Journal of Clinical and Translational Hepatology **2026**
DOI: 10.14218/JCTH.2026.00433

Original Article

Three Latent Immune-metabolic Dimensions Improve Prediction of ACLF and Mortality in Acute Decompensation of Cirrhosis

Ruiqi Li[1#], Hao Ren[1#], Junyi Wen[2], Jiale Shen[3], Jun Guo[3], Xiaoying Dong[1], Shuling Chen[4], Hongli Liu[1], Caiyun Zhang[5], Changhua Yi[6], Wei Chen[6], Qingfang Xiong[1], Wei Zhang[2], Xixuan Wang[1], Yuzheng Zhuge[2*] and Yongfeng Yang[1,3*]

[1]Department of Hepatology, The Second Hospital of Nanjing, Medical School, Southeast University, Nanjing, Jiangsu, China; [2]Department of Gastroenterology, Nanjing Drum Tower Hospital, the Affiliated Hospital of Nanjing University Medical School, Nanjing, Jiangsu, China; [3]Department of Hepatology, The Second Hospital of Nanjing, Medical School, Affiliated to Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China; [4]Department of Pharmacy, The Second Hospital of Nanjing, Affiliated to Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China; [5]Nanjing Key Laboratory of Hepatology, Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China; [6]Clinical Research Center, The Second Hospital of Nanjing, Affiliated to Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China

Received: June 01, 2026 | **Revised:** August 29, 2026 | **Accepted:** September 20, 2026 | **Published online:** November 13, 2026

Abstract

Background and Aims: Acute decompensation (AD) of cirrhosis has heterogeneous outcomes; existing scores incompletely capture the immune-metabolic mechanisms associated with progression to acute-on-chronic liver failure (ACLF). We aimed to construct and evaluate three latent immune-metabolic dimensions (hyperinflammation, immunoparalysis, and defense-reserve depletion), identify exploratory unsupervised phenotypes, and assess their incremental predictive value beyond The Model for End-Stage Liver Disease (MELD) and Chronic Liver Failure Consortium Acute Decompensation (CLIF-C AD) scores. ***Methods:*** In this ambispective cohort study, 485 prospectively enrolled patients formed the discovery cohort (January 2024–September 2025), and 623 patients formed the same-center historical evaluation cohort (December 2019–December 2023). Eleven biomarkers were modeled via exploratory and confirmatory factor analysis. Gaussian mixture modeling identified three phenotypes. Tiered models were compared using Fine-Gray regression and competing-risk-aware performance measures for 90-day ACLF (primary endpoint), and Cox regression and survival performance measures for 90-day mortality. ***Results:*** In the historical evaluation cohort, all three dimensions were independently associated with 90-day ACLF and mortality; immunoparalysis was most strongly associated with ACLF (subdistribution hazard ratio 1.92 per 1 standard deviation [SD], 95% confidence interval [CI] 1.56–2.35; $P < 0.001$). The Depleted phenotype (P3, approximately 5%) had 90-day ACLF cumulative incidences of 56.0% (discovery) and 66.7% (historical evaluation). In the historical evaluation cohort, FS2-MELD achieved an area under the curve (AUC) of 0.793 (95% CI 0.748–0.832) for 90-day ACLF, versus 0.711 (0.661–0.756) for FS0-MELD (ΔAUC = 0.082; $P < 0.001$). ***Conclusions:*** Three immune-metabolic dimensions derived from admission biomarkers improve ACLF and mortality prediction beyond conventional scores in AD of cirrhosis, with immunoparalysis showing the strongest prognostic association.

Citation of this article: Li R, Ren H, Wen J, Shen J, Guo J, Dong X, *et al*. Three Latent Immune-metabolic Dimensions Improve Prediction of ACLF and Mortality in Acute Decompensation of Cirrhosis. J Clin Transl Hepatol 2026. doi: 10.14218/JCTH.2026.00433.

Introduction

Acute decompensation (AD) of cirrhosis encompasses ascites, hepatic encephalopathy, variceal hemorrhage, and bacterial infections, and is the leading cause of hospitalization in chronic liver disease.[1] The natural history is heterogeneous: some patients stabilize, while others progress to acute-on-chronic liver failure (ACLF), defined by organ failures and short-term mortality exceeding 30%.[1] Portal hypertension, portosystemic shunting, and associated renal or pulmonary dysfunction contribute to the clinical complexity of decompensated cirrhosis.[2,3] The PREDICT study identified three clinical trajectories—pre-ACLF, unstable decompensated cirrhosis (UDC), and stable decompensated cirrhosis (SDC)—demonstrating that the AD-to-ACLF transition is a critical inflection point.[4] However, trajectory assignment is inherently

Keywords: Liver cirrhosis; Acute-on-chronic liver failure; Immunoparalysis; Factor analysis; Gaussian mixture model; Prognosis; Cytokines.
[#]Contributed equally to this work.
[*]**Correspondence to:** Yongfeng Yang, Department of Hepatology, The Second Hospital of Nanjing, Medical School, Southeast University, Nanjing, Jiangsu, China. ORCID: https://orcid.org/0000-0002-3214-0038. Tel: +86-139-5199-0210, E-mail: yyf1997@163.com; Yuzheng Zhuge, Department of Gastroenterology, Nanjing Drum Tower Hospital, Affiliated Medical College of Nanjing University, Nanjing, China. ORCID: https://orcid.org/0000-0002-3829-5831. Tel: +86-159-9628-9206, E-mail: yuzheng9111963@aliyun.com

Copyright: © 2026 The Author(s). This article has been published under the terms of Creative Commons Attribution-Noncommercial 4.0 International License (CC BY-NC 4.0), which permits noncommercial unrestricted use, distribution, and reproduction in any medium, provided that the following statement is provided. "This article has been published in *Journal of Clinical and Translational Hepatology* at https://doi.org/10.14218/JCTH.2026.00433 and can also be viewed on the Journal's website at http://www.jcthnet.com".

retrospective and cannot guide early decision-making.

Established severity scores—The Model for End-Stage Liver Disease (MELD) and Chronic Liver Failure Consortium Acute Decompensation (CLIF-C AD)—support prognostic assessment in cirrhosis and acute decompensation.[5,6] Recent work has compared conventional and newer prognostic scores for long-term mortality in hospitalized patients with cirrhosis.[7] Immune dysfunction in advanced cirrhosis involves inflammatory activation, impaired cellular immunity, and susceptibility to infection.[8] Depletion of hepatic synthetic reserves can further compromise the host response.[8] These observations suggest three interconnected but distinct pathophysiological mechanisms—hyperinflammation, immunoparalysis, and defense-reserve depletion—that may collectively govern progression from AD to ACLF.

Cao *et al.* classified inflammatory markers into type-1, -2, and -3 subtypes and demonstrated their association with AD severity.[9] The CLIF Consortium developed a 28-gene systemic inflammation score associated with clinical severity.[10] However, these approaches either rely on a priori classifications that do not formally test latent structure, or require specialized platforms not routinely available. To our knowledge, no study has applied formal latent variable methods to discover data-driven immunological dimensions from routine or near-routine biomarkers in AD, or demonstrated that such dimensions improve ACLF prediction beyond established scores.

We designed an ambispective cohort study with three objectives: to construct three mechanistic dimensions from 11 biomarkers using exploratory and confirmatory factor analysis in a prospective discovery cohort and evaluate them in a same-center historical cohort; to identify unsupervised immune-metabolic phenotypes and characterize their associations with 90-day ACLF and mortality; and to evaluate whether adding dimension scores to MELD or CLIF-C AD provides incremental predictive value, using Fine-Gray competing-risks and Cox proportional hazards models.

Methods

Study design and analysis populations

This ambispective, single-center study included a prospective discovery cohort (485 patients, January 2024–September 2025) and a retrospectively assembled historical evaluation cohort (623 patients, December 2019–December 2023). The later cohort provided complete coverage of the 11-marker domain and was used for dimension construction and prediction development; the primary evaluation applied the frozen models to the earlier cohort. Both cohorts and their outcomes had been examined in previous analyses. The complementary forward chronological analysis was also performed. The study followed the 1975 Declaration of Helsinki and received Second Hospital of Nanjing ethics approval (2023-LY-ky-056), covering prospective enrollment and retrospective secondary use; written informed consent was obtained. Patient flow, cohort completeness, and comparative baseline characteristics appear in Figure 1 and Supplementary Tables 1–2.

Eligible patients were aged 18–80 years with cirrhosis and admission for AD: grade 2–3 ascites, hepatic encephalopathy, gastrointestinal bleeding, clinically diagnosed bacterial infection, or combinations thereof. Infection required compatible symptoms/signs with imaging evidence and/or elevated inflammatory markers and was recorded as an admission trigger; microbiological confirmation was not required. Exclusions were admission ACLF by the European Association for the Study of the Liver (EASL)-CLIF criteria, prior liver transplantation, active systemic immunosuppression, hepatocellular carcinoma beyond Milan criteria,[12] extrahepatic malignancy, and pregnancy.[1] Screen-positive SARS-CoV-2 patients were routed outside the hepatology ward. Full cohorts supported descriptive and association analyses; primary complete-panel prediction evaluation used 515 historical patients. Construct-validity analyses used 196 historical patients with lymphocyte-subset measurements.

The exploratory external-center cohort comprised Drum Tower Hospital admissions during January–December 2025. Among 406 screened patients, exclusions comprised 6 with admission ACLF, 156 without complete 90-day follow-up, and 65 missing at least 6 of 11 model variables, leaving 179. The common 4-/5-marker evaluation set contained 166 patients. Ethics approval was obtained from the Ethics Committee of Nanjing Drum Tower Hospital (approval number: 2025-0479-02); broad written consent covered secondary use, with study-specific consent waived.

Outcomes and follow-up

Admission defined time zero. The primary endpoint was incident ACLF within 90 days, determined using EASL-CLIF criteria and the CLIF-C Organ Failure framework for liver, kidney, brain, coagulation, circulation, and respiration.[1,13] Incident 28-day ACLF was exploratory; 90-day all-cause mortality was secondary. Outcome ascertainment was independent of biomarker measurements. For ACLF, death or transplantation before ACLF remained competing events. For mortality, transplantation was a censoring event. Analyses used event and follow-up times within the corresponding 28- or 90-day window.

