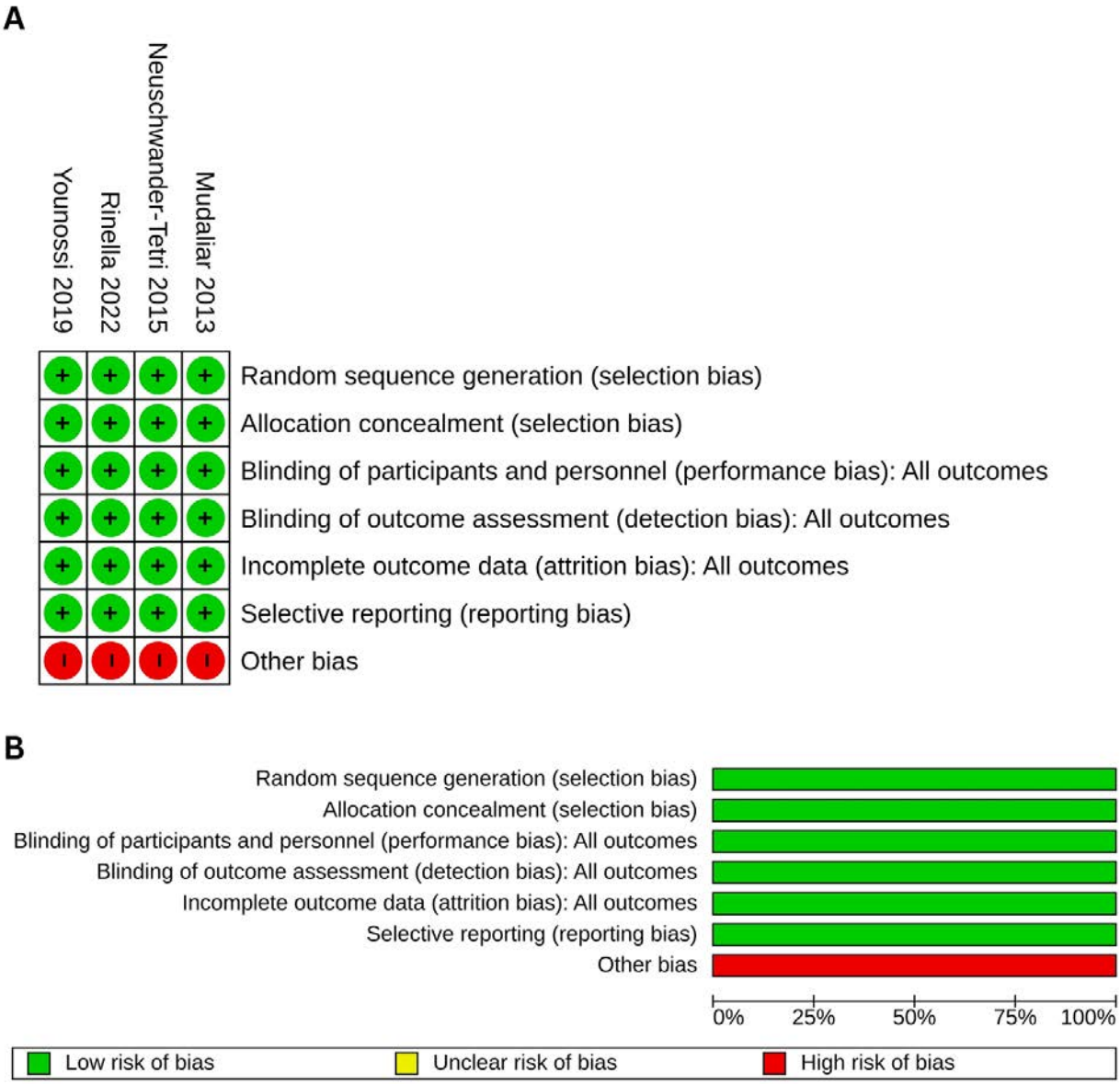
