

doi: 10.12052/gdutxb.230007

基于多任务循环神经网络带状回归模型的 乳腺癌个体生存分析

陈睿¹, 蔡念¹, 罗智浩¹, 刘璇^{2,3}, 黎剑^{2,3}

(1. 广东工业大学 信息工程学院, 广东 广州 510006; 2. 中山大学肿瘤防治中心 华南肿瘤学国家重点实验室, 广东 广州 510060; 3. 中山大学肿瘤防治中心 广东省恶性肿瘤临床医学研究中心, 广东 广州 510060)

摘要: 针对乳腺癌病程长、疾病发展较缓和的特点, 提出了一种多任务循环神经网络带状回归模型进行乳腺癌个体生存分析。首先, 提出一种基于循环神经网络的多任务带状回归模型, 通过识别各病理特征对不同患者之间影响的区别, 优化患者个体生存分析。其次, 对带状校验矩阵的形式进行拓展并研究其对患者风险分布的影响。最后, 在乳腺癌真实数据集上进行生存分析, 不同患者之间产生明显的差异性, 验证了模型的有效性。在2个乳腺癌真实数据集上进行的生存分析结果显示, 基于循环神经网络的多任务带状回归模型的一致性指数(Concordance Index, C-index)较医学上常用的Cox回归模型有较大提升, 并有着更小的95%置信区间。

关键词: 乳腺癌; 个体生存分析; 循环神经网络; 多任务带状回归

中图分类号: TP391

文献标志码: A

文章编号: 1007-7162(2024)01-0034-07

Individual Survival Analysis of Breast Cancer Based on Multi-task Recurrent Neural Network Banded Regression Model

Chen Rui¹, Cai Nian¹, Luo Zhi-hao¹, Liu Xuan^{2,3}, Li Jian^{2,3}

(1. School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China; 2. State Key Laboratory of Oncology in South China, Sun Yat-sen University Cancer Center, Guangzhou 510060, China; 3. Guangdong Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China)

Abstract: In view of the characteristics of long course and mild disease development of breast cancer, a multi-task recurrent neural network banded regression model is proposed to analyze individual survival of breast cancer. First, a multi-task banded regression model based on recurrent neural network is proposed to optimize the individual survival analysis of patients by identifying the difference in the effects of various pathological features on different patients. Then, the form of the banded check matrix is expanded and its effect on the hazard distribution of patient is investigated. Finally, the survival analysis on the real datasets of breast cancer shows obvious differences among different patients, which verifies the validity of the model. The survival analysis on two real datasets of breast cancer shows that the C-index of the multi-task banded regression model based on recurrent neural network is greatly improved compared with the Cox regression model commonly used in medicine, and has a smaller 95% confidence interval.

Key words: breast cancer; individual survival analysis; recurrent neural network; multi-task banded regression

乳腺癌已成为全球最常见的女性癌病, 是最致命的女性癌症之一^[1], 2020年约有68.5万女性死于患乳腺癌^[2]。研究发现部分乳腺癌患者在确诊后十年期间仍存活于复发可能性的阴影之下^[3], 因此针对乳腺

癌患者的预后管理显得尤为重要, 但其适用性和有效性又受限于临床医生颇具主观性的专业经验。近年来, 生存分析^[4]技术被用以辅助像乳腺癌这种发病不明显、病程较长的疾病制定预后措施, 研究人员将

收稿日期: 2023-01-13

基金项目: 国家自然科学基金资助面上项目(82172019); 广州市科技计划项目(202102010251)

作者简介: 陈睿(1998-), 男, 硕士研究生, 主要研究方向为机器学习、生存分析等

通信作者: 蔡念(1976-), 男, 教授, 博士, 主要研究方向为机器视觉、机器学习、数字信号处理等, E-mail: cainian@sina.com

预后和预测特征即患者病理协变量独立或整合以指导患者的治疗^[5]。鉴于个体化医学的重要性,研究人员结合当下个体化医学研究中流行的用于发现并识别复杂数据之间模式和关系的机器学习技术,对乳腺癌等长病程数据集进行个体生存分析。此类方法基于患者间的病理协变量表达差异实现个性化的生存风险估计,进而绘制患者个体生存曲线,以辅助实现精准医疗。