Measurements, preprocessing, and missing data

Admission blood (T0) was collected within approximately two hours and before treatment. Routine tests, including albumin, used this sample. The 12-cytokine panel was measured by multiplex fluorescence immunoassay[14] (Cellgene Biotech, Jiangxi; P110100403; NMPA 20192400359); complement component 3 (C3), C4, immunoglobulin A (IgA), IgG, and IgM by immunoturbidimetry (Abbott ARCHITECT c16000; Fosun reagents).[15] Samples were fresh and analyzed in the same central laboratory under standardized protocols. Total cholesterol, triglycerides, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), apolipoprotein A-I (ApoA-I), apolipoprotein B, and lipoprotein(a) came from the first fasting draw, usually hospital day 1–2. Medication/procedure records identified intervening albumin, systemic antibacterials, other intravenous fluid ≥ 500 mL, intravenous diuretics and/or large-volume paracentesis, vasoactive therapy, and artificial liver support and/or renal replacement therapy; categories could overlap. Child–Pugh class was derived from total bilirubin, albumin, international normalized ratio (INR), ascites, and hepatic encephalopathy for baseline severity description (Table 1).[16] MELD-Na incorporated serum sodium into MELD for descriptive between-cohort comparisons (Supplementary Table 2).[11]

Cytokines ≤ 0.01 pg/mL were replaced by 0.005 pg/mL.[17] Candidate cytokines, C-reactive protein (CRP), absolute neutrophil count (ANC), lactate, complement/immunoglobulins, lipid markers, and platelets underwent ln(1+x) transformation; albumin and sodium remained untransformed. Discovery means and SDs defined *z*-scores.[18] C3, C4, IgM, HDL-C, ApoA-I, and albumin were negated after standardization so higher scores represented worse status; IL-10 was not negated. Historical scoring retained these transformations and scaling parameters (Supplementary Tables 3–4). For full-

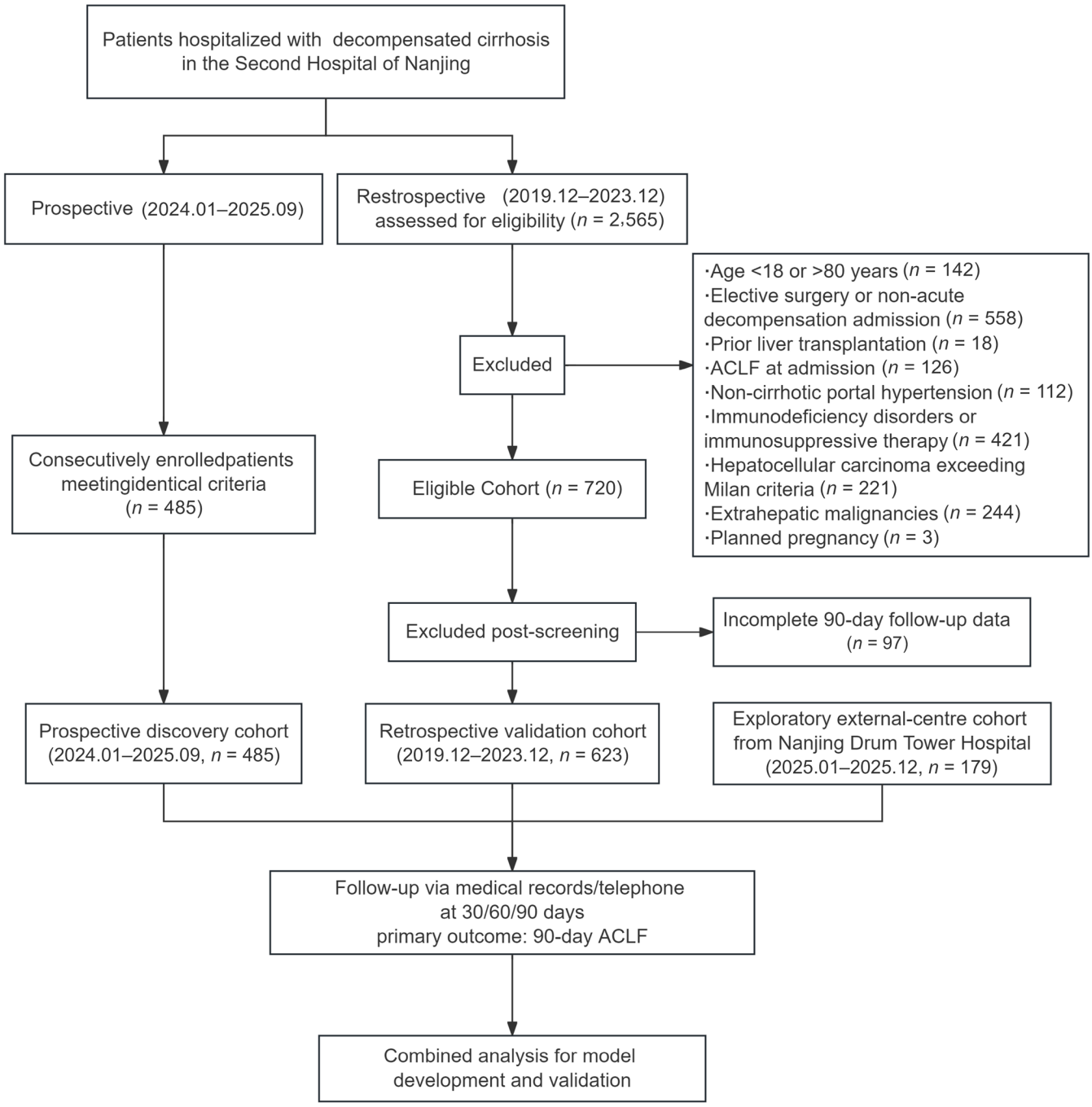


Fig. 1. Study design and patient flow. This ambispective study comprised a prospective discovery cohort and a preceding same-center historical evaluation cohort at the Second Hospital of Nanjing, plus an exploratory external-center cohort. In the retrospective arm (December 2019 to December 2023), 2,565 patients were assessed for eligibility; 720 were eligible, and 97 were excluded for incomplete 90-day follow-up, leaving 623 patients in the historical evaluation cohort. In the prospective arm (January 2024 to September 2025), 485 patients meeting the same eligibility criteria were consecutively enrolled as the discovery cohort. The exploratory external-center cohort comprised 179 patients from Nanjing Drum Tower Hospital enrolled between January and December 2025. The primary outcome was incident ACLF within 90 days. ACLF, acute-on-chronic liver failure.

cohort descriptive and association scoring with incomplete indicators, missing standardized indicators were replaced by medians from historical complete cases before applying frozen Bartlett weights.[19] This procedure was separate from the primary complete-case prediction evaluation.

The sole sensitivity analysis addressed historical biomarker missingness using multiple imputation by chained equations.[20] The full historical cohort was evaluated with 20 imputed datasets, predictive mean matching, 10 iterations, and seed set to 42. The imputation model used 73 predictors, encompassing demographic, clinical, and laboratory information while excluding outcomes, event/censoring times, and

Table 1. Baseline characteristics at admission (T0) in the discovery cohort, stratified by 90-day incident ACLF

Variable	Overall (*n* = 485)	Incident ACLF (*n* = 120)	No ACLF (*n* = 365)	*P*	SMD
Demographics					
Age, years	55.5 ± 11.6	54.3 ± 11.8	55.9 ± 11.5	0.192	0.140
Sex, female	207 (42.7)	52 (43.3)	155 (42.5)	0.952	0.018
Cirrhosis etiology and comorbidities					
Cirrhosis etiology				0.675	
Hepatitis B/C	239 (49.3)	55 (45.8)	184 (50.4)		0.092
Alcohol	77 (15.9)	17 (14.2)	60 (16.4)		0.062
AIH	28 (5.8)	8 (6.7)	20 (5.5)		0.051
MASLD	20 (4.1)	8 (6.7)	12 (3.3)		0.170
PBC	27 (5.6)	8 (6.7)	19 (5.2)		0.064
Multiple	23 (4.7)	5 (4.2)	18 (4.9)		0.036
Others	71 (14.6)	19 (15.8)	52 (14.2)		0.045
Previous AD of cirrhosis	164 (33.8)	44 (36.7)	120 (32.9)	0.516	0.080
Previous ACLF	5 (1.0)	2 (1.7)	3 (0.8)	0.601	0.084
HCC (within Milan)	26 (5.4)	5 (4.2)	21 (5.8)	0.663	0.070
Splenectomy	45 (9.3)	12 (10.0)	33 (9.0)	0.894	0.033
TIPS history	21 (4.3)	5 (4.2)	16 (4.4)	1.000	0.011
AD trigger events[a]					
Infection	151 (31.1)	46 (38.3)	105 (28.8)	0.064	0.207
Ascites	398 (82.1)	95 (79.2)	303 (83.0)	0.415	0.100
Gastrointestinal bleeding	50 (10.3)	10 (8.3)	40 (11.0)	0.517	0.086
Hepatic encephalopathy	47 (9.7)	21 (17.5)	26 (7.1)	0.002	0.351
Routine laboratory values					
WBC, $\times 10^9$/L	4.1 [2.8, 7.1]	5.3 [3.1, 7.9]	4.0 [2.7, 6.5]	0.013	0.242
PLT, $\times 10^9$/L	72 [48, 116]	62 [39, 90]	76 [50, 120]	0.002	0.318
CRP, mg/L	15.3 [9.0, 27.3]	20.2 [10.4, 39.2]	14.4 [8.5, 23.9]	<0.001	0.418
Neutrophil, %	63.5 ± 14.8	67.6 ± 15.0	62.2 ± 14.5	<0.001	0.369
Hemoglobin, g/L	109 ± 27	101 ± 29	112 ± 27	<0.001	0.401
INR	1.3 [1.1, 1.7]	1.5 [1.2, 1.8]	1.3 [1.1, 1.6]	<0.001	0.410
ALT, U/L	32.2 [12.3, 83.9]	38.7 [11.0, 125.0]	29.9 [12.9, 75.9]	0.074	0.173
AST, U/L	44.6 [21.9, 95.6]	63.5 [28.3, 134.6]	39.7 [20.2, 82.0]	<0.001	0.369
ALP, U/L	113.5 [76.0, 171.3]	129.4 [80.3, 194.3]	111.0 [74.3, 161.6]	0.055	0.202
GGT, U/L	54.3 [28.8, 111.6]	57.3 [34.0, 111.8]	52.8 [27.1, 111.6]	0.210	0.137

(continued)

Table 1. *(continued)*

Variable	Overall (*n* = 485)	Incident ACLF (*n* = 120)	No ACLF (*n* = 365)	*P*	SMD
Total bilirubin, μmol/L	47.6 [23.2, 105.6]	75.4 [34.6, 162.1]	41.9 [22.2, 90.9]	<0.001	0.433
Albumin, g/L	32.1 ± 6.0	29.6 ± 5.4	33.0 ± 6.0	<0.001	0.569
Creatinine, μmol/L	65.0 [41.7, 106.5]	80.3 [47.3, 129.6]	61.2 [39.8, 94.7]	<0.001	0.414
BUN, mmol/L	5.0 [3.0, 7.8]	5.5 [3.6, 8.8]	4.9 [2.8, 7.5]	0.026	0.219
Sodium, mmol/L	138.8 ± 5.6	136.6 ± 5.4	139.5 ± 5.5	<0.001	0.534
Lactate, mmol/L	2.7 [2.2, 3.5]	3.0 [2.3, 3.6]	2.6 [2.1, 3.4]	0.051	0.181
Glucose, mmol/L	5.3 [3.2, 8.6]	5.4 [3.2, 9.0]	5.3 [3.2, 8.4]	0.704	0.041
Uric acid, μmol/L	262 [191, 359]	298 [212, 364]	252 [182, 353]	0.026	0.218
LDH, U/L	200 [146, 304]	210 [150, 344]	197 [145, 282]	0.070	0.242
PT, seconds	15.1 [12.2, 18.7]	17.0 [12.7, 20.4]	14.4 [12.0, 18.1]	<0.001	0.383
Key mechanistic biomarkers[b]					
IL-6, pg/mL	15.1 [3.8, 44.6]	28.9 [9.4, 69.5]	11.9 [2.8, 38.3]	<0.001	0.455
IL-10, pg/mL	2.5 [0.7, 7.7]	4.3 [1.2, 12.4]	2.3 [0.6, 6.2]	<0.001	0.382
TNF-α, pg/mL	0.70 [0.13, 2.51]	1.44 [0.50, 2.91]	0.54 [0.07, 1.90]	<0.001	0.531
HDL-C, mmol/L	0.83 [0.58, 1.20]	0.73 [0.43, 0.98]	0.88 [0.61, 1.35]	<0.001	0.549
ApoA-I, g/L	0.78 [0.50, 1.04]	0.64 [0.23, 0.91]	0.82 [0.57, 1.10]	<0.001	0.567
C3, g/L	0.78 [0.58, 1.04]	0.66 [0.50, 0.96]	0.83 [0.63, 1.07]	<0.001	0.435
Clinical scores					
Child–Pugh class				<0.001	
A	92 (19.0)	12 (10.0)	80 (21.9)		0.304
B	259 (53.4)	51 (42.5)	208 (57.0)		0.290
C	134 (27.6)	57 (47.5)	77 (21.1)		0.590
MELD	16.8 ± 7.3	20.3 ± 7.8	15.7 ± 6.8	<0.001	0.659
CLIF-C AD	43.8 ± 12.0	48.4 ± 12.7	42.3 ± 11.4	<0.001	0.520
90-day follow-up outcomes (not baseline)					
90-day mortality	82 (16.9)	65 (54.2)	17 (4.7)		
Liver transplantation	5 (1.0)	1 (0.8)	4 (1.1)		