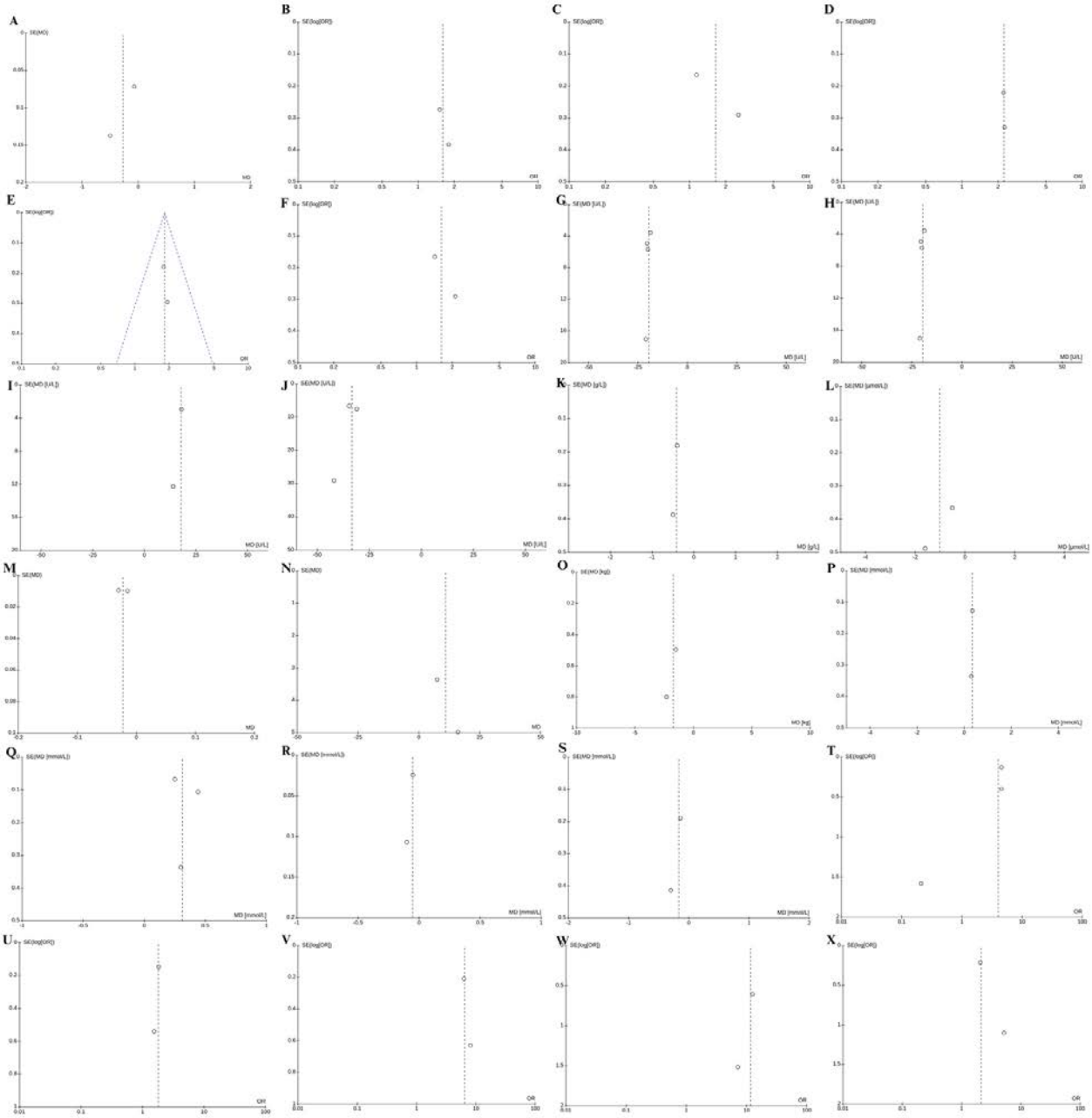