近年来,个体生存分析研究主要分为基于Cox比例风险假设的生存分析和无参数型离散生存分析。Cox比例风险模型假设独立协变量对生存时间的影响不随时间而变化,以评估多个协变量对生存的影响^[6]。Katzman等^[7]基于Cox对数比例风险假设搭建一个深度前馈神经网络DeepSurv,用于预测患者协变量对风险率的影响。Lee等^[8]使用一个共享子网络以处理并预测时变协变量,同时设计多个特定原因子网络捕获特定原因风险同协变量之间的关系,构成Dynamic-DeepHit模型用于动态风险预测。Tang等^[9]对删失数据进行最大似然估计并结合常微分方程进行生存分析建模。上述单任务回归模型遵循对数风险假设,导致乳腺癌患者间的风险函数永不相交,虽然可描绘出患者病情在总生存期中一个大致的走势,但难以实现阶段性且个性化的生存风险描述。Yu等^[10]提出多任务线性回归模型(Multi-task Logistic Regression, MTLR),其采用响应转换的方式绕过比例风险假设以实现生存函数预测,但回归器核心为线性变换,难以捕捉患者生存数据中的非线性关系。Fotso等^[11]通过往MTLR中引入多层感知器,实现灵活的建模,并命名为N-MTLR(Neural Multi-task Logistic Regression)。Hu等^[12]基于时下流行的Transformer网络架构,使用有序回归实现对时变生存概率的优化,但受限于多头注意力机制在病理协变量对风险概率贡献较均衡时难以分配权重。Wang等^[13]提出了多任务带状回归模型(Multi-task Banded Regression, MTBR),带状校验矩阵的引入为响应转换提供约束,解决风险概率下降缓慢导致的生存概率分配不均的情况。以上的多任务回归模型采用线性变换或多层感知器进行响应转换,使得协变量独立于多个序列之间,难以提取出病理协变量对不同患者生存时间之间影响的异同关系,导致模型一致性指数下降。

为此,本文提出了一种新的多任务回归模型RNN-MTBR(Multi-task Banded Regression via Recurrent Neural Network)用于乳腺癌个体生存分析。首先,针

对以往的多任务模型的回归器无法使序列之间产生前后联系的情况,使用循环神经网络(Recurrent Neural Network, RNN)^[14]作为多任务回归模型的核心,RNN可通过维护网络过去的记忆,让模型在提取病理协变量同生存时间的关系时能利用到序列的上下文信息。其次,将带状校验矩阵改进并探究其对风险分布的作用效果。最后,在乳腺癌真实数据集上实现了对个体生存分析区分度的优化,成功验证了模型的有效性。

1 乳腺癌生存分析

给定一个含有 M 名患者的乳腺癌生存数据集,设 $\forall i \in [1, M]$, (X_i, δ_i, t_i) 为一个数据点,其中包含生存时间、与生存时间有关联的 p 维独立变量 $X_i = [x_1^i, x_2^i, \dots, x_p^i]$ 和事件类型标签 δ_i (0或1)。生存数据分为完全数据、删失数据。完全数据指从观察起点到发生死亡事件所经历的时间均被观察到的生存数据。删失数据则分为左删失数据、区间删失数据和右删失数据^[15]。左删失表示事件发生在观察时间开始之前。区间删失表示只能得知患者在不连续的2次随访区间中发生终点事件与否,而无法得知终点事件发生的具体时间。右删失表示在随访观察中,已知患者被观察的起始时间,而无法得知患者发生终点事件的时间,其生存时间大于观察时间。本文实验中“删失”数据只包含“右删失”数据。

对于在观察期间内能够完整观察到事件发生的患者个体实例,其生存时间记为实际事件时间 t_0 。对于在观察期间内失去对患者的观察,或者患者事件发生时间超出观察时间的实例,其生存时间记为删失时间 t_c , t_0 为未删失时间。患者个体生存时间 $t_i = \min(t_0, t_c)$ 。生存数据如图1所示。患者1为完全数据,因乳腺癌死亡于观察区间之外的患者2为右删失数据,发生了其他事件死亡的患者3也为右删失数据。

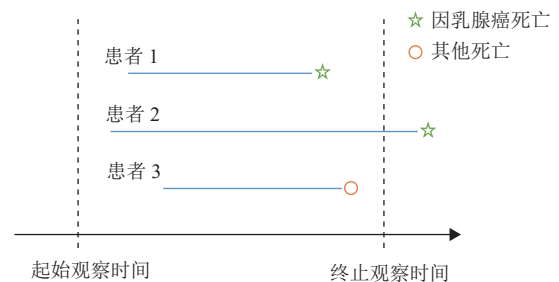


图1 患者生存数据形式

Fig.1 Forms of patient survival data

乳腺癌个体生存分析利用患者病理独立协变量 X_i 与生存时间 t_i 建立一个乳腺癌风险随时间变化的函数, 常见为风险函数 $h(t)$ 、风险概率分布函数 $f(t)$ 或概率累计函数 $F(t)$ 。

2 多任务循环神经网络带状回归

2.1 基本原理及流程

多任务方法通过在不同时间间隔上建立回归模型, 来计算乳腺癌患者死亡在每个时间间隔中发生的概率, 使不同患者间的结果具有更为明显的差异性。多任务回归模型以响应变量的形式来分别代表患者在每一时间间隔的状态(生存或死亡)。假设总生存期分为 N 个生存子区间, 则在第 j 个生存子区间

