在以生存结局为终点的临床试验中，期中分析时的生存数据成熟度如何?

Marianne A Jonker *健康证据系，生物统计学科，Radboudumc，奈梅亨，荷兰；Steven Teerenstra 健康证据系，生物统计学科，Radboudumc，奈梅亨，荷兰

arXiv:2305.04103v1 [stat.ME] 6 May 2023

2023年5月9日

摘要

在以生存结局为研究终点的临床试验中，通常会进行期中分析，以便在疗效显著时提前停止试验。如果期中分析发生在试验早期，研究者可能会得出新疗法比其他治疗方式（例如安慰剂）更有效的结论，并据此停止试验。然而，对于正在研究的研究问题而言，试验组的生存曲线可能尚未成熟，例如，此时的生存曲线仍接近于1。这意味着，该决策仅基于长期来看所发生事件中的一小部分；这些事件可能仅来自试验中体弱程度较高的患者，而这些患者可能无法代表整个患者群体。仅凭如此有限的信息得出有关有效性的结论，可能并不合理。因此，应谨慎选择确定期中分析实施时机的标准，并将期中分析时数据的成熟度纳入考虑。在这一过程中，期中分析时预期的生存率也发挥着重要作用。本文将推导在单组和双组临床试验中，期中分析实际实施时（随机时点）Kaplan-Meier 曲线的渐近分布。基于这一分布，本文将推导一个区间，使得Kaplan-Meier曲线以95%的概率落在该区间内。该区间可用于在试验设计阶段（即试验开始之前）规划期中分析的实施时机。

关键词：研究设计、期中分析、至事件时间终点、总生存期（OS）、无进展生存期（PFS）

1 引言

假设我们计划开展一项双臂临床试验，设A组A 和B组B，结局为生存（时间至事件）。例如，A组A 的癌症患者接受新疗法，B组B 的患者接受安慰剂或“常规治疗”。通过比较两组的总生存期或无进展生存期来研究新疗法的效应。我们考虑旨在进行确证性检验（例如采用对数秩检验）的试验。三期临床试验就是一个例子，因为其重点是确证疗效（和安全性）。疗效检验可能需要大量患者或较长时间的随访，因此通常会计划进行期中分析，以便因疗效而提前停止

*通讯作者：marianne.jonker@radboudumc.nl

试验。期中分析的时间根据一项或多项标准确定，这些标准应在试验开始前充分明确。例如，分析时间可以根据入组患者数或事件数确定（Floriani et al., 2008）。对于基于事件的期中分析，通常采用组序贯方法（Jennison and Turnbull, 2000），这需要选择一个α消耗函数（例如O'Brien-Fleming型或Pocock型）以及一个信息分数；它们在设计中所起的作用，类似于仅进行一次检验的设计在效能计算中显著性水平和样本量所起的作用（Lan et al., 1994）。

某一时点可获得的统计信息量取决于观察到的事件数，或该时点的信息分数；后者定义为观察到的事件数除以最终分析计划的事件总数。信息分数达到预先指定的值时，即进行期中分析。该值可自由选择，并应在设计阶段确定。然而，信息分数只决定对数秩/Cox 回归检验的效能，而期中分析时的生存数据不一定在 Kaplan-Meier 曲线方面达到成熟。这里的"数据成熟度"是指欧洲药品管理局（EMA）肿瘤学指南中的含义："事件在时间上的分布（早期–晚期）使得能够估计整个研究人群中的治疗效应。"例如，在一项 1:1 随机对照试验中，若要以双侧显著性水平 0.05 和 80% 的效能检出 0.63 的风险比，则需要 147 个事件（Schoenfeld, 1983）。期中分析可安排在信息分数为 68% 时进行，即出现 100 个事件时。如果试验有 200 名患者，那么与有 1000 名患者的试验相比，每组中观察到的 Kaplan-Meier 曲线会多得多。这一点在图 1 中可以清楚地看到：图中绘出了 200 名患者的试验（左）和 1000 名患者的试验（右）在期中分析时的 Kaplan-Meier 曲线（基于模拟数据）。更一般地说，如果期中分析时发生事件的患者 p 比例较低，

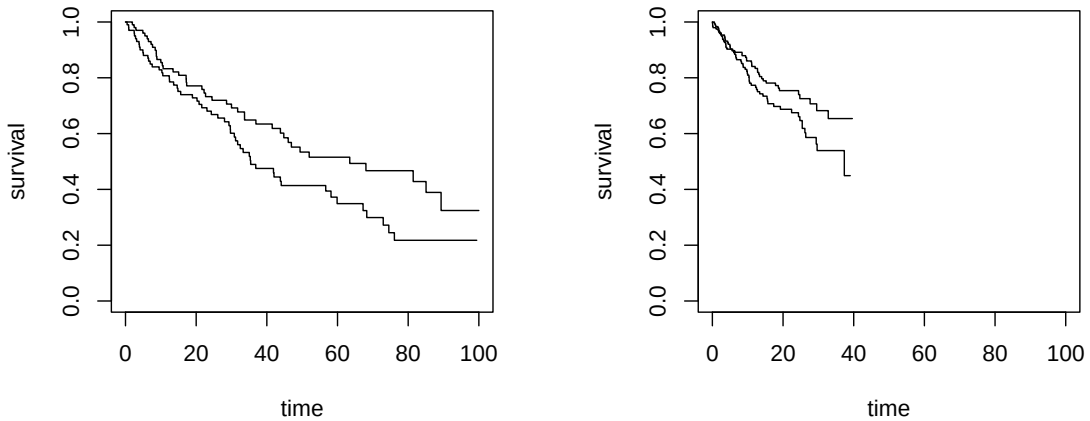


图 1：一项包含 200 名患者（左）和 1000 名患者（右）的试验在期中分析时（观察到 100 个事件后）的 Kaplan-Meier 曲线。两项试验中，组A 和 B 的真实生存分布均为均值分别为 51.9 和 82.4 的指数分布。

在期中分析时，事件数可能足以使统计检验达到统计学显著，并得出新疗法比安慰剂更有效的结论，而生存曲线仍接近 100%，例如均高于 80%。在这种情况下，统计学显著差异的结论仅基于治疗开始后早期死亡的患者；这些患者可能预后不良。因此，依据期中分析的结论来宣称治疗对整个人群有效，并不合理。因此，在规划期中分析的时间时，重要的是大致了解该时点的生存情况，以免期中分析在"生存曲线上"进行得过早。这促使我们提出本文的核心问题："我们能否推导出一个预测区间公式，用于估计期中分析时的生存率，并且可在试验开始前使用？"这并非一个简单的问题，因为期中分析的时间点是日历时间中的一个随机时点，而且在期中分析时点之前，可能并非所有患者都已入组研究。

为探究　　为了了解这一问题的现有研究情况，我们在 Pubmed 中进行了标题-摘要检索。由于♀♀♀♀

没有任何词语可以直接检索期中分析的时间安排，因此我们决定采用广泛的检索策略，在 Biometrical Journal、Biomet- rics、Biometrika、《临床试验》、《当代临床试验》、《对照临床试验》、《美国统计学会会刊》、《生存数据分析》、《药物统计学》、《医学研究中的统计方法》和《医学统计学》中检索“interim analysis”“interim analyses”“interim look”或“interim looks”（截至 2022 年 4 月 22 日共检索到 458 条结果）。检索到的大多数文献都探讨了特定情境下期中分析检验的临界值（例如，如果假设涉及固定随访时间的生存概率（Lin et al., 1996）、贝叶斯自适应设计、估计与偏倚校正、治疗组选择、最优设计以及样本量调整）。我们也找到了一些探讨与期中分析时间安排相关问题的论文。例如：基于事件数的期中分析在日历时间上的发生时点（Bagiella and Heitjan, 2001），或者反过来，如何在固定（非随机）日历时间更新对事件数的预测，从而更新对效能的预测（Royston and Barthel, 2010）；如何将日历时间尺度转换为信息时间尺度，即事件比例（Lan and Lachin, 1990）；或者估计给定日历时间的百分位数（医疗器械安全性（Murray et al., 2013））。我们发现了一条具体建议，见 Van Houwelingen 等人（2005）：他们建议，在非比例风险的情况下，只有在明确指定最终分析的最终时间范围，且期中分析时截至该时间范围的整个时间区间内均有足够信息时，才计划进行期中分析。不过，他们的论据和结果并未涉及期中分析时的 Kaplan-Meier估计。总之，看来文献中尚未探讨我们提出的问题。

