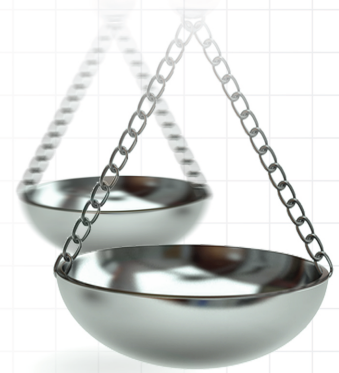




Network meta-analysis: relative clinical efficacy and safety of elafibranor versus seladelpar as second-line treatment for patients with primary biliary cholangitis

Journal of **Comparative Effectiveness Research**

David Jones¹ , Emily Combe², Harun Knight², Vicki Laskier-Owens^{*,2} , Shijie Ren³ ,

Tom Wright², Elaine A Böing⁴ & Alex Pashley⁵ 

¹Institute of Cellular Medicine & NIHR Newcastle Biomedical Research Centre, Newcastle University, Newcastle Upon Tyne, UK

²FIECON, London, UK

³Sheffield Centre for Health & Related Research, University of Sheffield, Sheffield, UK

⁴Ipsen, Cambridge, MA 02142, USA

⁵Costello Medical, Cambridge, UK

*Author for correspondence: vicki.laskier-owens@acumetisglobal.com

Aim: To indirectly compare the efficacy and safety of elafibranor and seladelpar, as second-line treatments for primary biliary cholangitis. **Materials & methods:** Bayesian network meta-analyses compared data from randomized-controlled studies of elafibranor and seladelpar identified by a systematic literature review up to June 2024: (a) elafibranor (n = 108) versus placebo (n = 53; ELATIVE [NCT04526665]) and (b) seladelpar (n = 128) versus placebo (n = 65; RESPONSE [NCT03301506]). Patients from ELATIVE not meeting the RESPONSE upper limit of normal (ULN) criteria for alkaline phosphatase (ALP) and total bilirubin were excluded (n = 16); summary statistics for ELATIVE were recalculated using the new dataset. Random-effects models assessed the outcomes of cholestasis response (ALP <1.67 × ULN, ALP reduction ≥15% from baseline and total bilirubin ≤ULN), ALP normalization, change from baseline in ALP and pruritus, pruritus as a treatment-emergent adverse event and all-cause discontinuation. **Results:** Elafibranor-treated patients had greater odds of achieving cholestasis response than placebo- (median odds ratio [95% credible interval]: 84.79 [12.49, 2513.00]) or seladelpar-treated patients (13.02 [1.45, 420.20]), with posterior probabilities ≥99% that odds were higher with elafibranor than seladelpar or placebo. Among patients with ALP ≥350 U/l, the median odds ratio [95% credible interval] of cholestasis response for elafibranor-treated patients versus seladelpar-treated patients was 18.71 [0.65, 10,610.00], with a 95.2% posterior probability that odds were higher with elafibranor than seladelpar. For all other outcomes, there was no strong evidence of a difference between treatments. **Conclusion:** Bayesian network meta-analyses found strong probabilistic evidence supporting the treatment benefit of elafibranor compared with seladelpar for the achievement of cholestasis response at 52 weeks, while the treatment effect on other outcomes was uncertain. Head-to-head studies are needed to validate results of these indirect comparisons.

Plain language summary: Comparison of elafibranor & seladelpar treatment in patients with primary biliary cholangitis following previous treatment with ursodeoxycholic acid

What is this article about? Primary biliary cholangitis (PBC) is a chronic disease that gradually damages the liver over time. This damage can lead to a build-up of symptoms and life-threatening complications, such as scarring (cirrhosis) and failure of the liver. Up to 40% of people living with PBC incompletely respond to first-line treatment with ursodeoxycholic acid. Recently, two new second-line treatments for PBC were launched, called elafibranor and seladelpar, which have similar ways of working. At present, there are no head-to-head trials of elafibranor and seladelpar, and we are not aware of any planned trials. Thus, this study aimed to compare the effectiveness of the two second-line treatments using a technique called an 'indirect treatment comparison'. The objective is to help doctors and patients decide which of these treatments might be more suitable for different people. The approach compared the likelihood of clinical events from data provided across individual studies of elafibranor and seladelpar.

What were the results? We find that patients treated with elafibranor had a significantly higher chance of improving liver blood tests (specifically, cholestasis response) compared with those on seladelpar. However, for other outcomes, there was no significant difference between the treatments.

Why is this important? These findings could support treatment decisions for patients with PBC who respond poorly to ursodeoxycholic acid.

First draft submitted: 16 December 2025; Accepted for publication: 26 March 2026; Published online: 17 April 2026

Keywords: autoimmune liver disease • liver • meta-analyses • outcomes research • primary biliary cholangitis

Patients with primary biliary cholangitis (PBC) experience a substantial clinical burden, with disease progression leading to accumulation of symptoms, comorbidities and life-threatening liver-related complications [1–3]. While most patients are asymptomatic at presentation [4,5], up to 70% and 80% of patients develop itching (pruritus) and fatigue, respectively [5–7], significantly impacting health-related quality-of-life. As PBC advances, the risk of complications rise, consequently necessitating transplantation in end-stage liver disease [8]. Indications for liver transplantation vary between countries, but broadly include hepatocellular carcinoma, persistent jaundice, hepatic decompensation and/or pharmacotherapy-refractory pruritus [9].

PBC is more therapy-responsive in its early stages [8]. Notably, alkaline phosphatase (ALP) and total bilirubin (TB) values under ursodeoxycholic acid (UDCA) treatment are predictive of liver-related events, and widely accepted biochemical surrogates of long-term outcomes [10,11]. This has been extrapolated in the development of second-line therapy, wherein biochemical cholestasis response (defined as an ALP $<1.67 \times$ upper limit of normal [ULN], $TB \leq ULN$ and ALP decrease from baseline of $\geq 15\%$) has been adopted as a primary efficacy outcome measure in clinical trials [12–14].

Multicenter and population-based data indicate that $\sim 40\%$ of patients inadequately respond to UDCA, with 3–5% of patients being UDCA-intolerant [10,15,16], highlighting a need for effective second-line treatments. Elafibranor and seladelpar have recently been approved by the US FDA, EMA and Medicines and Healthcare products Regulatory Agency in the UK [17–22], providing new second-line treatment options for patients who inadequately respond to, or who are UDCA-intolerant.

Elafibranor, a peroxisome proliferator-activated receptor (PPAR) agonist exerting effects on both α and δ isoforms, has demonstrated efficacy in multicenter randomized controlled trials (RCTs) [13,23]. Notably, significantly more elafibranor-treated patients achieved cholestasis response at 52 weeks than placebo-treated patients in the Phase III ELATIVE RCT [13]. Significant differences between elafibranor- and placebo-treated patients were also observed for ALP normalization. Similarly, seladelpar, a selective PPAR- δ agonist, was evaluated in the Phase III RESPONSE RCT, wherein significantly more seladelpar-treated patients achieved cholestasis response at 52 weeks than placebo-treated patients [14]. ALP normalization was also achieved by significantly more seladelpar-treated patients than placebo-treated patients. However, comparing efficacy and safety across these trials is challenging due to different definitions of ULN for ALP and TB in trial eligibility criteria and response outcomes.

The lack of head-to-head RCTs comparing elafibranor and seladelpar poses challenges for decision-makers in evaluating their relative efficacy and safety. Consequently, using indirect treatment comparison (ITC) methods and existing clinical trial data, this study compares the relative efficacy and safety of elafibranor and seladelpar in adult patients with PBC who have an inadequate response or intolerance to UDCA.

Materials & methods

Study selection

A systematic literature review (SLR) was conducted to identify evidence describing the clinical efficacy and/or safety of investigational treatments for adult patients with PBC. The SLR was conducted in accordance with the methodological principles to conduct SLRs recommended by the Centre for Reviews and Dissemination (CRD), and with a prespecified protocol registered on PROSPERO (ID: CRD42023382262) [24].