τ_j (左开右闭)内患者的响应变量为

$$y_i = \begin{cases} 1, \tau_j \text{ 为患者死亡事件发生所在生存子区间} \\ 0, \text{其余生存子区间} \end{cases} \quad (1)$$

且患者发生死亡事件后的所有时间间隔响应变量值将恒为1。因而患者的生存向量为

$$Y = [y_1 = 0, \dots, y_{j-1} = 0, y_j = 1, \dots, y_N = 1]^T \quad (2)$$

多任务循环神经网络带状回归模型的处理流程如图2所示。将患者的病理协变量 X 输入循环神经网络回归器中并以生存向量 Y 作为监督向量, 多个回归器输出的生存响应 $\Psi(X)$ 再经过带状响应转换, 则得到患者在每个生存子区间的风险概率, 即可拟合出患者的生存曲线。

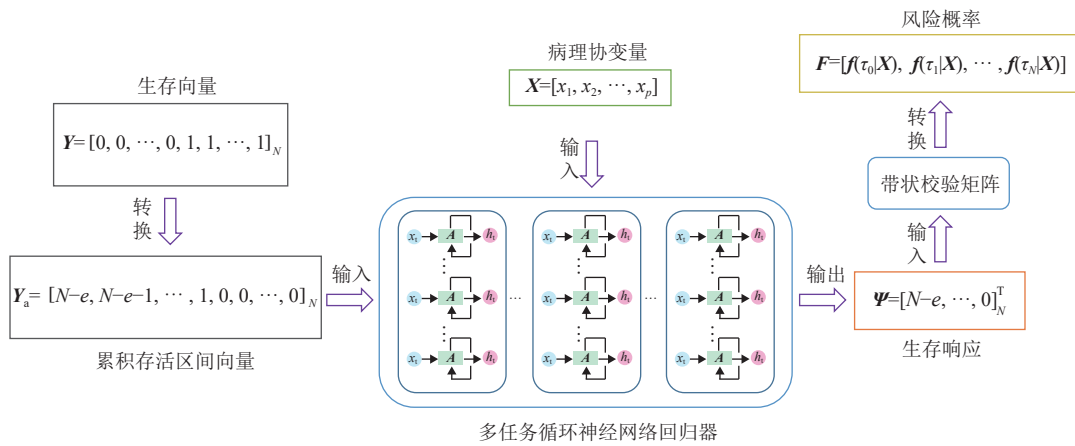


图2 多任务循环神经网络带状回归模型的操作流程

Fig.2 Operation process of RNN-MTBR

2.2 带状校验矩阵的优化

当乳腺癌患者的生存时间分为 N 个生存子区间, 则多任务回归模型的可信度矩阵^[6]为

$$A = \begin{bmatrix} A_1 & & & & \\ & \ddots & & & \\ & & A_i & & \\ & & & \ddots & \\ & & & & A_N \end{bmatrix} \quad (3)$$

式中: A_i 为组分 i 的系数矩阵。

MTBR 中将各组分系数矩阵设计为正二对角块非奇矩阵, 定义为 $A_i = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 。此种系数矩阵设计造成回归器生成每一子区间的生存响应只同上一子区间之间产生联系。为了在生成这一子区间的生存响应时对下一子区间的生存响应有所约束, 本文将带状校验系数矩阵优化为只含有主对角线元素和次对角线元素的 Toeplitz 矩阵^[7], 故带状校验矩阵 B 为

$$B = \begin{bmatrix} b_0 & b_1 & 0 & \dots & 0 \\ b_{-1} & b_0 & b_1 & \dots & 0 \\ 0 & b_{-1} & b_0 & \dots & 0 \\ \vdots & \vdots & \vdots & & b_1 \\ 0 & 0 & 0 & b_{-1} & b_0 \end{bmatrix} \quad (4)$$

且必须保持矩阵主对角线占优, 故矩阵需满足各行和(对称矩阵中同各列和)为1, 且 $b_0 > b_{-1} + b_1$ 。后续实验将研究 b_{-1} 、 b_1 变化对患者生存概率分布的影响。

2.3 生存函数计算及损失函数设计

本文中乳腺癌患者的风险分布函数为

$$f(\tau_s, X) = \frac{\exp(\psi(X) \mathcal{S} \Delta Y)}{\mathcal{Z}(\psi(X) \mathcal{S})} \quad (5)$$

