

基于限制平均生存时间的多组疗效比较*

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

【摘要】目的 在临床随访研究中,基于时间尺度指标的限制平均生存时间(restricted mean survival time, RMST)越来越受到关注,然而目前基于 RMST 的统计推断主要用于两组比较,缺少进行两组以上比较的方法。**方法** 本文提出 RMST 多组间的假设检验法,包括经典法(naive)、对数转换法(log)、双对数转换法(cloglog)三种检验法,并通过 Monte Carlo 模拟评价其 I 类错误和检验效能,最后进行实例分析。**结果** 综合 Monte Carlo 模拟的 I 类错误及检验效能结果,显示所提出的 RMST 检验可以处理多组比较的问题,特别是 cloglog 转换法最为稳健。**结论** 针对生存数据的多组比较问题,若考虑从时间尺度指标分析,推荐使用 cloglog 转换法的 RMST 多组检验。

【关键词】 限制平均生存时间 比例风险假定 多组比较 双对数转换

【中图分类号】 C81 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.03.016

Multiple Groups Comparison Test based on Restricted Mean Survival Time

Jiang Peiyu, Zhou Jiayi, Xu Yuntian, et al (Department of Biostatistics, School of Public Health, Southern Medical University (510515), Guangzhou)

【Abstract】Objective In clinical follow-up studies, it is the most common method to quantify treatment differences between groups using a hazard ratio (HR). Moreover, restricted mean survival time (RMST) has attracted more and more attention. However, the current statistical inference method based on RMST is mainly used for the comparison between two groups. **Methods** In this paper, three RMST tests between multiple groups are proposed, including naive, logarithmic transformation and complementary logarithmic transformation. Monte Carlo simulations were performed to evaluate the type I error and power, and a case study was performed. **Results** Based on the type I errors obtained by Monte Carlo simulation and the test performance results, it is shown that the proposed RMST test can deal with the problem of multiple sets of comparisons, especially the complementary logarithmic transformation method is the most robust. **Conclusions** For the multi-group comparison of survival data, if the time scale index is considered, the RMST multi-group test by the complementary logarithmic transformation method is recommended.

【Key words】 Restricted mean survival time; Proportional hazard assumption; Comparison of multiple groups; Complementary logarithmic transformation

在临床随访研究中,风险率比(hazard ratio, HR)常被用于量化治疗差异,但其有解释抽象、比例风险假定、缺少基线风险率值等各种局限^[1-2]。因此,基于时间尺度指标的限制平均生存时间(restricted mean survival time, RMST)被越来越多的讨论和使用^[3]。目前基于 RMST 的统计推断方法主要用于两组的比较,然而在现实研究中也需要考虑多组处理组间的治疗效果差异比较,因此我们考虑构造 RMST 的多组假设检验方法。由于 RMST 理解为到某固定时间点(τ)下的生存曲线下面积值,参考吕晶晶^[4]针对多组生存率统计量^[5]的构建方法,构建 RMST 多组比较的统计量,并参考其经典法(即原始的 RMST 多组检验),log 转换法和 cloglog 转换法构造了 RMST 多组比较的统计量

转化形式。本研究通过 Monte Carlo 模拟评价其一类错误和检验效能,最后用两个实例进行分析验证。

方法介绍

RMST^[6-7]指的是生存曲线到某固定时间点 τ 下曲线下面积,即截止某个特定时间点 τ 的平均生存时间。治疗组间的 RMST 之差(the difference in RMST, RMSTd)可以用于衡量组间效果的绝对大小。

