

不同纵向信息利用策略下肝细胞癌患者动态生存率预测

任嘉雯^{1*}, 姜子川²

¹青岛大学数学与统计学院, 山东 青岛

²悉尼大学理学院数学与统计系, 澳大利亚 悉尼

收稿日期: 2026年2月4日; 录用日期: 2026年2月27日; 发布日期: 2026年3月5日

摘要

动态生存预测能够随患者随访信息更新个体风险评估, 但不同纵向信息利用策略模型的预测表现仍有待系统比较。本文基于肝细胞癌患者的多次随访数据, 分别构建仅使用基线协变量的Cox比例风险模型、在随访时点更新协变量的时间依赖Cox模型以及联合建模纵向轨迹与生存结局的联合模型, 并在联合模型中比较当前值、变化率与累积效应三种关联结构。围绕动态预测任务, 设计多预测起点与多预测窗口的评价框架, 采用时间依赖C-index与Brier score综合评估模型区分能力与概率预测误差。研究表明, 纵向随访信息有助于提升动态风险刻画能力, 所提出的动态预测任务评价框架可为动态生存模型的系统比较提供参考。

关键词

生存分析, 动态生存预测, Cox比例风险模型, 联合模型, 肝细胞癌

Dynamic Survival Probability Prediction of Hepatocellular Carcinoma Patients under Different Longitudinal Information Utilization Strategies

Jiawen Ren^{1*}, Zichuan Jiang²

¹School of Mathematics and Statistics, Qingdao University, Qingdao Shandong

²Faculty of Science, The University of Sydney, Sydney, Australia

Received: February 4, 2026; accepted: February 27, 2026; published: March 5, 2026

*通讯作者。

文章引用: 任嘉雯, 姜子川. 不同纵向信息利用策略下肝细胞癌患者动态生存率预测[J]. 应用数学进展, 2026, 15(3): 110-122. DOI: 10.12677/aam.2026.153092

Abstract

Dynamic survival prediction allows individual risk estimates to be continuously updated as follow-up information accumulates, yet the predictive performance of different longitudinal information utilization strategies has not been systematically compared. Based on follow-up data from patients with hepatocellular carcinoma, this study constructed a Cox proportional hazards model using baseline covariates only, a time-dependent Cox model with covariates updated at follow-up time points, and joint models that explicitly link longitudinal trajectories with survival outcomes, and further compared three association structures within the joint modeling framework, including the current value, slope, and cumulative effect. To match the dynamic prediction task, a multi-prediction-time and multi-prediction-window evaluation framework was designed, and model performance was assessed using the time-dependent C-index and Brier score to quantify discrimination ability and probability prediction error. These findings demonstrate the value of longitudinal follow-up information for dynamic risk assessment and suggest that the proposed evaluation framework provides a practical reference for the systematic comparison of dynamic survival models.

Keywords

Survival Analysis, Dynamic Survival Prediction, Cox Proportional Hazards Model, Joint Model, Hepatocellular Carcinoma

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

肝细胞癌(hepatocellular carcinoma, HCC)是全球发病率和死亡率均较高的恶性肿瘤之一,疾病进展快,复发风险高,整体预后不理想[1] [2]。对于中期及以上分期患者,经动脉化疗栓塞(TACE)、局部消融、外科手术及其联合方案已成为常用治疗策略,同时需配合密集的影像与实验室随访以监测疾病进展并指导再治疗[3] [4]。在此背景下,患者预后风险呈现明显的时间动态变化特征,单次基线评估往往难以反映真实的疾病演化过程。因此,能够随患者随访信息更新的动态预后评估体系具有重要临床价值。

在生存率预测研究中, Cox 比例风险模型因能够处理删失并提供可解释的风险比而被广泛采用[5],但这种传统预后模型通常依赖于静态基线或平均生物标志物值,无法捕捉对个体化风险预测至关重要的时间变化[6]。为进一步利用随访过程中不断更新的临床与实验室测量信息,一种做法是将随时间变化的协变量直接纳入 Cox 模型,使风险评估能够随最新观测进行更新,例如时间依赖 Cox 模型(Time-dependent Cox) [7]。联合模型(Joint Model, JM)是一种先进的动态预后模型,通过纵向子模型刻画指标随时间变化的过程,再与生存子模型联立估计,从而在统一框架下同时建模纵向过程与生存结局[8] [9]。在预测应用层面,联合模型可输出随患者随访推进而更新的个体生存概率,近年综述表明联合模型的预测表现还可能受到关联结构设定,以及不同参数估计策略与计算实现选择的影响[10]。

针对生存分析任务,有多种评价模型的指标。一致性指数(C-index)用于衡量模型给出的风险排序与真实生存结局的一致程度,是生存预测中最常用的判别指标之一[11] [12]。AUC 也可在生存结局框架下作扩展,用于刻画模型对发生事件与未发生事件的区分能力[13]。Brier score (BS)用于量化生存概率预测与真实结局之间的整体误差,并可进一步汇总为 IBS 作为综合误差度量[14]。

本文使用肝细胞癌患者随访数据, 分别构建仅使用基线协变量的 Cox 模型、时间依赖 Cox 模型与联合模型, 并在联合模型中比较当前值、变化率与累积效应三类关联结构, 系统考察不同纵向信息利用策略对生存率预测的影响。围绕动态预测任务, 本文进一步建立与任务相匹配的时间依赖评价体系, 在多个预测起点与多个预测窗口的组合下, 使用时间依赖一致性指数(Time-dependent C-index)与时间依赖 Brier score 从区分能力与概率误差两方面对模型进行综合比较, 为纵向随访信息驱动的动态生存预测提供更实用的对比与评价范式。

