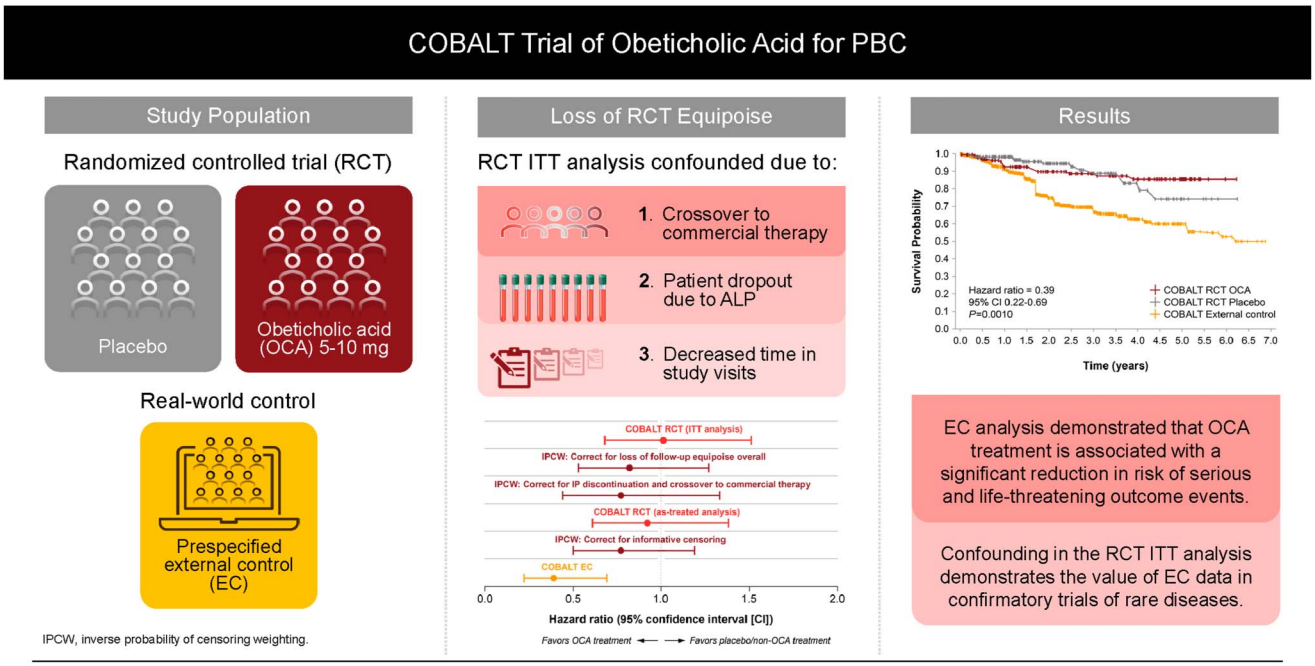
Open

COBALT: A Confirmatory Trial of Obeticholic Acid in Primary Biliary Cholangitis With Placebo and External Controls

Kris V. Kowdley, MD, FACG[1], Gideon M. Hirschfield, PhD[2], Charles Coombs, ME[3], Elizabeth S. Malecha, PhD[4], Leona Bessonova, PhD, MBA[4], Jing Li, PhD[4], Nuvan Rathnayaka, PhD[5], George Mells, PhD[6], David E. Jones, PhD[7], Palak J. Trivedi, MD, PhD[8], Bettina E. Hansen, PhD[9,10,11], Rachel Smith[12], James Wason, PhD[7], Shaun Hiu, MS[7], Dorcas N. Kareithi, MSc[7], Andrew L. Mason, MD[13], Christopher L. Bowlus, MD[14], Kate Muller, PhD[15], Marco Carbone, PhD[16], Marina Berenguer, MD[17], Piotr Milkiewicz, MD[18,19], Femi Adekunle, MD[20] and Alejandra Villamil, MD[21]

INTRODUCTION: **Obeticholic acid (OCA) treatment for primary biliary cholangitis (PBC) was conditionally approved in the phase 3 POISE trial. The COBALT confirmatory trial assessed whether clinical outcomes in patients with PBC improve with OCA therapy.**



Kowdley et al. *Am J Gastroenterol.* 2025. doi:10.14309/ajg.0000000000003029

[1]Liver Institute Northwest and Elson S. Floyd College of Medicine, Washington State University, Seattle, Washington, USA; [2]Toronto Centre for Liver Disease, Toronto General Hospital, University of Toronto, Toronto, Ontario, Canada; [3]Real World Evidence, Syneos Health, Morrisville, North Carolina, USA; [4]Intercept Pharmaceuticals, Morristown, New Jersey, USA; [5]Department of Population Health Sciences, Duke University, Durham, North Carolina, USA; [6]Cambridge University Hospitals NHS Foundation Trust MRC Clinical Academic Research Partner, Academic Department of Medical Genetics, University of Cambridge, Cambridge, UK; [7]Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, UK; [8]National Institute for Health and Care Research (NIHR) Birmingham Biomedical Research Centre, Centre for Liver and Gastrointestinal Research, Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham, UK; [9]Department of Epidemiology and Biostatistics, Erasmus MC, Rotterdam, Netherlands; [10]IHPME University of Toronto, Toronto, Ontario, Canada; [11]Toronto Centre for Liver Disease and TGHRI, University Health Network, Toronto, Ontario, Canada; [12]Cambridge Liver Unit, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK; [13]Division of Gastroenterology, University of Alberta, Edmonton, Alberta, Canada; [14]University of California Davis, Sacramento, California, USA; [15]Flinders Medical Centre, Adelaide, South Australia, Australia; [16]University of Milano-Bicocca, Milano, Italy; [17]La Fe University Hospital, IISLaFe, Ciberehd, University of Valencia, Valencia, Spain; [18]Medical University of Warsaw, Warszawa, Poland; [19]Translational Medicine Group, Pomeranian Medical University, Szczecin, Poland; [20]Intercept Pharmaceuticals, London, UK; [21]Hospital Italiano de Buenos Aires, Buenos Aires, Argentina. **Correspondence:** Kris V. Kowdley, MD. E-mail: kkowdley@liverinstitutenw.org.

Received June 6, 2024; accepted July 25, 2024; published online August 14, 2024

METHODS: **Patients randomized to OCA (5–10 mg) were compared with placebo (randomized controlled trial [RCT]) or external control (EC). The primary composite endpoint was time to death, liver transplant, model for end-stage liver disease score ≥15, uncontrolled ascites, or hospitalization for hepatic decompensation. A prespecified propensity score–weighted EC group was derived from a US healthcare claims database.**

RESULTS: **In the RCT, the primary endpoint occurred in 28.6% of OCA (n = 168) and 28.9% of placebo patients (n = 166; intent-to-treat analysis hazard ratio [HR] = 1.01, 95% confidence interval = 0.68–1.51), but functional unblinding and crossover to commercial therapy occurred, especially in the placebo arm. Correcting for these using inverse probability of censoring weighting and as-treated analyses shifted the HR to favor OCA. In the EC (n = 1,051), the weighted primary endpoint occurred in 10.1% of OCA and 21.5% of non-OCA patients (HR = 0.39; 95% confidence interval = 0.22–0.69; *P* = 0.001). No new safety signals were identified in the RCT.**

DISCUSSION: **Functional unblinding and treatment crossover, particularly in the placebo arm, confounded the intent-to-treat estimate of outcomes associated with OCA in the RCT. Comparison with the real-world EC showed that OCA treatment significantly reduced the risk of negative clinical outcomes. These analyses demonstrate the value of EC data in confirmatory trials and suggest that treatment with OCA improves clinical outcomes in patients with PBC.**

KEYWORDS: confounding variables; functional unblinding; inverse probability of censoring weighting; treatment crossover

SUPPLEMENTARY MATERIAL accompanies this paper at https://links.lww.com/AJG/D380

Am J Gastroenterol 2025;120:390–400. https://doi.org/10.14309/ajg.0000000000003029

INTRODUCTION

Primary biliary cholangitis (PBC) is a progressive liver disease characterized by destruction of intrahepatic bile ducts, which can lead to cirrhosis and complications of end-stage liver disease (1). The disease predominantly affects women and is typically diagnosed between 40 and 60 years of age (1). In the United States, the prevalence of PBC ranges from 16.0 to 40.9 per 100,000 persons (2–6).

