

小样本下竞争风险限制平均损失时间的转换法研究*

任嘉翘^{1#} 黑子健^{1#} 余磊^{1#} 张成凤¹ 巫宏基¹ 江珮瑜¹ 周佳仪¹ 徐筠天¹ 侯雅文^{2△} 陈征^{1△}

【摘要】目的 当竞争风险存在时,基于限制平均损失时间(restricted mean time lost, RMTL)的方法具有较少的模型假设条件和更直观的解释性。组间效应量为 RMTL 差值(RMTL difference, RMTLd),对应假设检验基于大样本下构建,而在小样本假设下的表现效果未知。**方法** 本文探讨 RMTLd 在小样本下的表现,并发展了几种 RMTL 的变量转换法以提高此时的统计性能,且通过 Monte Carlo 模拟评价它们在不同情形下的 I 类错误和检验效能。**结果** 在小样本下, RMTLd 检验原方法存在 I 类错误膨胀的现象,而四种转换法之一的逻辑转换法能够保持较好的统计性能。**结论** 在分析小样本竞争风险数据时,推荐使用 RMTL 的逻辑转换进行统计分析。

【关键词】 限制平均损失时间 竞争风险 假设检验 小样本 逻辑转换

【中图分类号】 R195.1 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.05.010

Research on the Transformation Method of Restricted Mean Time Lost with Competing Risks Data Under Small Samples

Ren Jiaqiao, Hei Zijian, Yu Lei, et al (Department of Biostatistics, School of Public Health, Southern Medical University, Guangzhou 510515)

【Abstract】Objective When the competition risk exists, the method based on restricted mean time lost (RMTL) has fewer model assumptions and more intuitive explanation. The intergroup effect size is RMTL difference (RMTLd). The corresponding hypothesis test is based on large samples, but the performance effect under small samples is unknown. **Methods** In this study, we explore the performance of RMTLd under small samples, and develop several variable conversion methods of RMTL to improve the statistical performance, and evaluate their type I error and power in different situations by Monte-Carlo simulation. **Results** In the case of small samples, the original method of RMTLd test has the phenomenon of type I error expansion, while the logic transformation method, which is one of the four transformation methods, can maintain good statistical performance. **Conclusion** When analyzing small sample competitive risk data, it is recommended to use the logical transformation of RMTL for statistical analysis.

【Key words】 Restricted mean time lost; Competing risks; Hypothesis test; Small sample size; Logit transformation

竞争风险型生存数据大量存在于生物医学研究中,进行组间差异比较时,若只考虑单终点事件的结局而忽略竞争事件的发生可能会导致结果产生部分偏倚^[1]。存在竞争风险型生存数据的临床研究常使用 Gray 检验^[2]作为组间差异检验方法,该检验需要满足部分分布风险率成比例假设。此外,由于基线值的缺失,两组风险率之比(hazard ratio, HR)作为组间效应进行解释时含义较为抽象,解释复杂。为了解决上述问题,一些研究者^[1,3-5]提出了限制平均损失时间(restricted mean time lost, RMTL)作为竞争风险型数据分析的统计描述指标及对应统计推断方法。

RMTL 可解释为在指定时间(τ)内由于特定原因(如:死于恶性黑色素瘤)造成的平均损失时间(或者

说失去的寿命), τ 通常可以设置为两组最大观测时间的最小值^[6],或者根据临床专业需要选择特定的有意义的时间点,例如新冠临床试验通常可以选择 28 天。RMTL 可以用累积发生率曲线(cumulative incidence function, CIF)下的面积来计算,是一种直接用时间概念来进行统计描述的方法,因此,无论医生还是病人对于结果都更容易理解,同时其无需满足等比例风险假设。进行组间差异比较时,一般使用限制平均损失时间的差值(RMTLd)来量化治疗差异。在已有研究报告中,统计量 RMTLd 对应假设检验通常基于大样本假设下构建^[1],而在小样本假设下的表现效果未知。本研究在模拟验证中发现小样本下 RMTLd 检验原方法(线性转换法)的 I 类错误稍显膨胀,因此本文发展了几种 RMTL 的变量转换法^[7-9],包括对数转换、双对数转换、逻辑转换以及反正弦平方根转换对线性法进行校正,以提高统计性能。通过 Monte Carlo 模拟比较上述五种变量转换下 I 类错误与检验效能,并进行实例分析验证。

