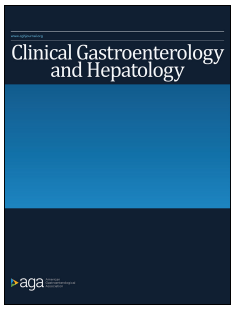
Journal Pre-proof



Development and validation of a scoring system to predict response to obeticholic acid in primary biliary cholangitis.

Antonio De Vincentis, Javier Ampuero, Francesca Terracciani, Daphne D'Amato, Alessio Gerussi, Laura Cristoferi, Nora Cazzagon, Emanuela Bonaiuto, Annarosa Floreani, Vincenza Calvaruso, Luca Cadamuro, Elisabetta Degasperi, Anna Morgando, Ester Vanni, Ana Lleo, Francesca Colapietro, Domenico Alvaro, Antonino Castellaneta, Sara Labanca, Mauro Viganò, Marco Distefano, Valeria Pace Palitti, Chiara Ricci, Nicoletta De Matthaeis, Marco Marzioni, Elena Gómez-Dominguez, Jose-Luis Montero, Esther Molina, Luisa Garcia-Buey, Marta Casado, Marina Berenguer, Isabel Conde, Miguel-Angel Simon, Javier Fuentes, Pedro Costa-Moreira, Guilherme Macedo, Francisco Jorquera, Rosa-Maria Morillas, Jose Presa, Jose-Manuel Sousa, Dario Gomes, Luis Santos, Antonio Olveira, Manuel Hernandez-Guerra, Leire Aburruza, Arsenio Santos, Armando Carvalho, Juan Uriz, Maria-Luisa Gutierrez, Elia Perez, Luchino Chessa, Adriano Pellicelli, Massimo Marignani, Luigi Muratori, Grazia Anna Niro, Maurizia Brunetto, Francesca Romana Ponziani, Maurizio Pompili, Fabio Marra, Andrea Galli, Alessandro Mussetto, Giuliano Alagna, Loredana Simone, Gaetano Bertino, Floriano Rosina, Raffaele Cozzolongo, Maurizio Russello, Leonardo Baiocchi, Carlo Saitta, Natalia Terreni, Teresa Zolfino, Cristina Rigamonti, Raffaella Vigano, Giuseppe Cuccorese, Pietro Pozzoni, Claudio Pedone, Simone Grasso, Antonio Picardi, Pietro Invernizzi, Rodolfo Sacco, Antonio Izzi, Conrado Fernandez-Rodriguez, Umberto Vespasiani-Gentilucci, Marco Carbone

PII: S1542-3565(24)00482-8
DOI: https://doi.org/10.1016/j.cgh.2024.05.008
Reference: YJCGH 59448

To appear in: *Clinical Gastroenterology and Hepatology*
Accepted Date: 7 May 2024

Brunetto M, Ponziani FR, Pompili M, Marra F, Galli A, Mussetto A, Alagna G, Simone L, Bertino G, Rosina F, Cozzolongo R, Russello M, Baiocchi L, Saitta C, Terreni N, Zolfino T, Rigamonti C, Vigano R, Cuccorese G, Pozzoni P, Pedone C, Grasso S, Picardi A, Invernizzi P, Sacco R, Izzi A, Fernandez-Rodriguez C, Vespasiani-Gentilucci U, Carbone M, Development and validation of a scoring system to predict response to obeticholic acid in primary biliary cholangitis., *Clinical Gastroenterology and Hepatology* (2024), doi: https://doi.org/10.1016/j.cgh.2024.05.008.

OCA Response Score (ORS)

To estimate the probability of response of patients with Primary Biliary Cholangitis treated with Obeticholic Acid.

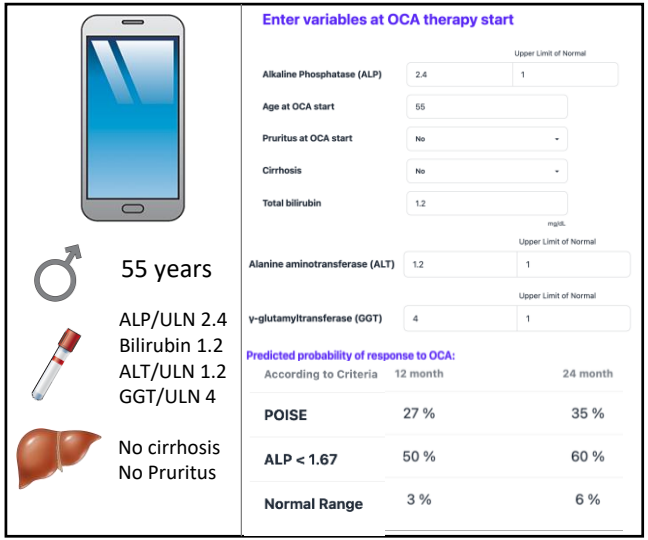
DERIVATION COHORT
RECAPITULATE cohort N 441

COX PROPORTIONAL REGRESSION MODELS
Reduced Model with backward selection (ORS)
ORS + ALP/ULN and bilirubin 6 months (ORS+)

EXTERNAL VALIDATION
IBER-PBC cohort (N 244)

ONLINE CALCULATOR →

https://ocaresponsescore.github.io/calculator/



DISCRIMINATION (c-statistics)

	POISE		ALP/ULN<1.67		NORMAL RANGE	
	ORS	ORS+	ORS	ORS+	ORS	ORS+
DERIVATION	0.75	0.83	0.78	0.88	0.72	0.81
VALIDATION	0.70	0.80	0.72	0.84	0.71	0.78

Clinical Gastroenterology and Hepatology

Development and validation of a scoring system to predict response to obeticholic acid in primary biliary cholangitis.

Short title: OCA response score

Antonio De Vincentis[1,2 *], Javier Ampuero[3 *], Francesca Terracciani[1,4 *], Daphne D'Amato[5], Alessio Gerussi[6], Laura Cristoferi[6], Nora Cazzagon[7], Emanuela Bonaiuto[7], Annarosa Floreani[7], Vincenza Calvaruso[8], Luca Cadamuro[8], Elisabetta Degasperi[9], Anna Morgando[5], Ester Vanni[5], Ana Lleo[10], Francesca Colapietro[10], Domenico Alvaro[11], Antonino Castellaneta[12], Sara Labanca[13], Mauro Viganò[14], Marco Distefano[15], Valeria Pace Palitti[16], Chiara Ricci[17], Nicoletta De Matthaeis[18], Marco Marzioni[19], Elena Gómez-Dominguez[20], Jose-Luis Montero[21], Esther Molina[22], Luisa Garcia-Buey[23], Marta Casado[24], Marina Berenguer[25], Isabel Conde[25], Miguel-Angel Simon[26], Javier Fuentes[27], Pedro Costa-Moreira[28], Guilherme Macedo[28], Francisco Jorquera[29], Rosa-Maria Morillas[30], Jose Presa[31], Jose-Manuel Sousa[3], Dario Gomes[32], Luis Santos[32], Antonio Olveira[33], Manuel Hernandez-Guerra[34], Leire Aburruza[35], Arsenio Santos[36], Armando Carvalho[36], Juan Uriz[37], Maria-Luisa Gutierrez[38], Elia Perez[38], Luchino Chessa[39], Adriano Pellicelli[40], Massimo Marignani[41], Luigi Muratori[42], Grazia Anna Niro[43], Maurizia Brunetto[44], Francesca Romana Ponziani[45], Maurizio Pompili[45], Fabio Marra[46], Andrea Galli[46], Alessandro Mussetto[47], Giuliano Alagna[48], Loredana Simone[49], Gaetano Bertino[50], Floriano Rosina[51], Raffaele Cozzolongo[52], Maurizio Russello[53], Leonardo Baiocchi[54], Carlo Saitta[55], Natalia Terreni[56], Teresa Zolfino[57], Cristina Rigamonti[58], Raffaella Vigano[59], Giuseppe Cuccorese[60], Pietro Pozzoni[61], Claudio Pedone[1, 62], Simone Grasso[63], Antonio Picardi[1,4], Pietro Invernizzi[6], Rodolfo Sacco[64], Antonio Izzi[65], Conrado Fernandez-Rodriguez[38,66 §], Umberto Vespasiani-Gentilucci[1,4 §] and Marco Carbone[6 §]

*shared first authorship; § shared last authorship

Affiliations

[1] Fondazione Policlinico Universitario Campus Bio-Medico, Via Alvaro del Portillo, 200, 00128 Roma, Italy; [2]Research Unit of Internal Medicine, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Via Alvaro del Portillo, 21, 00128 Roma, Italy; [3]Hospital Universitario Virgen del Rocio, Seville, Spain; [4]Research Unit of Hepatology, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Via Alvaro del Portillo, 21, 00128 Roma, Italy; [5]Gastroenterology Unit, Città della salute e della scienza, Torino, Italy; [6]Division of Gastroenterology, Centre for Autoimmune Liver Diseases, Department of Medicine and Surgery, University of Milano-Bicocca, European Reference Network on Hepatological Diseases, San Gerardo Hospital, Monza, Italy; [7]Gastroenterology Unit, Department of Surgery, Oncology and Gastroenterology, Padova University Hospital, Padova, Italy; [8]Gastroenterology and Hepatology Unit, University of Palermo, Palermo, Italy; [9]Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico – Division of Gastroenterology and Hepatology – CRC A.M. and A. Migliavacca Center for Liver Disease, Milan, Italy; [10]Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Italy and IRCCS Humanitas Research Hospital, Rozzano, Milan, Italy; [11]Department of Translational and Precision Medicine, University La Sapienza, Rome, Italy; [12]Gastroenterology Unit, Policlinico di Bari Hospital, Bari, Italy; [13]Gastroenterology Unit, Department of Internal Medicine, University of Genoa, IRCCS Ospedale Policlinico San Martino; [14]Hepatology Unit, ASST Papa Giovanni XXIII, Bergamo, Italy; [15]Department of Infectious Diseases, Umberto I Hospital, Siracusa, Italy; [16]Hepatology Unit, Santo Spirito Hospital, Pescara, Italy; [17]Internal Medicine, Spedali Civili, Brescia, Italy; [18]UOC of Gastroenterology, Gastroenterological, Endocrine-Metabolic and Nephro-Urological Sciences Department, Fondazione Policlinico Universitario A.Gemelli IRCCS, Università Cattolica del Sacro Cuore, Rome, Italy; [19]Clinic of Gastroenterology and Hepatology, Università Politecnica delle Marche, Ancona, Italy; [20]Hospital Universitario 12 de Octubre, Madrid, Spain; [21]Complexo Hospitalario Universitario Santiago de Compostela, Coruña, Spain; [22]Hospital Universitario Reina Sofia, Cordoba, Spain; [23]Hospital Universitario de La Princesa, Madrid, Spain; [24]Hospital de Torrecardenas, Almeria, Spain; [25]Hospital Universitario La Fe, Valencia, Spain; [26]Hospital Universitario Lozano-Blesa, Zaragoza, Spain; [27]Hospital Miguel Servet, Zaragoza, Spain; [28]Centro Hospitalar Sao Joao Porto, Portugal;

[29]Complejo Asistencial Universitario de Leon, Spain; [30]Hospital Germans Trias i Pujol, Badalona, Spain; [31]Centro Hospitalar De Trás-Os-Montes E Alto Douro, Portugal; [32]Dept de Gastroenterología, Centro Hospitalar Universitário de Coimbra, Portugal; [33]Hospital Universitario La Paz, Madrid, Spain; [34]Hospital Universitario de Canarias, Universidad de La Laguna, Tenerife, Spain; [35]Hospital Universitario de Donosti, Spain; [36]Centro Hospitalar e Universitário de Coimbra, Portugal; [37]Complejo Hospitalario de Navarra, Spain; [38]Hospital Universitario Fundación Alcorcón, Alcorcón, Madrid, Spain; [39]Liver Unit, University Hospital of Cagliari, Cagliari, Italy; [40]Liver Unit, San Camillo Hospital, Rome, Italy; [41]Digestive and Liver Disease Department, School of Medicine and Psychology University Sapienza, Azienda Ospedaliera S. Andrea, Rome, Italy; [42]DIMEC Università di Bologna, Policlinico di Sant'Orsola, Bologna, Italy; [43]Gastroenterology Unit, Fondazione Casa Sollievo Della Sofferenza IRCCS, San Giovanni Rotondo, Foggia, Italy; [44]University Hospital of Pisa, Pisa, Italy, Hepatology Unit; [45]Internal Medicine and Hepatology Unit, Policlinico Gemelli, Sapienza University, Rome, Italy; [46]Internal Medicine and Hepatology Unit, Department of Experimental and Clinical Medicine, University of Firenze, Firenze, Italy; [47]Gastroenterology Unit, Santa Maria Delle Croci Hospital, Ravenna, Italy; [48]UOC Medicina Interna Ospedale SS. Annunziata Sassari; [49]Department of Gastroenterology, University Hospital Sant'Anna, Ferrara, Italy; [50]Gastroenterology and Hepatology Unit, University Hospital Policlinico Vittorio Emanuele, Catania, Italy; [51]Medical Team Torino, Torino, Italy; [52]Gastroenterology Unit, National Institute of Gastroenterology S de Bellis Research Hospital, Castellana Grotte, Italy; [53]Liver Unit, Arnas Garibaldi, Catania, Italy; [54]Hepatology Unit, University of Rome "Tor Vergata", Rome, Italy; [55]Division of Medicine and Hepatology, University Hospital of Messina "Policlinico G. Martino", Messina, Italy; [56]Hepatology Unit, Valduce Hospital, Como, Italy; [57]Department of Gastroenterology, Brotzu Hospital, Cagliari, Italy; [58]Dipartimento di Medicina Traslazionale, Università del Piemonte Orientale, Novara, Italy and Division of Internal Medicine, Azienda Ospedaliero-Universitaria Maggiore della Carità, Novara, Italy; [59]Hepatology Unit, Niguarda Hospital, Milan, Italy; [60]Internal Medicine Ospedale "R.Dimiccoli", Barletta, Italy; [61]Hepatology Unit, Alessandro Manzoni Hospital, Lecco, Italy); [62]Research Unit of Geriatrics, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Via Alvaro del Portillo, 21, 00128 Roma, Italy; [63]Research Unit of Electronics for Sensor Systems, Università Campus Bio-Medico di Roma, Via Alvaro del Portillo, 21, 00128 Roma, Italy; [64]Gastroenterology Unit, Ospedali Riuniti, Foggia, Italy; [65]Department of Infectious Diseases, D. Cotugno Hospital, Napoli, Italy; [66]University Rey Juan Carlos Alcorcón, Madrid, Spain;