设 T 为一个非负的随机变量,表示为每个个体的生存时间,对于某个限制时间点 $\tau(\tau>0)$,使得 $Pr(T\geq\tau)>0$,设 X_i 是 T 和 τ 的最小值, $X_i=\min(T,\tau)$ 。 X 的期望即为 RMST,通常用 $\mu(\tau)$ 来表示:

$$\mu(\tau) = E(X_i) = E[\min(T, \tau)] = \int_0^{\tau} S(t) dt \quad (1)$$

其中 $S(t)$ 是生存函数; $\mu(\tau)$ 及其方差,通常可以基于非参数 Kaplan-Meier 估计法得到^[8]。

在两组 RMST 的比较中^[9],设 $\mu_g(\tau)$ 为 g 组的 RMST, $g=0,1$ (0 为对照组,1 为实验组), $RMSTd=\mu_1(\tau)-\mu_0(\tau)$ 。相应的假设检验为: $H_0:\Delta=\mu_1-\mu_0=0$, $H_1:\Delta\neq 0$ 。当 H_0 成立时,该检验统计量计算为:

* 基金项目:广东省基础与应用基础研究基金(2022A1515011525; 2024A1515011402);国家自然科学基金(82173622);广东省大学生创新创业训练计划项目(S202212121142;S202212121140)

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学科学系

#共同第一作者

△通信作者:陈征,E-mail:zchen@smu.edu.cn;侯雅文,E-mail:thouyw@jnu.edu.cn

$$s_D = \frac{\mu_1(\tau) - \mu_0(\tau)}{\sqrt{\text{Var}[\mu_1(\tau)] + \text{Var}[\mu_0(\tau)]}} \quad (2)$$

在大样本情况下,检验统计量近似服从标准正态分布。

1. RMST 多组比较检验统计量

给定 1, 2, ..., K 组, 此时的原假设为各组的 RMST 相等, 备择假设为各组的 RMST 不全相等, 即至少有一个有一组不等。即原假设 $H_0: \mu_1 = \mu_2 = \dots = \mu_K$, $H_1: \mu_1, \mu_2, \dots, \mu_K$ 不全相等。

多组检验的统计量形式为:

$$X = A \Sigma^{-1} A^T \quad (3)$$

其中, A 为向量 $[\varphi(\mu_1) - \varphi(\mu_2), \varphi(\mu_1) - \varphi(\mu_3), \dots, \varphi(\mu_1) - \varphi(\mu_K)]$, Σ 为向量 A 中的元素的协方差矩阵 $\Sigma_{(K-1) \times (K-1)}$, 矩阵主对角线上元素为 $V[\varphi(\mu_1)] + V[\varphi(\mu_p)]$, $p = 2, 3, \dots, K$, 非对角线上元素为 $V[\varphi(\mu_1)]$ 。在 H_0 成立的前提下, 检验统计量 X [公式 (3)] 服从自由度为 (K-1) 的 χ^2 分布。

小样本情况下, RMST 两组检验可能不服从正态分布, 且无法进行多组比较。本文基于固定点处生存率检验统计量的两组和多组检验方法, 针对多组 RMST 比较, 构造了三种不同的向量 A 以及对应的方差协方差矩阵 Σ , 主对角线和非主对角线元素构造不同, 最后代入公式 (3) 得到最终的检验统计量。

2. 检验统计量的三种方法

(1) 经典法 (naive)

naive 法中, 向量 A 中的 φ 变换为 $\varphi(\mu_p) = \mu_p$, 则 $A = (\mu_1 - \mu_2, \mu_1 - \mu_3, \dots, \mu_1 - \mu_K)$, Σ 矩阵为主对角线上元素为 $V(\mu_1) + V(\mu_p)$, $p = 2, 3, \dots, K$, 非对角线上元素为 $V(\mu_1)$ 的协方差矩阵。

在 H_0 成立的前提下, 最终检验统计量服从自由度为 (K-1) 的 χ^2 分布, 下列检验法也全部服从自由度为 (K-1) 的 χ^2 分布。

(2) 对数转换法 (log)

log 转换法中, 向量 A 中的 φ 变换为 $\varphi(\mu_p) = \log \mu_p$, Σ 矩阵的主对角线上元素为 $V(\log \mu_1) + V(\log \mu_p)$, 非对角线上元素为 $V(\log \mu_1)$, 其中 $V(\log \mu_p) = \frac{1}{\mu_p^2} \times \text{Var}(\mu_p)$ (参考附录推导)。

