Exploring the Relationship Between Surrogate Endpoints and Clinical Outcomes in Primary Biliary Cholangitis: A Systematic Literature Review and Meta-Analysis

CO170

Dilip Makhija[1], Marvin Rock[1], Chong H Kim[1], Mirko von Hein[2], Ryan Thaliffdeen[1], Oskar Eklund[3], Pankaj Rai[4], Howard Thom[5], Gianluca Baio[6], Barinder Singh[4]
[1]Gilead Sciences, Inc., Foster City, CA, USA, [2]Gilead Sciences, London, United Kingdom, [3]Gilead Sciences AB, Solna, Sweden, [4]Pharmacoevidence, Mohali, India, [5]Clifton Insight, Bristol, United Kingdom, [6]University College London, London, United Kingdom.

Key Findings

- The study findings emphasize the prognostic value of ALP reduction and normalization in guiding clinical decisions and monitoring treatment efficacy
- The consistent associations observed across diverse studies validate their utility as reliable indicators for adverse outcomes, including liver transplantation or death
- The study revealed that PBC patients with reduced ALP levels were less likely to experience adverse outcomes, including the need for liver transplantation or death
- These results support the integration of biomarker-based strategies to improve PBC management, potentially reducing the clinical and economic burden on patients

Conclusion

- This review underscores the critical role of surrogate endpoints, such as ALP response and ALP normalization in predicting significant clinical outcomes in liver diseases
- The analysis of ALP response criteria, including Barcelona, Toronto I, and ALP normalization, demonstrated significant associations with clinical outcomes such as liver transplantation and death across multiple studies
- The pooled effect estimates for ALP reduction and normalization were statistically significant, supporting their predictive value
- These findings highlight the potential of these markers to enhance clinical decision-making and improve the management of liver conditions by providing early insights into patient prognosis

References: 1. Kaplan MM, et al. Primary biliary cirrhosis. N Engl J Med. 2005 Sep 22;353(12):1261-73

2. Lammers WJ, 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. 2014 Dec;147(6):1338-49.

3. IntHout J, et al. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol. 2014 Feb 18;14:25

Acknowledgments: This study was funded by Gilead Sciences, Inc. Medical writing and editorial support were provided by Pharmacoevidence.

Disclosures: **BS** and **PR** are the employees of Pharmacoevidence. **DM, MR, CK, MH, RT,** and **OE** are employees of Gilead Sciences, Inc., **HT** is affiliated with Clifton Insight, and **GB** is associated with University College London

Correspondence: barinder.singh@pharmacoevidence.com

Introduction

- Primary biliary cholangitis (PBC) is a rare, chronic autoimmune liver disease characterized by slow progression, making it challenging to assess long-term clinical outcomes, such as liver transplantation (LT) and death[1]
- Surrogate markers, including alkaline phosphatase (ALP) and bilirubin levels, serve as valuable indicators of disease progression and treatment effectiveness. Lower ALP and bilirubin levels correlate strongly with improved transplant-free survival and reduced risk of adverse outcomes[2]
- Utilizing these surrogate endpoints allows for more timely and practical assessments of treatment efficacy, helping to guide treatment decisions while long-term outcomes remain difficult to predict accurately

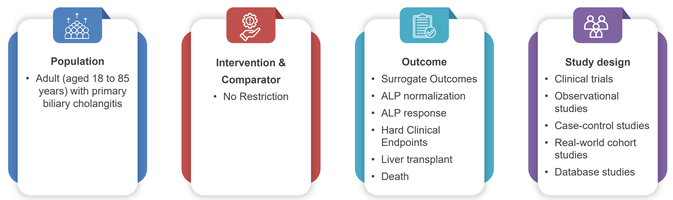
Objective

- To explore the association between ALP-based surrogate endpoints and long-term clinical outcomes in PBC through systematic literature review (SLR) and meta-analysis (MA)

