

A CHI3L1-containing diagnostic model for biopsy-confirmed significant fibrosis in chronic HBV infection: a retrospective study

Jinnian Lin

The First People's Hospital of Foshan, Southern University of Science and Technology

Zicong Gu

492755446@qq.com

The First People's Hospital of Foshan, Southern University of Science and Technology

Research Article

Keywords: hepatitis B, liver fibrosis, CHI3L1, liver biopsy, transient elastography, diagnostic model

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11243785/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Aim

To evaluate the association between serum chitinase-3-like protein 1 (CHI3L1) and biopsy-confirmed significant fibrosis in chronic hepatitis B virus (HBV) infection, and to explore CHI3L1-containing diagnostic models.

Methods

This retrospective study included 81 adults who underwent liver biopsy between January 2020 and December 2024. Significant fibrosis was defined as Scheuer stage S2-S4. Participants were allocated to training (n = 65) and hold-out validation (n = 16) cohorts before missing-value imputation. HBV DNA was log₁₀-transformed. FibroScan, a serum model (prothrombin time, alkaline phosphatase, CHI3L1 and log₁₀ HBV DNA), and a combined model were assessed.

Results

Fifty-six participants had significant fibrosis. Median CHI3L1 was numerically higher with significant fibrosis [31.73 (22.23–55.80) vs 27.41 (19.86–35.44) ng/mL; P = 0.105]. In the seven-variable model, CHI3L1 showed a positive adjusted association (odds ratio 1.061, 95% confidence interval 1.004–1.121; P = 0.036). Training/validation areas under the receiver operating characteristic curve were 0.783/0.808 for FibroScan, 0.856/0.750 for the serum model and 0.873/0.800 for the combined model.

Conclusions

CHI3L1 showed a positive adjusted association, but model ranking was inconsistent in the small validation cohort. These exploratory findings require rigorous internal validation and independent prospective multicenter validation before clinical use.

Introduction

Chronic hepatitis B virus (HBV) infection remains a major cause of progressive liver disease, cirrhosis and hepatocellular carcinoma. Fibrosis stage informs prognosis, treatment eligibility and follow-up, and international and Chinese guidance recommends systematic fibrosis assessment in people with chronic HBV infection [1–3]. Long-term antiviral suppression can be accompanied by histological regression of fibrosis and cirrhosis, reinforcing the importance of assessing disease stage [4, 5].

Liver biopsy provides direct histological assessment and remains an important reference standard, but its invasive nature, sampling variability and observer variation limit repeated use [6, 7]. Non-invasive

assessment is therefore an important component of clinical practice [8]. FibroScan-derived liver stiffness has demonstrated diagnostic value in biopsy-referenced HBV cohorts and meta-analysis [9, 10]. However, liver stiffness is not specific to collagen deposition: even mild-to-moderate alanine aminotransferase elevation can influence measurements in chronic hepatitis B [11]. Acquisition reliability also depends on patient and operator factors and the dispersion of valid measurements [12, 13].

Routine blood-based scores provide accessible comparators for new biomarkers. The aspartate aminotransferase-to-platelet ratio index (APRI) was developed in chronic hepatitis C, and the fibrosis-4 index (FIB-4) in HIV/HCV coinfection [14, 15]. Subsequent HBV studies show that diagnostic performance depends on the fibrosis endpoint and threshold [16, 17]. The large biopsy-referenced synthesis informing the 2024 WHO guidance likewise found threshold-dependent trade-offs for APRI and FibroScan [18]. A new biomarker should therefore be evaluated for information added to existing tests, rather than judged solely by its standalone discrimination.

Chitinase-3-like protein 1 (CHI3L1, also known as YKL-40) is a secreted glycoprotein associated with inflammation, tissue repair and extracellular-matrix remodeling [19]. Early biopsy-based work linked serum CHI3L1 with histological fibrosis stage [20, 21]. Recent experimental work in metabolic dysfunction-associated steatohepatitis implicates CHI3L1/interleukin-13 receptor alpha 2 signaling in stellate-cell activation [22]. Experimental CHI3L1 deficiency also attenuates liver fibrosis by promoting hepatic macrophage apoptosis [23]. These observations provide biological plausibility but do not establish an HBV-specific mechanism or diagnostic threshold.

Clinical studies have evaluated serum CHI3L1 alone and in combination with routine indicators in HBV-related liver disease [24–27]. More recent studies include a biopsy-based model addressing liver injury in the immune-tolerant phase and the CHILI multicenter study of CHI3L1-containing machine-learning models [28, 29]. A general-population nomogram has also been reported [30]. Differences in population, assay platform, reference standard and model specification limit direct transfer of these findings to a particular biopsy-confirmed HBV cohort. A CHI3L1 meta-analysis reported substantial between-study heterogeneity, including variation associated with age and liver-disease etiology [31].

We therefore evaluated the association between serum CHI3L1 and biopsy-confirmed significant fibrosis in chronic HBV infection. We also explored the discrimination, calibration and clinical utility of a CHI3L1-containing serum model, FibroScan alone and a combined model.

Methods

Study design and participants

This retrospective diagnostic-modeling study included adults with chronic HBV infection who underwent liver biopsy at Foshan First People's Hospital between January 2020 and December 2024. Chronic HBV infection was diagnosed according to the 2022 Chinese guideline [3]. Eligible participants were 18–60

years of age, had alanine aminotransferase and aspartate aminotransferase values no greater than twice the upper limit of normal, and had an interpretable liver-biopsy fibrosis stage.

Exclusion criteria were concomitant liver disease (including autoimmune, drug-induced or alcohol-related liver disease, steatohepatitis or hepatocellular carcinoma), coinfection with hepatitis C virus, hepatitis D virus or human immunodeficiency virus, other malignancy, active systemic infection, severe cardiac, pulmonary, neurological or renal disease, pregnancy or lactation, massive ascites, or body mass index greater than 28 kg/m².

The analytical cohort comprised 81 participants. Before imputation and model fitting, participants were randomly allocated in an approximately 8:2 ratio to training (n = 65) and hold-out internal validation (n = 16) cohorts. The supplied analytical workbooks were complete after imputation. Reporting was prepared with reference to TRIPOD + AI [32]. The study flow and analytical framework are shown in Fig. 1.

Histopathological assessment

Ultrasound-guided percutaneous liver-biopsy specimens were required to be at least 1.5 cm long and to contain at least six portal tracts. Fibrosis was staged from S0 to S4 using the Scheuer system [33]. Significant fibrosis was defined as stage S2-S4; stages S0-S1 were classified as non-significant fibrosis.

