

Construction and evaluation of a predictive model for acute mortality patients with chronic active Epstein-Barr virus infection: a dual-center retrospective cohort study

Fei Hu

Air Force Medical University

Jiayi Zhan

Air Force Medical University

Rui Li

Taiyuan Maternity and Child Care Hospital

Wen Tian

Air Force Medical University

Haifeng Hu

Air Force Medical University

Xuyang Zheng

Air Force Medical University

Hongkai Xu

Air Force Medical University

Denghui Yu

Air Force Medical University

Lina Yan

Air Force Medical University

Hong Du

Air Force Medical University

Fangfang Li

Air Force Medical University

Meijuan Peng

Air Force Medical University

Danfeng Ren

First Affiliated Hospital of Xi'an Jiaotong University

Xiaofei Yang

Air Force Medical University

Jingtao Hou

First Affiliated Hospital of Xi'an Jiaotong University

Yan Liang

Air Force Medical University

Xingchi Chen

Air Force Medical University

Pingping Zhu

Southern Theater Air Force Hospital

Yingli He

First Affiliated Hospital of Xi'an Jiaotong University

Jianqi Lian

Lianjq@fmmu.edu.cn

Air Force Medical University

Research Article

Keywords: Epstein-Barr virus, chronic active Epstein-Barr virus infection, Risk factors, Nomogram, Score scale, Mortality

Posted Date: October 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10555005/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

Background

Patients with chronic active Epstein-Barr virus infection (CAEBV) face a remarkably high rate of early mortality. We aimed at developing a nomogram and a risk scoring system able to predict one-year acute mortality in patients with CAEBV.

Methods

A retrospective, two-center analysis was conducted on 180 CAEBV patients from January 2014 to December 2024, and identified independent risk factors for 1-year death of CAEBV. Using LASSO regression analysis, while multifactorial Cox regression was utilized to establish the nomogram model (Nomo) and the risk stratification score (Score). Calibration plots, receiver operating characteristic (ROC) curves, and decision curve analysis (DCA) were used to assess model's clinical utility.

Results

Among 180 CAEBV cases, 49 patients (27.2%) died within 1 year and 131 patients (72.8%) survived. LASSO and COX regression together identified five independent risk factors (all $P < 0.05$): polyserositis (HR = 3.68, 95% CI: 1.89–7.18), cytopenia (HR = 3.29, 95% CI: 1.68–6.42), plasma EBV DNA (HR = 1.39, 95% CI: 1.16–1.66), LDH (HR = 1.01, 95% CI: 1.01–1.01), and TBIL (HR = 1.01, 95% CI: 1.01–1.02). Nomo and Score models for predicting 1-year acute mortality in CAEBV patients was subsequently developed. Both the Nomo and Score models demonstrated excellent discrimination (AUC > 0.8) and reliable calibration over a 1-year period. The Score model exhibited predictive power comparable to Nomo, as indicated by internal cross-validation and improvement-index analysis.

Conclusion

We developed and validated an evidence-based, weighted scoring system to stratify the mortality risk of CAEBV. These tools can assist clinicians in rapid patient triage and facilitate the implementation of tiered management, particularly during peak mortality periods.

Introduction

CAEBV is a T/NK cell lymphoproliferative disorder caused by persistent Epstein Barr Virus (EBV) infection, characterized by persistent infectious mononucleosis (IM) symptoms such as fever, lymphadenopathy, and hepatosplenomegaly, which may persist or heal slowly and are often accompanied by multiple organ damage, such as interstitial pneumonia, central nervous system infection, cytopenia, and other serious complications. While the 2017 WHO classification defined CAEBV as an EBV-associated T/NK-cell lymphoproliferative disorder, the 2022 revision narrows this definition to strictly T- or NK-cell lineage infection, thereby excluding EBV-driven lymphoproliferative disorders linked to immunodeficiency as well as those confined exclusively to B cells^{1, 2}. To control the inflammatory

response and proliferation of T/NK cells caused by EBV infection, treatment often involves the use of chemotherapy drugs and allogeneic hematopoietic stem cell transplantation (HSCT). However, despite these interventions, treatment outcomes for this condition have generally been suboptimal^{3, 4}. Previous studies have reported that the 3-year survival rate for this disease is only 58%, with most patients succumbing acutely within 1 year of onset^{4, 5}; however, the risk factors and predictors of acute mortality remain unclear. Therefore, we plan to conduct a two-center, retrospective cohort study to analyze the 1-year acute survival rate of CAEBV patients, identify risk factors for death, providing a reference for early clinical prediction.

Materials and methods

Study design and subjects

This study was a retrospective two-center analysis that collected data from 180 CAEBV patients between January 2015 and December 2024 through the electronic medical records system, tracked their 1-year post-discharge survival status through medical record review, outpatient visits, and telephone follow-ups, and subsequently categorized them into survival and death groups based on the outcome.

Inclusion criteria (suggested by the Research Group of Measures Against Intractable Diseases of the Ministry of Health, Labor, and Welfare of Japan): Elevated EBV-DNA load ($>10^{2.5}$ copies/ml DNA); EBV infection of T or NK cells; Systemic inflammatory symptoms (such as fever, lymphadenopathy, liver dysfunction, progressive skin lesions and cytopenia) persisting for 3 months. **Exclusion criteria:** Presence of autoimmune diseases or immunodeficiencies; progression of CAEBV to neoplastic diseases, such as extranodal NK/T-cell lymphoma and peripheral T-cell lymphoma; co-infections with other pathogens; lack of laboratory examination results and case data; and presence of other serious diseases.

Methods

The following clinical data at diagnosis were collected anonymously from electronic medical records: Epidemiological analysis included age, gender, area of onset, as well as personal history of smoking, alcohol, hypertension, and diabetes. Assessed clinical symptoms comprised fever, lymphadenopathy, abdominal pain, fecal occult blood, neurological manifestations, hepatomegaly, splenomegaly, pulmonary infection, polyserositis, cytopenia, and thrombocytopenic purpura. Laboratory indicators evaluated were platelet count (PLT), hemoglobin (HGB), alanine aminotransferase (ALT), albumin (ALB), total bilirubin (TBIL), rheumatoid factor (RF), C-reactive protein (CRP), prothrombin time (PT), activated partial thromboplastin time (APTT), fibrinogen (FIB), lactate dehydrogenase (LDH), triglycerides (TG), alpha-fetoprotein (AFP), ferritin (FER), and EBV DNA in plasma (pEBV DNA).

For variable selection, least absolute shrinkage and selection operator (LASSO) regression analysis was applied with 10-fold cross-validation. The optimal penalty parameter (λ) was defined using the 1-

standard error ($\lambda \cdot 1se$) criterion to ensure a parsimonious and robust model. Variables with non-zero coefficients were subsequently incorporated into a multivariable Cox proportional hazards regression model to calculate hazard ratio (HR) and 95% confidence interval (95% CI). Additionally, a scoring system (Score) was constructed based on the HR of independent predictors. Specifically, continuous variables were transformed into graded categorical variables according to clinical guidelines, clinically meaningful cut-off values and common usage practices. Each variable was assigned a score weight proportional to its HR, facilitating an intuitive and clinically applicable risk quantification system. Model performance was evaluated using ROC and time-ROC curve analyses, reporting AUC with 95% CI. Clinical utility was assessed via DCA, while calibration curves and internal cross-validation were applied to examine model consistency.