Randomized controlled trials (RCTs) of potential therapeutics for PBC have long been challenging to perform because of its rarity and slow rate of progression in some patients, which makes it difficult to recruit an adequate number of patients and accrue sufficient outcome events to detect statistically significant differences between treatments (7). Early trials of ursodeoxycholic acid (UDCA) were not powered to detect an effect of therapy on survival, which was not demonstrated until after its regulatory approval (7). In contrast to these early trials, which compared UDCA vs placebo in treatment-naive patients, trials performed after UDCA became first-line treatment for PBC also needed to randomize patients who did not respond to or who could not tolerate UDCA to both treatment arms, which further limited the pool of eligible patients (7,8).

In 2016, obeticholic acid (OCA) received accelerated approval from the United States Food and Drug Administration (US FDA) and European Medicines Agency (EMA) for the treatment of PBC based on reduction in alkaline phosphatase (ALP) as a surrogate endpoint in the phase 3 POISE trial (9,10). A significantly higher proportion of patients treated with OCA for 12 months (at a dose of 5 mg/d titrated to 10 mg/d or 10 mg/d fixed) achieved the primary endpoint of ALP <1.67 times the upper limit of normal (ULN) and normal total bilirubin (TB) than those who received placebo (approximately 46% vs 10%; $P < 0.001$) (9). In the subsequent open-label extension (OLE), the reductions in ALP and TB were sustained for an additional 5 years (11). At the time of accelerated approval, the US FDA required that an adequate and well-controlled trial be conducted to confirm the efficacy and safety of OCA in patients with PBC. COBALT (NCT02308111) was to be performed to assess the impact of OCA on a composite endpoint of progression to death, liver transplant, or hospitalization for hepatic decompensation in a predominantly cirrhotic patient population. In May 2021, the US FDA contraindicated OCA in patients with decompensated cirrhosis with evidence of portal hypertension, which included approximately 55% of patients enrolled in COBALT (12). This contraindication was included in the Boxed Warning section of the US Prescribing Information for OCA (13).

External control (EC) studies are potentially useful in clinical trials of therapeutic agents for rare diseases such as PBC when inclusion of a placebo control group may be considered unethical and patient recruitment and retention can be difficult (14–16). A previous study compared patients treated with OCA in the OLE arm of the POISE trial with an external cohort of non-OCA–treated individuals from 2 multicenter PBC registries (Global PBC and UK-PBC) to assess the long-term efficacy of OCA (16,17). This analysis showed that OCA treatment substantially reduced the likelihood of liver transplantation and death in patients with PBC (POISE vs Global PBC, hazard ratio [HR] = 0.29, 95% confidence interval [CI] = 0.10–0.83; POISE vs UK-PBC, HR = 0.30, 95% CI = 0.12–0.75) (16).

The phase 3b/4 COBALT study was designed to examine the association of OCA with clinical outcomes in a high-risk population of patients with PBC, namely those with baseline biomarker levels that are associated with more advanced disease or greater risk of progression, as described below. COBALT included a randomized placebo control arm and an EC arm. In this article, we describe the final results from COBALT and describe

potential factors leading to confounding and bias because of functional unblinding and treatment crossover to commercially available therapy in the RCT. We demonstrate a novel approach to address these sources of bias through use of an appropriately matched real-world–based EC to provide evidence of clinical benefit in rare, slowly progressing diseases such as PBC.

METHODS

Study design and oversight

COBALT was a phase 3b/4 double-blind RCT with supportive EC analyses. The EC analysis was prespecified in October 2014 (before the start of COBALT in February 2015) to mitigate the risks associated with inclusion of a long-term placebo arm in the study and to assess the efficacy of OCA vs a real-world comparator. In the RCT, patients received either once-daily oral placebo or OCA 5 mg/d, increased to 10 mg at 3 months if tolerated. Randomization was stratified by UDCA treatment (yes/no) and TB (≤ULN/>ULN) at baseline.

The study protocol was approved by appropriate local and national ethics and regulatory agencies and was conducted in accordance with Good Clinical Practice guidelines and the Declaration of Helsinki (Seoul, South Korea, October 2008 amendment). All study participants provided written informed consent. Details regarding trial oversight and author contributions are provided in the Supplementary Digital Content (see Supplementary Appendix, http://links.lww.com/AJG/D380).

Trial participants

Patients aged ≥18 years diagnosed with PBC were enrolled at 137 sites in 27 countries starting in February 2015 (see Supplementary Materials—Recruitment of study participants, http://links.lww.com/AJG/D380 for details). The original entry criteria included mean ALP >5× ULN and mean TB > ULN and ≤3× ULN. Subsequently, these criteria were revised to ALP >3× ULN and mean TB > ULN and ≤5× ULN to increase patient recruitment. Eligible patients included those who had either discontinued UDCA >3 months earlier or who were taking UDCA >12 months with an approved, stable dose ≥3 months before enrollment. Patients with concomitant liver disease, cirrhosis with complications in the previous 12 months, or other significant medical conditions were excluded (see Supplementary Table S1, http://links.lww.com/AJG/D380 for complete inclusion/exclusion criteria). The intent-to-treat (ITT) and safety analysis populations included all randomized patients who received at least 1 dose of OCA or placebo.

Primary efficacy endpoint

The composite endpoint for the ITT analysis was time to first occurrence of any of the following events: death (all-cause); liver transplant; model for end-stage liver disease (MELD) score ≥15; hospitalization ≥24 hours for new onset or recurrence of variceal bleed, hepatic encephalopathy (West Haven score ≥2), or spontaneous bacterial peritonitis (confirmed by diagnostic paracentesis); or uncontrolled ascites requiring therapeutic paracentesis ≥2 times in a month. All events were adjudicated by a blinded committee of experts with adjudication experience who were not involved in the study as investigators, data monitoring committee members, or consultants. Adjudication of suspected events was based on available source documentation, including, but not limited to, individual clinical study data, hospital records, histology, and/or death certificates.

Safety and tolerability

Safety was assessed by the incidence of treatment-emergent adverse events (TEAEs). Tolerability was assessed by comparing the proportion of patients in the OCA vs placebo arms who discontinued treatment.

External controls

The prespecified EC analysis used Komodo Healthcare Map, a large US healthcare claims database with approximately 330 million unique patients. Komodo was linked to other databases using Datavant tokenization, which enabled deidentified patient records to be matched across health plans in a US Health Insurance Portability and Accountability Act–compliant manner (18). These databases included laboratory data from LabCorp and Quest Diagnostics; transplant data from the Organ Procurement and Transplantation Network, a registry that contains US transplant information, including donors, waitlisted candidates, and transplant recipients; and vital statistics data including the US Social Security Death Index combined with a national obituary search. Eligible patients had 1 inpatient or 2 outpatient *International Classification of Diseases, 10th Revision* codes for PBC between January 1, 2014, and December 31, 2021 (n = 109,970); 1 year of baseline data, including all laboratory tests (ALP, TB, alanine aminotransferase [ALT], aspartate aminotransferase [AST], and platelet count; n = 7,693); and met modified COBALT study inclusion/exclusion criteria (see Supplementary Table S2, http://links.lww.com/AJG/D380; n = 1,233). Among all healthcare visits at which a patient met eligibility criteria, a randomly selected visit was defined as the index date.

The composite endpoint of the EC analysis was designed to be as consistent as possible with the composite endpoint of the COBALT RCT. Specifically, the composite endpoint for the EC analysis was time to first occurrence of any of the following events: death (all-cause); liver transplant; hospitalization for new onset or recurrence of gastroesophageal variceal bleed or hepatic encephalopathy; or uncontrolled ascites, including spontaneous bacterial peritonitis, transjugular intrahepatic portosystemic shunt procedure, or paracentesis ≥2 times per month for ≥2 months. MELD score was excluded from the primary composite endpoint because not all component laboratory data were available in the Komodo database.

Statistical analyses

In the COBALT RCT, the ITT analysis was a log-rank test of the randomized OCA and placebo cohorts with respect to the primary composite endpoint, stratified by the randomization stratification factors. Assuming 1:1 randomization, exponential survival curves, and a placebo survival estimate of 0.6 at 8 years with an HR of 0.60, 428 patients with 127 events were required to provide 80% power to detect an effect on a 2-sided test at a 5% level of significance.