Clinical, laboratory and elastography measurements

Demographic characteristics, body mass index, complete blood count, liver biochemical tests, coagulation indices, HBV serological markers, serum HBV DNA, serum CHI3L1 and FibroScan liver-stiffness measurements were obtained from the clinical record. Serum CHI3L1 was measured by enzyme-linked immunosorbent assay. Detailed assay-platform and acquisition-quality metadata were not available in the analytical files.

Data preprocessing and exploratory models

Random allocation was completed before missing values were imputed. HBV DNA was retained in IU/mL for descriptive reporting and log₁₀-transformed for modeling. Exploratory least absolute shrinkage and selection operator (LASSO) regression was used for screening [34]. In training data, LASSO at lambda.min retained the sex variable, red blood cell count, hemoglobin, prothrombin time (PT), alkaline phosphatase (ALP), HBeAg and CHI3L1. A seven-variable association model contained sex, red blood cell count, hemoglobin, PT, ALP, CHI3L1 and log₁₀ HBV DNA. Three diagnostic approaches were compared: FibroScan alone; a serum model containing PT, ALP, CHI3L1 and log₁₀ HBV DNA; and a combined model that added FibroScan. The cross-validation results and coefficient trajectories from exploratory LASSO screening are shown in Fig. 2.

Statistical analysis

Continuous variables were summarized as median and interquartile range and compared using the Mann-Whitney U test. Categorical variables were summarized as counts and percentages. Logistic-regression results are reported as odds ratios (ORs) with 95% confidence intervals (CIs). Discrimination was assessed with receiver operating characteristic curves and AUCs with DeLong CIs [35]. Calibration was evaluated graphically and by calibration intercept, calibration slope and Brier score [36, 37]. Decision-curve analysis was exploratory and evaluated net benefit relative to treat-all and treat-none strategies [38, 39]. All P values were two-sided; $P < 0.05$ was considered statistically significant.

Use of generative AI tools

ChatGPT and Codex (OpenAI) assisted with manuscript drafting, reference identification, code preparation and figure formatting. Numerical results were derived from the supplied analytical data; no patient observations were generated. The author group remains responsible for final scientific verification and approval.

Results

Participant characteristics

Among 81 participants, 56 (69.1%) had significant fibrosis and 25 (30.9%) had non-significant fibrosis. The training cohort contained 46 participants with significant fibrosis and 19 without; the validation cohort contained 10 and 6, respectively. The median age was 37 years in both fibrosis groups (Table 1).

Compared with the non-significant-fibrosis group, participants with significant fibrosis had higher red blood cell count, hemoglobin, PT, alanine aminotransferase, gamma-glutamyl transferase, ALP, creatinine and FibroScan liver stiffness and lower platelet count and prothrombin activity. Median CHI3L1 was numerically higher with significant fibrosis [31.73 (22.23–55.80) vs 27.41 (19.86–35.44) ng/mL], but the unadjusted difference was not statistically significant ($P = 0.105$).

Missing data and exploratory screening

The two supplied cohort workbooks contained complete post-imputation analytical records for all 81 participants, and no participant was moved between cohorts. The available files did not retain pre-imputation missingness counts or the historical imputation algorithm; these details therefore could not be reconstructed. Allocation remained fixed at 65 participants for training and 16 for hold-out validation.

At lambda.min, exploratory LASSO retained the sex variable, red blood cell count, hemoglobin, prothrombin time, alkaline phosphatase, HBeAg and CHI3L1. Log10 HBV DNA was not retained at lambda.min but remained in the fixed serum and combined model specifications. These screening results were treated as exploratory and were not interpreted as establishing a final transportable predictor set.

Table 1
Selected baseline characteristics according to biopsy-confirmed fibrosis status

Characteristic	Overall (n = 81)	S0-S1 (n = 25)	S2-S4 (n = 56)	P value
Sex category coded 1, n (%)	52 (64.2)	10 (40.0)	42 (75.0)	0.002
Age, years	37 (33–42)	37 (33–41)	37 (33–42)	0.963
BMI, kg/m ²	21.62 (19.97–23.65)	20.93 (19.97–23.63)	22.09 (20.02–23.66)	0.481
Red blood cell count, x10 ¹² /L	4.85 (4.37–5.27)	4.41 (4.09–4.63)	4.97 (4.60–5.41)	< 0.001
Haemoglobin, g/L	141 (127–152)	135 (125–144)	144 (132–156)	0.035
Platelet count, x10 ⁹ /L	210 (178–238)	226 (208–239)	189 (162–235)	0.008
Prothrombin time, s	13.10 (11.70–13.90)	11.80 (11.30–13.20)	13.50 (12.00–14.10)	0.001
Prothrombin activity, %	100 (89–133)	130 (96–149)	95 (86–124)	0.002
ALT, U/L	33.00 (20.20–47.00)	24.50 (16.00–35.00)	36.40 (24.77–54.40)	0.019
AST, U/L	26.00 (22.00–33.00)	24.00 (18.60–27.00)	27.20 (22.50–34.25)	0.059
GGT, U/L	24.00 (18.00–37.00)	20.00 (17.00–24.00)	28.50 (18.75–42.50)	0.006
ALP, U/L	77 (59–90)	61 (54–77)	82 (69–94)	< 0.001
Creatinine, umol/L	69.90 (56.00–79.30)	58.00 (48.00–74.80)	72.40 (62.40–79.58)	0.017
CHI3L1, ng/mL	30.95 (21.42–47.19)	27.41 (19.86–35.44)	31.73 (22.23–55.80)	0.105
HBV DNA, IU/mL	32,600 (2,440–1,710,000)	14,300 (452 – 93,700,000)	41,750 (2,498–588,500)	0.747
FibroScan liver stiffness, kPa	6.10 (5.00–8.90)	5.00 (4.40–6.00)	7.30 (5.70–9.60)	< 0.001

Values are median (interquartile range) or number (percentage). HBV DNA is presented in IU/mL. The label 'sex category coded 1' is retained because the supplied analytical files did not define the coding direction. P values are descriptive. ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; CHI3L1, chitinase-3-like protein 1; GGT, gamma-glutamyl transferase; HBV, hepatitis B virus.

Regression associations

In univariable analysis, the sex category coded 1, red blood cell count, hemoglobin, PT, ALP and CHI3L1 were associated with significant fibrosis. In the simultaneous seven-variable model, ALP (OR 1.052, 95% CI 1.000-1.106; P = 0.048) and CHI3L1 (OR 1.061, 95% CI 1.004–1.121; P = 0.036) had confidence intervals excluding 1 (Table 2). Log10 HBV DNA showed a negative but statistically inconclusive association (OR 0.792, 95% CI 0.537–1.167; P = 0.239).