Statistical Analysis

Continuous variables with normal distribution were expressed as mean $\pm$ standard deviation and compared using the independent Student's *t*-test or one-way ANOVA. Non-normally distributed variables were presented as median (interquartile range, IQR) and compared using the Mann–Whitney *U* test or Kruskal–Wallis *H* test. Categorical variables were presented as numbers (percentages) and compared using the chi-squared test. A two-sided *p*-value < 0.05 was considered statistically significant. All statistical analyses were performed using R software (version 4.3.2).

Results

Baseline and clinical characteristics of enrolled patients with CAEBV.

From an initial cohort of 934 adults with EBV infection, 557 cases of acute infection and 141 cases with EBV-associated neoplastic diseases were sequentially excluded. The final cohort comprised 180 patients who met the diagnostic criteria for CAEBV and were followed for one year. The primary endpoint was overall survival, with a median follow-up duration of 375 days (IQR 330–419). According to the pre-specified eligibility criteria, 180 patients were enrolled in this study (Fig. 1), among whom 49 patients (27.2%) died within one year. The distribution of primary causes of death was as follows: hemophagocytic syndrome (HLH) accounted for 19 cases (38.8%), multiple organ dysfunction (MODS) for 11 (22.4%), severe infection for 8 (16.3%), lymphoma or leukemia for 7 (14.3%), and other complications for the remaining 4 (8.2%). Relative to the survival group, the death group exhibited significant differences (all $p < 0.05$) in the distribution of area of onset, fecal occult blood, neurological symptoms, polyserositis, cytopenia, and thrombocytopenic purpura; with respect to laboratory parameters, levels of ALT, TBIL, LDH, and pEBV DNA were significantly elevated (all $p < 0.05$), whereas PLT, HGB, and ALB were markedly reduced (all $p < 0.05$) (Table 1).

Table 1

Comparison of baseline data and clinical characteristics between survival group and death group

Variables^a	Survival Group (n = 131)	Death Group (n = 49)	χ^2/Z	<i>P</i> value
Gender n (%)			0.105	0.746
Male	81 (61.83)	29 (59.18)		
Female	50 (38.17)	20 (40.82)		
Area of onset n (%)			4.291	0.038
Town	51 (38.93)	11 (22.45)		
Countryside	80 (61.07)	38 (77.55)		
Age (years old) n (%)			1.083	0.582
≤30	48 (36.64)	14 (28.57)		
30–45	29 (22.14)	13 (26.53)		
>45	54 (41.22)	22 (44.90)		
Smoking n (%)	40 (30.53)	18 (36.73)	0.628	0.428
Drinking n (%)	33 (25.19)	14 (28.57)	0.211	0.646
Surgical history n (%)	28 (21.37)	12 (24.49)	0.200	0.654
Hypertension history n (%)	13 (9.92)	5 (10.20)	0.003	0.957
Diabetes history n (%)	6 (4.60)	3 (6.10)	0.179	0.673
Fever n (%)	119 (90.84)	46 (93.88)	0.125	0.724
Head and neck lymphadenopathy n (%)	89 (67.94)	38 (77.55)	1.586	0.208
Multiple lymphadenopathy (≥ 3 regions) ^b n (%)	58 (44.27)	29 (59.18)	3.174	0.075
Abdominal pain ^c n (%)	60 (45.80)	23 (46.94)	0.019	0.892
Fecal occult blood ^d (%)	32 (24.43)	21 (42.86)	5.830	0.016
Neurological manifestations ^e n (%)	20 (15.27)	15 (30.61)	5.361	0.023
Hepatomegaly n (%)	61 (46.56)	29 (59.18)	2.271	0.132
Splenomegaly n (%)	99 (75.57)	43 (87.76)	3.178	0.075
Pulmonary infection ^f n (%)	80 (61.07)	32(65.31)	0.272	0.602

Variables^a	Survival Group (n = 131)	Death Group (n = 49)	χ^2/Z	<i>P</i> value
Polyserositis ^g n (%)	52 (39.69)	35 (71.43)	14.381	< 0.001
Cytopenia ^h n (%)	57 (43.51)	36 (73.47)	12.816	< 0.001
Thrombocytopenic purpura ⁱ n (%)	28 (21.37)	18 (36.73)	4.423	0.035
PLT, $\times 10^9/L$	110.00 (58.00, 205.00)	76.00 (54.00, 120.00)	-2.128	0.033
HGB, g/L	123.00 (108.50, 139.00)	112.00 (99.00, 126.00)	-2.342	0.019
ALT, U/L	57.00 (32.50, 91.55)	81.00 (47.00, 164.00)	-3.310	< 0.001
ALB, g/L	32.40 (28.30, 36.65)	29.20 (24.90, 32.40)	-3.288	0.001
TBIL, $\mu\text{mol/L}$	17.87 (10.70, 28.35)	23.45 (15.10, 55.80)	-3.222	0.001
RF, IU/mL	10.40 (9.75, 11.90)	10.80 (9.46, 11.60)	-0.189	0.851
CRP, mg/L	19.10 (9.29, 46.59)	25.20 (14.90, 40.50)	-0.736	0.462
PT, s	12.55 (12.00, 14.38)	12.65 (11.88, 15.70)	-0.670	0.504
APTT, s	37.50 (31.63, 45.00)	38.25 (32.43, 52.50)	-1.058	0.290
FIB, g/L	2.10 (1.37, 3.53)	1.67 (1.10, 2.97)	-1.599	0.109
LDH, U/L	376.00 (185.50, 602.50)	453.00 (339.00, 772.00)	-3.133	0.002
TG, mmol/L	1.99 (1.37, 2.82)	2.06 (1.23, 2.96)	-0.151	0.880
AFP, $\mu\text{g/L}$	1.67 (1.08, 2.41)	1.40 (0.97, 1.91)	-1.594	0.111
FER (log10, $\mu\text{g/L}$)	3.99 (3.66, 4.20)	4.18 (3.82, 4.20)	-0.857	0.392
pEBV DNA ^j (log10, copies/mL)	4.15 (3.06, 4.77)	4.70 (4.39, 6.17)	-4.250	< 0.001

