Win ratio法在临床试验中比较两组交叉生存曲线的应用*

张　铮[1#]　郑思睿[2#]　哈艳萍[1]　米白冰[1]　陈方尧[1]　康轶君[1]　Duolao　Wang[2△]　裴磊磊[1△]

【提　要】 目的　介绍win ratio方法在解决临床试验中生存曲线交叉问题方面的应用，包括基本原理和应用方法。**方法**　以一项临床试验为例，终点事件为肾透析患者发生首次感染，评估传统的外科手术置管与经皮下置管两种方法对肾透析患者的疗效，使用win ratio方法，在手术组和经皮组生存曲线互相交叉的情况下，比较两组疗效间的差异。**结果**　log-rank检验的P值为0.11，无法判定疗效优劣。Cox回归分析的P值为0.12，由于生存曲线交叉，风险比（hazard ratio，HR）不是恒定值，传统的Cox回归模型无法应用，结果无意义。win ratio法分析结果显示，P值大小为0.01，疗效间差异有统计学意义。手术置管组发生感染概率约是经皮置管组的3.17倍，即手术置管组的感染发生率较高，95%CI为1.27~7.87。**结论**　生存曲线交叉时，log-rank检验存在偏差，且Cox模型的比例风险（proportional hazards，PH）假设不成立，win ratio法不需要假设HR恒定，故win ratio法在检测疗效差异上优于log-rank检验和Cox模型。

【关键词】 win ratio　生存曲线交叉　生存分析　临床试验

【中图分类号】 R195.1　　**【文献标志码】** A　　**DOI** 10.11783/j.issn.1002-3674.2026.03.023

很多临床试验会将生存分析中的事件发生时间或者其他生存数据作为结局，在比较两种生存资料的分布时，研究者通常会采用log-rank检验评价两条生存曲线是否存在显著差异，即两种治疗方法的疗效是否不同。据此获得的相关P值虽然可以判断疗效之间是否存在差异，但无法衡量差异大小，因此，P值通常还伴随着Cox回归模型提供的风险比（hazard ratio，HR）及其95%CI[1]。然而，Cox模型成立的前提假设是两组的风险函数必须成比例，即HR应为一个固定值。当两组生存曲线出现交叉的情况时，HR不恒定，Cox回归模型可能会导致误导性的结果，log-rank检验也失去了统计效力[2]。如果依旧采用log-rank检验来检测两条交叉生存曲线之间的差异，显然是违背比例风险（proportional hazards，PH）假设的。在本文中，我们使用win ratio方法来解决生存曲线交叉问题。该方法由Pocock等[3]首次提出，win ratio不仅提供P值，还能提供两条生存曲线差异的点估计和95%CI，并对疗效之间的差异进行检验和衡量。

原理与方法

1. win ratio方法的应用

win ratio法可应用的场景很多，例如，分析临床试验中的复合终点或者连续非正态分布的结局等等[4]。有研究指出，对于存在独立删失的生存数据，可以采用基于倾向评分的逆概率加权方法计算调整后的win ratio[5]。另外，分层调整win ratio方法也可以用来控制患者基线特征不平衡[6]。

win ratio的计算主要有3种方法，包括匹配法、不匹配法以及匹配且分层法[3]。匹配法在使用时，要考虑到复合终点对患者的重要性，依次降序比较，非匹配法则适合没有危险因素分层的试验，匹配且分层法适用于随访时间较长的临床试验。由于很多研究在匹配患者时没有较为客观的匹配机制和方案，实际操作中也无法对参加试验的全部患者进行匹配，故大多研究者更推荐不匹配法，且该方法亦适用于小样本量的研究以及亚组分析。

例如，实验组的患者例数表示为N_A，对照组的患者例数表示为N_B，实验组中的每一位患者都与对照组中的每一位患者进行一次配对，所有可能的配对数为$N_A \times N_B$。对于每一对，根据治疗效果即是否出现感兴趣的终点事件，可以将实验组的患者标记为“winner”、“loser”或者“tied”。计算winner、loser以及tied的总数，分别表示为N_W、N_L、N_T，则有$N_W+N_L+N_T=N_A \times N_B$。win ratio的值表示为$R_W$，$R_W=N_W/N_L$，该值是临床试验中用于评估生存分析终点结局的治疗效果的统计量。如果假定实验组中第i号受试者（$i=1,2,\cdots,n_t$）与对照组中第j号受试者（$j=1,2,\cdots,n_c$）配对，结局分别为X_i和Y_j，输赢的统计指标可以列为一个$n_t \times n_c$矩阵，元素如下所示：

$$C_{ij}=C(X_i,Y_j)=\begin{cases}1, & \text{实验组的受试者}i\text{赢了对照组的受试者}j\\ -1, & \text{实验组的受试者}i\text{输给对照组的受试者}j\\ 0, & \text{实验组受试者}i\text{与对照组的受试者}j\text{平局}\end{cases}$$

