinically severe on the PBC-40 (i.e. scored 7+) were classified as mild or moderate

on the NRS (i.e. scored 1-7), and 87% of those classed as clinically severe on PBC-40 at 12 months were classed as mild or moderate on the NRS scale. In summary, the model will have underestimated the proportions of patients with severe itch (and hence the overall QALY impact of itch) by using the NRS rather than the PBC-40.

The EAG agrees that the survey data confirm that pruritus is highly prevalent amongst those with PBC, and that the symptoms appear worse amongst those receiving treatment, but the survey does not quantitively inform the health state utility associated with severe pruritus.

With respect to antipruritic effectiveness, whilst those with more severe itch are more likely to seek treatment, poor effectiveness is also a plausible explanation for more severe symptoms reported amongst those survey participants. Further data on which treatments participants were receiving would assist in determining plausibility of these hypotheses: the Company's model assumed all patients with pruritus would receive one of colestyramine, rifampicin, naltrexone, gabapentin or bezafibrate. Confirmation that all respondents in the survey receiving antipruritics were one of these rather than, for example, emollients or other treatments would improve confidence in the lack of efficacy argument.

3.2.2. Adelphi DSP data

The Company obtained Adelphi DSP data for PBC, which collects EQ-5D, PBC-40 and NRS data globally on people with PBC recruited to the study through routine clinical practice. The sample size was not reported but the Company conducted an OLS regression analysis of EQ-5D against pruritus by NRS score. Results presented in the Company's submission are reproduced in Table 4. The EAG noted that the unadjusted and adjusted incremental disutilities were substantially different for both mild and severe pruritus (but not for moderate). For mild the adjusted decrement was approximately one tenth smaller than the unadjusted ([redacted] vs [redacted]) and for severe the adjusted was almost twice that of the unadjusted ([redacted] vs [redacted]).

Table 4 Analysis of Adelphi DSP data, with stratification of pruritus severity on NRS scale

Pruritus category	Mean	Unadjusted disutility	Adjusted disutility	95% CI		P - Value
				Lower	Upper	
No Pruritus	0.86					
Mild Pruritus (NRS ≥1-<4)	0.796	[redacted]	[redacted]	[redacted]	[redacted]	0.8538
Moderate Pruritus (NRS ≥4–<7)	0.778	[redacted]	[redacted]	[redacted]	[redacted]	0.0573
Severe Pruritus (NRS ≥7)	0.707	[redacted]	[redacted]	[redacted]	[redacted]	<0.0001*

Abbreviations: CI, confidence internal; NRS, numerical rating scale

Source: Adapted from company post ACM2 additional evidence 16032026, Table 8 (with addition of unadjusted utility column).

Note: *statistically significant

Whilst the purpose of adjustment is to factor out confounding factors, the EAG noted the large discrepancy between adjusted and unadjusted results and requested further details from the Company. The Company provided the code and output of relevant data frames in R. Reviewing the coefficients and standard errors of all explanatory variables, the EAG noted that the intercept was above the logical maximum utility of 1 (1.48, SE 0.18), and that as well as severe pruritus ([redacted]), other statistically significant variables were presence of fatigue ([redacted]) and white ethnicity ([redacted]). Furthermore, health insurance status appeared to be disaggregated into more variables than required, with separate variables for Germany, UK, Italy and Spain for state provision with or without private top-up insurance, potentially reducing the statistical power of the analysis (Table 5).

The results suggest that ethnicity is a much bigger determinant of health state utility (with white respondents experiencing a disutility of [redacted]) than severe pruritus ([redacted]). The EAG considers this to lack face validity suggesting that the model may be mis-specified. In particular, there were no interaction terms (e.g. pruritus with fatigue) and a linear model is less suited to analysis of data with a logical upper limit of 1. The EAG noted that the ITCH-E study explored a large number of alternative model structures, concluding that OLS provided the best fit to map EQ-5D to PBC-40; similar explorations were not conducted with the Adelphi dataset, and it is unclear whether the same model structure would provide the same degree of fit.