Corresponding authors

Marco Carbone, MD, PhD

Department of Medicine and Surgery

Centre for Autoimmune Liver Disease

University of Milan-Bicocca

Milan, Italy

Tel. +39 039.2334515

Mobile: +39 3927333154

Email: marco.carbone@unimib.it

Umberto Vespasiani-Gentilucci, MD, PhD

Unit of Clinical Medicine and Hepatology, Fondazione Policlinico Universitario Campus Bio-Medico

Research Unit of Hepatology, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Rome, Italy

Tel. +39 06.225411207

Email: u.vespasiani@policlinicocampus.it

Conflict of interest

A Lleo: consulting fees from Advanz Pharma, AlfaSigma, Takeda, and Albireo Pharma, and speaker fees from Gilead, Abbvie, MSD, Advanz Pharma, AlfaSigma, GSK, and Incyte; F Colapietro: speaker fees from Advanz Pharma; U Vespasiani Gentilucci: consulting fee from Astra Zeneca, Advance Pharma, Ipsen, Novo NordisK; M Carbone: advisor for Advanz Pharma, Albireo, Ipsen, Cymabay, Kowa, Moderna, Genetics SpA, Perspectum, GSK, Mayoly Spindler and scientific board of Ipsen, Cymabay, Kowa, Albireo; E Gómez-Dominguez: Consulted and received fees for consultation and remunerate speeches from Intercept Pharma; E Molina: Consulted and received fees for consultation and remunerate speeches from Intercept Pharma; M Berenguer : Grants and received fees for advisory from Intercept-Advanz; M Simon : Consulted and received fees for consultation and remunerate speeches from Intercept Pharma; F Jorquera : Fees for advisory from Intercept-Advanz; R Morillas : Fees for advisory and speeches from Intercept-Advanz;J Ampuero : Fees for consultation and remunerate speeches from Intercept Pharma-Advanz;A Olveira : Fees for advisory and speeches from Intercept-Advanz;M Hernandez-Guerra : Fees for advisory and speeches from Intercept-Advanz; A Carvalho : Fees for advisory and speaker from Intercept-Advanz.; C Fernandez-Rodriguez : His Institution has received a grant from Advanz. Other authors deny any potential conflict of interest.

Collaborators

Miki Scaravaglio, Eugenia Nofit (*Division of Gastroenterology, Centre for Autoimmune Liver Diseases, Department of Medicine and Surgery, University of Milano-Bicocca, European Reference Network on Hepatological Diseases, San Gerardo Hospital, Monza, Italy*), Giovanni Galati (*Fondazione Policlinico Universitario Campus Bio-Medico, Via Alvaro del Portillo, 200, 00128 Roma, Italy*), Francesco Pezzato, Paolo Rollo, Erica D'Ovidio (*Gastroenterology Unit, Department of Surgery, Oncology and Gastroenterology, Padova University Hospital, Padova, Italy*), Barbara Coco (*Hepatology Unit, University Hospital of Pisa, Pisa, Italy*), Annalisa Tortora (*Internal Medicine and Hepatology Unit, Fondazione Policlinico Universitario A.Gemelli IRCCS, Università Cattolica del Sacro Cuore, Rome, Italy*), Cecilia Fiorini (*Internal Medicine and Hepatology Unit, Department of Experimental and Clinical Medicine, University of Firenze, Firenze, Italy*), Rosanna Venere (*Department of Translational and Precision Medicine, University La Sapienza, Rome, Italy*), Gaetano Scifo (*Department of Infectious Diseases, Umberto I Hospital, Siracusa, Italy*), Mariarita Cannavò (*Liver Unit, Arnas Garibaldi, Catania, Italy*), Valentina Feletti (*Gastroenterology Unit, Santa Maria Delle Croci Hospital, Ravenna, Italy*), Fabrizio Pizzolante (*UOC of Gastroenterology, Gastroenterological, Endocrine-Metabolic and Nephro-Urological Sciences Department, Fondazione Policlinico Universitario A.Gemelli IRCCS, Università Cattolica del Sacro Cuore, Rome, Italy*), Edoardo Giovanni Giannini (*Gastroenterology Unit, Department of Internal Medicine, University of Genoa, IRCCS Ospedale Policlinico San Martino, Genova, Italy*), Rosa Cotugno (*Gastroenterology Unit, Fondazione Casa Sollievo Della Sofferenza IRCCS, San Giovanni Rotondo, Foggia, Italy*), Silvia Fanella (*Digestive and Liver Disease Department, School of Medicine and Psychology University Sapienza, Azienda Ospedaliera S. Andrea, Rome, Italy*), Francesco Losito (*Gastroenterology Unit, National Institute of Gastroenterology S de Bellis Research Hospital, Castellana*

Grotte, Italy), Giuseppe Grassi (*Hepatology Unit, University of Rome Tor Vergata, Rome, Italy*), Giulia Francesca Manfredi (*Dipartimento di Medicina Traslazionale, Università del Piemonte Orientale, Division of Internal Medicine, Azienda Ospedaliero-Universitaria Maggiore della Carità, Novara, Italy*), Valerio Buzzanca (*Clinic of Gastroenterology and Hepatology, Università Politecnica delle Marche, Ancona, Italy*), Barbara Omazzi (*Gastroenterology Unit, Guido Salvini Hospital, Rho, Italy*), Silvia Casella, Francesca Zani (*Hepatology Unit, Spedali Civili Gardone Val Trompia, Brescia, Italy*), Chiara Ricci (*Internal Medicine, Spedali Civili, Brescia, Italy*), Valentina Bellia (*Hepatology Unit, Valtellina e Alto Lario Hospital, Sondrio, Italy*), Ludovico Abenavoli (*Department of Health Sciences, University "Magna Graecia" of Catanzaro, Italy*), Olivia Morelli (*Gastroenterology and Hepatology, Department of Medicine, Università degli Studi di Perugia, Perugia, Italy*), Lory Saveria Crocè (*Department of Medical Surgical and Health Sciences, University of Trieste, Liver Clinic, Trieste University Hospital, Trieste, Italy*), Paolo Scivetti, Antonio Panero (*Internal Medicine Unit, Azienda Sanitaria Locale di Biella, Biella, Italy*), Valentina Boano (*Department of Gastroenterology and Endoscopy, Cardinal Massaia Hospital, Asti, Italy*), Guido Poggi (*Oncology Unit, Istituto di Cura Città di Pavia, Pavia, Italy*), Giancarlo Gimignani, Alessandro Conforti (*Unit of Internal Medicine, San Paolo Hospital, Civitavecchia*), Evelise Frazzetto, Laura Rapisarda (*Gastroenterology and Hepatology Unit, University Hospital Policlinico Vittorio Emanuele, Catania, Italy*), Shrin Demma (*Liver Unit, San Camillo Hospital, Rome, Italy*).

Abstract (max 260 w)

Background & Aims. Obeticholic acid (OCA) is the only licensed second-line therapy for primary biliary cholangitis (PBC). With novel therapeutics in advanced development, clinical tools are needed to tailor the treatment algorithm. We aimed to derive and externally validate the *OCA response score* (ORS) for predicting the response probability of individuals with PBC to OCA.

Methods. We used data from the Italian RECAPITULATE (N 441) and the IBER-PBC (N 244) OCA real-world prospective cohorts to derive/validate a score including widely available variables obtained either pre-treatment (ORS), or also after 6 months of treatment (ORS+). Multivariable Cox's regressions with backward selection were applied to obtain parsimonious predictive models. The predicted outcomes were biochemical response according to POISE (ALP/ULN<1.67 with a reduction of at least 15%, and normal bilirubin), or ALP/ULN<1.67, or NORMAL RANGE criteria (NR: normal ALP, ALT and bilirubin) up to 24 months.

Results. Depending on the response criteria, ORS included age, pruritus, cirrhosis, ALP/ULN, ALT/ULN, GGT/ULN and bilirubin. ORS+ also included ALP/ULN and bilirubin after 6 months of OCA therapy. Internally validated c-statistics for ORS were of 0.75, 0.78 and 0.72 for POISE, ALP/ULN<1.67 and NR response, which raised to 0.83, 0.88, 0.81 with ORS+, respectively. The respective performances in validation were of 0.70, 0.72 and 0.71 for ORS, and 0.80, 0.84, 0.78 for ORS+. Results were consistent across groups with mild/severe disease.

Conclusions. We developed and externally validated a scoring system capable to predict OCA response according to different criteria. This tool will enhance a stratified second-line therapy model to streamline standard care and trial delivery in PBC.

Keywords

Obeticholic acid; predictive model; primary biliary cholangitis

Introduction

Obeticholic acid (OCA) is the only licensed second-line therapy for patients with primary biliary cholangitis (PBC) with inadequate biochemical response or intolerance to ursodeoxycholic acid (UDCA). In the registrative, randomised controlled trial (POISE) and its open-label extension, OCA induced a significant reduction of alkaline phosphatase (ALP) with a stabilisation of total bilirubin up to 48 months of treatment.[1,2] Subsequently, several post-marketing real-world studies confirmed these data by showing an effective biochemical response in ~40% of patients.[3–6] Furthermore, a recent study comparing patients from clinical trial setting with real-world external controls highlighted greater transplant-free survival in OCA-treated individuals.[7]

However, numerous challenges are still to be faced to optimize the management of patients with PBC not responding to UDCA. Firstly, upward to 50-70% of patients, particularly those with liver cirrhosis, are not rescued to an effective response even with OCA therapy. It must be also considered that OCA use has been restricted in subjects with cirrhosis with present or previous hepatic decompensation or signs of portal hypertension, and -as such- it cannot be an option in these cases. Moreover, accumulating evidence is supporting the switch of treatment goal in PBC from the simple amelioration of liver biochemistry (*e.g.* POISE, Toronto, Paris criteria) towards its complete normalization, as this is associated with the best patient outcomes.[8,9] However, this target is achieved only in a minority of subjects (~10-15%) taking OCA.[3]

New therapeutic agents, some of which are in advanced phase III investigation, might offer hope in the near future, namely seladelpar,[10] elafibranor,[11] saroglitazar,[12] within the peroxisome proliferator-activated receptors (PPARs) agonist family, and setanaxib[13], a NADPH oxidase (NOX) inhibitor. Moreover, fibrates (*e.g.* bezafibrate and fenofibrate), PPAR agonists used as a lipid-lowering agent to treat hyperlipidaemia, have already shown to provide an effective biochemical response[14] and improved clinical outcomes,[15] although their use currently remains off-label. In addition, preliminary evidence from ongoing trials suggests that the combination therapy with OCA and bezafibrate can induce high rates of ALP normalization with a better safety profile.[16]

In this evolving scenario, the allocation of individuals to the second line therapy with a higher likelihood of success will be central according to a personalized medicine approach allowing to save the costs of ineffective, and potentially harmful, therapies and improving individual patient's prognosis. However, at present, there are no clinical tools capable to forecast treatment response and failure and side effects to OCA, *i.e.* the only approved second-line drug so far. We have recently described the most impacting predictive factors for biochemical response, *i.e.* liver cirrhosis, pruritus, and higher baseline ALP and total bilirubin.[3,4] However, to date, a tool capturing the predictive information of all these factors, and providing a response probability based relevant baseline information, is still lacking.

Therefore, in the present study, we aimed to derive and validate a score (i.e., the *OCA response score, ORS*) that, based on easily available baseline characteristics, would accurately predict the probability of an individual patient to respond to OCA therapy.

Methods

This study was conducted and reported in accordance with the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prediction or Diagnosis) guidelines.[17]

Derivation cohort (RECAPITULATE)

The "REal-world obetiCholic Acid theraPy in ITaly recapitULATion of Efficacy and safety" (RECAPITULATE) cohort was leveraged to derive the ORS. The RECAPITULATE is a prospective study from centres belonging to the *Italian PBC Registry* and/or the *Club Epatologi Ospedalieri* (CLEO) and/or the *Associazione Italiana Gastroenterologi e Endoscopisti Digestivi Ospedalieri* (AIGO) PBC group, the *Sicilian PBC Network* and *PBC Project Piemonte-Liguria-Valle D'Aosta*. All adult patients with PBC consecutively starting OCA in 51 Italian centres from September 2017 to February 2022 were included. More details can be found in Supplementary materials. The complete RECAPITULATE cohort included 487 patients. After excluding 10 subjects with unavailable baseline ALP and 36 with a follow-up <6 months, a final cohort of 441 individuals was used to derive the ORS.

Validation cohort (IBER-PBC)

The IBER-PBC cohort was used to externally validate the predictive performance of the ORS. This is a prospective, observational, multicentre study collecting real-world data on patients with PBC from 25 institutions in Spain and Portugal.[5] All adults patients prescribed with OCA in the participating centres were included (see Supplementary materials). A total of 244 subjects with available baseline ALP and with at least 6 months' follow-up were constituted the IBER-PBC validation cohort.

Definition of predicted outcomes

The study outcome was the attainment of an effective biochemical response to OCA therapy, as defined by different criteria:

- POISE (alkaline phosphatase(ALP)/upper limit of normal(ULN) <1.67 with a reduction of at least 15%, and total bilirubin≤1 mg/dL);
- ALP/ULN<1.67;
- NORMAL RANGE (NR, ALP/ULN≤1 and ALT/ULN≤1 and total bilirubin≤1 mg/dL).

Biochemical response was adjudicated when the above-mentioned criteria were attained in at least 2 consecutive follow-ups, with no more than one isolate violation of the criteria thereafter. Follow-up commenced at the date of the start of OCA therapy. Patients not attaining biochemical response during follow-up were censored at the time either of OCA discontinuation for any cause, or of fibrate start or of the last database update (July 31st 2022 for RECAPITULATE and April 30th 2023 for IBER-PBC), whichever occurred first.