The discovery cohort was a single-center prospective cohort enrolled between January 2024 and September 2025 (*n* = 485). Incident ACLF was defined as development of ACLF (EASL-CLIF criteria) within 90 days of admission. Continuous variables are presented as mean ± SD (approximately normally distributed) or median [IQR] (skewed). Categorical variables are presented as *n* (%). *P* values were calculated using Student's *t* test or Wilcoxon rank-sum test for continuous variables, and X^2 or Fisher's exact test for categorical variables, as appropriate. *P* values are descriptive and not adjusted for multiple comparisons. For multi-category variables, the global *P* value is shown on the variable header row; SMD is reported for each category separately. For skewed continuous variables, SMD was computed on log-transformed values. [a]AD trigger events are not mutually exclusive; patients may have more than one event at admission. [b]Key mechanistic biomarkers were selected as representative anchors for three mechanistic dimensions; full biomarker panels including 12 cytokines, lipid metabolism markers, and complement/immunoglobulins are provided in Supplementary Table 7. Cytokine values of 0.01 pg/mL represent measurements below the lower limit of detection. Previous AD: history of prior AD-related hospitalization. Multiple: ≥2 concurrent etiologies. Others: cryptogenic cirrhosis, drug-induced liver injury, and less common etiologies. ACLF, acute-on-chronic liver failure; AD, acute decompensation of cirrhosis; AIH, autoimmune hepatitis; ALP, alkaline phosphatase; ALT, alanine aminotransferase; ApoA-I, apolipoprotein A-I; AST, aspartate aminotransferase; BUN, blood urea nitrogen; C3, complement component 3; CLIF-C AD, Chronic Liver Failure Consortium Acute Decompensation score; CRP, C-reactive protein; GGT, gamma-glutamyl transferase; HCC, hepatocellular carcinoma; HDL-C, high-density lipoprotein cholesterol; IL, interleukin; INR, international normalized ratio; IQR, interquartile range; LDH, lactate dehydrogenase; MASLD, metabolic dysfunction-associated steatotic liver disease; MELD, Model for End-Stage Liver Disease; PBC, primary biliary cholangitis; PLT, platelet count; PT, prothrombin time; SD, standard deviation; SMD, standardized mean difference; TIPS, transjugular intrahepatic portosystemic shunt; TNF-α, tumor necrosis factor alpha; WBC, white blood cell count.

follow-up-window indicators. Discovery preprocessing, scoring weights, regression coefficients, and baseline functions remained fixed in every dataset; changes in area under the curve (AUC) were compared with complete-case evaluation.

Externally, cytokines were measured using the Jiangxi Saiji 12-plex assay (NMPA 20192400359) on a BD FACSCanto II[14]; complement/immunoglobulins using Siemens BN II immunonephelometry and manufacturer reagents.[15] No numerical unit conversion was required.

Latent dimensions and exploratory clustering

Outcome variables were not included in factor construction or indicator selection. The three hypothesized dimensions represented innate inflammatory activation, impaired immune/complement competence, and depleted hepatic synthetic/lipid reserves. Exploratory factor analysis (EFA) considered 24 candidate biomarkers: CRP, ANC, interleukin-6 (IL-6), IL-8, IL-1β, tumor necrosis factor alpha (TNF-α), interferon-γ (IFN-γ), IL-17A, lactate, IL-10, C3, C4, IgM, IgA, IgG, IL-4, IL-5, IL-2, IFN-α, IL-12p70, HDL-C, ApoA-I, albumin, and platelets (Supplementary Table 4). Maximum-likelihood EFA used promax rotation (κ = 4).[21,22] Factor-number evidence included the scree plot,[23] Horn parallel analysis with 1,000 random reference datasets,[24] and Velicer's minimum average partial criterion,[25] considered alongside the three hypothesized mechanisms (Supplementary Fig. 1).

Indicator retention required $|\lambda| \geq 0.30$, a cross-loading gap ≥ 0.15, and bootstrap stability > 0.70. Albumin (λ = 0.29) and IgM (λ = 0.25) were retained through clinical override based on stability > 0.90, model identification, and biological roles. Six cytokines with > 60% of measurements at/below detection (IFN-α, IL-1β, IL-5, IL-12p70, IL-2, IL-4) were excluded from the final model. Loading stability used 1,000 bootstrap samples[26] with Procrustes alignment to the original EFA solution.[27] Maximum-likelihood confirmatory factor analysis (CFA) specified three correlated factors using the 11 retained indicators: CRP/ANC/IL-6/IL-8; IL-10/C3/C4/IgM; and HDL-C/ApoA-I/albumin, respectively, without cross-loadings or correlated residuals.[28] Fit assessment used Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR).[29] Historical CFA retained this structure and identification constraints to assess measurement agreement; its estimates did not replace the discovery scoring weights.

Bartlett scores were weighted sums of standardized, direction-harmonized indicators, using discovery-derived weights for both cohorts.[19] Gaussian mixture modeling (GMM) then clustered the three continuous scores.[30] Candidate solutions used K = 2–6 with full or tied covariance matrices; Bayesian Information Criterion (BIC)[30] selected the three-class full-covariance solution (Supplementary Fig. 2). Labels P1–P3 followed increasing mean reserve depletion. Assignment used maximum posterior probability (p_max); p_max < 0.60 flagged uncertainty without removing patients. Mixture means, covariances, and class probabilities were frozen for historical assignment. Phenotypes supported description and association; prediction used continuous scores.

Association analyses

Cohort-specific associations were estimated separately from prediction development. Fine–Gray regression estimated subdistribution hazard ratios (sHRs) for 28-/90-day ACLF with competing death/transplantation; Cox regression estimated hazard ratios for 90-day mortality with transplant censoring.[31,32] Dimension models included all three scores, scaled per discovery-cohort standard deviation (SD), and adjusted for age, sex, and admission infection. Phenotype models used P1 as reference with the same adjustment. Estimates and 95% confidence intervals (CIs) describe associations rather than frozen predictive performance. Phenotype-specific cumulative incidence functions used competing-risk estimation and Gray tests[33,34]; survival curves used Kaplan–Meier estimation[18] and log-rank tests.[18]

Global and covariate-specific proportional-hazards diagnostics were reported for Cox and Fine–Gray Cox diagnostics used scaled Schoenfeld residuals and tests of their relationship with time.[35] Retained diagnostic analyses included Cox models split at day 30 with separate baseline hazards by period.[18,32] Phenotype-specific restricted mean survival time, the area under the survival curve through day 90, was estimated using survRM2; pairwise differences were expressed in days, with two-sided inference from 1,000 bootstrap.[26,36]

Prediction development and frozen evaluation

Five prediction specifications evaluated MELD and CLIF-C AD in parallel[5,6]: FS0-MELD and FS0-CLIF-C AD contained the respective clinical score; FS1 contained the three dimensions; FS2-MELD and FS2-CLIF-C AD combined dimensions with the respective score. Each specification additionally included age, sex, and admission infection, selected for their established prognostic relevance. Models used raw Bartlett scores; per-SD scaling applied to association reporting. Fine–Gray models predicted 28-/90-day ACLF, and Cox models predicted 90-day mortality. The reported implementation used survival version 3.5-8, with finegray-expanded data and coxph for Fine–Gray estimation and coxph for mortality (Supplementary Table 5).[31,32]

All coefficients and baseline functions were estimated in discovery. For covariate vector X and coefficient vector β, ACLF risk was $1-[1-\mathrm{CIF}_0(t)]^{\exp(X\beta)}$, and mortality risk was $1-\mathrm{S}_0(t)^{\exp(X\beta)}$.[31,32] $\mathrm{CIF}_0(t)$ and $\mathrm{S}_0(t)$ are the baseline cumulative incidence and survival functions at $X = 0$ under the reported variable coding, rather than risks for a clinically representative patient. The same discovery preprocessing, Bartlett mapping, coefficients, and baseline functions generated historical complete-case predictions without refitting or recalibration. Calibration models evaluated these fixed risks; their estimates were not used to update predictions. Reduced-panel score definitions and prediction parameters followed the same discovery-to-evaluation freezing principle. External outcomes were not used for marker reselection, fitting, recalibration, or threshold adjustment.

Performance assessment and statistical inference

Discrimination and prediction error were evaluated at the corresponding horizon with time-dependent AUC and Brier scores. ACLF metrics were measured using competing-risk methods, retaining deaths/transplants before ACLF in evaluation; mortality metrics using survival methods with transplant censoring.[37,38] Thus, ACLF performance concerned predicted cumulative incidence rather than a binary analysis after excluding competing events. Brier scores summarized squared prediction error, with lower values indicating better overall prediction. Calibration assessment included observed-to-expected event ratios, slopes, calibration-in-the-large (CITL), and calibration curves.[39] Ideal O/E and slope values were 1, and ideal CITL was 0. Historical ACLF calibration additionally used Locally Estimated Scatterplot Smoothing (LOESS) to obtain the smoothed observed-risk curve.[40] Absolute differences between this curve and predicted risks were summarized by their mean (integrated calibration index, ICI), median (E50), and 90th percentile (E90), following the definitions of these calibration-error summaries; lower

values indicated less calibration error.[41] These were diagnostic summaries of the frozen predictions.

Decision curves evaluated net benefit across risk thresholds against treating all or no patients at the same horizon.[42] Incremental performance used paired AUC and Brier differences between models evaluated in the same patients. Patient-level bootstrap resampling used 1,000 samples unless otherwise specified, preserving each patient's paired predictions and outcome record; 95% CIs summarized sampling uncertainty.[26] Positive ΔAUC and negative ΔBrier favored the augmented model; reduced-panel comparisons were defined as reduced minus full-panel performance.