Table 2
Univariable and multivariable logistic regression for significant fibrosis

Variable	Univariable OR (95% CI)	P value	Multivariable OR (95% CI)	P value
Sex category coded 1	3.896 (1.267–11.984)	0.018	2.380 (0.184–30.716)	0.507
Red blood cell count, per 1 x10 ¹² /L	3.417 (1.208–9.668)	0.021	2.885 (0.847–9.829)	0.090
Haemoglobin, per 1 g/L	1.034 (1.003–1.065)	0.033	1.024 (0.953–1.099)	0.519
Prothrombin time, per 1 s	1.746 (1.126–2.708)	0.013	1.783 (0.891–3.566)	0.102
Alkaline phosphatase, per 1 U/L	1.075 (1.031–1.120)	< 0.001	1.052 (1.000-1.106)	0.048
CHI3L1, per 1 ng/mL	1.033 (1.001–1.066)	0.043	1.061 (1.004–1.121)	0.036
HBV DNA, per 1-log10 IU/mL increase	0.904 (0.699–1.170)	0.442	0.792 (0.537–1.167)	0.239

The multivariable model included all seven variables shown. The HBV DNA effect is per 1-log10 IU/mL increase, corresponding to a tenfold increase in viral load. CI, confidence interval; OR, odds ratio.

Model discrimination

The FibroScan model had an AUC of 0.783 (95% CI 0.663–0.903) in training and 0.808 (95% CI 0.576-1.000) in validation. Corresponding AUCs were 0.856 (95% CI 0.757–0.954) and 0.750 (95% CI 0.500-1.000) for the serum model, and 0.873 (95% CI 0.787–0.959) and 0.800 (95% CI 0.566-1.000) for the combined model (Table 3 and Fig. 3). The combined model ranked highest in training, whereas FibroScan ranked highest in validation. AUCs were compared descriptively.

Table 3
Discrimination of the three diagnostic approaches

Model	Predictors	Training AUC (95% CI)	Validation AUC (95% CI)
FibroScan model	FibroScan liver stiffness	0.783 (0.663–0.903)	0.808 (0.576-1.000)
Serum model	PT + ALP + CHI3L1 + log10 HBV DNA	0.856 (0.757–0.954)	0.750 (0.500-1.000)
Combined model	PT + ALP + CHI3L1 + log10 HBV DNA + FibroScan	0.873 (0.787–0.959)	0.800 (0.566-1.000)

AUCs were compared descriptively; no formal pairwise superiority claim is made. ALP, alkaline phosphatase; AUC, area under the receiver operating characteristic curve; CI, confidence interval; PT, prothrombin time.

Calibration and exploratory clinical utility

The combined model had an apparent training calibration intercept of 0.000, slope of 1.000 and Brier score of 0.129. In validation, the intercept was – 0.916, the slope was 1.476 and the Brier score was 0.221. The training calibration curve was broadly aligned with the ideal line, whereas validation calibration was irregular because the hold-out cohort contained only 16 participants (Fig. 4).

Exploratory decision-curve analysis showed positive net benefit for the combined model across most displayed thresholds in training. The validation curve was jagged and crossed below zero in part of the displayed range, indicating sensitivity to individual observations (Fig. 5). No probability-threshold interval should be considered established.

Discussion

In this biopsy-confirmed chronic HBV cohort, serum CHI3L1 was numerically higher in participants with Scheuer stage S2-S4 fibrosis. Although the unadjusted group comparison was not statistically significant, the positive adjusted association persisted after inclusion of other model variables. This supports CHI3L1 as a candidate marker for further evaluation, while the small sample and wide confidence intervals require cautious interpretation.

The serum and combined models showed modest-to-good discrimination, but model ranking changed between cohorts. The combined model ranked highest in training, whereas FibroScan ranked highest in validation. This instability argues against claiming superiority. The model components may represent complementary domains: CHI3L1 may reflect inflammatory and matrix-remodeling activity; PT reflects hepatic synthetic function; ALP may capture cholestatic or ductular processes; and HBV DNA reflects virological activity [2, 19]. Liver stiffness provides a physical measure related to fibrosis but is also

influenced by inflammatory activity [9, 11]. This biological rationale is not evidence that each variable provides independent incremental diagnostic value.

The direction of the CHI3L1 association is broadly consistent with earlier HBV studies, but their reference standards and clinical spectra differ. Jin et al. assessed CHI3L1 in 134 patients using an elastography-defined fibrosis endpoint [24], whereas Das et al. examined histological severity in 50 treatment-naïve patients with compensated HBV-related disease [26]. Jiang et al. investigated a broader HBV-related disease spectrum that included cirrhosis and hepatocellular carcinoma [25]. Huang et al. developed a CHI3L1/alpha-fetoprotein/platelet model in 337 patients with Scheuer-staged fibrosis [27]. These studies support further investigation of CHI3L1, but neither their thresholds nor their model performance can be assumed to apply to the present cohort.

Newer reports further clarify the context for this exploratory study. Wang et al. evaluated a CHI3L1-containing model for hepatic injury, including patients in the immune-tolerant phase [28]. The 2026 CHILI study included 347 participants, with 44 in an external-validation cohort, and used a chemiluminescent CHI3L1 assay with machine-learning models [29]. Jin et al. studied 6,361 general-population participants and defined significant fibrosis using FibroTouch liver stiffness rather than liver biopsy [30]. These studies differ from our small cohort in case mix, measurement and modeling; cross-study AUC differences should not be interpreted as head-to-head evidence of superiority.

The negative, inconclusive coefficient for log₁₀ HBV DNA should also be interpreted within the selected biopsy cohort. In the prospective REVEAL-HBV cohorts, higher baseline viral load predicted subsequent cirrhosis and hepatocellular carcinoma [40, 41]. Those longitudinal outcomes differ from contemporaneous significant fibrosis; our coefficient neither contradicts the prognostic importance of viral replication nor supports a protective effect of higher HBV DNA.

Mechanistic findings also warrant restraint. The experimentally defined CHI3L1/interleukin-13 receptor alpha 2 pathway and CHI3L1-deficiency data support a link with inflammatory fibrogenesis [22, 23]. However, these non-HBV findings do not demonstrate that serum CHI3L1 specifically measures HBV-induced collagen deposition. Variation in inflammation, comorbidity and assay calibration may contribute to differences between clinical studies.