Selection of candidate predictors and missing data

The primary candidate variables of interest were easily and readily available clinical and biochemical parameters known to influence OCA response probability based on previously published studies.[3,4,6] In the derivation cohort, baseline variables with >5% missingness were excluded from score derivation: platelet count (missing 132, 30%), albumin (missing 190, 43%), prothrombin time (missing 188, 43%), creatinine (missing 189, 43%) and body mass index (missing 99, 22%). Variables with ≤5% missing values were conversely imputed with random forests (see Supplementary Materials). Finally, 13 candidate variables at baseline (sex, age, disease duration, diabetes, UDCA not treated, PBC/AIH overlap, cirrhosis, ALP/ULN, AST/ULN, ALT/ULN, GGT/ULN and total bilirubin), and other two collected after 6 months of OCA therapy (ALP/ULN and total bilirubin) were considered for score derivation (Table 1).

Statistical analyses

Two types of predictive Cox regression models for each outcome were developed: one based only on baseline variables (ORS), and one possibly including ALP/ULN and total bilirubin after 6 months of OCA (ORS+). From full multivariable models, parsimonious models were obtained with automated backward selection procedure to derive the ORS for each outcome (ORS_{POISE}, $ORS_{ALP/ULN<1.67}$, ORS_{NR}). Then, the ORS+ was derived in the subset of subjects not responding/censored at 6 months by adding the relative change from baseline of ALP/ULN and/or total bilirubin after 6 months of OCA therapy[(value at 6 months - value at baseline)/value at baseline]. A penalized maximum likelihood estimation was used to account for overfitting of the models.[18] The ORS/ORS+ were calculated as the regression linear predictor (βX). Predicted response probabilities were consequently estimated accordingly. More details on the statistical procedure can be found in the Supplementary materials. An online ORS calculator can be found at https://ocaresponsescore.github.io/calculator/.

Discrimination was measured with Harrell's c-statistics, 12- and 24-month time-dependent area under the curve (AUC), and visually shown by plotting the cumulative incidence curves according to quartiles of the score. Overfitting was evaluated by 300-bootstrapped calibration slopes. Calibration was studied through calibration plots. External validation was performed in the IBER-PBC cohort. All analyses were carried out using R version 4.2.0.

Results

Study cohorts

The RECAPITULATE included 441 individuals (women 88%, mean age 57.8), while the IBER-PBC comprised 244 individuals (women 93%, mean age 56.6) with a lower proportion of subjects with cirrhosis (23% *vs* 34%). Apart from this, the two cohorts presented similar clinical and biochemical features at OCA start, and a comparable change of ALP/ULN and of total bilirubin after 6 months of therapy (Table 1). Median follow-up time was 18 and 23 months in the RECAPITULATE and IBER-PBC, respectively. The observed OCA response rate according to POISE, ALP/ULN<1.67 and NR was 38%, 58%, 10% at 12 months, and 46%,

66%, 16% at 24 months in the RECAPITULATE, and 36%, 51%, 7% and 46%, 63%, 10% at 12 and 24 months in the IBER-PBC.

Models development, performance, and internal validation.

The phases of the model-building procedures are detailed in the Supplementary materials. The final multivariable models for ORS included age at OCA start, pruritus, cirrhosis, ALP/ULN, ALT/ULN, GGT/ULN and total bilirubin for the prediction of POISE, and pruritus, cirrhosis, ALP/ULN, GGT/ULN and total bilirubin for the prediction of ALP/ULN<1.67 response (Table 2). For the prediction of NR response, only pruritus, ALP/ULN and Total Bilirubin were retained (Table 2). ORS+ models also included the relative change of ALP/ULN for all the outcomes, and also of total bilirubin after 6 months of OCA therapy for POISE. We did not identify significant interaction terms in the final models. However, we found that the main effect of selected variables (*i.e.* ALT and GGT) was consistently less pronounced in patients with higher disease activity and fibrosis stage, as indicated by higher ALP and bilirubin values (Supplementary Figure 3).

ORS had an optimism-corrected Harrell's c-statistics of 0.75, 0.78 and 0.72 for POISE, ALP/ULN<1.67 and NR response, with an apparent 12- and 24-months AUC of 0.78, 0.80 for POISE, of 0.83, 0.83 for ALP/ULN<1.67, and of 0.79, 0.72 for NR response (Table 3). With ORS+, the optimism-corrected c-statistics raised to 0.83, 0.88, 0.81, respectively, and apparent 12- and 24-months AUC were 0.87, 0.85 for POISE, 0.91, 0.89 for ALP/ULN<1.67, and 0.85, 0.80 for NR response (Table 3). Calibration slopes of ORS and ORS+ on 300 bootstrapped samples were 0.92, 0.93 for POISE, 0.96, 0.96 for ALP/ULN<1.67, and 0.93, 0.91 for NR response, suggesting modest overfitting. Internal validation disclosed mean |errors| in prediction in the range of 0.02-0.05 (0.11 only for prediction of ALP/ULN<1.67 at 12 months; Supplementary Figure 2), indicating good general calibration.

Subgroup analyses were conducted in men, patients with cirrhosis, subjects starting with ALP/ULN values above 3 or in therapy with OCA and UDCA (i.e. after excluding 9 and 22 individuals intolerant to UDCA or already taking fibrates at OCA start, respectively), and disclosed comparable discriminative performances to those observed in the complete cohort (Table 3).

Example.

For a 55-year-old patient with ALP/ULN of 2.4, total bilirubin level of 1.2 mg/dL, ALT/ULN of 1.2, GGT/ULN of 4 without pruritus or advanced liver disease: ORS_{POISE}= -0.15, $ORS_{ALP/ULN<1.67}$=-0.18, ORS_{NR}=-0.77. The corresponding probabilities of OCA response at 24 months are 35% for POISE, 60% for ALP/ULN<1.67 and 6% for NR. After 6 months, the patient attains response to OCA according to ALP/ULN<1.67, since reporting a drop of ALP/ULN to 1.5 (-37.5%) with total bilirubin of 1.1 mg/dL (-8.3%). The residual predicted probability of attaining also POISE or NR response in the following 18 months (*i.e.* within 24 months from OCA start) is 38% and 5%, respectively.

External validation.

In the IBER-PBC cohort, ORS showed Harrell's c-statistics of 0.70, 0.72 and 0.71 for POISE, ALP/ULN<1.67 and NR response, with 12- and 24-months AUCs of 0.77, 0.73 for POISE, of 0.80, 0.82 for ALP/ULN<1.67, and of 0.74, 0.71 for NR response. ORS+ improved c-statistics to 0.80, 0.84, 0.78, respectively, with 12- and 24-months AUCs of 0.89, 0.82 for POISE, of 0.91, 0.89 for ALP/ULN<1.67, and of 0.78, 0.80 for NR response. Predicted response probabilities well corresponded with the observed ones (mean |error|~0.02-0.08; Figure 1). While NR predictions were globally well calibrated with mean |error| of 0.03-0.04, calibration plots evidenced at tendency for overestimation for higher observed risk.

Risk stratification.

The models for POISE and ALP/ULN<1.67 could identify quartiles of subjects with progressively increasing cumulative incidence of OCA response both in the derivation and validation cohorts (Figure 2). This was confirmed by increasing hazard ratios (Figure 2). For NR, while correctly identifying groups with lower and higher response (I and IV quartiles), risk was less clearly stratified in the intermediate classes (II and III quartile). An ORS_{POISE} below -1.3 identified a small proportion of subjects (N 39 [9%] and 23 [9%] in the derivation and validation cohorts, respectively) with a low 24-month observed POISE response probability (6% and 7%) and high negative predictive value (91% and 92%, respectively).

Discussion.

In the present study, we analysed data from two large and independent real-world cohorts, including a total of 685 PBC subjects treated with OCA, to develop and validate a scoring system (ORS/ORS+) that accurately predicts OCA response according to different criteria. The final clinical score incorporates readily available clinical and biochemical parameters, such as age, presence of pruritus and of cirrhosis, ALP, total bilirubin, ALT and GGT. ORS/ORS+ performances were good/excellent and comparable between subgroups with milder and more severe disease. In clinical practice, the ORS is expected to help in driving treatment allocations based on a personalized medicine approach, and its usefulness will increase further as soon as alternative approved second-line options will become available.

Previous studies have suggested the prognostic importance of the individual components of the ORS.[3,4,6] In particular, the probability of response to OCA sharply declines with increasing baseline values of ALP and total bilirubin[3,4,6], which were confirmed as the strongest predictors in the ORS. This is consistent with ALP and bilirubin being the reference parameters of biochemical response, but it likely results also from their association with the severity of biliary injury and ductopenia, which reflect a more aggressive PBC phenotype. Indeed, the presence of cirrhosis was another negative predictive factor of response to OCA,[4] as well as advanced disease stage was a key negative determinant in the case of response to UDCA.[19] It is also not surprising that pruritus was associated with a lower likelihood of response to OCA. Indeed, being potentially further worsened by OCA, baseline pruritus predicts a higher probability of drug discontinuation and, ultimately, treatment failure. To note, ALT and GGT had a marginal but still prognostically meaningful role in predicting OCA response, at least with respect to POISE criteria. In PBC, the elevation of ALT represents

interface hepatitis activity, which can be relevant even in absence of a definite AIH overlap. While higher ALT levels are directly associated with response to UDCA at the time of PBC diagnosis[19], here we observed an inverse association between ALT levels and response to OCA. This finding is possibly due to the fact that the presence of persistently elevated ALT during long-term treatment with UDCA circumscribes a subgroup of patients with more aggressive and less responsive disease phenotype. Instead, there was a slightly direct association between GGT levels and the likelihood of attaining response according to POISE and ALP<1.67 criteria, which was somehow unexpected considering the well-known association of GGT with ALP levels and with a worse PBC prognosis.[20] Notably, differently from ALP, GGT levels are affected by oxidative stress due to drug/alcohol exposure and, more frequently, to the coexistence of steatotic liver disease and metabolic comorbidities (*e.g.* diabetes or obesity). As such, it is possible that the relative elevation of GGT with respect to that of ALP might identify a subgroup of PBC patients with a dysmetabolic background which is somehow sensitive to some of OCA pharmacologic activities. Indeed, it is well known that the activation of FXR by OCA reduces liver fat and has potent metabolic effects.[21] Notably, the association of GGT, but also that of ALT, are dependent on disease severity, being evident only in case of milder disease (*i.e.* lower ALP and bilirubin values; see Supplementary Figure 3). Conversely, at higher ALP and bilirubin levels, their predictive potential is lost, since strong disease activity and advanced fibrosis force out any other surrogate predictive index.

The ORS was derived by synthesizing the prognostic information conveyed by the abovementioned parameters, either alone or with the inclusion of ALP and total bilirubin after 6 months (ORS+), and externally validated in an independent large real-world cohort. Different definitions of biochemical response to OCA were considered to embrace progressively more stringent criteria. Indeed, ALP/ULN<1.67 and POISE are the criteria traditionally applied to define UDCA and OCA response, respectively, in clinical studies. Conversely, response according to NR criteria (i.e. complete normalization of ALP, ALT and total bilirubin) represents the new treatment goal in PBC, as it has been shown to provide the best disease outcomes.[8] ORS/ORS+ displayed fair to good discrimination in both derivation and validation. Consistent results were obtained for time-dependent predictions at 12-24 months, and in subgroup of individuals with mild and severe disease (men, presence/absence of cirrhosis or starting ALP/ULN>/≤3). Performances were good for all differently defined outcomes and their reliability is supported by the consistent results obtained in two unselected and independent real-world cohorts. Notably, the additional information on ALP and bilirubin response at 6 month (ORS+) further increased the observed performances to c-statistics in the range of 0.8 or more both in derivation and validation. Altogether, ORS/ORS+ will allow the treating physician to carve therapeutic strategies with flexibility, estimating the chances of OCA treatment success starting from the minimum (ALP/ULN<1.67) and up to the most ambitious (NR) target.

Since 2016, regulatory agencies have licensed the use of OCA in patients with PBC and inadequate response or intolerance to UDCA. Thereafter, also bezafibrate was disclosed to be effective in ameliorating liver biochemistry and improving survival in these patients.[14,15] However, to date, bezafibrate use as second-line agent for the treatment of PBC still remains off-label, and it largely depends on local availability and practice.

Concerning OCA, post-marketing studies confirmed its capability to induce biochemical response according to POISE criteria in ~40% of cases,[3–6] but the complete normalization of liver biochemistry is achieved only in ~10-15%. Moreover, its use is not free from unpleasant side effects (*e.g.* pruritus), and it has been associated with occurrence of severe adverse events when prescribed to patients with advanced cirrhosis (Child class B and C, previous decompensation). New second-line approaches including either new drugs (seladelpar,[10] elafibranor,[11] saroglitazar[12], and setanaxib[13]) or combinational strategies (OCA and bezafibrate) are currently under-evaluations with very promising results. Given the forthcoming new therapeutic options, a structured algorithm will be needed to allocate the most effective therapy to the patient with the highest response chances according to a personalized medicine approach. Tools like ORS/ORS+ will likely play a central role in this context. Indeed, the possibility of stratifying patients according to OCA response probabilities paves the way to personalized approaches for prioritizing the prescription of OCA in subjects with high response chances or, conversely, for fast-tracking the switch to a combinational therapy with bezafibrate or to another drug in those with low response chances (as displayed in Figure 2).

The relatively large sample size used for ORS derivation, along with the robust validation in an independent cohort, are the main strengths of this study. The real-world setting makes the obtained scoring system directly exportable to the intended target population, and it is expected to increase its generalizability. Moreover, the incorporation of easily available clinical variables makes its use in clinical practice highly feasible, and amenable to further external validation. The complex calculation has been simplified by the development of a web calculator (https://ocaresponsescore.github.io/calculator/) to improve its usage in clinical practice.