Continuous variables were summarized as mean ± SD or median (interquartile range [IQR]), and categorical variables as n (%), using available observations. Descriptive comparisons used Student's *t* or Wilcoxon rank-sum tests, Kruskal–Wallis tests across phenotypes, and χ^2 or Fisher tests for categorical variables.[18] Standardized mean differences used Cohen's *d*,[43] with skewed variables assessed after log transformation and categorical differences by category. Baseline *P* values were descriptive without multiplicity adjustment. For construct-validity analyses, Benjamini–Hochberg false-discovery-rate (FDR) correction was used to the 6 tests within each dimension.[44] Analyses were performed using R 4.5.2 and Python 3.11 with deterministic seeds. Reporting followed Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD).[45,46]

Additional retained analyses

Reduced panels comprised 6 markers (CRP, IL-6, IL-10, C3, HDL-C, ApoA-I), five markers (CRP, ANC, C3, HDL-C, albumin), four markers (CRP, C3, HDL-C, albumin), and a three-marker lower-bound panel (CRP, C3, albumin). Discovery-defined score mappings were compared with the original scores by Pearson correlation in both cohorts.[18] Frozen discrimination, Brier score, calibration, and paired differences for ACLF and mortality were evaluated in the common 515-patient historical set. External evaluation used the internally defined 4-/5-marker models in the common 166 patients; 6-/11-marker availability was described without full-panel external validation. External case-mix comparisons were descriptive, without inferential between-cohort testing.

The complementary chronological analysis developed preprocessing, measurement scores, and prediction models in the 515 earlier complete cases, then applied their fixed parameters to the 485 later prospectively collected patients without refitting/recalibration. Measurement agreement included Tucker congruence; frozen performance covered 90-day ACLF and mortality (Supplementary Table 18).[47] This was a supplementary evaluation in a later single-center temporal cohort, not the original primary design.

Exploratory heterogeneity analyses separated association interactions from subgroup prediction.[18,31,32] Etiology was viral hepatitis, alcohol-related liver disease, or other causes; metabolic dysfunction-associated steatotic liver disease and less common etiologies remained in the other category because event support was insufficient for separate strata. Joint 6-degree-of-freedom, dimension-by-etiology tests were carried out using pooled and cohort-specific models, with adjustment for MELD, age, sex, and admission infection, with cohort period added in pooled models; frozen subgroup prediction using historical complete cases (Supplementary Table 20). Infection analyses were performed in 1,000 complete-panel patients, using models including all three dimensions and their infection interactions, with adjustment for MELD, age, sex, and cohort period; a joint test evaluated the three interaction terms for ACLF and mortality. Historical policy-period analyses distinguished pre-February 2020, universal pre-ward testing during February 2020–December 2022, and symptom/clinical-feature-based screening during January–December 2023. Joint dimension-by-period tests and period-specific prediction were exploratory; incomplete individual SARS-CoV-2 status was not directly adjusted. Period comparisons of cytokines were performed using Kruskal–Wallis tests and Benjamini–Hochberg correction across the 12 cytokines.[44]

Construct validity related scores to flow-cytometry T-, B-, and natural killer-cell counts and CD4/CD8 ratio[14] in 196 historical patients using Spearman correlations and Hodges–Lehmann contrasts between quartile 4 and quartile 1 of the score within each subgroup.[18,48] Percentile bootstrap CIs were generated using 1,000 samples,[26] with FDR correction as described above. Descriptive PREDICTing Acute-on-Chronic Liver Failure (PREDICT) trajectory cross-tabulations mapped phenotypes to pre-ACLF (ACLF within 90 days), stable decompensated cirrhosis (no ACLF, death, transplantation, or readmission within 90 days), or unstable decompensated cirrhosis (remaining patients), retaining uncertain phenotype assignments.[4]

Results

Patient characteristics

Between January 2024 and September 2025, 485 patients with acutely decompensated cirrhosis were prospectively enrolled in the discovery cohort; 623 patients admitted between December 2019 and December 2023 comprised the same-center historical evaluation cohort. In the discovery cohort, hepatitis B/C was the leading etiology (49.3%), mean age was 55.5 ± 11.6 years, and 57.3% were male. A bacterial infection diagnosed clinically and recorded as an AD trigger was present in 151/485 patients (31.1%) in the discovery cohort and 189/623 patients (30.3%) in the historical evaluation cohort (Supplementary Tables 6 and 7). Alcohol-related cirrhosis accounted for 15.9% of the discovery cohort and 17.3% of the historical evaluation cohort.

Between T0 and the first fasting blood draw, 443/1,108 patients (40.0%) received at least one recorded treatment. The recorded categories comprised albumin infusion, systemic antibacterial therapy, other intravenous fluid ≥ 500 mL, intravenous diuretic therapy and/or large-volume paracentesis, vasoactive therapy, and artificial liver support and/or renal replacement therapy; 112/1,108 patients (10.1%) received albumin infusion (Supplementary Table 8).

During 90-day follow-up, 120 patients (24.7%) in the discovery cohort developed incident ACLF, 82 (16.9%) died, and 5 (1.0%) underwent liver transplantation. Event rates were comparable in the historical evaluation cohort (90-day ACLF, 154 [24.7%]; mortality, 112 [18.0%]). All pairwise standardized mean differences were below 0.20, indicating broadly similar profiles across the measured admission characteristics (Table 1; Supplementary Tables 2 and 7).

Three mechanistic dimensions

EFA reduced 24 candidate biomarkers to 11 indicators organized into three dimensions: hyperinflammation, immunoparalysis, and defense-reserve depletion. The three-factor CFA showed acceptable fit in the discovery cohort (CFI = 0.954, TLI = 0.938, RMSEA = 0.066). Model fit was maintained and numerically improved in the historical evaluation cohort (CFI = 0.974, TLI = 0.965, RMSEA = 0.050; ΔCFI = +0.020).

Table 2. Indicators defining the three immune-metabolic dimensions: standardized CFA loadings in the discovery cohort

Dimension (mechanism)	Indicator	Standardized loading (λ)	R^2
Hyperinflammation *(innate / acute-phase activation)*	CRP	0.662	0.438
	ANC	0.672	0.451
	IL-6	0.678	0.460
	IL-8	0.590	0.348
Immunoparalysis *(adaptive & complement competence)*	IL-10	0.551	0.303
	C3	0.735	0.540
	C4	0.561	0.315
	IgM	−0.366	0.134
Defense-reserve depletion *(hepatic synthetic & lipid transport)*	HDL-C	0.946	0.895
	ApoA-I	0.971	0.943
	ALB	0.488	0.238

Standardized CFA loadings (λ) and explained variance (R^2) for the 11 retained indicators on their assigned dimension in the discovery cohort (three-factor model: CFI = 0.954, TLI = 0.938, RMSEA = 0.066). A negative loading (IgM) indicates that lower indicator values correspond to a higher dimension score. Dimension scores were computed by Bartlett's method; scoring weights were estimated in the discovery cohort and frozen for historical evaluation. Full exploratory/confirmatory factor analysis results and frozen scoring weights are provided in Supplementary Table 4. ALB, albumin; ANC, absolute neutrophil count; ApoA-I, apolipoprotein A-I; CRP, C-reactive protein; C3/C4, complement components 3 and 4; CFA, confirmatory factor analysis; CFI, Comparative Fit Index; HDL-C, high-density lipoprotein cholesterol; IgM, immunoglobulin M; IL, interleukin; TLI, Tucker-Lewis Index; RMSEA, root mean square error of approximation.

Correlations between the dimensions were moderate (hyperinflammation–immunoparalysis, 0.483; immunoparalysis–defense-reserve depletion, 0.409; hyperinflammation–defense-reserve depletion, 0.276). The dimensions therefore shared information but were not interchangeable, supporting their interpretation as complementary rather than redundant immune-metabolic features (Table 2; Fig. 2A–D; Supplementary Table 4).[4,18,19,28]

Unsupervised phenotypes

GMM identified three phenotypes with a graded severity pattern: GMM translated the continuous dimension scores into three patient phenotypes, with $K = 3$ providing the optimal solution (BIC = 3,432.5; Supplementary Fig. 2). In the discovery cohort, P1, P2, and P3 included 140 (28.9%), 320 (66.0%), and 25 (5.2%) patients, respectively. Corresponding historical evaluation counts were 152 (24.4%), 447 (71.7%), and 24 (3.9%), showing a similar phenotype distribution across cohorts. P3 patients had the highest scores across all three dimensions and a markedly higher median MELD score than P1 patients (27.2 versus 13.0). Assignment uncertainty (p_max < 0.60) was the highest in P1 (20.7% in discovery and 39.5% in historical evaluation) and the lowest in P3 (0.0% and 4.2%, respectively). The low uncertainty in P3 indicates that the Depleted phenotype was the most distinctly separated group in the three-dimensional score space (Fig. 3A, B; Supplementary Table 9).[18,30,34]

Phenotype–outcome associations

In mutually adjusted Fine-Gray models, all three dimension scores were independently associated with 90-day ACLF in both cohorts (Table 3). In the discovery cohort, each 1-SD increase in hyperinflammation and immunoparalysis was associated with sHRs of 1.30 (95% CI, 1.07–1.58; $P = 0.008$) and 1.44 (95% CI, 1.15–1.81; $P = 0.002$), respectively. Defense-reserve depletion showed a similar association (sHR, 1.40; 95% CI, 1.14–1.71; $P = 0.001$). The corresponding historical evaluation estimates were 1.74 (95% CI, 1.50–2.02), 1.92 (95% CI, 1.56–2.35), and 1.40 (95% CI, 1.19–1.65), respectively (all $P < 0.001$). The simultaneous associations of all three dimensions indicate that each captured a complementary component of ACLF risk.

Compared with P1, P3 was associated with substantially higher ACLF risk in both the discovery cohort (sHR, 5.49; 95% CI, 2.94–10.27) and the historical evaluation cohort (sHR, 5.58; 95% CI, 3.22–9.66; both $P < 0.001$). The 90-day cumulative incidence of ACLF increased from 12.1% in P1 to 27.8% in P2 and 56.0% in P3 in the discovery cohort. The same ordering was observed in the historical evaluation cohort (9.9%, 27.5%, and 66.7%, respectively; Fig. 3C).

Associations with 90-day mortality followed a similar pattern. Immunoparalysis showed the largest dimension-level association in both the discovery cohort (HR, 1.78; 95% CI, 1.34–2.39) and the historical evaluation cohort (HR, 2.29; 95% CI, 1.80–2.91; both $P < 0.001$). At the phenotype level, P3 was also associated with markedly higher mortality than P1 in the discovery cohort (HR, 14.06; 95% CI, 4.98–39.71) and the historical evaluation cohort (HR, 4.05; 95% CI, 2.22–7.39; Table 3; Fig. 3D; Supplementary Figs. 3–4).