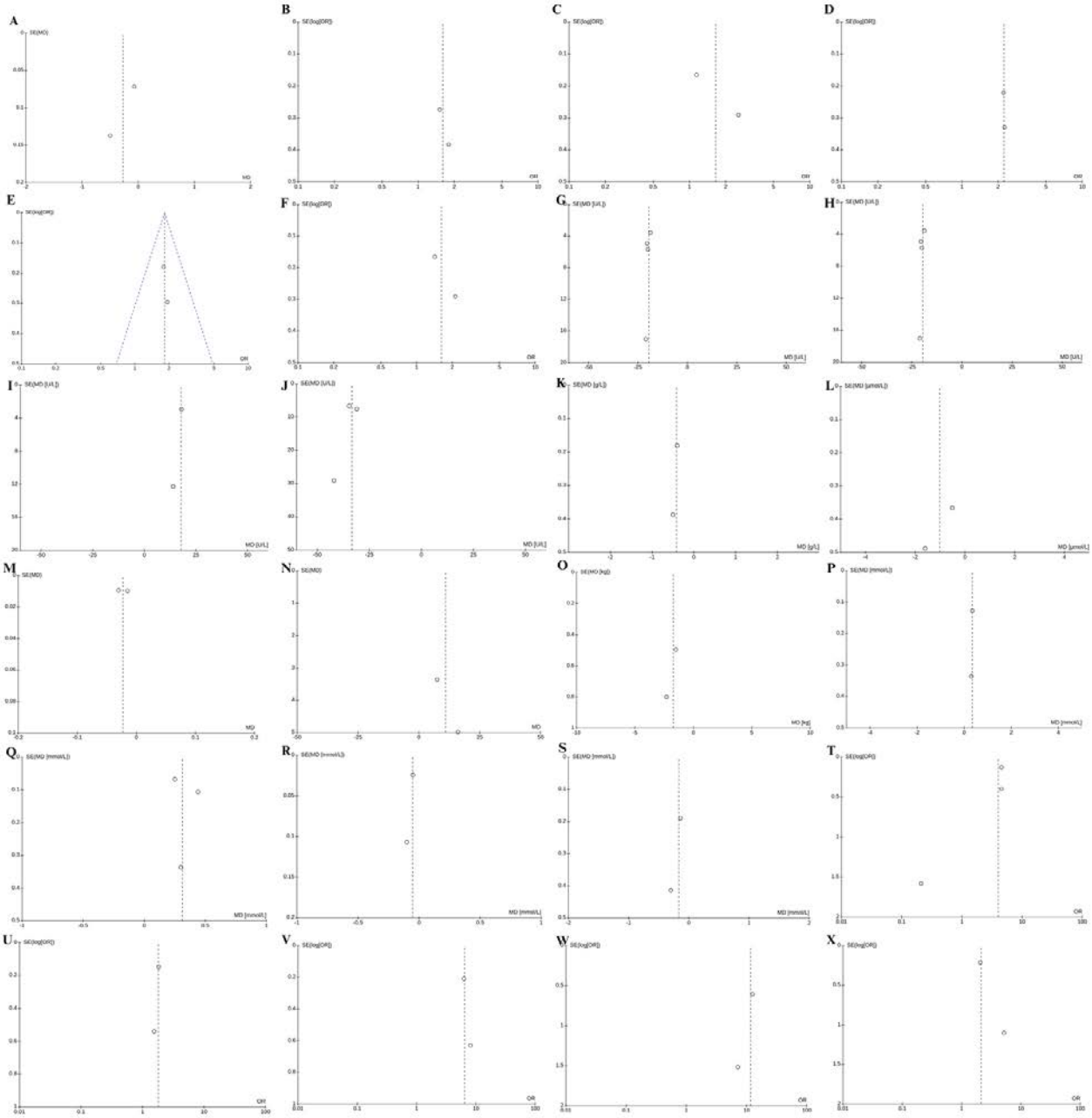