实验组胜利总数为$TW=\sum_{i=1}^{n_t}\sum_{j=1}^{n_c} I\left\{C(X_i,Y_j)=1\right\}$

相应地，输的总数为

$$TL=\sum_{i=1}^{n_t}\sum_{j=1}^{n_c} I\left\{C(X_i,Y_j)=-1\right\}$$

* 基金项目：国家自然科学基金（72174167）

1. 西安交通大学医学部公共卫生学院（710061）

2. Department of Clinical Sciences，Liverpool School of Tropical Medicine

#共同第一作者

△通信作者：Duolao Wang，E-mail：Duolao. Wang@lstmed. ac. uk；裴磊磊，E-mail：peileilei424@163. com

win ratio被定义为：$R_W=TW/TL$，在临床试验中win ratio可以用来评价不同疗法的治疗效果，如果$R_W\neq1$，说明两种疗法的效果是有差异的；如果$R_W=1$，则说明两组的疗效之间并无统计差异。win ratio法的计算结果通常涉及三个统计量，win ratio，net benefit和win odds。win ratio如前文所述，可以理解为“胜率”，是指在比较两组疗效时，一组相对于另一组的获胜次数比；net benefit可以理解为“净增益”，该指标量化了一种治疗方法相对另一种治疗方法所带来的整体收益；win odds代表胜利赔率，表示一种治疗方法相对于另一种方法的胜利概率比。

2. win ratio的统计推断

win ratio法在应用时的检验假设为H_0：$R_W=1$，两组之间“winner”的数量没有差异。H_1：$R_W\neq1$，两组之间“winner”的数量确实存在差异。将两组的研究对象进行配对，在出现的所有可能的配对中，如果有一对发生了感兴趣的终点事件，则标记实验组的研究对象为“winner”或者“loser”，win ratio的值就是winner总数与loser总数的比值。统计过程采用统计软件R，使用WINS包分析win ratio等统计指标及其95%*CI*，若win ratio的95%*CI*不包含1，则拒绝原假设，可以认为两组治疗方法的疗效之间差异有统计学意义，反之不拒绝原假设，尚不能认为疗效之间差异有统计学意义。

实例分析

以Nahman等[7]进行的一项临床试验为例，该研究旨在评估肾功能不全的患者发生首次感染的时间（以月为单位），43例患者使用了传统的外科手术置入导管的方法，76例患者采用经皮置入导管的方法，定义终点事件为患者发生首次感染。随后对手术组和经皮组的患者进行随访，研究经皮置管的患者与手术置管的患者在发生感染的时间上是否有差异，经皮组与手术组的Kaplan-Meier曲线如图1所示。两条生存曲线相互交叉，表明在初期，外科手术置管相比较经皮置管对患者生存是有利的，但后期是有害的。

对该生存资料进行log-rank检验（表1），χ^2=2.53，P=0.11，表明不能拒绝原假设（P>0.05），此时无法判断两组方法优劣。

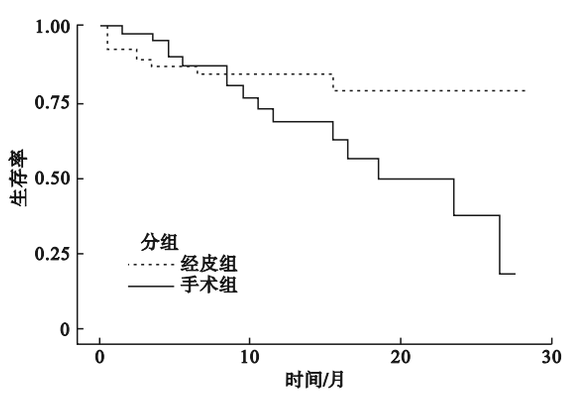


图1 经皮置管与手术置管患者的Kaplan-Meier估计

表1 利用传统生存分析方法对生存资料进行分析

统计方法	相关统计指标	数值
log-rank检验	卡方（χ^2）	2.53
	自由度	1
	显著性（P）	0.11
Cox回归分析	自由度	1
	显著性（P）	0.12
	HR	0.54
	95%*CI*	0.25~1.18