Variables ^a	Survival Group (n = 131)	Death Group (n = 49)	χ^2/Z	P value
* PLT, platelet; HGB, hemoglobin; ALT, alanine aminotransferase; ALB, albumin; TBIL, total bilirubin; RF, rheumatoid factor; CRP, C-reactive protein; PT, prothrombin time; APTT, activated partial thromboplastin time; FIB, fibrinogen; LDH, lactate dehydrogenase; TG, triglyceride; AFP, alpha-fetoprotein; FER, ferritin; EBV DNA, Epstein-Barr Virus Deoxyribonucleic Acid.				
a All laboratory test results and clinical characteristics considered for diagnosis were obtained within one week before and after the diagnosis.				
b Multiple lymphadenopathy (≥ 3 regions) is defined as the involvement of three or more of the following sites: head, neck, axilla, chest cavity, abdominal cavity, pelvic cavity, groin, etc.				
c Abdominal pain, the patient-reported symptom of pain localized to the abdominal region.				
d Fecal occult blood, a positive fecal occult blood test indicates the presence of hidden (non-visible) blood in the stool, typically when the amount of bleeding exceeds 5 milliliters.				
e Neurological manifestations, clinical signs or symptoms indicating dysfunction of the central or peripheral nervous system.				
f Pulmonary infection, the presence of an infectious process involving the lower respiratory tract, such as pneumonia.				
g Polyserositis, polyserositis is defined as the inflammation of at least two serous membranes (e.g., pleura, peritoneum, or pericardium), diagnosed by imaging studies such as ultrasound or computed tomography.				
h Cytopenia (reduction in at least two of the three peripheral blood lines) is defined as hemoglobin < 90 g/L, platelets $< 100 \times 10^9/L$, and neutrophils $< 1.0 \times 10^9/L$.				
i Thrombocytopenic purpura, a condition characterized by a low platelet count (thrombocytopenia) leading to the appearance of purpura (small bleeding spots under the skin).				
j pEBV DNA refers to the viral DNA load of Epstein-Barr virus detected and quantified in the plasma compartment, usually measured by real-time polymerase chain reaction (PCR) and expressed as copies per milliliter (copies/mL) or $\log_{10}$ transformed copies/mL.				

LASSO analysis of the impact on acute death within 1 year in CAEBV

The 13 variables mentioned above were included in the LASSO analysis, wherein, as the penalty factor increases, the coefficients of the initially included variables are progressively shrunk. The results showed that when the lambda value was selected as lambda.min (0.042), 9 variables with non-zero coefficients were screened out. When the lambda value was selected as lambda.1se (0.074), 7 variables were identified as predictors: polyserositis, hypocytosis, pEBV DNA, LDH, PLT, ALT, and TBIL, all of which exhibited non-zero coefficients in the LASSO regression model (Fig. 2). Subsequently, multivariate Cox regression analysis confirmed that polyserositis (HR = 3.68, 95% CI = 1.89–7.18), cytopenia (HR = 3.29, 95%CI = 1.68–6.42), pEBV DNA (HR = 1.39, 95%CI = 1.16–1.66), LDH (HR = 1.01, 95%CI = 1.01–1.01), TBIL

(HR = 1.01, 95%CI = 1.01–1.02) were independent risk factors for CAEBV patients to predict acute mortality within 1 year.

Establishment of scoring systems for predicting acute death in CAEBV within 1 year.

A nomogram model for predicting acute death within one year in CAEBV patients was constructed based on five independent influencing factors: polyserosal effusion, cytopenia, pEBV DNA, LDH, and TBIL. Each parameter was scored individually, and the sum of these scores calculated the total score to estimate the risk of acute death in CAEBV patients (Fig. 3A). Each variable was assigned a weighted score proportional to its regression coefficient. For improved clinical applicability, continuous variables were grouped into predefined clinically relevant categories (e.g., pEBV DNA levels of ≤ 3 , 3–5, 5–7, and > 7 log copies/mL), and a simplified risk stratification score (range: 0–11 points) was developed (Fig. 3B).

Evaluation of a nomogram model for predicting acute death in CAEBV within 1 year.

Both the nomogram and the scoring system demonstrated sustained predictive performance for mortality across all time points (4, 8, and 12 months). Time-dependent ROC analysis yielded AUCs (95% CI) of 0.813 (0.733–0.893), 0.869 (0.802–0.936), and 0.894 (0.835–0.953) for the nomogram, and 0.805 (0.722–0.888), 0.838 (0.761–0.915), and 0.843 (0.768–0.917) for the score model, respectively, with all AUCs > 0.7 (Fig. 4A–C). Calibration curves showed good agreement between predicted and observed death probabilities (Fig. 4D–F). Decision curve analysis indicated favorable net clinical benefit for both models at clinically relevant threshold ranges: 7%–34%, 9%–79%, and 11%–87% for the nomogram at 4, 8, and 12 months, and 7%–36%, 10%–73%, and 11%–87% for the scoring system at corresponding time points (Fig. 4G–I). Internal cross-validation confirmed consistent performance, with C-indices of 0.813, 0.815, 0.823 (nomogram) and 0.796, 0.792, 0.785 (score model), and AUCs of 0.822, 0.879, 0.905 (nomogram) and 0.816, 0.844, 0.850 (score model), respectively, indicating excellent discrimination (Fig. 5). The net reclassification index and integrated discrimination improvement further demonstrated that the score model's discriminative ability was comparable to that of the nomogram (Table 2).

Table 2

Net reclassification improvement and integrated discrimination improvement index of the nomo diagram model and score model

Time	Item	Est.	Lower	Upper	p
4-month	IDI	-0.025	-0.093	0.056	0.468
	NRI	0.019	-0.310	0.332	0.374
	Median improvement	-0.141	-0.485	0.268	0.348
8-month	IDI	-0.025	-0.087	0.056	0.448
	NRI	0.006	-0.300	0.270	0.281
	Median improvement	-0.250	-0.545	0.173	0.209
1-year	IDI	-0.063	-0.117	0.018	0.129
	NRI	-0.133	-0.332	0.156	0.146
	Median improvement	-0.393	-0.647	0.114	0.119
*Abbreviations: IDI, Integrated Discrimination Improvement; NRI, Net Reclassification Improvement. The IDI and median improvement values were not statistically significant at the 4-month, 8-month, and 1-year time points. ($p > 0.05$), The NRI and median improvement show non-significant results at the 4-month, 8-month, and 1-year time points. ($p > 0.05$).					

Discussion

EBV infects more than 90% of adults worldwide, typically as an asymptomatic infection. Primary infection with this virus typically causes IM, which is usually self-limiting and has a favorable prognosis^{6, 7}. However, chronic active EBV infection is a serious condition characterized by the persistent proliferation of NK or T lymphocytes, systemic inflammation, and a high EBV DNA load in peripheral blood. Moreover, given that the clinical manifestations of CAEBV closely resemble those of EBV-related lymphomas, and considering that its treatment primarily involves chemotherapy and hormonal therapy—it is no longer classified merely as an infectious condition. Instead, it has been redefined as Epstein-Barr virus-associated Lymphoproliferative Disorders. Initially considered a disease primarily affecting children, CAEBV has been found through further studies to present more frequently in adults, who tend to exhibit a more severe prognosis and less favorable response to therapeutic interventions compared to children^{8, 9}. Although the annual incidence of CAEBV is extremely low, approximately 0.18 per 100,000 people in Japan, its mortality rate exceeds 80%^{9, 10}. The high mortality rate of CAEBV is primarily attributable to disease progression and severe complications rather than the viral infection itself^{11, 12}. The leading causes of death include hemophagocytic lymphohistiocytosis (HLH), multiple organ dysfunction syndrome (MODS), lymphoma, and others^{11, 13}. Our observed 1-year mortality rate of 27.2% is closely consistent with the nationwide Japanese survey, which reported a 3-year survival of only 58% and nearly 40% of deaths within 20 months of onset. Together, these findings confirm that CAEBV carries