2. 数据与方法

2.1. 数据来源与研究对象

Table 1. Baseline characteristics of the three datasets

表 1. 基线特征描述

变量	取值	训练集(%)	内部测试集(%)	P 值	多中心外部测试集(%)	P 值
年龄	≤55	487 (52.2)	301 (51.8)	0.924	235 (59.8)	0.013
	>55	446 (47.8)	280 (48.2)		158 (40.2)	
性别	男	848 (90.9)	445 (76.6)	<0.001	331 (84.2)	<0.001
	女	85 (9.1)	136 (23.4)		62 (15.8)	
AFP (ng/ml)	≤25	259 (27.8)	148 (25.5)	0.360	129 (32.8)	0.074
	>25	674 (72.2)	433 (74.5)		264 (67.2)	
ALB (g/L)	≤35	243 (26.0)	129 (22.2)	0.104	106 (27.0)	0.778
	>35	690 (74.0)	452 (77.8)		287 (73.0)	
TBIL (umol/L)	≤20	535 (57.3)	370 (63.7)	0.017	225 (57.3)	1.000
	>20	398 (42.7)	211 (36.3)		168 (42.7)	
AST (U/L)	≤40	238 (25.5)	169 (29.1)	0.142	116 (29.5)	0.150
	>40	695 (74.5)	412 (70.9)		277 (70.5)	
主要肿瘤直径 (mm)	≤50	315 (33.8)	200 (34.4)	0.835	128 (32.6)	0.721
	>50	618 (66.2)	381 (65.6)		265 (67.4)	
肝内病灶数	>3	565 (60.6)	357 (61.4)	0.772	242 (61.6)	0.775
	≤3	368 (39.4)	224 (38.6)		151 (38.4)	
病灶部位	单叶	539 (57.8)	330 (56.8)	0.750	240 (61.1)	0.292
	双叶	394 (42.2)	251 (43.2)		153 (38.9)	
腹水	≥1	41 (4.4)	27 (4.6)	0.918	19 (4.8)	0.836
	0	892 (95.6)	554 (95.4)		374 (95.2)	

本研究基于既往发表的一项临床研究数据开展二次分析[15]。原始数据包括三个相互独立的数据集, 分别用于模型训练、内部测试及多中心外部测试。训练集纳入 2007 年 1 月至 2012 年 5 月期间在中山大学肿瘤防治中心(SYSUCC)初次诊断并接受规范治疗的肝细胞癌(Hepatocellular carcinoma, HCC)患者; 内部测试集来自 2012 年 6 月至 2015 年 12 月同一中心后续入组患者; 多中心外部测试集则包含 2010 年 1

月至 2016 年 12 月期间来自中山大学第五附属医院、中山大学第三附属医院及广州医科大学附属第二医院的病例。

在不改变原始数据集划分方式的前提下, 本研究进一步剔除了随访信息缺失严重的个体, 以保证建模分析所需数据完整性。最终纳入训练集、内部测试集和多中心外部测试集的样本量分别为 933 例、581 例和 393 例。三组数据的中位随访时间分别为 41.8 个月(0.1~115.3)、25.9 个月(0.1~67.0)和 46.0 个月(0.2~84.2)。随访期间, 三组数据中分别观察到死亡事件 530 例(56.8%)、229 例(39.4%)和 219 例(55.7%), 其余病例在随访截止时未发生结局事件, 记为删失(Censoring)。三组数据集入组时的基线临床特征分布情况见表 1。

表 1 同时给出了不同数据集之间基线特征分布差异的统计检验结果。可以观察到, 部分人口学特征及实验室指标在三组数据之间存在统计学差异, 而肿瘤负荷相关核心变量整体分布相对稳定。这一现象反映了真实世界多中心数据在来源构成上的异质性, 同时也为后续在独立测试集与外部测试集上评估模型泛化性能提供了现实背景。

除基线信息外, 数据同时包含随访过程中动态更新的纵向临床随访变量, 涵盖实验室检测结果及疾病状态相关信息, 具有纵向生存数据结构。对分布右偏明显的连续变量(包括 AFP、AST、TBIL 及主要肿瘤直径)统一进行 $\log(1+x)$ 变换以减弱极端值影响, 并对进入模型的连续变量进一步进行标准化处理, 以提高数值稳定性并避免不同量纲对参数估计的影响。