因此，本文将考虑以下情形。假设在p100%的患者发生事件后进行期中分析。我们将推导出一些区间，使其以较高概率包含期中分析时的Kaplan-Meier曲线。这些区间的边界取决于对两组生存曲线预期形态的假设、入组率以及发生事件的患者比例（p）。虽然可以通过开放更多研究中心来影响入组率，但它通常是既定的后勤限制。因此，在设计阶段，研究者只能自由选择比例p 。在试验设计阶段了解p 与预测区间边界之间的关系后，便可选择p 的取值，以确保期中分析时的Kaplan-Meier曲线远低于100%。预测区间的边界是根据期中分析进行时这一随机时间点上Kaplan-Meier曲线的渐近分布推导得出的。

本文结构如下。第2节通过用数学术语描述研究目标，使其具体化。为此，本文将引入相关符号。接下来，第3节给出双臂和单臂研究中，期中分析时 Kaplan-Meier 曲线的渐近分布。这些定理的证明见附录。然后，第4节介绍一系列设定下模拟研究的结果，以确认渐近理论可用于有限样本。结果以表格形式列出，以便直接将理论用于规划；附录中还提供了 R代码。此外，第5节阐述如何在规划试验时将渐近理论应用于实践。本文最后在第6节进行讨论。

2 符号与具体目标

符号

我们考虑一项具有生存结局并进行期中分析的双臂临床试验（A 组和B组）。有两条时间线很重要：随访时间和日历时间。在生存分析中，研究人员通常关注随访时间线，因为它描述了从预先指定的起始点（例如开始治疗）到发生关注事件的时间。日历时间线从试验开始（时间零）持续到试验结束。这条时间线很重要，因为期中分析的实施时点以日历时间定义。在期中分析时估计生存曲线，只能使用该时点已有的信息。

在临床试验中，通过比较A 组和B组的生存结局（例如总生存期），来研究一种新治疗的效应。为在观测值及其

分布函数的记号中区分两组，使用下标 A 或 B。对于 A 组的一名患者，将 T_A 和 C_A 分别定义为从进入研究到发生关注事件（即生存时间）以及删失的时间，二者均以随访时间计。假设 T_A 和 C_A 在随机意义上相互独立。用 F_A表示生存时间 T_A 的分布，其连续密度为 f_A，生存函数为 $S_A = 1 - F_A$。T_A的风险函数定义为 $\lambda_A(t) = f_A(t)/(1 - F_A(t)) = f_A(t)/S_A(t)$，相应的累积风险函数定义为 $\Lambda_A(t) = \int_0^t \lambda_A(s)ds$。B 组也采用类似的记号。以日历时间计，研究开始时刻为零。令 E_A 表示 A 组患者进入研究的时间（自研究开始起算），其分布函数为 G_{Acc}。以日历时间计，关注事件发生在 $E_A + T_A$时（即自研究开始起算的时间），或者患者在 $E_A + C_A$ 时被删失，取两者中较早者。根据定义，$E_A + C_A = L$，其中 L 为研究临时暂停（进行期中分析时）或正式结束的时刻（日历时间）（L 将在后文说明）。在本文中及下文中，我们会交替使用“临时暂停”和“期中分析时点”等说法，因为从分析角度来看，期中分析之后招募的患者或发生的事件不会影响分析。从试验实施的角度来看，招募和随访通常会继续进行（除非存在需要处理的安全性问题）。

A组和B组中的患者人数分别为n和m。若$E_{A,i} < L$，则A组患者i的观测值为$(E_{A,i} \wedge T_{A,i}, C\Delta_{A,i,A,i})$，其中 $T_{A,i} \wedge C_{A,i} = \min\{T_{A,i}, C_{A,i}\}$，且$\Delta_{A,i} = 1\{T_{A_i} \leq C_{A,i}\}$ 在$T_{A,i} \leq C_{A,i}$ 时等于1，否则等于0。若$E_{A,i} > L$，则患者i在研究于L时（暂时）停止之前尚未进入研究，因此没有可用的观测值。假设不同患者的观测值相互独立。B组采用类似的记号。

具体目标

如引言所示，期中分析时的 Kaplan-Meier 曲线并不取决于（统计）信息分数，而是取决于发生事件的患者比例。因此，假设在 100p% 的患者发生事件后进行期中分析，而不论这些患者来自 A组还是 B组。比例 p 是事件数（通常根据效能考量选定）和患者总数（通常根据后勤可行性选定）的直接结果。请注意，事件数和患者数是在试验的设计阶段选定的，因此可以在设计阶段对其进行调整，使 p 达到一个能让期中分析时的数据成熟度具有实际意义的值。期中分析的随机时刻记为 $\hat{t}_{p,n+m}$，它是日历时间上的一个时刻。在期中分析时，只能使用截至该时刻的观测值进行估计。具体而言，只有在期中分析进行时刻之前进入研究的患者，才会纳入分析。图2显示了六名在期中分析之前进入研究的患者的随访时间。左图以日历时间表示事件。右图显示了这些患者的随访时间。同时也标出了时间点 $\hat{t}_{p,n+m}$。从图中可以立即看出，几乎肯定没有患者的随访时间至少达到 $\hat{t}_{p,n+m}$。这意味着，即使样本量很大，在（随访时间上的）随机时刻 $\hat{t}_{p,n+m}$ 对生存曲线 S_A 和 S_B 进行的估计仍不可靠。因此，研究的目的是考察在 $\delta > 0$ 为预先指定值时，两个组在时间点 $\hat{t}_{p,n+m} - \delta$的 Kaplan-Meier曲线。设计研究时可以考虑 δ 的不同取值。如果样本量 n 和 m 趋于无穷大，$\hat{t}_{p,n+m}$ 依概率收敛于 t_p，该值定义为某个混合分布（稍后将予以定义）的第p分位数；在 n 和 m 较大但有限时，假设该混合分布至少在 t_p的某个邻域内严格递增，则会有正比例的患者其随访时间大于 $\hat{t}_{p,n+m} - \delta$。

将在第3节（及附录）中证明，$\sqrt{n}(\tilde{S}_{A,n}(\hat{t}_{p,n+m} - \delta) - S_A(t_p - \delta))$ 渐近正态，其均值为零，方差为 $\sigma^2_{A,\delta}$，其中 $\tilde{S}_{A,n}$是基于截至时间 $\hat{t}_{p,n+m}$进行的期中分析观测值所得到的 A组 Kaplan-Meier估计量。渐近方差 $\sigma^2_{A,\delta}$是生存函数 S_A、S_B、G_{Acc}（G_{Acc}的定义见第2节开头）以及比例 p 的函数（$\sigma_{A,\delta}$的表达式见第3节）。由这一渐近行为可知，

$$P\left(S_A(t_p - \delta) - \xi_{\alpha/2}\frac{\sigma_{A,\delta}}{\sqrt{n}} \leq \tilde{S}_{A,n}(\hat{t}_{p,n+m} - \delta) \leq S_A(t_p - \delta) + \xi_{\alpha/2}\frac{\sigma_{A,\delta}}{\sqrt{n}}\right) \approx 1 - \alpha,$$