The eligibility criteria were developed using the Population, Intervention, Comparator, Outcomes, Study design (PICOS) framework (Supplementary Table 1). Eligible studies investigated unselected adult patients with PBC who received monotherapy or combination therapy in any treatment line including, but not limited to elafibranor, UDCA, obeticholic acid, fibrates and seladelpar. Clinical studies that reported clinical effectiveness, pruritus, or

safety outcomes were eligible for inclusion. Comprehensive search strategies were defined according to the PICOS framework, combining subject headings and test words for PBC, plus terms for interventional and observational studies based on validated search filters ([Supplementary Tables 2 & 3](#)). Searches were conducted within Medline, EMBASE, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews and the Database of Abstracts of Reviews of Effects from their inception to identify peer-reviewed publications in English. Database searches were conducted up to 19 June 2024. Conference proceedings for the American Association for the Study of Liver Diseases Liver Meeting, European Association for the Study of the Liver International Liver Congress, European Association for the Study of the Liver Digital Liver Cancer Summit, International Conference on Hepatology and Liver Disease and World Conference on Gastroenterology and Hepatology from 2021 to 2024 were last searched on 22 July 2024 ([Supplementary Tables 4–6](#)). Clinical trials registry searches were also last performed on 12 July 2024 in ClinicalTrials.gov ([Supplementary Table 7](#)). Further details on the dates of searches for the SLR are included in [Section S1](#) of the Supplementary materials. The bibliographies of relevant SLRs and network meta-analyses (NMAs) identified during the SLR were also hand-searched to identify any additional, relevant studies for inclusion.

Studies were screened by two independent reviewers (AP, SS, EW; see Acknowledgements) for inclusion against the predefined eligibility criteria in two stages: first, the titles and abstracts of the search results; second, the full texts of potentially relevant articles, in order to obtain the final list of included studies. Disagreements were adjudicated by a third reviewer to arbitrate any discrepancies.

Due to the high volume of evidence identified in the SLR, a two-stage prioritization strategy was implemented. Interventional and observational studies in the second-line or later settings were prioritized following the title/abstract review phase. Subsequently, any studies of elafibranor and RCTs in the second-line or later setting were prioritized for extraction. Data extraction were performed by a single reviewer and verified independently by a second reviewer (AP, SS, EW; see Acknowledgements).

Feasibility assessment

As the objective of the ITC was to compare elafibranor and seladelpar for second-line treatment of PBC, only studies including elafibranor and seladelpar were retained. Eligibility criteria for the ITC, developed using the PICOS framework, are reported in [Supplementary Table 8](#) ([Supplementary Section 2](#)). Eligible studies were RCTs reporting on outcomes after at least 52 weeks of treatment as earlier outcomes were considered incomparable to the ELATIVE outcomes. Outcomes of interest were clinical efficacy, pruritus and safety outcomes, with the clinical efficacy outcomes including, but not limited to, survival, response rates or biochemical response. Pruritus outcomes included, but were not limited to, experience of pruritus, pruritus-specific health-related quality-of-life or ItchyQoL. Safety outcomes included, but were not limited to, adverse events, deaths or discontinuation due to adverse events.

The comparability of studies was assessed by identifying heterogeneity in study design, population, outcomes and treatments considered. Two clinical experts advised that age at diagnosis, ALP levels, TB, cirrhosis and antinuclear antibody positive status at baseline were treatment effect modifiers, and validated that they were comparable across the included studies for patients with PBC; the feasibility assessment was modified based on their insights.

Risk-of-bias assessment

Risk-of-bias (RoB) in the studies included in the network meta-analysis (NMA) was assessed using Version-2 of the Cochrane RoB tool for randomized trials (RoB 2) [25].

Statistical analysis

Despite both studies specifying $ALP \geq 1.67 \times ULN$ and $TB \leq 2 \times ULN$ as key eligibility criteria, the clinical heterogeneity assessment found that definitions of ULN for ALP and TB varied. The ALP ULN was 129 U/l and 104 U/l for males and females in ELATIVE, respectively, compared with 116 U/l in RESPONSE. The TB ULN was 1.2 mg/dl and 1.1 mg/dl in ELATIVE and RESPONSE, respectively. Individual patient data (IPD) was available for ELATIVE, whereas only published aggregate data was available for RESPONSE. To improve similarity of baseline characteristics across respective trial populations and end points, the ELATIVE population was truncated to align with the RESPONSE inclusion criteria:

- $ALP \geq 1.67 \times ULN$ (defined as 1.67×116 U/l)

- $TB \leq 2 \times ULN$ (defined as $2 \times 1.1 \text{ mg/dl}$)

Prior to performing the ITC, all outcomes in ELATIVE were re-estimated in the truncated ELATIVE population ($n = 145$) applying the definitions of ULN for ALP and TB from RESPONSE to the outcome definitions. Further methodological details are reported in [Section S3](#) of the Supplementary materials.

An anchored ITC was conducted via a Bayesian NMA using aggregate data from the truncated ELATIVE and RESPONSE populations in accordance with methodological framework for Pairwise and Network Meta-Analysis of Randomized Controlled Trials reported in the NICE Decision Support Unit (DSU) Technical Support Document 2 [26]. The network consisted of elafibranor, seladelpar and placebo, with placebo serving as the common comparator from the ELATIVE and RESPONSE studies. As clinical validation confirmed key treatment effect modifiers at baseline were sufficiently similar in the aggregated data, it was unnecessary to pursue techniques such as simulated or matching-adjusted treatment comparisons. All analyses were performed using outcomes in the intention-to-treat (ITT) populations after 52 weeks.

Aligning with ELATIVE and RESPONSE, the primary analysis was the likelihood of cholestasis response. An additional analysis of cholestasis response in patients with $ALP \geq 350 \text{ U/l}$ at baseline was performed, aligning with the predefined subpopulation from RESPONSE. Further binary end points assessed included likelihood of ALP normalization, pruritus as a treatment-emergent adverse event (TEAE), and all-cause discontinuation. The continuous end points assessed included the mean change from baseline (CFB) in ALP and mean CFB in pruritus according to 5-D Itch, the PBC-40 itch domain and a numerical rating scale (NRS; pruritus NRS and worst-itch NRS for RESPONSE and ELATIVE, respectively). The latter three analyses were also performed in the pruritus ITT subpopulation (patients in the ITT population with a NRS ≥ 4 at baseline). For binary outcomes, odds ratios (ORs) were estimated, while for continuous outcomes, the difference in least-square mean in CFB were estimated; median outcomes with 95% credible intervals (CrIs) were reported in line with typical Bayesian methodology. Probabilities, derived from the full posterior distribution for each Bayesian model (referred to as ‘posterior probabilities’), were generated for all outcomes to quantify the likelihood that elafibranor treatment was more effective than placebo or seladelpar. To quantify statistical heterogeneity between the studies, the between-study standard deviation (SD) was calculated for random-effects models. Both random-effects (base case) and fixed-effect models (sensitivity analyses) were assessed. Informative priors were used for between-study SD and vague priors were used for treatment effects and baseline study effects.

OpenBUGS version 3.2.3 and R version 4.3.1 via RStudio were used to perform the NMA [27–29]. Additional methodological details are reported in [Section S4](#) of the Supplementary materials.

Results

Systematic literature review

[Figure 1](#) presents the results of each stage of the screening process in the SLR. Searches of electronic databases retrieved 9,888 records, of which 3,772 were duplicates, resulting in 6,116 novel records that were screened at the title/abstract review stage. Following full-text screening, 125 publications reporting on 43 unique studies were included for extraction in the SLR following prioritization of RCTs in the second-line or later setting.

Feasibility assessment

Of the 43 unique studies extracted, there were two elafibranor and five seladelpar studies. Five studies were excluded: three according to the NMA eligibility criteria, one could not provide a feedback loop in subsequent ITCs and one was a pooled analysis of already excluded studies ([Supplementary Section 5](#)). A summary of all studies excluded from the NMA is reported in [Supplementary Table 9](#) ([Supplementary Section 5](#)).

The network of evidence comprised ELATIVE (elafibranor; NCT04526665) and RESPONSE (seladelpar; NCT03301506) with the common comparator of placebo ([Figure 2](#)). Study design, population, outcomes and treatments were found to be sufficiently homogenous across both trials ([Table 1](#)). However, it was found that the inclusion and exclusion criteria of ELATIVE and RESPONSE were not aligned due to differences in the ULN definitions in ALP and TB. After the alignment of ULNs, the feasibility assessment concluded that comparisons between ELATIVE and RESPONSE were feasible and a NMA was appropriate.

RoB assessment

The assessment found a low RoB across the two trials ([Table 1](#)).