为进一步比较不同数据集总体生存结局差异, 图 1 绘制了三组数据的 Kaplan-Meier 生存曲线。

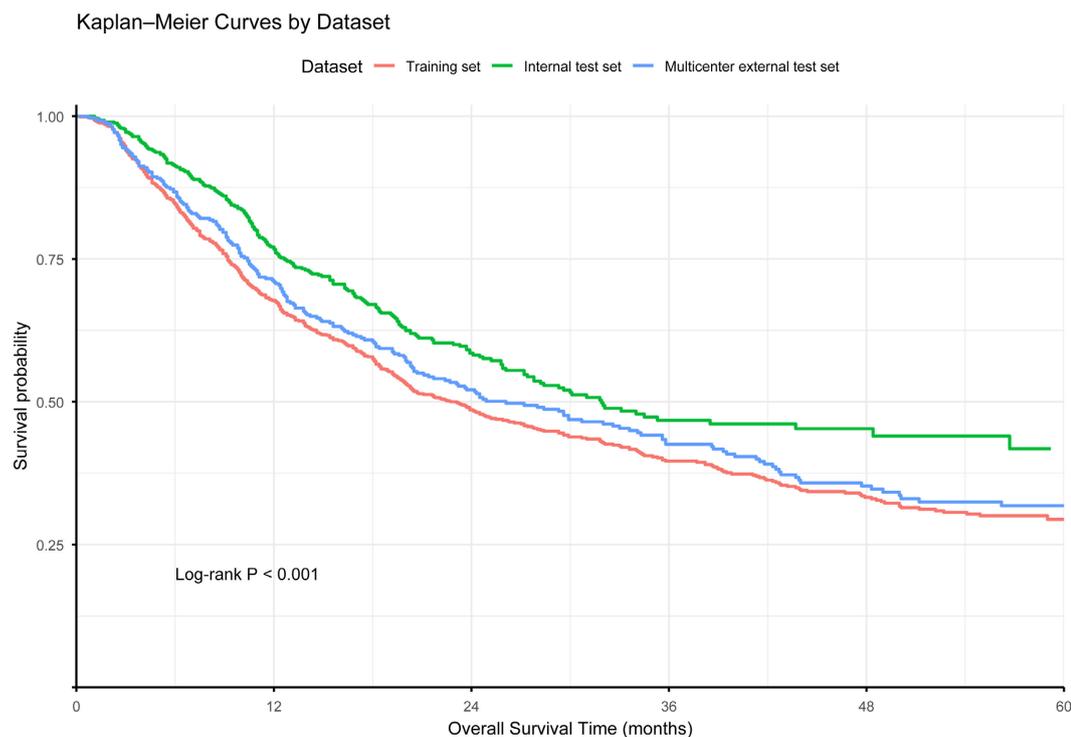


Figure 1. Kaplan-Meier survival curves
图 1. KM 曲线

三组数据的生存曲线随着随访时间延长逐渐出现一定程度分离, 表明不同中心来源患者总体生存水平存在差异。

2.2. 研究设计与模型构建

本研究围绕纵向随访信息在生存预测中的利用方式, 构建统一的统计建模与比较框架, 对不同建模策略的预测表现进行系统比较。

在模型设计上, 按照纵向信息利用程度由浅入深的思路, 设置三类对照策略。首先构建仅使用入组时基线协变量的 Cox 比例风险模型, 作为不引入随访信息时的静态预测基线。其次构建时间依赖 Cox 模型, 在生存模型中引入随访过程中更新的协变量取值, 使个体风险随时间变化。该方法实现简便, 但未对纵向测量过程本身进行显式建模。在此基础上, 引入联合模型作为核心比较方法。联合模型通过在同一统计框架下同时刻画纵向随访过程与生存结局过程, 实现对个体风险随时间更新的联立建模。

纵向部分用于描述个体随时间变化的潜在真实轨迹, 并显式区分个体间差异与测量误差; 生存部分用于刻画事件发生的风险水平。两部分通过共享结构建立关联, 使得模型能够在综合利用个体历史随访信息的基础上, 对未来风险进行更稳定的估计。

以单个纵向指标为例, 纵向子模型采用线性混合效应模型(Linear mixed-effects model, LME)描述个体潜在真实轨迹:

$$y_{ij} = m_i(t_{ij}) + \varepsilon_{ij},$$

$$m_i(t) = x_i(t)^T \beta + z_i(t)^T b_i,$$

其中 y_{ij} 表示个体 i 在时间 t_{ij} 的观测值, $m_i(t)$ 为潜在真实轨迹, $\varepsilon_{ij} \sim N(0, \sigma^2)$ 为测量误差; $x_i(t)$ 与 $z_i(t)$ 分别为固定效应与随机效应的协变量向量; β 为固定效应参数, 描述总体平均轨迹; $b_i \sim N(0, D)$ 为个体随机效应, 其中 D 是随机效应的协方差阵。

生存子模型采用 Cox 比例风险结构, 并引入纵向过程相关特征:

$$h_i(t|b_i) = h_0(t) \exp\{\gamma^T w_i + \alpha m_i(t)\},$$

其中 $h_0(t)$ 为基线风险函数, w_i 为生存部分纳入的协变量, γ 是对应的回归系数向量, α 为纵向过程与事件风险之间的关联参数。当存在多个纵向指标时, 上述纵向子模型可对每个指标分别建立对应的混合效应模型, 并在生存子模型中同时引入各纵向过程的关联项, 从而形成多纵向联合模型结构。

考虑到纵向过程可能通过不同机制影响生存风险, 本研究进一步将纵向与生存关联结构作为比较维度, 分别构建基于当前值、变化率及累积效应的多种关联形式。

(1) 当前值关联

当前值关联假设个体在时间 t 的事件风险主要由该时刻纵向指标的水平决定, 即风险取决于潜在轨迹的即时值 $m_i(t)$ 。对应的生存子模型可写为:

$$h_i(t) = h_0(t) \exp\{\gamma^T w_i + \alpha m_i(t)\}。$$

(2) 变化率关联

斜率关联在当前值基础上进一步引入纵向轨迹的一阶导数, 用于刻画指标变化速度对风险的附加影响:

$$h_i(t) = h_0(t) \exp\{\gamma^T w_i + \alpha_1 m_i(t) + \alpha_2 m'_i(t)\}。$$

(3) 累积效应关联

该结构假设风险受纵向过程历史累积暴露的影响, 更强调长期水平对结局风险的贡献:

$$h_i(t) = h_0(t) \exp\{\gamma^T w_i + \alpha \int_0^t m_i(s) ds\}。$$

本研究采用时间依赖一致性指数 $C(t_0, \Delta)$ 与时间依赖 Brier score $BS(t_0, \Delta)$ 评价模型动态预测性能, 分别反映风险排序能力与概率预测误差。模型在预测时间点 t_0 利用截至该时点的随访信息, 对未来预测窗口 Δ 内的结局风险进行预测, 并据此计算对应指标。为减小删失对评价结果的影响, 所有指标均采用基于逆概率删失加权(IPCW)的修正方法。预测时间点设为 $t_0 \in \{1, 6, 12\}$ 个月, 预测窗口长度设为 $\Delta \in \{3, 6, 9, 12, 18, 24, 36, 48\}$ 个月, 以系统评估模型在不同随访阶段与预测窗口下的表现。

3. 结果

3.1. 模型拟合结果

本研究原始数据包含多项随访过程中重复测量的临床与实验室指标, 理论上均可作为纵向过程纳入联合模型。然而, 在实际建模中, 联合模型对纵向子模型数量与复杂度较为敏感, 当同时引入过多纵向结局时, 模型计算开销与收敛不稳定风险显著增加。综合考虑模型稳定性、计算可行性及既往肝细胞癌相关研究中对关键实验室指标的选择经验, 本研究在建模前预先设定纵向指标集合, 未采用数据驱动的变量筛选方法, 在联合模型框架中仅对甲胎蛋白(AFP)、白蛋白(ALB)、总胆红素(TBIL)和天冬氨酸转氨酶(AST)四项指标建立纵向子模型, 用于刻画个体随访过程中生理状态与疾病负担的动态变化: AFP 反映肿瘤负荷及其生物学行为, ALB 与 TBIL 共同表征肝功能状态并构成常用的肝功能分层体系, 而 AST 作为肝细胞损伤指标, 在肝细胞癌风险评估与随访管理中亦被广泛应用[16][17]。其余临床变量以基线协变量形式进入生存子模型, 用于风险调整。本文中主要变量缩写为: 年龄(AGE)、主要肿瘤直径(TD)、肝内病灶数(LN)、病灶部位(LS)及腹水(ASC)。