Methodology

- This study adhered to National Institute for Health and Care Excellence (NICE) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for SLRs, following standard methodology with a transparent, reproducible, and unbiased approach
- EMBASE® and PubMed® were searched for English-language articles from database inception to September 2024, to identify publications evaluating the association between ALP-based surrogate endpoints and long-term clinical outcomes in PBC. The prespecified eligibility criteria are presented in Figure 1
- Two independent reviewers performed the data collection and data extraction activities, with conflicts resolved by a third independent reviewer
- The MA was performed using Stata 17, employing the DerSimonian-Laird (DL) + Hartung-Knapp-Sidik-Jonkman (HKSJ)[3] method for hazard ratio (HR) estimates, which combines the DL random-effects model with the HKSJ adjustment to improve confidence interval (CI) accuracy; Mantel-Haenszel (MH) method was employed for categorical data (i.e., n/N data)
- Statistical heterogeneity was assessed using the parameters such as Cochrane's Q, I-squared (I^2), and Tau-squared (τ^2)

Figure 1. Eligibility criteria of the SLR



Results

- Overall, 38 studies were included, reporting an association between ALP response or normalization and clinical outcomes. The PRISMA flow for the SLR and MA is provided in Figure 2
- Of these, 28 studies were peer-reviewed journal articles, while 10 were conference abstracts. The study design reported in most studies was retrospective observational (n=27), followed by prospective observational (n=11)
- Nine studies were conducted globally, with five each from the United States and China, four from Japan, three from the United Kingdom, two each from Canada and South Korea, and one each from Austria, France, Iceland, the Netherlands, Singapore, Spain, and Taiwan; details were unclear in one study (Figure 3)
- Across these studies, death or LT were the most frequently reported hard clinical endpoints (n=15), with most patients treated with UDCA alone or with fibrates
- The most commonly used ALP response criteria were Barcelona (ALP reduction ≥40% or ALP ≤1 × ULN; n=18) and Toronto I (ALP ≤1.67 × ULN; n=18), followed by ALP normalization (ALP ≤1 × ULN; n=7) and Toronto II (ALP ≤1.76 × ULN; n=3)
- Of the five studies examining the correlation between ALP response (Barcelona criteria) and LT or death, two showed statistically significant results; however, the pooled effect estimate of all studies was statistically significant (HR: 2.85, 95% CI: 1.11, 7.34) (Figure 4)

Figure 2: Flow of studies in the SLR

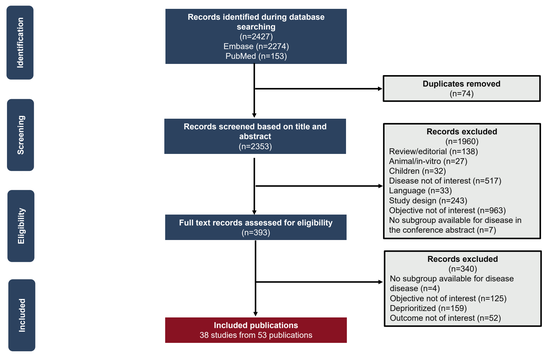


Figure 3: Study characteristics of included studies

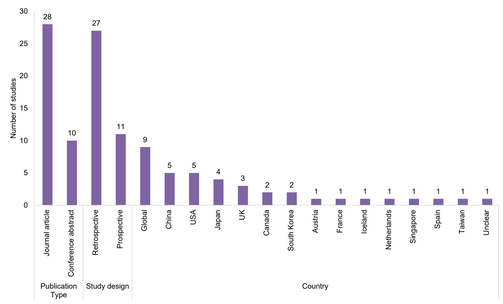


Figure 4: Association of ALP Response (Barcelona/Toronto 1 criteria) or ALP normalization with LT or death (HR estimates)