(3) 双对数转换法 (cloglog)

cloglog 转换法中, 由于互补双对数定义域的特殊性, 我们对 μ_p 做一个处理, 令 $x = \frac{\mu_p}{\tau}$ (具体公式推导见文末附录), 则向量 A 变换为 $\varphi\left(\frac{\mu_p}{\tau}\right) = \log\left(-\log\left(\frac{\mu_p}{\tau}\right)\right)$, Σ 矩阵的主对角线上元素为 $V\left[\log\left(-\log\left(\frac{\mu_1}{\tau}\right)\right)\right] + V$

$\left[\log\left(-\log\left(\frac{\mu_p}{\tau}\right)\right)\right]$, 非对角线上元素为 $V\left[\log\left(-\log\left(\frac{\mu_1}{\tau}\right)\right)\right]$ 。其中 $V\left[\log\left(-\log\left(\frac{\mu_p}{\tau}\right)\right)\right] = \frac{1}{\left(\frac{\mu_p}{\tau} \times \left(-\log\left(\frac{\mu_p}{\tau}\right)\right)\right)^2} \times \text{Var}\left(\frac{\mu_p}{\tau}\right)$ (参考附录推导)。

模拟研究

采用 Monte-Carlo 模拟研究三种 RMST 检验法的 I 类错误和检验效能, 检验组数为 3。

在 I 类错误方面, 共有四种删失情况, 生存时间均由参数为 $\lambda = 0.2$ 的指数分布产生, 特定时间点选为三组最长生存时间的最小值。在检验效能方面, 为控制生存曲线开口大小, 当比例风险假定成立时, 令三组的生存时间分别服从参数为 $\lambda^T \sim (0693, 0.462, 0.315)$ 的指数分布 (见图 1A); 当比例风险假定不成立时, 假设生存曲线为前期开口, 令三组的生存时间在前期 ($t \leq 2$) 时, 服从形状参数 $k \sim (2, 4, 10)$, 尺度参数 $\lambda = 2$ 的 weibull 分布; 在后期 ($t > 2$) 时, 服从形状参数 $k = 2$, 尺度参数 $\lambda = 2$ 的 weibull 分布 (见图 1B)。

I 类错误和检验效能的删失时间均由服从 $U(0, C)$ 的均匀分布产生, 共有四种删失情况, 平均删失率约为 0%、15%、30%、45%。最后获得的个体生存时间为 $t \sim \min(T, C)$ 。当 $t = T$ 时, 即个体在随访时间没有发生删失, 也就是发生终点事件, 设生存状态 status 为 1, 否则为 0。

考虑样本量均衡 (n_1, n_2, n_3 均为 25, 50, 100, 250, 500) 和不均衡 [(n_1, n_2, n_3) 为 (25, 25, 50)、(25, 50, 50)、(25, 50, 100)、(50, 50, 100)、(50, 100, 150)、(100, 250, 250)、(100, 250, 500)、(250, 250, 500)、(250, 500, 500)] 的情形, 每种参数组合下模拟 10000 次, 显著水平 $\alpha = 0.05$ 。

由于本模拟研究中样本量、删失率组合较多, 为了使结果更清晰的展示, 本文采用了方差分析 (ANOVA) 方法^[10-11]。其中, I 类错误的评价指标为平均偏差值, 其值越接近 0, 检验方法表现越好; 检验效能的评价指标为平均拒绝率, 其值越大, 说明检验方法表现越好。

模拟结果

表 1 显示三种 RMST 检验方法的 I 类错误均随着删失率的增加而增加。其平均偏差值的可接受区间为 (-0.427, 0.427), 三种方法的整体平均偏差值均在可接受区间范围内, 其中 cloglog 转换法的平均偏差值较小, 除了在样本量均为 25 时偏离可接受区间, 其余