以 BIC 作为主要模型选择准则。最终确定的纵向子模型形式如下:

(1) AFP

$$AFP_{ij} = \beta_0^{(1)} + \beta_1^{(1)}TD + \beta_2^{(1)}AGE + \beta_3^{(1)}LS + \beta_4^{(1)}t_{ij} + b_{0i}^{(1)} + b_{1i}^{(1)}t_{ij} + \varepsilon_{ij}^{(1)}$$

(2) ALB

$$ALB_{ij} = \beta_0^{(2)} + \beta_1^{(2)}AGE + \beta_2^{(2)}ASC + \beta_3^{(2)}LN + \beta_4^{(2)}TD + \beta_5^{(2)}t_{ij} + b_{0i}^{(2)} + b_{1i}^{(2)}t_{ij} + \varepsilon_{ij}^{(2)}$$

(3) TBIL

$$TBIL_{ij} = \beta_0^{(3)} + \beta_1^{(3)}t_{ij} + b_{0i}^{(3)} + b_{1i}^{(3)}t_{ij} + \varepsilon_{ij}^{(3)}$$

(4) AST

$$AST_{ij} = \beta_0^{(4)} + \beta_1^{(4)}t_{ij} + b_{0i}^{(4)} + b_{1i}^{(4)}t_{ij} + \varepsilon_{ij}^{(4)}$$

Table 2. Longitudinal submodel results

表 2. 纵向子模型

纵向子模型	变量	估计(95% CrI)	p
AFP	TD	0.288 (0.233, 0.344)	<0.001
	AGE	-0.194 (-0.250, -0.139)	<0.001
	LS	0.196 (0.087, 0.307)	<0.001
	Time	-0.009 (-0.014, -0.002)	0.004
ALB	AGE	-0.181 (-0.224, -0.138)	<0.001
	ASC	-0.382 (-0.522, -0.240)	<0.001

续表

	LN	-0.066 (-0.112, -0.018)	0.008
	TD	-0.083 (-0.126, -0.039)	<0.001
	Time	0.011 (0.004, 0.017)	0.002
TBIL	Time	-0.006 (-0.012, -0.000)	0.040
AST	Time	-0.031 (-0.036, -0.025)	<0.001

本研究的联合模型均在 R (version 4.5.2) 环境下, 采用 JMbayes2 包进行拟合。同时, 为与临床研究中常见的显著性呈现方式保持一致, 表中列示的 p 为基于后验样本的双侧尾概率, 由 JMbayes2 给出, 用于辅助展示效应证据强弱。

四个纵向指标(AFP、ALB、TBIL、AST)的纵向子模型参数估计结果汇总见表 2。总体来看, 各指标的潜在轨迹均与随访时间相关, 并与部分基线临床特征存在关联, 表明纵向子模型能够有效刻画纵向指标随疾病进程演化的整体趋势。AFP 与 ALB 的纵向轨迹同时受到时间效应与肿瘤负荷等临床状态相关协变量的共同影响, 而 TBIL 与 AST 的变化主要由时间趋势项主导, 说明不同指标的纵向动态模式存在差异。

鉴于不同关联结构下纵向子模型估计结果高度一致, 本文仅汇总报告一次纵向子模型结果, 重点比较各关联结构下的生存子模型与预测性能。在确定纵向子模型形式后, 分别构建当前值关联 JM (current)、变化率关联 JM (slope) 与累积效应关联 JM (area) 三类联合模型, 并对生存子模型参数进行估计。进入联合模型的部分连续变量已在建模前进行 $\log(1+x)$ 变换并统一标准化处理, 因此 HR 的解释均应结合对应变量的实际尺度。

Table 3. Joint model with current-value association

表 3. 联合模型当前值关联

变量	估计	HR (95% CrI)	p
AGE	-0.091	0.913 (0.783, 1.070)	0.243
TD	0.502	1.652 (1.462, 1.876)	<0.001
LN	0.162	1.176 (1.048, 1.318)	0.005
LS	0.090	1.094 (0.884, 1.342)	0.405
ASC	0.198	1.219 (0.912, 1.579)	0.163
AFP	0.526	1.692 (1.490, 1.929)	<0.001
ALB	-0.834	0.434 (0.331, 0.56)	<0.001
TBIL	0.223	1.250 (1.007, 1.545)	0.043
AST	-0.340	0.712 (0.497, 0.994)	0.053

在当前值关联结构(表 3)下, 主要肿瘤直径与肝内病灶数与死亡风险升高显著相关。纵向指标方面, AFP 当前水平与死亡风险呈正相关, 而 ALB 当前水平与死亡风险呈负相关, 二者在风险评估中具有相对稳定的指示作用。

在变化率关联结构(表 4)中, 同时纳入四项纵向指标的变化率会明显增加模型复杂度并影响收敛稳定性。综合模型稳定性与解释性考虑, 本文仅对 AFP 与 ALB 引入变化率项, 而对 AST 与 TBIL 仅保留当

前值项。结果显示, 主要肿瘤直径与肝内病灶数仍与更高的死亡风险相关, 表明肿瘤负荷相关因素在引入动态项后依旧保持稳定的风险指示作用。ALB 上升趋势与死亡风险降低相关, 而 AFP 的变化率未表现出稳定的附加预测信息, 这代表着不同纵向指标在变化率层面的预后贡献存在差异。