A clinically useful extension would compare a routine-marker base model with the same model after adding CHI3L1, alongside established scores such as APRI and FIB-4. Their variable performance in HBV cohorts makes disease-specific comparison important [16, 17]. Discrimination alone is insufficient: calibration and net benefit at clinically justified thresholds are also needed [36, 37, 39]. The present three-model comparison does not establish the incremental value of CHI3L1 or a threshold for routine clinical implementation.

Limitations

Several limitations require emphasis. The retrospective setting and unresolved distinction between one center, two campuses or two centers may limit transportability. The cohort was small relative to the number of candidate predictors; model-development sample size should account for outcome frequency and anticipated model performance, not only total enrollment [42]. Simulation evidence also cautions against treating a fixed events-per-variable rule as a universal adequacy criterion [43]. The hold-out cohort contained only 16 participants, including six without significant fibrosis, resulting in wide AUC confidence intervals and unstable calibration and decision curves. A full-cohort bootstrap validation would use the available observations more efficiently than a single random split, although it would not replace external validation [44].

The fixed four-variable serum model and exploratory LASSO results may be sample-dependent. Data-driven variable selection can affect coefficient stability and the interpretation of conventional P values and confidence intervals [45]. Small samples, predictor handling and analytical choices are relevant to the quality and risk-of-bias domains emphasised by PROBAST + AI [46]; no formal PROBAST + AI assessment was conducted here. Pre-imputation missingness and the exact historical imputation algorithm could not be reconstructed from the post-imputation analytical files, so preprocessing uncertainty was not propagated. Missing-data assumptions and procedures require transparent reporting because they can influence bias and precision [47]. Detailed CHI3L1 assay characteristics, FibroScan acquisition criteria, pathology-review procedures and HBV DNA below-quantification handling were unavailable. External validation, longitudinal outcomes and formal assessment of CHI3L1's incremental value were not performed.

Future studies should prespecify a compact predictor set, document preprocessing completely, use bootstrap optimism correction during development and test a locked equation in an adequately sized independent multicenter sample [42, 44, 48]. Model development and validation should be treated as distinct stages within a prespecified workflow [49]. Reporting should include calibration intercept and slope, Brier score, clinically justified decision thresholds and assay reproducibility [32, 36, 37, 39]. Such work is needed to determine whether CHI3L1 adds enough information to routine laboratory tests or FibroScan to change patient management.

Conclusions

After log₁₀ transformation of HBV DNA, the CHI3L1-containing serum and combined models showed modest-to-good discrimination for biopsy-confirmed significant liver fibrosis in chronic HBV infection. CHI3L1 retained a positive adjusted association, but model ranking was inconsistent and validation estimates were imprecise. These exploratory findings require rigorous internal validation and independent prospective multicenter validation before clinical implementation.

Abbreviations

ALP

alkaline phosphatase
ALT
alanine aminotransferase
AST
aspartate aminotransferase
AUC
area under the receiver operating characteristic curve
CHI3L1
chitinase-3-like protein 1
CI
confidence interval
GGT
gamma-glutamyl transferase
HBV
hepatitis B virus
LSM
liver stiffness measurement
OR
odds ratio
PT
prothrombin time.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of Foshan First People's Hospital (approval No. 2025216). Written informed consent was waived by the Ethics Committee of Foshan First People's Hospital due to the retrospective nature of the study. The consent statement will be verified against the ethics approval record before submission.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests. The funders had no role in study design, data collection and analysis, interpretation of results, or the decision to publish.

Author Contribution

Jinnian Lin: Conceptualization, methodology, data curation, formal analysis, writing–original draft, validation, review & editing. Zicong Gu: Conceptualization, methodology, project administration, funding acquisition, supervision, review & editing. All authors read and approved the final manuscript and agree to be accountable for the work.

Acknowledgments and funding

This work was supported by the Foshan Self-funded Science and Technology Innovation Project (Grant No. 2420001004073), and the Medical Research Project of Foshan Municipal Health Bureau (Grant No. 20260133).

Data Availability

The datasets generated and/or analyzed during the study are not publicly available because they contain potentially identifiable clinical information, but may be available from the corresponding author on reasonable request and subject to institutional ethics and data-governance approval.

References

1. World Health Organization. Guidelines for the prevention, diagnosis, care and treatment for people with chronic hepatitis B infection. Geneva: World Health Organization. 2024. Available from: <https://www.who.int/publications/i/item/9789240090903>. Accessed 31 Aug 2026.
2. European Association for the Study of the Liver. EASL Clinical Practice Guidelines on the management of hepatitis B virus infection. *J Hepatol*. 2025;83(2):502–83. <https://doi.org/10.1016/j.jhep.2025.03.018>.
3. You H, Wang F, Li T, Xu X, Sun Y, Nan Y, et al. Guidelines for the Prevention and Treatment of Chronic Hepatitis B (version 2022). *J Clin Transl Hepatol*. 2023;11(6):1425–42. <https://doi.org/10.14218/jcth.2023.00320>.
4. Marcellin P, Gane E, Buti M, Afdhal N, Sievert W, Jacobson IM, et al. Regression of cirrhosis during treatment with tenofovir disoproxil fumarate for chronic hepatitis B: a 5-year open-label follow-up study. *Lancet*. 2013;381(9865):468–75. [https://doi.org/10.1016/s0140-6736\(12\)61425-1](https://doi.org/10.1016/s0140-6736(12)61425-1).
5. Chang TT, Liaw YF, Wu SS, Schiff E, Han KH, Lai CL, et al. Long-term entecavir therapy results in the reversal of fibrosis/cirrhosis and continued histological improvement in patients with chronic hepatitis B. *Hepatology*. 2010;52(3):886–93. <https://doi.org/10.1002/hep.23785>.
6. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD, American Association for the Study of Liver Diseases. Liver biopsy *Hepatol*. 2009;49(3):1017–44. <https://doi.org/10.1002/hep.22742>.