Table 5: Regression coefficients from OLS analysis of Adelphi DSP data

Coefficients	Estimate	Std. Error	Pr(>| t|)
(Intercept)	[redacted]	[redacted]	<0.0000000001***
Mild Pruritus (NRS ≥1-<4)	[redacted]	[redacted]	0.85328
Moderate Pruritus (NRS ≥4–<7)	[redacted]	[redacted]	0.05735
Severe Pruritus (NRS ≥7)	[redacted]	[redacted]	<0.00001***
Fatigue	[redacted]	[redacted]	0.00106**
Age	[redacted]	[redacted]	0.77972
Male	[redacted]	[redacted]	0.18422
White ethnicity	[redacted]	[redacted]	0.0034**

Coefficients	Estimate	Std. Error	Pr(>| t|)
Statutory insurance with private topup (DE)	[redacted]	[redacted]	0.89033
State health insurance (UK)	[redacted]	[redacted]	0.81495
Private health insurance (DE)	[redacted]	[redacted]	0.39714
State health insurance (IT)	[redacted]	[redacted]	0.53515
Statutory insurance with private topup (IT)	[redacted]	[redacted]	0.87045
NHS (SNS, ES)	[redacted]	[redacted]	0.55075
NHS with private topup (IT)	[redacted]	[redacted]	0.05472
BMI	[redacted]	[redacted]	0.20935
Receiving antipruritics	[redacted]	[redacted]	0.82384
ALP in U/L	[redacted]	[redacted]	0.3505

Abbreviations: ALP, alkaline phosphatase; BMI, body mass index; DE, Germany; ES, Spain; IT, Italy; NRS, numerical rating scale; SNS, Sistema Nacional de Salud; Std, standard; UK, United Kingdom; U/L, units per litre.

*** = p<0.001, ** = p<0.01, * = p<0.05.

Overall, the EAG considered the Adelphi DSP analysis to be poorly conducted and generated results that lacked face validity.

3.3. EAG opinion

In summary there are two issues to consider with respect to the impact of pruritus on the cost-effectiveness of seladelpar.

Firstly, the Company used the NRS score rather than the PBC-40 score to define severity of itch. The Company stated that this mis-classified most patients in the model as mild/moderate rather than severe.

Secondly, the Company criticised the ITCH-E study on the basis of the disutilities not reflecting the results of the PBC survey, showing that pruritus affected the majority of those with PBC, with around 1/3 of patients reporting an impact either ‘sometimes’ or ‘most of the time’.

The EAG considers the first issue can be readily corrected by modifying the Company’s model to base the severity of itch on PBC-40 score rather than NRS. The EAG is unable to do this as it requires access to the individual patient data.

With respect to the second issue, the EAG considers that the survey shows the prevalence of itch is high but does not quantitatively inform the disutility associated with itch per se and so

does not provide grounds for rejection of the ITCH-E study. The EAG is also concerned that the Company's critique of the ITCH-E study amounts to post-hoc rationalisation when the study was designed a priori to answer the relevant question. However, the EAG is minded to agree that the disutility of severe pruritus is likely higher than suggested by the MMRM2 algorithm from the ITCH-E study (-0.0345), and that the PBC foundation survey could be interpreted as showing a lack of efficacy of antipruritic medications. Nonetheless, the EAG has substantial reservations as to the quality of the Adelphi DSP analysis due to the lack of face validity of the complete results and unclear coding of parameters. The EAG therefore conducted a sensitivity analysis with a disutility of -0.05, -0.1 and -0.15 for the impact of severe pruritus.

4. EAG SUGGESTED BASE CASE

Taking into account the real-world evidence on discontinuation and the data on disutility associated with severe pruritus, the EAG suggests a base case with a 40% discontinuation rate at one year and a disutility for severe pruritus of -0.15. However, the EAG cautions that the Adelphi disutility analysis was poorly conducted and reported, and contained results that lacked face validity, whilst the ITCH-E study yielded an implausibly low estimate, and this itself lacked face validity. The discontinuation percentage is driven by the IQVIA data. The EAG has not viewed the data, inclusion criteria, indication or setting in which the treatments were prescribed, and NICE / NHS England may wish to monitor discontinuation rates to ensure the estimates are plausible.

5. ADDITIONAL SCENARIO ANALYSES

Table 6 and Table 7 below show the EAG analyses for UDCA tolerant and intolerant patients respectively. The scenarios can be considered in 3 groups:

Scenarios 1-3 reiterate the EAG and Committee preferences at AC2. Scenario 1 shows the base case at ACM2 which includes outcome recalculation for differing definitions of response by ALP levels. Following ACM2 the committee requested the ITC to be unadjusted for this (scenario 2). In their post AC2 submission, the Company offered a further discount off the price, the results of which are shown in scenario 3.

