

中文医学期刊编辑对生存分析的核查要点

罗云梅 蒲素清 李缨来[†]

四川大学华西医院《中国普外基础与临床杂志》编辑部,610041,成都

摘要 目前医学期刊中生存分析的应用十分广泛。在处理生存分析类稿件时,编辑需审查生存曲线是否仅有横纵坐标,生存曲线是否存在交叉,多因素分析时样本量是否足够,多因素分析时变量纳入方法是否正确,多因素分析结果是否完整,RR值、RR 95% CI及P值之间是否一致等。本文详细阐述了上述生存分析中的审查要点,以供编辑人员参考。

关键词 医学期刊;生存分析;审查要点

Key points for survival analysis in the Chinese medical journals//LUO Yunmei, PU Suqing, LI Yinglai

Abstract Survival analysis is commonly reported in the Chinese medical journals. When we deal with survival analysis - related papers, we should pay attention to the following points: whether the survival curve has only horizontal or vertical coordinate, whether the survival curve has overlapping, whether the sample and variables for Cox proportional hazard model are suitable or not, whether relative risk value, 95% confidence interval, and P value coincide fairly well, and so on. This paper elaborates on these key points in order to provide reference to the editors of Chinese medical journals.

Keywords medical journal; survival analysis; key points

Authors' address Editorial Department of Chinese Journal of Bases and Clinics in General Surgery, West China Hospital, 610041, Chengdu, China.

DOI:10.16811/j.cnki.1001-4314.2018.01.009

目前,医学论文中的生存资料(survival data)非常常见。笔者在万方数据库中以“生存分析或预后分析或Cox”进行检索(检索条件为题名),共检索出文献2万6476篇(其中2016—2017年的文献1773篇),这还不包括其他中文数据库中收录的其他文献,或者题目中没有涵盖这些表述但实际进行了生存分析的文献,因此实际的生存分析文章远不止检索出的数量。这些论文中,生存分析涵盖了医学^[1]、农业^[2]、交通^[3]等领域,可见目前生存分析的应用十分广泛。由于生存资料蕴涵有结局和时间2个方面的信息,单纯针对结局或时间的统计分析方法都不恰当,需采用特殊分析方法即生存分析。随着生存分析的广泛应用,医学论文中关于一些生存分析不恰当的表述、甚至生存分析统计方法的误用情况不断出现,严重降低了文章的质量。因此,笔者总结了医学期刊生存分析中的审查

要点,以供编辑人员参考。

1 统计描述部分

1.1 注意生存概率和生存率的正确表述 首先我们来看一下生存概率和生存率的定义。生存概率表示在某单位时段开始时存活的个体到该时段结束时仍存活的可能性大小。某年生存概率的计算公式为: $p = N_1 / N_0$,其中 N_1 为某年活满1年人数, N_0 为某年年初观察人数;而生存率是指观察对象活过某时刻的概率,一般采用概率乘法原理估计生存率(如3年生存率等于第1年、第2年和第3年生存概率的连乘积)。医学论文中“1年生存率”“3年生存率”“5年生存率”的表述非常常见。但这种表述是否正确呢?这还要看作者是如何得出该数值的。若作者是根据生存曲线,取值1、3及5年对应的生存率,则表述正确。但若作者是根据观察例数和存活例数计算的,如3年生存率是计算的活满3年的患者数量在随访患者中所占的比例,则其本质是3年生存概率,则3年生存率的表述就欠妥。

1.2 生存曲线的审查要点 既往有学者^[4]详细总结了生存曲线的常见错误,包括曲线形状不规则、曲线上没有截尾数据、一批患者不同的分层曲线上截尾数据的有无不一致或例数不一致、曲线在后期或中期出现很长的平台及生存曲线横坐标为“0”时生存率不是“100%”,笔者在此不再赘述。本文总结了生存曲线的其他错误,如生存曲线包括了横纵坐标以外的坐标、生存曲线存在交叉、生存曲线对应随访时间和文字结果不一致的问题。

1.2.1 生存曲线仅有横、纵坐标 生存曲线的本质是线图。线图的要求是:纵轴一般以0点作起点,否则需作特殊标记或说明;各测定值标记点间以直线连接,不可修匀成光滑曲线。在实际工作中,编辑往往要求作者提供可供编辑的矢量图,通常是“emf”或“wmf”格式。而统计软件输出的统计图,如最常用的SPSS,其生存分析图片都是不符合编辑规范的,因此要求编辑在后期的校对时,将无关线条予以删除。而目前的医学论文中,很多生存分析曲线图都不是规范的统计图,如徐冉等^[1]、傅文祯等^[5]、张鹏等^[6]的论文中,生存分析曲线图中都有一些无关的线条。