This study has also some limitations. First of all, the heterogeneity of patient characteristics, which is inevitable in a real-world scenario. Indeed, patients with PBC/AIH overlap, not taking UDCA, or taking fibrates together with OCA were retained in the analysis. However, subgroups analyses were performed, and no sensible deviations were evidenced from what observed in the overall cohort. Secondly, the high rate of missing values for certain variables (i.e. platelets, albumin, liver stiffness measurements) hampered their inclusion in the models, even though some of them could likely play a role in the prognostic prediction based on the *a priori* knowledge. Finally, the scoring system has been both derived and validated in cohorts from Southern Europe. As such, validation in other cohorts from North Europe/America is warranted to confirm its full generalizability.

In conclusion, by analysing two independent large real-world cohorts of patients with PBC started with OCA treatment, we have derived and externally validated a score capable of predicting the probabilities of response to the drug according to different meaningful criteria. Together with the UDCA response scores, and with other scores which will be hopefully developed to predict the response to the new incoming second-line drugs, the ORS/ORS+ will enhance the background knowledge needed to face PBC treatment according to a personalized medicine approach.

References

1. *Nevens F, Andreone P, Mazzella G, et al. A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis. N Engl J Med 2016;375:631-643.*

2. *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. The Lancet Gastroenterology & Hepatology 2019;4:445-453.*

3. *D'Amato D, De Vincentis A, Malinverno F, et al. Real-world experience with obeticholic acid in patients with primary biliary cholangitis. JHEPReport 2021;3:100248.*

4. *De Vincentis A, D'Amato D, Cristoferi L, et al. Predictors of serious adverse events and non-response in cirrhotic patients with primary biliary cholangitis treated with obeticholic acid. Liver Int 2022;42:2453-2465.*

5. *Gomez E, Garcia Buey L, Molina E, et al. Effectiveness and safety of obeticholic acid in a Southern European multicentre cohort of patients with primary biliary cholangitis and suboptimal response to ursodeoxycholic acid. Aliment Pharmacol Ther 2021;53:519-530.*

6. *Abbas N, Culver EL, Thorburn D, et al. UK-wide Multicenter Evaluation of Second-line Therapies in Primary Biliary Cholangitis. Clinical Gastroenterology and Hepatology 2022;0.*

7. *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 September 2022:S0016-5085(22)01060-5.*

8. *Murillo Perez CF, Harms MH, Lindor KD, et al. Goals of Treatment for Improved Survival in Primary Biliary Cholangitis: Treatment Target Should Be Bilirubin Within the Normal Range and Normalization of Alkaline Phosphatase. Official journal of the American College of Gastroenterology | ACG 2020;115:1066-1074.*

9. *Corpechot C, Lemoinne S, Soret P-A, et al. Adequate vs. deep response to UDCA in PBC: To what extent and under what conditions is normal ALP level associated with complication-free survival gain? Hepatology July 2023.*

10. *Hirschfield GM, Bowlus CL, Mayo MJ, et al. A Phase 3 Trial of Seladelpar in Primary Biliary Cholangitis. New England Journal of Medicine 2024;390:783-794.*

11. *Kowdley KV, Bowlus CL, Levy C, et al. Efficacy and Safety of Elafibranor in Primary Biliary Cholangitis. New England Journal of Medicine 2024;390:795-805.*

12. *Zydus Therapeutics Inc. A Multicenter, Randomized, Double-Blind, Placebo Controlled, Phase 2b/3 Study to Evaluate the Efficacy and Safety of Saroglitazar Magnesium in Subjects With Primary Biliary Cholangitis. clinicaltrials.gov, 2023. https://clinicaltrials.gov/study/NCT05133336. Accessed July 17, 2023.*

13. *Calliditas Therapeutics Suisse SA. TRANSFORM: A 52-Week, Randomized, Placebo-Controlled, Double-Blind, Adaptive Phase 2b/3 Trial of Setanaxib With a 52-Week Extension Phase in Patients With Primary Biliary Cholangitis (PBC) and Elevated Liver Stiffness. clinicaltrials.gov, 2023. https://clinicaltrials.gov/study/NCT05014672. Accessed July 17, 2023.*

14. *Corpechot C, Chazouilleres O, Rousseau A, et al. A placebo-controlled trial of bezafibrate in primary biliary cholangitis. New England Journal of Medicine 2018;378:2171-2181.*

15. *Tanaka A, Hirohara J, Nakano T, et al. Association of bezafibrate with transplant-free survival in patients with primary biliary cholangitis. Journal of Hepatology 2021;75:565-571.*

16. *Hejda V, Louvet A, Civitarese A, et al. Results from a planned interim analysis of a randomized, double-blind, active-controlled trial evaluating the effects of obeticholic acid and bezafibrate on serum biomarkers in primary biliary cholangitis. Journal of Hepatology 2023;78:S45.*

17. *Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015;350:g7594.*

18. *Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. Biom J 2010;52:70-84.*

19. *Carbone M, Nardi A, Flack S, et al. Pretreatment prediction of response to ursodeoxycholic acid in primary biliary cholangitis: development and validation of the UDCA Response Score. The Lancet Gastroenterology & Hepatology 2018;3:626-634.*

20. *Gerussi A, Bernasconi DP, O'Donnell SE, et al. Measurement of Gamma Glutamyl Transferase to Determine Risk of Liver Transplantation or Death in Patients With Primary Biliary Cholangitis. Clin Gastroenterol Hepatol 2021;19:1688-1697.e14.*

21. *Neuschwander-Tetri BA, Loomba R, Sanyal AJ, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. Lancet 2015;385:956-965.*

Table 1. Baseline characteristics of the study cohorts.

	Derivation cohort	Validation cohort
	RECAPITULATE	**IBER-PBC**
Country	**Italy**	**Spain-Portugal**
N	441	244
Sex, female	390 (88%)	226 (93%)
Age at OCA start, years	57.8 (10.7)	56.6 (10.2)
Duration of disease before OCA start, years	7.0 (3.0-12.0)	7.8 (3.4-13.1)
Diabetes mellitus	30 (7%)	n.a.
UDCA not treated	9 (2%)	0 (0%)
PBC/AIH overlap[*]	59 (13%)	40 (16%)
Cirrhosis [£]	152 (34%)	56 (23%)
Pruritus at baseline	141 (32%)	107 (44%)
ALP/ULN at baseline	2.0 (1.7-2.9)	2.1 (1.7-2.8)
ALT/ULN at baseline	1.1 (0.8-1.7)	1.2 (0.8-2)
AST/ULN at baseline	1.1 (0.8-1.6)	1.3 (1-1.9)
GGT/ULN at baseline	4.1 (2.3-6.4)	4 (2.3-7.3)
Total Bilirubin at baseline	0.7 (0.5-1.0)	0.7 (0.5-0.9)
Change ALP/ULN at 6 months, relative[”]	-0.3 (-0.4--0.1)	-0.2 (-0.4--0.1)
Change Total Bilirubin at 6 months, relative[”]	0.0 (-0.2-0.1)	0.0 (-0.3-0.2)

Data reported as means and standard deviation (SD) or median with 25th– 75th centile for continuous variables, as appropriate, and as numbers with percentage frequency for categorical variables. For Fibrate therapy: 22 and 41 subjects from the RECAPITULATE and IBER-PBC were already on fibrate therapy at the time of OCA and were kept in the analyses, while subjects starting fibrates after OCA were censored at fibrate start.
OCA prescribed dose in the RECAPITULATE was 5 mg daily in 221 patients (50%), 5 mg uptitrated to 10 mg daily in 170 patients (39%), 5 mg every other day uptitrated to 5 mg daily in 15 patients (3%), other dosages in 35 patients (8%).
[*]PBC/AIH overlap syndrome was defined by histological evidence and all included patients were on a stable immunosuppressive therapy for at least 6 months.
[£]presence of cirrhosis was ascertained by 1) liver histology and/or 2) liver stiffness measurement assessed by vibration-controlled transient elastography ≥16.9 kPa and/or 3) radiological (surface nodularity, caudate lobe hypertrophy, enlarged spleen or other sign of portal hypertension at ultrasound scan), and/or clinical features (presence of gastro-esophageal varices or previous decompensating events, such as ascites, variceal bleeding, encephalopathy) eventually supported by laboratory findings (low platelets, low albumin, prolonged prothrombin time).
[”]calculated as (value at 6 months – value at baseline)/value at baseline.
Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma glutamyl transferase; n.a., not available; PBC,/AIH primary biliary cholangitis/autoimmune hepatitis; OCA, obeticholic acid; UDCA, ursodeoxycholic acid; ULN, upper limit of normal.

Table 2. Derivation models for OCA response scores (ORS) according to different response criteria.

A) MODEL DERIVATION	POISE RESPONSE		ALP/ULN<1.67 RESPONSE		NR RESPONSE	
Score	**ORS_{POISE}**	**ORS_{POISE}+**	**$ORS_{ALP/ULN<1.67}$**	**$ORS_{ALP/ULN<1.67}$+**	**ORS_{NR}**	**ORS_{NR}+**
Candidate Predictor	aHR (95%CI), χ^2	aHR (95%CI), χ^2	aHR (95%CI), χ^2	aHR (95%CI), χ^2	aHR (95%CI), χ^2	aHR (95%CI), χ^2
Sex, female						
Age at OCA start	0.79 (0.62-1.01), 3.7	0.85 (0.67-1.09), 2				
Duration of disease						
UDCA not treated						
Diabetes mellitus						
PBC/AIH overlap						
Pruritus at baseline	0.52 (0.36-0.76), 11	0.70 (0.48-1.03), 3	0.59 (0.44-0.79), 12	0.73 (0.55-0.98), 4	0.64 (0.34-1.21), 2	0.87 (0.46-1.66), 2
ALP/ULN at baseline	0.43 (0.28-0.66), 27	0.42 (0.28-0.64), 42	0.33 (0.26-0.42), 84	0.21 (0.16-0.27), 150	0.34 (0.20-0.56), 18	0.15 (0.08-0.29), 31
Cirrhosis	0.77 (0.55-1.07), 2.4	0.74 (0.53-1.05), 3	0.80 (0.61-1.04), 2.7	0.82 (0.63-1.07), 2		
Total Bilirubin at baseline	0.52 (0.41-0.67), 26	0.50 (0.38-0.66), 24	0.78 (0.67-0.91), 9.4	0.84 (0.71-1.00), 4	0.53 (0.35-0.80), 9	0.50 (0.32-0.79), 9
ALT/ULN at baseline	0.82 (0.58-1.16), 1.2	0.77 (0.54-1.11), 2				
AST/ULN at baseline						
GGT/ULN at baseline	1.30 (1.10-1.52), 10	1.23 (1.05-1.45), 6	1.09 (0.96-1.25), 1.8	1.01 (0.88-1.15), 0.9		
Change ALP/ULN at 6 months, relative		0.43 (0.33-0.57), 37		0.43 (0.35-0.52), 73		0.22 (0.12-0.40), 25
Change Total Bilirubin at 6 months, relative		0.88 (0.73-1.06), 2				

ORS models were obtained from full models including all candidate predictors with automated backward selection procedure using the Akaike Information Criteria as stopping rule, and the Wald χ^2 of individual variables as the statistics on which to base the stopping rule. ORS+ models are fitted only in subjects not responding/censored at 6 months (N 264, N 190 and N 354 for POISE, ALP/ULN<1.67 and NR), with the addition to reduced models 1 of the relative change of ALP/ULN and of total bilirubin at 6 months of OCA therapy. The relative change is calculated as [(value at 6 months – value at baseline) / value at baseline]. A penalized maximum likelihood estimation was used to account for overfitting. Variables have been transformed as detailed in Supplementary table 1, and hazard ratios (HR) are reported for the comparison of the third vs first quartile for continuous variables, and for categories for categorical variables. Wald χ^2 is reported for indicating the contribution of each variable in the predictive scores.

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the curve; CI, confidence intervals; GGT, gamma glutamyl transferase; PBC,/AIH primary biliary cholangitis/autoimmune hepatitis; NR, normal range; OCA, obeticholic acid; ORS, OCA response score; sbjs, subjects; UDCA, ursodeoxycholic acid; ULN, upper limit of normal.

Table 3. Discriminative performance of the OCA response scores (ORS) according to different response criteria.