an exceptionally high early fatality risk and follows a relentlessly progressive course, strongly emphasising the urgent need for timely therapeutic intervention. Based on accumulated evidence, earlier HSCT is universally recognized as the only strategy that improves survival in CAEBV patients, both within and outside Asia¹⁴. For patients who do not undergo transplantation, the prognosis remains extremely poor—one cohort of 149 non-transplanted CAEBV patients reported a 3-year mortality rate of 90.9%¹⁵. Even among those who receive HSCT, early outcomes have not always been encouraging; interestingly, a 2008 study documented that all seven transplanted patients in a small cohort died between 1 and 16 months post-transplantation, with a median survival of only 5 months¹⁶. In stark contrast, a 2020 study reported a survival rate as high as 83%, suggesting a significant improvement in outcomes for CAEBV⁹. We hypothesize that this apparent paradox may be attributed to two main factors: first, advances in pre-transplantation chemotherapeutic regimens; and second, a selection bias introduced by the evolving criteria for HSCT eligibility, leading to imbalances between the patient groups compared⁹. Together with the above findings, this figure strongly underscores that CAEBV is a rapidly progressive and extremely life-threatening disease. Hence, there is an urgent need to develop novel therapeutic approaches and more refined evaluation tools—not only to reduce early mortality but also to serve as a bridge to HSCT, ultimately improving overall survival in this challenging patient population.

Previous studies in CAEBV have primarily focused on comparing demographic differences, clinical features and treatment outcomes^{9, 11}. A study has identified age > 8 years as an independent predictor of poor prognosis³. However, an unresolved question remains: how to define risk factors specifically for patients older than 8 years, or even for adult populations? While some reports indicate a female predominance among elderly patients, a separate study of 42 T/NK cell proliferative disorders reported 79% male patients; overall, however, no consistent gender bias has been established^{9, 17}. In our cohort, we observed markedly elevated transaminase levels, especially in fatal cases, which aligns with prior evidence of EBV-related hepatic involvement^{9, 14}. Smoking is highly correlated with EBV capsid antigen and transcription activator protein seropositivity^{18, 19}. Furthermore, although a recent study identified smoking as a independent risk factor²⁰, our findings revealed no significant association between smoking history and prognosis. This suggests that while smoking may contribute to disease initiation, it maybe not a major driver of subsequent progression or severity. Cytopenias were the most common presenting sign (93/180, 51.7%), occurring at a slightly lower frequency than the 80% reported in non-Asian populations, a discrepancy potentially attributable to later-stage disease data collection in this cohort¹⁴. This was particularly notable in the acute mortality group in adult, underscoring the urgent need for novel early warning factors or predictive models.

The 2022 revised Japanese diagnostic criteria to the diagnostic criteria has recommended incorporating reduced NK-cell activity and EBV DNA > 10,000 IU/mL into the guidelines^{21, 22}. Previously, the diagnosis relied more on exclusion of other conditions, but the focus has now shifted toward obtaining evidence of viral proliferation and active disease. Notably, persistently high EBV-DNA loads and chronic, uncontrolled active viral replication are consistently linked to adverse outcomes. The clinical management of CAEBV,

encompassing screening, prevention, diagnosis, and treatment, is heavily reliant on EBV DNA load assessment^{20, 23, 24}, and our study focuses on plasma EBV DNA load, which demonstrates a stronger correlation with active disease burden than PBMC^{23, 25}. A study from Beijing Friendship Hospital has established an association between plasma EBV DNA and CAEBV complicated by hemophagocytic lymphohistiocytosis (HLH), further underscoring the clinical significance of plasma viremia in this setting^{5, 23}. As an EBV-associated malignancy, NPC has shown consistent correlation between post-treatment EBV clearance and improved survival outcomes. Serial cEBV DNA monitoring reveals progressive viral activity decline during treatment, with viral clearance predicting longer survival^{26, 27}. Furthermore, prior studies have also demonstrated that an EBV DNA load $> 2.88 \times 10^4$ IU/mL offers high diagnostic accuracy for lymphoma-associated HLH²⁸. CAEBV exhibits considerable heterogeneity in its cellular origin and clonal architecture, complicating its pathological classification. Emerging opinions even propose infection of a common progenitor cell^{29, 30}. This very complexity makes the use of conventional clinical and laboratory markers a more feasible and immediately applicable strategy for predicting patient outcomes in most healthcare settings. In patients from Asia and Central and South America, EBV primarily infects T or NK cells, with the T-cell type being more common and associated with a poorer prognosis^{8, 31}. Interestingly, another study indicates the mortality rates were comparable between two cell types¹⁴. A nationwide survey in Japan reported that only 14 of 91 patients were pathologically diagnosed with CAEBV⁹, highlighting the diagnostic challenges. Moreover, latest criteria still require a disease duration of more than three months. Although CAEBV diagnosis relies on EBER-ISH and EBV-DNA assays¹¹, post-diagnostic assessment maybe depend more heavily on virological and clinical indicators. Therefore, we developed a prognostic model based on a large cohort of 180 Chinese patients, which represents the largest such series reported to date. We prospectively assessed 1-year survival outcomes at the infectious disease departments of two major hospitals. Multivariate Cox regression analysis identified polyserositis, cytopenia, EBV DNA, lactate dehydrogenase (LDH), and total bilirubin (TBIL) as significant predictors of death risk, highlighting their combined diagnostic and prognostic value in this patient population. To facilitate the transition from risk identification to early intervention, we advocate the following: for patients with a total score > 5 , prompt organ support and vigilant monitoring are strongly advised; for those scoring exactly 5 (just below the threshold), regular check-ups and systematic therapy are also reasonable options.

While our study represents the largest reported Asian cohort to date with decade-long data, this study has several limitations. Its retrospective, two-center design may introduce inherent biases related to medical resources, economic status, and evolving diagnostic and treatment standards over its decade-long timeframe. Most notably, we were unable to compare pathological subtypes due to missing classification data, a direct consequence of regional healthcare disparities. Due to the limitations of local medical resources, this study did not conduct an in-depth investigation based on pathological types, genetic levels, or cytokine profiles. Despite these limitations, we successfully constructed a highly accurate predictive nomogram for CAEBV. We hope this model will aid clinicians in patient management, improve outcomes, and ultimately help reduce mortality. Future studies should integrate pathological,

genetic, and cytokine profiling with these robust clinical indicators to further refine prognostic stratification.

Conclusions

A nomogram and risk score model based on polyserosal effusion, cytopenia, pEBV DNA, LDH, and TBIL serve as effective, practical tools for individualized risk assessment and early intervention.