This supplementary material was peer reviewed with the manuscript but has not been subject to our in-house editing or production processes.



This supplementary material was peer reviewed with the manuscript but has not been subject to our in-house editing or production processes.

SUPPLEMENTARY TABLES

Supplementary Table S1. Summary of the randomized-controlled trials included in the meta-analysis

Trial registration no. and Trial name	Study ID	Study participants		Study arms	N	Age (y)	Male/ Female	Outcomes studied	Study duration
		Major traits	Inclusion criteria						
NCT00501592	Mudaliar 2013 [16]	NAFLD plus T2DM	≥1 of the following criteria: ALT 47 U/L for female and 56 U/L for male, AST 47 U/L for female and 60 U/L for male, enlarged liver in imaging and diagnostic histologic findings shown on prior biopsy (in the prior 5 years)	Placebo	23	53.1±12.1	10/13	**Primary:** glucose infusion rate **Secondary:** serum FGF19, C4, bile acids, caspase-cleaved keratin-18 levels, ELF score	6 weeks
				OCA 25 mg	20	52.2±8.7	14/6		
				OCA 50 mg	21	50.5±10.8	9/12		
NCT01265498 FLINT trial	Neuschwander-Tetri 2015 [17]	Biopsy-evidenced NASH	≥18 years at the time of screening, histological evidence of definite or borderline NASH based upon a liver biopsy obtained ≤90 days	Placebo	142	51±12	53/89	**Primary:** improvement in centrally scored liver histology defined as a decrease in NAFLD activity score by at least	72 weeks
				OCA 25 mg	141	52±11	43/98		

			before randomization, and a histological NAFLD activity score of ≥4 with a score of ≥1 in each component of the score					2 points without worsening of fibrosis from baseline to the end of treatment **Secondary:** resolution of non-alcoholic steatohepatitis, change in NAFLD activity score, and changes in the individual scores for hepatocellular ballooning, steatosis, lobular and portal inflammation, and fibrosis	
NCT02548351 REGENERATE Trial (non-invasive evaluation)	Rinella 2022 [18]	NASH	≥18 years of age with histologic evidence of steatohepatitis; NAFLD activity score ≥4 (including ≥1 point for	Placebo	311	55±12	124/187	**Exploratory end-points:** change in non-invasive tests results of liver fibrosis	18 months
				OCA 10 mg	312	55±11	135/177		
				OCA 25 mg	308	55±11	133/175		

			steatosis, lobular inflammation, and hepatocellular ballooning); fibrosis stage F2 or F3 per NASH Clinical Research Network criteria or F1 with ≥1 comorbidity (obesity, type 2 diabetes mellitus, or ALT >1.5x the upper limit of normal)						
NCT02548351 REGENERATE trial (invasive evaluation)	Younossi 2019 [19]	Similar to Rinella 2022 [18]	Similar to Rinella 2022 [18]	Placebo	311	55±12	124/187	**Primary endpoints:** improvement in fibrosis with no worsening of NASH, or NASH resolution with no worsening of fibrosis. **Secondary endpoints:** improvement of	18 months
				OCA 10 mg	312	55±11	135/177		
				OCA 25 mg	308	55±11	133/175		

								fibrosis by at least one stage or resolution of NASH, or both, without worsening of either, histological improvement of features of NASH as well as NAS, and liver biochemistry.	

NAFLD = Non-alcoholic fatty liver disease, T2DM = Type 2 Diabetes Mellitus, MAFLD = Metabolic dysfunction-associated fatty liver disease, OCA = Obeticholic acid, FGF19 = Fibroblast growth factor 19, C4 = 7alphahydroxy-4-cholesten-3-one, NASH = Non-alcoholic steatohepatitis, ALT = Alanine Transaminase, AST = Aspartate Transaminase