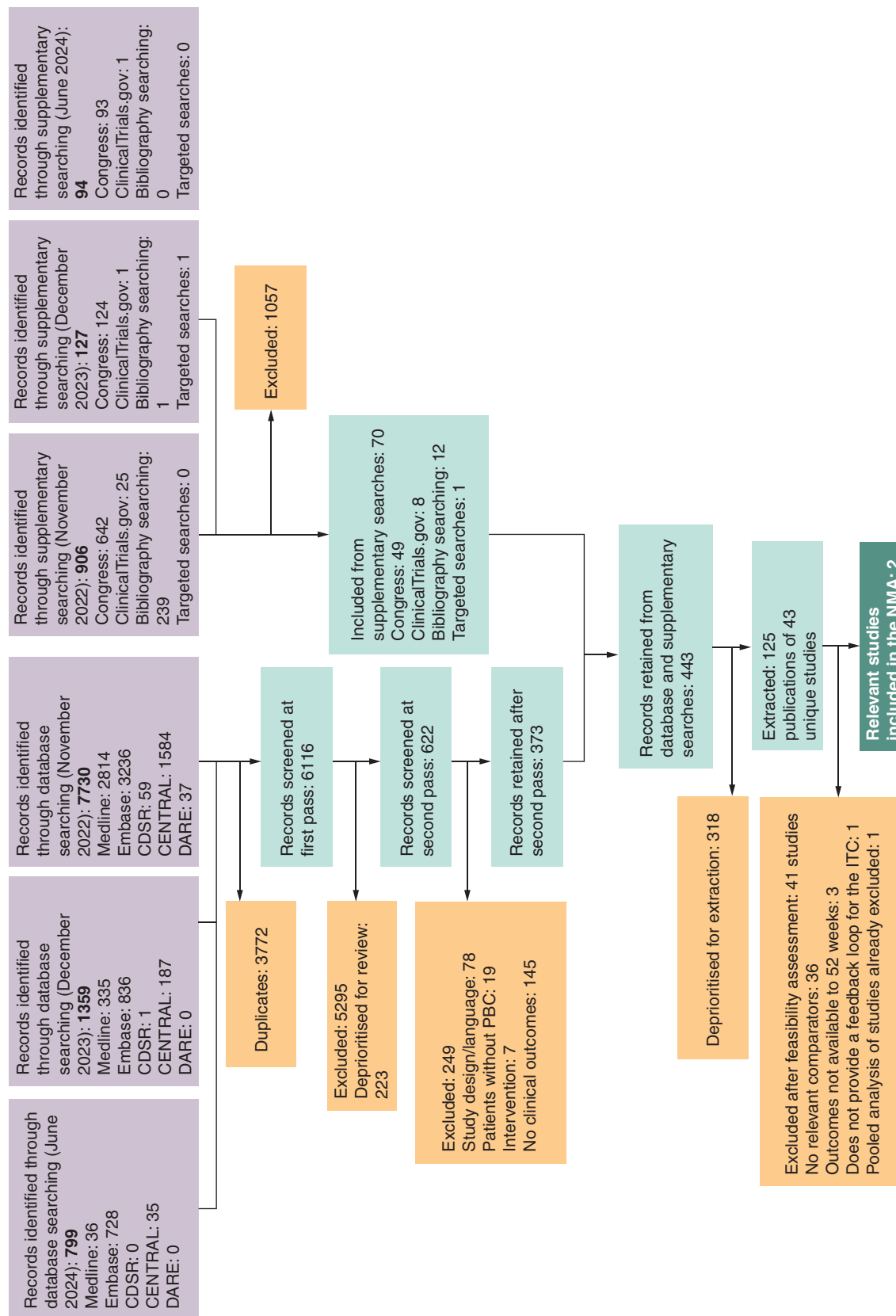


Figure 1. PRISMA diagram of studies identified in the systematic literature review.
 CDSR: Cochrane Database of Systematic Reviews; CENTRAL: Cochrane Central Register of Controlled Trials; DARE: Database of Abstracts of Reviews of Effects; NMA: Network meta-analysis; PBC: Primary biliary cholangitis.

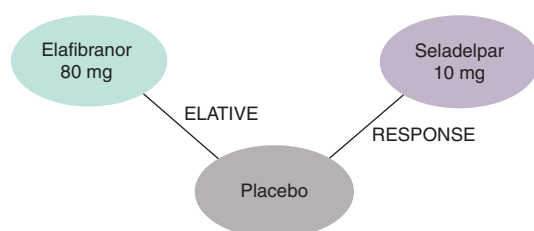


Figure 2. Network of evidence for the network meta-analysis.

NMA: Network meta-analysis.

Network meta-analysis results

Following IPD analysis of ELATIVE to harmonize the ELATIVE and RESPONSE patient populations in terms of trial eligibility criteria, 16 patients were excluded as their baseline ALP did not meet the RESPONSE eligibility criteria. The results of the IPD analysis of ELATIVE and the data used in the ITC analyses are reported in [Sections S6 & S7](#) of the Supplementary materials, respectively.

Using placebo as the common comparator in the random-effects NMA, there were greater odds of achieving cholestasis response at 52 weeks in the ITT population for elafibranor-treated patients than placebo- (median OR [95% CrI]: 84.79 [12.49, 2513.00]) or seladelpar-treated patients (13.02 [1.45, 420.20]; [Figure 3](#)). The posterior probabilities that the odds were higher with elafibranor than placebo or seladelpar were 100.0% and 99.1%, respectively ([Table 2](#)).

Consistent with the ITT population, there were greater odds of achieving cholestasis response at 52 weeks in the ALP ≥ 350 U/l subpopulation for elafibranor-treated patients than placebo- (33.52 [2.17, 17,040.00]) or seladelpar-treated patients (18.71 [0.65, 10,610.00]; [Figure 3](#)). The posterior probabilities that the odds were higher with elafibranor than placebo or seladelpar were 99.7% and 95.2%, respectively ([Table 2](#)).

The median OR (95% CrI) of ALP normalization at 52 weeks for elafibranor-treated patients compared with placebo- or seladelpar-treated patients were 40.54 (2.89, 18,880.25) and 0.42 (0.00, 282.30), respectively ([Figure 4](#)). There were 99.9% and 33.8% posterior probabilities that the odds were higher with elafibranor than placebo or seladelpar, respectively ([Table 2](#)). Likewise, the median difference in least-square mean CFB (95% CrI) in ALP at 52 weeks for elafibranor-treated patients compared with placebo- or seladelpar-treated patients were -115.20 (-151.10, -77.34) and 1.87 (-48.59, 53.77), respectively ([Figure 5](#)). There were 100.0% and 47.2% posterior probabilities that the reduction was greater for elafibranor than placebo or seladelpar, respectively ([Table 2](#)).

The median relative treatment effects and 95% CrIs for pruritus as a TEAE, all-cause discontinuation or CFB in pruritus when measured using a NRS, 5-D Itch or the itch domain of the PBC-40 questionnaire at 52 weeks indicated substantial uncertainty in elafibranor's relative effectiveness versus seladelpar ([Figures 6–8](#)). The posterior probabilities that elafibranor-treated patients had preferred outcomes to seladelpar-treated patients in these outcomes ranged from 23.6 to 64.9% ([Table 2](#)), supporting the uncertainty in relative treatment effects. Relative effect matrices are reported in [Section S9](#) of the Supplementary materials.

The models were a good fit to the data, as demonstrated by the total residual deviance remaining below the number of data points in all but three analyses; total residual deviance was marginally greater than the number of data points in the analyses of pruritus as a TEAE, CFB in pruritus measured using 5-D Itch in the pruritus ITT subgroup and cholestasis response in the ALP ≥ 350 U/l subpopulation ([Table 2](#) & [Supplementary Section 8](#)). Using between-study SD to assess heterogeneity, high levels of heterogeneity were not identified, while moderate heterogeneity was identified for cholestasis response and ALP normalization. There were low levels of heterogeneity identified for the remaining binary outcomes. Low levels of heterogeneity were observed for the continuous outcomes when assessing the between-study SD on the standardized mean difference scale. The burn-in, thinning interval and number of iterations for each analysis are reported in [Supplementary Table 16](#) ([Supplementary Section 10](#)).

Results from the fixed-effect models ([Supplementary Section S11](#)) were consistent with results from the base case random-effects models. As expected, the fixed-effect models had narrower CrIs for the relative treatment effects. The random-effects models were retained as the base-case because the difference in total residual deviance did not suggest a better fit to the data with the fixed-effect model.

Table 1. Summary of primary publications of the studies included in the network meta-analysis.