7. Bedossa P, Dargère D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology*. 2003;38(6):1449–57. <https://doi.org/10.1016/j.hep.2003.09.022>.
8. European Association for the Study of the Liver. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis – 2021 update. *J Hepatol*. 2021;75(3):659–89. <https://doi.org/10.1016/j.jhep.2021.05.025>.
9. Marcellin P, Ziol M, Bedossa P, Douvin C, Poupon R, de Ledinghen V, et al. Non-invasive assessment of liver fibrosis by stiffness measurement in patients with chronic hepatitis B. *Liver Int*. 2009;29(2):242–7. <https://doi.org/10.1111/j.1478-3231.2008.01802.x>.
10. Qi X, An M, Wu T, Jiang D, Peng M, Wang W, et al. Transient Elastography for Significant Liver Fibrosis and Cirrhosis in Chronic Hepatitis B: A Meta-Analysis. *Can J Gastroenterol Hepatol*. 2018;2018:3406789. <https://doi.org/10.1155/2018/3406789>.
11. Fung J, Lai CL, Cheng C, Wu R, Wong DK, Yuen MF. Mild-to-moderate elevation of alanine aminotransferase increases liver stiffness measurement by transient elastography in patients with chronic hepatitis B. *Am J Gastroenterol*. 2011;106(3):492–6. <https://doi.org/10.1038/ajg.2010.463>.
12. Castéra L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology*. 2010;51(3):828–35. <https://doi.org/10.1002/hep.23425>.
13. Boursier J, Zarski JP, de Ledinghen V, Rousselet MC, Sturm N, Lebaill B, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. *Hepatology*. 2013;57(3):1182–91. <https://doi.org/10.1002/hep.25993>.
14. Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology*. 2003;38(2):518–26. <https://doi.org/10.1053/jhep.2003.50346>.
15. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology*. 2006;43(6):1317–25. <https://doi.org/10.1002/hep.21178>.
16. Xiao G, Yang J, Yan L. Comparison of diagnostic accuracy of aspartate aminotransferase to platelet ratio index and fibrosis-4 index for detecting liver fibrosis in adult patients with chronic hepatitis B virus infection: a systemic review and meta-analysis. *Hepatology*. 2015;61(1):292–302. <https://doi.org/10.1002/hep.27382>.
17. Sonneveld MJ, Brouwer WP, Chan HL, Piratvisuth T, Jia JD, Zeuzem S, et al. Optimisation of the use of APRI and FIB-4 to rule out cirrhosis in patients with chronic hepatitis B: results from the SONIC-B study. *Lancet Gastroenterol Hepatol*. 2019;4(7):538–44. [https://doi.org/10.1016/s2468-1253\(19\)30087-1](https://doi.org/10.1016/s2468-1253(19)30087-1).
18. Liguori A, Zoncapè M, Casazza G, Easterbrook P, Tsochatzis EA. Staging liver fibrosis and cirrhosis using non-invasive tests in people with chronic hepatitis B to inform WHO 2024 guidelines: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol*. 2025;10(4):332–49. [https://doi.org/10.1016/s2468-1253\(24\)00437-0](https://doi.org/10.1016/s2468-1253(24)00437-0).

19. Wang S, Hu M, Qian Y, Jiang Z, Shen L, Fu L, et al. CHI3L1 in the pathophysiology and diagnosis of liver diseases. *Biomed Pharmacother.* 2020;131:110680. <https://doi.org/10.1016/j.biopha.2020.110680>.
20. Huang H, Wu T, Mao J, Fang Y, Zhang J, Wu L, et al. CHI3L1 Is a Liver-Enriched, Noninvasive Biomarker That Can Be Used to Stage and Diagnose Substantial Hepatic Fibrosis. *OMICS.* 2015;19(6):339–45. <https://doi.org/10.1089/omi.2015.0037>.
21. Li H, Yan T, Zhu Z, Bao ZW, Xu TM, Xie J, et al. [Diagnostic value of serum chitinase-3-like protein 1 in chronic liver disease of significant fibrosis and cirrhosis]. *Zhonghua Gan Zang Bing Za Zhi.* 2018;26(5):337–41. <https://doi.org/10.3760/cma.j.issn.1007-3418.2018.05.005>.
22. Zheng Q, Cao Y, Jiang X, Wang X, Wang X, He Y, et al. Decoding the CHI3L1/IL-13R α 2 signaling nexus in MASH-fibrosis pathogenesis. *Sci Adv.* 2025;11(50):eadz3223. <https://doi.org/10.1126/sciadv.adz3223>.
23. Higashiyama M, Tomita K, Sugihara N, Nakashima H, Furuhashi H, Nishikawa M, et al. Chitinase 3-like 1 deficiency ameliorates liver fibrosis by promoting hepatic macrophage apoptosis. *Hepatol Res.* 2019;49(11):1316–28. <https://doi.org/10.1111/hepr.13396>.
24. Jin X, Fu B, Wu ZJ, Zheng XQ, Hu JH, Jin LF, et al. Serum chitinase-3-like protein 1 is a biomarker of liver fibrosis in patients with chronic hepatitis B in China. *Hepatobiliary Pancreat Dis Int.* 2020;19(4):384–9. <https://doi.org/10.1016/j.hbpd.2020.05.009>.
25. Jiang Z, Wang S, Jin J, Ying S, Chen Z, Zhu D, et al. The clinical significance of serum chitinase 3-like 1 in hepatitis B-related chronic liver diseases. *J Clin Lab Anal.* 2020;34(5):e23200. <https://doi.org/10.1002/jcla.23200>.
26. Das A, Kamrul-Hasan A, Kabir MR, Das S, Zaki K, Al Mahtab M. Evaluation of Chitinase 3-like 1 (CHI3L1) as a noninvasive biomarker of hepatic fibrosis in patients with Hepatitis B virus-related compensated chronic liver disease. *J Family Med Prim Care.* 2021;10(4):1694–8. https://doi.org/10.4103/jfmpc.jfmpc_1922_20.
27. Huang Q, Wu J, Huang C, Wang X, Xu Z. A noninvasive diagnostic model for significant liver fibrosis in patients with chronic hepatitis B based on CHI3L1 and routine clinical indicators. *Ann Palliat Med.* 2021;10(5):5509–19. <https://doi.org/10.21037/apm-21-957>.
28. Wang JL, Jiang SW, Hu AR, Shi XJ, Zhou AW, Lin K, et al. A model based on chitinase 3-like protein for expecting liver severity of hepatitis B virus infections in the immune tolerance phase. *Clin Chim Acta.* 2025;567:120085. <https://doi.org/10.1016/j.cca.2024.120085>.
29. Cai Q, Liu J, Xu W, Fei J, Li J, Mao L, et al. Explainable machine learning incorporating CHI3L1 enhances liver fibrosis staging in chronic hepatitis B: The CHILI multicenter study. *Dig Liver Dis.* 2026;58(7):925–34. <https://doi.org/10.1016/j.dld.2026.03.017>.
30. Jin P, Li N, Liu C, Sheng D, Qiu L, Zhao C, et al. A chitinase-3-like protein-1-based nomogram for identification of significant liver fibrosis in the general population. *BMC Gastroenterol.* 2025;25(1):688. <https://doi.org/10.1186/s12876-025-04277-0>.