Supplementary Table S2. Summary of the trials excluded from the meta-analysis

Trial Reg. no.	Authors	Type of study/analysis	Sample size and Study arms	Primary endpoint
NCT01265498	Hameed et al. [20]	Sub-analysis of the FLINT trial	200 patients of the FLINT trial having baseline and end-of-treatment liver biopsies	• Weight loss occurred in 44% of OCA and 32% of placebo-treated patients (p = 0.08). • NAS improved more in those with than without weight loss in both the OCA group (-2.4 vs. -1.2, p<0.001) and the placebo group (-1.2 vs. -0.5, p 0.03). • ALT levels also improved in those with vs without weight loss in the OCA group (- 43 vs. - 34 U/L, p = 0.12) and placebo group (-29 vs. -10 U/L, p = 0.02) • Among those with weight loss, OCA was associated with opposite effects from placebo. Changes in ALP (+21 vs. -12 U/L, p<0.001), LDL cholesterol (+18 vs. -12 mg/dl, p = 0.01), and Hba1c (-01% vs. -0.4%, p = 0.01)
NCT01265498	Siddiqui et al. [21]	Sub-analysis of the FLINT trial	196 patients (99 OCA group and 97 placebo group) were enrolled in the FLINT trial and had samples available for lipid analysis and liver biopsies at enrolment and end-of-treatment at 72 weeks.	• ↓Large VLDL particle concentration at 12 weeks (baseline-adjusted mean: 6.8 vs. 8.9 nmol/L; p = 0.002). • ↑Small VLDL particle concentration (33.9 vs. 28.0 nmol/L; p = 0.02). • ↑Total LDL in the OCA group vs placebo (1,667 vs. 1,329 nmol/L; p < 0.0001). • ↑Both less atherogenic, large-buoyant LDL (475 vs. 308 nmol/L; p < 0.001) and more atherogenic small-dense LDL particles (1,015 vs. 872 nmol/L; p = 0.002). • Similar levels of LDL concentrations in OCA and placebo groups at 24 weeks likely d/t improvement in the OCA cohort

				• ↓ HDL in the OCA group compared to placebo → resolved after drug discontinuation.
NCT02548351	Younossi et al. [22]	Sub-analysis of REGENERATE: 18-Month Interim Analysis	1218 patients assigned randomly to 10 mg (n = 407) or 25 mg (N = 404) OCA or placebo (N = 407)	• Nineteen (1.6%) patients discontinued therapy (protocol mandated) because of grade 3 pruritus. • There is no difference in patient-reported outcomes (PROs) evaluating HRQoL assessed using Chronic Liver Disease Questionnaire–NASH and EuroQol EQ-5D-5L between OCA and placebo at baseline. • Patients receiving 25 mg OCA experienced mild worsening of itch scores primarily in the first months of treatment (mean change from baseline –0.66, –0.44, and –0.42 at 6, 12, and 18 months, respectively (all P < .01). • No worsening of other PROs with OCAs at any point. • Patients experiencing fibrosis improvement, NAS decrease by >= 2 points, or NASH resolution had greater PRO improvements in some domains.

NASH = Non-alcoholic steatohepatitis, NAFLD = Non-alcoholic fatty liver disease, OCA = Obeticholic acid, DM = Diabetes Mellitus, Alt = Alanine Transaminase, AST = Aspartate Transaminase, PRO = Patient reported outcomes, HRQoL = Health Related Quality of Life, NAS = NAFLD activity score

Supplementary Table S3. Risk of bias assessment

Mudaliar 2013 [16]	**Risk of bias**	**Author judgment**
Random sequence generation (selection bias)	Low risk	This was a multicenter, double-blind, randomized, placebo-controlled, multiple-dose, parallel-group exploratory study to evaluate the safety and efficacy of OCA
Allocation concealment (selection bias)	Low risk	Patients who met all inclusion and exclusion criteria were randomly assigned to receive 25 mg OCA, 50 mg OCA, or matching placebo orally once daily
Blinding of participants & personnel (performance bias)	Low risk	Double-blind RCT.
Blinding of outcome assessment (detection bias)	Low risk	Double-blind RCT. Insufficient information about blinding of outcome assessment, but the review authors judge that the outcome measurement is not likely to be influenced by lack of blinding.
Incomplete outcome data (attrition bias)	Low risk	~88% completed the trial.
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes are reported.
Other biases	High risk	Supported by a research grant from Intercept Pharmaceuticals, inc.
Neuschwander-Tetri 2016 [17]	**Risk of bias**	**Author judgment**
Random sequence generation (selection bias)	Low risk	The study randomly assigned (1:1) patients meeting eligibility criteria to oral OCA, 25 mg once daily or placebo using a computer-generated, centrally administered procedure, stratified by the clinical center and diabetes status and blocked by calendar date
Allocation concealment (selection bias)	Low risk	Treatment was assigned centrally using a web-based application
Blinding of participants & personnel (performance bias)	Low risk	Double-blind, placebo-controlled, parallel-group, randomized clinical trial
Blinding of outcome assessment (detection bias)	Low risk	Double-blind, placebo-controlled, parallel-group, randomized clinical trial. There is insufficient information about blinding of outcome assessment, but the review authors judge that the outcome measurement is not likely to be influenced by lack of blinding.