Construct validity and clinical trajectories

Among 196 patients with lymphocyte-subset data, higher immunoparalysis scores were associated with lower $CD3^+$ T-cell ($\rho = -0.44$), B-cell ($\rho = -0.39$), and NK-cell counts ($\rho = -0.25$; all q_FDR < 0.001). The inverse association was strongest for $CD3^+$ T cells. By contrast, the CD4/CD8 ratio was not associated with the immunoparalysis score, a pattern consistent with broadly proportional reductions across lymphocyte compartments rather than a selective shift between $CD4^+$ and $CD8^+$ T cells (Fig. 2E; Supplementary Table 10).

P1 patients predominantly followed stable decompensated cirrhosis trajectories, whereas P3 patients were most frequently classified as pre-ACLF (Fig. 2F). Together, the immune-cell correlations and clinical trajectory mapping support the biological and clinical interpretability of the three-dimensional phenotypes. Detailed construct-validity statistics are provided in the Supplementary File 1.

Prediction performance

Adding dimension scores improved discrimination in the historical evaluation cohort, while calibration varied by outcome. For 90-day ACLF, the combined FS2-MELD model achieved an

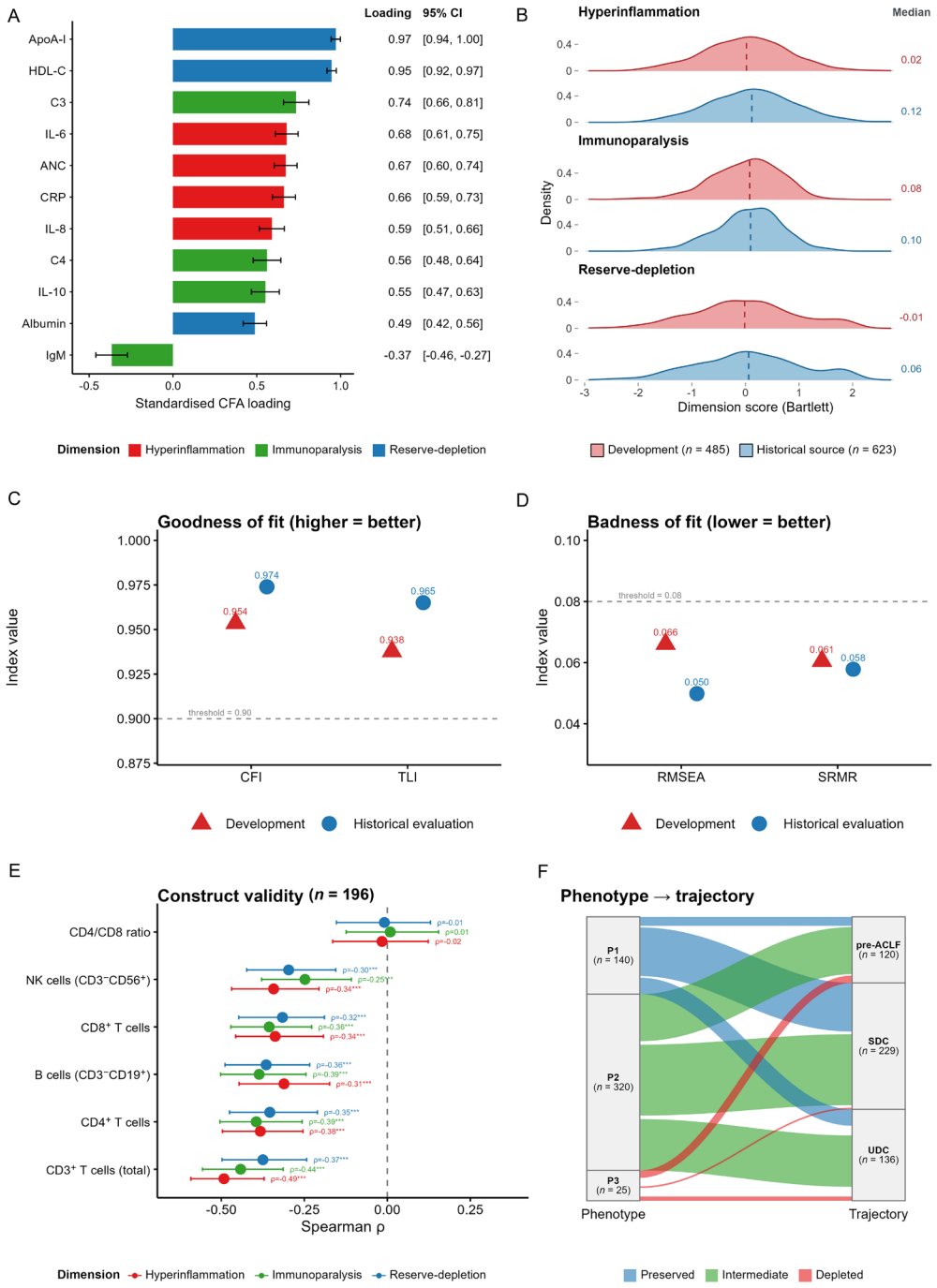


Fig. 2. Immune-metabolic dimensions and clinical trajectories in patients with acute decompensation of cirrhosis. (A) Bars show standardized CFA loadings for 11 admission biomarkers in the development cohort. Error bars indicate 95% CIs. The loadings and CIs are listed at right. (B) Paired density curves show the distribution of Bartlett factor scores for hyperinflammation, immunoparalysis, and reserve depletion. For each dimension, the upper red curve represents the development cohort (n = 485) and the lower blue curve represents the same-center historical evaluation cohort (n = 623). All curves share the same score and density scales. Vertical dashed lines mark the cohort medians. (C, D) Red triangles and blue circles show CFA fit indices for the development and historical evaluation cohorts, respectively. Higher CFI and TLI values (C) and lower RMSEA and SRMR values (D) indicate better fit. (E) Points show Spearman correlations between dimension scores and lymphocyte subset counts or the CD4/CD8 ratio in historical cohort patients with available measurements (n = 196). Horizontal bars indicate 95% bootstrap CIs. Three asterisks indicate an FDR-adjusted q value < 0.001. (F) The Sankey diagram links GMM-defined phenotypes to observed PREDICT trajectories in the development cohort.[4,18,19,28] Ribbon colors identify the preserved (P1, blue), intermediate (P2, green), and depleted (P3, red) phenotypes. Ribbon width is proportional to the number of patients with each phenotype–trajectory combination. Node labels give the total number of patients in each group. ACLF, acute-on-chronic liver failure; ANC, absolute neutrophil count; ApoA-I, apolipoprotein A-I; C3 and C4, complement components 3 and 4; CFA, confirmatory factor analysis; CFI, comparative fit index; CI, confidence interval; CRP, C-reactive protein; FDR, false discovery rate; GMM, Gaussian mixture modeling; HDL-C, high-density lipoprotein cholesterol; IgM, immunoglobulin M; IL, interleukin; NK, natural killer; RMSEA, root mean square error of approximation; SDC, stable decompensated cirrhosis; SRMR, standardized root mean square residual; TLI, Tucker–Lewis index; UDC, unstable decompensated cirrhosis.

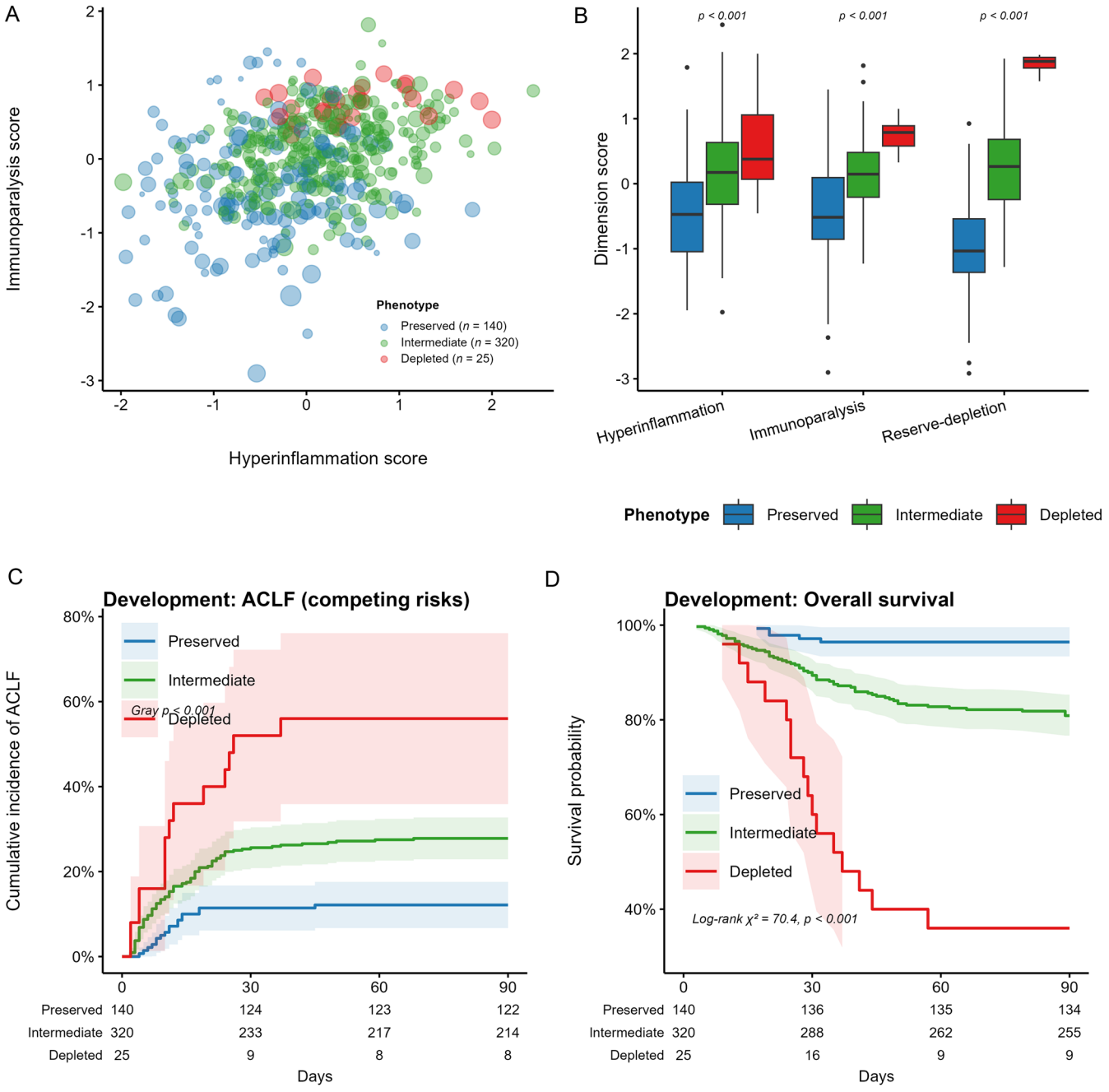


Fig. 3. Unsupervised immune-metabolic phenotypes and outcome associations. (A) Scatter plot of hyperinflammation vs. immunoparalysis scores, colored by phenotype. (B) Box plots of dimension scores by phenotype; Kruskal–Wallis $P < 0.001$ for all three dimensions. (C) Cumulative incidence of 90-day ACLF by phenotype; Gray's test $P < 0.001$. (D) Kaplan–Meier curves for 90-day survival; log-rank $\chi^2 = 70.4$, $P < 0.001$.[18,30,34] ACLF, acute-on-chronic liver failure.