31. Huang X, Zhuang J, Yang Y, Jian J, Ai W, Liu C, et al. Diagnostic Value of Serum Chitinase-3-Like Protein 1 for Liver Fibrosis: A Meta-analysis. *Biomed Res Int.* 2022;2022:3227957. <https://doi.org/10.1155/2022/3227957>.
32. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ.* 2024;385:e078378. <https://doi.org/10.1136/bmj-2023-078378>.
33. Scheuer PJ, Standish RA, Dhillon AP. Scoring of chronic hepatitis. *Clin Liver Dis.* 2002;6(2):335–47. [https://doi.org/10.1016/s1089-3261\(02\)00009-0](https://doi.org/10.1016/s1089-3261(02)00009-0).
34. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B Methodol.* 1996;58(1):267–88. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>.
35. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44(3):837–45. <https://doi.org/10.2307/2531595>.
36. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010;21(1):128–38. <https://doi.org/10.1097/ede.0b013e3181c30fb2>.
37. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Topic Group ‘Evaluating diagnostic tests and prediction models’ of the STRATOS initiative. Calibration: the Achilles heel of predictive analytics. *BMC Med.* 2019;17(1):230. <https://doi.org/10.1186/s12916-019-1466-7>.
38. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak.* 2006;26(6):565–74. <https://doi.org/10.1177/0272989x06295361>.
39. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res.* 2019;3:18. <https://doi.org/10.1186/s41512-019-0064-7>.
40. Iloeje UH, Yang HI, Su J, Jen CL, You SL, Chen CJ, et al. Predicting cirrhosis risk based on the level of circulating hepatitis B viral load. *Gastroenterology.* 2006;130(3):678–86. <https://doi.org/10.1053/j.gastro.2005.11.016>.
41. Chen CJ, Yang HI, Su J, Jen CL, You SL, Lu SN, et al. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. *JAMA.* 2006;295(1):65–73. <https://doi.org/10.1001/jama.295.1.65>.
42. Riley RD, Ensor J, Snell KIE, Harrell FE Jr, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ.* 2020;368:m441. <https://doi.org/10.1136/bmj.m441>.
43. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol.* 2007;165(6):710–8. <https://doi.org/10.1093/aje/kwk052>.
44. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol.* 2001;54(8):774–81. [https://doi.org/10.1016/s0895-4356\(01\)00341-9](https://doi.org/10.1016/s0895-4356(01)00341-9).

45. Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. *Biom J.* 2018;60(3):431–49. <https://doi.org/10.1002/bimj.201700067>.
46. Moons KGM, Damen JAA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P, et al. PROBAST + AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ.* 2025;388:e082505. <https://doi.org/10.1136/bmj-2024-082505>.
47. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ.* 2009;338:b2393. <https://doi.org/10.1136/bmj.b2393>.
48. Riley RD, Archer L, Snell KIE, Ensor J, Dhiman P, Martin GP, et al. Evaluation of clinical prediction models (part 2): how to undertake an external validation study. *BMJ.* 2024;384:e074820. <https://doi.org/10.1136/bmj-2023-074820>.
49. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J.* 2014;35(29):1925–31. <https://doi.org/10.1093/eurheartj/ehu207>.

Figures

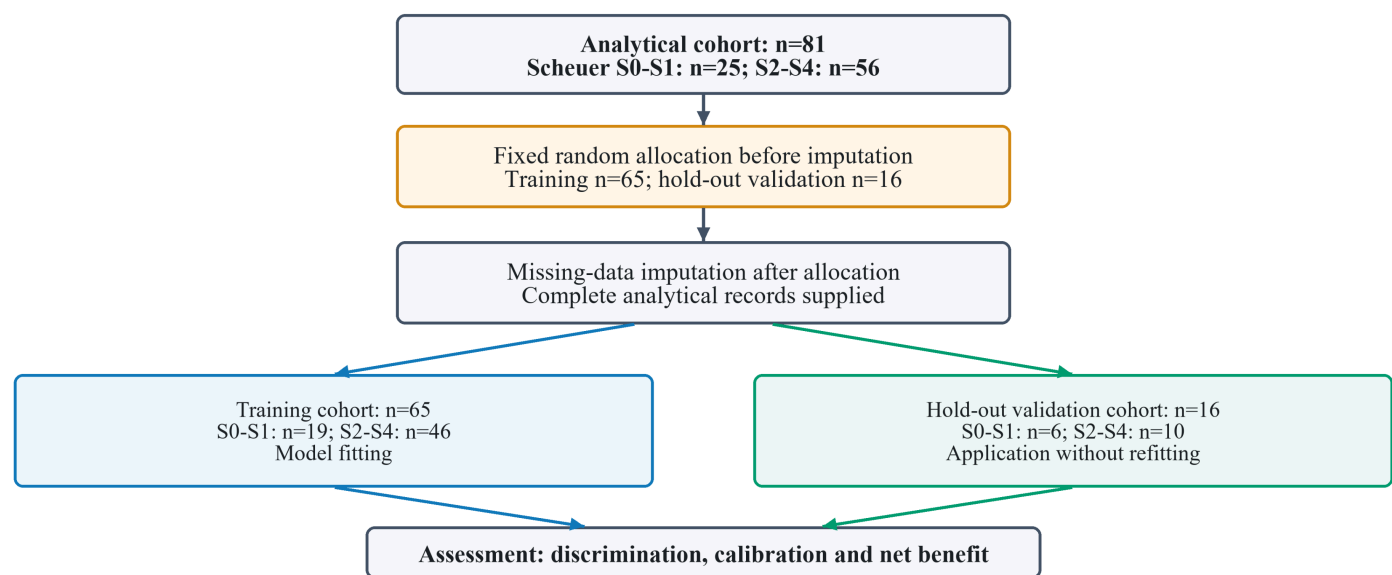


Figure 1

Study flow and analytical framework. Cohort allocation preceded missing-data imputation; the training and validation cohorts remained fixed throughout the analysis