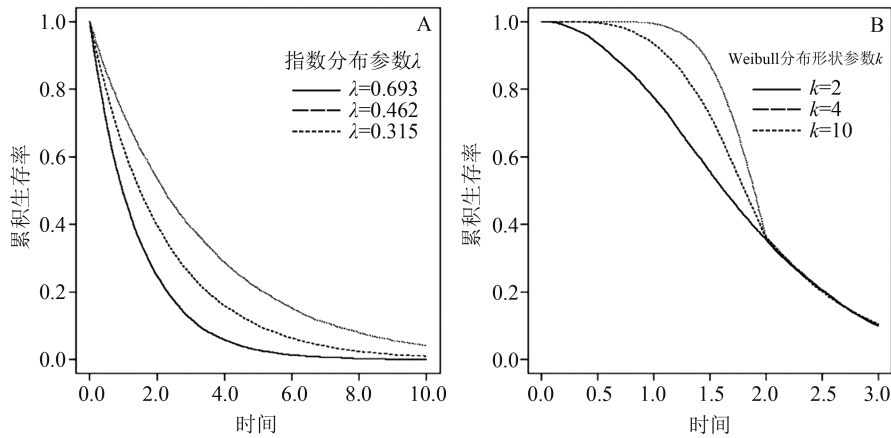


图 1 模拟情形图

情况处于可接受区间内。在删失率为 45% 时, naive 法和 log 转换法的平均偏差值较大。综合结果看, cloglog 转换法的平均偏差值最小, 意味着其 I 类错误最接近 0.05。

表 2 展现了检验效能的结果, 随着样本量的增加, 各检验法得到的检验效能均呈上升趋势; 随着删失率的减少, RMST 检验及其转换法的检验效能均呈上升趋势。在不同情形下, 三种检验法的检验效能都表现良好。根据综合结果, cloglog 转换法的检验效能最高, naive 法次之。

表 1 三种检验方法的 I 类错误(平均偏差值)

	TEST	naive	log	cloglog
样本量	(25, 25, 25)	0.5075	-0.2875	-0.7025
	(50, 50, 50)	0.5150	0.3575	-0.0675
	(100, 100, 100)	0.1175	0.0350	-0.2375
	(250, 250, 250)	0.0925	0.1475	0.0050
	(500, 500, 500)	0.1425	0.1975	0.0950
	(25, 25, 50)	0.6650	0.1700	-0.4275
	(25, 50, 50)	0.7125	0.4700	-0.0550
	(25, 50, 100)	0.5875	0.5075	-0.0250
	(50, 50, 100)	0.1025	0.0200	-0.2550
	(50, 100, 150)	0.3250	0.3450	0.0050
	(100, 250, 250)	-0.0175	-0.0600	-0.1775
	(100, 250, 500)	-0.0050	-0.0075	-0.1525
	(250, 250, 500)	0.2075	0.2725	0.1750
	(250, 500, 500)	0.2800	0.3275	0.2575
	删失率	0%	0.0036	-0.0179
15%		0.0486	0.0293	-0.2321
30%		0.3771	0.2400	-0.0736
45%		0.7800	0.4614	0.0050
综合结果		0.3023	0.1782	-0.1116

注: 平均偏差值的可接受区间为(-0.427, 0.427)

表 2 三种检验方法的检验效能(平均拒绝率)

	TEST	naive	log	cloglog
样本量	(25, 25, 25)	0.4571	0.4219	0.4369
	(50, 50, 50)	0.7048	0.6850	0.7017
	(100, 100, 100)	0.9038	0.8955	0.9048
	(250, 250, 250)	0.9967	0.9963	0.9967
	(500, 500, 500)	1.0000	1.0000	1.0000
	(25, 25, 50)	0.5575	0.4955	0.5641
	(25, 50, 50)	0.5587	0.5086	0.5580
	(25, 50, 100)	0.6580	0.5965	0.6711
	(50, 50, 100)	0.7943	0.7656	0.8015
	(50, 100, 150)	0.8406	0.8189	0.8473
	(100, 250, 250)	0.9625	0.9584	0.9638
	(100, 250, 500)	0.9796	0.9763	0.9808
	(250, 250, 500)	0.9990	0.9988	0.9990
	(250, 500, 500)	0.9992	0.9991	0.9992
	删失率	0%	0.8463	0.8283
15%		0.7979	0.7793	0.7986
30%		0.7969	0.7742	0.7977
45%		0.7837	0.7585	0.7852
情形 II: 比例风险假设成立		0.8890	0.8718	0.8816
	情形 III: 比例风险假设失效	0.7412	0.7163	0.7505
综合结果		0.8151	0.7940	0.8161