Abbreviations

CAEBV chronic active Epstein-Barr virus infection

HLH hemophagocytic lymphohistiocytosis

HSCT hematopoietic stem cell transplantation

IM infectious mononucleosis

ALT alanine aminotransferase

ALB albumin

TBIL total bilirubin

RF rheumatoid factor

CRP C-reactive protein

PT prothrombin time

APTT activated partial thromboplastin time

FIB fibrinogen

TG triglycerides

AFP alpha-fetoprotein

FER ferritin

pEBV DNA plasma Epstein-Barr virus DNA

LASSO least absolute shrinkage and selection operator

AUC area under the curve

DCA decision curve analysis

NRI net reclassification index

IDI integrated discrimination improvement

CI confidence interval

HR hazard ratio

ROC receiver operating characteristic

WHO World Health Organization

Declarations

Ethics statement and consent to participate

The current study was conducted by the Declaration of Helsinki. The study protocol were approved by the ethics committee of Tangdu Hospital of Fourth Military Medical University (K202407-33) and First Affiliated Hospital of Xi 'an Jiaotong University (XJTU1AF2025LSYY-635). All patients provided written informed consent.

Consent for publication

Not applicable.

Clinical trial registration

The study was registered on National Medical Research Registration Information System (registration number: MR-61-25-085353, release date: December 30, 2025)

Data availability statement

The data that support the findings of this study are available from the corresponding author [Jian-Qi Lian] upon reasonable request.

Competing interests

They authors declare no potential conflicts of interests.

Funding

The authors declare that they have no conflicts of interest. This work was supported by the National Natural Science Foundation of China (82402516). The funders had no role in the design and implementation of the study or preparation of the manuscript.

Authors' contributions

Study conception, design, and methodology were contributed by Fei Hu, Jiayi Zhan, Rui Li, Yingli He, and Jianqi Lian. Data collection was performed by Wen Tian, Haifeng Hu, Pingping Zhu, Xingchi Chen, and Meijuan Peng,. Data analysis and interpretation were performed by Fei Hu, Fangfang Li, Jingtao Hou, Xiaofei Yang, and Xuyang Zheng. Graph and table visualization was performed by Rui Li, Danfeng Ren, Yan Liang, and Lina Yan. Study supervision was provided by Jianqi Lian, Yingli He, and Xiaofei Yang. The original draft of the manuscript was written by Fei Hu, Hong Du, Denghui Yu, and Hongkai Xu. All authors have participated in the revision process, critically reviewed the manuscript, and given final approval of the version to be published.

Acknowledgment

We are grateful to thank all of the participants, the staff, and the other study investigators for their valuable contributions. Additionally, we also thank the entire health-care team for their diagnosis and treatment of patients with CAEBV.

References

1. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375–90.
2. Alaggio R, Amador C, Anagnostopoulos I, Attygalle A, Araujo I, Berti E, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia*. 2022;36(7):1720–48.
3. Kimura H, Ito Y, Kawabe S, Gotoh K, Takahashi Y, Kojima S, et al. EBV-associated T/NK-cell lymphoproliferative diseases in nonimmunocompromised hosts: prospective analysis of 108 cases. *Blood*. 2012;119(3):673–86.
4. Ichiro Y, Chizuko S, Ken-Ichi I, Tohru K, Masahide Y, Akihisa S et al. Nationwide survey of systemic chronic active EBV infection in Japan in accordance with the new WHO classification. *Blood Adv*. 2020;4(13).
5. Jun-Ichi K, Yoshinori I, Koichi O, Masaki Y, Shinsuke K, Hideki M et al. Updated guidelines for chronic active Epstein-Barr virus disease. *Int J Hematol*. 2023;118(5).
6. Nader Kim E-M, Cholladda VC, Carl EA. Haemophagocytic lymphohistiocytosis and Epstein-Barr virus: a complex relationship with diverse origins, expression and outcomes. *Br J Haematol*. 2021;196(1).
7. Graham ST, Heather ML, Jill MB, Alan BR, Andrew DH. The immunology of Epstein-Barr virus-induced disease. *Annu Rev Immunol*. 2015;33.
8. Bollard CM, Cohen JL. How I treat T-cell chronic active Epstein-Barr virus disease. *Blood*. 2018;131(26):2899–905.
9. Arai A, Fujiwara S, Kimura H, Ohga S, Miura O, Ishimura M, et al. Nationwide survey of systemic chronic active EBV infection in Japan in accordance with the new WHO classification. *Blood Adv*. 2020;4(13):2918–26.

10. Kimura H, Ito Y, Kawabe S, Gotoh K, Takahashi Y, Kojima S, et al. EBV-associated T/NK-cell lymphoproliferative diseases in nonimmunocompromised hosts: prospective analysis of 108 cases. *Blood*. 2012;119(3):673–86.
11. Nobuya A, Yuichiro F. Chronic active Epstein-Barr virus infection. *Blood*. 2020;136(18).
12. Anders F, Christen Lykkegaard A. Epstein-Barr virus and its association with disease - a review of relevance to general practice. *BMC Fam Pract*. 2019;20(1).
13. Qazi B, Tang K, Qazi AJ. Recent advances in underlying pathologies provide insight into interleukin-8 expression-mediated inflammation and angiogenesis. *Int J Inflam*. 2011;2011:908468.
14. Blachy JDS, Tami J, Challice B, David B, Sharat C, Shanmuganathan C et al. High risk of relapsed disease in patients with NK/T-cell chronic active Epstein-Barr virus disease outside of Asia. *Blood Adv*. 2021;6(2).
15. Zhang Q, Zhao Y-Z, Wang D, Ma H-H, Cui L, Li W-J, et al. Chronic active Epstein-Barr virus disease in children: Clinical outcomes and prognostic risk factors from a 149-patient cohort. *Br J Haematol*. 2026;208(5):1704–12.
16. Kensei G, Yoshinori I, Yukiko S-W, Jun-ichi K, Yoshiyuki T, Hiroshi Y et al. Clinical and virological characteristics of 15 patients with chronic active Epstein-Barr virus infection treated with hematopoietic stem cell transplantation. *Clin Infect Dis*. 2008;46(10).
17. Paik JH, Choe J-Y, Kim H, Lee J-O, Kang HJ, Shin HY, et al. Clinicopathological categorization of Epstein–Barr virus-positive T/NK-cell lymphoproliferative disease: an analysis of 42 cases with an emphasis on prognostic implications. *Leuk Lymphoma*. 2016;58(1):53–63.
18. Aguayo F, Boccardo E, Corvalán A, Calaf GM, Blanco R. Interplay between Epstein-Barr virus infection and environmental xenobiotic exposure in cancer. *Infect Agent Cancer*. 2021;16(1):50.
19. Dolcetti R, Dal Col J, Martorelli D, Carbone A, Klein E. Interplay among viral antigens, cellular pathways and tumor microenvironment in the pathogenesis of EBV-driven lymphomas. *Semin Cancer Biol*. 2013;23(6):441–56.
20. Xiameng L, Liru Z, Bibo T, Zhuowei L, Kaihua C, Suning H et al. Early circulating EBV-DNA clearance predicts enhanced short-term treatment response and survival benefit in locally advanced nasopharyngeal carcinoma. *BMC Cancer*. 2025;25(1).
21. Quintanilla-Martinez L, Swerdlow SH, Tousseyn T, Barrionuevo C, Nakamura S, Jaffe ES. New concepts in EBV-associated B, T, and NK cell lymphoproliferative disorders. *Virchows Arch*. 2022;482(1):227–44.
22. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al. editors. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Revised 4th ed. Lyon: IARC Press; 2017.
23. Jun-Ichi K, Yasuko K, Akihisa S, Keiji I, Koji I, Yuka T et al. Viral DNA Loads in Various Blood Components of Patients With Epstein-Barr Virus-Positive T-Cell/Natural Killer Cell Lymphoproliferative Diseases. *J Infect Dis*. 2019;220(8).
24. Shaoyan L, Huaqiang Z, Gang C, Jinhui X, Qianwen L, Jianing L et al. Early change of plasma Epstein-Barr virus DNA load and the viral lytic genome level could positively predict clinical outcome in