方法介绍

1. 限制平均损失时间 RMTL

在竞争风险事件中,假设 j 为事件类型($j = 1$ 或

* 基金项目:国家自然科学基金(82173622;81903411;81673268);广东省基础与应用基础研究基金(2022A1515011525);广东省大学生创新创业训练计划项目(S202212121140;S202212121142)

1. 南方医科大学公共卫生学院(广东省热带病研究重点实验室)生物统计学系(510515)

2. 暨南大学经济学院统计与数据科学系

#共同第一作者

△通信作者:陈征, E-mail: zchensmu@163.com;侯雅文, E-mail: ywzhou@163.com

2); k 表示不同的组别 ($k=1$ 或 0 , 分别表示试验组和对照组); t_i 表示在事件发生在第 i 个时刻; d_{ij} 表示在 t_i 时刻, 事件 ($j=1$ 或 2) 的事件发生数; $Y(t_i)$ 表示在 t_i 时刻风险集人数; $\hat{S}(t_{i-1})$ 表示复合事件下由 Kaplan-Meier 估计的生存率; F_j 为兴趣事件对应的累积事件发生率 CIF; F_{j-} 代表竞争事件对应的累积事件发生率 CIF; τ 代表选择的限定时间点; μ 表示为兴趣事件的 RMTL。

$$\mu = \int_0^\tau F_j(t) dt \quad (1)$$

基于非参数估计得到的兴趣事件 RMTL 估计值

$$\hat{\mu} = \int_0^\tau \hat{F}_j(t) dt = \sum_{t_i \leq \tau} (d_{ij}/Y(t_i)) \hat{S}(t_{i-1}) (t_i - t_{i-1}) \quad (2)$$

RMTL 估计值的方差:

$$\begin{aligned} Var(\hat{\mu}) = & \int_0^\tau \left\{ (\tau - t) \frac{1 - \hat{F}_{j-}(t)}{Y(t)} - \frac{1}{Y(t)} \int_t^\tau \hat{F}_j(\mu) d\mu \right\}^2 \frac{Y(t)}{\hat{S}(t)} d\hat{F}_j(t) \\ & + \int_0^\tau \left\{ (\tau - t) \frac{\hat{F}_j(t)}{Y(t)} - \frac{1}{Y(t)} \int_t^\tau \hat{F}_j(\mu) d\mu \right\}^2 \frac{Y(t)}{\hat{S}(t)} d\hat{F}_{j-}(t) \end{aligned} \quad (3)$$

假设 $\mu_k(\tau)$ 表示为第 k 组 ($k=1$ 或 0) 的兴趣事件 RMTL。RMTLd 的估计值用符号 $\hat{\Delta}$ 来表示, RMTLd 的表达式与方差如下所示:

$$\hat{\Delta} = \hat{\mu}_1(\tau) - \hat{\mu}_0(\tau) \quad (4)$$

$$Var(\hat{\Delta}) = Var(\hat{\mu}_1(\tau)) + Var(\hat{\mu}_0(\tau)) \quad (5)$$

2. 基于 RMTLd 两组疗效比较的假设检验原假设与备择假设分别是:

$$H_0: \Delta = \mu_1(\tau) - \mu_0(\tau) = 0$$

$$H_1: \Delta = \mu_1(\tau) - \mu_0(\tau) \neq 0$$

在原假设 H_0 下, RMTLd 的检验统计量为:

$$Z = \frac{\hat{\Delta}}{\sqrt{Var(\hat{\Delta})}} \quad (6)$$

在大样本下, 该统计量服从标准正态分布。

3. RMTL 的转换方法

通过 RMTL 的对数转换、双对数转换、逻辑转换以及反正弦平方根转换, 对转换后得到的 RMTL 做差得到 RMTLd。

(1) 对数转换

假设 $g(\cdot)$ 表示一种转换函数, 它可以将函数 x 转换为其他形式, 例如对数转换函数 $g_1(x) = \log(x)$, 它可以将函数 x 转化为它的对数形式。由此, 可以将 RMTL 转换为:

