Supplementary material 1

Linear Mixed-Effect Model (LME):

The linear mixed effects model is an important analytical method for processing longitudinal data, which allows estimation of model parameters that describe the inter-subject and intra-subject variability of individual responses. Its basic form is as follows:

$$y_{ij} = m_i(t) + \varepsilon_{ij} = b_{0i} + b_{1i}t_{ij} + \sum_{k=1}^{k} b_k x_k + \varepsilon_{ij} = b_{00} + \mu_{0i} + (b_{10} + \mu_{0i})t_{ij} + \sum_{k=1}^{k} b_k x_k + \varepsilon_{ij}$$

y_{ij} represents the observed value of subject i at the JTH follow-up visit, $m_i(t)$ is the true measurement approximate by y_{ij}, ε_{ij} is the random error, satisfying $\varepsilon_{ij} \sim N(0,\sigma 2)$, b_{0i} and b_{1i} are the outcome and slope of time t of subject i . b_{0i} can be decomposed into b_{00} and μ_{0i}, b_{00} is the average intercept of all subjects, μ_{0i} is the random effect of intercept, and can be expressed as the variation of intercept between individuals. b_{1i} can be decomposed into b_{10} and μ_{1i}, where b_{10} is the average slope of all subjects and μ_{1i} is the random effect of slope, representing the variation of slope among individual subjects. b_{00} and b_{10} are the fixed effects part of the model, μ_{0i}, μ_{1i}, and ε_{ij} represent the random effects part of the model, x_k represents the fixed covariates that do not change over time, such as gender, baseline age at enrollment, and b_k is its slope. The JM package in R is used to build a linear mixed effect model.

Joint modelling of longitudinal and survival data (JMLS):

The JMLS model is a Joint model that connects the LME model with the Cox model using the association coefficient, which allows the longitudinal repeated covariates and the event occurrence time data to be modeled together. It can analyze the impact of data measures at the subject's baseline on survival risk, and the impact of changes in longitudinal data measures over time on survival risk. Combining these two seed models, the basic form of the combined model is as follows:

$$\begin{cases} y_{ij} = m_i(t) + \varepsilon_{ij} = b_{0i} + b_{1i}t_{ij} + \sum_{k=1}^{k} b_k x_k + \varepsilon_{ij} = b_{00} + \mu_{0i} + (b_{10} + (b_{10} + \mu_{0i})t_{ij} + \sum_{k=1}^{k} b_k x_k + \varepsilon_{ij} \\ h_i\left[[t|M_i(t), x_i\right] = h_0(t)exp\left[\sum \beta_i x_i + \alpha m_i(t)\right] \end{cases}$$

$h_0(t)$ represents the risk function under the subject's baseline state, that is, the risk probability of an individual subject having a intercept event at time t when all

independent variables x are 0 , $\sum\beta_i x_i$ is the cumulative risk function. Represents the cumulative risk of an outcome event occurring until time t is terminated, and α is defined as the joint coefficient of the joint model, meaning the effect of changes in mi(t) on $h_i(t)$.

Supplementary material 2

Variables settings:

LME and JMLS models for the total population:

LME model uses the ratio as the dependent variable, with time, Cage, state, gender, liver cirrhosis, and their interactions (Time:cage, Time:state) as independent variables.

JMLS approach consists of two components: (1). The LME sub-model, which takes the ratio as the dependent variable and includes time, Cage, gender, state, liver cirrhosis, and their interactions (Time:cage) as independent variables. (2). The COX sub-model, which uses state and time as the dependent variables, with baseline ratio, Cage, Gender, and liver cirrhosis as independent variables.

LME and JMLS models for the males and females:

LME model uses the ratio as the dependent variable, with time, Cage, state, liver cirrhosis, and their interactions (Time:cage, Time:state) as independent variables.