MODEL DISCRIMINATION	ORS_{POISE}	ORS_{POISE}+	$ORS_{ALP/ULN<1.67}$	$ORS_{ALP/ULN<1.67}$+	ORS_{NR}	ORS_{NR}+
DERIVATION COHORT (N 441)						
Harrell's C-statistics, apparent	0.77	0.84	0.79	0.88	0.73	0.82
Harrell's C-statistics, optimism-corrected*	0.75	0.83	0.78	0.88	0.72	0.81
Time-dependent AUC at 12 months	0.78 (0.73-0.84)	0.87 (0.83-0.90)	0.83 (0.78-0.88)	0.91 (0.87-0.95)	0.79 (0.71-0.87)	0.85 (0.82-0.89)
Time-dependent AUC at 24 months	0.80 (0.74-0.86)	0.85 (0.79-0.90)	0.83 (0.77-0.88)	0.89 (0.83-0.95)	0.72 (0.64-0.78)	0.80 (0.73-0.87)
VALIDATION COHORT (N 244)						
Harrell's C-statistics, apparent	0.70	0.80	0.72	0.84	0.71	0.78
Time-dependent AUC at 12 months	0.77 (0.71-0.83)	0.89 (0.83-0.95)	0.80 (0.73-0.86)	0.91 (0.86-0.96)	0.74 (0.59-0.9)	0.78 (0.55-0.97)
Time-dependent AUC at 24 months	0.73 (0.65-0.80)	0.82 (0.76-0.88)	0.82 (0.75-0.89)	0.89 (0.83-0.96)	0.71 (0.57-0.82)	0.80 (0.66-0.93)
Subgroup analyses in the derivation cohort						
Men (N 51)						
Harrell's C-statistics, apparent	0.79	0.83	0.77	0.82	0.77	0.79
Time-dependent AUC at 12 months	0.80 (0.65-0.96)	0.90 (0.79-0.99)	0.86 (0.69-1.02)	-[a]	-[a]	-[a]
Time-dependent AUC at 24 months	0.85 (0.76-0.95)	0.89 (0.68-0.99)	0.90 (0.79-1.02)	-[a]	-[a]	-[a]
No Cirrhosis (N 289)						
Harrell's C-statistics, apparent	0.75	0.81	0.80	0.88	0.73	0.79
Time-dependent AUC at 12 months	0.75 (0.68-0.82)	0.84 (0.78-0.90)	0.83 (0.77-0.89)	0.91 (0.86-0.96)	0.81 (0.71-0.92)	0.85 (0.80-0.91)
Time-dependent AUC at 24 months	0.78 (0.7-0.86)	0.82 (0.74-0.90)	0.87 (0.82-0.92)	0.92 (0.86-0.97)	0.70 (0.60-0.80)	0.77 (0.68-0.87)
Cirrhosis (N 152)						
Harrell's C-statistics, apparent	0.81	0.88	0.76	0.86	0.72	0.78
Time-dependent AUC at 12 months	0.84 (0.76-0.92)	0.91 (0.86-0.97)	0.83 (0.75-0.91)	0.87 (0.72-1.02)	0.74 (0.62-0.86)	0.86 (0.72-0.96)
Time-dependent AUC at 24 months	0.80 (0.69-0.90)	0.88 (0.80-0.96)	0.75 (0.63-0.87)	0.80 (0.66-0.93)	0.71 (0.56-0.86)	0.74 (0.64-0.89)
ALP/ULN≤3 (N 339)						
Harrell's C-statistics, apparent	0.74	0.81	0.76	0.86	0.71	0.81
Time-dependent AUC at 12 months	0.76 (0.70-0.82)	0.85 (0.80-0.90)	0.81 (0.75-0.87)	0.89 (0.83-0.94)	0.73 (0.63-0.84)	0.84 (0.80-0.88)
Time-dependent AUC at 24 months	0.80 (0.73-0.87)	0.83 (0.76-0.89)	0.86 (0.81-0.92)	0.89 (0.81-0.96)	0.72 (0.64-0.80)	0.80 (0.73-0.88)
ALP/ULN>3 (N 102)						
Harrell's C-statistics, apparent	0.72	0.84	0.70	0.80	0.62	0.70
Time-dependent AUC at 12 months	0.75 (0.58-0.92)	0.90 (0.79-0.99)	0.69 (0.55-0.83)	0.82 (0.72-0.92)	-[a]	-[a]
Time-dependent AUC at 24 months	0.67 (0.49-0.85)	0.88 (0.77-0.99)	0.71 (0.52-0.89)	0.84 (0.71-0.96)	-[a]	-[a]

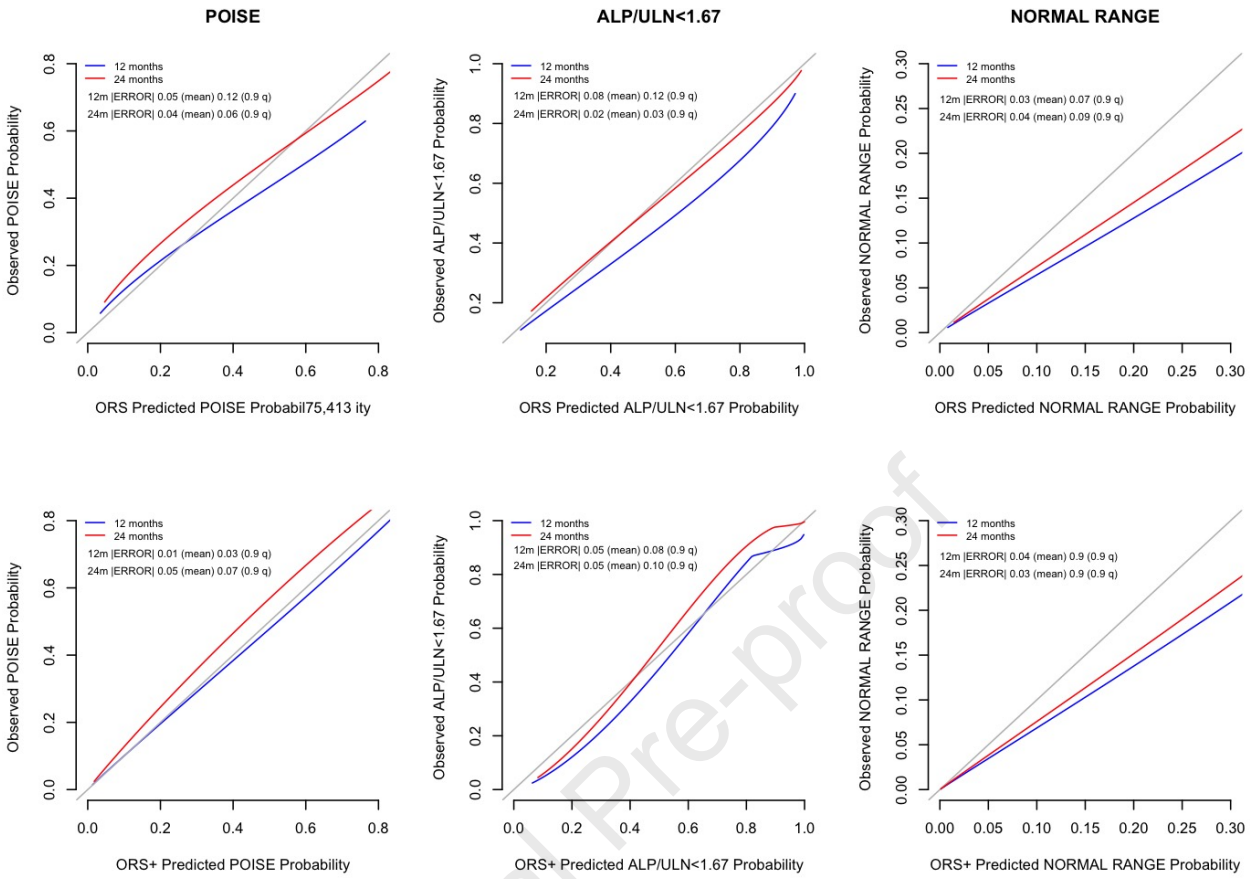
Only OCA+UDCA** (N 410)						
Harrell's C-statistics, apparent	0.77	0.84	0.80	0.89	0.72	0.81
Time-dependent AUC at 12 months	0.80 (0.74-0.85)	0.88 (0.84-0.92)	0.85 (0.80-0.89)	0.94 (0.90-0.97)	0.77 (0.68-0.86)	0.86 (0.84-0.89)
Time-dependent AUC at 24 months	0.80 (0.73-0.86)	0.85 (0.80-0.91)	0.83 (0.77-0.88)	0.91 (0.86-0.95)	0.70 (0.61-0.79)	0.8 (0.73-0.88)

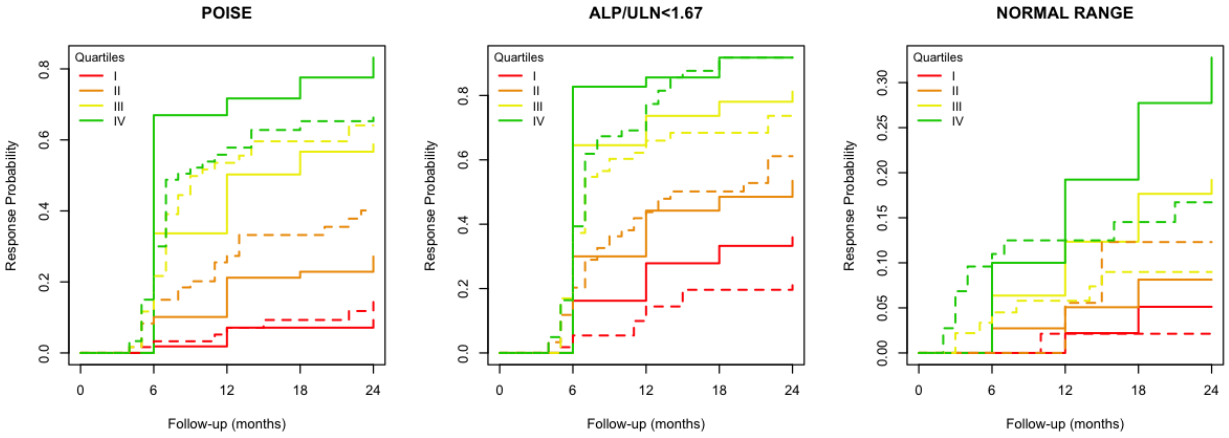
* Determined by bootstrapping 300 samples of the derivation data. ** Excluding 9 subjects with intolerance to UDCA and 22 subjects who started Fibrate before OCA and continued it during OCA therapy.
-[a] inaccurate estimates, not reported, due to few observed response events during follow-up.

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the curve; CI, confidence intervals; GGT, gamma glutamyl transferase; PBC,/AIH primary biliary cholangitis/autoimmune hepatitis; NR, normal range; OCA, obeticholic acid; ORS, OCA response score; sbjs, subjects; UDCA, ursodeoxycholic acid; ULN, upper limit of normal.

Figure 1. External calibration of the OCA Response Score (ORS and ORS+) for the occurrence of response according to POISE, ALP/ULN<1.67 and NORMAL RANGE criteria in the v cohort, at 12 and 24 months of OCA therapy. Reported curves are for the observed *vs* predicted response probabilities. The absolute error in prediction is reported as mean and 90th quantile. Calibration is reported for ORS (upper panels) and for ORS+ (lower panels).

Figure 2. Response probability to Obeticholic acid (OCA) according to quartiles of the OCA response score (ORS). Cumulative incidence curves (upper panels) for OCA response in the derivation (*solid lines*) and validation cohorts (*dashed lines*), and accompanying hazard ratios (lower panel) between the risk groups.





	POISE		ALP/ULN<1.67		NORMAL RANGE	
ORS QUARTILES	DERIVATION	VALIDATION	DERIVATION	VALIDATION	DERIVATION	VALIDATION
I	ref	ref	ref	ref	ref	ref
II	3.31 (1.52-7.23)	2.42 (1.25-4.68)	1.74 (1.13-2.69)	2.81 (1.58-4.99)	1.47 (0.47-4.62)	1.47 (0.25-8.79)
III	9.65 (4.76-19.54)	5.07 (2.72-9.47)	4.37 (2.92-6.55)	4.22 (2.41-7.4)	4.2 (1.58-11.13)	2.31 (0.5-10.68)
IV	19.2 (9.66-38.18)	5.81 (3.1-10.9)	6.69 (4.47-10.02)	6.69 (3.83-11.68)	6.58 (2.55-17.01)	4.88 (1.13-21.03)

Risk estimates are Hazard Ratios with 95% confidence intervals

Supplementary Materials

Development and validation of a scoring system to predict response to obeticholic acid in primary biliary cholangitis.

Index

Description of the study cohorts

Derivation cohort – RECAPITULATE

The "REal-world obetiCholic Acid theraPy in ITaly recapitULATion of Efficacy and safety" (RECAPITULATE) is a prospective study collecting data from centers belonging to the Italian PBC Registry and/or the Club Epatologi Ospedalieri (CLEO) and/or the Associazione Italiana Gastroenterologi e Endoscopisti Digestivi Ospedalieri (AIGO) PBC group, the Sicilian PBC Network and PBC Project Piemonte-Liguria-Valle D'Aosta. All adult patients with PBC consecutively starting OCA in 51 Italian centres from September 2017 to February 2022 were included. Diagnosis of PBC and of PBC/AIH overlap was established according to EASL guidelines[1]. The cohort included also patients with histologically defined PBC/autoimmune hepatitis (AIH) overlap syndrome and on a stable immunosuppressive therapy for at least 6 months. Patients who had been previously enrolled in a sponsored trial with OCA were excluded. Presence of cirrhosis was defined by: 1) liver histology and/or 2) liver stiffness measurement assessed by vibration-controlled transient elastography ≥16.9 kPa and/or 3) radiological (surface nodularity, caudate lobe hypertrophy, enlarged spleen or other sign of portal hypertension at ultrasound scan), and/or clinical features (presence of gastro-esophageal varices or previous decompensating events, such as ascites, variceal bleeding, encephalopathy) eventually supported by laboratory findings (low platelets, low albumin, prolonged prothrombin time)[2–4]. Eligibility for OCA treatment was based on physician judgment and on Italian prescriptive rules: ALP/ULN>1.5 and/or total bilirubin more than 1 but less than 2 after at least 12 months of treatment with UDCA, or intolerance to UDCA. The administration and dosage of OCA therapy was managed independently by each physician, based on patient characteristics and the package insert.

Data collection was opened from February till July 2022. Data capture was performed through informatized case record forms, completed by physicians in each participating centre. Demographic, clinical and biochemical data were collected at baseline (immediately before starting OCA therapy), and every 6 months of OCA therapy up to July 31st 2022. OCA dose adjustment and OCA discontinuation were collected. Pruritus was systematically assessed at baseline and at every follow-up visit. Other adverse events were not systematically assessed but registered when they led to permanent drug discontinuation. Data underwent quality control for completeness and accuracy at the University of Milan - Bicocca, Milan and University Campus Bio Medico, Rome. Missing, inaccurate or implausible data were systematically queried with the treating physicians. The study was conducted in accordance with the Declaration of Helsinki guidelines and the principles of good clinical practice. The study was approved by the University of Milan-Bicocca research ethics committee (Study name: PBC322), coordinator of the Italian National Registry and by the Research and Development Department of each collaborating hospital.

The complete RECAPITULATE cohort included 487 patients. After excluding 10 subject with not available baseline ALP and 36 without at least 6 months of follow-up, a sample of 441 individuals was used to derive the OCA response score.