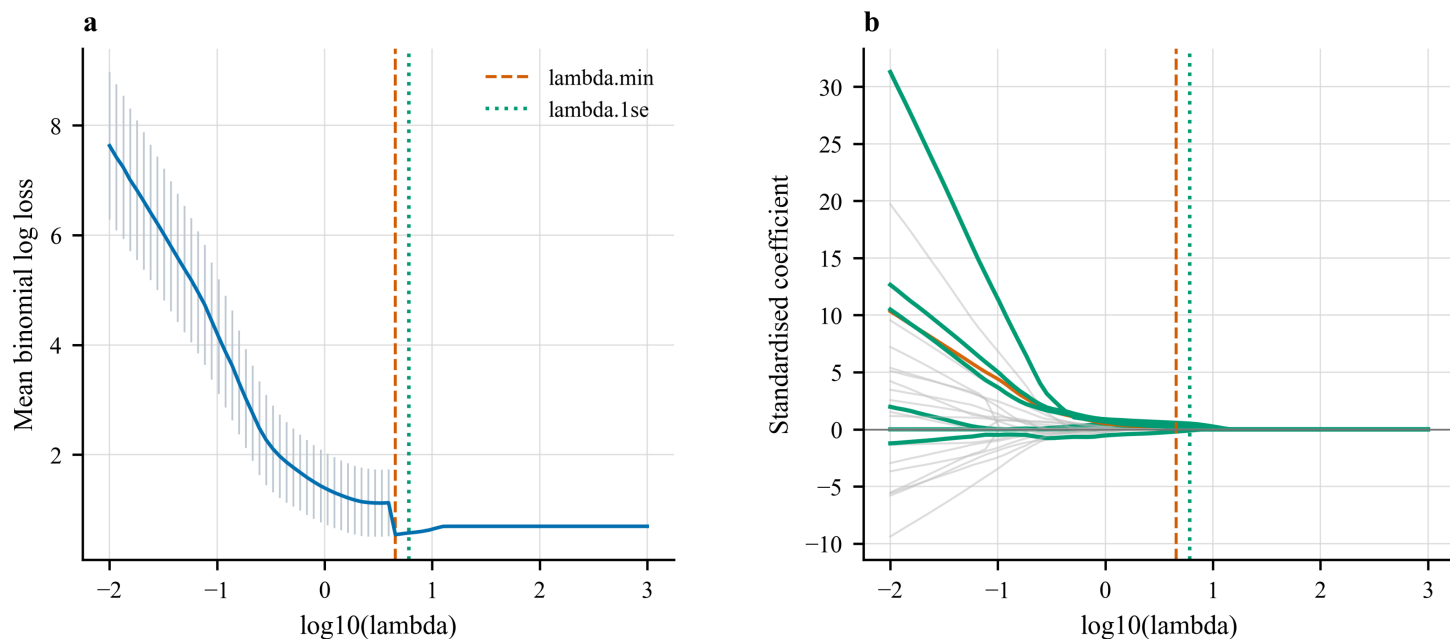


Figure 2

Exploratory LASSO screening in the training cohort. (a) Cross-validated binomial log loss against $\log_{10}(\lambda)$. Vertical lines indicate $\lambda_{\min}$ and λ_{1se} . (b) Standardized coefficient trajectories for candidate predictors

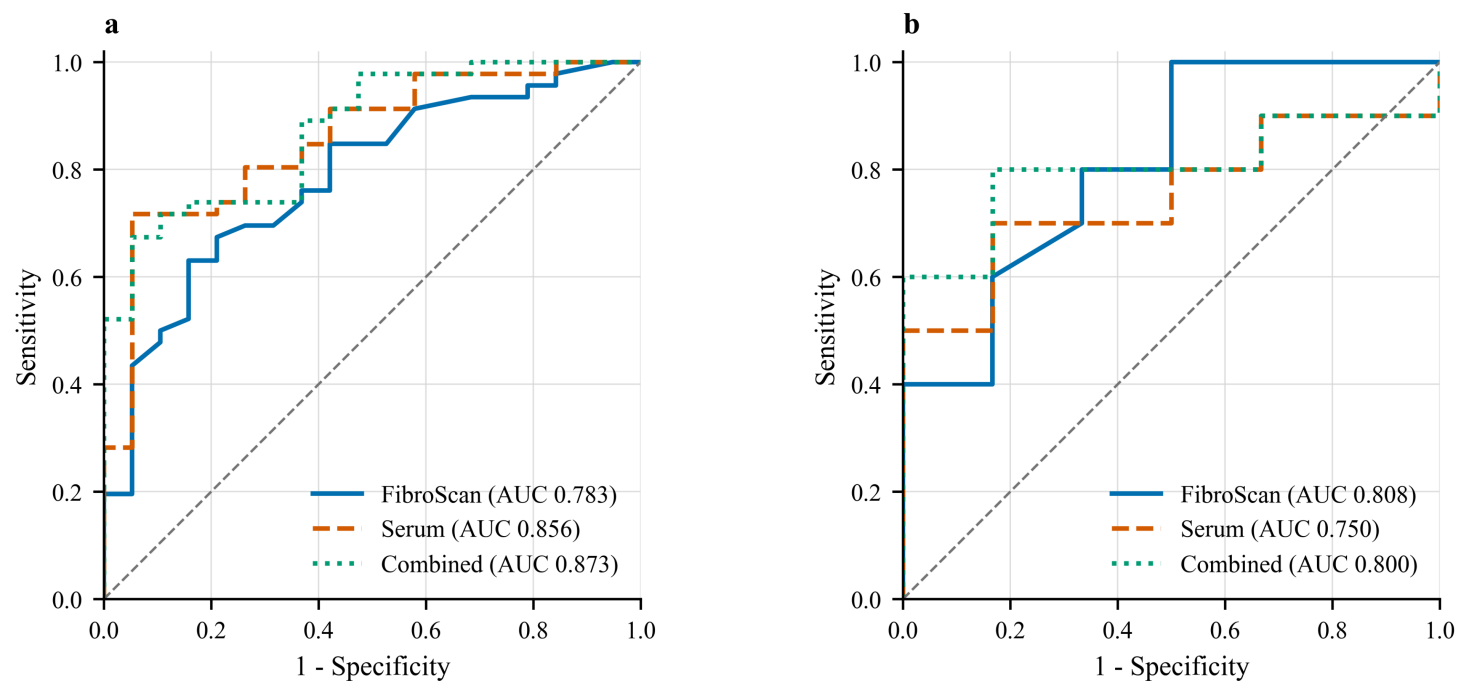


Figure 3

Receiver operating characteristic curves for FibroScan, serum and combined models in the training cohort (a) and hold-out validation cohort (b). The serum model contained PT, ALP, CHI3L1 and $\log_{10}$