Scenarios 4-6 represent the Company's preferred base case for AC3. Scenarios 4 and 5 are the Company's preferred scenarios for discontinuation rates and pruritus disutility respectively (all equal discontinuation at 15.6% at year 1 and Adelphi DSP disutility set). Scenario 6 combines these into the Company base case.

Scenarios 7-14 represent the EAG's explorations and suggested base case. Scenario 7 includes the coding modification to set the relative rate of discontinuation equal across all three comparators rather than the absolute rate post 12m being equal. Scenarios 8-10 are as per scenario 7 but assume equal discontinuation rates across all comparators of 20%, 30% and 40% respectively. Scenarios 11-13 are as per scenario 7 but assume the disutility for severe pruritus is -0.05, -0.10 and -0.15 respectively. Scenario 14 represents a suggested base case.

Table 6: Post ACM2 scenarios: UCDA tolerant

Scenario	Treatment	Total costs (£)	Total QALYs	Incremental ICER	NHB at £25k per QALY*
1. EAG base case at ACM2 (including outcome recalculation)	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
2. Committee base case at ACM2 excluding outcome recalculation (=EAG base without outcome recalculation)	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
3. Committee base case at ACM2 (=2) with	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]

Scenario	Treatment	Total costs (£)	Total QALYs	Incremental ICER	NHB at £25k per QALY*
additional price reduction offered by Company	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
4. Committee base case at ACM2 (=3) + Company equal discontinuation rates 15.6% in year 1.	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
5. Committee base case at ACM2 (=3) + Company DSP pruritus disutilities.	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
6. Company Base case post ACM2 (=4+5)	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
7. Committee base case (=3) + Correction to discontinuation 12m+	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
8. Committee base case (=7) + 20% year 1 discontinuation	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
9. Committee base case (=7) + 30% year 1 discontinuation	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
10. Committee base case (=7) + 40% year 1 discontinuation	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
11. Committee base case (=7) + 0.05 pruritus disutility	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
12. Committee base case (=7) + 0.10 pruritus disutility	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
13. Committee base case (=7) + 0.15 pruritus disutility	SEL + UDCA	[redacted]	[redacted]	-	[redacted]
	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	SEL + UDCA	[redacted]	[redacted]	-	[redacted]

Scenario	Treatment	Total costs (£)	Total QALYs	Incremental ICER	NHB at £25k per QALY*
14. EAG suggested base case: 0.15 disutility, 40% discontinuation (=10+13)	OCA + UDCA	[redacted]	[redacted]	Dominated	[redacted]
	ELA + UDCA	[redacted]	[redacted]	Dominated	[redacted]

*NHB calculations are valid only under the assumption that £25,000 represents the opportunity cost of generating a QALY in the NHS; empirical estimates vary. Abbreviations: ACM, appraisal committee meeting; ICER, incremental cost effectiveness ratio; NHB, net health benefit; OCA, Obeticholic acid; QALYs, quality adjusted life years; UDCA, Ursodeoxycholic acid.

Table 7: Post ACM2 scenarios: UCDA intolerant

Scenario	Treatment	Total costs (£)	Total QALYs	Incremental ICER	NHB at £25k per QALY*
1. EAG base case at ACM2 (including outcome recalculation)	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
2. Committee base case at ACM2 excluding outcome recalculation (=EAG base without outcome recalculation)	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
3. Committee base case at ACM2 (=2) with additional price reduction offered by Company	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
4. Committee base case at ACM2 (=3) + Company equal discontinuation rates 15.6% in year 1.**	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
5. Committee base case at ACM2 (=3) + Company DSP pruritus disutilities.**	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
6. Company Base case post ACM2 (=4+5)**	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
	Seladelpar	[redacted]	[redacted]	-	[redacted]

Scenario	Treatment	Total costs (£)	Total QALYs	Incremental ICER	NHB at £25k per QALY*
7. Committee base case (=3) + Correction to discontinuation 12m+	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
8. Committee base case (=7) + 20% year 1 discontinuation	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
9. Committee base case (=7) + 30% year 1 discontinuation	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
10. Committee base case (=7) + 40% year 1 discontinuation	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
11. Committee base case (=7) + 0.05 pruritus disutility	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
12. Committee base case (=7) + 0.10 pruritus disutility	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
13. Committee base case (=7) + 0.15 pruritus disutility	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]
14. EAG suggested base case: 0.15 disutility, 40% discontinuation (=10+13)	Seladelpar	[redacted]	[redacted]	-	[redacted]
	OCA	[redacted]	[redacted]	Dominated	[redacted]
	Elafibranor	[redacted]	[redacted]	Dominated	[redacted]