- recurrent or metastatic nasopharyngeal carcinoma receiving anti-programmed cell death 1 monotherapy. *BMC Cancer*. 2024;24(1).
25. Baldanti F, Gatti M, Furione M, Paolucci S, Tinelli C, Comoli P, et al. Kinetics of Epstein-Barr virus DNA load in different blood compartments of pediatric recipients of T-cell-depleted HLA-haploidentical stem cell transplantation. *J Clin Microbiol*. 2008;46(11):3672–7.
 26. Shao-Bo L, Ning Z, Dan-Ming C, Xing-Li Y, Bin-Hong C, Hai Z et al. Prognostic value of gross tumor regression and plasma Epstein Barr Virus DNA levels at the end of intensity-modulated radiation therapy in patients with nasopharyngeal carcinoma. *Radiother Oncol*. 2018;132.
 27. Sha-Sha H, Yan W, Yong B, Xiu-Yu C, Xing-Li Y, Dan-Ming C et al. Dynamic changes in plasma Epstein-Barr virus DNA load during treatment have prognostic value in nasopharyngeal carcinoma: a retrospective study. *Cancer Med*. 2018;7(4).
 28. Zhang J, Qin S, Jin Z, Chen Q, Xing L, Qiu T, et al. The Clinical Significance and Prognostic Role of Whole-Blood Epstein-Barr Virus DNA in Lymphoma-Associated Hemophagocytic Lymphohistiocytosis. *J Clin Immunol*. 2023;43(6):1302–10.
 29. Okuno Y, Murata T, Sato Y, Muramatsu H, Ito Y, Watanabe T, et al. Defective Epstein-Barr virus in chronic active infection and haematological malignancy. *Nat Microbiol*. 2019;4(3):404–13.
 30. Murata T, Okuno Y, Sato Y, Watanabe T, Kimura H. Oncogenesis of CAEBV revealed: Intragenic deletions in the viral genome and leaky expression of lytic genes. *Rev Med Virol*. 2020;30(2):e2095.
 31. Leticia Q-M, Steven HS, Thomas T, Carlos B, Shigeo N, Elaine SJ. New concepts in EBV-associated B, T, and NK cell lymphoproliferative disorders. *Virchows Arch*. 2022;482(1).

Figures

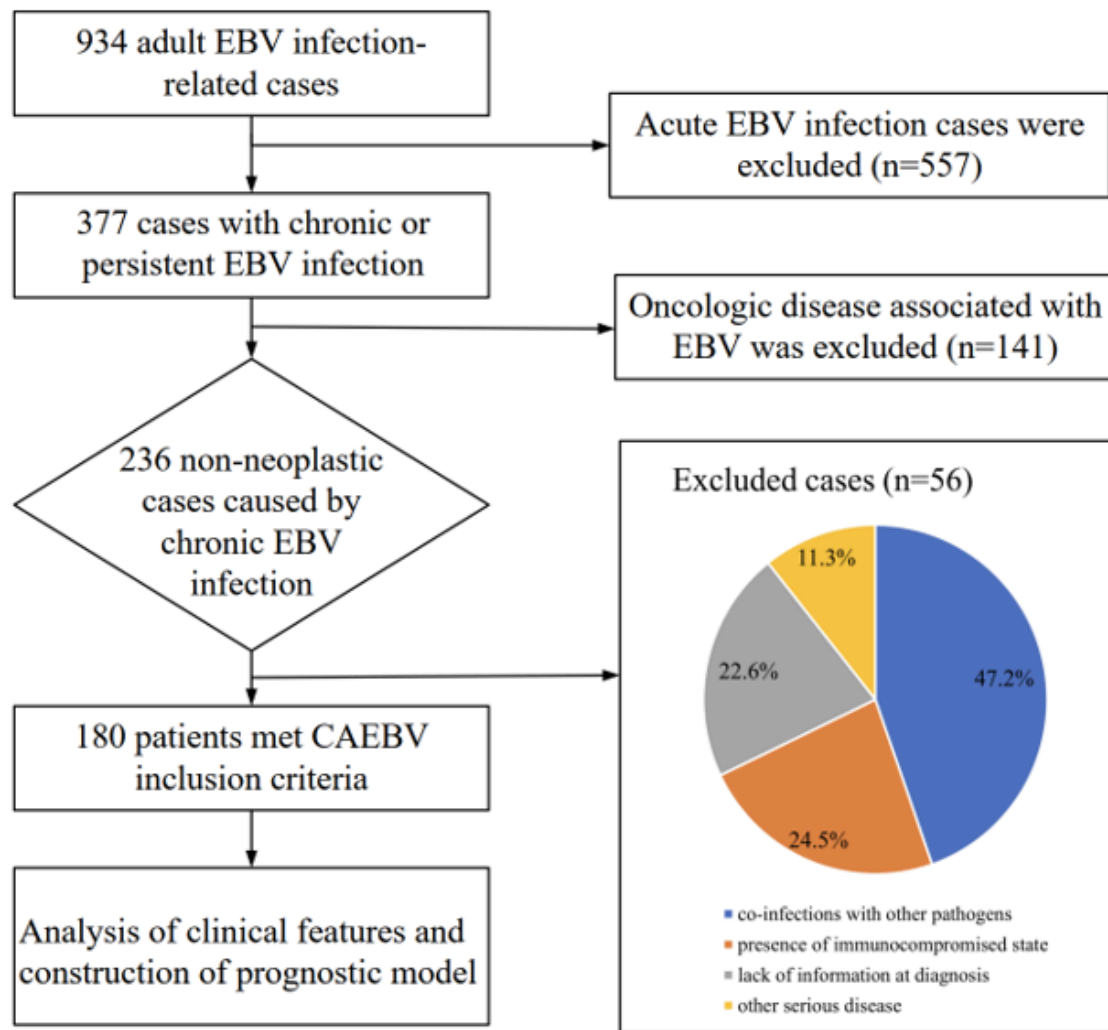


Figure 1

Flowchart of patient collection and reasons for the exclusion of some samples.