ELATIVE (NCT04526665)		RESPONSE (NCT03301506)	
Summary of key study characteristics			
Key eligibility criteria	• Aged 18–75 • Diagnosis of PBC with inadequate response to or unacceptable side effects with UDCA • Taking UDCA for at least 12 months or unable to tolerate UDCA • ALP $\geq 1.67 \times \text{ULN}$ (ULN = 104 U/l for females and 129 U/l for males) • TB $\leq 2 \times \text{ULN}$ (ULN = 1.2 mg/dl) • ALP $\geq 1.67 \times \text{ULN}$ (ULN = 116 U/l) • TB $< 2 \times \text{ULN}$ (ULN = 1.1 mg/dl)		
Interventions	• Elafibranor 80 mg once daily (N = 108) • Placebo once daily (N = 53)	• Seladelpar 10 mg once daily (N = 128) • Placebo once daily (N = 65)	
Concomitant UDCA usage	• 95.0%	• 93.8%	
Baseline characteristics, mean (SD)	ITT population		
	Before truncation	After truncation	
Age at diagnosis (years)	49 (8.2)	48.4 (8.3)	49.2 (14.7)
ALP (U/l)	322 (150.9)	338.0 (150.5)	314.3 (170.2)
TB (mg/dl)	0.56 (0.3)	0.57 (0.3)	0.76 (0.4)
Liver stiffness (kPa)	10.1 (8.2)	10.3 (8.3)	9.4 (7.5)
Antinuclear antibody status	NR	NR	NR
	Pruritus ITT population		
Age at diagnosis (years)	47.1 (8.3)	46.9 (8.4)	NR
ALP (U/l)	335.1 (192.1)	357.1 (195.0)	NR
TB (mg/dl)	0.61 (0.3)	0.63 (0.3)	NR
Liver stiffness (kPa)	9.9 (6.8)	10.1 (6.5)	NR
Antinuclear antibody status	NR	NR	NR
Outcomes analyzed	• Cholestasis response • ALP normalization • Change in pruritus across various end points (NRS, 5-D Itch and PBC-40 Itch) • Frequency of adverse events, including pruritus • Patient disposition • Change in non-invasive markers of fibrosis		
Study type	Phase III, double-blind, placebo-controlled, multicenter, randomized study		
Risk-of-bias assessment: Cochrane's risk-of-bias tool for randomized trials 2.1			
1.Randomization			
1.1. Was the allocation sequence random?	Yes	Yes	
1.2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Yes	Yes	
1.3. Did baseline difference between intervention groups suggest a problem with the randomization process?	No	No	
2. Deviations from intended interventions?			
2.1. Were participants aware of their assigned intervention during the trial?	No	No	
2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No	No	
2.3. Were there deviations from the intended intervention that arose because of the trial context? [†]	N/A	N/A	
The subdomains of the RoB tool do not need completing based on the following conditions.			
[†] Where the response to 2.2 is no.			
[‡] Where the response to 2.6 is yes.			
[§] Where the response to 3.1 is no.			
[¶] Where the response to 4.3 is no.			
ALP: Alkaline phosphatase; ITT: Intention-to-treat; N/A: Not applicable; NRS: Numerical rating scale; PBC: Primary biliary cholangitis; TB: Total bilirubin; UDCA: Ursodeoxycholic acid; ULN: Upper limit of normal.			
Data taken from [25].			

Table 1. Summary of primary publications of the studies included in the network meta-analysis (cont.).

	ELATIVE (NCT04526665)	RESPONSE (NCT03301506)
2.4. Were these deviations likely to have affected the outcome? [†]	N/A	N/A
2.5. Were these deviations from intended deviation balanced between groups? [†]	N/A	N/A
2.6. Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes	Yes
2.7. Was there potential for a substantial impact (on the result) of the failure to analyze participants in the group to which they were randomized? [‡]	N/A	N/A
3. Missing outcome data		
3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Yes	Yes
3.2. Is there evidence that the result was not biased by missing outcome data? [§]	N/A	N/A
3.3. Could missingness in the outcome depend on its true value? [§]	N/A	N/A
3.4. Is it likely that missingness in the outcome depended on its true value? [§]	N/A	N/A
4. Measurement of the outcome		
4.1. Was the method of measuring the outcome inappropriate?	No	No
4.2. Could measurement or ascertainment of the outcome have differed between intervention groups?	No	No
4.3. Were outcome assessors aware of the intervention received by study participants?	No	No
4.4. Could assessment of the outcome have been influenced by knowledge of intervention received? [¶]	N/A	N/A
4.5. Is it likely that assessment of the outcome was influenced by knowledge of intervention received? [¶]	N/A	N/A
5. Selection of the reported result		
5.1. Were the data that produced this result analyzed in accordance with a prespecified analysis plan that was finalized before unblinded outcome data were available for analysis?	Yes	Yes
5.2. Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible outcome measurements (e.g., scales, definitions, time points) within the outcome domain?	No	No
5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible analyses of the data?	No	No
Overall risk of bias	Low	Low
The subdomains of the RoB tool do not need completing based on the following conditions. [†]Where the response to 2.2 is no. [‡]Where the response to 2.6 is yes. [§]Where the response to 3.1 is no. [¶]Where the response to 4.3 is no. ALP: Alkaline phosphatase; ITT: Intention-to-treat; N/A: Not applicable; NRS: Numerical rating scale; PBC: Primary biliary cholangitis; TB: Total bilirubin; UDCA: Ursodeoxycholic acid; ULN: Upper limit of normal. Data taken from [25].		

Discussion

Despite wide CrIs suggesting uncertainty in the magnitude of the relative effects of elafibranor compared with seladelpar, there were 99.1% and 95.2% probabilities that elafibranor was more effective in achieving cholestasis response at week 52 in the ITT and ALP ≥ 350 U/l at baseline populations, respectively. This provides strong evidence for elafibranor treatment benefit compared with seladelpar for this outcome. For all other outcomes including ALP normalization, reduction in ALP from baseline, all-cause discontinuation, reduction in pruritus from baseline in pruritus, and pruritus as an adverse event, probabilities that outcomes were more favorable with

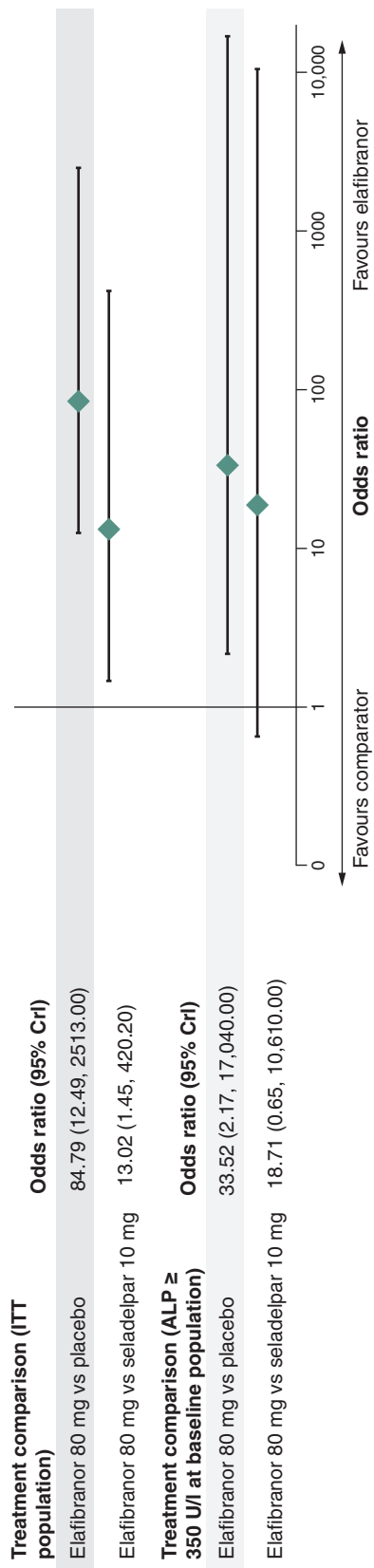


Figure 3. Odds ratio of achieving cholestasis response at 52 weeks (random-effects model). ALP: Alkaline phosphatase; CFI: Change from baseline; CrI: Credible interval; ITT: Intention-to-treat; mg: Milligram; U/l: Units per liter.

Table 2. Summary statistics from the base-case network meta-analysis[†].