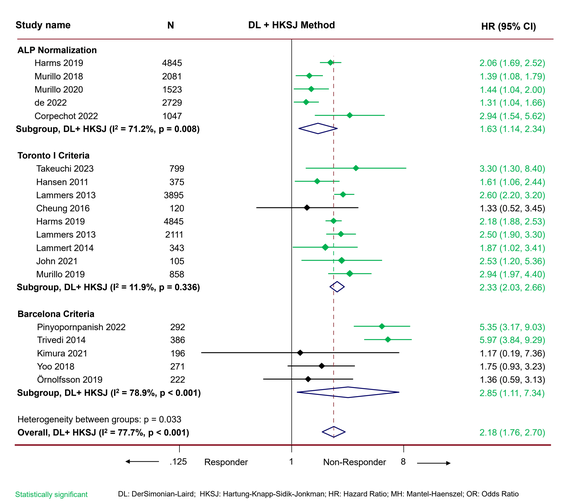
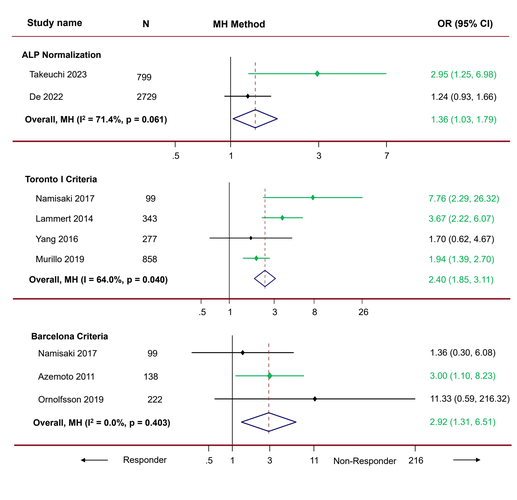


Figure 5: Association of ALP Response (Barcelona Barcelona/Toronto 1 criteria) or ALP normalization with LT or death (n/N data)



Statistically significant

DL: DerSimonian-Laird; HKSJ: Hartung-Knapp-Sidik-Jonkman; HR: Hazard Ratio; MH: Mantel-Haenszel; OR: Odds Ratio

Results (Contd.)

- Among three studies with categorical data, only one showed a significant association between non-responders and responders for these outcomes (Figure 5)
- Similarly, a statistically significant association was found in eight of nine studies examining the correlation between ALP response (Toronto I criteria) and LT/death, with one study, i.e., Cheung 2016, reported a statistically non-significant association. The pooled effect estimate was statistically significant (HR: 2.33, 95% CI: 2.03, 2.66) (Figure 4). Among four studies with categorical data, all except Yang 2016 showed significant associations for these outcomes (Figure 5)
- Five studies investigating the correlation between ALP normalization (ALP ≤ 1 X ULN) and LT, death, or LT/death found a statistically significant association in all cases. The pooled effect estimate of these studies was also statistically significant (HR: 1.63, 95% CI: 1.14, 2.34) (Figure 4)
- Of the two studies with categorical data included in the ALP normalization analysis, one (Takeuchi 2023) found a statistically significant association between non-responders and responders, while in the other study, statistically significance was not attained (De 2022) (Figure 5)
- Overall, patients who did not meet the ALP response or normalization criteria showed a significantly higher risk of LT or death compared to responders (HR: 2.18, 95% CI [1.76-2.70]) (Figure 4)

Limitations

- The focus on ALP as the main surrogate endpoint is emphasized; it is important to note that the results of other common biomarkers were also studied, but not included in the analysis, which may limit the overall scope of predicting clinical outcomes
- The majority of included studies involved patients treated with UDCA alone or in combination with fibrates. The findings may not be generalizable to patients receiving other treatments
- The quality of data and reporting standards across the included studies can vary, potentially impacting the reliability of the meta-analysis results
- The study may not fully account for all potential confounding factors that could influence the relationship between ALP levels and clinical outcomes, such as patient comorbidities and variations in treatment adherence