JMLS approach consists of two components: (1). The LME sub-model, which takes the ratio as the dependent variable and includes time, Cage, state, liver cirrhosis, and their interactions (time:Cage) as independent variables. (2). The COX sub-model, which uses state and time as the dependent variables, with baseline ratio, Cage, gender, and liver cirrhosis as independent variables

Supplementary Table 1 fundamental descriptions of the variables in the model

Variables	Description
Id	Number
Time	Interval between each detection time and the first detection time,(year)
Etime	Time interval between last test and first test, (year)
Cage	Baseline age after centralization (Cage= Baseline age−Median age of subjects)

Liver cirrhosis	1(cirrhosis at baseline), 0 (no cirrhosis at baseline)
Gender	1(male),2(female)
State	1(HCC),0(non-HCC)
Ratio	Ln(10APR)

Supplementary Table 2 Model equations for all models

	LME models
LME_1	$\ln(10\text{APR}) = \beta_0 + \beta_1(Liver\ cirrhosis)_{ij} + \beta_2 gender_{ij} + \beta_3 \text{Cage}_{ij} + \beta_4 \text{state}_{ij} + \beta_5(\text{time:Cage})_{ij} + \beta_6(\text{time:state})_{ij} + \varepsilon_{ij}$
LME_2	$\ln(10\text{APR}) = \beta_0 + \beta_1(Liver\ cirrhosis)_{ij} + \beta_2 \text{Cage}_{ij} + \beta_3 \text{state}_{ij} + \beta_4(\text{time:Cage})_{ij} + \beta_5(\text{time:state})_{ij} + \varepsilon_{ij}$
LME_3	$\ln(10\text{APR}) = \beta_0 + \beta_1(Liver\ cirrhosis)_{ij} + \beta_2 \text{Cage}_{ij} + \beta_3 \text{state}_{ij} + \beta_4(\text{time:Cage})_{ij} + \beta_5(\text{time:state})_{ij} + \varepsilon_{ij}$
	JMLS Models
$JMLS_1$	$\begin{cases} \ln(10\text{APR}) = \beta_0 + \beta_1(Liver\ cirrhosis)_{ij} + \beta_2 gender_{ij} + \beta_3 \text{Cage}_{ij} + \beta_4 \text{state}_{ij} + \beta_5(\text{time:Cage})_{ij} + \beta_6(\text{time:state})_{ij} + \varepsilon_{ij} \\ h_i[[t\|M_i(t), x_i] = h_0(t)exp[\beta_0 \ln(10\text{APR}) + \beta_1(Liver\ cirrhosis)_{ij} + \beta_2 gender_{ij} + \beta_3 \text{C} \end{cases}$
$JMLS_2$	$\begin{cases} \ln(10\text{APR}) = \beta_0 + \beta_1(Liver\ cirrhosis)_{ij} + \beta_2 \text{Cage}_{ij} + \beta_3 \text{state}_{ij} + \beta_4(\text{time:Cage})_{ij} + \beta_5(\text{time:state})_{ij} + \varepsilon_{ij} \\ h_i[[t\|M_i(t), x_i] = h_0(t)exp[\beta_0 \ln(10\text{APR}) + \beta_1(Liver\ cirrhosis)_{ij} + \beta_2 gender_{ij} + \beta_3 \text{C} \end{cases}$
$JMLS_3$	$\begin{cases} \ln(10\text{APR}) = \beta_0 + \beta_1(Liver\ cirrhosis)_{ij} + \beta_2 \text{Cage}_{ij} + \beta_3 \text{state}_{ij} + \beta_4(\text{time:Cage})_{ij} + \beta_5(\text{time:state})_{ij} + \varepsilon_{ij} \\ h_i[[t\|M_i(t), x_i] = h_0(t)exp[\beta_0 \ln(10\text{APR}) + \beta_1(Liver\ cirrhosis)_{ij} + \beta_2 gender_{ij} + \beta_3 \text{C} \end{cases}$

LME: Linear Mixed-Effect Mode; LME1: Linear Mixed-Effect Mode for the overall population; LME2: Linear Mixed-Effect Mode for males; LME3: Linear Mixed-Effect Mode for females; JMLS: Joint Modelling of Longitudinal and Survival data; JMLS1: Joint Modelling of Longitudinal and Survival data for the overall population; JMLS2: Joint Modelling of Longitudinal and Survival data for males; JMLS3: Joint Modelling of Longitudinal and Survival data for females.