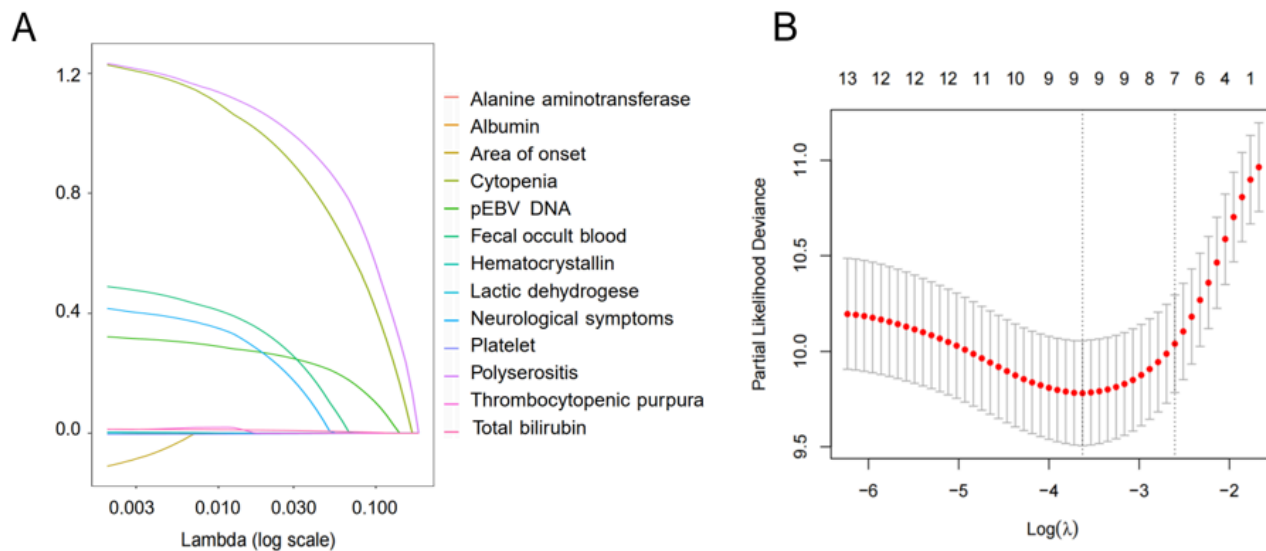
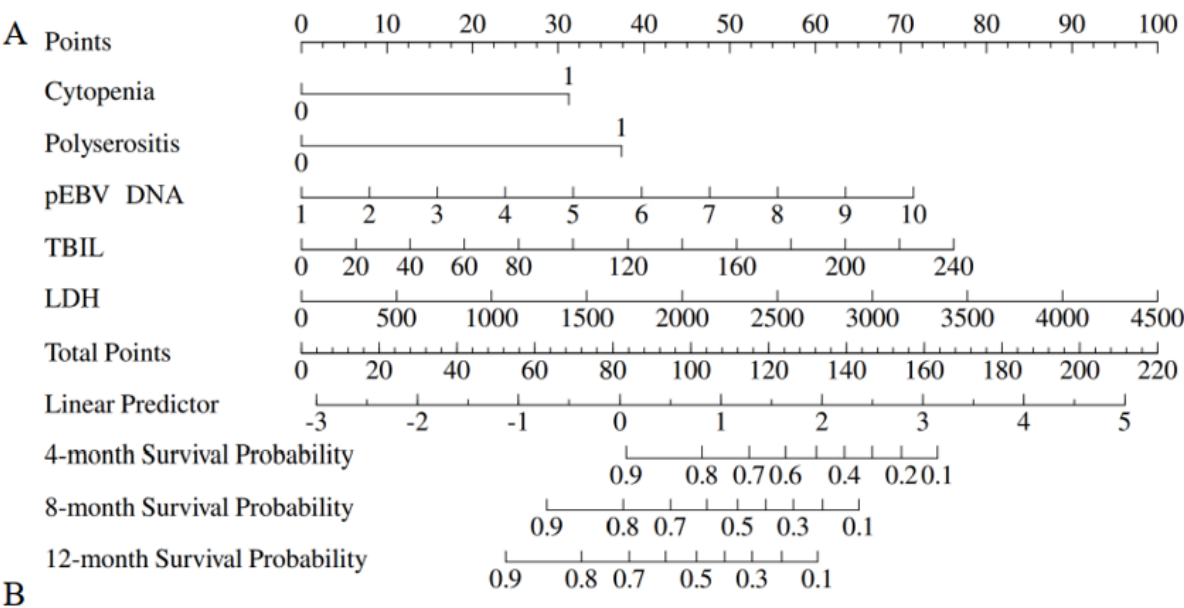


Figure 2

Texture feature selection using the least absolute shrinkage and selection operator (LASSO) binary logistic regression model. (A) Coefficient paths of different variables using the LASSO model: seven variables with nonzero coefficients were chosen by the optimal lambda. (B) LASSO shrinkage parameter λ selected by 10- fold cross-validation with minimum criteria. The Mean-Squared Error (red dots) and its 95% error bar (grey interval) was plotted against $\ln(\lambda)$ sequence. Vertical line shown the optimal value of λ selected by the minimum criteria (1-SE criteria: the regularized model that the cross-validated error is within one standard error of the minimum). A shrinkage factor λ of 0.074, with $\ln(\lambda)=- 2.60$ was chosen for optimal parameter.

(C) Nomogram model predicts acute death in CAEBV patients within 1 year.



B

Score	0	1	2	3	Total Score and Death Risk				
Predictor	0	1	2	3					
Cytopenia	no	yes	/	/	Scale	Time	4-month	8-month	12-month
Polyserositis	no	yes	/	/		Score			
TBIL	≤20	20-100	100-200	>200		3	7.7%	14.3%	18.7%
LDH	≤245	245-490	490-735	>735		5	19.3%	33.8%	42.4%
pEBV-DNA	≤3	3-5	5-7	>7		7	42.4%	65.3%	75.8%

Figure 3

The nomogram model and risk scoring scale were developed to predict progression from CAEBV to death within 1 year. (A) Nomogram of the prediction model: Points on the top scale assign weights to each predictor. The sum of all predictor points yields the total risk score, which is then projected onto the bottom scale to estimate the 1-year probability of death in CAEBV patients. (B) Risk stratification score: Variable scores are weighted according to their corresponding nomogram points, and the total score is

calculated by summing all individual variable scores. The 1-year risk probabilities of death (displayed on the right) were derived from Cox regression analysis. Based on this scoring scale, the risk of death in CAEBV patients is stratified into distinct clinical categories.