Analysis	Between-study SD on mean difference or OR scale [‡]	Total residual deviance	Posterior probability of elafibranor being preferred to placebo	Posterior probability of elafibranor being preferred to seladelpar
Cholestasis response (ITT population)	0.301	3.495	1.000	0.991
Cholestasis response (ALP $\geq$ 350 U/l at baseline population)	0.312	4.371	0.997	0.952
ALP normalization	0.301	3.851	0.999	0.380
Change from baseline in ALP	11.810	3.333	1.000	0.472
Pruritus as a TEAE	0.305	4.062	0.816	0.649
All-cause discontinuation	0.194	3.430	0.586	0.368
Change from baseline in a pruritus NRS (ITT)	0.205	3.365	0.850	0.236
Change from baseline in a pruritus NRS (Pruritus ITT)	0.248	3.356	0.950	0.333
Change from baseline in 5-D Itch (ITT)	0.366	3.344	0.969	0.233
Change from baseline in 5-D Itch (Pruritus ITT)	0.491	4.013	0.995	0.448
Change from baseline in PBC-40 Itch (ITT)	0.272	3.347	0.980	0.642
Change from baseline in PBC-40 Itch (Pruritus ITT)	0.313	3.358	0.992	0.640

[†] Between-study SD on mean difference scale was used for continuous outcomes, while the OR scale was used for binary outcomes.

[‡] The base-case model used a random-effects model.

ALP: Alkaline phosphatase; CFB: Change from baseline; ITT: Intention-to-treat; LSM: Least-square mean; NMA: Network meta-analysis; NRS: Numerical rating scale; OR: Odds ratio; PBC: Primary biliary cholangitis; SD: Standard deviation; TEAE: Treatment-emergent adverse event; U/l: Units per liter.

elafibranor ranged between 23.6% and 64.9%, providing no strong evidence of different treatment effects between elafibranor and seladelpar.

Trial comparability was assessed by examining heterogeneity in study design, population, outcomes and treatments administered. Both ELATIVE and RESPONSE had high-quality trial design, with double-blind, placebo-controlled, multicenter, multinational trials assessed as low RoB. Although ULN definitions for ALP and TB differed, access to IPD allowed exclusion of ELATIVE patients not meeting the stricter RESPONSE eligibility criteria and re-estimation of outcomes using RESPONSE thresholds. This approach taken avoided the limitations of simulated or matching-adjusted treatment comparisons where differences in outcome definitions cannot be accounted for and facilitated a robust NMA using aggregated data without need for an IPD meta-analysis.

In aligning to RESPONSE, patients with the lowest baseline ALP were excluded from ELATIVE. Therefore, the cohort considered was less likely to achieve cholestasis response than the ELATIVE ITT cohort, as demonstrated by a monotonically decreasing likelihood of cholestasis response with increasing baseline ALP (≤ 2 , $2 < \leq 2.5$, $2.5 < \leq 3$, $3 < \leq 4$ and $> 4 \times \text{ULN}$) for elafibranor- or placebo-treated patients in subgroup analyses of ELATIVE [30]. It is also not reported which central laboratory or which assays were used in RESPONSE to determine biological markers. Therefore, it is unclear whether there are differences leading to the estimation of response in both studies. These differences underscore the challenges in comparing outcomes across trials with varying inclusion criteria and response thresholds. Indeed, between-study SD demonstrated moderate heterogeneity for cholestasis response and ALP normalization, suggesting potential underlying differences which cannot be accounted for.

For TB, no patients were excluded from ELATIVE for not meeting the RESPONSE TB eligibility criteria. Given this, mean baseline TB remained lower in ELATIVE compared with RESPONSE. However, as the mean TB in both trials was below the ULN, no additional adjustment was performed. This decision was guided by clinical input and supported by major prognostic models, which demonstrate risk stratification primarily when TB exceeds the ULN [10].

The feasibility assessment examined outcome definitions in ELATIVE and RESPONSE as part of the heterogeneity assessment and found that outcomes were similarly defined, supporting the validity of comparing these end points across trials. In addition, because ALP and TB were identified as treatment effect modifiers their similarity at baseline was confirmed. This reduces the risk of bias when comparing response-based end points, as patients had comparable potential for improvement on these biomarkers. Conversely, pruritus, which was assessed using the patient reported outcomes of a NRS, 5-D Itch and the itch domain of the PBC-40 questionnaire, was not identified as a treatment effect modifier. As a result, differences between studies in baseline levels could influence the achievable amount of improvement that may be observed. Despite this, there was a comparable proportion of patients with a NRS ≥ 4 at baseline in the two trials (41% in ELATIVE compared with 36% in RESPONSE). However, as a

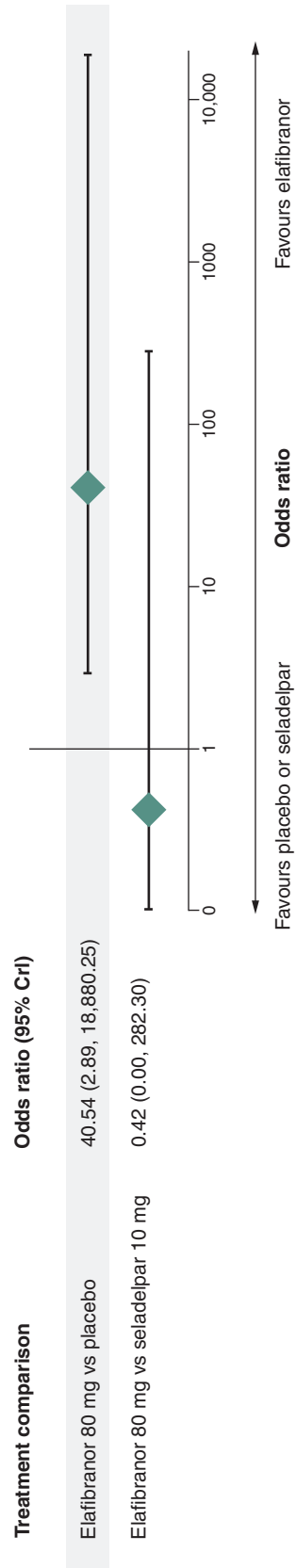


Figure 4. Odds ratio of achieving alkaline phosphatase normalization at 52 weeks in the intention-to-treat population (random-effects model).
ALP: Alkaline phosphatase; CrI: Credible interval; ITT: Intention-to-treat; mg: Milligram.

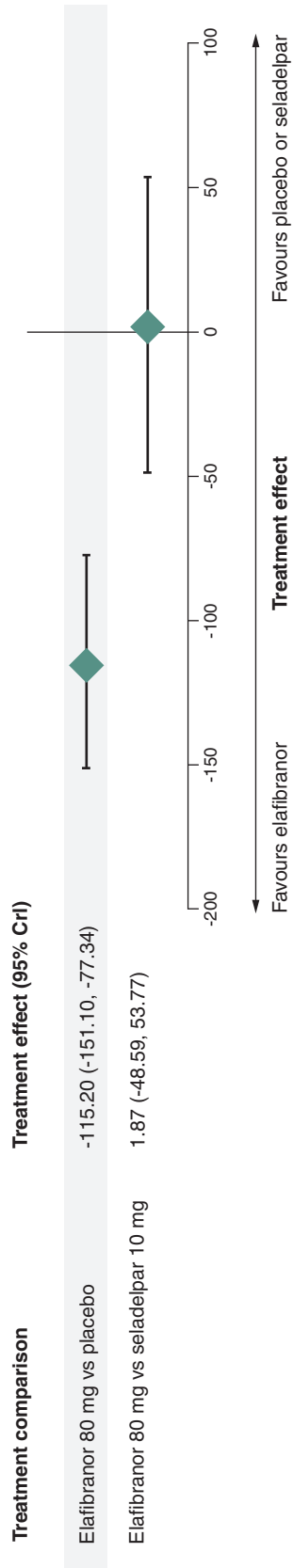


Figure 5. Median difference in least-square mean change from baseline in alkaline phosphatase at 52 weeks in the intention-to-treat population (random-effects model). ALP: Alkaline phosphatase; CFB: Change from baseline; CrI: Credible interval; ITT: Intention-to-treat; LSM: Least-square mean; mg: Milligram.