$$g_1(\hat{\mu}_k(\tau)/\tau) = \log(\hat{\mu}_k(\tau)/\tau) \quad (7)$$

注意, (7) 中 RMTL 除以 τ 是将转换对象的值控

制在区间 $[0, 1]$, 否则函数不满足下面的双对数转换以及逻辑转换的定义域。然后通过 delta 法得到转换后的方差为

$$Var(g_1(\hat{\mu}_k(\tau)/\tau)) = Var(\hat{\mu}_k(\tau)) / [(\hat{\mu}_k(\tau)/\tau)^2 \tau^2] \quad (8)$$

$$\hat{\Delta}_1 = g_1(\hat{\mu}_1(\tau)/\tau) - g_1(\hat{\mu}_0(\tau)/\tau)$$

$$Var(\hat{\Delta}_1) = Var(g_1(\hat{\mu}_1(\tau)/\tau)) + Var(g_1(\hat{\mu}_0(\tau)/\tau))$$

基于对数转换后的 RMTLd 的假设检验的统计量为

$$Z_1 = \frac{\hat{\Delta}_1}{\sqrt{Var(\hat{\Delta}_1)}}$$

该统计量服从标准正态分布。其它的转换法在 RMTLd 的方差估计和检验统计量的构造方法均相同。故下面只给出另外三种转换法通过 delta 法得到的转换后的方差。

(2) 双对数转换

双对数转换函数: $g_2(x) = \log\{-\log(x)\}$

$$g_2(\hat{\mu}_k(\tau)/\tau) = \log(-\log(\hat{\mu}_k(\tau)/\tau))$$

$$Var(g_2(\hat{\mu}_k(\tau)/\tau)) =$$

$$Var(\hat{\mu}_k(\tau)) / [(\hat{\mu}_k(\tau)/\tau)^2 \log(\hat{\mu}_k(\tau)/\tau)^2 \tau^2]$$

(3) 逻辑转换

逻辑转换函数: $g_3(x) = \log\{x/(1-x)\}$

$$g_3(\hat{\mu}_k(\tau)/\tau) = \log\left(\frac{\hat{\mu}_k(\tau)/\tau}{1-\hat{\mu}_k(\tau)/\tau}\right)$$

$$Var(g_3(\hat{\mu}_k(\tau)/\tau)) = Var(\hat{\mu}_k(\tau)) / [(\hat{\mu}_k(\tau)/\tau)^2 (1-\hat{\mu}_k(\tau)/\tau)^2 \tau^2]$$

(4) 反正弦平方根转换

反正弦平方根转换函数: $g_4(x) = \arcsin\sqrt{x}$

$$g_4(\hat{\mu}_k(\tau)/\tau) = \arcsin\sqrt{\hat{\mu}_k(\tau)/\tau}$$

$$Var(g_4(\hat{\mu}_k(\tau)/\tau)) = Var(\hat{\mu}_k(\tau)) /$$

$$[4(\hat{\mu}_k(\tau)/\tau)(1-\hat{\mu}_k(\tau)/\tau)\tau^2]$$

模拟研究

通过 Monte Carlo 模拟, 探究小样本下 RMTLd 的表现情况。比较线性转换、对数转换、双对数转换、逻辑转换、反正弦平方根转换这 5 种变量转换方法下得到的 RMTLd 的 I 类错误与检验效能, 并且通过改变样本的删失率, 研究 I 类错误与检验效能在不同情况下的变化, 以此说明在小样本的情况下, 使用所研究的变量转换方法的可行性与不同变量转换方法的性能比较。

通过 Monte Carlo 模拟, 比较五种变量转换方法下 RMTLd 的 I 类错误和不同情形下 (图 1) 的检验效能。这些情形分别是: a. 两组 CIF 曲线成比例,

PH ≈ 0.905; b. 两组 CIF 曲线成比例, PH ≈ 0.74; c. 衰减效应(前期开口); d. 延迟疗效(后期开口, 在时间 $t=1$ 处开口); e. 延迟疗效(后期开口, 在时间 $t=2$ 处开口)。