*NHB calculations are valid only under the assumption that £25,000 represents the opportunity cost of generating a QALY in the NHS; empirical estimates vary. ** Company did not provide calculations for its base case and scenarios for the UDCA-intolerant group. The EAG has calculated these (scenarios 4, 5 and 6). Abbreviations: ACM, appraisal committee meeting; ICER, incremental cost effectiveness ratio; NHB, net health benefit; OCA, Obeticholic acid; QALYs, quality adjusted life years; UDCA, Ursodeoxycholic acid.

6. REFERENCES

1. Hirschfield GM, Bowlus CL, Mayo MJ, Kremer AE, Vierling JM, Kowdley KV, et al. A Phase 3 Trial of Seladelpar in Primary Biliary Cholangitis. NEJM. 2024;390(9):783-94.

2. Kowdley KV, Bowlus CL, Levy C, Akarca US, Alvares-da-Silva MR, Andreone P, et al. Efficacy and Safety of Elafibranor in Primary Biliary Cholangitis. NEJM. 2024;390(9):795-805.

3. Nevens F, Andreone P, Mazzella G, Strasser SI, Bowlus C, Invernizzi P, et al. A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis. NEJM. 2016;375(7):631-43.

4. Ronca V, De Vincentis A, Terracciani F, Vespasiani-Gentilucci U, Cristoferi L, Gerussi A, et al. Fibrate-Oca (FI-OCA) Study - a Global snapshot of PBC practice around the Globe. Digestive and Liver Disease. 2025;57:S35-S6.

5. Roberts SB, Ismail M, Kanagalingam G, Mason AL, Swain MG, Vincent C, et al. Real-World Effectiveness of Obeticholic Acid in Patients with Primary Biliary Cholangitis. Hepatol Commun. 2020;4(9):1332-45.

6. Abbas N, Culver EL, Thorburn D, Halliday N, Crothers H, Dyson JK, et al. UK-Wide Multicenter Evaluation of Second-line Therapies in Primary Biliary Cholangitis. Clin Gastroenterol Hepatol. 2023;21(6):1561-70.e13.

7. Jones A, Greener K, Cripps S, Rees F, Shah S, Crook J, et al. P34 Initial clinical response and safety of elafibranor in primary biliary cholangitis: a multi-centre pharmacist-led real-world evaluation. Gut. 2025;74(Suppl 4):A35-A.

8. Gilead. Data on File - Direct and Indirect Mapping of PBC-40 onto EQ-5D-5L Utilities from RESPONSE trial. 2024.

9. Smith H, Fettiplace J, Von Maltzahn R, Das S, McLaughlin MM, Jones D. More Than Just an Itch: Impact of Cholestatic Pruritus in Primary Biliary Cholangitis (PBC) on Health-Related Quality of Life (HRQoL). Am J Gastroenterol. 2022;117(10):S78.

10. Levy C, Kendrick S, Bowlus CL, Tanaka A, Jones D, Kremer AE, et al. GLIMMER: A Randomized Phase 2b Dose-Ranging Trial of Linerixibat in Primary Biliary Cholangitis Patients With Pruritus. Clin Gastroenterol Hepatol. 2023;21(7):1902-12.e13.

Seladelpar for previously treated primary biliary cholangitis [ID6429]

Additional note on EAG suggested base case for ACM3

April, 2026

Produced by	Peninsula Technology Assessment Group (PenTAG) University of Exeter Medical School
Authors	Caroline Farmer[1] Jemma Perks[1] Alan Lovell[1] Maxwell S. Barnish[1] Valdemar Santo[1] Ahmed Abdelsabour[1] Saul Stevens[1] Dawn Lee[1] Taha Khan[2] Rebecca Jones[3] Gwilym Webb[4] Edward CF Wilson[1] [1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter [2] St Leonard's Practice, Exeter [3] Leeds Teaching Hospital NHS Trust [4] Cambridge University Hospitals NHS Foundation Trust
Correspondence to	Caroline Farmer; c.farmer@exeter.ac.uk

Source of funding	This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR175575.
Declared competing interests of the authors	None

This report should be referenced as follows: Farmer, Perks, Lovell, Barnish, Santo, Abdelsabour, Lee, Khan, Jones, Webb, Wilson. Seladelpar for previously treated primary biliary cholangitis [ID6429]: EAG response to additional company submission prior to ACM3. Peninsula Technology Assessment Group (PenTAG), 2026.