式中: \mathcal{S} 为带状校验矩阵 B 的逆矩阵, Δ 为下三角矩阵, $\mathcal{Z}(\cdot)$ 为归一化操作。乳腺癌患者生存函数为

$$S(\tau_s, X) = 1 - \sum_{k=1}^s \frac{\exp(\psi(X) \mathcal{S} \Delta y)}{\mathcal{Z}(\psi(X) \mathcal{S})} \quad (6)$$

模型的损失函数由两部分组成, 第1部分为累计损失函数, 公式为

$$L_1 = \sum_{i=1}^K \left(\delta(k_i \neq \emptyset) \sum_{j=0}^N |\hat{F}(\tau_j | \mathbf{X}) - y_j| + \delta(k_i = \emptyset) \sum_{j=0}^{\tau_c} |\hat{F}(\tau_j | \mathbf{X}) - y_j| \right) \quad (7)$$

式中: $\hat{F}(\cdot)$ 为在离散生存子区间上的生存估计, K 为患者总数, k 为样本事件情况, τ 为生存子区间, τ_c 为事件发生或观察到删失时所在生存子区间。

第 2 部分为排序损失函数, 公式为

$$L_2 = \sum_{k=1}^K \sum_{i \neq j} \delta(k_i = k_j, y_i < y_j) \exp(\hat{F}(\tau_k | \mathbf{X}_j) - \hat{F}(\tau_k | \mathbf{X}_i)) \quad (8)$$

3 实验与讨论

3.1 数据集与实验预设置

本文采用 METABRIC^[18] 和 GBSG^[19] 2 种常用的乳腺癌真实数据集来评估模型的性能。METABRIC 数据集包含 1 980 例乳腺癌患者的 9 项特征记录, 其中约 58% 的患者死于患乳腺癌, 中位生存时间为 116 个月。GBSG 数据集包含 2 232 例乳腺癌患者的记录, 其中 1 546 例为淋巴结阳性乳腺癌患者, 以及 686 例随机临床试验的患者完整数据, 删失数据约占 57%。

本文采用生存分析研究中常用的一致性指数 (Concordance index, C-index)^[20] 作为评估指标, C-

index 常被用于评价预后模型的预测精度, 判断出模型的区分能力, 公式为

$$C\text{-index} = \frac{\sum_{i,j} \varphi(t_i < t_j) \varphi(\eta_i < \eta_j) \delta_i}{\sum_{i,j} \varphi(t_i < t_j) \delta_i} \quad (9)$$

式中: $\varphi(\cdot)$ 为患者对的比较, η 为生存子区间的风险值。

为了验证模型的可靠性, 将数据集以 6:2:2 的比例分为训练集、验证集和测试集, 并采用更换随机种子的方式进行 10 次实验, 基于正态分布法计算 C-index 的 95% 置信区间。RNN-MTBR 的模型采用循环神经网络作为核心, 隐藏层维度为 32, 选用 Adam 优化器, Dropout 概率设为 0.1, 学习率设为 0.01。

3.2 实验结果

3.2.1 带状校验矩阵系数对患者生存概率分布的影响

带状校验矩阵系数变化会对 RNN-MTBR 模型性能产生影响, 本文通过固定带状校验矩阵主对角线系数 b_0 , 变更次对角线系数 b_{-1} 、 b_1 的大小关系进行实验, 其中 $b_0 = \frac{5}{9}$, $b_{-1} \in \left[0, \frac{4}{9}\right]$, $b_1 \in \left[0, \frac{4}{9}\right]$, 且 $b_0 + b_{-1} + b_1 = 1$ 。为了更明显地看出带状校验矩阵系数对乳腺癌患者生存概率的作用效果, 实验将得到的患者相邻生存子区间的生存概率进行线性 10 等分, 并随机抽取其中 5 名患者的生存曲线进行绘制。如图 3 所示,