The EC analysis was a log-rank test of OCA patients in COBALT and comparable non-OCA–treated EC individuals with respect to the primary composite endpoint (excluding MELD score). Comparability was achieved by using propensity scores and standardized morbidity ratios to weight EC patients to closely resemble randomized OCA patients (see Supplementary Table S3, http://links.lww.com/AJG/D380) (15,19). Propensity scores were estimated using a logistic regression model with treatment arm (OCA vs EC) as the dependent variable and covariates (age at index; sex; presence/absence of portal hypertension, hepatic

decompensation, or cirrhosis; laboratory values at index; use of UDCA at index; and dichotomous calendar year [pre-/post-COVID]) associated with the composite endpoint as independent variables (see Supplementary Table S4, http://links.lww.com/AJG/D380). The balance among covariates at baseline was assessed using Wilcoxon rank-sum tests and the standardized mean difference (SMD) between treatment and control (see Supplementary Figure S1, http://links.lww.com/AJG/D380). The potential for residual confounding was also assessed using SMDs between the randomized control and EC groups.

In both ITT and EC analyses, HR and 95% CI were calculated using stratified Cox regression models. In time-to-event analyses of the EC arm, weighted control patients were censored at initiation of commercial OCA or database disenrollment. Patients treated with OCA were censored 90 days after discontinuation of investigational product (IP), consistent with outcomes analyses of real-world claims data in recent trial emulation studies, such as the RCT-DUPLICATE initiative (20). Censoring also occurred at study discontinuation, when patients were lost to follow-up, or when consent was withdrawn.

Safety assessments were conducted at baseline and each visit. TEAEs were summarized using Medical Dictionary for Regulatory Activities (version 23.1) system organ class, preferred term, and severity (as defined per protocol).

Sensitivity analyses

Sensitivity analyses were performed to assess whether differential discontinuation and treatment crossover occurred, and if so, whether they were due to treatment response and functional unblinding. To examine this possibility, the mean ALP in the OCA and placebo arms was compared over time among patients who discontinued study visits or started commercial therapy. In addition, the relationship between IP discontinuation and initiation of commercial PBC therapies was evaluated by comparing mean ALP by treatment arm up to 6 months before vs up to 12 months after initiation of commercial therapy. All patients provided informed consent for the measurement and analysis of ALP levels.

Because ALP measures were collected and observed during follow-up in the blinded RCT, functional unblinding may also have led to differential dropout in the placebo and OCA arms. The impact of potential confounding from informative censoring was assessed by inverse probability of censoring weighting (IPCW) analysis. Weights were derived by estimating the probability of remaining uncensored at the end of each of 10 intervals of follow-up time adjusted for prespecified baseline variables and post-baseline time-varying covariates using a pooled logistic regression model. Prespecified baseline predictors of both censoring and clinical outcomes included treatment arm (OCA/placebo); age at screening; sex; treatment with UDCA at screening (yes/no); and baseline levels of ALP, ALT, AST, TB, albumin, and platelet count. Postbaseline time-varying covariates included ALP, ALT, AST, and TB levels. The HR and 95% CI from the IPCW analysis were estimated using a weighted Cox proportional hazards model with the treatment arm as an independent variable and a sandwich variance estimator.

IPCW sensitivity analysis was applied to a subset of the ITT population (excluding data from >90 days after patients discontinued IP) to assess the impact of bias from IP discontinuation. In addition, a subset that excluded data after the initiation of commercial OCA, UDCA (if not treated with UDCA at baseline), or fibrates (if not treated with fibrates at baseline) was analyzed to determine the impact of bias from initiation of second-line therapy (i.e., in addition to bias from IP discontinuation).

Patient follow-up was also assessed in the randomized cohorts by comparing the proportion of patients who had regularly scheduled on-site study visits vs those who had follow-up through biannual telephone calls and medical record reviews (patients provided informed consent for these activities). Because the latter precluded collection of information needed to ascertain the primary efficacy endpoint (e.g., laboratory blood tests to calculate MELD score) and functional unblinding and treatment crossover compromised the ITT analysis of the RCT (explained in the Results section), we performed a post-hoc as-treated analysis by classifying placebo crossover patients as "OCA-exposed." An IPCW sensitivity analysis was also applied to this as-treated subset to assess the impact of bias from informative censoring.

RESULTS

Study populations and baseline characteristics in the COBALT RCT

The COBALT trial was terminated in December 2021 as recommended by the data monitoring committee and in alignment with health authorities because it did not seem feasible to continue the study as designed owing to the impossibility of conducting a placebo-controlled randomized trial of long-term outcomes in the setting of commercially available therapies. At that time, 631 patients had been screened, with 334 patients randomized (placebo, n = 166; OCA, n = 168; 78% of recruitment target). Demographics (mean age ~53 years, ~90% female, ~86% White) and clinical characteristics (~88% prescribed UDCA at baseline, ~20% with a history of hepatic decompensation) were generally balanced in the placebo and OCA arms (Table 1). The proportion of patients with a history of portal hypertension was slightly higher in the placebo (n = 98; 59.0%) vs OCA arms (n = 82; 48.8%).

Endpoint analyses in the COBALT RCT

In the ITT analysis of COBALT, there were 48 events each in the OCA (28.6%) and placebo (28.9%) arms (HR, 1.01; 95% CI, 0.68–1.51; Figure 1a). Of the 48 patients in the OCA arm who had an event, 31 (64.6%) discontinued IP a median of 9 months (interquartile range [IQR], 2–17 months) before the event, and 1 initiated commercial OCA. Among the 48 patients in the placebo arm who had an event, 29 (60.4%) discontinued placebo a median of 4 months (IQR, 1–10 months) before the event. Overall, in the OCA arm, 22.6% initiated commercial treatments, namely, 7.7% commercial OCA, 11.9% fibrates, and 3.0% UDCA. In the placebo arm, 32.5% initiated commercial treatments, namely, 15.7% commercial OCA, 12.7% fibrates, and 4.2% UDCA. Details about patient disposition in the ITT analysis are shown in Supplementary Digital Content (see Supplementary Table S5, http://links.lww.com/AJG/D380). Analyses of individual components of the composite endpoint are shown in Supplementary Digital Content (see Supplementary Tables S7 to S12 and Supplementary Figures S2 and S3, http://links.lww.com/AJG/D380).

In the as-treated analysis, 194 patients were classified as OCA-exposed, and 140 patients were classified as non-OCA–exposed.

Table 1. Baseline patient demographic and disease characteristics

	COBALT		Komodo[a] Untreated control (unweighted n = 1,051) (weighted n = 165)
	Placebo (n = 166)	OCA (n = 168)	
Age, y			
Mean (SD)	53.9 (10.4)	53.4 (10.3)	53.8 (5.2)
Sex, n (%)			
Female	149 (89.8)	151 (89.9)	146.5 (88.8)
Race, n (%)			
American Indian or Alaska Native	2 (1.2)	1 (0.6)	ND
Asian	9 (5.4)	11 (6.5)	ND
Black or African American	2 (1.2)	2 (1.2)	ND
White	143 (86.1)	146 (86.9)	ND
Multiple races	1 (0.6)	4 (2.4)	ND
Not reported	9 (5.4)	4 (2.4)	ND
Baseline/index laboratory value, mean (SD)			
ALP, U/L	499.3 (294.5)	481.3 (276.7)	463.8 (104.4)
TB, mg/dL	1.7 (0.8)	1.6 (0.8)	1.6 (0.5)
ALT, U/L	84.0 (50.2)	81.5 (51.9)	77.1 (35.9)
AST, U/L	81.9 (36.3)	81.1 (41.1)	79.1 (35.8)
ALB, g/dL	3.96 (0.40)	4.00 (0.41)	4.00 (0.20)
Platelets ×1,000/μL	197.2 (102.6)	209.9 (101.8)	194.8 (37.3)
Disease history at baseline, n (%)			
Decompensation	37 (22.3)	32 (19.0)	36.3 (22.0)
Portal hypertension	98 (59.0)	82 (48.8)	90.1 (54.6)
Use of UDCA, n (%)			
At baseline/index	147 (88.6)	147 (87.5)	143.2 (86.8)
Ever used	163 (98.2)	161 (95.8)	165 (100)
Never	3 (1.8)	7 (4.2)	0

Percentages are based on the number of patients in the intent-to-treat population within each treatment arm.
ALB, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ND, no data; OCA, obeticholic acid; TB, total bilirubin; UDCA, ursodeoxycholic acid.
[a]N represents weighted populations.

There were 55 events in the OCA arm (28.4%) and 41 events in the placebo arm (29.3%) (HR, 0.92; 95% CI, 0.61–1.38).