Incomplete outcome data (attrition bias)	Low risk	Of 283 patients, randomized, primary intention-to-treat analysis was done on 219 patients (22.6% attrition)
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes are reported.
Other biases	High risk	Partial funding for the trial, OCA and an identical placebo were provided by Intercept Pharmaceuticals under a collaborative research and development agreement
Rinella 2022 [18]	**Risk of bias**	**Author judgment**
Random sequence generation (selection bias)	Low risk	Randomization was based on a predefined randomization code generated by electronic data capture and done using an interactive web response system;
Allocation concealment (selection bias)	Low risk	Patients with NASH and fibrosis stage F2 or F3 (n = 931) were randomized (1:1:1) to receive a placebo, OCA 10 mg, or OCA 25 mg once daily.
Blinding of participants & personnel (performance bias)	Low risk	Double-blind RCT.
Blinding of outcome assessment (detection bias)	Low risk	Double-blind RCT. Insufficient information about blinding of outcome assessment, but the review authors judge that the outcome measurement is not likely to be influenced by lack of blinding.
Incomplete outcome data (attrition bias)	Low risk	765 of 931 patients completed the study (17.8% attrition)
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes are reported.
Other biases	High risk	The regenerate study was funded by Intercept Pharmaceuticals
Yonossi 2019 [19]	**Risk of bias**	**Author judgment**
Random sequence generation (selection bias)	Low risk	Randomization was based on a predefined randomization code generated by electronic data capture and done using an interactive web response system;
Allocation concealment (selection bias)	Low risk	Eligible patients were randomly assigned in a 1:1:1 ratio to receive a daily placebo, OCA 10 mg or OCA 25 mg orally
Blinding of participants & personnel (performance bias)	Low risk	Double-blind RCT.

Blinding of outcome assessment (detection bias)	Low risk	Double-blind RCT. Insufficient information about blinding of outcome assessment, but the review authors judge that the outcome measurement is not likely to be influenced by lack of blinding.
Incomplete outcome data (attrition bias)	Low risk	At the time of the interim analysis, 73 (23%) patients in the placebo group, 71 (23%) in the obeticholic acid 10 mg group, and 77 (25%) in the obeticholic acid 25 mg group had discontinued treatment
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes are reported.
Other biases	High risk	The regenerate study was designed in collaboration with the funder, Intercept Pharmaceuticals, which was involved in data collection, analysis, and interpretation

Supplementary Table S4. The summary of findings table

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	No. of participants (studies)	Certainty of the evidence (GRADE)
	Risk with PL	Risk with OCA-25			
ELF score	The mean ELF score was **9.64**	MD **0.27 lower** (0.69 lower to 0.15 higher)	-	662 (2 RCTs)	⊕⊕⊕⊕ High
Resolution of NASH	93 per 1,000	**141 per 1,000** (96 to 203)	**OR 1.60** (1.04 to 2.48)	819 (2 RCTs)	⊕⊕⊕⊕ High
Improvement of steatosis	379 per 1,000	**502 per 1,000** (317 to 688)	**OR 1.65** (0.76 to 3.61)	819 (2 RCTs)	⊕⊕◯◯ Low[a]
Improvement of fibrosis	137 per 1,000	**261 per 1,000** (198 to 337)	**OR 2.23** (1.56 to 3.20)	819 (2 RCTs)	⊕⊕⊕⊕ High
Improvement of hepatocellular ballooning	249 per 1,000	**378 per 1,000** (310 to 451)	**OR 1.83** (1.35 to 2.47)	819 (2 RCTs)	⊕⊕⊕⊕ High
Improvement of lobular inflammation	355 per 1,000	**471 per 1,000** (383 to 560)	**OR 1.62** (1.13 to 2.32)	819 (2 RCTs)	⊕⊕⊕⊕ High
ALT	The mean ALT was **63.08** U/L	MD **19.47 U/L lower** (24.44 lower to 14.5 lower)	-	1538 (4 RCTs)	⊕⊕⊕◯ Moderate[b]
AST	The mean AST was **48.0** U/L	MD **11.82 U/L lower** (15.32 lower to 8.32 lower)	-	1538 (4 RCTs)	⊕⊕⊕◯ Moderate[b]
ALP	The mean ALP was **75.3** U/L	MD **17.79 U/L higher** (12.25 higher to 23.32 higher)	-	300 (2 RCTs)	⊕⊕⊕⊕ High
GGT	The mean GGT was **78.28** U/L	MD **33.34 U/L lower** (42.92 lower to 23.77 lower)	-	919 (3 RCTs)	⊕⊕⊕◯ Moderate[b]
Albumin	The mean albumin was **4.34** g/L	MD **0.42 g/L lower** (0.74 lower to 0.1 lower)	-	876 (2 RCTs)	⊕⊕⊕⊕ High