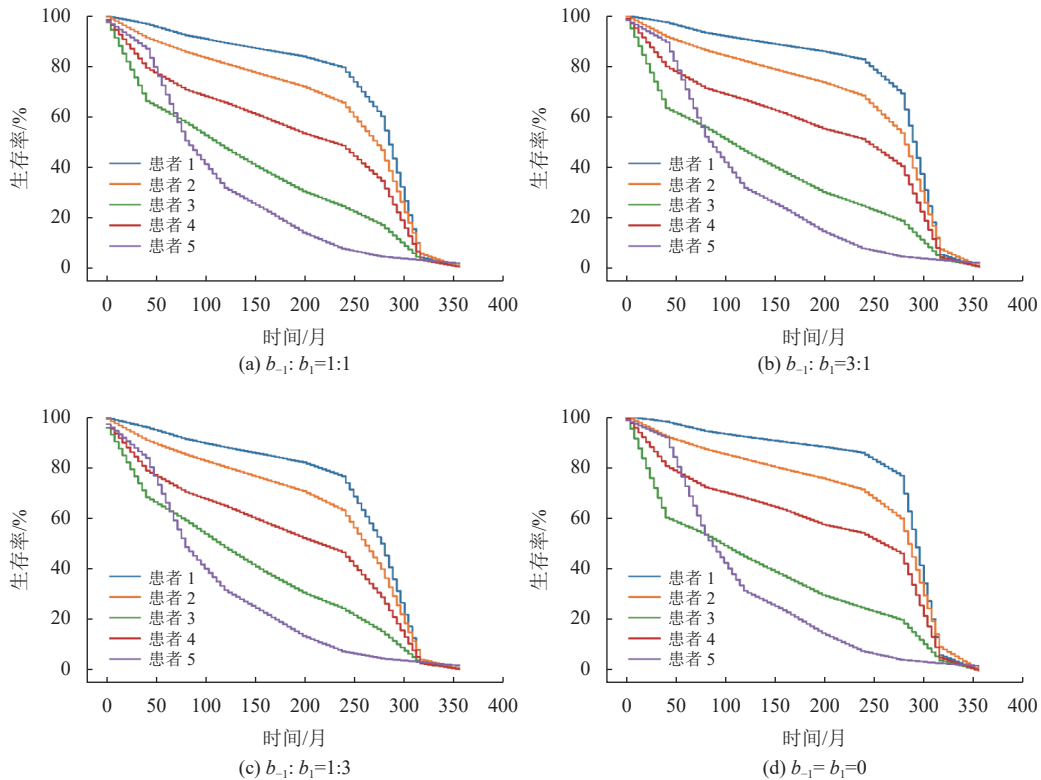


图 3 不同带状校验矩阵转换生成的患者生存曲线

Fig.3 Survival curves generated by different banded verification matrices

随着 b_{-1} 的增大,患者生存风险呈现逐步前移的现象,反之,随着 b_1 的增大,患者生存风险则出现逐步后移。而且能明显看出,在引入带状矩阵后,患者生存率的骤变情况能够得到明显的缓解。

3.2.2 带状校验矩阵优化对性能的影响

图4为各组分为正二对角块非奇矩阵的带状校验矩阵与优化为只含有主对角线元素和次对角线元素的Toeplitz矩阵生存模型C-index的对比。经比较,优化后带状校验矩阵对模型性能的提升更大。

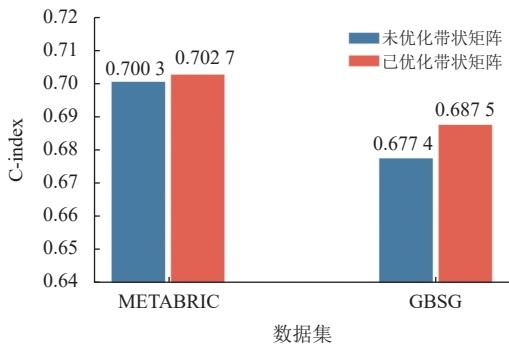


图4 带状校验矩阵优化前后模型C-index对比图

Fig.4 Comparison of C-index in survival models before and after banded check matrix optimization

3.2.3 生存分析模型性能对比

CoxCC^[21]、DeepSurv^[7]、DeepHit^[8]、SODEN^[9]、N-MTLR^[11]、Transformer-Based DSA^[12]和MTBR^[13]共7种生存分析模型作为RNN-MTBR的实验对比模型。表1为各生存分析模型在乳腺癌数据集测试集上的C-

index值及95%置信区间。从表1可以看出,本文提出的RNN-MTBR模型效果在变更回归器核心为循环神经网络和优化了带状校验矩阵后,模型在2个数据集上的预测性能相比原先的MTBR模型均有所提升,且同其他主流的生存分析模型相比具有最好的预测性能。

图5给出了几种模型分别在GBSG数据集中拟合相同的5例乳腺癌患者的生存曲线的效果。CoxCC和DeepSurv属于比例风险模型,该种模型默认所有个体拥有相同的基线风险,且实际风险与基线风险成正比,使得计算得到的患者间风险不会出现交叉,这与现实相悖。DeepHit没有考虑对连续时间的尺度离散化,导致患者间生存曲线差异化不明显。N-MTLR将多层感知器引入多任务回归中,能提取患者生存数据中的非线性关系,但存在响应转换在阈值附近波动的现象,导致生存曲线中较多出现生存率下降幅度减小的情况。MTBR通过引入带状校验矩阵有效抑制了响应转换波动在阈值附近出现的现象,使得患者生存曲线能呈现较明显的下降趋势,但患者之间差异性的体现仍不明显。RNN-MTBR将MTBR核心替换为RNN,输入数据则得以在各时间步之间实现上下文关联,并能在不同时间步之间共享权重,并合理调整带状校验矩阵,使得患者生存概率下降幅度随时间变化显得更为合理且患者之间的生存曲线能有较为明显的差异性,从而方便个性化预后管理的施行。