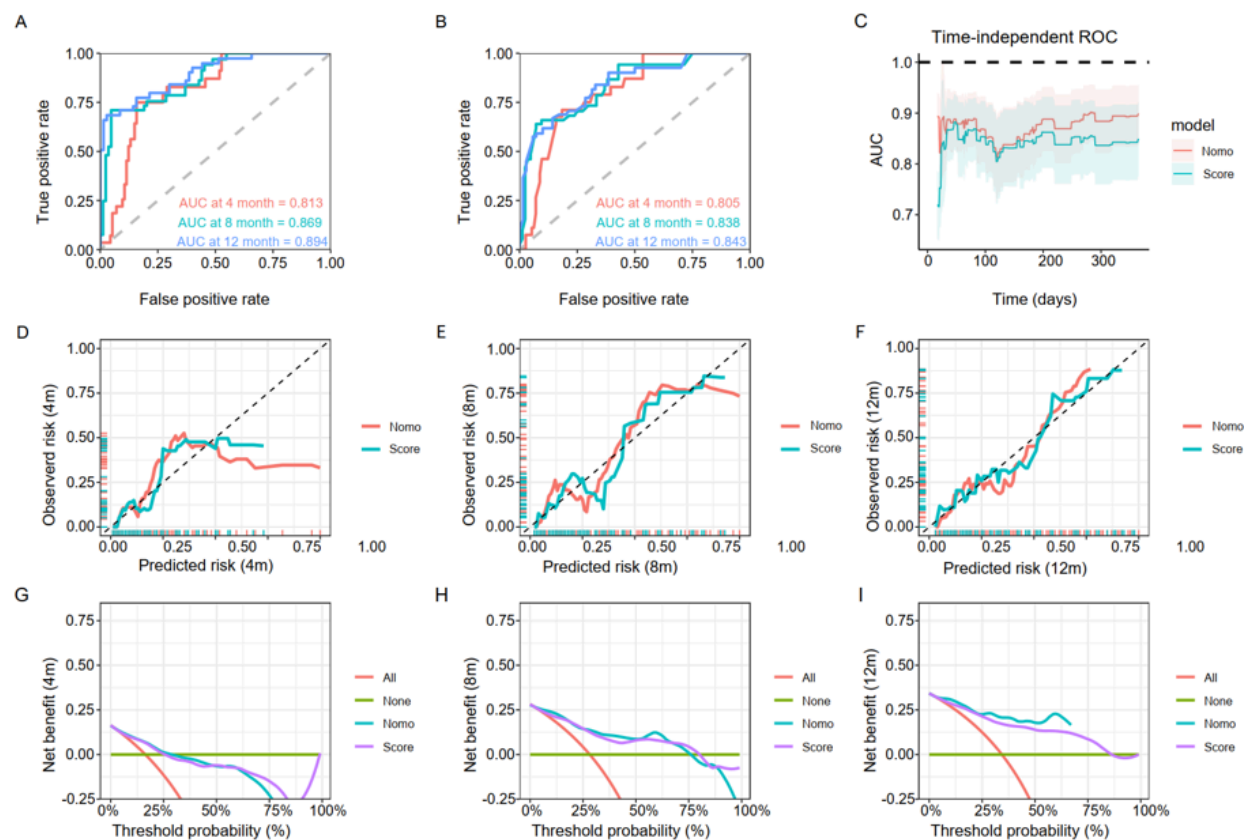


Figure 4

Comparison of the Nomo and Score Models Using ROC, Calibration, and Decision Curve Analyses (A-B) Receiver operating characteristic (ROC) curves showing the discriminative performance of Nomo and Score models at 4, 8, and 12 months in the CAEBV cohort; (C) Time-dependent ROC curve of the nomogram model. (D–F) Calibration plots of the nomogram at 4 months (D), 8 months (E), and 12 months (F) in the CAEBV cohort. (G–I) Decision curve analysis plots of the nomogram at 4 months (G), 8 months (H), and 12 months (I) in the CAEBV cohort. The decision curves plot the standardized net benefit (y-axis) across a range of risk thresholds (x-axis) for three strategies: intervene on all patients ("All"), intervene on none ("None"), and intervene based on the nomogram's predicted risk ("Nomogram").

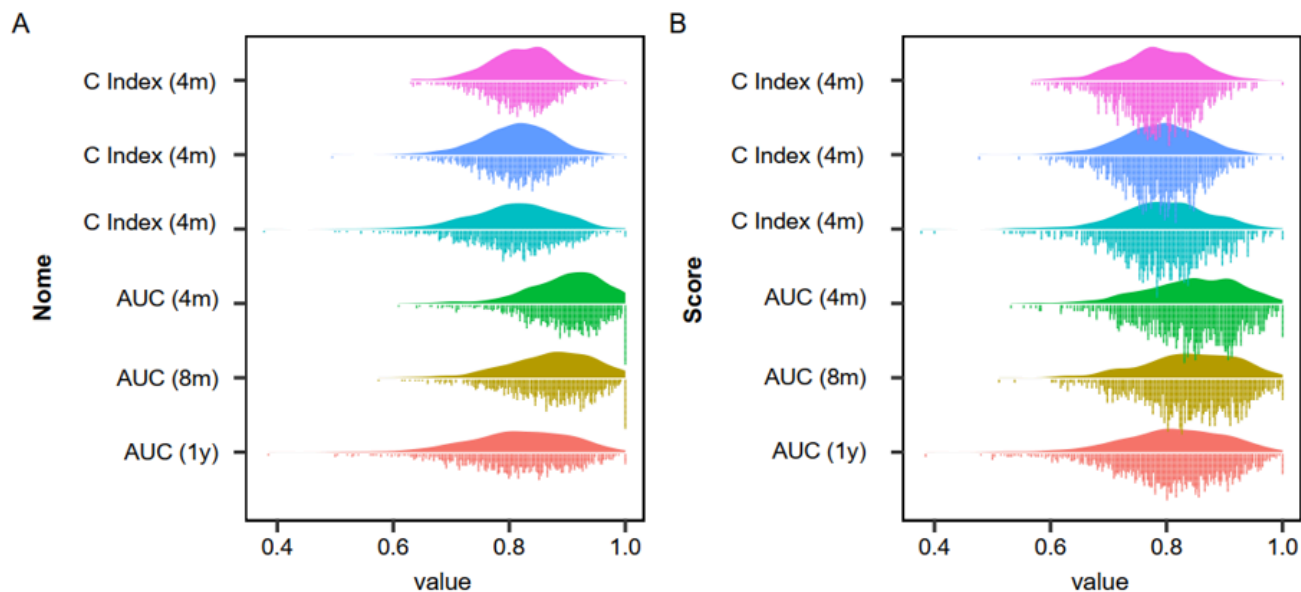


Figure 5

The internally cross-validated cloud-and-rain plot compares the Nomo and Score models. Using 10-fold cross-validation repeated 2000 times, we evaluated the predictive performance of both models for death in CAEBV patients. The internally validated C-index (concordance index) measures the discriminatory ability of each model at 4 months, 8 months, and 1 year. Additionally, the time-dependent AUC derived from internal cross-validation quantifies the models' discriminatory performance across multiple data splits over the same three prediction horizons (4 months, 8 months, and 1 year).