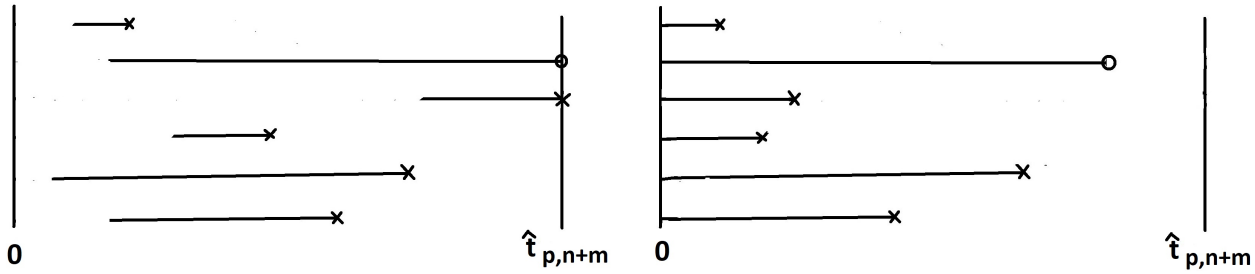


图2：六名患者在期中分析前的随机时刻进入研究。左图：六名患者的日历时间。右图：相同六名患者的随访时间。关注事件用叉号表示，删失用圆圈表示。尽管部分患者在期中分析时仍处于风险中（也就是说，处于日历时间尺度上），但他们的随访时间并未达到 $\hat{}t_{p,n+m}$。

对于标准正态分布的上侧 $\xi_{\alpha/2}$ $\alpha/2$ 分位数。因此，概率 $\tilde{}S_{A,n}(\hat{}t_{p,n+m}-\delta)$ 将位于区间内

$$\left[S_A(t_p-\delta)-\xi_{\alpha/2}\frac{\sigma_{A,\delta}}{\sqrt{n}}\,;\,S_A(t_p-\delta)+\xi_{\alpha/2}\frac{\sigma_{A,\delta}}{\sqrt{n}}\right] \tag{1}$$

约等于 $1-\alpha$。该区间的边界取决于真实的生存函数 S_A,S_B、$G_{Acc}\hat{}$ 以及比例 p （通过 $\sigma_{A,\delta}$）。Breslow估计量 $S_{A,n}$ 也适用相同的边界。对于B 组的生存曲线，可以构造类似的区间。

在规划试验时，为S_A、S_B 和G_{Acc} 所选择的分布函数在确定样本量方面发挥着重要作用。值t_p 也是这些曲线等因素的函数，并且在设计研究时可以计算S_A（$t_p-\delta$）和S_B（$t_p-\delta$）。这意味着，在试验的设计阶段，可以选择p （以及t_p），使S_A（$t_p-\delta$）和S_B（$t_p-\delta$）及相应的界限充分低于1，从而使生存曲线对于当前目标而言足够成熟。在第5节中，将解释如何在实际情境中使用这一间隔。

3 渐近结果

本节给出双臂临床试验（定理1）和单臂试验（定理2）中，在随机时间 $\hat{}t_{p,\ n+m}-\delta$ （随访时间）处计算的 Kaplan-Meier 和 Breslow 估计量的渐近分布。这些估计量基于期中分析时刻 $\hat{}t_{p,\ n+m}$ （日历时间）之前发生的所有观测值。定理的证明见附录。定理中使用的符号，尤其是渐近方差 $\sigma^2_{A,\ \delta}$ 的定义，在定理下方给出。

定理1：双臂临床试验
假设期中分析在事件发生比例达到 100p% 时进行，不论事件发生在哪个组，时点为 $\hat{}t_{p,\ n+m}$。令 $\hat{}S_{A,\ n}$ 和 $\tilde{}S_{A,\ n}$ 分别为S_A 的 Breslow 估计量和 Kaplan-Meier 估计量。令 $\delta>0$ 为预先指定的，$\hat{}t^{\delta}_{p,\ n+m}=$、$\hat{}t_{p,\ n+m}-\delta$ 及其极限$t^{\delta}_p=t_p-\delta$。Breslow 估计量的渐近分布如下：

$$\sqrt{n}(\hat{S}_{A,n}(\hat{t}^{\delta}_{p,n+m})-S_A(t^{\delta}_p))\rightsquigarrow\mathcal{N}(0,\sigma^2_{A,\delta}),$$

而且，Kaplan-Meier估计量的渐近分布相同：

$$\sqrt{n}(\tilde{S}_{A,n}(\hat{t}^{\delta}_{p,n+m})-S_A(t^{\delta}_p))\rightsquigarrow\mathcal{N}(0,\sigma^2_{A,\delta}),$$

（其中 $\rightsquigarrow$ 是依分布收敛的记号），当$n,m \to \infty$且

$$\sigma^2_{A,\delta} = S_A(t_p^\delta)^2 \int_0^{t_p^\delta} \frac{\mathrm{d}\Lambda_A(s)}{1-H_{A,t_p}(s)} + \frac{q_A f_A(t_p^\delta)^2}{(h^{uc,\star}_{mix,t_p}(t_p))^2}\, p(1-p)$$

$$- 2S_A(t_p^\delta)\frac{q_A\sqrt{q_A} f_A(t_p^\delta)}{h^{uc,\star}_{mix,t_p}(t_p)}\left((1-H^{uc,\star}_{A,t_p}(t_p))\Lambda_A(t_p^\delta) + \int_0^{t_p^\delta} \frac{(H^{uc}_{A,t_p}(s) - H^{uc,\star}_{A,t_p}(t_p)H_{A,t_p}(s))\mathrm{d}\Lambda_A(s)}{1-H_{A,t_p}(s)}\right).$$

ˆ 该定理的证明见附录 B。对于 B 组的 Breslow 估计量 $S_{B,m}$ 和 Kaplan-Meier 估计量 $\tilde{S}_{B,m}$，也有类似的结果。方差 $\sigma^2_{A,\delta}$ 取决于多个分布函数和参数。下文将解释相关记号。与日历时间分析一样，只使用截至期中分析时的观测值，因此渐近地有 $L = t_P$，且观测值在日历时间 t_P处删失：因此 $E_A + C_A = t_P$。在随访时间尺度上，（次）分布函数 H_{A,t_p} 和 H^{uc}_{A,t_p} 定义为 $H_{A,t_p}(t) = P(T_A \wedge C_A \le t)$ 和 $H^{uc}_{A,t_p}(t) = P(T_A \wedge C_A \le t\ \Delta_A = 1) = P(T_A \le t\ \Delta_A = 1)$，其中后者针对未删失观测（上标“uc”表示“未删失”）。采用 B 组的类似记号，两组分布的混合分布定义为

$$H_{mix,t_p}(t) = q_A H_{A,t_p}(t) + q_B H_{B,t_p}(t), \qquad H^{uc}_{mix,t_p}(t) = q_A H^{uc}_{A,t_p}(t) + q_B H^{uc}_{B,t_p}(t),$$