AUC of 0.793 (95% CI, 0.748–0.832), compared with 0.711 (95% CI, 0.661–0.756) for the clinical FS0-MELD model (ΔAUC = 0.082). FS2-CLIF showed similar discrimination (AUC, 0.791; 95% CI, 0.746–0.830). For 90-day mortality, the AUC increased from 0.754 (95% CI, 0.701–0.805) with FS0-MELD to 0.824 (95% CI, 0.777–0.863) with FS2-MELD (ΔAUC = 0.070). Results for 28-day ACLF were consistent (Table 4; Supplementary Table 11; Fig. 4A, B).[37,38,42]

Cox mortality calibration slopes ranged from 0.940 to 1.243, whereas the Fine-Gray ACLF calibration slopes were higher, ranging from 1.359 to 1.748. The mortality slopes were closer to 1, while the ACLF slopes indicated less agreement between the predicted and observed spread of risk. Detailed calibration metrics are provided in Supplementary Table 12. Global Schoenfeld tests provided no evidence of violation of the proportional hazards assumption for any model (all $P > 0.05$; Supplementary Table 13). Exploratory time-stratified and RMST results are provided in Supplementary Table 14.

Decision-curve analysis showed higher net benefit for FS2

Table 3. Association of mechanistic dimension scores and phenotypes with 90-day ACLF and 90-day mortality in the discovery and historical evaluation cohorts

Exposure	Discovery cohort		Historical evaluation cohort	
	sHR (95% CI)	***P* value**	**sHR (95% CI)**	***P* value**
90-day ACLF (Fine–Gray subdistribution hazard model)[a]				
Dimension scores (per 1-SD)				
Hyperinflammation	1.30 (1.07–1.58)	0.008	1.74 (1.50–2.02)	<0.001
Immunoparalysis	1.44 (1.15–1.81)	0.002	1.92 (1.56–2.35)	<0.001
Defense-reserve depletion	1.40 (1.14–1.71)	0.001	1.40 (1.19–1.65)	<0.001
Phenotypes (ref = P1 Preserved)				
P1 Preserved	1 (ref)	—	1 (ref)	—
P2 Intermediate	2.53 (1.50–4.25)	<0.001	3.19 (1.86–5.45)	<0.001
P3 Depleted	5.49 (2.94–10.27)	<0.001	5.58 (3.22–9.66)	<0.001

Exposure	Discovery cohort		Historical evaluation cohort	
	HR (95% CI)	***P* value**	**HR (95% CI)**	***P* value**
90-day mortality (Cox proportional hazards model)[b]				
Dimension scores (per 1-SD)				
Hyperinflammation	1.45 (1.15–1.84)	0.002	1.68 (1.41–2.00)	<0.001
Immunoparalysis	1.78 (1.34–2.39)	<0.001	2.29 (1.80–2.91)	<0.001
Defense-reserve depletion	1.76 (1.36–2.26)	<0.001	1.24 (1.03–1.50)	0.022
Phenotypes (ref = P1 Preserved)				
P1 Preserved	1 (ref)	—	1 (ref)	—
P2 Intermediate	5.84 (2.35–14.54)	<0.001	3.19 (1.71–5.96)	0.011
P3 Depleted	14.06 (4.98–39.71)	<0.001	4.05 (2.22–7.39)	<0.001

Values are sHRs for ACLF and HRs for mortality with 95% CIs. [a]90-day ACLF: Fine–Gray subdistribution hazard model with death and liver transplantation (without prior ACLF) as competing events. Discovery: *n* = 485; 120 events and 21 competing events. Historical evaluation: *n* = 623; 154 events and 33 competing events. [b]90-day mortality: Cox proportional hazards model with transplant-free survival. Patients undergoing liver transplantation were censored at transplant date. Discovery: *n* = 485; 82 deaths. Historical evaluation: *n* = 623; 112 deaths. All models adjusted for age, sex, and infection at admission. Dimension scores are mutually adjusted (all three dimensions entered simultaneously). Per 1-SD increment based on discovery cohort standard deviations. Phenotype models: P1 (Preserved) as reference; adjusted for age, sex, and infection at admission. ACLF, acute-on-chronic liver failure; CI, confidence interval; HR, hazard ratio; sHR, subdistribution hazard ratio; ref, reference; SD, standard deviation.

across clinically relevant thresholds (Fig. 4C, D).

After multiple imputation of missing biomarker values in the historical evaluation cohort, AUC changes across the augmented models were between −0.009 and −0.004, while the FS2 > FS1 > FS0 ordering was preserved (Supplementary Table 15). Supplementary Figures 5–6 provide the missing-data pattern and construct-validity displays, respectively.

Simplified marker panels: In the 515-patient historical evaluation set, 6-marker plus MELD and 11-marker plus MELD yielded ACLF AUCs of 0.797 (95% CI, 0.750–0.840) and 0.796 (95% CI, 0.750–0.840), respectively (paired ΔAUC, 0.001; 95% CI, −0.011 to 0.012), although the 6-marker panel still required IL-6 and IL-10. The 5-marker panel retained minimum score correlations of 0.856 in discovery and 0.851 in historical evaluation and achieved an AUC of 0.781 (95% CI, 0.730–0.828). Removing ANC reduced 4-marker hyperinflammation fidelity from 0.904 to 0.769 in discovery and from 0.931 to 0.814 in historical evaluation, leaving CRP as the sole indicator. Thus, the 4-marker panel was a parsimonious lower bound and the 5-marker panel the preferred clinically practical candidate; complete results are provided in Supplementary Table 16.

The frozen 4- and 5-marker models were applied without refitting or recalibration to the external-center cohort from Drum Tower Hospital, the Affiliated Hospital of Nanjing University Medical School (179 patients overall; 166 in the common performance set). For 90-day ACLF, the AUC and Brier score were 0.770 (95% CI, 0.686–0.847) and 0.188 (95% CI, 0.146–0.233) for MELD, 0.798 (95% CI, 0.710–0.877) and 0.171 (95% CI, 0.134–0.213) for 4-marker plus MELD, and 0.814 (95% CI, 0.724–0.894) and 0.168 (95% CI, 0.130–0.209) for 5-marker plus MELD, respectively. These exploratory results provide directional support for ACLF prediction by the simplified panels; complete marker availability, event support, paired differences, mortality performance, and calibration are reported in Supplementary Table 17.

Complementary chronological analysis: Models rebuilt in 515 earlier historical patients with complete 11-marker data were locked and applied to 485 later prospective patients without refitting or recalibration; Tucker congruence was 0.997. Paired ΔAUCs were 0.053 (95% CI, −0.001 to 0.107) for 90-day ACLF and 0.047 (95% CI, 0.003–0.093) for mortality, providing directional support for the primary findings (Supplementary Table 18).

Across the three SARS-CoV-2 screening and ward-routing policy periods, the joint dimension-by-period interaction did

Table 4. Prediction performance of mechanistic dimension scores (FS1) and combined models (FS2) over clinical scores alone (FS0) for 28-day and 90-day ACLF and 90-day mortality in the discovery and historical evaluation cohorts

Model	Discovery AUC[a]	Historical evaluation AUC (95% CI)[b]	ΔAUC vs. FS0-MELD	ΔAUC vs. FS0-CLIF	Discovery Brier	Historical evaluation Brier	Historical evaluation Cal. slope	Historical evaluation CITL
90-day ACLF (Fine–Gray subdistribution hazard model)								
FS0: MELD	0.686	0.711 (0.661–0.756)	Ref	—	0.1759	0.1902	1.359	−0.121
FS0: CLIF-C AD	0.674	0.742 (0.695–0.784)	—	Ref	0.1787	0.1824	1.599	−0.108
FS1: Dimensions	0.725	0.776 (0.730–0.817)	0.066	0.035	0.1698	0.1708	1.482	0.148
FS2: MELD + Dim.	0.740	0.793 (0.748–0.832)	0.082	—	0.1670	0.1496	1.748	−0.009
FS2: CLIF-C AD + Dim.	0.727	0.791 (0.746–0.830)	—	0.049	0.1695	0.1504	1.695	0.055
28-day ACLF (Fine–Gray subdistribution hazard model)								
FS0: MELD	0.671	0.688 (0.636–0.736)	Ref	—	0.1677	0.1667	1.252	−0.071
FS0: CLIF-C AD	0.656	0.744 (0.695–0.786)	—	Ref	0.1702	0.1584	1.699	−0.057
FS1: Dimensions	0.698	0.780 (0.732–0.821)	0.092	0.036	0.1650	0.1762	1.550	0.195
FS2: MELD + Dim.	0.713	0.790 (0.744–0.830)	0.102	—	0.1636	0.1408	1.757	0.039
FS2: CLIF-C AD + Dim.	0.703	0.796 (0.749–0.835)	—	0.052	0.1653	0.1465	1.781	0.100
90-day mortality (Cox proportional hazards model)								
FS0: MELD	0.744	0.754 (0.701–0.805)	Ref	—	0.1239	0.1540	1.148	−0.134
FS0: CLIF-C AD	0.712	0.756 (0.705–0.801)	—	Ref	0.1272	0.1499	1.160	−0.128
FS1: Dimensions	0.781	0.794 (0.743–0.837)	0.040	0.038	0.1169	0.1235	0.940	0.082
FS2: MELD + Dim.	0.802	0.824 (0.777–0.863)	0.070	—	0.1128	0.1201	1.243	−0.062
FS2: CLIF-C AD + Dim.	0.784	0.811 (0.762–0.852)	—	0.054	0.1162	0.1147	1.172	−0.022

All models were adjusted for age, sex, and infection at admission. For ACLF, time-dependent AUC and Brier scores retained death and liver transplantation without prior ACLF as competing events. For mortality, liver transplantation was treated as censoring. [a]Discovery metrics are reported from the original development analysis. [b]Historical evaluation AUCs are shown with 95% CIs from frozen-model application without refitting or recalibration. ΔAUC is the incremental change relative to the corresponding FS0 model. ACLF, acute-on-chronic liver failure; AUC, area under the curve; Cal. slope, calibration slope; CITL, calibration-in-the-large; CI, confidence interval; CLIF-C AD, Chronic Liver Failure Consortium Acute Decompensation score; Dim., mechanistic dimension scores; FS0, baseline clinical score only; FS1, mechanistic dimension scores only; FS2, clinical score plus dimension scores; MELD, Model for End-Stage Liver Disease; ref, reference.

not indicate clear heterogeneity in the 90-day ACLF associations (P = 0.353; Supplementary Table 19).

Exploratory etiology-stratified analyses: Joint dimension-by-etiology tests did not detect clear heterogeneity in the 90-day ACLF associations (pooled P = 0.224; discovery cohort, P = 0.235; historical evaluation cohort, P = 0.368). In the alcohol-related historical subgroup, the AUCs for MELD and the augmented model were 0.815 and 0.810, respectively, with a paired ΔAUC of −0.004 (95% CI, −0.119 to 0.111). The alcohol-related subgroup contained only 17 ACLF events in each cohort; complete subgroup ACLF and exploratory mortality results are provided in Supplementary Table 20.