1.2.2 仅进行单因素分析时生存曲线不能交叉 有

[†] 通信作者

些医学论文^[5,7]中呈现的生存曲线是存在交叉的,这是生存曲线最常见的错误之一。关于生存曲线不能交叉的原因及应对策略,将在第二部分予以总结,此处不再赘述。

1.2.3 生存曲线最末一点对应的随访时间不应与文字相悖 生存曲线依据生存时间绘制。医学论文中往往报道了随访时间,而最长生存时间和最长随访时间的关系是,最长生存时间 \leq 随访时间。因此,若生存曲线中的生存时间长于文中的随访时间,则数据就存在较大的问题,需要编辑提出质疑,请作者核查。

2 统计推断部分

2.1 单因素分析时审查生存曲线是否存在交叉 目前 SPSS 软件用于比较生存曲线的方法有 log-rank 检验、Breslow (BS) 检验(即 Gehan-Wilcoxon 检验)和 Tarone-Ware (TW) 检验。目前医学期刊的“主流”方法是 log-rank 检验。其实 3 种检验方法有其适用条件,各有侧重:log-rank 检验更注重“远期效应”,即生存曲线开始差距小,随时间推移逐渐分开的曲线适用 log-rank 检验(图 1);BS 检验更注重“近期效应”,即生存曲线开始差距大,随时间推移逐渐接近的曲线适用 BS 检验(图 2);TW 检验介于两者之间。由此可知,由于 TW 和 BS 检验在前期赋予较大的权重,因此该 2 种方法对生存曲线后期的差异不敏感,因此不适合后期差异较大的生存曲线的比较。由于医学上对患者的随访,常常是开始生存情况较好,随时间延长生存率降低,生存曲线逐渐分开,因此 log-rank 检验在医学期刊中最常见。但 log-rank 检验有一重要条件(也是 BS 检验和 TW 检验的重要条件),就是生存曲线不能交叉^[8-9],因此进行生存曲线的绘制非常有必要。假设 2 条生存曲线有 1 个交叉点,交叉之前和交叉之后的生存情况完全相反,则哪组的生存情况更好呢?无法判定。

生存曲线交叉提示存在混杂因素。针对生存曲线交叉的情况,目前的处理方法有 Reny (RE)、Lin-Wang (LW) 和 Two-stage (TS) 法,其中尤以 TS 法具有较好的适用性与稳健性^[10-12],也可进行分层分析或构建 Cox 比例风险回归模型控制混杂因素的影响。

2.2 多组生存曲线有统计学意义时审查是否需进行两两比较 当涉及 3 个组别及以上的生存曲线的比较时,若总体有统计学意义,只能得出至少有 2 组比较差异有统计学意义的结论,但具体哪组好?哪组差?则需要通过两两比较来回答。生存分析的两两比较一般通过校正 α 来实现,通过 Bonferroni 过程、Dunn-Šidák 过程^[13]、Benjamini-Hochbe 过程、Benjamin-Yekutieli 过