综合 I 类错误和检验效能, 三种 RMST 检验法性能都不错, 特别是 cloglog 转换法既有接近 0.05 的 I 类错误, 又有较高的检验效能, 综合来看 cloglog 转换法是三种里面最为稳健的。

实例分析

本文提供两个实例分析, 用于进行 RMST 的多组检验法的验证。

【例 1】一个关于白血病异体骨髓移植的例子^[12], 共纳入 137 名患者。从患者成功移植骨髓后进行随

访,将死亡和复发作为复合终点事件,其余定义为删失。根据骨髓移植时的疾病状态进行分组,分别为:急性淋巴细胞白血病(acute lymphoblastic leukemia, ALL),低危急性骨髓性白血病(acute myeloid leukemia-low risk, AML-Low Risk),高危急性骨髓性白血病(AML-High Risk)三组。这三组满足成比例风险假定($P=0.14$)。图2为三组的生存曲线图。

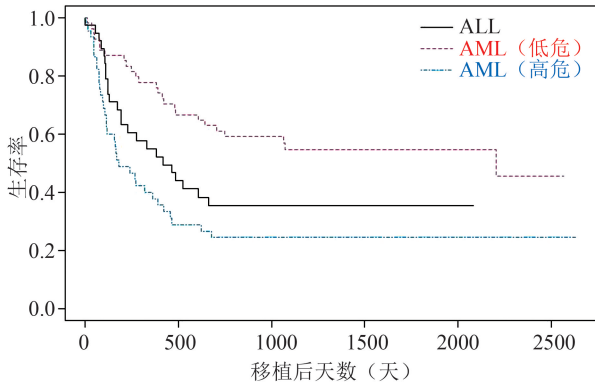


图2 例1的生存曲线图

将三组最长随访时间的最小值设为限定时间点^[13],经计算可得限制时间点为2081天,即 $\tau=2081$ 天。设原假设为三组RMST相等,备择假设为三组RMST不全相等。经计算,三种方法的结论一致($P<0.05$),三组的治疗效果存在统计学差异(表3)。

表3 例1分析结果

检验方法	P值(χ^2)
RMST(naive)检验	<0.0010(15.20)
RMST(log)检验	0.0015(13.07)
RMST(cloglog)检验	0.0014(13.15)

【例2】三种疗法治疗66例白血病患者数据^[14]。该数据是探究三种疗法对白血病的治疗效果有无差异,终点事件是疾病恶化,生存时间即术后到疾病进展之间的缓解期时间。根据疗法不同分为3组,分别为:A疗法组、B疗法组、C疗法组,共66名患者纳入研究。从图3中可以看出,三组的生存曲线存在明显交叉,这三组不满足比例风险假定($P=0.035$)。

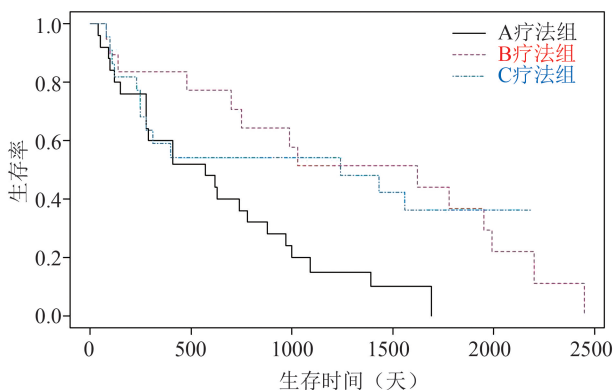


图3 例2的生存曲线图

将三组最长随访时间的最小值设为限定时间点,经计算可得限制时间点为169天,即 $\tau=169$ 天。设原假设为三组RMST相等,备择假设为三组RMST不全相等。经计算,三种RMST检验结果一致(表4)。