表 5 显示, 在累积效应关联结构下, 主要肿瘤直径与肝内病灶数仍与死亡风险显著相关。纵向累积项中, AFP 的累积暴露与死亡风险升高相关, ALB 的累积暴露与死亡风险降低相关, 说明二者的风险信息不仅体现在单一时点水平或短期变化趋势, 也能够通过长期累积暴露稳定地反映个体风险差异。TBIL 的累积项在该结构下呈现出明确的风险关联, 而 AST 的累积项未观察到清晰证据, 说明 TBIL 更可能通过长期累积暴露体现预后影响, 而 AST 的作用不表现为累积机制。

Table 4. Joint model with slope association

表 4. 联合模型变化率关联

变量	估计	HR (95% CrI)	p
AGE	-0.109	0.897 (0.757, 0.944)	0.203
TD	0.640	1.900 (1.614, 2.255)	<0.001
LN	0.194	1.214 (1.064, 1.391)	0.002
LS	0.101	1.106 (0.865, 1.413)	0.405
ASC	0.291	1.338 (0.966, 1.837)	0.081
AFP	0.408	1.504 (1.257, 1.784)	<0.001
ALB	-0.852	0.427 (0.309, 0.571)	<0.001
TBIL	0.206	1.229 (0.908, 1.647)	0.159
AST	-0.567	0.567 (0.366, 0.862)	0.008
AFP (slope)	4.543	0.956 (0.943, 1.154)	0.332
ALB (slope)	-10.484	0.900 (0.795, 0.993)	0.038

Table 5. Joint model with cumulative-effect association

表 5. 联合模型累积效应关联

变量	估计	HR (95% CrI)	p
AGE	-0.075	0.928 (0.800, 1.079)	0.332
TD	0.429	1.536 (1.373, 1.730)	<0.001
LN	0.139	1.149 (1.033, 1.281)	0.013
LS	0.095	1.100 (0.900, 1.343)	0.346
ASC	0.181	1.198 (0.912, 1.545)	0.185
AFP (area)	0.585	1.794 (1.582, 2.040)	<0.001
ALB (area)	-0.814	0.443 (0.343, 0.571)	<0.001
TBIL (area)	0.207	1.230 (1.013, 1.489)	0.038
AST (area)	-0.046	0.955 (0.702, 1.281)	0.768

3.2. 预测性能

图 2 展示了三个数据集在不同预测时间点 t_0 与预测窗口 Δ 下的时间依赖一致性指数 $C(t_0, \Delta)$ 。总体来看, 各模型的区分度随预测窗口延长呈现逐步下降趋势, 表明长期预测任务的难度更高。