Validation cohort – IBER-PBC

The IBER-PBC cohort is a prospective, observational, multicentre, real practice study collecting data on patients with PBC from 25 institutions in Spain and Portugal.[5] All adults patients prescribed with OCA in the participating centres were included. Diagnosis of PBC was made in presence of intrahepatic cholestasis with positive AMA at a titre ≥ 1:80 or, in case of negative AMA, with positive ANA (positivity for GP210 and/or SP100 antibodies) or with a liver biopsy suggestive of PBC.[1] Diagnosis of PBC/AIH overlap was established according to EASL guidelines.[1] Patients achieving response by the Paris II criteria at baseline regardless of liver fibrosis stage, with intolerance to UDCA, previous liver transplantation transplanted patients, or pregnant women were excluded. Presence of liver cirrhosis was ascertained with 1) liver histology and/or 2) liver stiffness measurement assessed by vibration-controlled transient elastography ≥16.9 kPa and/or 3) radiological, clinical and laboratory features.[2–4] Eligible patients were consecutive patients with PBC not responding to UDCA according to Paris II Criteria (*i.e.* patients with ALP ≥1.5x ULN or ALT ≥1.5x ULN or bilirubin ≥1 mg/dl) who received OCA-based therapy as second line treatment. Demographic, clinical and biochemical data were collected at baseline (immediately before starting OCA therapy) and at each visit thereafter. All patients underwent visits every 3-6 months at local investigator discretion. Blood count, liver biochemistry including aminotransferases, ALP, GGT, serum bilirubin, IgG and IgM were determined at baseline and at each visit. The occurrence of adverse events were monitored at each visit. Pruritus at baseline or during follow-up was assessed by verbal rating scale (VRS; mild, moderate or severe). Discontinuation of OCA was also collected. The study was conducted according to the principles of the updated declaration of

Helsinki and approved by the Institutional Research Board of the corresponding centres, in accordance with local regulations. A total of 244 subjects with available baseline ALP and with at least 6 months' follow-up were used as validation cohort.

Variable selection and model development

Missing values

In the derivation cohort, the following variables presented >5% missing values: platelet count (N missing 132, 30%), serum albumin (N missing 190, 43%), prothrombin time (N missing 188, 43%), serum creatinine (N missing 189, 43%) and body mass index (N missing 99, 22%). Due to the high missingness rate, these variables were not considered for model derivation procedures. Conversely, missing values for AST/ULN (N missing 8, 1.8%), ALT/ULN (N missing 7, 1.6%), GGT/ULN (N missing 10, 2.3%) and total bilirubin (N missing 9, 2%) were imputed with random forests using the *missRanger* package[6] in R-software. In the validation IBER-PBC cohort AST/ULN (N missing 1, 0.4%), GGT/ULN (N missing 11, 4.5%) and total bilirubin (N missing 1, 0.4%) were imputed with the same method. Missing values were predicted based on these same variables as well as age at OCA start, sex, ALP/ULN, diabetes mellitus and presence of cirrhosis.

Model building

Finally, 13 candidate variables at baseline (sex, age, disease duration, diabetes, UDCA not treated, PBC/AIH overlap, cirrhosis, ALP/ULN, AST/ULN, ALT/ULN, GGT/ULN and total bilirubin), and other two collected after 6 months of OCA therapy (ALP/ULN and total bilirubin) were considered for score derivation.

Model building procedures were performed as described in Ewout W. Steyemberg's (in particular Chapter 23) and Frank E. Harrell's textbooks.[7,8] Two types of risk prediction models for each outcome were developed: one based only on variables collected at OCA start (ORS), and one possibly including also the values of ALP/ULN and total bilirubin after 6 months of OCA therapy (ORS+). To note, the ORS+ was derived in the subset of subjects not responding/censored at 6 months (N 264, N 190 and N 354 for POISE, ALP/ULN<1.67 and NR criteria).

Candidate variables were tested by uni- and multivariable Cox regression analyses with OCA response criteria (POISE, ALP/ULN<1.67, NORMAL RANGE criteria) as outcomes. The proportional hazards assumption of the Cox models was checked using Schoenfeld residuals, and no violations were detected.

Since highly correlated with values at OCA start (Spearman's ρ 0.63 and 0.76, respectively), ALP/ULN and total bilirubin after 6 months were expressed as relative change from baseline [(value at 6 months - value at baseline)/value at baseline]. Possible non-linear relationships between continuous predictors and the log hazard of the outcomes were checked and visually explored by means of restricted cubic splines. In case of manifest non-linearity, different types of variable transformation were tested. The optimal variable transformation was chosen when maximizing the model goodness-of-fit, as evaluated by the Wald χ^2 (Supplementary table 1).

For POISE response, we modelled ALP/ULN with a restricted cubic spline (4 knots, Supplementary Figure 1), which showed a significantly improved χ^2 of 42 compared to the linear fit (χ^2 29) and to other types of transformation. For ALP/ULN<1.67 and NORMAL RANGE response, ALP/ULN was modelled with a restricted cubic spline (3 knots, Supplementary Figure 1), displaying the best model fit. Non linearity was also found for ALT/ULN and AST/ULN. By applying restricted cubic spline, we visually detected a decrease of the log hazard of POISE and ALP/ULN<1.67 response and a plateau for values of 1.5 and higher (Supplementary Figure 1). We then fitted a linear model up to the value of 1.5 and then a plateau, finding χ^2 values similar to restricted cubic spline. With similar χ^2 values, the transformation with lower degrees of freedom was preferred and a linear fit up to 1.5 was considered for ALT/ULN and AST/ULN for both outcomes. For NORMAL RANGE, the threshold effect at 1.5 for ALT/ULN and AST/ULN was less evident, and a natural logarithmic transformation showed the best model fit with the lowest degrees of freedom.

For age at OCA start, the linear fit resulted in a model χ^2 of 0.53 (POISE response), of 0.67 (ALP/ULN<1.67 response), and of 0.3 (NORMAL RANGE response). If we fit a model with restricted cubic spline (4k, 3 df), the χ^2 raised to 1.33, 3.13 and 1.01, respectively. The difference between the linear and the spline fit was 0.8, 2.47 and 0.71 which was not significant at 2 df (3 - 1 df, p values of 0.67, 0.29 and 0.72, respectively). As

such, the linear fit was considered to age at OCA start. A similar approach was also applied for disease duration and GGT/ULN (Suppl Fable 1).

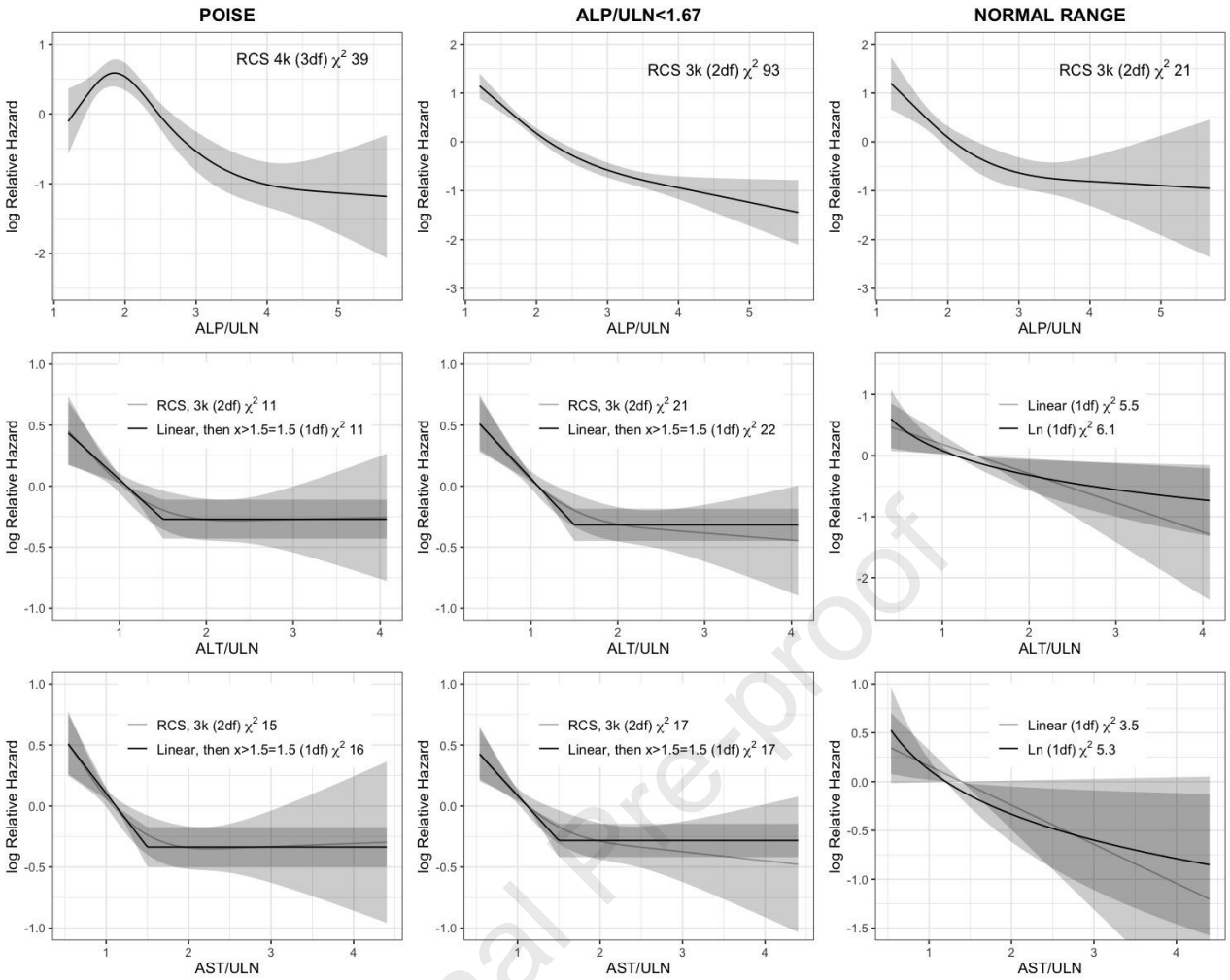
Supplementary Table 1. Impact of various codings of continuous predictors in univariate Cox regression models.

	Response Criteria	**POISE**	**ALP/ULN<1.67**	**NORMAL RANGE**
Predictor	**Coding**	Wald χ^2 (*df*)	Wald χ^2 (*df*)	Wald χ^2 (*df*)
ALP/ULN	Linear	27 (1)	66 (1)	12 (1)
	RCS, 3k	23 (2)	**93 (2)**	**21 (2)**
	RCS, 4k	**39 (3)**	90 (3)	19 (3)
	Log	26 (1)	87 (1)	17 (1)
	Sqrt	27 (1)	77 (1)	15 (1)
	Quadratic	27 (2)	66 (2)	12 (1)
	x<2.1=0, then x-2.1	26 (1)	47 (1)	8 (1)
	x<2.1=0, then (x-2.1)^-1	31 (1)	65 (1)	11 (1)
	x<2.1=0, then (x-2.1)^-2	31 (1)	70 (1)	12 (1)
Total Bilirubin	Linear	**39 (1)**	**12 (1)**	**9 (1)**
	RCS,3k	31 (2)	11 (2)	7 (2)
	RCS,4k	28 (3)	11 (3)	6 (3)
	Sqrt	37 (1)	11 (1)	8 (1)
	Ln	37 (1)	11 (1)	7 (1)
	Quadratic	39 (2)	12 (2)	7 (1)
Age at OCA start	Linear	**0.53 (1)**	**0.67 (1)**	**0.3 (1)**
	RCS,3k	0.85 (2)	0.99 (2)	0.57 (2)
	RCS,4k	1.33 (3)	3.13 (3)	1.01 (3)
	Sqrt	0.46 (1)	0.53 (1)	0.31 (1)
	Ln	0.39 (1)	0.62 (1)	0.31 (1)
	Quadratic	0.53 (2)	0.47 (2)	0.3 (1)
Disease Duration	Linear	**0.16 (1)**	**1.23 (1)**	**0.41 (1)**
	RCS,3k	0.49 (2)	1.23 (2)	1.10 (2)
	RCS,4k	0.95 (3)	1.27 (3)	1.15 (3)
	Sqrt	0.03 (1)	1.37 (1)	0.27 (1)
	Ln	0.02 (1)	1.15 (1)	0.19 (1)
	Quadratic	0.16 (2)	1.23 (2)	0.41 (1)
ALT/ULN	Linear	5 (1)	13 (1)	5.5 (1)
	RCS,3k	11 (2)	21 (2)	5.5 (2)
	RCS,4k	10 (3)	21 (3)	5.5 (3)
	Ln	9 (1)	19 (1)	**6.1 (1)**
	Sqrt	7 (1)	17 (1)	6.0 (1)
	Quadratic	5 (2)	13 (2)	5.5 (1)
	linear, then x>1.5=1.5	**11 (1)**	**22 (1)**	5.4 (1)
AST/ULN	Linear	7 (1)	11 (1)	3.5 (1)
	RCS,3k	15 (2)	17 (2)	5.8 (2)
	RCS,4k	15 (3)	17 (3)	6.4 (3)
	Ln	12 (1)	15 (1)	**5.3 (1)**
	Sqrt	10 (1)	14 (1)	4.6 (1)
	Quadratic	7 (2)	11 (2)	3.5 (1)
	linear, then x>1.5=1.5	**16 (1)**	**17 (1)**	5.6 (1)
GGT/ULN	Linear	**0.5 (1)**	**7.27 (1)**	**1.31 (1)**

	RCS,3k	0.51 (2)	9.68 (2)	2.82 (2)
	RCS,4k	2.00 (3)	11.08 (3)	3.83 (3)
	Ln	0.31 (1)	9.22 (1)	0.84 (1)
	Sqrt	0.4 (1)	8.51 (1)	0.99 (1)
	Quadratic	0.5 (2)	7.27 (2)	1.01 (1)
Relative change ALP/ULN at 6 months	Linear	**27 (1)**	**20 (1)**	**10 (1)**
	RCS,3k	24 (2)	20 (2)	12 (2)
	RCS,4k	20 (3)	20 (3)	12 (3)
	Sqrt	4 (1)	5 (1)	3 (1)
	Quadratic	27 (2)	20 (2)	10 (1)
	x<-0.2=-0.2, then linear	20 (1)	19 (1)	2 (1)
Relative change Total bilirubin at 6 months	Linear	**27 (1)**	**20 (1)**	**0.3 (1)**
	RCS,3k	24 (2)	20 (2)	0.7 (2)
	RCS,4k	20 (3)	20 (3)	1.5 (3)
	Sqrt	4 (1)	5 (1)	0.1 (1)
	Quadratic	27 (2)	20 (2)	0.3 (1)
	x<0=0, then linear	20 (1)	19 (1)	0.1 (1)

Relative change of ALP/ULN and total bilirubin calculated as (value at 6 months – value at baseline)/value at baseline.
Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; df, degree of freedom; k, knot; RCS, restricted cubic spline.