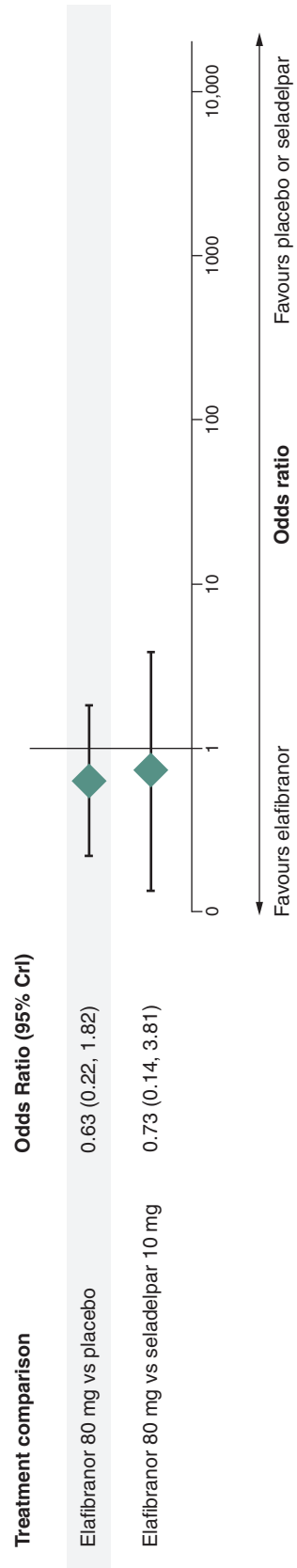


Figure 6. Odds ratio of occurrence of pruritus as a treatment-emergent adverse event at 52 weeks in the intention-to-treat population (random-effects model).
CrI: Credible interval; ITT: Intention-to-treat; mg: Milligram; TEAE: Treatment-emergent adverse event.

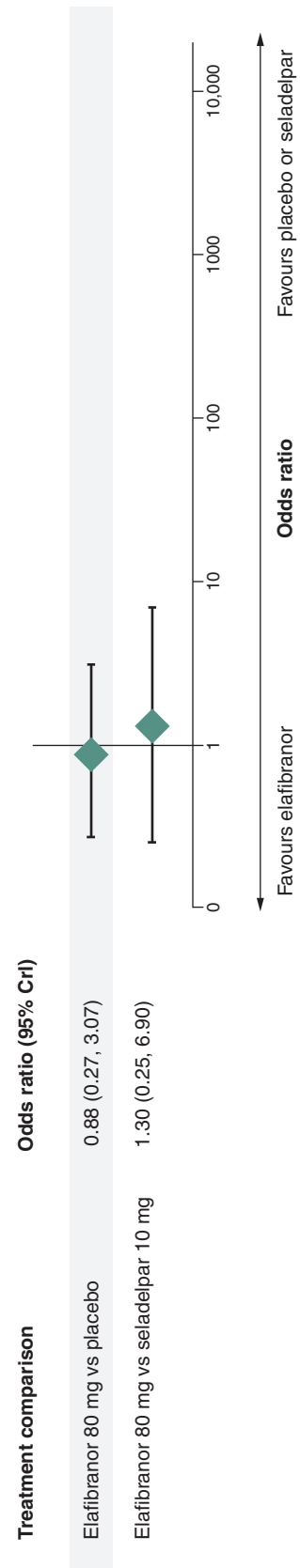


Figure 7. Odds ratio of all-cause discontinuation at 52 weeks in the intention-to-treat population (random-effects model). CrI: Credible interval; ITT: Intention-to-treat; mg: Milligram.

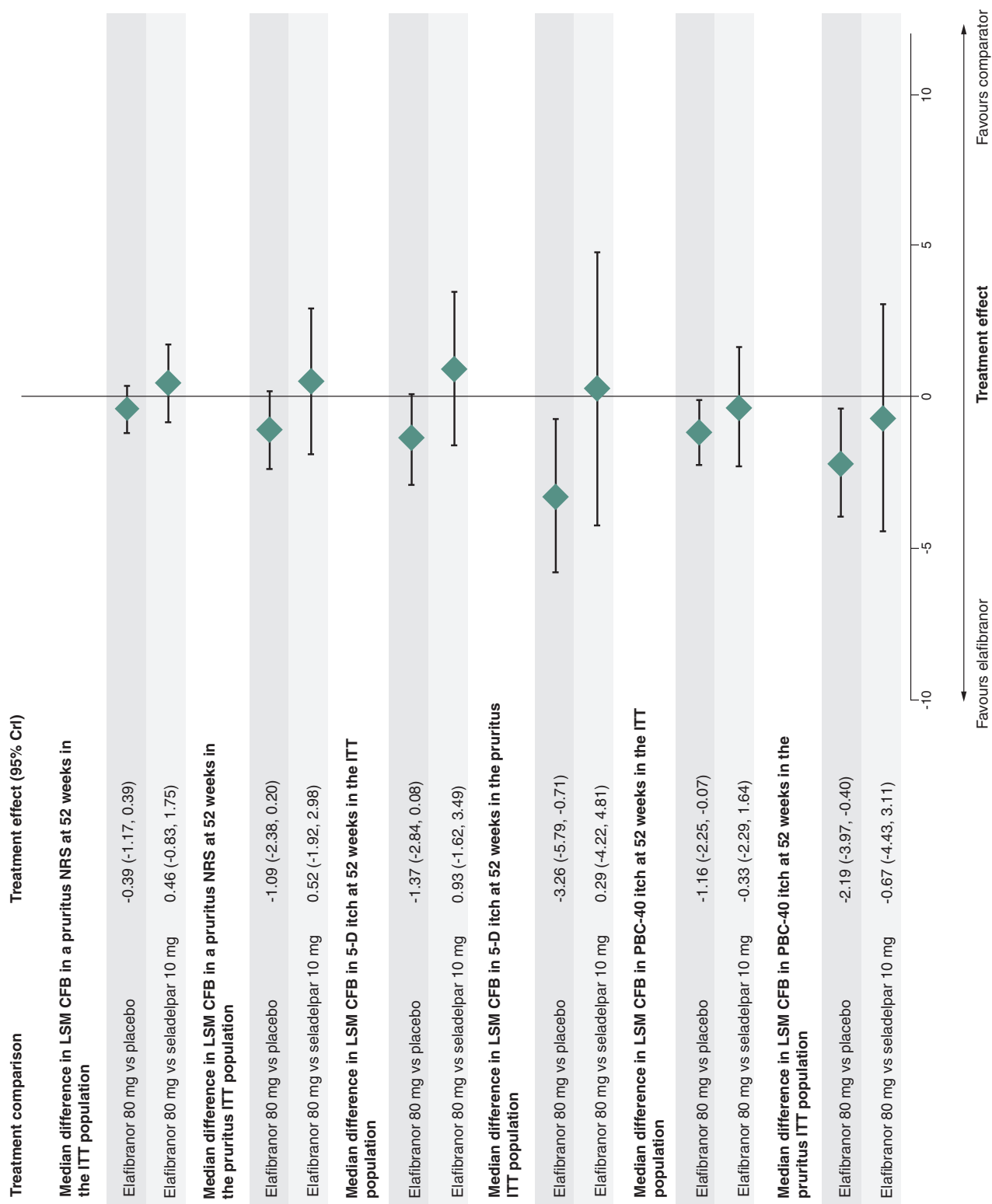


Figure 8. Median difference in least-square mean change from baseline for pruritus outcomes at 52 weeks (random-effects models).

ALP: Alkaline phosphatase; CFB: Change from baseline; CrI: Credible interval; ITT: Intention-to-treat; LSM: Least-square mean; mg: Milligram; NRS: Numerical rating scale; PBC: Primary biliary cholangitis.

subjective outcome, pruritus may be associated with more between-study variation than response-based end points. Though informed priors for between-study SD were used to mitigate this, limited study data restricted the model's ability to fully account for potential between-study differences and findings should be interpreted with caution. However, by prespecifying the analyses for this NMA in a statistical analysis plan and focusing on aligning end point definitions and analytic approaches aligned with those of the original trials, the risk of confounding in the analyses was minimized.

Considering the evidence informing the NMA, a small network led to uncertain relative treatment effects. Notably, recruiting a large study population of patients with advanced PBC was difficult, a subgroup of an already small patient population [31]. Therefore, though the studies were adequately powered to detect differences in their key end points versus placebo, the ability to achieve statistical power for ITCs may be limited [10,32]. Furthermore, excluding patients from ELATIVE not meeting the eligibility criteria of RESPONSE may have restricted the power of the analyses.

In addition to small numbers of patients contributing to the network of evidence, the network contained only one study per treatment comparison. This made it infeasible to estimate between-study SD. However, due to the limited data available for the analysis, random effects models with informed priors for between-study SD were used in the base case analyses throughout, unless there was evidence favoring the fixed effect analysis. Though this avoided overstating the precision of the results, it is a limitation of the analyses.