为进一步有力验证PH假设，对分组变量与依时协变量的交互项进行Cox回归分析（表2），结果显示交互项有统计学意义（P<0.05），即生存分析过程中，分组变量对结局的影响强度随着时间变化而变化，*HR*不恒定，PH假设不成立，故Cox回归模型无法应用，所得结果没有统计学意义。

表2 验证Cox回归模型的PH假设

检验方法	相关统计指标	数值
含依时协变量的Cox回归	自由度	1
	显著性（P）	0.01
	HR	0.23
	95%*CI*	0.07~0.73

因试验没有危险因素分层，终点类型是二元，为非复合终点，故win ratio的3种计算方法中，我们选择使用不匹配的win ratio法（表3），所有可能的配对一共有3268种，win ratio即R_W值为3.17，95%*CI*为1.27~7.87（P=0.01），两组疗效差异有统计学意义。分析结果表明手术组比经皮组更易发生感染，即经皮置管的方法能够降低患者发生感染的概率。除win ratio外还有两个相关统计量，win odds为1.51［95%*CI*（1.06~2.10），P=0.02］，意味着手术组的感染发生概率比经皮组高约1.51倍，net benefit为0.20［95%*CI*（0.03~0.38），P=0.03］，说明手术组比经皮组高出了20%的感染风险。

表3 利用win ratio法对生存数据进行分析

类目	终点数据	win ratio（R_w）及其95%*CI*
配对总数	3268（100%）	3.17（95%*CI*：1.27~7.87）
手术组“发生感染”winners	975（29.8%）	
经皮组“发生感染”winners	308（9.4%）	
两组平局tied	1985（60.8%）	

讨 论

log-rank检验和Cox回归模型是使用较为广泛的比较生存分析资料的统计学方法[8]。在PH假设成立的条件下，log-rank检验和Cox回归模型具有足够的

统计效力检验两组之间治疗效果的差异。但当生存曲线交叉，PH 假设不成立时，log-rank 检验检测这种差异的能力较差，因为两条曲线在通过交叉点之后，一种治疗的优势会被另一种治疗的优势所抵消[9]。当 *HR* 随时间变化的时候，Cox 模型统计效力下降，对治疗效果的估计会产生一定程度的偏差[10]。win ratio 可以在违反 PH 假设的情况下，提供一个估算治疗效果的替代统计指标，其检验疗效差异的能力优于同等情况下的 log-rank 检验和 Cox 回归模型。此外，win ratio 法还可以为心血管试验的研究设计提供实际的指导，包括如何确定试验规模等[11]。

win ratio 法在生存数据分析中有一定优势。首先，win ratio 及其 95%*CI*，包括 *P* 值的统计计算不要求 *HR* 为固定值作为前提条件。但对于 log-rank 检验和 Cox 回归模型来说，*HR* 恒定却是非常重要的基本要求，因此，这是一种无假设的方法，在某些特定情况下（如生存曲线交叉）具有较高的统计效能和灵活性。其次，win ratio 在复合终点的组成事件分析中起着非常重要的作用，与传统生存分析方法不同，win ratio 法可以通过对复合终点的组成事件赋予排序优先级来进行分析，这种差异化处理方式，可以更灵活地反映各组成部分的重要性，从而提供更全面的分析结果。在实践的研究中，win ratio 法已开始被应用于临床试验的设计中，例如心血管试验[12]、肾移植的随机试验[13]以及高血压试验[14]等等。在 *HR* 恒定时，win ratio 法与 log-rank 检验和 Cox 回归模型具有相似的功效，而当生存曲线交叉，*HR* 不恒定时，该方法在检验疗效差异方面优于传统的 log-rank 检验以及 Cox 回归模型[15]。