其中，$q_A = \lim_{n,m\to\infty} n/(n+m)$，$q_B = \lim_{n,m\to\infty} m/(n+m)$。在日历时间下，定义非常相似：定义（次）分布函数 $H^\star_{A,t_p}(t) = P(E_A + (T_A \wedge C_A) \le t)$ 和 $H^{uc,\star}A,tp\,(t) = P(E_A + T_A \le t\ \Delta_A = 1)$。其密度函数分别记为 $h^\star_{A,t_p}$ 和 $h^{uc,\star}A,tp$。两组中分布的混合分布定义为 $H^\star_{mix,t_p}(t) = q_A H^\star_{A,t_p}(t) + q_B H^\star_{B,t_p}(t)$，以及 $H^{uc,\star}$ mix,tp(t) $= q_A H^{uc,\star}$ A,tp (t) $+ q_B H^{uc,\star}$ B,tp (t)。请记住，当所有患者中有 100p% 发生了事件时，即在时间点 $\hat{t}_{p,n+m}$，将进行期中分析。这个随机时间点 $\hat{t}_{p,n+m}$ 依概率收敛于 t_P，其定义为混合分布 $H^{uc,\star}$ mix,tp 的第p分位数。

在下一个定理中，我们考虑单臂情形。由于无需区分组别，因此在记号中省略A和B。

定理2：单组临床试验

定义 $\hat{t}_{p,n}$ 为患者中 100p% 发生事件的时刻。令 $\hat{S}_n$ 和 $\tilde{S}_n$ 分别为基于截至时间 $\hat{t}_{p,n}$的观测值计算得到的 Breslow 估计量和 Kaplan-Meier 估计量。令 $\delta > 0$为固定值，$\hat{t}^\delta_{p,n} = \hat{t}_{p,n} - \delta$及其极限 $t_p^\delta = t_p - \delta$。对于 Breslow 估计量，有如下结论：

$$\sqrt{n}(\hat{S}_n(\hat{t}^\delta_{p,n}) - S(t_p^\delta)) \rightsquigarrow \mathcal{N}(0, \sigma^2_\delta),$$

以及对于Kaplan-Meier估计量彩票平台注册

$$\sqrt{n}(\tilde{S}_n(\hat{t}^\delta_{p,n}) - S(t_p^\delta)) \rightsquigarrow \mathcal{N}(0, \sigma^2_\delta),$$

作为 $n \to \infty$，其中

$$\sigma^2_\delta = S(t_p^\delta)^2 \int_0^{t_p^\delta} \frac{\mathrm{d}\Lambda(s)}{1-H_{t_p}(s)} + \frac{f(t_p^\delta)^2}{(h^{uc,\star}_{t_p}(t_p))^2}\, p(1-p)$$

$$- 2S(t_p^\delta)\frac{f(t_p^\delta)}{h^{uc,\star}_{t_p}(t_p)}\left((1-p)\Lambda(t_p^\delta) + \int_0^{t_p^\delta} \frac{(H^{uc}_{t_p}(s) - H^{uc,\star}_{t_p}(t_p)H_{t_p}(s))\mathrm{d}\Lambda(s)}{1-H_{t_p}(s)}\right). \tag{2}$$

定理2由定理1推得，令$m = 0$，因此$q_A = 1$ 且$q_B = 0$。

式 (2) 中的渐近方差由三项组成。第一项可视为生存函数 S 估计所导致的方差，第二项源于时间点t_P的估计，最后一项

源于这两项之间的协方差。由于 $\hat{S}_n$ 和 $\hat{t}_{p,n}$ 呈负相关，因此该协方差为负。如果显示式中第二项与第三项之和为负，则用于估计 $S(t_p)$ 的估计量 $\hat{S}_n(\hat{t}_{p,n})$ 的渐近方差小于估计量 $\hat{S}_n(t_p)$ 的渐近方差，尽管后一个估计量是在固定时间点t_p确定的。这并不令人意外，原因可由以下示例说明。考虑不存在删失的情况。在这种情况下，Kaplan-Meier 曲线等于经验生存曲线，并且根据定义，$\tilde{S}_n(\hat{t}_{p,n})$ 等于 $1-p$，因此渐近方差将等于零。$\tilde{S}_n(t_p)$ 的渐近方差等于 $p(1-p)$，其值大于零。

4 模拟研究

渐近结果与模拟结果的比较

在定理1中，给出了双臂试验中 Kaplan-Meier 估计量和 Breslow 估计量的渐近分布。根据该渐近分布可知，概率 $\tilde{S}_{A,\,n}(\hat{t}_{p,\,n+m})$ 将落在区间内

$$\left[S_A(t_p^\delta) - \xi_{\alpha/2}\frac{\sigma_{A,\delta}}{\sqrt{n}}\,;\, S_A(t_p^\delta) + \xi_{\alpha/2}\frac{\sigma_{A,\delta}}{\sqrt{n}}\right]$$

约等于 $1-\alpha$。本小节旨在通过将渐近区间与蒙特卡罗模拟得到的区间进行比较，研究渐近区间的准确性。

在所有情景中，均假设两组中发生关注事件的时间服从指数分布。总共考虑八种不同的设定，通过改变两组的风险比、疾病严重程度（以中位生存时间表示）以及疾病罕见程度（以纳入患者所需的入组期表示）获得。具体而言，我们考虑

- 治疗效应以风险比表示：我们考虑两种情况：– 强效应：风险比等于 0.65。这意味着 $\theta_B/\theta_A = 0.65$.– 中等效应：风险比等于 0.75。这意味着 $\theta_B/\theta_A = 0.75$。
- 疾病严重程度：
  – 进展迅速：中位生存时间为 6 个月；$\theta_A = -\log(0.5)/6 = 0.12$。– 进展缓慢：中位生存时间为 36 个月；$\theta_A = -\log(0.5)/36 = 0.019$。
- 疾病罕见程度：
  – 罕见：每月累积 4 名患者；入组时间等于 $(n+m)/4$ 个月。– 常见：每月累积 20 名患者；入组时间等于 $(n+m)/20$ 个月。

在临床试验中，样本量通常根据研究结束时对数秩检验具有足够效能来确定。事实上，关键是患者人数与随访相结合，能够确保获得足够的（预期）事件数。对数秩检验（Schoenfeld, 1991）或 Cox 回归（Schoenfeld, 1983）所需的预期事件数通过 Schoenfeld 公式计算：$d = \#\text{events} = (\xi_{\alpha/2}+\xi_\beta)^2/((\log \text{HR})^2 q_A q_B)$，其中 HR 为风险比。对于 1:1 随机分配、$\alpha = 0.05$（双侧）和 80% 效能，$\#\text{events} = 31.4/(\log \text{HR})^2$。一种常见做法是选择患者人数$n+m$、入组分布函数$G_{Acc}$以及最后一名患者招募后的随访时长$FU$的组合，使得预期事件数（最后一名患者入组并完成$FU$ 随访后）等于所需的预期事件数d。如果招募在日历时间上从 0 到R 的招募期内均匀分布（即G_{Acc} 在 0 到R的区间内的密度为$1/R$，在其他位置为 0），那么试验时长为$L = R + FU$，在 0 到L 期间，A 组的预期事件数为：

$$\begin{aligned} n\,H^{uc,\star}_{A,L}(L) &= n\,\mathrm{P}(T_A \le L - E_A\,,\,\Delta_A = 1) = n/R\int_0^R \mathrm{P}(T_A \le L - s)\,\mathrm{d}s \\ &= n/R\int_0^R \Big(1 - \exp(-\lambda_A(L-s))\Big)\,\mathrm{d}s = n\left[1 - \frac{\exp(-\lambda_A L)}{\lambda_A R}\,\big(\exp(\lambda_A R) - 1\big)\right], \end{aligned} \tag{3}$$

组 B 也有类似的表达式。样本量、招募时间 R 和随访 FU 的组合选择满足以下条件：$nH^{uc}_{,\star A,L}(L) + mH^{uc}_{,\star B,L}(L) = d$。