Supplementary Table 3 Abbreviations in English

Abbreviation	Full name in English
CHB	Chronic Hepatitis B
HCC	Hepatocellular carcinoma
non-HCC	non-Hepatocellular carcinoma
ALP	Alkaline phosphatase
PA	Prealbumin
APR	Alkaline phosphatase to prealbumin ratio
PLT	Thrombocyte
WBC	white blood cell
RBC	Red blood cells
Hb	Hemoglobin
NE	Neutrophil
LYM	lymphocyte
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
TbIL	Total bilirubin
DbIL	Direct Bilirubin
GGT	Gamma-glutamyltransferase
PY	Person-year

Supplementary Table 4 Baseline comparison of the effectively cohort with the baseline cohort

Variable	Effectively cohort (n = 4843)	Original cohort (n = 5143)	P value
Male （%）	63.74	63.80	0.96
Liver cirrhosis (%)	35.13	34.51	0.52
Age	48.83	48.85	0.93

	(48.83±11.58)	(48.85±11.63)	
PLT(g/L)	148.93 (148.93±69.26)	149.82 (149.82±69.26)	0.55
WBC(10^9/L)	4.86 (4.87±1.87)	4.88 (4.88±1.88)	0.74
RBC(10^12/L)	4.54 (4.54±0.68)	4.60 (4.53±0.68)	0.70
Hb(g/L)	140.04 (140.04 ± 21.91)	139. 90 (139.9±21.95)	0.75
NE(10^9/L)	2.73 (2.73±1.42)	2.74 (2.74±1.43)	0.75
LYM(g/L)	1.66 (1.66±0.73)	1.66 (1.66±0.73)	0.97
ALT(U/L)	84.06 (84.06±177.18)	87.59 (87.62±184.13)	0.33
AST(U/L)	74.44 (74.44±124.11)	75.49 (75.49±123.75)	0.67
TbIL(umol/L)	22.17 (22.17±29.13)	23.30 (23.29 ± 32.11)	0.07
DbIL(umol/L)	11.28 (11.28±22.52)	12.27 (12.27±25.34)	0.04
GGT(U/L)	83.96 (83.96±119.63)	85.37 (48.9±11.69)	0.56
ALP(U/L)	81.95 (81.95±48.03)	82.56 (82.56±47.96)	0.53
PA(mg/L)	178.09 (178.09±81.61)	177.51 (177.51±82.82)	0.33

NOTES: The numbers in the table are mean (mean ± SD).

Supplementary Table 5 Alkaline phosphatase to prealbumin ratio levels in

hepatocellular carcinoma and non-hepatocellular carcinoma groups during follow-up

Follow-up years	Male		Female	
	HCC	**non-HCC**	**HCC**	**non-HCC**
0	0.69 (0.37, 1.35)	0.37 (0.25, 0.68)	0.74 (0.49, 1.42)	0.39 (0.27, 0.64)
1	0.56 (0.34, 1.07)	0.30 (0.22, 0.46)	0.72 (0.45, 1.21)	0.35 (0.25, 0.51)
2	0.44 (0.29, 0.94)	0.28 (0.21, 0.39)	0.56 (0.41, 1.06)	0.32 (0.24, 0.45)
3	0.41 (0.28, 0.71)	0.27 (0.21, 0.37)	0.55 (0.42, 1.02)	0.31 (0.24, 0.44)
4	0.45 (0.30, 0.68)	0.27 [0.20, 0.38]	0.50 (0.41, 1.16)	0.31 (0.24, 0.44)
5	0.37 (0.27, 0.58)	0.26 (0.20, 0.35)	0.48 (0.35, 0.94)	0.31 (0.23, 0.42)
≧6	0.43 (0.29, 0.69)	0.26 (0.20, 0.36)	0.56 (0.38, 0.84)	0.30 (0.22, 0.42)