HBV DNA; the combined model additionally contained FibroScan

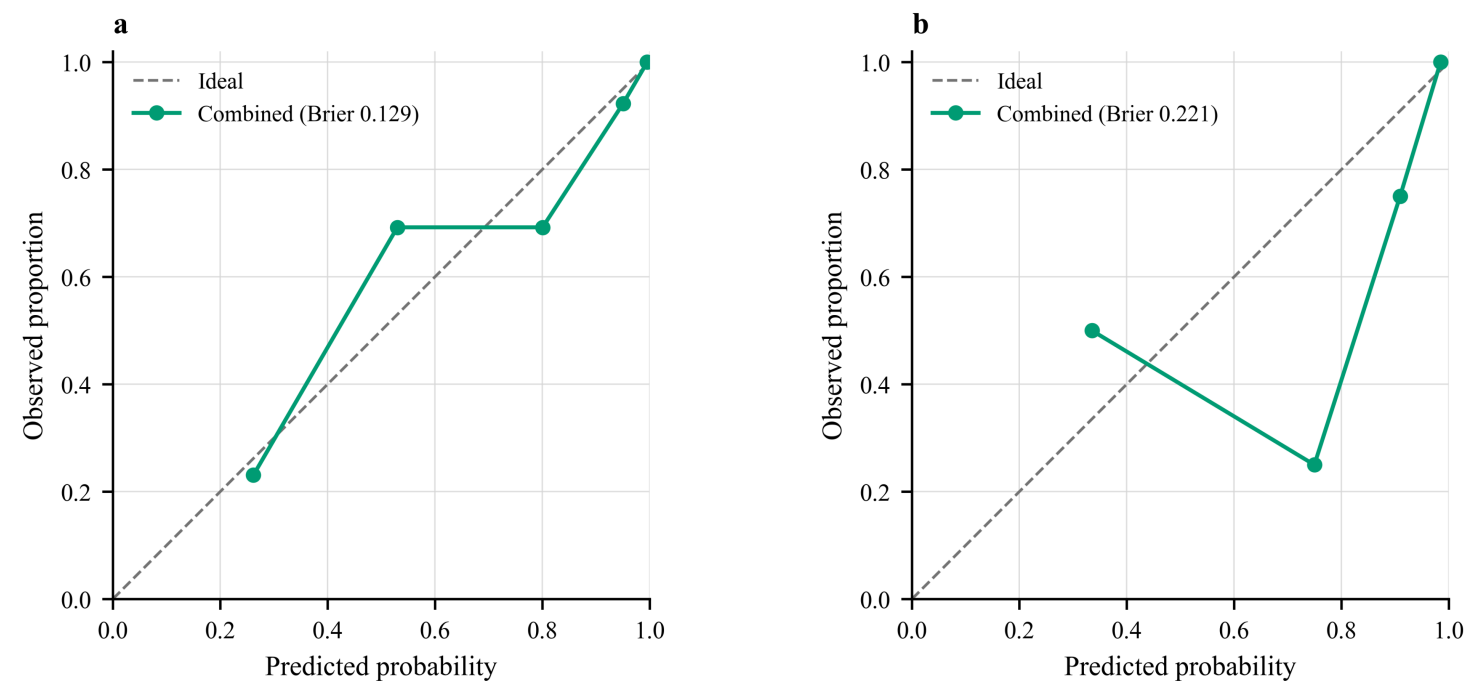


Figure 4

Calibration of the combined model in the training cohort (a) and hold-out validation cohort (b). The diagonal line represents perfect calibration. Validation estimates are imprecise because only 16 participants were available

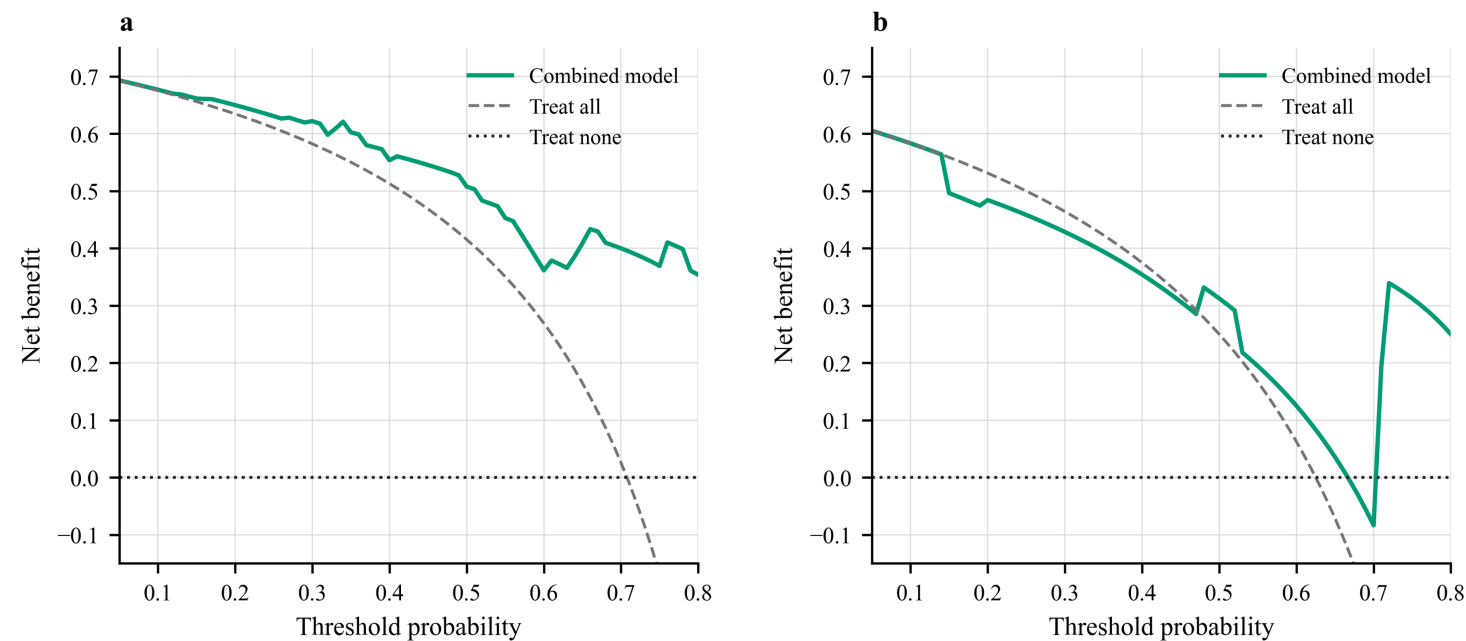


Figure 5

Exploratory decision-curve analysis for the combined model in the training cohort (a) and hold-out validation cohort (b). Treat-all and treat-none strategies are shown; no threshold interval is asserted from the small validation cohort