External control analysis

A total of 1,051 non-OCA patients who met the eligibility criteria were identified in the Komodo Healthcare Map database. After weighting, the OCA and EC arms were well balanced (Table 1). The SMDs of all covariates except platelet count and clinical evidence of portal hypertension were less than the prespecified ±0.10 threshold (see Supplementary Figure S1, http://links.lww.com/AJG/D380). Because the SMDs for the 2 exceptions were only slightly above the threshold (0.15 and 0.12, respectively), all covariates were included in the Cox regression model.

In the EC analysis, there were 17 events (10.1%) in the COBALT OCA arm and 35.4 events (21.5%) among weighted non-OCA individuals (HR, 0.39; 95% CI, 0.22–0.69; $P = 0.0010$; Figure 1b). The proportion of events in the Komodo OCA arm was similar to that in the COBALT OCA arm (see Supplementary Materials—Comparison of randomized vs Komodo OCA patients, http://links.lww.com/AJG/D380).

Details about patient disposition in the EC analysis are shown in Supplementary Digital Content (see Supplementary Table S6, http://links.lww.com/AJG/D380). Analyses of individual components of the composite endpoint are shown in Supplementary Digital Content (see Supplementary Tables S13 to S16 and Supplementary Figure S4, http://links.lww.com/AJG/D380).

Sensitivity analyses of potential sources of bias in the ITT analysis of the COBALT RCT

Observable ALP levels. In the COBALT RCT, most placebo patients who discontinued study visits or initiated commercial therapy before an endpoint event did so before year 2 (Figure 2) and tended to have higher ALP levels than patients continuing in study visits.

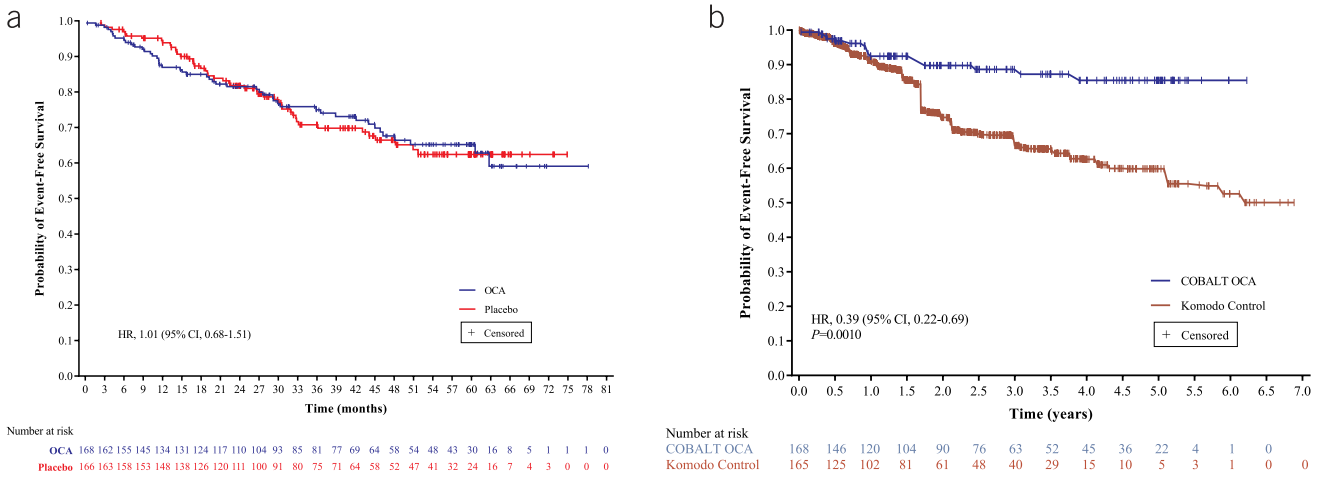


Figure 1. Kaplan–Meier survival curves for the COBALT primary composite endpoint. This figure illustrates the survival probabilities and number of OCA and non-OCA patients with primary biliary cholangitis who are at risk of death, liver transplant, or hepatic decompensation over time. (**a**) Intent-to-treat analysis of OCA vs placebo in the COBALT RCT. (**b**) COBALT OCA arm vs external control from the Komodo Healthcare Map database. The endpoint of this analysis did not include model for end-stage liver disease score ≥15 because not all component laboratory data were available in the Komodo database. Censoring is indicated by + symbols. HR, hazard ratio; OCA, obeticholic acid; RCT, randomized controlled trial.

Initiation of commercial PBC therapies. In the placebo arm of the COBALT RCT, treatment discontinuation before an endpoint event was highest during months 12–24 (n = 41 [24.7%]; Figure 3a). In the OCA arm, it peaked during the first 6 months (n = 25 [14.9%]; Figure 3b). The proportion of patients who initiated commercial treatment increased faster in the placebo arm than in the OCA arm, reaching 25% of the cohort in approximately 19 vs 48 months after randomization, respectively (dashed lines in Figure 3a,b; red and blue dotted lines in Figure 4). During the entire study, placebo patients were significantly more likely to initiate non-IP therapy compared with OCA patients (HR, 1.59; 95% CI, 1.02–2.50; $P = 0.040$). In addition, placebo patients who initiated commercial PBC treatment had higher ALP before switching from randomized therapy than OCA patients (568.8 vs 355.9 U/L) and had a more substantial decrease in ALP after crossover (106.3 vs 67.5 U/L; Figure 3c). Of note, placebo patients who initiated commercial therapy, which lowered ALP, were not censored and were included in the ITT analysis as per US FDA requirement.

Remote patient follow-up. In the COBALT RCT, the proportion of patients at risk of an endpoint event who had regularly scheduled on-site (in-clinic) study visits was greater in the OCA arm than in the placebo arm >12 months after randomization (Figure 4, solid lines). From months 21–51, the proportion remained approximately 90% in the OCA arm, whereas it decreased to approximately 70% in the placebo arm. Placebo patients spent 12% of follow-up time outside of regular study visits compared with 8.1% for OCA patients.

IPCW analyses. As shown in Figure 4, the number of patients in the randomized ITT population of COBALT who discontinued IP (dashed lines) or initiated second-line therapy (dotted lines) increased with time, with a greater dropout in the placebo vs OCA cohorts. Using IPCW to adjust for this differential dropout reduced the HR from the ITT analysis of the randomized cohorts in favor of OCA and trended toward the HR from the EC analysis

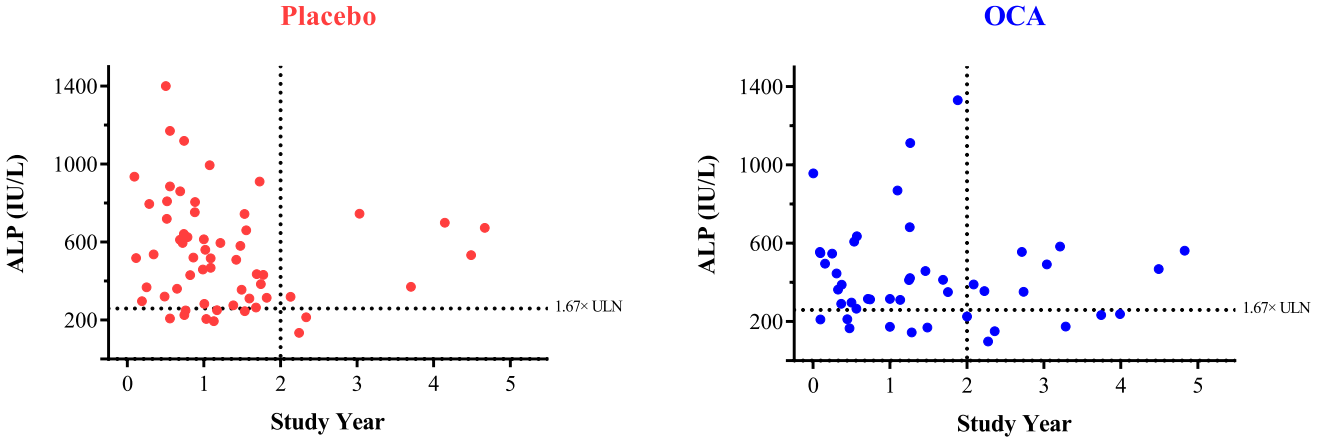


Figure 2. ALP levels vs time after randomization for placebo and OCA patients who discontinued study visits or started commercial therapy before an endpoint event. Each data point represents a patient's ALP level at the last study visit before an endpoint event. ALP, alkaline phosphatase; OCA, obeticholic acid; ULN, upper limit of normal.