假设每组的竞争风险数据只有一个兴趣事件(类型一)和一个竞争事件(类型二)。事件类型的产生分别服从二项分布 $B(N, p_1)$ 和 $B(N, 1-p_1)$, 其中 N 为每组的样本量, p_1 为兴趣事件的发生率。本研究中取值兴趣事件 $p_1=0.7$, 每组的结局时间 T_j (其中 $j=1, 2$ 分别对应不同的两个组)。两组的删失时间 C 分别服从于均匀分布 $U(0, a)$ 和 $U(0, b)$ 。定义观测到的生存时间 $T = \min(T_j, C)$, 结局指标 $\delta_j = I(T_j > C)$ ($I=0$ 为删失, 1 为兴趣事件, 2 为竞争事件)。通过改变 a 和 b 的值来控制两组的删失率相同且近似等于 0%, 15%, 30%。两组的样本量设置为 (30, 30), (60, 60), (100, 100)。每一种情形的检验水平 α 设置为 0.05, 特定时间点 τ 设置为两组最大观测时间的最小值, 模拟次数设置为 50000 次。

由于该模拟研究中参数组合众多, 为了能够更清

晰地展示模拟结果以及评价上述变量转换方法的综合性能, 本文采用方差分析技术^[10]构造 I 类错误以及综合检验效能结果的评价指标。本研究共考虑 4 个影响因素, 分别是变量转换方法 (method)、样本量 (num)、删失率 (cen) 和模拟情形 (sit), 以下是使用方差分析技术拟合的 4 个模型:

模型 1: $E(Y) = method \times num + cen + sit$

模型 2: $E(Y) = method \times cen + num + sit$

模型 3: $E(Y) = method \times sit + num + cen$

模型 4: $E(Y) = method + num + cen + sit$

模型 1, 2, 3 分别研究不同样本量、不同删失和不同模拟情形下, 各变量转换方法的表现。模型 4 则是综合了所有影响因素来评价各变量转换方法的表现。其中 I 类错误的评价指标为平均偏差值, 用 I 类错误的模拟结果减去所定义的显著水平 0.05 所得, 平均偏差值越接近 0, I 类错误表现越稳定。检验效能的评价指标为平均拒绝率, 由于检验效能表示拒绝实际不成立的原假设的能力, 因此平均拒绝率越高, 检验效能越大。