Exploratory infection-stratified analyses: Among 1,000 patients with complete 11-marker data, all three dimension scores were associated with 90-day ACLF both with and without clinically diagnosed bacterial infection at admission; the joint infection-by-dimension interaction did not indicate clear heterogeneity (P = 0.806). Thus, the dimension–ACLF associations were not confined to patients with clinically recognized infection; complete ACLF and mortality results are provided in Supplementary Table 6.

Discussion

This study derived three correlated but non-interchangeable immune-metabolic dimensions from 11 early-hospitalization biomarkers in the prospective discovery cohort, with preservation of the three-factor measurement structure in the same-center historical evaluation cohort. All three dimensions were independently associated with 90-day ACLF and mortality, with immunoparalysis showing the largest dimension-level associations. Adding dimension scores to MELD or CLIF-C AD improved discrimination, although ACLF calibration was less close than mortality calibration. The exploratory GMM phenotypes provided a complementary categorical summary: P3 was a small, distinctly separated group with the greatest conventional disease severity and outcome burden.

Cao *et al.* classified inflammatory markers into type-1, -2, and -3 subtypes using a priori biological categories and evaluated their associations with AD severity.[9] The CLIF Consortium's 28-gene systemic inflammation score achieved an AUC of 0.803 for distinguishing high- from low-severity clinical groups,[10] but requires transcriptomic profiling not routinely available. Our approach differs in that it employs formal latent variable methods to discover dimension structure from routine and near-routine biomarkers and quantifies incremental predictive gain over conventional scores. HDL-C has also shown prognostic value in patients with established ACLF.[49]

The PREDICT study defined three trajectories (pre-ACLF,

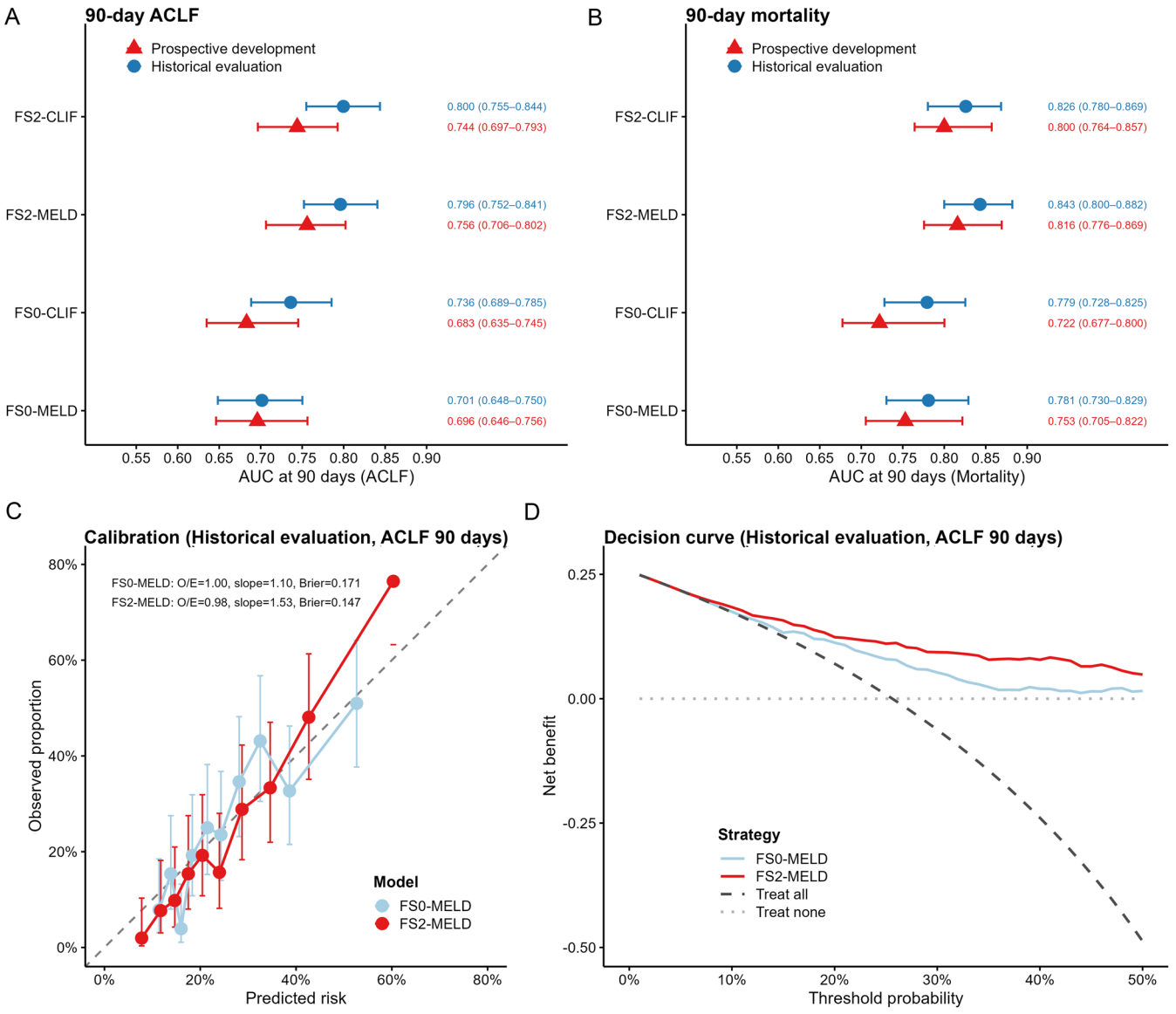


Fig. 4. Prediction performance of tiered models for 90-day ACLF and 90-day mortality. (A) AUC forest plot for ACLF. (B) AUC forest plot for mortality. (C) Calibration plot, historical evaluation cohort. (D) Decision curve analysis, historical evaluation cohort.[37,38,42] ACLF, acute-on-chronic liver failure; AUC, area under the curve; CLIF, Chronic Liver Failure; MELD, Model for End-Stage Liver Disease.

UDC, SDC) based on 90-day outcomes,[4] but trajectory assignment requires complete follow-up. Our Sankey analysis showed that the majority P2 group (66%–72%) distributed across all three trajectories, illustrating a clinical "gray zone" where dimension-based phenotyping may identify higher-risk patients prospectively. P1 patients predominantly followed SDC trajectories, whereas P3 patients were most frequently classified as pre-ACLF, confirming that admission phenotype assignment provides early prognostic information.

These findings currently support prognostic risk stratification rather than treatment selection. P3 represented approximately 5% of admissions and had the highest ACLF and mortality burden, while decision-curve analysis indicated higher net benefit for the dimension-augmented model than for the clinical model alone across the evaluated threshold range. Immunoparalysis remained associated with ACLF among patients without a clinically recorded bacterial infection serving as an AD trigger at admission, indicating that its prognostic information was not confined to clinically recognized infection at presentation. Prospective intervention studies can evaluate the score as a stratification or enrichment variable for immune-directed therapies or structured infection-surveillance strategies. Because ACLF calibration was less close than mortality calibration, absolute-risk use in new settings requires further evaluation.

The immunoparalysis dimension was not reducible to simple lymphopenia. It was constructed from IL-10, C3, C4, and IgM without incorporating lymphocyte counts, while circulating T-, B-, and NK-cell counts were examined only in the cytological subgroup. Moderate inverse correlations with each lymphocyte subset ($\rho = -0.25$ to -0.44), together with a preserved CD4/CD8 ratio, were consistent with proportional lymphocyte depletion accompanying the soluble-marker pattern. This combination aligns with cirrhosis-associated

immune dysfunction[8] and supports interpretation of the dimension as a broader immune-regulatory state rather than a measure of absolute lymphocyte number alone.

The exploratory stratified analysis did not detect clear etiology-related heterogeneity. Viral hepatitis predominated, and the alcohol-related subgroup contributed 17 ACLF events in each cohort, limiting the precision of etiology-specific estimates. Gut-barrier dysfunction and microbial translocation can contribute to systemic immune activation in cirrhosis.[8] Because alcohol exposure patterns, concomitant alcohol-related hepatitis, endotoxemia, and gut permeability were not characterized in the present dataset, the relative contributions of these pathways could not be distinguished. Future prospective multicenter studies with adequate representation of alcohol-related cirrhosis should combine serial immune-metabolic profiling with detailed characterization of alcohol exposure and alcohol-related hepatitis and direct assessment of gut-barrier function, microbial translocation, and innate immune activation to determine whether these pathways modify the level and prognostic contribution of the hyperinflammation dimension.

This study has several limitations. First, the historical evaluation cohort was retrospective and drawn from the same center. The complementary chronological frozen-model analysis provided directional support, but residual differences in case mix, clinical practice, calendar period, and unmeasured COVID-era factors remain. Second, the complete 11-marker analysis excluded records with missing biomarkers. The MICE analysis yielded similar results, although missingness may still affect generalizability. Third, HDL-C and ApoA-I were obtained from the first fasting blood draw after admission rather than at T0 and may have been influenced by treatment before sampling. Repeated measurements of all 11 markers at multiple time points during hospitalization were available for very few patients, particularly for cytokines and complement components, precluding assessment of score-level test-retest reliability and the additional prognostic information of serial scores.

IL-8 and IL-10 are not routinely available in many centers. Although the 5-marker panel balanced construct fidelity and clinical availability, it remains an exploratory practical candidate requiring prospective evaluation. The external-center analysis was also exploratory and restricted to reduced panels because complete cytokine and complement testing was uncommon. The number of mortality events yielded an events-per-variable ratio of 11.7, and the small P3 group produced imprecise phenotype-specific estimates. Assignment uncertainty was greatest in P1, reinforcing that the GMM phenotypes should be interpreted as exploratory rather than as fixed clinical categories.

ACLF definitions identify overlapping but non-identical populations.[13,50] The present findings apply to EASL-CLIF-defined ACLF. The Asian Pacific Association for the Study of the Liver (APASL) framework additionally requires a qualifying acute hepatic insult followed within four weeks by ascites and/or hepatic encephalopathy; these temporal elements were not prospectively adjudicated in the current cohorts. Performance for APASL-defined ACLF therefore remains to be established.

Conclusions

In conclusion, three data-driven immune-metabolic dimensions derived from routine and near-routine admission biomarkers improve prediction of ACLF and mortality beyond conventional clinical scores in AD of cirrhosis. Immunoparalysis showed the strongest independent association with adverse outcomes, supported by its inverse correlation with circulating immune cell counts. Dimension-based phenotyping at admission may help prospectively identify immunologically at-risk patients whose trajectories cannot be distinguished by conventional assessment alone. Multicenter prospective validation across diverse etiological populations and evaluation of dimension-informed risk stratification in prospective multicenter settings are warranted.

Supporting information

Supplementary material for this article is available at https://doi.org/10.14218/JCTH.2026.00433.

Acknowledgments

The authors would like to thank all the patients who participated in this study and the clinical staff who contributed to data collection and patient care. During the preparation of this work, the authors used ChatGPT (OpenAI) to assist with language polishing. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of this article.

Funding

This work was supported by the Natural Science Foundation of Jiangsu Province (BK20251727), the Innovation Center for Infectious Disease of Jiangsu Province (No. CXZX202232), and the SEU Innovation Capability Enhancement Plan for Doctoral Students (CXJH_SEU 26239).