win ratio 法在实际应用中依然存在着局限性。首先，本文建议用于生存数据分析的 win ratio 方法是一种单变量分析方法，生成的统计量是粗略且未调整过的，如果某些重要的基线预后因素不平衡，还需估计调整后的治疗效果。另外，我们在计算独立删失情况下的 win ratio 统计量时，排除了 tied 的观测值，当删失变量依赖于协变量时，得到的 win ratio 可能是不可靠的。对此已有研究团队提出[5]，可使用逆概率删失加权法（inverse-probability-of-censoring weighting，IPCW）调整因删失数据带来的偏差，更准确地计算 win ratio。对于基线协变量或时依协变量，也有研究人员提出了针对性的 IPCW 方法，简称为 CovIPCW[16]，该方法扩展了 IPCW 调整 win ratio 的概念，不仅减小了删失数据造成的偏差，还考虑到了其他混杂因素的影响，使得 win ratio 的估算更加全面精准。尽管如此，但目前尚缺乏 win ratio 法的样本量计算和检验效能的公认方法，很多临床医生和研究人员对此方法并不熟悉，这限制了其推广和应用[17]。

随着研究人员对 win ratio 法认识的深入和应用经验的累积，预计其在国内医学研究领域将会得到更加广泛的认可，进一步促进医学研究方法的创新和发展。

参 考 文 献

[1] Gregson J, Sharples L, Stone GW, et al. Nonproportional Hazards for Time-to-Event Outcomes in Clinical Trials: JACC Review Topic of the Week [J]. J Am Coll Cardiol, 2019, 74(16): 2102-2112.

[2] Mao L. On the alternative hypotheses for the win ratio [J]. Biometrics, 2019, 75(1): 347-351.

[3] Pocock SJ, Ariti CA, Collier TJ, et al. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities [J]. European Heart Journal, 2012, 33 (2): 176-182.

[4] Wang D, Pocock S. A win ratio approach to comparing continuous non-normal outcomes in clinical trials [J]. Pharmaceutical Statistics, 2016, 15(3): 238-245.

[5] Dong G, Hoaglin DC, Qiu J, et al. The win ratio: On Interpretation and Handling of Ties [J]. Statistics in Biopharmaceutical Research, 2020, 12(1): 99-106.

[6] Gasparyan SB, Folkvaljon F, Bengtsson O, et al. Adjusted win ratio with stratification: Calculation methods and interpretation [J]. Stat Methods Med Res, 2021, 30(2): 580-611.

[7] Nahman NS Jr, Middendorf DF, Bay WH, et al. Modification of the percutaneous approach to peritoneal dialysis catheter placement under peritoneoscopic visualization clinical results in 78 patients [J]. J Am Soc Nephrol, 1992, 3(1) 103-107.

[8] Koziol JA, Jia Z, Xia X. Weighted Lin-Wang Tests for Crossing Hazards [J]. Computational and Mathematical Methods in Medicine, 2014, 2014: 643455-643457.

[9] Lin X, Xu Q. A new method for the comparison of survival distributions [J]. Pharm Stat, 2010, 9(1): 67-76.

[10] Lin RS, Leon LF. Estimation of treatment effects in weighted log-rank tests [J]. Contemporary Clinical Trials, 2017, 8: 147-155.

[11] Ratitch B, Goel N, Mallinckrodt C, et al. Defining Efficacy Estimands in Clinical Trials: Examples Illustrating ICH E9(R1) Guidelines [J]. Ther Innov Regul Sci, 2020, 54(2): 370-384.

[12] Ferreira JP, Jhund PS, Duarte K, et al. Use of the winratio in Cardiovascular Trials [J]. JACC Heart Failure, 2020, 8(6): 441-450.

[13] Fergusson NA, Ramsay T, Chassé M, et al. The win ratio approach did not alter study conclusions and may mitigate concerns regarding unequal composite end points in kidney transplant trials [J]. Journal of Clinical Epidemiology, 2018, 98:9-15.

[14] Kandzari DE, Hickey GL, Pocock SJ, et al. Prioritised endpoints for device-based hypertension trials: the win ratio methodology [J]. EuroIntervention, 2021, 16(18): e1496-e1502.

[15] Zheng S, Wang D, Qiu J, et al. A win ratio approach for comparing crossing survival curves in clinical trials [J]. Journal of Biopharmaceutical Statistics, 2023, 33(4): 488-501.

[16] Dong G, Huang B, Wang D, et al. Adjusting win statistics for dependent censoring [J]. Pharm Stat, 2021, 20(3): 440-450.

[17] Butler J, Stockbridge N, Packer M. Win ratio: A Seductive But Potentially Misleading Method for Evaluating Evidence from Clinical Trials [J]. Circulation, 2024, 149(20): 1546-1548.

（责任编辑：邓　妍）