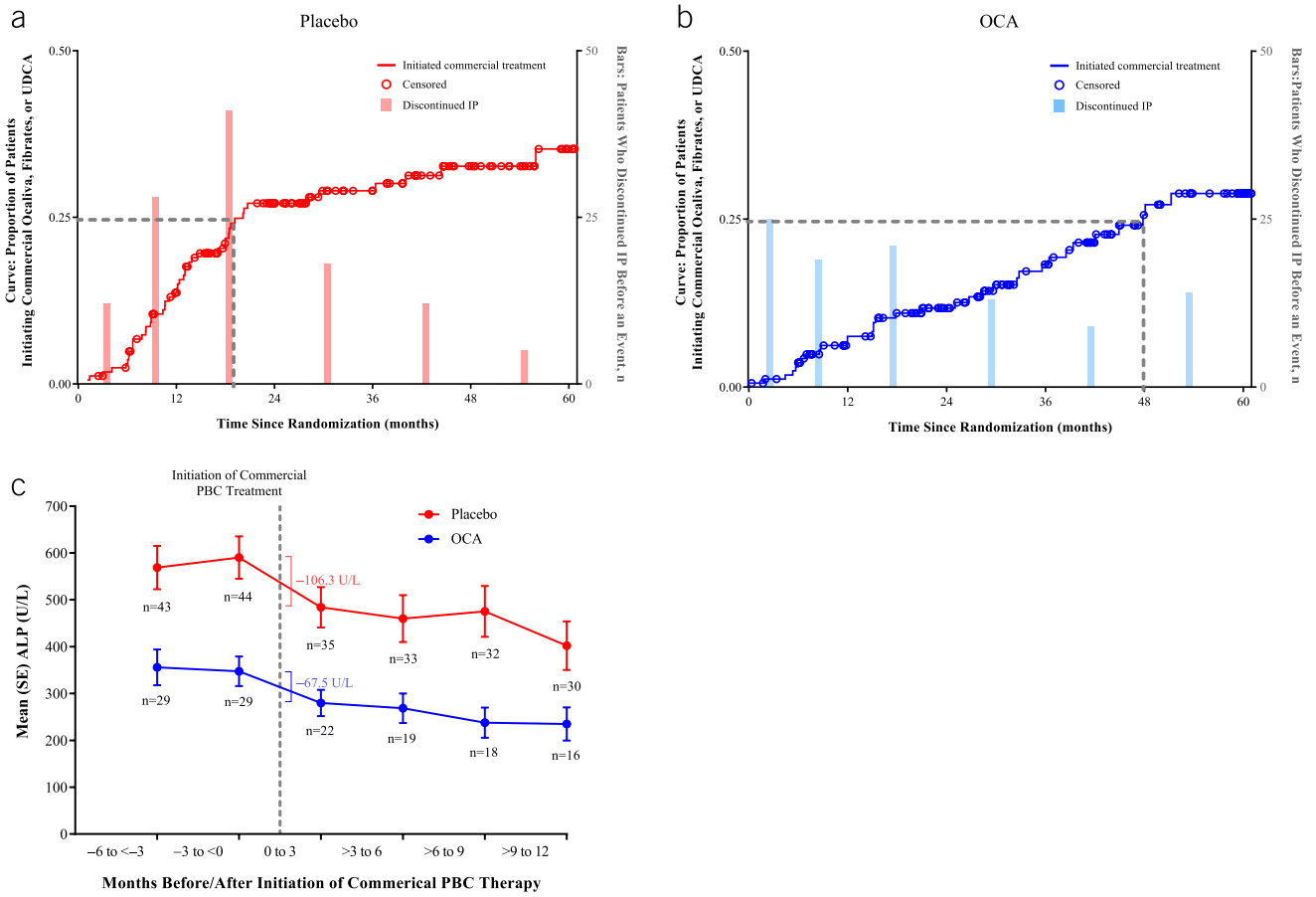


Figure 3. Evidence of confounding in the ITT analysis of the COBALT RCT. Panels (**a**) and (**b**) illustrate the number of patients who discontinued IP before an endpoint event (bars) over time vs the proportion of patients who initiated commercial PBC treatments (curves) over time in the placebo arm (**a**) and OCA treatment arm (**b**). Commercial PBC treatments included commercial OCA, fibrates, or UDCA (if the patients were not receiving UDCA at baseline). The dotted lines show the time since randomization for 25% of patients to initiate commercial treatment in the placebo and OCA cohorts. Panel (**c**) shows mean ALP levels before and after initiating commercial treatment by treatment arm. ALP, alkaline phosphatase; IP, investigational product; ITT, intent-to-treat; OCA, obeticholic acid; PBC, primary biliary cholangitis; RCT, randomized controlled trial; UDCA, ursodeoxycholic acid.

(Figure 5). Specifically, IPCW analysis of the ITT population reduced the HR point estimate by 18.8% to 0.82% (95% CI, 0.53–1.27). Exclusion of data >90 days after IP discontinuation or after initiation of second-line therapy reduced this by a further 6.1% to 0.77% (95% CI, 0.44–1.33), indicating a trend for benefit in favor of OCA. IPCW analysis of the as-treated subset resulted in a similar HR of 0.77 (95% CI, 0.50–1.19).

Safety and tolerability of OCA treatment in the COBALT RCT. Treatment with OCA was generally well tolerated in patients with PBC. In the COBALT trial, the most common TEAE was pruritus, which was reported in 78.6% of OCA patients and 51.2% of placebo patients (Table 2). Other TEAEs (≥10% incidence) that were reported more often in the OCA vs placebo arms included peripheral edema (OCA, 18.5%; placebo, 10.8%), upper abdominal pain (OCA, 14.9%; placebo, 7.2%), abdominal pain (OCA, 12.5%; placebo, 10.8%), nausea (OCA, 14.9%; placebo, 12.7%), headache (OCA, 13.7%; placebo, 12.7%), constipation (OCA, 11.3%; placebo, 6.0%), and nasopharyngitis (OCA, 10.7%; placebo, 8.4%).

Both treatment arms had similar rates of serious TEAEs (OCA, 31.5%; placebo, 31.9%). Severe TEAEs occurred in <2% of patients in both arms, except for pruritus (OCA, 30.4%; placebo, 9.0%) and esophageal variceal hemorrhage (OCA, 2.4%; placebo, 1.8%). In addition, hepatic TEAEs occurred less often in OCA vs placebo patients overall (47.6% vs 58.4%), including increased bilirubin (11.9% vs 15.1%), esophageal varices (11.9% vs 16.9%), and ascites (10.7% vs 12.7%). Approximately one-third of all patients, 62 (36.9%) in the OCA arm and 45 (27.1%) in the placebo arm discontinued IP because of a TEAE. Of these discontinuations, 19 of 62 patients (30.6%) in the OCA arm and 3 of 45 patients (6.7%) in the placebo arm were due to pruritus.

DISCUSSION

COBALT was designed as a randomized, placebo-controlled, confirmatory clinical outcomes trial to assess the efficacy and safety of OCA in patients with PBC. Despite extensive efforts to recruit and retain study participants after the commercial availability of OCA as well as multiple consultations with the US FDA and EMA Committee for Medicinal Products for Human Use, the data monitoring committee recommended against further patient enrollment in September 2020 and concluded that the study was futile in October 2021. Shortly thereafter, in alignment with the US FDA and EMA, the COBALT study was terminated early.