Conflict of interest

The authors have no conflict of interests related to this publication.

Author contributions

RL: Investigation, Data curation, Visualization, Software, Writing – original draft; HR: Formal analysis, Methodology, Validation, Follow-up; JW: Investigation, Data curation; JS: Investigation, Data curation, Validation, Follow-up; JG: Investigation; XD: Investigation; SC: Investigation; HL: Formal analysis; CZ: Investigation; CY: Resources; WC: Investigation; QX: Resources, Supervision; WZ: Resources, Supervision; XW: Investigation, Data curation, Conceptualization, Visualization, Methodology, Supervision, Writing – review & editing; YZ: Resources, Supervision, Project administration, Writing – review & editing; YY: Funding acquisition, Project administration, Resources. All authors have approved the final version and publication of the manuscript.

Ethical statement

The study protocol conformed to the ethical guidelines of the Declaration of Helsinki (as revised in 2024), was approved by the Ethics Committee of the Second Hospital of Nanjing (approval number: 2023-LY-ky-056; the approval explicitly covered both prospective enrollment and retrospective secondary use of clinical data), and written informed consent was obtained from participants in the Second Hospital of Nanjing cohorts. The exploratory external-center cohort at Nanjing Drum Tower Hospital was approved by the Ethics Committee of Nanjing Drum Tower Hospital (approval number: 2025-0479-02); broad written consent covered secondary use, with study-specific consent waived.

Data sharing statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Moreau R, Jalan R, Gines P, Pavesi M, Angeli P, Cordoba J, *et al*. Acute-on-chronic liver failure is a distinct syndrome that develops in patients with acute decompensation of cirrhosis. Gastroenterology 2013;144(7):1426–1437.e9. doi:10.1053/j.gastro.2013.02.042, PMID:23474284.

[2] Abdelwahed AH, Aboeldahb M, Wu GY. Effects of Transjugular Intrahepatic Portosystemic Shunt on Renal and Pulmonary Function in Hepatic Decompensation with and without Hepatorenal and Hepatopulmonary Syndromes: A Review. J Clin Transl Hepatol 2024;12(9):780–791. doi:10.14218/JCTH.2024.00188, PMID:39280072.

[3] Anastasio NM, Wu GY. Non-variceal Extrahepatic Portosystemic Shunts: A Review of Pathogenesis, Diagnosis, and Treatment. J Clin Transl Hepatol 2025;13(5):425–433. doi:10.14218/JCTH.2024.00315, PMID:40385942.

[4] Trebicka J, Fernandez J, Papp M, Caraceni P, Laleman W, Gambino C, *et al*. The PREDICT study uncovers three clinical courses of acutely decompensated cirrhosis that have distinct pathophysiology. J Hepatol 2020;73(4):842–854. doi:10.1016/j.jhep.2020.06.013, PMID:32673741.

[5] Jalan R, Pavesi M, Saliba F, Amorós A, Fernandez J, Holland-Fischer P, *et al*. The CLIF Consortium Acute Decompensation score (CLIF-C ADs) for prognosis of hospitalised cirrhotic patients without acute-on-chronic liver failure. J Hepatol 2015;62(4):831–840. doi:10.1016/j.jhep.2014.11.012, PMID:25463539.

[6] Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, *et al*. A model to predict survival in patients with end-stage liver disease. Hepatology 2001;33(2):464–470. doi:10.1053/jhep.2001.22172, PMID:11172350.

[7] Wang S, Guo G, Wang H, Zhang X, Yang W, Yang J, *et al*. Improved Discrimination and Predictive Ability of Novel Prognostic Scores for Long-term Mortality in Hospitalized Patients with Cirrhosis. J Clin Transl Hepatol 2025;13(6):484–492. doi:10.14218/JCTH.2025.00004, PMID:40474888.

[8] Albillos A, Martin-Mateos R, Van der Merwe S, Wiest R, Jalan R, Álvarez-Mon M. Cirrhosis-associated immune dysfunction. Nat Rev Gastroenterol Hepatol 2022;19(2):112–134. doi:10.1038/S41575-021-00520-7, PMID:34703031.

[9] Cao Z, Yao Y, Cai M, Zhang C, Liu Y, Xin H, *et al*. Blood markers for type-1, -2, and -3 inflammation are associated with severity of acutely decompensated cirrhosis. J Hepatol 2025;82(5):836–850. doi:10.1016/j.jhep.2024.10.028, PMID:39490592.

[10] Trebicka J, Aguilar F, Queiroz Farias A, Lozano JJ, Sánchez-Garrido C, Usón-Raposo E, *et al*. Gene score to quantify systemic inflammation in patients with acutely decompensated cirrhosis. Gut 2025;74(8):1293–1307. doi:10.1136/gutjnl-2024-333876, PMID:40011033.

[11] Kim WR, Biggins SW, Kremers WK, Wiesner RH, Kamath PS, Benson JT, *et al*. Hyponatremia and mortality among patients on the liver-transplant waiting list. N Engl J Med 2008;359(10):1018–1026. doi:10.1056/NEJMoa0801209, PMID:18768945.

[12] Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, *et al*. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. N Engl J Med 1996;334(11):693–699. doi:10.1056/NEJM199603143341104, PMID:8594428.

[13] European Association for the Study of the Liver. EASL Clinical Practice Guidelines on acute-on-chronic liver failure. J Hepatol 2023;79(2):461–491. doi:10.1016/j.jhep.2023.04.021, PMID:37364789.

[14] Cossarizza A, Chang HD, Radbruch A, Abrignani S, Addo R, Akdis M, *et al*. Guidelines for the use of flow cytometry and cell sorting in immunological studies (third edition). Eur J Immunol 2021;51(12):2708–3145. doi:10.1002/eji.202170126, PMID:34910301.

[15] Price CP, Spencer K, Whicher J. Light-scattering immunoassay of specific proteins: a review. Ann Clin Biochem 1983;20 Pt 1:1–14. doi:10.1177/000456328302000101, PMID:6340592.

[16] Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. Br J Surg 1973;60(8):646–649. doi:10.1002/bjs.1800600817, PMID:4541913.

[17] Hornung RW, Reed LD. Estimation of average concentration in the presence of nondetectable values. Appl Occup Environ Hyg 1990;5(1):46–51. doi:10.1080/1047322X.1990.10389587.

[18] Altman DG. Practical statistics for medical research. Chapman and Hall/CRC. 1990.

[19] Bartlett MS. The statistical conception of mental factors. Br J Psychol 1937;28(1):97–104. doi:10.1111/j.2044-8295.1937.tb00863.x.

[20] van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw 2011;45(3):1–67. doi:10.18637/jss.v045.i03.

[21] Lawley DN. The estimation of factor loadings by the method of maximum likelihood. Proc R Soc Edinb 1940;60(1):64–82. doi:10.1017/s0370164600020006x.

[22] Hendrickson AE, White PO. Promax: a quick method for rotation to oblique simple structure. Br J Stat Psychol 1964;17(1):65–70. doi:10.1111/j.2044-8317.1964.tb00244.x.

[23] Cattell RB. The Scree Test For The Number Of Factors. Multivariate Behav Res 1966;1(2):245–276. doi:10.1207/s15327906mbr0102_10, PMID:26828106.

[24] Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrika 1965;30(2):179–185. doi:10.1007/bf02289447, PMID:14306381.

[25] Velicer WF. Determining the number of components from the matrix of partial correlations. Psychometrika 1976;41(3):321–327. doi:10.1007/bf02293557.

[26] Efron B. Bootstrap methods: another look at the jackknife. Ann Statist 1979;7(1):1–26. doi:10.1214/aos/1176344552.

[27] Schönemann PH. A generalized solution of the orthogonal Procrustes problem. Psychometrika 1966;31(1):1–10. doi:10.1007/bf02289451.

[28] Jöreskog KG. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika 1969;34(2):183–202. doi:10.1007/bf02289343.

[29] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Modeling 1999;6(1):1–55. doi:10.1080/10705519909540118.

[30] Fraley C, Raftery AE. Model-based clustering, discriminant analysis, and density estimation. J Am Stat Assoc 2002;97(458):611–631. doi:10.1198/016214502760047131.

[31] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94(446):496–509. doi:10.1080/01621459.1999.10474144.

[32] Cox DR. Regression models and life-tables. J R Stat Soc Series B 1972;34(2):187–202. doi:10.1111/j.2517-6161.1972.tb00899.x.

[33] Aalen OO, Johansen S. An empirical transition matrix for non-homogeneous Markov chains based on censored observations. Scand J Stat 1978;5(3):141–150.

[34] Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. Ann Statist 1988;16(3):1141–1154. doi:10.1214/aos/1176350951.

[35] Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994;81(3):515–526. doi:10.1093/biomet/81.3.515.

[36] Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol 2013;13:152. doi:10.1186/1471-2288-13-152, PMID:24314264.

[37] Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med 2013;32(30):5381–5397. doi:10.1002/sim.5958, PMID:24027076.

[38] Schoop R, Beyersmann J, Schumacher M, Binder H. Quantifying the predictive accuracy of time-to-event models in the presence of competing risks. Biom J 2011;53(1):88–112. doi:10.1002/bimj.201000073, PMID:21259311.

[39] Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. Stat Methods Med Res 2016;25(4):1692–1706. doi:10.1177/0962280213497434, PMID:23907781.

[40] Cleveland WS. Robust locally weighted regression and smoothing scatterplots. J Am Stat Assoc 1979;74(368):829–836. doi:10.1080/01621459.1979.10481038.

[41] Austin PC, Putter H, Giardiello D, van Klaveren D. Graphical calibration curves and the integrated calibration index (ICI) for competing risk models. Diagn Progn Res 2022;6(1):2. doi:10.1186/s41512-021-00114-6, PMID:35039069.

[42] Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inform Decis Mak 2008;8:53. doi:10.1186/1472-6947-8-53, PMID:19036144.

[43] Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

[44] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B 1995;57(1):289–300. doi:10.1111/j.2517-6161.1995.tb02031.x.

[45] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, *et al*. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370(9596):1453–1457. doi:10.1016/S0140-6736(07)61602-X, PMID:18064739.

[46] Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med 2015;162(1):55–63. doi:10.7326/M14-0697, PMID:25560714.

[47] Lorenzo-Seva U, ten Berge JMF. Tucker's congruence coefficient as a meaningful index of factor similarity. Methodology 2006;2(2):57–64. doi:10.1027/1614-2241.2.2.57.

[48] Hodges JL, Lehmann EL. Estimates of location based on rank tests. Ann Math Statist 1963;34(2):598–611. doi:10.1214/aoms/1177704172.

[49] Fu Y, Wu K, Yang S, Fan Y, Qiu J, Zhu Q, *et al*. Exploring the predictive value of serum lipids for short-term prognosis in patients with acute-on-chronic liver failure. Port Hypertens Cirrhos 2024;3(4):184–195. doi:10.1002/poh2.93.

[50] Sarin SK, Choudhury A, Sharma MK, Maiwall R, Al Mahtab M, Rahman S, *et al*. Acute-on-chronic liver failure: consensus recommendations of the Asian Pacific association for the study of the liver (APASL): an update. Hepatol Int 2019;13(4):353–390. doi:10.1007/s12072-019-09946-3, PMID:31172417.