程等^[14]对 α 进行校正。但笔者认为这些方法较复杂,可通过构建多因素 Cox 比例风险回归模型来解决,直接实现类别间的比较,而不需单独校正 α 。

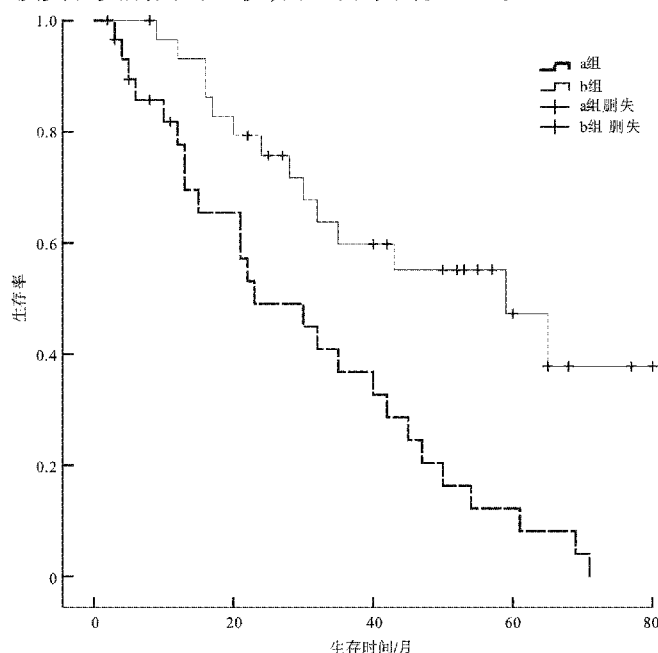


图1 生存曲线示例 1

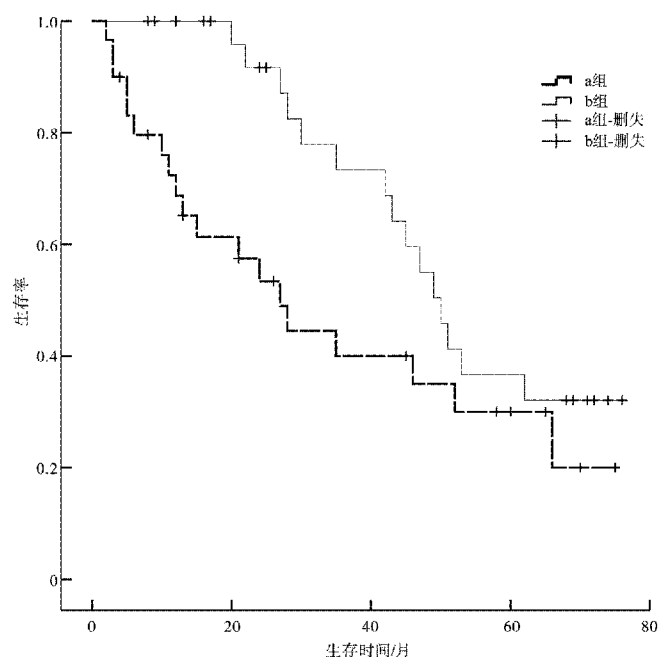


图2 生存曲线示例 2

2.3 多因素分析 Cox 比例风险回归模型为最常用的多因素生存分析方法。医学期刊编辑在处理此类文章时,需从以下几个方面进行核查。

2.3.1 多因素分析审查样本量是否足够 进行 Cox 比例风险回归模型分析时,样本含量不宜过小。随着纳入变量的增加,样本量应有所增加。孙振球等^[9]认为,Cox 比例风险回归模型分析时,样本量应该为纳入

变量的5~20倍。如《28例同时多原发肺癌的生存分析及预后影响因素研究》^[15]中,拟合了Cox比例风险回归模型,显然样本量过少,结论有待商榷。

2.3.2 变量纳入方法审查 Cox比例风险回归模型和其他多因素模型一样,变量的纳入需考虑专业、临床实际、既往研究、变量之间的关系,进行综合选择。尤其需注意的是,变量的纳入不能单纯依靠单因素分析结果,即只纳入单因素分析时有统计学意义的变量^[7, 16-17],这是医学期刊多因素分析文章中最常出现的问题。例如,众所周知,肿瘤的TNM分期对于患者预后的影响是肯定的,分期越晚,预后越差。但若单因素分析时,TNM分期并无统计学意义,多因素分析就不纳入了吗?答案是否定的。单因素分析之所以没有统计学意义,即是因为单因素分析时只分析了TNM分期一个变量,并未对其他混杂因素进行控制,因此导致了和临床实际相悖的结论。那据此就否定了单因素分析的价值吗?答案也是否定的。当病例数量较少,难以进行多因素分析时,单因素分析作为一种探索方法,对于各因素对预后影响的探索,也是很有必要的。