The lack of long-term clinical trials was also noted. To address the small number of studies included in the analysis, extra studies could have been included if analyses were conducted comparing outcomes after 12 weeks of treatment. However, given the progressive nature of PBC, longer follow-up was required to demonstrate the impact of treatment and these were prioritized for inclusion. Ongoing studies may provide further data in the future (Supplementary Table 19). Additionally, per-protocol population analyses were not possible due to insufficient data availability from RESPONSE.

The ITC results should be considered within the context of these limitations and the potential for unobserved biases. Moreover, for the binary outcomes, despite ORs being the natural output of an NMA, caution is warranted to avoid interpreting the ORs as risk ratios. Odds ratios, when used as a proxy, may overestimate risk ratios, though a ratio of 1 indicates no evidence of relative difference between treatments. Despite these limitations, the credibility of the NMA findings is supported by the similarity in key eligibility criteria, concomitant UDCA usage, baseline characteristics and outcome specifications across studies (Table 1). Further, the NMA results are validated by the consistency of the results with findings from ELATIVE and RESPONSE.

A recently published frequentist NMA by Giannini *et al.* [33] also evaluated the relative efficacy of elafibranor and seladelpar via a frequentist NMA with cholestasis response as the primary outcome, confirming that cholestasis response was statistically significantly more likely in elafibranor-treated patients than seladelpar-treated patients. Like this analysis, it also showed no significant differences in the occurrence of pruritus between elafibranor and seladelpar. A methodological advantage that this analysis has over that reported by Giannini *et al.* is that our Bayesian model provides full posterior distributions for treatment effects. While frequentist analyses produce *p*-values, which can be misinterpreted, Bayesian analyses enable probabilistic statements for direct interpretation of the probability that outcomes are better with one treatment compared with another. Bayesian models are also particularly advantageous in sparse networks, where frequentist methods can suffer from unstable variance estimation. Moreover, use of IPD allowed end point and trial eligibility harmonization for differences in ULN threshold definitions between studies. This harmonization is not possible in aggregate-data NMAs, and as such this analysis addresses a key limitation of Giannini *et al.*'s by leveraging IPD from ELATIVE to align definitions for the ULN of ALP and TB to make ELATIVE and RESPONSE data more comparable.

Another NMA comparing treatments for patients with PBC refractory to UDCA has been published by Lin *et al.* which, like this analysis, found elafibranor to be associated with better biochemical response outcomes than seladelpar [34]. However, as the NMA performed by Lin *et al.* was informed by a SLR conducted in 2023, the RESPONSE study was not included. Therefore, the data informing comparisons of elafibranor and seladelpar differs across this analysis and that published by Lin *et al.*. Additionally, there are significant methodological differences between the two analyses, which limit the ability to reliably compare the conclusions of the two ITCs.

In conclusion, this Bayesian NMA provided strong probabilistic evidence that elafibranor-treated patients were more likely to achieve cholestasis response than seladelpar-treated patients. There was no strong evidence of different treatment effects between elafibranor and seladelpar for all other outcomes. These findings suggest that elafibranor is a promising second-line treatment option for adults with PBC, helping to alleviate disease burden. To strengthen

the findings from this study, further analyses could be explored including additional end points and subgroup analyses to further evaluate the comparative effectiveness of elafibranor and seladelpar. Moreover, direct comparator studies are needed to validate the analysis' findings.

Summary points

- Summary of established knowledge on this subject:
 - Uncontrolled, primary biliary cholangitis (PBC) may progress with accumulation of symptoms such as pruritus (itching) and fatigue with significant impact on health-related quality-of-life. As PBC advances, the risk of complications necessitating liver transplant rise, such as hepatocellular carcinoma, persistent jaundice, hepatic decompensation and pruritus refractory to medical therapy.
 - Up to 40% of patients with PBC do not adequately respond to ursodeoxycholic acid (UDCA), the only recommended first-line treatment, and a further 3–5% of patients are UDCA-intolerant, highlighting a need for effective second-line treatments.
 - Elafibranor and seladelpar – both peroxisome proliferator-activated receptor agonists – have recently been approved by the US FDA, EMA and Medicine and Healthcare products Regulatory Agency to treat patients with PBC who have inadequately responded to, or are unable to tolerate, UDCA.
 - At present, the lack of head-to-head trials of elafibranor and seladelpar poses challenges for decision-makers in evaluating their relative efficacy and safety. Given this, an indirect treatment comparison (ITC) can be useful for synthesizing evidence to inform decision-making.
- What are the significant and/or new findings of this study?
 - The ELATIVE and RESPONSE randomized controlled trials were included in a novel, high-quality ITC to compare the relative efficacy and safety of elafibranor and seladelpar. Before undertaking the ITC with aggregated data, individual patient data from ELATIVE were used to align definitions for upper limit of normal of alkaline phosphatase (ALP) and total bilirubin across studies to make them more comparable. Results from the ITC analyses were consistent with existing evidence.
 - Elafibranor-treated patients had greater odds of achieving cholestasis response at 52 weeks than seladelpar-treated patients, with $\geq 99\%$ posterior probability that the odds were higher for elafibranor than seladelpar, providing strong evidence supporting elafibranor's treatment benefit. Results for this outcome were consistent in the subgroup of patients with baseline ALP ≥ 350 U/l.
 - The credible intervals around median treatment effects for all other outcomes after 52 weeks of treatment, including ALP normalization, mean change from baseline in ALP, all-cause discontinuation, mean change from baseline in pruritus, and pruritus as an adverse event, demonstrated considerable uncertainty in the relative effects of elafibranor versus seladelpar and provided no strong evidence of a difference between treatments.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: <https://becarispublishing.com/doi/pdf/10.57264/cer-2025-0206>

Author contributions

All authors have reviewed and approved the final version of the manuscript, including the authorship list. Substantial contributions to study conception/design, or acquisition/analysis/interpretation of data: All authors; Drafting of the publication, or reviewing it critically for important intellectual content: All authors; Final approval of the publication: All authors.

Acknowledgments

The authors thank Dr. Palak Trivedi, NIHR Birmingham Biomedical Research Centre, Centre for Liver and Gastrointestinal Research, University of Birmingham, Birmingham, UK, for providing valuable contributions to the network meta-analyses and development of the manuscript.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the NHS, the NIHR or the Department of Health and Social Care.

Financial disclosure

This study was sponsored by Ipsen, France.

Competing interests disclosure

D Jones has received consulting fees from Intercept, Ipsen and Umeocrine; received payment or honoraria from GlaxoSmithKline and Ipsen; received support for attending meetings and/or travel from Advanz; participates on a Data Safety Monitoring Board

or Advisory Board for Cymabay and Ipsen; is an employee of the Institute of Cellular Medicine and National Institute for Health Research (NIHR) Newcastle Biomedical Research Centre at Newcastle University. E Combe, H Knight, T Wright and V Laskier-Owens are employees of FIECON Ltd, a health economics and outcomes research agency, who have received consulting fees from Ipsen, including fees for conducting the network meta-analysis presented in the manuscript. S Ren has received consulting fees from Ipsen and is an employee of the Sheffield Centre for Health and Related Research at the University of Sheffield. EA Böing was an employee of Ipsen at the time of the analysis. A Pashley is an employee of Costello Medical, who received payment to conduct the SLR. The authors have no other competing interests or relevant affiliations with any organization or entity with the subject matter or materials discussed in the manuscript apart from those disclosed.

Writing disclosure

The authors thank Shimaila Siddiqui, MBA, of Costello Medical, Manchester, UK for providing review management support, which was sponsored by Ipsen in accordance with Good Publication Practice guidelines.

The authors thank Vicki Laskier-Owens at FIECON for providing medical writing support, which was sponsored by Ipsen in accordance with Good Publication Practice guidelines.

Ethical conduct of research

The trials informing this study were conducted in accordance with the International Council for Harmonisation Good Clinical Practice guidelines and the principles of the Declaration of Helsinki.

Data availability statement

The data from the ELATIVE study which was used in the analyses reported in this study are available within the article's Supplementary Material. The data from the RESPONSE study which was used in the analyses reported in this study are openly available in *The New England of Journal Medicine* at 10.1056/NEJMoa2312100, reference number 14.