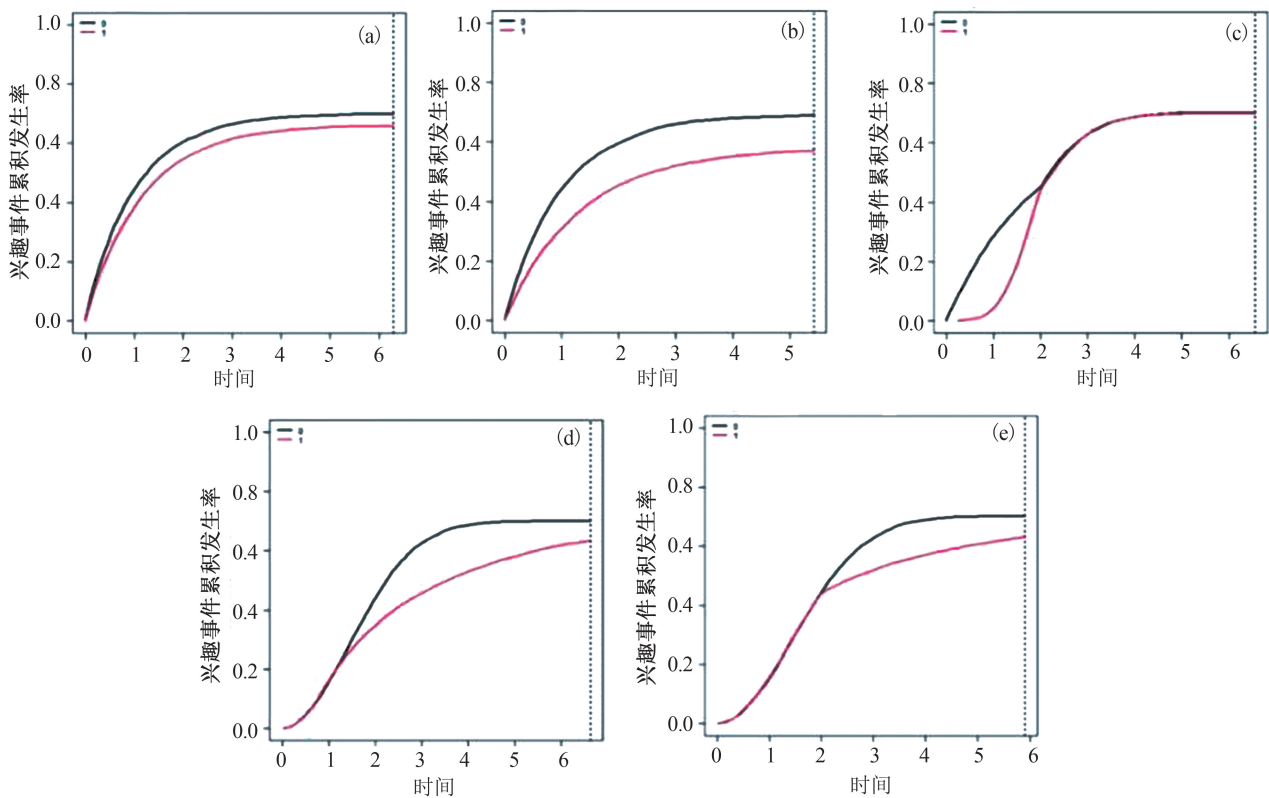


图 1 假设检验模拟下兴趣事件累积发生率图

表 1 展示的是 I 类错误(平均偏差值)的模拟结果, 平均偏差值的可接受区间为 (-0.191, 0.191)。模拟结果表明, 小样本下 ($n=30$ 或 60), 原方法(线性转换)下的 RMTLd 统计量其平均偏差值的绝对值大于 0.191, 存在 I 类错误膨胀的现象。在其余四种变量转换方法下, 除对数转换, 其余转换后均比原方法的 I 类错误

有所改善。结果显示, 在小样本下, 只有逻辑转换得到的平均偏差值全部属于可接受区间内, 双对数转换的平均偏差值除了在 30 样本下微微偏离可接受区间, 其余情况均处于可接受区间内, 而对数转换和反正弦平方根转换在小样本下的平均偏差值均处于可接受区间以外, 因此这两种变量转换不纳入检验效能的比较。

表 1 各种变量转换方法的平均偏差值(%) (I类错误)

		线性 转换	对数 转换	双对数 转换	逻辑 转换	反正弦 平方根 转换
样本量	(30,30)	0.753	-0.800	0.200	-0.003	0.423
(n_0, n_1)	(60,60)	0.387	-0.400	0.100	0.020	0.217
	(100,100)	0.273	-0.217	0.107	0.050	0.170
删失率	0%	0.480	-0.373	0.163	0.093	0.290
	15%	0.443	-0.467	0.107	0.010	0.250
	30%	0.490	-0.577	0.137	-0.037	0.270
总体		0.471	-0.472	0.136	0.022	0.270

注:此表所采用的数据兴趣事件发生率为70%。

可接受区间 $(-0.191, 0.191)$, 基于公式 $[(0.05 \pm 1.96\sqrt{0.05(1-0.05)/50000}) - 0.05] \times 100$ 算得。

表 2 是检验效能(平均拒绝率)的模拟结果。结果显示,五种方法的检验效能相差不大,结合 I 类错误和检验效能,逻辑转换方法的性能相对最稳健。

表 2 各种变量转换方法的平均拒绝率(检验效能)

		线性 转换	对数 转换	双对数 转换	逻辑 转换	反正弦 平方根 转换
样本量	(30,30)	0.149	0.130	0.141	0.139	0.146
(n_0, n_1)	(60,60)	0.231	0.218	0.225	0.224	0.228
	(100,100)	0.330	0.322	0.326	0.326	0.328
删失率	0%	0.251	0.237	0.245	0.244	0.248
	15%	0.237	0.224	0.231	0.230	0.234
	30%	0.222	0.209	0.216	0.215	0.220
模拟情形	A	0.096	0.084	0.091	0.090	0.093
	B	0.398	0.374	0.388	0.386	0.393
	C	0.296	0.307	0.292	0.300	0.299
	D	0.289	0.262	0.282	0.274	0.283
	E	0.105	0.091	0.101	0.097	0.102
总体		0.237	0.223	0.231	0.229	0.234