2.3.3 不同类型变量纳入形式核查 在变量纳入过程中,对于计量资料,直接纳入模型即可,但对于计数资料,不同的计数资料类型有不同的纳入形式。如闫丽等^[18]构建了高血压患者脑卒中发病影响因素的Cox比例风险回归模型,得到表1;张小鹏等^[19]构建了鼻咽癌生存影响因素的Cox比例风险回归模型,得到表2。根据表1和表2可知,年龄和性别为二分类变量,文化程度为无序多分类资料,临床分期为有序多分类资料即等级资料。那为何文化程度对应2行数据,而临床分期对应1行数据呢?原因是无序多分类是作为哑变量纳入模型(哑变量数量为分类数量-1,如三分类资料有2行数据、四分类资料有3行数据,以此类推);而临床分期,作者是将其以线性变量纳入模型,即将等级资料类似于计量资料纳入模型,因此只有1行结果。若临床分期是作为哑变量纳入的,则也需有3行结果。但是,无序多分类不能以线性形式纳入模型。编辑在编修论文时,需注意审查作者变量纳入和结果表述的一致性,尤其对于无序多分类资料以哑变量形式纳入模型时,需注意结果呈现是否完整或作者变量纳入形式是否有误。此外,对于以计数资料形式纳入的变量,一定要写明对照,如表1中,性别是以男性为对照,文化程度是以小学及以下为对照;表2中对照不是很明确,虽然作者给出了具体变量赋值(一般软件拟合模型时常规以赋值为0的类别作为对照),但从科学性角度来讲,明确对照也是必要的。

2.3.4 Cox比例风险回归模型的模型诊断 开展Cox

比例风险回归模型时,需重点考察比例风险假设是否满足,因只有在满足比例风险假设条件下进行拟合时,其参数估计值才是有效的。对于风险比例假设的验证,常用的方法包括图形法和构建时间依存性协变量的方法,前者只是粗略的评估方法,更严谨的是后者,一般采用软件计算。笔者认为,比例风险性的检验结果作为模型拟合的前提,可常规呈现。

表1 高血压患者脑卒中发病影响因素的Cox比例风险回归模型(节选)

因素	β 值	Wald χ^2 值	RR 值 ⁽¹⁾	95% CI 值	P 值
男性,文化 小学及以下	-0.45	12.39	0.64	0.50~0.82	<0.001
初中	-0.34	5.49	0.71	0.53~0.95	0.019
高中及以上	-1.07	11.36	0.35	0.19~0.64	0.001

注:1)原文统计量为OR,笔者将其正确修改为RR。

表2 鼻咽癌生存影响因素的Cox比例风险回归模型(节选)

因素	回归系数	标准误	Wald χ^2 值	相对危险度	置信区间	P 值
年龄	0.516	0.237	9.726	1.597	1.171~1.978	0.002
临床分期	0.323	0.167	7.908	1.364	0.853~1.924	0.003

注:年龄赋值:0为20~45岁,1为>45岁;临床分期赋值:0为I期,1为II期,2为III期,3为IV期。

2.3.5 多因素分析结果是否呈现完整 表1和表2呈现内容不同的地方在于,表1呈现了Wald χ^2 值,而表2呈现了标准误信息。笔者认为,Cox比例风险回归模型结果的呈现应同时包括回归系数(β 值)、Wald χ^2 值、标准误、RR值及其95%置信区间(95% CI)、P值及拟合优度信息。这样一方面可全面呈现分析结果,另一方面在一定程度上有助于审查模型的合理性。若某个变量标准误特别大(如>100),则在很大程度上提示模型存在问题,需请作者进一步核查。模型的拟合优度即模型拟合实际观测值的程度,建议常规呈现。

2.3.6 审查RR值、RR 95% CI及P值之间的一致性

RR值、RR 95% CI及P值之间的关系是,RR值一定在95% CI区间范围内;P值小于0.05时,RR 95% CI是一定不包括1的。因此,表2中的临床分期结果不正确(P值小于0.05,但95% CI却包括了1),需请作者排查。

2.3.7 审查多因素分析方法是否优选Cox比例风险回归模型 生存分析中的死亡时间是一个广义的概念,不单是指通常意义上的生物体死亡,而是泛指标志某种处理措施失效的特征事件,如肿瘤的复发^[20]、患者的无复发生存^[21]、股动脉内膜剥脱术后的再狭窄^[22]等,都可使用生存分析。有时作者采用logistic回归对这些资料进行分析,但该方法只利用了结局信

息,而未利用时间信息,并不是最优的统计分析方法。编辑可对作者提出优选方法建议。

随着生存分析在医学领域中应用的逐渐广泛,涉及生存分析的医学论文越来越多。医学期刊编辑在处理涉及生存分析的文章时,需从统计描述、统计推断两部分进行审核,多方位把关,遇到疑问之处务必告知作者核查,以尽量降低出错概率,提升论文质量。