Open access

This work is licensed under the Creative Commons Attribution 4.0 License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Axley P, Mudumbi S, Sarker S, Kuo Y-F, Singal A. Patients with stage 3 compared to stage 4 liver fibrosis have lower frequency of and longer time to liver disease complications. *PLoS ONE* 13(5), e0197117 (2018).
2. Trivedi PJ, Lammers WJ, van Buuren HR *et al.* Stratification of hepatocellular carcinoma risk in primary biliary cirrhosis: a multicentre international study. *Gut* 65(2), 321–329 (2016).
3. Cançado GGL, Lleo A, Levy C, Trauner M, Hirschfield GM. Primary biliary cholangitis and the narrowing gap towards optimal disease control. *Lancet Gastroenterol. Hepatol.* 10(9), 855–870 (2025).
4. Lindor KD, Bowlus CL, Boyer J, Levy C, Mayo M. Primary biliary cholangitis: 2018 practice guidance from the American Association for the Study of Liver Diseases. *Hepatology (Baltimore)* 69(1), 394–419 (2019).
5. Galoosian A, Hanlon C, Zhang J, Holt EW, Yimam KK. Clinical updates in primary biliary cholangitis: trends, epidemiology, diagnostics, and new therapeutic approaches. *J. Clin. Transl. Hepatol.* 8(1), 49–60 (2020).
6. Jung HE, Jang JY, Jeong SW *et al.* Prognostic indicators in primary biliary cirrhosis: significance of revised IAHG (International Autoimmune Hepatitis Group) score. *Clin. Mol. Hepatol.* 18(4), 375–382 (2012).
7. Milovanovic T, Popovic D, Stojkovic Lalosevic M, Dumic I, Dragasevic S, Milosavljevic T. Quality of life in patients with primary biliary cholangitis: a single-center experience in Serbia. *Dig. Dis.* 38(6), 515–521 (2020).
8. Hirschfield GM, Dyson JK, Alexander GJM *et al.* The British Society of Gastroenterology/UK-PBC primary biliary cholangitis treatment and management guidelines. *Gut* 67(9), 1568–1594 (2018).
9. Akamatsu N, Sugawara Y. Primary biliary cirrhosis and liver transplantation. *Intractable Rare Dis. Res.* 1(2), 66 (2012).
10. Lammers WJ, van Buuren HR, Hirschfield GM *et al.* Levels of alkaline phosphatase and bilirubin are surrogate end points of outcomes of patients with primary biliary cirrhosis: an international follow-up study. *Gastroenterology* 147(6), 1338–1349.e5 (2014).
11. Momah N, Silveira MG, Jorgensen R, Sinakos E, Lindor KD. Optimizing biochemical markers as endpoints for clinical trials in primary biliary cirrhosis. *Liver Int.* 32(5), 790–795 (2012).
12. Nevens F, Andreone P, Mazzella G *et al.* A placebo-controlled trial of obeticholic acid in primary biliary cholangitis. *N. Engl. J. Med.* 375(7), 631–643 (2016).

13. Kowdley KV, Bowlus CL, Levy C *et al.* Efficacy and safety of elafibranor in primary biliary cholangitis. *N. Engl. J. Med.* 390(9), 795–805 (2024).
- **The primary reference for the Phase III ELATIVE trial reporting elafibranor versus placebo in primary biliary cholangitis (PBC).**
14. Hirschfield GM, Bowlus CL, Mayo MJ *et al.* A Phase III trial of seladelpar in primary biliary cholangitis. *N. Engl. J. Med.* 390(9), 783–794 (2024).
- **The primary reference for the Phase III RESPONSE trial reporting seladelpar versus placebo in PBC.**
15. Invernizzi P, Floreani A, Carbone M *et al.* Primary biliary cholangitis: advances in management and treatment of the disease. *Dig. Liver Dis.* 49(8), 841–846 (2017).
16. Eruzun H, Bossen L, Gökçe DT *et al.* Clinical and biochemical characteristics of a Danish and Turkish cohort of incident and prevalent patients with primary biliary cholangitis. *Turkish J. Gastroenterol.* 36(4), 241 (2025).
17. Ipsen. Ipsen's Iqirvo[®] receives U.S. FDA accelerated approval as a first-in-class PPAR treatment for primary biliary cholangitis. (2024). <https://www.ipсен.com/press-release/ipsens-iqirvo-receives-u-s-fda-accelerated-approval-as-a-first-in-class-ppar-treatment-for-primary-biliary-cholangitis/>
18. Gilead Sciences Inc. Gilead's Livdelzi (seladelpar) granted accelerated approval for primary biliary cholangitis by U.S. FDA. (2024). <https://www.gilead.com/news/news-details/2024/gileads-livdelzi-seladelpar-granted-accelerated-approval-for-primary-biliary-cholangitis-by-us-fda>
19. Ipsen. U.K. MHRA grants marketing authorisation for Ipsen's Iqirvo[®] (elafibranor), a first-in-class peroxisome proliferator-activated receptor (PPAR) treatment for primary biliary cholangitis (PBC). (2024). <https://www.ipсен.com/uk-ireland/press-releases/u-k-mhra-grants-marketing-authorisation-for-ipsens-iqirvo-elafibranor-a-first-in-class-peroxisome-proliferator-activated-receptor-ppar-treatment-for-primary-biliary-cholangitis-p/>
20. Medicines and Healthcare products Regulatory Agency. Seladelpar approved to treat adult patients in the UK with liver illness. (2025). <https://www.gov.uk/government/news/seladelpar-approved-to-treat-adult-patients-in-the-uk-with-liver-illness>
21. Gilead Sciences Inc. Gilead's seladelpar granted conditional European marketing authorization for the treatment of primary biliary cholangitis. (2025). <https://www.gilead.com/news/news-details/2025/gileads-seladelpar-granted-conditional-european-marketing-authorization-for-the-treatment-of-primary-biliary-cholangitis>
22. Ipsen. Ipsen's Iqirvo[®] (elafibranor) approved in the European union as first new treatment for primary biliary cholangitis in nearly a decade. (2024). <https://www.ipсен.com/press-release/ipsens-iqirvo-elafibranor-approved-in-the-european-union-as-first-new-treatment-for-primary-biliary-cholangitis-in-nearly-a-decade-2949674/>
23. Schattenberg JM, Pares A, Kowdley KV *et al.* A randomized placebo-controlled trial of elafibranor in patients with primary biliary cholangitis and incomplete response to UDCA. *J. Hepatol.* 74(6), 1344–1354 (2021).
- **The primary reference for the Phase II ELATIVE trial reporting elafibranor 80 and 120 mg versus placebo in PBC.**
24. Akers J, Aguiar-Ibáñez R, Baba-Akbari A. *Systematic reviews: CRD's guidance for undertaking reviews in health care.* Centre for Reviews and Dissemination, University of York, UK (2009).
25. Sterne JAC, Savović J, Page MJ *et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 366, 14898 (2019).
26. Dias S, Welton NJ, Sutton AJ, Ades AE. *Decision Support Unit, NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials.* National Institute for Health and Care Excellence, UK (2014).
- **The NICE DSU Technical Support Document outlines the methodology used to conduct this network meta-analysis (NMA).**
27. Spiegelhalter D, Thomas A, Best N, Lunn D. *OpenBUGS User Manual, Version 3.0.2.* MRC Biostatistics Unit, UK (2007).
28. The R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing (2024).
29. Sturtz S, Ligges U, Gelman A. R2WinBUGS: a package for running WinBUGS from R. *J. Stat. Softw.* 12, 1–16 (2005).
30. Levy C, Bowlus C, Lawitz E *et al.* P250 beyond the mean: exploring the impact of baseline alkaline phosphatase levels on endpoints in primary biliary cholangitis. *Gut* 74, A239–A241 (2025).
- **Reports the impact of baseline alkaline phosphatase (ALP) levels on biochemical response and ALP normalization in ELATIVE, outlining limitations to dichotomous endpoints that limit comparability of trials with varying baseline levels.**
31. UK-PBC. Epidemiology of PBC. (2025). <https://www.uk-pbc.com/about/aboutpbc/epidemiology-of-pbc/>
32. Bahar R, Wong KA, Liu CH, Bowlus CL. Update on new drugs and those in development for the treatment of primary biliary cholangitis. *Gastroenterol. Hepatol.* 14(3), 154–163 (2018).
33. Giannini EG, Pasta A, Calabrese F *et al.* Second-line treatment for patients with primary biliary cholangitis: a systematic review with network meta-analysis. *Liver Int.* 45(1), e16222 (2025).
- **Reports an NMA conducted for OCA, seladelpar and elafibranor versus placebo as second-line treatments for PBC.**
34. Lin W, Wang J-X, Liu Y-J. Optimal drug regimens for improving ALP biochemical levels in patients with primary biliary cholangitis refractory to UDCA: a systematic review and Bayesian network meta-analysis. *Sys. Rev.* 13(1), 46 (2024).