为了减少不同设定的数量，本文考虑以下情形。首先，随机化比例为 1:1：公式中的$n = m$。此外，最后一名患者招募后FU 随访时间固定为 6 个月。即便如此，仍有若干招募期R与总样本量 $2n$ 的组合，能够在随访结束时达到所需的事件数。根据疾病发生率，考虑每月招募 4 名或 20 名患者的招募率。这些招募率确定了样本量 $2n$ 与招募期R 的比例，从而得到一组 $2n$与R的组合。就疾病严重程度（即生存曲线）而言，考虑侵袭性疾病和惰性疾病。对于两种治疗效应情景（$hr = 0.65$ 或$hr = 0.75$），所需事件数分别为 170 和 380。

期中分析通常在观察到最终分析所需事件数的一定比例时进行。该比例称为信息分数（IF）。如果（且仅当）患者总数已确定，IF 与比例p一一对应；比例p 表示期中分析时发生事件的患者比例（为与信息分数相区别，也可称为患者分数）。

在模拟研究中，我们考虑两个期中分析时点：

- 早期期中分析：在发生40%的事件后：如果$= 0.40$。
- 晚期期中分析：在发生60%的事件后：如果$= 0.60$。

所考虑的不同设置见表1。

渐近区间通过根据所选设置计算其上下界得出。有限样本区间的计算方法如下：对研究中的$n + m$名患者进行抽样。根据抽样数据计算估计值 $\hat{S}_{A,n}(\hat{t}^{\delta}_{p,n+m})$。重复此过程1000次。以估计值的样本均值以及第2.5%和第97.5%分位数构建一个95%的预测区间。

nr.	疾病/治疗			试验				中期			
	效应 (HR)	严重程度 （中位数）	罕见性 （患者/监测）	$n+m$ 患者	总计 # 事件数	累积 （月）	FU （月）	IF	# 事件	p	t_p
1.	0.65	6	4	196	170	49	6	0.40	68	0.35	27
2.	0.65	6	20	260	170	13	6	0.40	68	0.26	10
3.	0.65	36	4	344	170	86	6	0.40	68	0.20	53
4.	0.65	36	20	620	170	31	6	0.60	102	0.16	27
5.	0.75	6	20	480	380	24	6	0.40	152	0.32	16
6.	0.75	36	4	580	380	145	6	0.40	152	0.26	82
7.	0.75	36	20	1000	380	50	6	0.40	152	0.15	33
8.	0.75	36	20	1000	380	50	6	0.60	228	0.23	41

表1：模拟研究的不同情景。在所有情况下，$n = m$；两组患者人数相等。变量“事件总数”表示根据 Schoenfeld 公式，为 $\alpha = 0.05$ 且效能为0.80所需的事件总数。变量“事件数”（用于期中分析）表示期中分析时观察到的事件数。“FU”是最后一位患者入组后的随访时间。

模拟研究的结果（表2）显示，渐近区间是准确的，因为它们与基于蒙特卡罗模拟构建的区间非常相似。此外，在所有设定下，区间都足够窄，可用于规划期中分析。

δ值 δ 与预测区间的宽度

请记住，期中分析是在日历时间$\hat{t}_{p,n+m}$ 进行的。由于没有患者的随访时间能够达到或超过$\hat{t}_{p,n+m}$ （见图2），因此在随访时间$\hat{t}_{p,n+m} - \delta$处估计两组的生存曲线，其中 $\delta > 0$。如果 δ较小，随访时间达到或超过$\hat{t}_{p,n+m} - \delta$的患者可能只有少数，且

	组 A				组 B			
nr.	$\hat{S}_{A,n}(\tilde{t}^{\delta}_{p,n+m})$	95% 区间	$S_A(t_p^{\delta})$	渐近区间	$\hat{S}_{B,m}(\tilde{t}^{\delta}_{p,n+m})$	95%区间	$S_B(t_p^{\delta})$	渐近区间
1.	0.07	[0.00；0.20]	0.06	[-0.06；0.18]	0.17	[0.00；0.36]	0.16	[-0.02；0.34]
2.	0.36	[0.20；0.50]	0.36	[0.20；0.52]	0.52	[0.35；0.66]	0.52	[0.36；0.68]
3.	0.39	[0.23；0.53]	0.40	[0.25；0.55]	0.55	[0.38；0.69]	0.55	[0.40；0.70]
4.	0.63	[0.53；0.70]	0.63	[0.54；0.72]	0.74	[0.66；0.81]	0.74	[0.66；0.82]
5.	0.20	[0.07；0.30]	0.20	[0.09；0.31]	0.30	[0.16；0.41]	0.30	[0.17；0.42]
6.	0.24	[0.13；0.36]	0.24	[0.13；0.35]	0.34	[0.21；0.46]	0.34	[0.22；0.46]
7.	0.57	[0.48；0.65]	0.57	[0.48；0.65]	0.65	[0.56；0.73]	0.65	[0.57；0.73]
8.	0.49	[0.41；0.56]	0.49	[0.41；0.57]	0.58	[0.50；0.66]	0.58	[0.51；0.66]

表2：模拟结果。第一列中的数字对应表1中的模拟情景。在所有情况下，均取 δ 等于 $\delta = 0.1tp$。

	组 A				组 B			
δ	$\hat{S}_{A,n}(\tilde{t}^{\delta}_{p,n+m})$	95% 区间	S_A（t_p^{δ}）	渐近区间	$\hat{S}_{B,m}(\tilde{t}^{\delta}_{p,n+m})$	95% 区间	S_B（t_p^{δ}）	渐近区间
0.01	0.37	[0.00；0.53]	0.36	[0.14；0.58]	0.51	[0.26；0.67]	0.52	[0.30；0.73]
0.10	0.40	[0.22；0.53]	0.40	[0.25；0.55]	0.55	[0.37；0.69]	0.55	[0.40；0.70]
0.25	0.47	[0.34；0.58]	0.46	[0.34；0.59]	0.61	[0.50；0.72]	0.61	[0.49；0.73]
0.01	0.54	[0.39；0.64]	0.54	[0.41；0.66]	0.62	[0.49；0.72]	0.63	[0.51；0.74]
0.10	0.57	[0.48；0.65]	0.57	[0.48；0.65]	0.65	[0.57；0.73]	0.65	[0.57；0.73]
0.25	0.62	[0.56；0.69]	0.62	[0.56；0.69]	0.70	[0.65；0.76]	0.70	[0.64；0.77]

表 3：模拟结果。 δ 的值： $\delta = x\,tp$ 与第一列中的值相同。

$\tilde{t}_{p,n+m} - \delta$处的Kaplan-Meier估计量可能不准确，从而导致$S_A(t_p - \delta)$和$S_B(t_p - \delta)$的区间较宽。另一方面，如果 δ较大，则会在随访时间较早的时点计算Kaplan-Meier（或Breslow）估计量，可能无法反映期中分析时可获得的全部信息。这些相互制衡的考量类似于偏差-方差权衡。研究者可以选择 δ 的取值，因此，研究 δ 对预测区间宽度的影响很有意义。针对 δ的不同取值，考察了设置3和7（见表1）。从表3可以看出，随着 δ增大，$t_p - \delta$的值变小，而$t_p^{\delta} = t_p - \delta$处的生存曲线则上升（按定义如此）。此外，还可以看出，预测区间的宽度会随着 δ增大而减小。这是因为仍处于风险集中的患者越多，生存曲线就能估计得越准确。我们还考察了更多设置，结论相同（结果未显示）。