3 参考文献

- [1] 徐冉,蒋宏毅,赵晓昆,等. Boari 膀胱肌瓣术治疗中下段输尿管癌患者的生存分析[J]. 中南大学学报(医学版), 2014, 39(8): 857
- [2] 杜运苏,陈小文. 我国农产品出口贸易关系的生存分析: 基于 Cox PH 模型[J]. 农业技术经济, 2014(5): 101
- [3] 张洪宾,孙小端,贺玉龙,等. 公交站和出口组合影响的交通流生存分析[J]. 北京工业大学学报, 2015, 41(1): 104
- [4] 庄爱华,甘可建,柳青,等. 肿瘤学研究论文中的生存曲线图谬误辨析[J]. 编辑学报, 2008, 20(6): 491
- [5] 傅文祿,严信祺,陈拥军,等. 腹腔镜胆囊切除术发现意外胆囊癌患者生存分析[J]. 中华腔镜外科杂志, 2014, 7(2): 79
- [6] 张鹏,杨明,刘涛,等. 三阴性乳腺癌的临床病理特征及生存分析[J]. 中国老年学杂志, 2012, 32(8): 1591
- [7] 李青楠,张恒,董喆,等. 肾癌合并静脉癌栓的生存分析[J]. 中国医科大学学报, 2014, 43(9): 848
- [8] 李晓松. 医学统计学[M]. 2版. 北京:高等教育出版社, 2008: 190
- [9] 孙振球,徐勇勇. 医学统计学[M]. 2版. 北京:人民卫生出版社, 2008: 367
- [10] 李慧敏,韩栋,陈征,等. 生存曲线交叉时统计推断方法的比较和选择[J]. 中国卫生统计, 2013, 30(5): 672
- [11] 李慧敏. 生存曲线交叉时统计推断方法的比较研究[D]. 广州:南方医科大学, 2011
- [12] 陈金宝,邱李斌,王北琪,等. 固定点处组间生存率比较的统计检验法[J]. 中华流行病学杂志, 2015, 36(2): 188
- [13] Heo M, Leon A C. Comparison of post hoc multiple pairwise testing procedures as applied to small k-group logrank tests [J]. Int J Statist Med Res, 2013, 2(2): 104
- [14] 曹佩华,郑一男,欧春泉. 基于 α 校正的生存曲线多重比较方法 [C]. 2011 年中国卫生统计学年会论文集, 2011: 25-27
- [15] 梁朝阳,刘德若,郭永,等. 2015 中华中医药学会外科年会论文集[C]. 2015: 483
- [16] 李秋华,葛亚楠,刘兆喆,等. 乳腺癌辅助化疗诱发 2 型糖尿病患者预后因素及生存分析[J]. 临床误诊误治, 2016, 29(10): 85
- [17] 赵梦龙,郭凯,张源波,等. 心脏骤停自主循环恢复患者预后的多因素 Cox 生存分析[J]. 中国急救医学, 2017, 37(2): 151
- [18] 闫丽,曹焯,汪秋实,等. 高血压患者脑卒中发病影响因素 Cox 回归分析[J]. 中国公共卫生, 2016, 32(6): 751
- [19] 张小鹏,李秋捷,边枫,等. 110 例鼻咽癌患者预后因素及生存分析[J]. 中国耳鼻咽喉颅底外科杂志, 2014, 20(6): 509
- [20] 王伟英,桑文文,焉双梅,等. 急性缺血性脑卒中患者 1 年复发危险因素 Cox 回归分析[J]. 中华老年心脑血管病杂志, 2016, 18(1): 49
- [21] 杨等霞,周婵,王心悦,等. 空洞性肺癌发生相关因素及无进展生存分析[J]. 中华肿瘤杂志, 2015, 37(7): 538
- [22] 李立强,谷涌泉,佟铸,等. 股动脉内膜剥脱术后再狭窄危险因素的 Cox 比例风险回归模型分析[J]. 中国普外基础与临床杂志, 2017, 24(9): 1073
(2017-06-26 收稿;2017-08-25 修回)

编辑缩影(一字诗 2 首)

周桂莲/西北农林科技大学《麦类作物学报》编辑部

(一)

一心一意一期刊,一载难偷一日闲
一字一词一句读,一丝不妥一魂牵

(二)

一支朱笔一荧屏,一摞新篇一管灯
一稿齐清一轮月,一头华发一生情