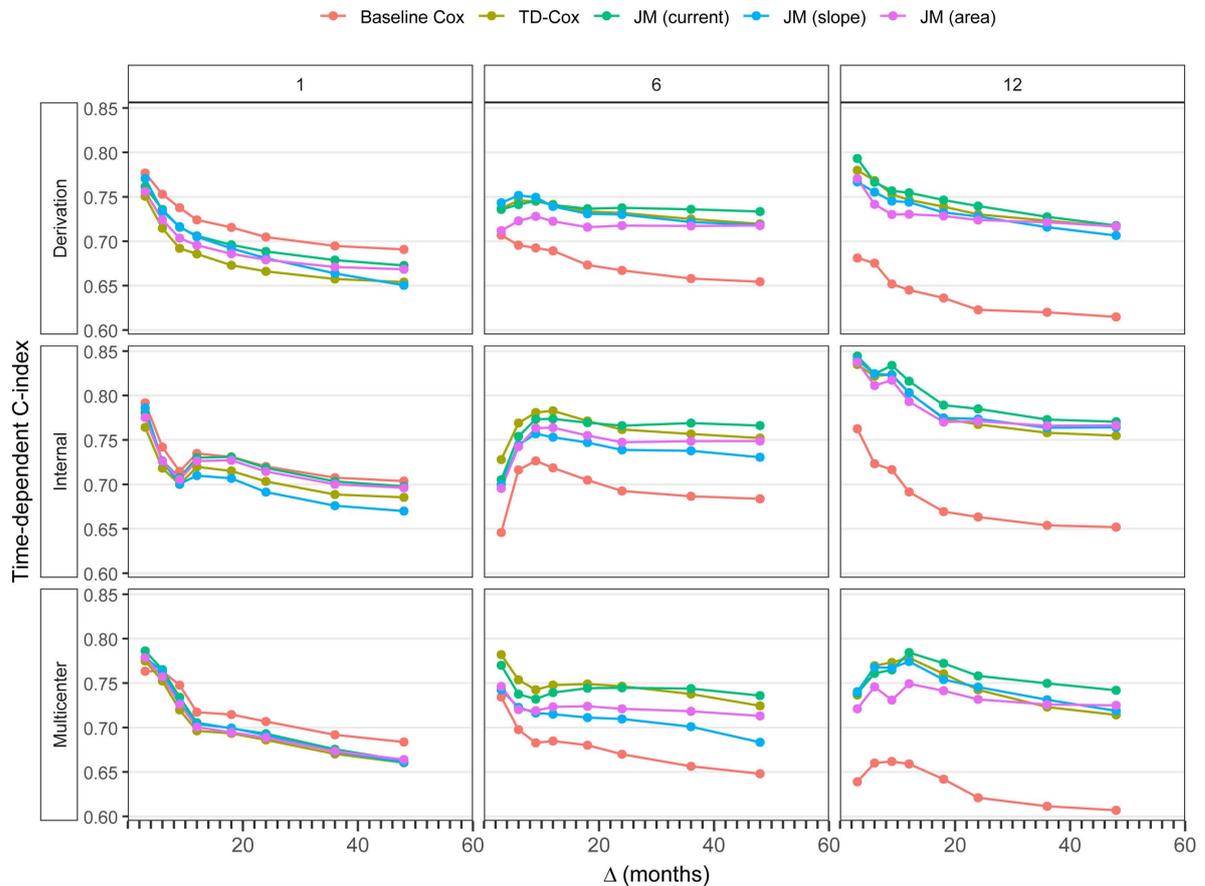


Figure 2. Time-dependent C-index
图 2. 时间依赖 C-index

在随访早期 $t_0 = 1$ 时, 各模型之间差异相对有限, 基线 Cox (Baseline Cox) 在多数预测窗口下取得最高或接近最高的区分性能, 表明在纵向随访信息尚未充分积累时, 基线协变量已能够较好刻画个体风险差异。当 $t_0 = 6$ 月时, 引入纵向随访信息的模型整体优于基线 Cox, 但不同方法之间已出现分化。JM (current) 在多数预测窗口内保持较高且稳定的区分度, 而时间依赖 Cox (TD-Cox) 仅在部分较短预测窗口下表现最优, 随 Δ 增大其区分性能下降更明显。当 $t_0 = 12$ 月时, 上述差异进一步扩大。JM (current) 在三个数据集的大多数预测窗口中持续保持领先表现, TD-Cox 的区分性能整体弱于 JM (current), 仅在个别短窗口场景下具有有限优势。该结果表明, 随访信息充分积累后, 基于纵向轨迹联立建模的联合模型在中长期风险排序任务中具有更稳定的表现。

相比之下, JM (slope) 与 JM (area) 在多数场景下未表现出额外优势, 其性能提升幅度有限, 表明在本研究数据条件下, 更复杂的关联结构对区分度的改善较小。这一现象可能与纵向随访数据的时间结构特征有关。变化率关联结构依赖于相邻随访点之间的斜率估计, 而在真实世界肝细胞癌随访中, 患者随访次数有限且时间间隔往往不规则, 使得基于局部观测计算的瞬时斜率易受到测量误差与时间噪声的影响,

从而掩盖了真实的风险信号。类似地, 累积效应虽在一定程度上整合了历史纵向信息, 但在随访点数有限的情况下往往需由离散观测近似积分, 其估计同样会受到时间跨度差异与噪声累积的影响。

图 3 展示了不同预测时间点下随预测窗口变化的时间依赖 Brier score。总体来看, 各模型的 BS 随预测窗口延长呈持续上升趋势, 表明预测时间跨度增加会带来更大的总体预测误差。

在模型比较方面, 基线 Cox 普遍表现出更高的 BS, 且在较长预测窗口时与其余模型的差距进一步扩大, 说明仅依赖基线协变量难以在随访过程中持续更新个体结局概率, 导致中长期窗口下总体误差累积更明显。相比之下, 引入纵向随访信息的模型整体保持更低的 BS, 说明随访信息能够有效降低窗口预测误差。

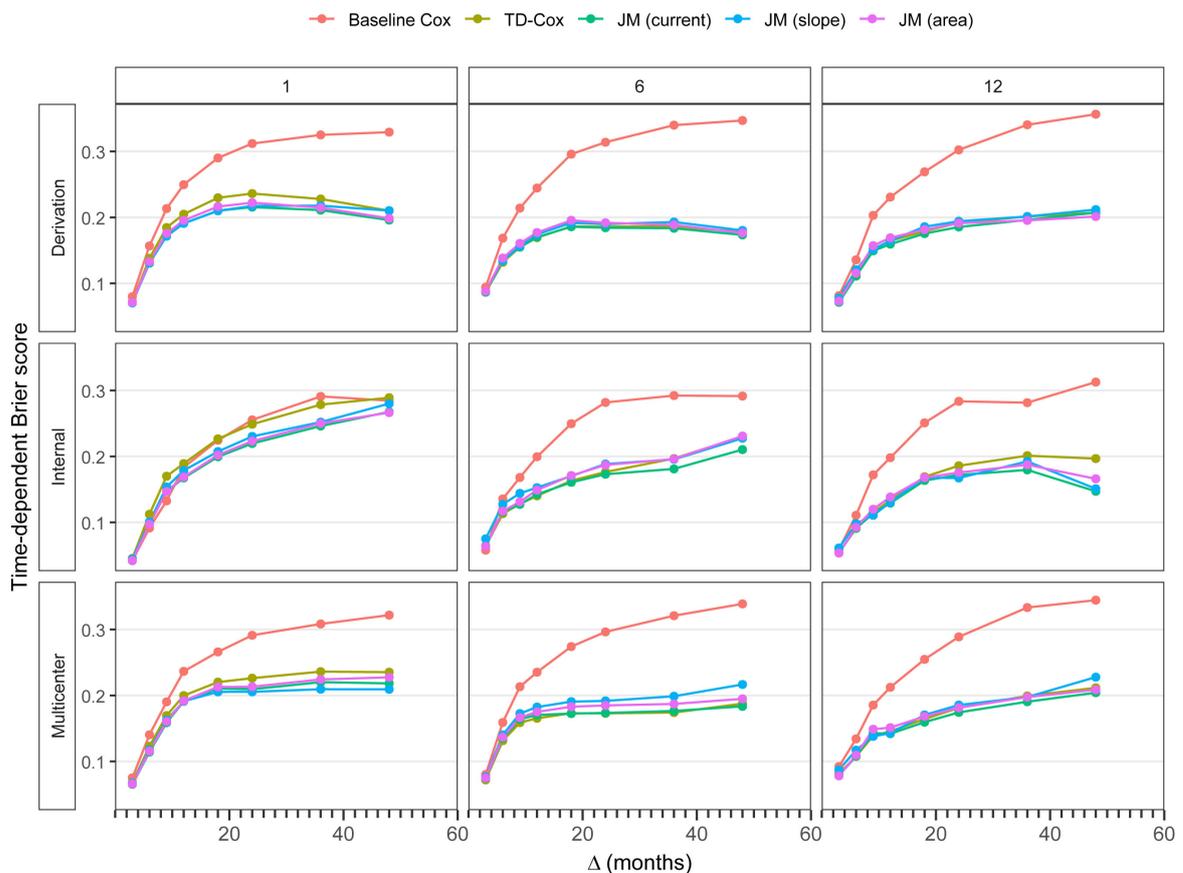


Figure 3. Time-dependent Brier score

图 3. 时间依赖 BS

进一步比较 TD-Cox 与 JM 系列模型可以发现, 多数时间组合下 JM (current) 的 BS 最低或接近最低, 在中长期预测窗口中表现出更稳定的误差水平, 而 JM (slope) 与 JM (area) 几乎未呈现额外优势。该结果与区分度分析相一致, 表明在本研究数据条件下, 基于纵向轨迹当前水平构建的联合模型能够在保持较好风险排序能力的同时, 更稳定地控制总体预测误差, 而更复杂的关联结构对误差改善的收益有限。