表4 例2分析结果

检验方法	P值(χ^2)
RMST(naive)检验	0.0155(8.333)
RMST(log)检验	0.0208(7.745)
RMST(cloglog)检验	0.0279(7.156)

从检验结果可以看出,RMST三种检验法的检验结果均认为三种疗法之间的差异具有统计学意义($P<0.05$),结合模拟研究结果,因此可认为A、B、C三种疗法对白血病的影响存在统计学差异。

讨论

在临床随访过程中,常聚焦于两组间比较研究^[15],而多组间疗效差异的比较也是重要的研究内容,已有大量研究显示当比例风险假定不成立时,log-rank检验的结果并不可靠^[16-17]。相比log-rank检验、HR指标等风险率体系方法,基于RMST的组间效应量RMST差值有更好解释性和较高的检验效能^[13],对应的RMST假设检验也没有比例风险假定成立的应用前提条件。所以本文主要是基于RMST的多组检验方法,通过模拟和实例研究可以得出所提方法,特别是cloglog检验法,具有较好的稳健性。

本研究存在一定的局限性。通过模拟发现, RMST多组检验法在样本量较少时,由于事件数较少,有效信息也相应减少,检验效能也较低,在实际应用中建议结合相关专业背景综合考虑检验结果。同时,特定时间点 τ 的选择,对于RMST的估计存在直接的影响,不同的取值可能导致不同的统计量和假设检验结果,因此, τ 值的选取仍是待解决的问题。

参考文献

- [1] Conner SC, Sullivan LM, Benjamin EJ, et al. Adjusted restricted mean survival times in observational studies. *Statistics in Medicine*, 2019, 38(20): 3832-3860.
- [2] Zhang C, Huang B, Wu H, et al. Restricted mean survival time regression model with time-dependent covariates. *Statistics in Medicine*, 2022, 41(21): 4081-4090.
- [3] Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Medical Research Methodology*, 2013, 13(1): 152.
- [4] 吕晶晶, 陈金宝, 吴天敏, 等. 固定点处k组生存率比较的非参数统计方法研究. *中国卫生统计*, 2017, 34(4): 538-541.
- [5] Klein JP, Logan B, Harhoff M, et al. Analyzing survival curves at a fixed point in time. *Statistics in Medicine*, 2007, 26(24): 4505-4519.
- [6] Irwin JO. The standard error of an estimate of expectation of life,

with special reference to expectation of tumourless life in experiments with. *Journal of Hygiene*, 1949, 47(2): 188-189.

[7] 杨紫荆, 吕晶晶, 侯雅文, 等. 限制平均生存时间在临床随访研究中的应用. *中华流行病学杂志*, 2019, 40(2): 247-250.

[8] Royston P, Parmar MK. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Statistics in Medicine*, 2011, 30(19): 2409-2421.

[9] Hasegawa T, Misawa S, Nakagawa S, et al. Restricted mean survival time as a summary measure of time-to-event outcome. *Pharmaceutical Statistics*, 2020, 19(4): 436-453.

[10] 黄兴辉. 删失数据中组间疗效比较的统计推断方法研究. 南方医科大学, 2019.

[11] 黄兴辉, 陈金宝, 杨紫荆, 等. 基于两条生存曲线间面积的非参数统计推断方法研究. *中国卫生统计*, 2019, 36(1): 8-12.

[12] Klein JP, Moeschberger ML. *Survival Analysis: Techniques for Censored and Truncated Data*. Germany: Springer, 1997: 3-6.

[13] 吕晶晶, 巫宏基, 侯雅文, 等. 竞争风险存在时基于限制平均损失时间的统计分析. *中国卫生统计*, 2021, 38(5): 661-664.

[14] 陈平雁, 黄浙明主编. *IBM SPSS 19 统计软件应用教程*. 第 2 版. 北京: 人民卫生出版社, 2012: 310-311.

[15] Huang B, Kuan PF. Comparison of the restricted mean survival time with the hazard ratio in superiority trials with a time-to-event end point. *Pharmaceutical Statistics*, 2018, 17(3): 202-213.