Bilirubin	The mean bilirubin was **11.35** µmol/L	MD **1.01 µmol/L lower** (2.07 lower to 0.06 higher)	-	876 (2 RCTs)	⊕⊕⊕◯ Moderate[c]
INR	The mean INR was **1.05**	MD **0.02 lower** (0.04 lower to 0.01 lower)	-	876 (2 RCTs)	⊕⊕⊕⊕ High
Platelets	The mean platelets was **240.6** x109	MD **10.96 x109 higher** (2.77 higher to 19.14 higher)	-	876 (2 RCTs)	⊕⊕⊕◯ Moderate[c]
Body weight	The mean body weight was **95.01** kg	MD **1.72 kg lower** (2.55 lower to 0.9 lower)	-	876 (2 RCTs)	⊕⊕⊕⊕ High
TC	The mean TC was **4.59** mmol/L	MD **0.34 mmol/L higher** (0.11 higher to 0.58 higher)	-	300 (2 RCTs)	⊕⊕⊕⊕ High
LDLC	The mean LDLC was **2.77** mmol/L	MD **0.31 mmol/L higher** (0.18 higher to 0.44 higher)	-	919 (3 RCTs)	⊕⊕⊕◯ Moderate[b]
HDLC	The mean HDLC was **1.11** mmol/L	MD **0.05 mmol/L lower** (0.1 lower to 0.01 lower)	-	300 (2 RCTs)	⊕⊕⊕◯ Moderate[b]
TG	The mean TG was **1.93** mmol/L	MD **0.17 mmol/L lower** (0.51 lower to 0.17 higher)	-	300 (2 RCTs)	⊕⊕⊕⊕ High
Pruritus (any)	163 per 1,000	**439 per 1,000** (296 to 594)	**OR 4.02** (2.16 to 7.50)	1641 (3 RCTs)	⊕⊕⊕◯ Moderate[b]
Pruritus (Grade 1)	120 per 1,000	**198 per 1,000** (158 to 246)	**OR 1.81** (1.37 to 2.39)	1598 (2 RCTs)	⊕⊕⊕◯ Moderate[b]
Pruritus (Grade 2)	41 per 1,000	**217 per 1,000** (158 to 290)	**OR 6.44** (4.37 to 9.49)	1598 (2 RCTs)	⊕⊕⊕◯ Moderate[b]
Pruritus (Grade 3)	4 per 1,000	**42 per 1,000** (14 to 117)	**OR 11.69** (3.89 to 35.07)	1598 (2 RCTs)	⊕⊕⊕⊕ High

Constipation	45 per 1,000	**91 per 1,000** (62 to 131)	**OR 2.12** (1.41 to 3.20)	1641 (3 RCTs)	⊕⊕⊕◯ Moderate[b]

Explanations

a. High heterogeneity among the studies present.

b. The funnel plot is suggestive of the asymmetrical presence of research on each side of the central line; hence, it is likely that significant publication bias is present.

c. Moderate heterogeneity among the studies present.