The ITT analysis of the RCT was largely compromised by functional unblinding and treatment crossover, which prevented the prespecified primary objective from being addressed adequately. Because measurement of ALP levels is part of standard-of-care

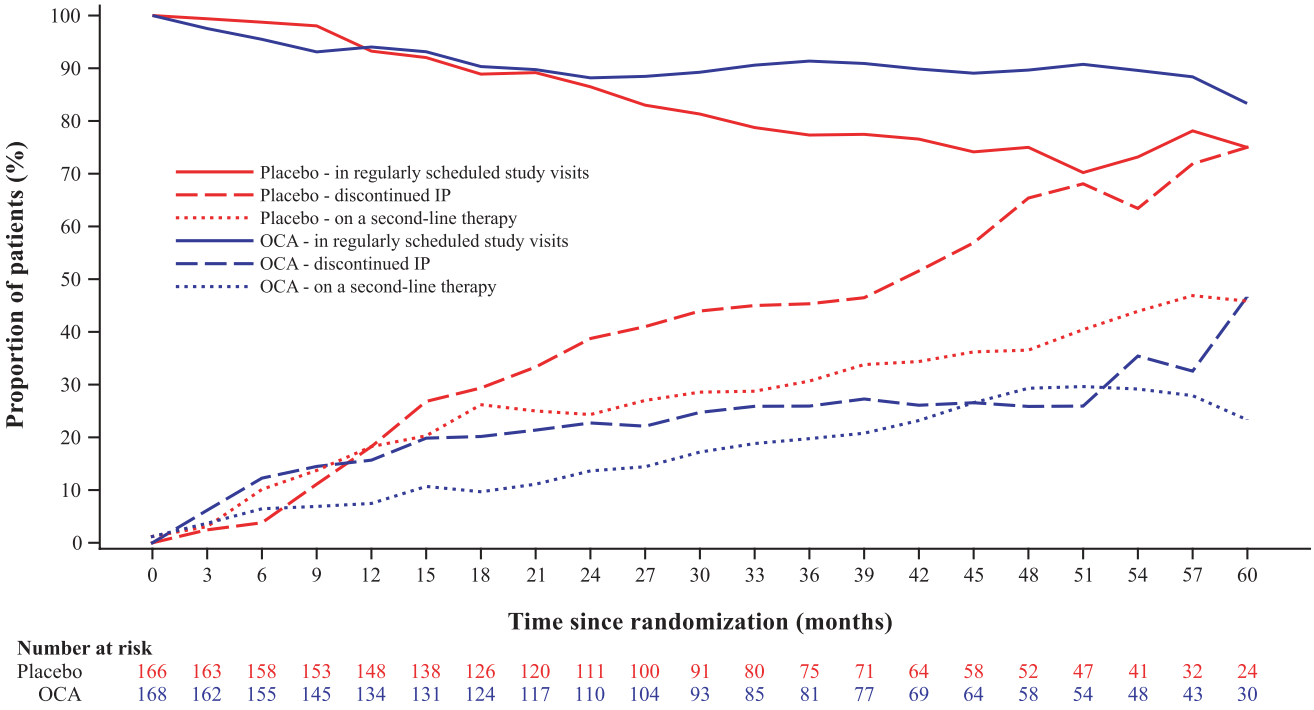


Figure 4. Patients at risk of a primary composite endpoint event during the COBALT RCT. The proportion of placebo patients who had regularly scheduled study visits (solid lines) steadily decreased >12 months after randomization and was less than that in OCA patients during the remainder of the study. Patients who did not have follow-up study visits were followed by telephone calls and medical record review (i.e., proportion = 1 − proportion who had study visits), as described in the main text. The proportion of patients who discontinued IP (dashed lines) and initiated second-line therapy (dotted lines) steadily increased >12 months since randomization and was higher in placebo vs OCA patients during the remainder of the study. IP, investigational product; OCA, obeticholic acid; RCT, randomized controlled trial.

management for patients with PBC, we suspected that observable ALP levels may have compromised the uncertainty principle of double-blinded, randomized clinical trials (i.e., neither patients nor their physicians can be certain about which trial arm is likely to benefit most) (21,22). During the first 24 months, especially among the placebo group, patients who discontinued study visits had higher levels of ALP, and many initiated commercial PBC therapies, leading to lower levels of ALP. Spontaneous reduction of ALP levels in the absence of active treatment is counter to the natural history of PBC and was not observed in the placebo arm of the POISE trial or other trials of investigational therapeutics for PBC (e.g., seladelpar and elafibranor) but was uncharacteristically observed in the RCT

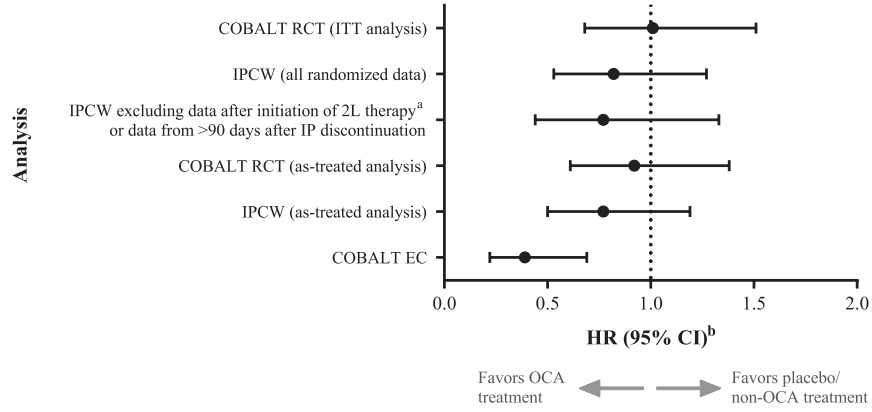


Figure 5. Hazard analyses of the composite endpoint in COBALT data sets. Forest plot comparing hazard ratios for the composite endpoint among patients treated with OCA vs placebo in the ITT analysis of the RCT, IPCW analyses to adjust for potential sources of bias in the RCT, the as-treated analysis of the RCT, and the real-world EC analysis. 2L, second line; EC, external control; HR, hazard ratio; IP, investigational product; IPCW, inverse probability of censoring weights; ITT, intent-to-treat; OCA, obeticholic acid; RCT, randomized controlled trial; UDCA, ursodeoxycholic acid. [a]Defined as initiation of commercial OCA, UDCA for patients who were not receiving UDCA at baseline, or fibrates for patients who were not administered fibrates at baseline. [b]Ratio of the hazard of OCA treatment to the hazard of placebo treatment.

Table 2. Incidence of treatment-emergent adverse events occurring in ≥10% of patients in the COBALT randomized placebo-controlled trial

Preferred term	Placebo, n (%) (n = 166)	OCA, n (%) (n = 168)
Pruritus	85 (51.2)	132 (78.6)
Edema, peripheral	18 (10.8)	31 (18.5)
Urinary tract infection	30 (18.1)	20 (11.9)
Arthralgia	29 (17.5)	23 (13.7)
Varices, esophageal	28 (16.9)	20 (11.9)
Diarrhea	26 (15.7)	21 (12.5)
Fatigue	25 (15.1)	18 (10.7)
Blood bilirubin increased	25 (15.1)	20 (11.9)
Nausea	21 (12.7)	25 (14.9)
Abdominal pain and upper	12 (7.2)	25 (14.9)
Headache	21 (12.7)	23 (13.7)
Ascites	21 (12.7)	18 (10.7)
Abdominal pain	18 (10.8)	21 (12.5)
Constipation	10 (6.0)	19 (11.3)
Nasopharyngitis	14 (8.4)	18 (10.7)
Serious adverse event	53 (31.9)	53 (31.5)

OCA, obeticholic acid.

(9,23–25). This suggested that differential dropout in the placebo arm and treatment crossover to commercial OCA biased the study results toward the null (because patients who discontinued study visits or initiated commercial therapies were those with higher observed ALP) and confounded the ITT estimate of effect (26). It is plausible that these severe patients who were no longer eligible for OCA per the May 2021 label update were at high risk of having observable excursions of ALP. The ITT analysis may have also been confounded by fact that placebo patients had nearly 50% more follow-up time outside of regular study visits than OCA patients, which may have reduced the amount of detailed information that was captured.

The IPCW sensitivity analyses and as-treated analysis, which were performed to address the confounding in the ITT analysis of the RCT, were directionally consistent with the real-world–based EC analysis to favor OCA over placebo. Specifically, the sensitivity analyses progressively decreased the HR point estimate from the ITT analysis by approximately 25% overall (1.01–0.77) in favor of OCA. These analyses adjusted the effect estimate for the primary composite endpoint in favor of OCA, similar to observations from other real-world studies of the benefits of OCA on clinical outcomes, and brought the effect estimate closer to the statistically significant and clinically meaningful reduction observed in the COBALT EC (HR, 0.39; 95% CI, 0.22–0.69; $P = 0.001$) (16,27).