Supplementary Figure 1. Unadjusted relations between continuous variables and the log hazard of response to OCA according to different response criteria. Only variables showing non-linear relationships with outcomes have been displayed according to their optimal (black lines) and nearly optimal fit (grey lines). *Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; df, degree of freedom; k, knot; Ln, natural logarithm; RCS, restricted cubic spline.*



Stepwise variable selection

Full multivariable Cox regression models were firstly fitted with all the candidate variables (Supplementary table 2). Then, parsimonious models were obtained by means of automated backward stepwise selection procedures using the Akaike Information Criteria (AIC, the lower the better) as stopping rule, and the Wald χ^2 of individual variables as the statistics on which to base the stopping rule (Supplementary table 2). Finally, a penalized maximum likelihood estimation was used to account for overfitting.[9]

The following variables were retained in the final models for

ORS$_{POISE}$: age at OCA start, pruritus at OCA start, ALP/ULN, cirrhosis, total bilirubin, ALT/ULN, GGT/ULN

ORS$_{ALP/ULN<1.67}$: pruritus at OCA start, ALP/ULN, cirrhosis, total bilirubin, GGT/ULN

ORS$_{NORMAL\ RANGE}$: pruritus at OCA start, ALP/ULN, total bilirubin

The relative change ([value at 6 months – value at baseline]/value at baseline) of ALP/ULN and/or of total bilirubin after 6 months of OCA therapy were subsequently added for the derivation -in the subset of subjects not responding/censored at 6 months- of an updated ORS (ORS+), if improving the goodness-of-fit:

ORS+$_{POISE}$: age at OCA start, pruritus at OCA start, ALP/ULN, cirrhosis, total bilirubin, ALT/ULN, GGT/ULN + relative change of ALP/ULN at 6 months and relative change of total bilirubin at 6 months

ORS+$_{ALP/ULN<1.67}$: pruritus at OCA start, ALP/ULN, cirrhosis, total bilirubin, GGT/ULN + relative change of ALP/ULN at 6 months and relative change of total bilirubin at 6 months

ORS+$_{NORMAL\ RANGE}$: pruritus at OCA start, ALP/ULN, total bilirubin + relative change of ALP/ULN at 6 months and relative change of total bilirubin at 6 months

Supplementary table 2. Multivariable Cox proportional hazard models for response to Obeticholic acid in the derivation cohort (RECAPITULATE), according to different criteria.

Outcome	POISE RESPONSE			ALP/ULN<1.67 RESPONSE			NORMAL RANGE RESPONSE		
Score		ORS_{POISE}	$ORS+_{POISE}$		$ORS_{ALP/ULN<1.67}$	$ORS+_{ALP/ULN<1.67}$		$ORS_{NORMALRANGE}$	$ORS+_{NORMALRANGE}$
Model	*Full*	*Reduced*	*Reduced 2*	*Full*	*Reduced*	*Reduced 2*	*Full*	*Reduced*	*Reduced 2*
Predictor	aHR (95%CI)	aHR (95%CI), χ^2	aHR (95%CI), χ^2	aHR (95%CI)	aHR (95%CI), χ^2	aHR (95%CI), χ^2	aHR (95%CI)	aHR (95%CI), χ^2	aHR (95%CI), χ^2
Sex, female	1.09 (0.69-1.74)			0.83 (0.58-1.19)			2.28 (0.80-6.52)		
Age at OCA start	0.81 (0.63-1.05)	0.79 (0.62-1.01), 3.7	0.85 (0.67-1.09), 2	0.98 (0.79-1.22)			0.83 (0.53-1.31)		
Duration of disease before OCA start	0.86 (0.70-1.05)			0.99 (0.83-1.19)			0.80 (0.54-1.17)		
Diabetes mellitus	0.77 (0.37-1.6)			0.80 (0.47-1.37)			1.43 (0.48-4.23)		
UDCA not treated	0.43 (0.10-1.78)			0.41 (0.15-1.14)			1.53 (0.32-7.29)		
PBC/AIH overlap	0.92 (0.56-1.49)			1.00 (0.68-1.49)			1.19 (0.48-2.94)		
Pruritus at baseline	0.48 (0.33-0.70)	0.52 (0.36-0.76), 11	0.70 (0.48-1.03), 3	0.61 (0.46-0.83)	0.59 (0.44-0.79), 12	0.73 (0.55-0.98), 4	0.60 (0.32-1.16)	0.64 (0.34-1.21), 2	0.87 (0.46-1.66), 2
ALP/ULN at baseline	0.34 (0.22-0.54)	0.43 (0.28-0.66), 27	0.42 (0.28-0.64), 42	0.32 (0.25-0.42)	0.33 (0.26-0.42), 84	0.21 (0.16-0.27), 150	0.27 (0.16-0.48)	0.34 (0.20-0.56), 18	0.15 (0.08-0.29), 31
Cirrhosis	0.79 (0.56-1.13)	0.77 (0.55-1.07), 2.4	0.74 (0.53-1.05), 3	0.81 (0.61-1.06)	0.80 (0.61-1.04), 2.7	0.82 (0.63-1.07), 2	0.66 (0.35-1.23)		
Total Bilirubin/ULN at baseline	0.49 (0.38-0.63)	0.52 (0.41-0.67), 26	0.50 (0.38-0.66), 24	0.77 (0.66-0.91)	0.78 (0.67-0.91), 9.4	0.84 (0.71-1), 4	0.53 (0.34-0.82)	0.53 (0.35-0.80), 9	0.5 (0.32-0.79), 9
ALT/ULN at baseline	0.84 (0.54-1.30)	0.82 (0.58-1.16), 1.2	0.77 (0.54-1.11), 2	0.76 (0.52-1.12)			0.70 (0.39-1.26)		
AST/ULN at baseline	0.87 (0.55-1.38)			1.28 (0.86-1.92)			1.08 (0.59-1.99)		
GGT/ULN at baseline	1.38 (1.17-1.63)	1.30 (1.10-1.52), 10	1.23 (1.05-1.45), 6	1.12 (0.97-1.30)	1.09 (0.96-1.25), 1.8	1.01 (0.88-1.15), 0.9	1.28 (0.92-1.77)		
Change ALP/ULN at 6 months, relative			0.43 (0.33-0.57), 37			0.43 (0.35-0.52), 73			0.22 (0.12-0.4), 25
Change Total Bilirubin at 6 months, relative			0.88 (0.73-1.06), 2						

Reduced models 1 (ORS) were obtained from full models with automated backward selection procedure using the Akaike Information Criteria as stopping rule, and the Wald $\chi 2$ of individual variables as the statistics on which to base the stopping rule. Reduced models 2 were fitted only in subjects not responding/censored at 6 months (N 264, N 190 and N 354 for POISE, ALP/ULN<1.67 and NR), with the addition to reduced models 1 of the relative change of ALP/ULN and of total bilirubin at 6 months of OCA therapy. The relative change is calculated as [(value at 6 months – value at baseline) / value at baseline]. A penalized maximum likelihood estimation was used to account for overfitting. Variables have been transformed as detailed in Supplementary table 1, and hazard ratios (HR) are reported for the comparison of the third vs first quartile for continuous variables, and for categories for categorical variables. Wald $\chi 2$ is reported for indicating the contribution of each variable in the predictive scores.

Abbreviations: AIH, autoimmune hepatitis; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma glutamyl transferase; ORS, OCA response score; OCA, Obeticholic acid; PBC, primary biliary cholangitis; UDCA, ursodeoxycholic acid; ULN, upper limit of normal to check

Collinearity and interactions between variables

Collinearity and multicollinearity between variables were explored in the final models by computing the Spearman ρ correlation coefficients and the variance inflation factor (VIF). Correlation coefficients were <0.6, and VIF<2, demonstrating no significant collinearity/multicollinearity (Supplementary table 3). Similarly, no interactions were evident between variables in the final models (Supplementary table 3).

Supplementary table 3. Two-way interactions, correlations and multicollinearity between variables included in the final model for the OCA response score (ORS).

POISE RESPONSE

	Age at OCA start	Pruritus at OCA start	ALP/ ULN	Cirrhosis	Total bilirubin	ALT/ULN	GGT/ULN	Rel change ALP/ULN at 6 months	Rel change total bilirubin at 6 months
Age at OCA start	-	P interact 0.57	P interact 0.57	P interact 0.77	P interact 0.5	P interact 0.19	P interact 0.19	P interact 0.7	P interact 0.7
Pruritus at OCA start	Spearman rho -0.11	-	P interact 0.72	P interact 0.21	P interact 0.91	P interact 0.69	P interact 0.06	P interact 0.32	P interact 0.77
ALP/ULN	Spearman rho -0.09	Spearman rho 0.14	-	P interact 0.47	P interact 0.19	P interact 0.32	P interact 0.19	P interact 0.02	P interact 0.95
Cirrhosis	Spearman rho 0.18	Spearman rho 0.05	Spearman rho 0.07	-	P interact 0.15	P interact 0.31	P interact 0.85	P interact 0.9	P interact 0.78
Total bilirubin	Spearman rho -0.07	Spearman rho -0.01	Spearman rho 0.07	Spearman rho 0.22	-	P interact 0.16	P interact 0.81	P interact 0.58	P interact 0.38
ALT/ULN	Spearman rho -0.3	Spearman rho 0.19	Spearman rho 0.37	Spearman rho 0.03	Spearman rho 0.24	-	P interact 0.31	P interact 0.58	P interact 0.27
GGT/ULN	Spearman rho -0.09	Spearman rho 0.09	Spearman rho 0.43	Spearman rho 0.11	Spearman rho 0.15	Spearman rho 0.55	-	P interact 0.01	P interact 0.75
Rel change ALP/ULN at 6 months	Spearman rho 0.12	Spearman rho 0.07	Spearman rho -0.37	Spearman rho -0.04	Spearman rho 0.01	Spearman rho -0.15	Spearman rho -0.23	-	P interact 0.63
Rel change total bilirubin at 6 months	Spearman rho 0.12	Spearman rho -0.05	Spearman rho -0.01	Spearman rho 0.14	Spearman rho -0.26	Spearman rho -0.07	Spearman rho -0.1	Spearman rho 0.05	-
VIF									
ORS	1.14	1.04	1.23	1.04	1.05	1.48	1.47		
ORS+	1.15	1.07	1.34	1.08	1.15	1.53	1.46	1.3	1.08

ALP/ULN<1.67 RESPONSE

	Pruritus at OCA start	ALP/ ULN	Cirrhosis	Total bilirubin	GGT/ULN	Rel change ALP/ULN at 6 months	Rel change total bilirubin at 6 months
Pruritus at OCA start	-	P interact 0.13	P interact 0.32	P interact 0.82	P interact 0.32	P interact 0.78	P interact 0.11
ALP/ULN	Spearman rho 0.14	-	P interact 0.29	P interact 0.65	P interact 0.53	P interact 0.20	P interact 0.94
Cirrhosis	Spearman rho 0.05	Spearman rho 0.07	-	P interact 0.47	P interact 0.42	P interact 0.35	P interact 0.58
Total bilirubin	Spearman rho -0.01	Spearman rho 0.07	Spearman rho 0.22	-	P interact 0.92	P interact 0.55	P interact 0.78
GGT/ULN	Spearman rho 0.09	Spearman rho 0.43	Spearman rho 0.11	Spearman rho 0.15	-	P interact 0.32	P interact 0.97
Rel change ALP/ULN at 6 months	Spearman rho 0.07	Spearman rho -0.37	Spearman rho -0.04	Spearman rho 0.01	Spearman rho -0.23	-	P interact 0.54
Rel change total bilirubin at 6 months	Spearman rho -0.05	Spearman rho -0.01	Spearman rho 0.14	Spearman rho -0.26	Spearman rho -0.1	Spearman rho 0.05	-
VIF							
ORS	1.01	1.11	1.04	1.05	1.11		
ORS+	1.03	1.14	1.09	1.13	1.12	1.14	1.09

NORMAL RANGE RESPONSE

	Pruritus at OCA start	ALP/ ULN	Total bilirubin	Rel change ALP/ULN at 6 months	Rel change total bilirubin at 6 months
Pruritus at OCA start	-	P interact 0.89	P interact 0.40	P interact 0.72	P interact 0.52
ALP/ULN	Spearman rho 0.14	-	P interact 0.75	P interact 0.63	P interact 0.95

Total bilirubin	Spearman rho -0.01	Spearman rho 0.07	-	P interact 0.81	P interact 0.96
Rel change ALP/ULN at 6 months	Spearman rho 0.07	Spearman rho -0.37	Spearman rho 0.01	-	P interact 0.54
Rel change total bilirubin at 6 months	Spearman rho -0.05	Spearman rho -0.01	Spearman rho -0.26	Spearman rho 0.05	-
VIF					
ORS	1.01	1.01	1.00		
ORS+	1.03	1.28	1.06	1.30	1.05

P for interaction are from Wald χ^2 tests.
Abbreviations: ALP, alkaline phosphatase; GGT, gamma-glutamyl transferase; interact, interaction; ORS, OCA response score; ULN, upper limit of normal; Rel, Relative; VIF, variance inflation factor.

OCA response score (ORS) formula.
The ORS/ORS+ were calculated as the sum of the variables included in the final models (Supplementary Table 2), weighted for their β regression coefficients (linear predictor, βX). The ORS/ORS+ were centered on the mean in the derivation cohort. Predicted response probabilities at 12 and 24 months of OCA therapy were then estimated using baseline survival estimates $S_0(t)$, where t is the time at which predicting OCA response, using the formula: $1 - S_0(t)^{e^{\beta X}}$. Full formulas for computing ORS/ORS+ according to different OCA response criteria, along with $S_0(t)$ estimates at 12 and 24 months are reported in Supplementary table 4. An online calculator can be found at https://ocaresponsescore.github.io/calculator/.