表1 各生存分析模型所获取的C-index(95%置信区间)¹⁾

Table 1 C-index in 95% confidence interval achieved by each survival analysis model

模型	C-index	
	METABRIC	GBSG
CoxCC	0.670 7(0.660 8~0.680 6)	0.672 8(0.661 1~0.684 5)
DeepSurv	0.654 5(0.642 6~0.666 4)	0.676 3(0.663 5~0.689 1)
DeepHit	0.683 5(0.675 9~0.691 1)	0.620 6(0.606 4~0.634 8)
SODEN ^[9]	0.658*(0.650~0.662)	—
N-MTLR	0.678 7(0.671 4~0.685 9)	0.618 2(0.601 3~0.632 6)
Transformer-Based DSA ^[12]	0.640*(0.624~0.656)	—
MTBR	0.700 3(0.692 6~0.708 0)	0.677 4(0.663 7~0.691 1)
RNN-MTBR	0.705 3(0.699 9~0.710 8)	0.697 5(0.689 9~0.705 0)

1) 加粗项为最佳结果; *项为论文引用结果。

4 结论

本文采用循环神经网络作为多任务回归器,优化设计了带状校验矩阵,提出了一种多任务循环神经网络带状回归模型,并应用于乳腺癌个体生存分

析。在2个乳腺癌真实数据集的生存分析中,多任务循环神经网络带状回归模型的C-index较医学界常用的Cox比例风险模型提升了0.02,且有效缩小了95%置信区间。研究还通过可视化的生存曲线验证了新模型对患者间区分效果的增益性。