Externally controlled trials have been conducted for other rare diseases and are recognized by the US FDA as adequate and well-controlled clinical investigations for providing evidence of therapeutic effectiveness (15,28). The COBALT EC result is consistent with findings from several real-world–based studies in different populations, methodologies, and geographies. The long-term safety extension of the POISE trial, which compared OCA patients and ECs in the Global PBC registry, showed a reduction in event-free survival (HR, 0.42; 95% CI, 0.21–0.85) (16). Similarly, the RECAPITULATE study of OCA-treated patients and an EC in the Italian PBC registry demonstrated a significant improvement in both liver transplant-free survival (HR, 0.32; 95% CI, 0.15–0.66; $P < 0.0001$) and event-free survival (HR, 0.33; 95% CI, 0.20–0.54; $P < 0.0001$) (29).

The incidence of TEAEs in the COBALT study was consistent with the known safety profile of OCA (9,13). Of note, hepatic TEAEs occurred in a smaller proportion of OCA patients than in placebo patients. Approximately 10% of patients in the OCA arm discontinued IP because of pruritus during follow-up.

The COBALT study experience demonstrates the necessity of prespecifying the inclusion of ECs in confirmatory trials when traditional trial design may not be feasible, especially for rare, slowly progressive diseases. Rigorous initiatives to replicate clinical trial findings using observational data such as RCT-DUPLICATE have demonstrated that well-designed and well-executed real-world–based studies can produce valid results and reach similar conclusions as RCTs (20,30).

Our study has several limitations that are common in real-world data sources and EC studies (15,31). First, real-world claims data reflect routine patient encounters with healthcare professionals, which may differ from those of patients enrolled in an RCT (32). We mitigated this limitation by matching patient characteristics and markers of disease progression to minimize potential baseline differences between the 2 arms. Second, event rates in clinical trials and real-world data can differ owing to the quality of data capture, although the observed consistency in event rates between randomized OCA patients and real-world patients treated with OCA in the Komodo database lends confidence in the estimated treatment effect. Third, healthcare claims data may lack clinical details, such as MELD score components in the Komodo Healthcare Map database. However, the large size and broad geographic coverage of the Komodo database coupled with nationwide laboratory data from LabCorp and Quest Diagnostics increases the confidence that key biochemical markers of disease progression and the objective clinical outcomes in this study (e.g., death, liver transplantation, and hospitalization for decompensation) were captured accurately and that the study results can be generalized.

The ITT analysis of the COBALT RCT was compromised by loss of equipoise because of functional unblinding, initiation of commercial OCA in placebo patients, and differences in the follow-up between treatment arms (notably, increased study discontinuation in the placebo arm). As a result, the RCT ITT analysis could not adequately assess the primary objective to demonstrate a difference in clinical outcomes between OCA and non-OCA patients. Adjusting for potential confounders with IPCW and as-treated analyses demonstrated evidence for bias in the RCT and shifted the HR in favor of OCA, in alignment with the totality of evidence from real-world–based studies that have shown improved clinical outcomes among patients with PBC who are treated with OCA. The COBALT EC analysis showed that OCA treatment is associated with a significant and clinically meaningful reduction in the likelihood of serious and life-threatening outcome events.

After receiving accelerated approval based on a surrogate endpoint, the US FDA requires confirmation of clinical benefit with an adequate and well-controlled trial (33). The use of real-

world evidence for confirmatory trials in PBC merits further consideration because conducting randomized, placebo-controlled trials to determine long-term clinical outcomes is challenging and often not feasible (22). EC studies, such as the COBALT EC analysis, POISE EC (16), and the fully real-world HEROES study (27), are better alternatives to RCTs, replicating the improved event-free and transplant-free survival among patients treated with OCA. The value of real-world evidence in confirming RCTs is not without precedent in the development of treatments for PBC, as illustrated by the real-world confirmation of the clinical benefits of first-line UDCA (22). The data presented here have demonstrated that timely monitoring and intervention with OCA for appropriate patients with PBC is critical to reduce the risk of hepatic decompensation, liver transplant, and death.

ACKNOWLEDGEMENTS

The authors thank the patients who participated in the study. Alexander Simon and Katherine O'Connor, PhD, of MedLogix Communications, LLC, a Citrus Health Group, Inc., company (Chicago, IL) provided medical writing and editorial support, respectively, which were funded by Intercept Pharmaceuticals.

CONFLICTS OF INTEREST

Guarantor of the article: Leona Bessonova, PhD, MBA.

Specific author contributions: All authors participated in the development of the manuscript, approved the submitted versions, and confirmed the accuracy and completeness of the data and the fidelity of the trial to the protocol. K.V.K., G.M.H., C.C., E.S.M., J.L., N.R., G.M., D.E.J., P.J.T., B.E.H., R.S., J.W., S.H., D.N.K., A.L.M., C.L.B., K.M., M.C., M.B., P.M., F.A., A.V.: conceived and designed the data. K.V.K., G.M.H., G.M., D.E.J., P.J.T., A.L.M., C.L.B., K.M., M.C., M.B., P.M., A.V.: collected the data. C.C., E.S.M.: performed the analysis. K.V.K., G.M.H., C.C., E.S.M., J.L., N.R., G.M., D.E.J., P.J.T., B.E.H., R.S., J.W., S.H., D.N.K.: reviewed the analysis. K.V.K., G.M.H., C.C., E.S.M., J.L., N.R., G.M., D.E.J., P.J.T., B.E.H., R.S., J.W., S.H., D.N.K., A.L.M., C.L.B., K.M., M.C., M.B., P.M., F.A., A.V.: wrote, reviewed, and edited the manuscript.

Financial support: This study was funded by Intercept Pharmaceuticals.

Potential competing interests: K.V.K. has received research support from CymaBay; grants and/or contracts from 89Bio, Genfit, Gilead, GSK, Hanmi, HighTide, Intercept, Madrigal, Mirum, NGM, Pfizer, Pliant, and Viking; royalties/licenses from UpToDate; consulting fees from 89Bio, Calliditas, CymaBay, Genfit, Gilead, Inipharm, Intercept, Madrigal, Mirum, NGM, and Pliant; payment/honoraria for lectures, presentations, speaker bureaus, manuscript writing, or educational events from AbbVie, Gilead, and Intercept; payment for expert testimony from the Department of Justice. He has also participated in a data safety monitoring board or advisory board for CTI and Medpace, has stock or options in Inipharm, and has received equipment from Sonic Incytes. G.M.H. serves as a consultant for CymaBay, Escient, Gilead, GSK, HighTide, Intercept, Ipsen, Mirum, and Pliant. C.C. is a former employee of Syneos Health and may hold stock in the company. E.S.M., L.B., and J.L. are employees of Intercept. N.R. is an employee and shareholder of Target RWE. G.M. has received research funding from Intercept. D.E.J. has received grant funding, speaker fees, and consultancy from Intercept; speaker fees from Falk and Ipsen; consultancy fees from Advanz; and consultancy and speaker fees from Abbott and GSK. P.J.T. receives institutional salary support from the National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre (BRC). This article presents independent research supported by the Birmingham NIHR BRC based at the University Hospitals Birmingham National Health Service Foundation Trust and the University of Birmingham. The views expressed are those of the author(s) and not necessarily those of the National Health Service, the NIHR, or the Department of Health. P.J.T. has received grant support from Bristol Myers Squibb, Falk, Gilead GSK, Guts UK, Intercept, LifeArc, Medical Research Foundation, PSC Support, and Wellcome Trust; speaker fees from Albireo, Falk, and Intercept; and has been an advisory board member/consultant for ChemoMab, CymaBay, Falk, GSK, Intercept, and Pliant. B.E.H. has received grants from Albireo, Calliditas, CymaBay, Eiger, Gilead, Intercept, and Mirum. She also serves as a consultant for Albireo, Calliditas, CymaBay, Eiger, Enyo, Intercept, Ipsen, Mirum, and Pliant. J.W. has had research funded by Intercept. S.H. has received research support from Intercept. A.L.M. has received research funding and medications from Merck. C.L.B. has received grant/research support from Boston Scientific, Bristol Myers Squibb, Calliditas, Cara, Cour, CymaBay, Genfit, Gilead, GSK, Intercept, Novo Nordisk, Pliant, and Viking. K.M. serves as an advisor for Chiesi. M.C. has received grant/research support from Genetic Spa and Intercept and served on an advisory board and/or been a consultant (paid or unpaid) for Advanz, Albireo, AMAF, Calliditas, CymaBay, Echosens, EpaC, Falk, Genetic Spa, GSK, Ipsen, Kowa, Mayoly, Mirum, Perspectum, PSC Pediatric Foundation, and Zydus. M.B. has received grant/research support (to institution) from AbbVie, Gilead, and Intercept and served on a paid advisory board for AbbVie, Advanz, Alexion, Chiesi, Deep Genomics, Gilead, and Orphalan. P.M. has received speaker fees from Alfa Wasserman, Astellas, Chiesi, and Polfarma and served as a consultant for Orphalan. F.A. is a former employee of Intercept and holds stock options in the company. A.V. serves as a speaker for Intercept. R.S. and D.N.K. have no disclosures to report.