5 实践中

旨在采用对数秩检验或Cox回归分析设计确证性临床试验的研究者，通常首先关注的是获得足够的效能。在组序贯方法中，这类检验在最终分析或期中分析时的效能取决于这些分析时的事件数和临界值（α消耗函数）。在实践中，期中分析通常在达到最终分析所需事件数的一定比例后进行，以确保最终分析具有足够的效能（这称为信息分数）。为说明信息分数（仅）与效能相关，请看以下示例：当最终分析在某一效应量下的效能为 80%（双侧显著性水平为 0.05），且计划采用 O'Brien-Fleming 界值进行一次期中分析时，在信息分数（IF）为 40%、IF=50%、IF=60%、IF=70%、IF=80% 和 IF=90% 时，该次期中分析中检验统计量达到统计学显著性的效能分别为 6%、18%、34%、51%、66% 和 76%。反过来说，只有在 the

真实效应大于最终分析时预期的效应。在对数风险比尺度上，该效应应增大的倍数分别为：IF=40% 时为 2.136，IF=50% 时为 1.864，IF=60% 时为 1.554，IF=70% 时为 1.330，IF=80% 时为 1.166，IF=90% 时为 1.051。

除了效能之外，期中分析时生存数据的成熟度也会产生影响。在推导出的公式中（第3节），用于估计期中分析时Kaplan-Meier曲线预测区间的输入参数包括：各组的生存分布、入组分布函数G_{Acc}、各组的相对样本量、比例p以及接近度参数 δ。分布S_A,S_B的选择，理想情况下应基于此前开展的类似（例如探索性）研究结果；否则，则应基于临床判断。通过选择招募中心和/或根据后勤可行性，研究者可以在较大程度上掌控入组率和入组时间（由G_{Acc}概括）、各组的患者人数n和m，以及最短随访时间。

效能分析无需指定参数p，但选定患者总数后即可计算其值。该参数反映了期中分析将在何时进行（以随访时间表示），因此p 决定了截至该时点预期的生存曲线。p的值越高，期中分析进行得越晚，可获得的生存曲线信息也越多。期中分析时通常需要有足够的信息，因此应据此选择p。例如，信息充分可能意味着，期中分析时点之前的生存曲线足以合理代表整个患者群体的生存情况。另一方面，期中分析也不应离试验结束时间过近，否则依据期中分析采取行动便失去意义。研究者可以并且应该考虑 δ 的不同取值。通过改变 δ，可以权衡期中分析时间的接近程度与获得对当前目标而言有意义的精确度。

总之，我们设想采用以下策略来确定期中分析的时间。首先，像往常一样考虑效能（即事件数）。为了评估生存曲线的成熟度，可使用例如我们的方法，估计期中分析时间附近的预期生存率。接下来，可调整分析时间，使得在期中分析时，不仅预期能达到足够的效能，也能获得足够成熟的数据，以满足当前目的。建议针对两组一系列合理的生存曲线进行考察。

一项应用

我们使用所提出的方法，评估 Keynote 204研究中预先计划的无进展生存期（PFS）期中分析是否有望事先获得成熟的数据。Keynote 204是一项研究，旨在对比帕博利珠单抗与维布妥昔单抗对照治疗复发或难治性经典型霍奇金淋巴瘤患者的疗效。相关设计参数为：300名患者按1:1随机分组；入组期为12个月；假设生存时间服从指数分布，维布妥昔单抗组的中位无进展生存期为5.6个月；风险比为0.622；计划在疗效（即确证性）期中分析时观察到176个事件（见 Kuruvilla等（2021）在线附录中的方案第93页和第96页）。方案未说明入组率随时间变化的形态，因此我们假设300名患者在12个月内均匀入组。此外，方案未规定最短随访时间。因此，我们假设在所有患者均完成招募后进行期中分析；此时为$t_p = 14.87$ 个月。

现在，我们评估接近期中分析预期时间的不同时间点上的 Kaplan-Meier曲线。从 $\delta = 0.1t_p$（这相当于比期中分析计划的预期时间早 6.4 周）开始，维布妥昔单抗组和帕博利珠单抗组的 Kaplan-Meier估计（95%预测区间）分别为 19%（9%-29%）和 36%（24%-47%）。在 $= 0.05t_p$（比期中分析早 3.2 周）时，维布妥昔单抗组和帕博利珠单抗组的估计值分别为 17%（6%- 29%）和 34%（21%-47%）。观察 $= 0.01t_p$（即期中分析前 4.5 天）时，我们预计维布妥昔单抗组为 16%（2%-30%），帕博利珠单抗组为 32%（16%-49%）。因此，预计两组均可观察到相当一部分生存曲线。特别是，预计可观察到早期和晚期事件，计划中的期中分析能够评估绝大多数人群的生存获益（如果存在）。

6 总结与讨论

在本文中，我们推导了在达到预先设定的事件总数时（该时间为随机时间），生存曲线预测区间的渐近边界；同时也涵盖了在计划招募的患者总数尚未达到之前就已达到该事件数的情况。推导公式的输入参数包括假定的生存曲线、患者累积、预先设定的事件总数以及患者总数。这些本就是规划生存结局试验时常用的参数，因此无需新增信息。因此，该推导公式可轻松应用于常规的试验规划实践。尽管我们关注的是双臂试验中的期中分析，但最终分析中的预期生存估计值也同样可以估计。

所提出理论的一项典型应用是III期随机临床试验，其中计划进行期中分析，以因疗效而提前停止，并采用对数秩检验或Cox模型检验。期中分析通常计划在发生指定数量的事件时进行。这只能涵盖统计检验是否有望达到足够的效能。由于该检验通常针对整条曲线是否相等，期中分析达到统计学显著性可能仅仅是由生存曲线早期的差异所致。这并不能保证生存数据已足够成熟，可以对研究目的得出结论。例如，欧洲药品管理局的肿瘤学指南（EMA）指出，期中分析“只有在现有生存数据能够提供充分评估获益/风险所需的信息时，才应进行”。同一指南还指出：“如果已观察到长期预期事件总数中的明显多数，且已记录到差异，通常便会接受这一点，将其视为研究已达到合理成熟度、且研究结果在长期随访期间将保持稳定的指标。”值得注意的是，“预期事件总数”这一措辞指的是长期（较长的随访时间），而非最终分析时的事件总数。尽管如此，在规划阶段，人们往往不会明确考虑期中分析时生存曲线的成熟度，或者只进行定性考量。我们的结果可以定量规划生存数据的成熟度。它们还能揭示患者分数 p （由研究者选定）与期中分析时预期可获得的信息量之间的关系：Kaplan-Meier生存估计值及其相应的预测区间。因此，本文的结果有助于试验设计。

在模拟研究中，假设生存时间来自指数分布，入组时间来自均匀分布。这是设计具有生存结局的试验时最常采用的设定。然而，渐近分布是针对任意分布推导的。应通讯作者要求，可提供R代码；该代码也可轻松调整以适用于其他分布（例如Weibul分布或Parmar-Royston模型（Royston和Parmar，2002年）），从而使应用更加广泛。

利益冲突
作者声明不存在潜在的利益冲突。

数据可用性声明
本出版物未使用真实世界数据。

ORCID
Marianne Jonker：https://orcid.org/ 0000-0003-0134-8482

参考文献

[1] Aalen O., Borgan O., Gjessing H.. (2008) 生存分析与事件史分析。Springer-Verlag.[2] Bagiella E., Heitjan D.F.. (2001) 预测随机临床试验的分析时间。*Statist.Med.*, 20: 2055–2063.[3] Breslow N.E.. (1972) 对Cox教授论文的讨论。*Royal Stat. Soc. B*,34: 216-217.