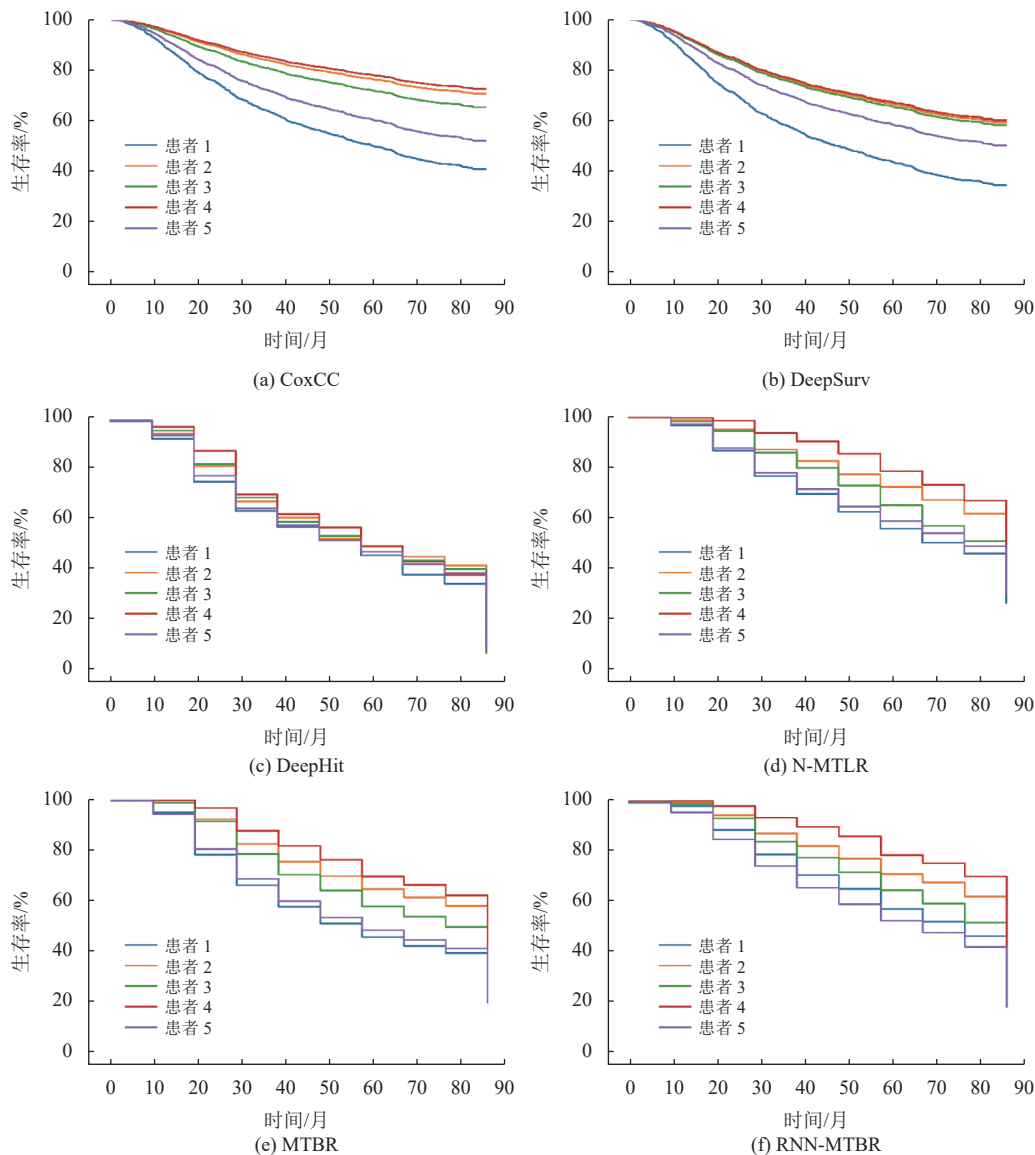


图 5 各生存分析模型拟合GBSG数据集中五名患者的生存曲线

Fig.5 Survival curves of five patients in GBSG dataset fitted from each survival analysis model

参考文献:

- [1] ZHOU X, LI C, RAHAMAN M M, *et al.* A comprehensive review for breast histopathology image analysis using classical and deep neural networks [J]. *IEEE Access*, 2020, 8: 90931-90956.
- [2] ARNOLD M, MORGAN E, RUMGAY H, *et al.* Current and future burden of breast cancer: global statistics for 2020 and 2040 [J]. *The Breast*, 2022, 66: 15-23.
- [3] LOH S Y, YIP C H. Breast cancer as a chronic illness: implications for rehabilitation and medical education [J]. *Journal of Health and Translational Medicine*, 2006, 9(2): 3-11.
- [4] GEORGE B, SEALS S, ABAN I. Survival analysis and regression models [J]. *Journal of Nuclear Cardiology*, 2014, 21(4): 686-694.
- [5] GANGGAYAH M D, TAIB N A, HAR Y C, *et al.* Predicting factors for survival of breast cancer patients using machine learning techniques [J]. *BMC Medical Informatics and Decision Making*, 2019, 19(1): 1-17.
- [6] SCHOBEL P, VETTER T R. Survival analysis and interpretation of time-to-event data: the tortoise and the hare [J]. *Anesthesia and Analgesia*, 2018, 127(3): 792.
- [7] KATZMAN J L, SHAHAM U, CLONINGER A, *et al.* DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network [J]. *BMC Medical Research Methodology*, 2018, 18(1): 1-12.
- [8] LEE C, YOON J, VAN DER SCHAAR M. Dynamic-deep-hit: a deep learning approach for dynamic survival analysis with competing risks based on longitudinal data [J]. *IEEE Transactions on Biomedical Engineering*, 2019, 67(1): 122-133.
- [9] TANG W, MA J, MEI Q, *et al.* SODEN: a scalable continuous-time survival model through ordinary differential equation networks[J]. *J Mach Learn Res*, 2022, 23: 34: 1-34: 29.

- [10] YU C N, GREINER R, LIN H C, *et al.* Learning patient-specific cancer survival distributions as a sequence of dependent regressors [J]. *Advances in Neural Information Processing Systems*, 2011, 24: 1845-1853.
- [11] FOTSO S. Deep neural networks for survival analysis based on a multi-task framework[EB/OL]. arXiv: 1801.05512 (2018-01-17) [2023-01-09]. <https://doi.org/10.48550/arXiv.1801.05512>.
- [12] HU S, FRIDGEIRSSON E, VAN WINGEN G, *et al.* Transformer-based deep survival analysis[C]//Survival Prediction-Algorithms, Challenges and Applications. New York: PMLR, 2021: 132-148.
- [13] 王慧恒, 蔡念, 陈睿, 等. 面向癌症个体生存分析的多任务带状回归模型[J]. *计算机工程与应用*, 2023, 59(10): 299-305.
WANG H H, CAI N, CHEN R, *et al.* Multi-task banded regression model for individual cancer survival analysis [J]. *Computer Engineering and Applications*, 2023, 59(10): 299-305.
- [14] MEDSKER L R, JAIN L C. Recurrent neural networks [J]. *Design and Applications*, 2001, 5: 64-67.
- [15] WU Y, HALABI S. Interval censoring[M]. Boca Raton: Chapman and Hall/CRC, 2019: 493-508.
- [16] BUHLMANN H, GISLER A. Credibility in the regression case revisited (A late tribute to Charles A. Hachemeister) [J]. *ASTIN Bulletin: The Journal of the IAA*, 1997, 27(1): 83-98.
- [17] YE K, LIM L H. Every matrix is a product of toeplitz matrices [J]. *Foundations of Computational Mathematics*, 2016, 16(3): 577-598.
- [18] CURTIS C, SHAH S P, CHIN S F, *et al.* The genomic and transcriptomic architecture of 2000 breast tumours reveals novel subgroups [J]. *Nature*, 2012, 486(7403): 346-352.
- [19] SCHUMACHER M, BASTERT G, BOJAR H, *et al.* Randomized 2 × 2 trial evaluating hormonal treatment and the duration of chemotherapy in node-positive breast cancer patients [J]. *Journal of Clinical Oncology*, 1994, 12(10): 2086-2093.
- [20] ANTOLINI L, BORACCHI P, BIGANZOLI E. A time-dependent discrimination index for survival data [J]. *Statistics in Medicine*, 2005, 24(24): 3927-3944.
- [21] KVAMME H, BORGAN Ø, SCHEEL I. Time-to-event prediction with neural networks and Cox regression[EB/OL]. arXiv: 1907.00825 (2019-09-13) [2023-01-09]. <https://doi.org/10.48550/arXiv.1907.00825>.