Supplementary table 4. ORS/ORS+ formulas and baseline survival estimates.

	OCA response score (ORS)	
Outcome	**ORS**	**ORS+**
POISE	1.5530021 -0.014619049*Age at OCA start +0.64567037*Pruritus +0.68734539*ALP/ULN -1.3477237*$(ALP/ULN-1.38)_+^3$ +2.3394176*$(ALP/ULN-1.8)_+^3$ -1.0333051*$(ALP/ULN-2.471429)_+^3$ +0.041611269*$(ALP/ULN-4.87)_+^3$ -0.26753893*Cirrhosis -1.2938781*TotalBilirubin −0.26917078*min(ALT/ULN,1.5) +0.063070033*GGT/ULN and $(x)_+ = x$ if $x > 0$, 0 otherwise	If ALP/ULN6months <1.67 & (ALP/ULN6months- ALP/ULN)/ ALP/ULN < -0.15 & TotalBilirubin6months ≤1, then POISE criterium is already attained, else calculate: 2.9645625-0.010079541*Age at OCA start -0.35115548*Pruritus -0.69673211* ALP/ULN -0.29445138*Cirrhosis -1.37536*TotalBilirubin −0.34807907*min(ALT/ULN,1.5) +0.050888012*GGT/ULN -2.9107294*[(ALP/ULN6months – ALP/ULN)/ALP/ULN] −0.34485537*[(TotalBilirubin6months – TotalBilirubin)/TotalBilirubin] and $(x)_+ = x$ if $x > 0$, 0 otherwise
ALP/ULN<1.67	3.2969418 -0.52175514*Pruritus - 1.3061516*ALP/ULN +0.2527601*$(ALP/ULN-1.5)_+^3$ -0.32479018*$(ALP/ULN-2.05)_+^3$ +0.07203008*$(ALP/ULN-3.98)_+^3$ -0.22389756*Cirrhosis - 0.49623968*TotalBilirubin +0.021747561*GGT/ULN	If ALP/ULN6months <1.67, then the ALP/ULN<1.67 criterium is already attained, else calculate 3.6024346 -0.31340847*Pruritus 1.8927908*ALP/ULN +0.37438631*$(ALP/ULN-1.5)_+^3$ -0.4810767*$(ALP/ULN-2.05)_+^3$

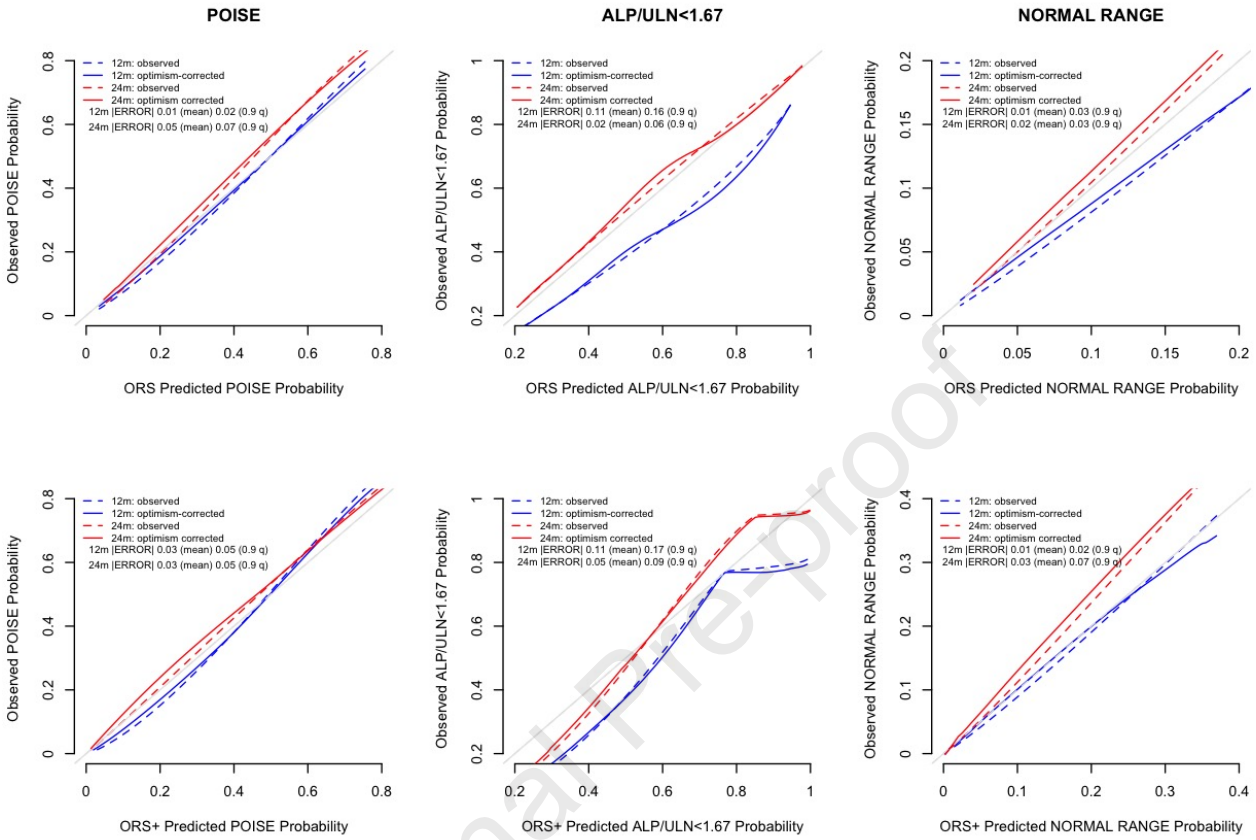
	and $(x)_+ = x$ if $x > 0$, 0 otherwise	$+0.1066904*(ALP/ULN - 3.98)_+^3$ -0.1994449*Cirrhosis $-0.34343259*TotalBilirubin$ $+0.0022129224*GGT/ULN$ $-2.9341961*[(ALP/ULN6months - ALP/ULN)/ALP/ULN]$ and $(x)_+ = x$ if $x > 0$, 0 otherwise
NORMAL RANGE	4.4364308 -0.44185524*Pruritus - 1.6629714*ALP/ULN $+0.4761417*(ALP/ULN - 1.5)_+^3$ $-0.61182975*(ALP/ULN - 2.05)_+^3$ $+0.13568805*(ALP/ULN - 3.98)_+^3$ $-1.2762865*TotalBilirubin$ and $(x)_+ = x$ if $x > 0$, 0 otherwise	If ALP/ULN6months ≤ 1 & ALT/ULN6months ≤ 1 & TotalBilirubin6months ≤1, then the NR criterium is already attained, else calculate 5.2190166 -0.13479947*Pruritus -2.7592853*ALP/ULN $+0.7405286*(ALP/ULN-1.5)_+^3$ $-0.95156006*(ALP/ULN-2.05)_+^3$ $+0.21103147*(ALP/ULN - 3.98)_+^3$ $-1.380134*TotalBilirubin$ $-5.1873971*[(ALP/ULN6months - ALP/ULN)/ALP/ULN]$ and $(x)_+ = x$ if $x > 0$, 0 otherwise
	Baseline Survival – $S_0(t)$	
POISE		
12 months	0.6954680	0.7026726
24 months	0.6115819	0.6179623
ALP/ULN<1.67		
12 months	0.4327055	0.3945538
24 months	0.3369205	0.2949377
NORMAL RANGE		
12 months	0.9285800	0.9551407
24 months	0.8760027	0.9183093
OCA predicted response probability $1 - S_0(t)^{e^{ORS}}$		

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; GGT, gamma-glutamyl transferase; ULN, upper limit of normal.

Internal validation.
Internal validation was performed by comparing the observed vs predicted response probability, after bootstrapping 300 sample of the derivation cohort. Absolute error in prediction were all in the range of 0.02-0.05, indicating good calibration. The prediction of ALP/ULN<1.67 at 12 months presented the highest mean error (0.11 for both ORS and ORS+). Calibration plots are reported as Supplementary Figure 2.

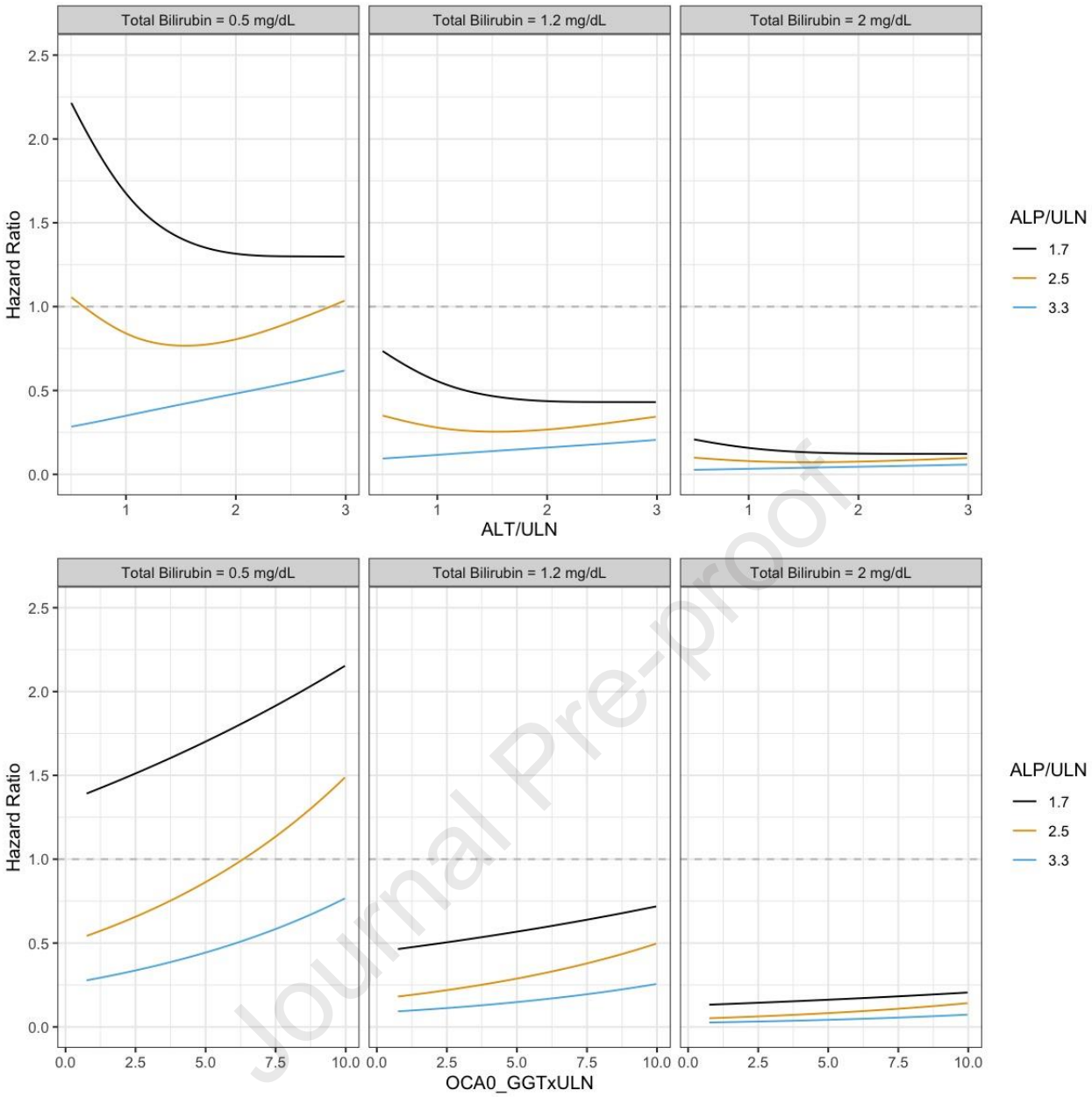
Supplementary Figure 2. Internal calibration of the OCA Response Scores (ORS and ORS+) for the occurrence of response according to POISE, ALP/ULN<1.67 and NORMAL RANGE criteria in the

derivation cohort, at 12 and 24 months of OCA therapy. Reported curves are for the observed *vs* predicted response probabilities and for optimism-corrected values, after bootstrapping 300 samples of the derivation cohort. The absolute error in prediction is reported as mean and 90th quantile.



Relation between ALT and GGT and the probability of OCA response according to POISE
Supplementary Figure 3 highlights the effect of ALP and bilirubin on the hazard of POISE response according to ALT and GGT values. An inverse association can be observed for ALT/ULN particularly at lower ALP and bilirubin levels (upper left panel). With increasing ALP/ULN values, the relation progressively gets blunted. With higher bilirubin levels (upper central and right panels), the relation of ALT with OCA response is no more evident and not influenced by ALP/ULN.
For GGT/ULN, a direct relation is observable particularly at lower ALP/ULN and bilirubin levels, which progressively declines at increasing ALP values (lower left panel). At high bilirubin levels (2 mg/dL, right panel), GGT effect on OCA response is close to flat and not influenced by ALP.

Supplementary Figure 3. Shape of the relationship between the hazard of OCA response according to POISE and ALT and GGT, at different levels of ALP/ULN and total bilirubin at baseline.



What You Need to Know

This content should include three headings (Background, Findings, Implications for patient care) with 1-2 brief sentences (25-30 words; exclude nonstandard abbreviations) per heading.

Background.

With novel second-line therapies in advanced development, a tool for predicting biochemical response to OCA is needed to tailor the treatment algorithm of PBC patients who are unresponsive or intolerant to UDCA.

Findings.

The OCA response score (ORS) was derived and externally validated for predicting biochemical response according to POISE, ALP/ULN<1.67 and normal range criteria. This incorporates the age, the presence of pruritus, a cirrhotic disease stage, the serum level of ALP, ALT, GGT and bilirubin, and the ALP and bilirubin change after 6 months of OCA therapy (https://ocaresponsescore.github.io/calculator/).

Implications for patient care.

In the evolving landscape of clinical practice, as more second-line therapies loom on the horizon, the ORS could enhance personalized treatment allocations.