The numbers in the Table are median [P25, P75]

Supplementary Table 6 Linear Mixed-Effect Mode for the overall population Results

Variable	**β**	**SE**	**t-value**	***P* value**
Intercept	1.147	0.025	46.787	$P < 0.01$
Liver cirrhosis	0.455	0.017	26.506	$P < 0.01$
Gender	0.085	0.016	5.261	$P < 0.01$
Time	−0.093	0.002	−38.908	$P < 0.01$
Cage	−0.002	0.001	−3.106	$P < 0.01$

State	0.314	0.037	8.583	$P < 0.01$
Time:cage	0.003	0.002	17.336	$P < 0.01$
Time:state	0.006	0.009	0.657	0.511

Intercept: The average intercept of all subjects; Liver cirrhosis: Whether cirrhosis had occurred at baseline; Time: Interval between each detection time and the first detection time(year); Cage: Baseline age after centralization; State: Outcome state; Time:Cage: The interaction of time and Cage; Time:State: The interaction of time and state; β: Fixed-effect regression coefficient.

Supplementary Table 7 Linear Mixed-Effect Mode for male Results

Variable	**β**	**SE**	**t-value**	***P* value**
Intercept	1.234	0.015	83.421	*P*<0.01
Liver cirrhosis	0.446	0.022	20.173	*P*<0.01
Time	−0.093	0.003	−29.938	*P*<0.01
Cage	0.006	0.001	−6.526	*P*<0.01
State	0.325	0.045	7.172	*P*<0.01
Time:Cage	0.004	0.000	16.075	*P*<0.01
Time:state	0.006	0.011	0.554	0.580

Intercept: The average intercept of all subjects; Liver cirrhosis: Whether cirrhosis had occurred at baseline; Time: Interval between each detection time and the first detection time(year); Cage: Baseline age after centralization; State: Outcome state; Time:Cage: The interaction of time and Cage; Time:State: The interaction of time and state; β: Fixed-effect regression coefficient.

Supplementary Table 8 Linear Mixed-Effect Mode for female model Results

Variable	β	SE	*t* value	*P* value
Intercept	1.292	0.016	81.834	*P*<0.01
Liver cirrhosis	0.458	0.027	16.987	*P*<0.01
Time	−0.089	0.004	−23.770	*P*<0.01
Cage	0.005	0.004	4.693	*P*<0.01
State	0.317	0.006	5.092	*P*<0.01
Time:Cage	0.002	0.000	6.863	*P*<0.01
Time:state	−0.002	0.016	−0.130	0.897

Intercept: The average intercept of all subjects; Liver cirrhosis: Whether cirrhosis had occurred at baseline; Time: Interval between each detection time and the first detection time(year); Cage: Baseline age after centralization; State: Outcome state; Time:Cage: The interaction of time and Cage; Time:State: The interaction of time and state; β: Fixed-effect regression coefficient.

Supplementary Table 9 Joint Modelling of Longitudinal and Survival data for the overall population model results

Variable	β	SE	*Z* value	*P* value	HR	HR 95%CI
Longitudinal sub-model						
fixed effects						
Intercept	1.187	0.018	65.996	*P*<0.01		

Liver cirrhosis	0.531	0.012	45.583	*P*<0.01		
Gender	0.053	0.013	4.191	*P*<0.01		
Time	−0.091	0.002	−37.784	*P*<0.01		
Cage	−0.002	0.001	−2.619	*P*<0.01		
Time:cage	0.003	0.000	17.841	*P*<0.01		
Random effects						
∂b_{00}	0.603					
∂b_{10}	0.114					
ε	0.358					
Survival sub-model						
Ratio	0.058	0.070	0.819	0.413	1.069	(0.923-1.216)
Cage	0.067	0.006	11.648	*P*<0.01	1.069	(1.057-1.216)
Gender	−0.735	0.122	−6.007	*P*<0.01	0.480	(0.377-0.610)
Liver cirrhosis	1.519	0.155	9.836	*P*<0.01		(1.749-2.439)
α	0.725	0.085	8.537	*P*<0.01	1.069	(0.923-1.216)

Intercept: The average intercept of all subjects; Liver cirrhosis: Whether cirrhosis had occurred at baseline; Time: Interval between each detection time and the first detection time(year); Cage: Baseline age after centralization; State: Outcome state; Time:Cage: The interaction of time and Cage; Time:State: The interaction of time and state; β: Fixed-effect regression coefficient.