(责任编辑: 杨耀辉 英文审核: 熊荣斌)



(上接第26页)

- [25] JOSHI P, MASILAMANI V, MUKHERJEE A. A knowledge graph embedding based approach to predict the adverse drug reactions using a deep neural network [J]. *Journal of Biomedical Informatics*, 2022, 132: 104122-104131.
- [26] ZHANG F, SUN B, DIAO X L, *et al.* Prediction of adverse drug reactions based on knowledge graph embedding [J]. *BMC Medical Informatics and Decision Making*, 2021, 21(1): 38-48.
- [27] WANG M, MA X Y, SI J W, *et al.* Adverse drug reaction discovery using a tumor-biomarker knowledge graph[J]. *Frontiers in Genetics*, 2021, 11: 625659
- [28] DEY S, LUO H, FOKOUE A, *et al.* Predicting adverse drug reactions through interpretable deep learning framework [J]. *BMC Bioinformatics*, 2018, 19(Suppl 21): 476-488.
- [29] BEAN D M, WU H H, IQBAL E, *et al.* Knowledge graph prediction of unknown adverse drug reactions and validation in electronic health records [J]. *Scientific Reports*, 2017, 7(1): 16416.
- [30] WISHART D S, FEUNANG Y D, GUO A C, *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018 [J]. *Nucleic Acids Research*, 2018, 46(D1): D1074-D1082.
- [31] KUHN M, LETUNIC I, JENSEN L J, *et al.* The SIDER database of drugs and side effects [J]. *Nucleic Acids Research*, 2016, 44(D1): D1075-D1084.
- [32] YANG B S, YIH W T, HE X D, *et al.* Embedding entities and relations for learning and inference in knowledge bases[EB/OL]. arXiv: 1412.6575 (2015-08-29) [2023-02-22]. <https://doi.org/10.48550/arXiv.1412.6575>.
- [33] NICKEL M, LORENZO R, TOMASO P. Holographic embeddings of knowledge graphs[C]//Thirtieth AAAI Conference on Artificial Intelligence. California: AAAI, 2016: 1955-1961.
- [34] 高飞萌, 宋智慧. 艾司奥美拉唑镁肠溶片致肾损伤[J]. *药物不良反应杂志*, 2020, 22(6): 381-382.
GAO F M, SONG Z H. Kidney injury caused by esomeprazole magnesium enteric-coated tablets [J]. *Adverse Drug Reactions Journal*, 2020, 22(6): 381-382.
- [35] DHAKRIT R, IVES A L, HARRIOTT N G, *et al.* Comparative incidence of acute kidney injury in patients on vancomycin therapy in combination with cefepime, piperacillin-tazobactam or meropenem [J]. *Journal of Chemotherapy (Florence, Italy)*, 2021, 34(2): 1-7.
- [36] NIITSU T, HAYASHI T, UCHIDA J, *et al.* Drug-induced kidney injury caused by osimertinib: report of a rare case [J]. *Nephron*, 2021, 146(1): 58-63.
- [37] WAKABAYASHI K, YAMAMOTO S, HARA S, *et al.* Nivolumab-induced membranous nephropathy in a patient with stage IV lung adenocarcinoma [J]. *CEN Case Rep*, 2022, 11(2): 171-176.
- [38] 苏伟, 苏培. 碘海醇注射液致患者获得性急性肾功能衰竭的不良反应[J]. *中国生化药物杂志*, 2016, 36(10): 92-94.
SU W, SU P. Effect of adverse reaction of acquired acute renal failure caused by iohexol injection [J]. *Chinese Journal of Biochemical and Pharmaceuticals*, 2016, 36(10): 92-94.

(责任编辑: 杨耀辉 英文审核: 熊荣斌)