3.3. 校准曲线

校准曲线用于评估模型给出的窗口事件预测概率与实际观测发生率之间的一致性。本研究在给定预测时间点 $t_0 = 6$ 月时, 模型输出未来预测窗口 $\Delta = 6, 12, 24$ 月事件发生概率, 并按预测概率从低到高将个体

划分为 5 个风险组。以各组的平均预测概率作为横坐标, 采用 Kaplan-Meier 方法估计对应窗口内的观测事件发生率作为纵坐标, 绘制分组散点并以 $y = x$ 作为理想校准参考线。

如图 4 所示, 在三个数据集中, 基线 Cox 的校准点整体偏离对角线, 且在较大预测窗口下偏离程度更为明显。相比之下, TD-Cox 及 JM 系列模型的校准点整体更接近参考线, 表明引入随访信息后, 模型对窗口事件概率的预测与实际发生率之间的一致性得到改善。

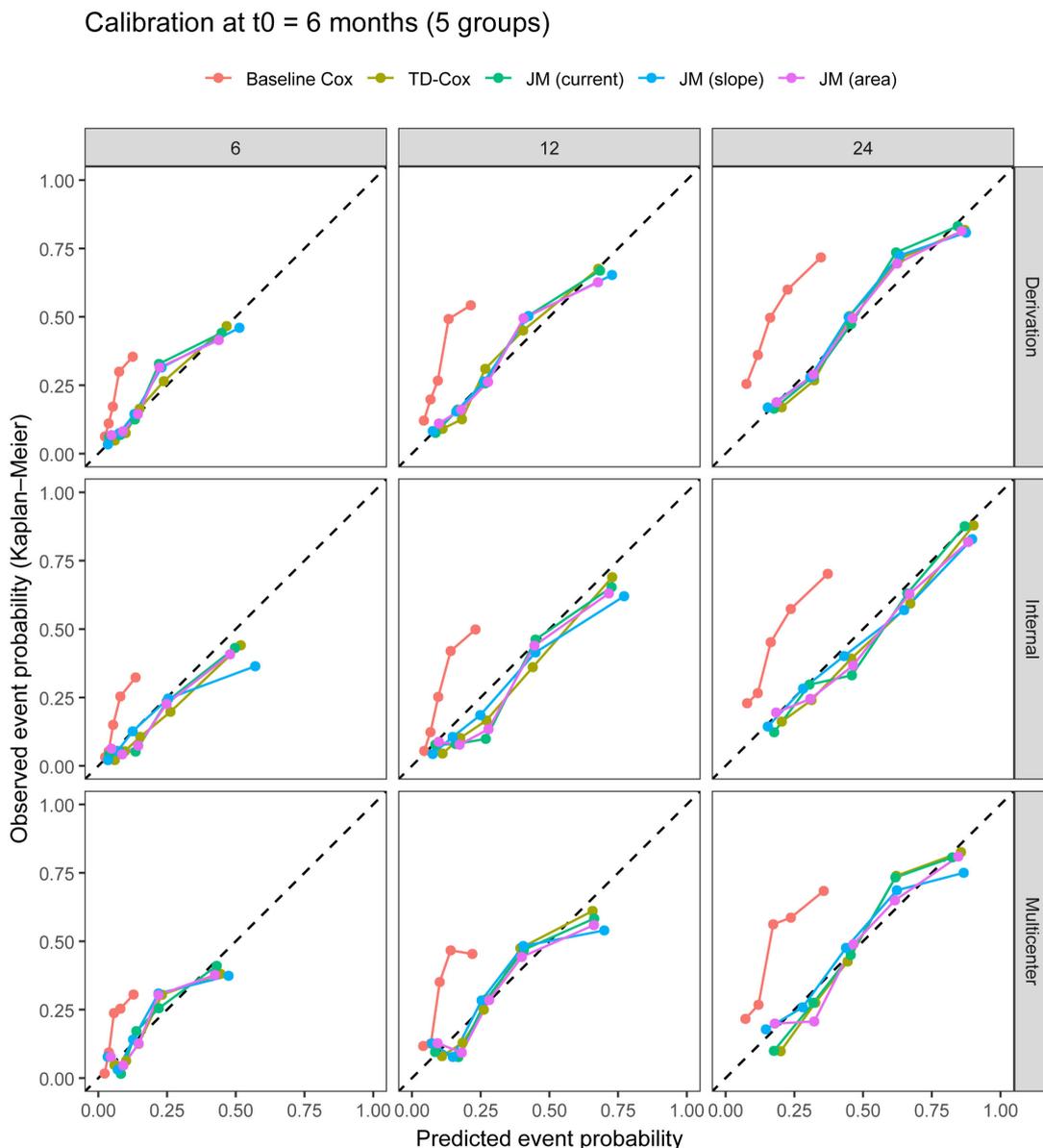


Figure 4. Calibration curves
图 4. 校准曲线

4. 结论

本研究基于纵向随访数据, 系统比较了不同统计建模策略在动态生存预测任务中的表现。结果表明, 相较仅使用基线协变量的传统 Cox 模型, 引入纵向随访信息能够更有效刻画个体风险随时间的演化过程,

提高预测结果的准确性。在多种利用纵向随访信息的建模方法中, 时间依赖 Cox 模型在短期预测窗口内对近期风险变化较为敏感, 联合模型在对纵向轨迹进行平滑建模的基础上, 对于中长期预测表现出更稳定的预测性能。进一步比较联合模型中不同关联结构的结果显示, 关联结构复杂度的增加未带来稳定的预测性能改善。