[4]EMA肿瘤学指南《人体抗癌药物评估指南》。EMA/CHMP/205/95 Rev.5.*https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-anticancer-medicinal-products-man-revision-5 en.pdf*[5] Floriani I., Rotmensz N., Albertazzi E., Torri V., De Rosa M., Tomino C., de Braud F.. (2008) 癌症随机临床试验至事件时间终点期中分析方法：意大利国家临床试验监测中心的一项调查。*Trials*, 9: 46. doi: 10.1186/1745-6215-9-46. [6] Freidlin B., Othus M., Korn E.L.. (2016) 随机临床试验期中分析的信息时间尺度。*ClinTrials.*, 13(4):391-9. doi: 10.1177/1740774516644752[7]Jennison C., Turnbull B.W.. (2000) 组序贯方法及其在临床试验中的应用。CRC/Chapman & Hall, 英国.[8] Kuruvilla J., Ramchandren R., Santoro A., Paszkiewicz-Kozik E., Gasiorowski R., Johnson N.A., Fogliatto L.M., Goncalves I., de Oliveira J.S.R., Buccheri V., Perini G.F., Goldschmidt N., Kriachok I., Dickinson M., Komarnicki M., McDonald A., Ozcan M., Sekiguchi N., Zhu Y., Nahar A., Marinello P., Zinzani P.L.; KEYNOTE-204研究者. (2021) 帕博利珠单抗与维布妥昔单抗治疗复发或难治性经典型霍奇金淋巴瘤（KEYNOTE-204）：一项多中心、随机、开放标签3期研究的期中分析。*Lancet Oncol.*22:512-524. doi: 10.1016/S1470-2045(21)00005-X.[9] Lan K.K.G., Lachin J.M.. (1990) 在最大持续时间试验中实施组序贯对数秩检验。*Biometrics*, 46: 759-770.[10] Lan K.K.G., Reboussin D.M., DeMets D.L.. (1994) 临床试验设计与序贯监测的信息及信息分数。*Communications in Statistics*, 23(2): 403-420.[11] Lin D.Y., Shen L., Ying Z., Breslow N.E.. (1996) 监测生存概率的组序贯设计 *Biometrics*, Vol 52, No 3, 1033–1041..[12] Murray T.A., Bradley P.C.. Lystig T.C.. (2013) 用于器械监测的贝叶斯自适应设计。*Clin Trials*. 10: 5-18.[13] Royston P., Parmar M.K.B. (2002) 针对删失生存数据的灵活参数比例风险与比例优势模型及其在预后建模和治疗效应估计中的应用。*Statist.Med.*, 21: 2175–2197.[14] Royston P., Barthel F.M.. (2010) 至事件时间结局临床试验中的效能与事件数预测。*TheStata Journal*, 10: 386–394.[15] Schoenfeld D.A.. (1981) 比较生存分布的非参数检验的渐近性质。*Biometrika*, 68: 316-319.[16] Schoenfeld D.A.. (1983) 比例风险回归模型的样本量公式。*Biometrics*39:499-503.[17] van Houwelingen H.C., van der Velde C.J., Stijnen T.. (2005) 生存数据的期中分析：其潜在偏倚及修正方法。*StatMed*. 24: 2823-2835. [18] Therneau T.M., Grambsch P.M.. (2000) 生存数据建模：扩展*Cox*模型. Springer-Verlag New York, Inc..[19] van der Vaart A.W.. (1998) 渐近统计学. 剑桥大学出版社，剑桥.[20] van der Vaart A.W., Wellner J.A.. (1996) 弱收敛与经验过程。Springer-Verlag. 纽约。

附录：渐近分布的推导

附录 B 给出了定理1的证明。该证明依赖于在随机时间点推导 Breslow 估计量和 Kaplan-Meier估计量（基于所有观测值）的渐近分布。附录A给出了该渐近分布的推导。下面首先介绍新的记号。

记号

假设进行一项包含 A 组和 B 组的双臂临床试验。令 T_A 表示 A 组中任意一位患者从进入研究到发生关注事件的时间。其分布函数记为 F_A，生存函数记为 $S_A = 1 - F_A$，密度记为 f_A，风险函数记为 $\lambda_A(t) = f_A(t)/(1 - F_A(t)) = f_A(t)/S_A(t)$。两组的删失时间 C_A（从进入研究到研究结束的时间）的分布记为 G_L，该分布取决于 L，其中 L 是试验暂时暂停或正式结束的时间（日历时间）。入组时间 E_A（日历时间）的分布记为 G_{Acc}，G_{Acc} 与 G_L 之间的关系如下：

$$G_L(c) = P(C_A \le c) = 1 - G_{Acc}(L - c).$$

其中我们利用了 $E_A + C_A = L$。

此外，定义 $H_{A,L}(t) = P(T_A \wedge C_A \le t)$ 和 $H^{uc}_{A,L}(t) = P(T_A \wedge C_A \le t \ \Delta_A = 1) = P(T_A \le t \ \Delta_A = 1)$。采用这一记号，$T_A$ 的累积风险函数可写为

$$\begin{aligned}\Lambda_A(t) &= \int_0^t \lambda_A(s)ds = \int_0^t \frac{1}{S_A(u)} dF_A(u) \\ &= \int_0^t \frac{1 - G_L(u)}{(1 - F_A(u))(1 - G_L(u))} dF_A(u) = \int_0^t \frac{1}{1 - H_{A,L}(u)} dH^{uc}_{A,L}(u), \end{aligned} \tag{4}$$

其中，最后一个等式由 $1 - H_{A,L}(t) = P(T_A \wedge C_A > t) = P(T_A > t, C_A > t) = (1 - F_A(t))(1 - G_L(t))$ 以及 $H^{uc}_{A,L}(t) = \int_0^t 1 - G_L(u) dF_A(u)$ 这一事实推出。生存函数 S_A 与累积风险函数 Λ_A 之间的关系为

$$S_A(t) = \exp(-\Lambda_A(t)). \tag{5}$$

考虑包含 n 个独立同分布抽样患者的标准右删失情形。（随访）观测值为 $(T_{A,1} \wedge C_{A,1}, \Delta_{A,1}) \ldots, (T_{A,n} \wedge C_{A,n}, \Delta_{A,n})$。基于这些观测值，分布$H_{A,L}$和次分布$H^{uc}_{A,L}$由其经验分布函数$\mathbb{H}_{A,n}$和 $\mathbb{H}^{uc}_{A,n}$ 估计，定义如下：

$$\mathbb{H}_{A,n}(t) = \frac{1}{n} \sum_{i=1}^n 1\{T_{A,i} \wedge C_{A,i} \le t\}, \tag{6}$$

$$\mathbb{H}^{uc}_{A,n}(t) = \frac{1}{n} \sum_{i=1}^n 1\{T_{A,i} \wedge C_{A,i} \le t, \Delta_{A,i} = 1\} = \frac{1}{n} \sum_{i=1}^n \Delta_{A,i} 1\{T_{A,i} \wedge C_{A,i} \le t\}.$$

通过将（4）中的$H_{A,L}$和$H^{uc}_{A,L}$替换为其经验分布函数$\mathbb{H}_{A,n}$和 $\mathbb{H}^{uc}_{A,n}$，即可得到 Nelson-Aalen估计量，从而估计累积风险函数 Λ_A。

$$\hat{\Lambda}_{A,n}(t) = \int_0^t \frac{1}{1 - \mathbb{H}_{A,n}(s)} d\mathbb{H}^{uc}_{A,n}(s), \tag{7}$$

求得了针对Breslow估计量（Breslow，1972）的S_A：

$$\hat{S}_{A,n}(t) = \exp(-\hat{\Lambda}_{A,n}(t)) = \exp\Big(-\int_0^t \frac{1}{1 - \mathbb{H}_{A,n}(u)} d\mathbb{H}^{uc}_{A,n}(u)\Big). \tag{8}$$