Supplementary Table 10 Joint Modelling of Longitudinal and Survival data ($JMLS_2$) for

male model Results

Variable	β	SE	*Z* value	*P* value	HR	HR 95%CI
Longitudinal sub-model						
fixed effects						
Intercept	1.124	0.012	107.562	*P*<0.01		
Liver cirrhosis	0.489	0.014	35.412	*P*<0.01		
Time	−0.091	0.003	−28.211	*P*<0.01		
Cage	−0.005	0.001	−6.371	*P*<0.01		
time:Cage	0.004	0.000	16.667	*P*<0.01		
Random effects						
∂b_{00}	0.640					
∂b_{10}	0.122					
ε	0.371					
Survival sub-model						
Ratio	0.038	0.081	0.416	0.677	1.034	(0.882-1.213)
Cage	0.066	0.006	10.190	*P*<0.01	1.068	(1.055,1.082)
Liver cirrhosis	1.589	0.182	8.732	*P*<0.01	4.896	(3.428,6.994)
α	0.695	0.098	7.068	*P*<0.01	2.005	(1.653-2.431)

Intercept: The average intercept of all subjects; Liver cirrhosis: Whether cirrhosis had occurred at baseline; Time: Interval between each detection time and the first detection time(year); Cage: Baseline age after centralization; State: Outcome state; Time:Cage: The interaction of time and Cage; Time:State: The interaction of time and state; β: Fixed-effect

regression coefficient.

Supplementary Table 11 Joint Modelling of Longitudinal and Survival data ($JMLS_2$) for female model Results

Variable	β	SE	*Z* value	*P* value	HR	HR 95%CI
Longitudinal sub-model						
fixed effects						
Intercept	1.287	0.013	100.722	*P*<0.01		
Liver cirrhosis	0.533	0.014	37.920	*P*<0.01		
Time	-0.093	0.003	-27.907	*P*<0.01		
Cage	-0.007	0.001	-7.723	*P*<0.01		
Time:cage	0.002	0.000	7.564	*P*<0.01		
Random effects						
∂b_{00}	0.537					
∂b_{10}	0.101					
ε	0.371					
Survival sub-model						
Ratio	0.139	0.145	0.958	0.338	1.149	(0.865-1.526)
Cage	0.068	0.012	5.549	*P*<0.01	1.071	(1.045-1.098)
Liver cirrhosis	1.589	0.182	4.446	*P*<0.01	3.690	(2.075-6.561)
α	0.695	0.098	4.748	*P*<0.01	2.273	(1.620-3.190)

Intercept: The average intercept of all subjects; Liver cirrhosis: Whether cirrhosis had

occurred at baseline; Time: Interval between each detection time and the first detection time(year); Cage: Baseline age after centralization; State: Outcome state; Time:Cage: The interaction of time and Cage; Time:State: The interaction of time and state; β: Fixed-effect regression coefficient.

Supplementary Table 12 Survival function in a male patient

Times	Mean	Median	Lower	Upper
7.88	1.00	1.00	1.00	1.00
8.00	0.98	0.98	0.96	0.99
9.00	0.83	0.84	0.72	0.91
10.00	0.72	0.72	0.54	0.85

Supplementary Table 13 Survival function in a female patient

Times	Mean	Median	Lower	Upper
7.15	1.00	1.00	1.000	1.00
8.00	0.98	0.98	0.96	0.99
9.00	0.96	0.97	0.91	0.99
10.00	0.95	0.95	0.87	0.98

Supplementary Table 14 Summary of Studies on Inflammatory Indices and Fibrosis Progression in chronic hepatitis B