Trial registration: COBALT ClinicalTrials.gov identifier NCT02308111.

Study Highlights

WHAT IS KNOWN

- ✓ Approximately 40% of patients with primary biliary cholangitis are ursodeoxycholic acid–intolerant or inadequate responders.
- ✓ Obeticholic acid (OCA) received accelerated approval as a second-line treatment of primary biliary cholangitis in 2016.
- ✓ The COBALT trial compared outcomes of treatment vs placebo or external control.

WHAT IS NEW HERE

- ✓ In the randomized controlled trial, the primary endpoint was not different between the OCA and placebo groups (intent-to-treat analysis hazard ratio = 1.01).
- ✓ Functional unblinding and treatment crossover biased the intent-to-treat analysis of the randomized controlled trial.
- ✓ The real-world external control analysis showed that OCA significantly reduced the risk of negative outcomes.

REFERENCES

1. Trivella J, John BV, Levy C. Primary biliary cholangitis: Epidemiology, prognosis, and treatment. Hepatol Commun 2023;7(6):e0179–16.
2. Lu M, Li J, Haller IV, et al. Factors associated with prevalence and treatment of primary biliary cholangitis in United States health systems. Clin Gastroenterol Hepatol 2018;16(8):1333–41.e6.
3. Kim WR, Lindor KD, Locke GR III, et al. Epidemiology and natural history of primary biliary cirrhosis in a US community. Gastroenterology 2000;119:1631–6.
4. Hurlburt KJ, McMahon BJ, Deubner H, et al. Prevalence of autoimmune liver disease in Alaska Natives. Am J Gastroenterol 2002;97(9):2402–7.
5. Younossi ZM, Stepanova M, Golabi P, et al. Factors associated with potential progressive course of primary biliary cholangitis: Data from real-world US database. J Clin Gastroenterol 2019;53(9):693–8.
6. Buchanan-Peart K-A, MacEwan JP, Levine A, et al. United States prevalence of diagnosed primary biliary cholangitis: 41 per 100,000 adults with wide regional variability. Hepatology 2023;78:S2088.
7. Silveira MG, Brunt EM, Heathcote J, et al. American Association for the Study of Liver Diseases endpoints conference: Design and endpoints for clinical trials in primary biliary cirrhosis. Hepatology 2010;52(1): 349–59.
8. Poupon RE, Lindor KD, Cauch-Dudek K, et al. Combined analysis of randomized controlled trials of ursodeoxycholic acid in primary biliary cirrhosis. Gastroenterology 1997;113(3):884–90.
9. Nevens F, Andreone P, Mazzella G, et al. A placebo-controlled trial of obeticholic acid in primary biliary cholangitis. N Engl J Med 2016;375(7): 631–43.
10. United States Food and Drug Administration. FDA approves Ocaliva for rare, chronic liver disease [media release] (https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm503964.htm) (2016). Accessed July 2, 2024.
11. Nevens F, Shiffman ML, Drenth JPH, et al. Durable response in the markers of cholestasis through 5 years of open-label extension study of obeticholic acid in primary biliary cholangitis. Dig Liver Dis 2020; 52:e30.
12. United States Food and Drug Administration. Due to risk of serious liver injury, FDA restricts use of Ocaliva (obeticholic acid) in primary biliary cholangitis (PBC) patients with advanced cirrhosis (https://www.fda.gov/drugs/drug-safety-and-availability/due-risk-serious-liver-injury-fda-restricts-use-ocaliva-obeticholic-acid-primary-biliary-cholangitis) (2021). Accessed July 2, 2024.
13. Intercept Pharmaceuticals. Ocaliva® [prescribing information]. https://www.interceptpharma.com/wp-content/uploads/2022/06/Ocaliva_USPI_Clean_VV-REG-040954.pdf (2022, Accessed January 12, 2024).
14. Khachatryan A, Read SH, Madison T. External control arms for rare diseases: Building a body of supporting evidence. J Pharmacokinet Pharmacodyn 2023;50(6):501–6.
15. Seeger JD, Davis KJ, Iannacone MR, et al. Methods for external control groups for single arm trials or long-term uncontrolled extensions to randomized clinical trials. Pharmacoepidemiol Drug Saf 2020;29(11): 1382–92.
16. Murillo Perez CF, Fisher H, Hiu S, et al. Greater transplant-free survival in patients receiving obeticholic acid for primary biliary cholangitis in a clinical trial setting compared to real-world external controls. Gastroenterology 2022;163(6):1630–42.e3.
17. Trauner M, Nevens F, Shiffman ML, et al. Long-term efficacy and safety of obeticholic acid for patients with primary biliary cholangitis: 3-year results of an international open-label extension study. Lancet Gastroenterol Hepatol 2019;4(6):445–53.
18. Bernstam EV, Applegate RJ, Yu A, et al. Real-world matching performance of deidentified record-linking tokens. Appl Clin Inform 2022;13(4):865–73.
19. Brookhart MA, Wyss R, Layton JB, et al. Propensity score methods for confounding control in nonexperimental research. Circ Cardiovasc Qual Outcomes 2013;6(5):604–11.
20. Wang SV, Schneeweiss S, Franklin JM, et al; RCT-DUPLICATE Initiative. Emulation of randomized clinical trials with nonrandomized database analyses: Results of 32 clinical trials. JAMA 2023;329(16):1376–85.
21. Fries JF, Krishnan E. Equipoise, design bias, and randomized controlled trials: The elusive ethics of new drug development. Arthritis Res Ther 2004;6(3):R250–5.
22. Jones DE, Beuers U, Bonder A, et al. Primary biliary cholangitis drug evaluation and regulatory approval: Where do we go from here? Hepatology 2024 (doi: 10.1097/HEP.0000000000000864).
23. Gatselis NK, Goet JC, Zachou K, et al. Factors associated with progression and outcomes of early stage primary biliary cholangitis. Clin Gastroenterol Hepatol 2020;18(3):684–92.e6.
24. Hirschfield GM, Shiffman ML, Gulamhusein A, et al. Seladelpar efficacy and safety at 3 months in patients with primary biliary cholangitis: ENHANCE, a phase 3, randomized, placebo-controlled study. Hepatology 2023;78(2):397–415.
25. Kowdley KV, Bowlus CL, Levy C, et al. Efficacy and safety of elafibranor in primary biliary cholangitis. N Engl J Med 2024;390(9):795–805.
26. Rimawi M, Hilsenbeck SG. Making sense of clinical trial data: Is inverse probability of censoring weighted analysis the answer to crossover bias? J Clin Oncol 2012;30(4):453–8.
27. Brookhart MA, Coombs C, Breskin A, et al. Results of the HEROES study: Treatment efficacy of obeticholic acid on hepatic real-world outcomes in patients with primary biliary cholangitis. Hepatology 2022;76:S195.
28. United States Food and Drug Administration. Considerations for the design and conduct of externally controlled trials for drug and biological products: Guidance for industry (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-design-and-conduct-externally-controlled-trials-drug-and-biological-products) (2023). Accessed January 12, 2024.
29. Terracciani F, De Vincentis A, D'Amato D, et al. Long-term results from the Italian real-world experience on obeticholic acid treatment in primary biliary cholangitis: The RECAPITULATE study. Dig Liver Dis 2023;55: S44–S45.
30. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. Cochrane Database Syst Rev 2014;2014:MR000034.
31. Collins R, Bowman L, Landray M, et al. The magic of randomization versus the myth of real-world evidence. N Engl J Med 2020;382(7):674–8.
32. Chodankar D. Introduction to real-world evidence studies. Perspect Clin Res 2021;12(3):171–4.
33. Code of Federal Regulations. Accelerated approval of new drugs for serious or life-threatening illnesses. CFR Title 21, Part 314, Subpart H. Department of Health and Human Services. 1992. Edition: June 2024 (https://www.ecfr.gov/current/title-21/chapter-I/subchapter-D/part-314/subpart-H). Accessed July 2, 2024.