在评价方法方面, 本研究构建了与动态预测任务相匹配的时间依赖评价框架, 通过多预测时间点 t_0 与多预测窗口 Δ 的组合设计, 结合时间依赖 C-index 与 Brier score, 从风险排序能力与概率预测误差两个维度对模型进行系统评估。该评价策略能够更加全面地反映模型在不同随访阶段与预测窗口下的预测性能, 为不同动态生存模型之间的公平比较提供统一基准。

本研究的不足之处在于, 在联合模型构建中仅纳入了有限数量的纵向指标用于轨迹建模, 尚未评估更大规模纵向变量集对模型性能与稳定性的影响。此外, 本文在建模与分析过程中未进一步探讨不同指标之间的相关性, 可能在一定程度上影响风险估计与预测结果的解释。未来工作可在保持模型可解释性的前提下, 进一步探索更多纵向过程联合建模策略及其对动态预测表现的影响。同时, 可将本文采用的动态预测评价框架推广至更复杂的动态生存模型, 用于评估不同模型在多时间点、多预测窗口设定下的性能。

参考文献

- [1] Siegel, R.L., Giaquinto, A.N. and Jemal, A. (2024) Cancer Statistics, 2024. *CA: A Cancer Journal for Clinicians*, **74**, 12-49. <https://doi.org/10.3322/caac.21820>
- [2] Konyon, P., Ahmed, A. and Kim, D. (2021) Current Epidemiology in Hepatocellular Carcinoma. *Expert Review of Gastroenterology & Hepatology*, **15**, 1295-1307. <https://doi.org/10.1080/17474124.2021.1991792>
- [3] Brown, Z.J., Tsilimigras, D.I., Ruff, S.M., Mohseni, A., Kamel, I.R., Cloyd, J.M., et al. (2023) Management of Hepatocellular Carcinoma: A Review. *JAMA Surgery*, **158**, Article 410. <https://doi.org/10.1001/jamasurg.2022.7989>
- [4] Elshaarawy, O., Gomaa, A., Omar, H., Rewisha, E. and Waked, I. (2019) Intermediate Stage Hepatocellular Carcinoma: A Summary Review. *Journal of Hepatocellular Carcinoma*, **6**, 105-117. <https://doi.org/10.2147/jhc.s168682>
- [5] Cox, D.R. (1972) Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **34**, 187-202. <https://doi.org/10.1111/j.2517-6161.1972.tb00899.x>
- [6] Zhou, Q., Chen, Z. and Peng, S. (2025) Evaluating the Application of Dynamic Prediction Models in Oncological Prognostic Studies with Repeated Measurement Predictors. *NPJ Precision Oncology*, **9**, Article No. 376. <https://doi.org/10.1038/s41698-025-01162-7>
- [7] Andersen, P.K. and Gill, R.D. (1982) Cox's Regression Model for Counting Processes: A Large Sample Study. *The Annals of Statistics*, **10**, 1100-1120. <https://doi.org/10.1214/aos/1176345976>
- [8] Henderson, R. (2000) Joint Modelling of Longitudinal Measurements and Event Time Data. *Biostatistics*, **1**, 465-480. <https://doi.org/10.1093/biostatistics/1.4.465>
- [9] Rizopoulos, D. (2011) Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics*, **67**, 819-829. <https://doi.org/10.1111/j.1541-0420.2010.01546.x>
- [10] Wang, J. and Zhong, Q. (2025) Joint Modeling of Longitudinal and Survival Data. *Annual Review of Statistics and Its Application*, **12**, 449-476. <https://doi.org/10.1146/annurev-statistics-112723-034334>
- [11] Harrell, F.E., Lee, K.L., Califf, R.M., Pryor, D.B. and Rosati, R.A. (1984) Regression Modelling Strategies for Improved Prognostic Prediction. *Statistics in Medicine*, **3**, 143-152. <https://doi.org/10.1002/sim.4780030207>
- [12] Gerds, T.A., Kattan, M.W., Schumacher, M. and Yu, C. (2013) Estimating a Time-Dependent Concordance Index for Survival Prediction Models with Covariate Dependent Censoring. *Statistics in Medicine*, **32**, 2173-2184. <https://doi.org/10.1002/sim.5681>
- [13] Heagerty, P.J. and Zheng, Y. (2005) Survival Model Predictive Accuracy and ROC Curves. *Biometrics*, **61**, 92-105. <https://doi.org/10.1111/j.0006-341x.2005.030814.x>
- [14] Gerds, T.A. and Schumacher, M. (2006) Consistent Estimation of the Expected Brier Score in General Survival Models with Right-censored Event Times. *Biometrical Journal*, **48**, 1029-1040. <https://doi.org/10.1002/bimj.200610301>
- [15] Shen, L., Zeng, Q., Guo, P., Huang, J., Li, C., Pan, T., et al. (2018) Dynamically Prognosticating Patients with Hepatocellular Carcinoma through Survival Paths Mapping Based on Time-Series Data. *Nature Communications*, **9**, Article No.

2230. <https://doi.org/10.1038/s41467-018-04633-7>
- [16] Zheng, Y., Hu, J., Mei, Y., Xiang, Y., Lu, H., Zhu, J., *et al.* (2025) Trajectories of α -Fetoprotein and Unresectable Hepatocellular Carcinoma Outcomes Receiving Lenvatinib: A Retrospective, Multicenter Cohort Study. *BMC Cancer*, **25**, Article No. 1137. <https://doi.org/10.1186/s12885-025-14516-y>
- [17] Lee, S.K., Song, M.J., Kim, S.H. and Park, M. (2019) Comparing Various Scoring System for Predicting Overall Survival According to Treatment Modalities in Hepatocellular Carcinoma Focused on Platelet-Albumin-Bilirubin (PALBI) and Albumin-Bilirubin (ALBI) Grade: A Nationwide Cohort Study. *PLOS ONE*, **14**, e0216173. <https://doi.org/10.1371/journal.pone.0216173>