[16] Huang X, Lyu J, Hou Y, et al. A nonparametric statistical method for two crossing survival curves. *Communications in Statistics - Simulation and Computation*, 2022, 51(9): 5041-5050.

(上接第 408 页)

[14] 袁丽霞, 孟瑞琳, 李佳莉, 等. 热浪与广东省居民脑卒中寿命损失年的关系: 基于贝叶斯时空模型. *环境与职业医学*, 2022, 3(39): 268-274.

[15] 蔡玥, 周脉耕, 李小洪, 等. 2013 年中国居民预期寿命和去死因预期寿命分析. *中华流行病学杂志*, 2017, 38(8): 1001-1004.

[16] Yang J, Yin P, Sun JM, et al. Heatwave and mortality in 31 major Chinese cities: Definition, vulnerability and implications. *Science of The Total Environment*, 2019, 649: 695-702.

[17] Gronlund CJ, Zanobetti A, Schwartz JD, et al. Heat, Heat Waves, and Hospital Admissions among the Elderly in the United States, 1992-2006. *Environmental Health Perspectives*, 2014, 122(11): 1187-1192.

[18] Huang CR, Jian C, Dung P, et al. Mortality burden attributable to

[17] Uno H, Claggett B, Tian L, et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *Journal of Clinical Oncology*, 2014, 32(22): 2380-2385.

附录: RMST 两种转换法的方差公式推导

将 RMST 看作随机变量 $\mu(\tau)$, φ 是 $\mu(\tau)$ 的转换函数, 即 $\varphi = \varphi(\mu(\tau))$ 。已知 $\mu(\tau)$ 的期望, 方差为 $Var[\mu(\tau)] = \sum_{j=1}^D [\sum_{i=j}^D (t_{i+1} - t_i) S(t_i)]^2 \frac{d_j}{Y_j(Y_j - d_j)}$ 。由 delta 法得, 当 $\mu(\tau) = \mu$ 时,

$$\varphi = \varphi(\mu(\tau)) \approx \varphi(\mu) + \varphi(\mu)'(\mu(\tau) - \mu)$$

则:

$$Var(\varphi) = Var[\varphi(\mu) + \varphi(\mu)'(\mu(\tau) - \mu)] = [\varphi(\mu)']^2 Var(\mu(\tau))$$

由此可得, 当 $\varphi(\mu(\tau)) = \log(\mu(\tau))$ 时,

$$Var(\varphi(\mu(\tau))) = [\log(\mu(\tau))']^2 Var(\mu(\tau)) = \frac{1}{\mu(\tau)^2} Var(\mu(\tau))$$

由于互补双对数定义域的特殊性, 这里对 RMST 进行一个处理, 令 $x = \frac{\mu(\tau)}{\tau}$, 当 $\varphi(x) = \log(-\log(x))$ 时,

$$Var(\varphi(x)) = \frac{1}{\left(\frac{\mu(\tau)}{\tau}(-\log(\frac{\mu(\tau)}{\tau}))\right)^2} Var\left(\frac{\mu(\tau)}{\tau}\right)$$

(责任编辑: 邓妍)

heatwaves in Thailand: A systematic assessment incorporating evidence-based lag structure. *Environment International*, 2018; 121 (pt1): 41-50.

[19] 谷少华, 金永高, 陆蓓蓓, 等. 2013-2018 年宁波市高温热浪致超额死亡风险评价. *预防医学*, 2021, 9(33): 897-901.

[20] Xu ZW, Jian C, Hu WB, et al. Heatwave and health events: A systematic evaluation of different temperature indicators, heatwave intensities and durations. *Science of The Total Environment*, 2018, 630: 679-689.

[21] Aboubakri O, Khanjani N, Jahani Y, et al. The impact of heat waves on mortality and years of life lost in a dry region of Iran (Kerman) during 2005-2017. *Int J Biometeorol*, 2019, 63(9): 1139-1149.

(责任编辑: 郭海强)