1. INTRODUCTION

At the request of the committee chair, the EAG provided some additional clarification around the EAG's suggested base case to inform ACM3 for the appraisal of seladelpar for previously treated biliary cholangitis, specifically the choice of discontinuation rate and disutility for pruritus.

Above all the EAG wishes to emphasise that the EAG base case is *suggested*, not *recommended* or explicitly preferred by the EAG. This is due to the uncertainties in the two relevant parameters.

1.1. Discontinuation rate

Following AC2, the Committee expressed a preference for differing discontinuation rates, whilst the Company preferred equal discontinuation rates across all three comparators. It argued the apparently high discontinuation rate for OCA was due to the unexpectedly low discontinuation in the control arm of the POISE study (see Table 2 of company post ACM2 evidence, reproduced as Table 2 in the EAG's response). This gave a high relative risk of 2.7. The EAG reviewed the baseline characteristics of the control arms for RESPONSE, ELATIVE and POISE (pivotal RCTs for seladelpar, elafibranor and OCA respectively), and could identify no differences that would lead to a different discontinuation rate in control arms between studies. On this basis (comparing discontinuation rates across RCTs) it may be reasonable to conclude equal discontinuation rates across all three comparators.

However, the Company also presented real-world data (IQVIA), which showed very high and substantially different discontinuation rates over a *6-month* period (34% elafibranor, 46% seladelpar). As stated in its report, the EAG did not have sight of these data or the analysis and notes that there are no data for OCA. In particular, the sample sizes are unknown.

Given that discontinuation of medicines in the real world tends to be higher than that observed in trials, the EAG suggested a higher rate of 40% per annum.

The EAG conducted several ad hoc explorations manually varying discontinuation rates in the model. It observed that when all else equal:

- Cost-effectiveness of seladelpar improves the shorter time it is taken (i.e. the higher the discontinuation rate). This is because incremental costs accrue faster than incremental

benefits over time: if discontinuation with seladelpar is higher than the comparators it is more likely to be cost-effective.

- Cost-effectiveness of seladelpar deteriorates when discontinuation rate for OCA is higher than for seladelpar: if discontinuation with OCA is higher than for seladelpar then seladelpar is less likely to be cost-effective. The EAG noted that the IQVIA data did not report discontinuation rates for OCA. These data would have been useful to inform relative discontinuation rates.

1.2. Pruritus disutility

On review of relevant evidence, the EAG considered the ITCH-E study's point estimate of the impact of severe pruritus to be lower than expected: the EAG conducted a rapid search of the Tuft's CEA Registry for utility estimates with the keyword 'pruritus' and identified three economic evaluations using such estimates, of which two[1 2] provided relevant data, although both drew on the same data source:

Thokola et al.[1] estimated the cost-effectiveness of a treatment for pruritus associated with chronic kidney disease (CKD) in people receiving haemodialysis in a UK setting, estimating the utility of an individual with CKD without pruritus at 0.62 and for one with severe pruritus at 0.42, a disutility of 0.2, based on mapping from the 5D itch scale to EQ-5D-5L. Manenti et al.[2] adapted the model to an Italian setting using the same sources, although reported the utility of severe pruritus at 0.37 (disutility of 0.25).

If the impact of CKD-driven pruritus is considered of similar magnitude to that arising from primary biliary cholangitis, then the EAG's estimate of 0.15 may be considered conservative.

2. BIBLIOGRAPHY

1. Thokala P, Hnynn Si PE, Hernandez Alava M, et al. Cost Effectiveness of Difelikefalin Compared to Standard Care for Treating Chronic Kidney Disease Associated Pruritus (CKD-aP) in People with Kidney Failure Receiving Haemodialysis. *PharmacoEconomics* 2023;41(4):457-66. doi: 10.1007/s40273-022-01237-4
2. Manenti L, Marcellusi A, Di Brino E, et al. Cost effectiveness of difelikefalin for the treatment of patients with chronic kidney disease-associated pruritus undergoing hemodialysis in Italy. *J Nephrol* 2025;38(1):251-59. doi: 10.1007/s40620-024-02144-x [published Online First: 20241108]