注:此表所采用的数据兴趣事件发生率为70%。

实例分析

一项关于恶性黑色素瘤患者手术的预后情况,共纳入了 37 名进行根治性手术的恶性黑色素瘤患者,研究分组因素为有无溃疡。其中有溃疡组为 23 人,无溃疡组 14 人。兴趣事件为患者死于恶性黑色素瘤,竞争事件为患者死于除恶性黑色素瘤的其它因素,研究结束时仍未发生终点事件的记为删失。两组的兴趣事件占比分别为 69.5% 和 42%,删失率分别为 13% 和 28.6%。限制时间点 τ 为两组最大观测时间的最小值, $\tau=2431$ (天)。表 3 为五种转换形式 RMTLd 检验的结果, P 值均小于所设定显著水平 0.05,且 P 值的大小与模拟验证部分小样本下平均偏差值的大小相对应。平均偏差值越大, I 类错误膨胀,该实例中相应转换下的 P 值越小。而基于部分分布风险率 Gray 检验

的 P 值大于 0.05。其中组间效应量 $RMTLd=-611$,可理解为对于恶性黑色素瘤患者而言,限制在随访 2431 天内,有溃疡的患者相对于无溃疡的患者的平均生存时间减少了 611 天。

表 3 黑色素瘤患者的组间比较结果

	有溃 疡组	无溃 疡组	比值/ 差值 ^b	P (统计量) ^c
SHR	/	/	2.24	0.084(2.979)
RMTL				
线性转换 ^a	332.295	943.761	-611.466	0.005(-2.764)
对数转换 ^a				0.016(-2.406)
双对数转换 ^a				0.008(2.650)
逻辑转换 ^a				0.011(-2.532)
反正弦平方根转换 ^a				0.007(-2.712)

注:SHR: subdistribution hazard ratios(部分分布风险率比);^aRMTL 的转换方式;^b比值对应 SHR,差值对应 RMTLd;^cSHR 对应 Gray 检验,括号内为 P 值对应的 χ^2 统计量,RMTL 各种转换对应转换的 RMTL 检验,括号内为 P 值对应的 Z 统计量。

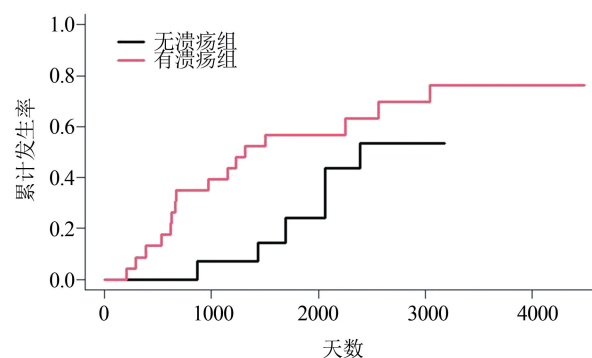


图 2 不同分组情况下黑色素瘤患者的累计发生率图

结论

组间比较是临床随访研究中的重要部分,分析竞争风险数据时,为了避免模型的假设以及增加解释上的便利,可使用组间效应量为 RMTLd 的统计方法。本研究模拟结果表明,在小样本下使用逻辑转换 RMTLd 检验能较好的控制 I 类错误,同时具有较好的检验效能,建议在实际应用中使用。

参考文献

[1] Wu H, Yuan H, Yang Z, et al. Implementation of an alternative method for assessing competing risks: restricted mean time lost[J]. American Journal of Epidemiology, 2022, 191(1): 163-172.

[2] Gray RJ. A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk[J]. The Annals of Statistics, 1988, 16(3): 1141-1154.

[3] Zhao L, Tian L, Claggett B, et al. Estimating Treatment Effect With Clinical Interpretation From a Comparative Clinical Trial With an End Point Subject to Competing Risks [J]. JAMA Cardiology, 2018, 3(4): 357-358.

(下转第 690 页)