Ref.	Target population	Main Fings	Conclusion

Piñero *et al*[1], 2020	HCC	Reviews the clinical utility of HCC biomarkers for surveillance, diagnosis, prognosis, and post-treatment assessment.	HCC biomarkers may have an increased clinical role in the near future, particularly for prognostic and predictive purposes.
Zhou *et al*[2], 2017	CHB Patients	Investigates whether serum anti-HBc could serve as a biomarker for liver inflammation in CHB patients, especially those with normal ALT levels.	Serum anti-HBc may be a strong indicator for assessing the hepatic inflammatory degree and used for antiviral treatment decisions in CHB patients with normal ALT levels.
Argirion *et al*[3], 2021	CHB Patients	Investigated the association between immunologic markers and cirrhosis in individuals with chronic hepatitis B.	The study identified several immunologic markers associated with the development of cirrhosis, which showed high accuracy in predicting cirrhosis.
Valiente *et al*[4], 2022	CHB Patients, AHB Patients	Compared the inflammatory cytokine profiles of patients with chronic hepatitis B to those with acute hepatitis B..	Patients with chronic hepatitis B had a broader and stronger inflammatory signature, indicating a more extensive immune activation in chronic hepatitis B.
Yao *et al*[5], 2019	CHB Patients	The study found that the Hepatic Inflammation Model (HIM), combining GP73, AST, and GGT, significantly improves diagnostic accuracy for liver necroinflammation, especially in patients with normal ALT levels.	The HIM, including GP73, AST, and GGT, is a promising non-invasive index for identifying patients with chronic hepatitis B who have moderate to severe liver necroinflammation, particularly when ALT is less than 40 U/l.

CHB: Chronic hepatitis B; HCC: Hepatocellular carcinoma.

REFERENCES

1 **Piñero F**, Dirchwolf M, Pessôa MG. Biomarkers in Hepatocellular Carcinoma: Diagnosis, Prognosis and Treatment Response Assessment. *Cells* 2020; **9** [PMID: 32492896 DOI: 10.3390/cells9061370]

2 **Zhou J**, Song L, Zhao H, Yan L, Ma A, Xie S, Zhang X, Zhang D, Xie Q, Zhang G, Shang J, Cheng J, Zhao W, Zou Z, Zhang M, Xia N, Wang G. Serum hepatitis B core antibody as a biomarker of hepatic inflammation in chronic hepatitis B patients with normal alanine aminotransferase. *Sci Rep* 2017; **7**: 2747 [PMID: 28584279 DOI: 10.1038/s41598-017-03102-3]

3 **Argirion I**, Pfeiffer RM, Lam TK, O'Brien TR, Yu K, McGlynn KA, Petrick JL, Pinto L, Chen CJ, Lee MH, Hildesheim A, Yang HI, Koshiol J. Association between immunologic markers and cirrhosis in individuals with chronic hepatitis B. *Sci Rep* 2021; **11**: 21194 [PMID: 34782638 DOI: 10.1038/s41598-021-00455-8]

4 **Johnson Valiente A**, Liem KS, Schwarz KB, Rosenthal P, Murray KF, Mogul D, Teckman J, Rodriguez-Baez N, Schwarzenberg SJ, Feld JJ, Wong DK, Lewis-Ximenez LL, Lauer G, Hansen BE, Ling SC, Janssen HLA, Gehring AJ. The Inflammatory Cytokine Profile Associated With Liver Damage Is Broader and Stronger in Patients With Chronic Hepatitis B Compared to Patients With Acute Hepatitis B. *J Infect Dis* 2022; **225**: 470-475 [PMID: 34286845 DOI: 10.1093/infdis/jiab373]

5 **Yao M**, Wang L, You H, Wang J, Liao H, Yang D, Liu S, Xu Q, Chen X, Jia J, Zhao J, Lu F. Serum GP73 combined AST and GGT reflects moderate to severe liver inflammation in chronic hepatitis B. *Clin Chim Acta* 2019; **493**: 92-97 [PMID: 30796901 DOI: 10.1016/j.cca.2019.02.019]