分布函数F_A估计为$\hat{F}_{A,n} := 1 - \hat{S}_{A,n}$。

生存曲线S_A 也通过乘积积分与 Λ_A 相关

$$S_A(t) = \prod_{u \le t}(1 - \mathrm{d}\Lambda_A(u)),$$

其中，$\prod$符号表示乘积积分。将累积风险 Λ_A 替换为 Nelson-Aalen 估计量$\Lambda_{A,n}$即可得到S_A 的 Kaplan-Meier 估计量：

$$\tilde{S}_{A,n}(t) = \prod_{u \le t}(1 - \mathrm{d}\hat{\Lambda}_{A,n}(u)), \tag{9}$$

~ ~ ^，而F_A的相应估计量记为$F_{A,n} = 1 - S_{A,n}$。Breslow 估计量$S_{A,n}$与 Kaplan-~Meier 估计量$S_{A,n}$并不完全一致，但渐近等价（Therneau 和 Grambsch，2000年）。

分布 $\mathrm{H}^{\star}_{A,L}(\mathrm{t}) = \mathrm{P}(\mathrm{E}_A + (\mathrm{T}_A \wedge \mathrm{C}_A) \le \mathrm{t})$、$\mathrm{H}^{uc}{}_{,\star}\mathrm{A,L}\,(\mathrm{t}) = \mathrm{P}(\mathrm{E}_A + (\mathrm{T}_A \wedge \mathrm{C}_A) \le \mathrm{t}\ \Delta_A = 1)$，以及 B组对应的分布，均通过相应的经验分布函数进行估计；对于 $\mathrm{H}^{\star}_{A,L}$ 和 $\mathrm{H}^{uc}{}_{,\star}$ A,L，其估计值分别记为$\mathbb{H}^{\star}_{A,n}$ 和 $\mathbb{H}^{uc}{}_{,\star}$ A,n。此外，将相应组特异性分布替换为其经验分布函数，可得到 $\mathrm{H}_{mix,L}$、$\mathrm{H}^{uc}_{mix,L}$、$\mathrm{H}^{\star}_{mix,L}$ 和 $\mathrm{H}^{uc}{}_{,\star}$ mix,L 的估计量，并以类似方式记号表示。

附录A：完整观测值、随访时间

本附录讨论单臂试验中按随访时间进行的情形：总计 n 名患者进入试验，并持续随访，直至关注事件发生或试验结束时删失。研究目的是估计随机时点 $\hat{t}_{n,p}$处的生存曲线；该时点是p100%的患者已发生事件的时刻；并推导该时点处 Brelow 估计量和 Kaplan-Meier估计量的渐近分布。众所周知，在固定的非随机时点处计算的 Kaplan-Meier估计量和 Breslow 估计量渐近正态（例如参见 Aalen 等人（2008年）；Therneau 和 Grambsch（2000年））， 但对于随机时点 $\hat{t}_{p,n}$，其渐近分布尚不明确。请注意，本附录讨论的情形与定理2中的情形不同；在定理2中（同样是单臂试验），试验在进行期中分析时停止，即在时点 $\hat{t}_{p,n}$停止；删失时间是随机的，并且依赖于观测值，而且在时点 $\hat{t}_{p,n}$之前，可能并非所有患者都已进入试验。此处有意对情形作了简化；这一简化情形下的证明将用于说明证明思路，该思路也将用于附录 B 中定理1（及定理2）的证明。

符号沿用前文的定义，但由于此处考虑的是单组试验，省略了下标中的A （用于表示研究的组别）。研究中n位患者的观测值由$(T_1 \wedge C1, \Delta_1) \ldots, (T_n \wedge C_n, \Delta_n)$组成，并假设这些观测值独立同分布。

定理：单组，随访时间

^ ~设 S_n 和 S_n 分别为基于截至日历时间 L 的所有观测值 $(\mathrm{T}_1 \wedge \mathrm{C1}, \Delta_1) \ldots, (\mathrm{T}_n \wedge \mathrm{C}_n, \Delta_n)$ 得到的 Breslow 估计量和 Kaplan-Meier 估计量，其中 S 为事件时间的生存分布。对于 $\hat{\mathrm{t}}_{p,n}$，即 $\mathbb{H}^{uc}_n$ 的样本第 p 分位数，以及满足 $\hat{\mathrm{t}}_{p,n} \to \mathrm{t}_p$ 依概率收敛（当 $\mathrm{n} \to \infty$）的 t 值，则 $\sqrt{\mathrm{n}}(\hat{\mathrm{S}}_n(\hat{\mathrm{t}}_{p,n}) - \mathrm{S}(\mathrm{t}_p)) \rightsquigarrow \mathcal{N}(0, \sigma^2)$ 且 $\sqrt{\mathrm{n}}(\tilde{\mathrm{S}}_n(\hat{\mathrm{t}}_{p,n}) - \mathrm{S}(\mathrm{t}_p)) \rightsquigarrow \mathcal{N}(0, \sigma^2)$，当 $\mathrm{n} \to \infty$ 时，其中

$$\sigma^2 = S(t_p)^2 \int_0^{t_p} \frac{\mathrm{d}\Lambda(s)}{1 - H(s)} + \frac{f(t_p)^2}{(h^{uc}(t_p))^2}\, p(1-p)$$
$$- 2S(t_p)\frac{f(t_p)}{h^{uc}(t_p)}\Big((1-p)\Lambda(t_p) + \int_0^{t_p} \frac{(H^{uc}(s) - H^{uc}(t_p)H(s))\mathrm{d}\Lambda(s)}{1 - H(s)}\Big). \tag{10}$$

证明：Breslow 估计量，$\hat{S}_n(\hat{t}_{p,n})$

将 $\sqrt{n}(\tilde{S}_n(\hat{t}_{p,n}) - S(t_p))$ 写成三个项之和：

$$\sqrt{n}(\hat{S}_n(\hat{t}_{p,n}) - S(t_p))$$
$$= \sqrt{n}(\hat{S}_n(t_p) - S(t_p)) + \sqrt{n}(S(\hat{t}_{p,n}) - S(t_p)) + \sqrt{n}((\hat{S}_n(\hat{t}_{p,n}) - S(\hat{t}_{p,n})) - (\hat{S}_n(t_p) - S(t_p))). \tag{11}$$

在证明结束时，将证明（11）中的第三项依概率收敛于零。由此并应用斯卢茨基引理可知，只需证明（11）中第一项与第二项之和，即 $\sqrt{n}(\hat{S}_n(t_p) - S(t_p)) + \sqrt{n}(S(\hat{t}_{p,n}) - S(t_p))$，渐近正态，均值为零，方差为 σ^2。接下来分别推导这两项的渐近分布，然后再考察它们的和。

第一项的渐近分布：$\sqrt{n}(\hat{S}_n(t_p) - S(t_p))$

尽管对于任意固定的t，$\sqrt{n}(\hat{S}_n(t) - S(t))$的渐近分布已为人所知（例如参见 Aalen 等人(2008年)；Therneau 和 Grambsch, (2000年)），但这里仍明确推导该分布，以简化后续证明。证明的关键在于应用泛函Delta方法（例如参见 van der Vaart (1998年)；van der Vaart 和 Wellner (1996年)）。

由于 $\mathbb{H}_n^{uc}$和 $\mathbb{H}_n$ 是 $1\{T_i \wedge C_i \le t\}$ 和 $\Delta_i 1\{T_i \wedge C_i \le t\}$ 的经验分布函数，对$t \ge 0$而言，主要的经验中心极限定理（Donsker 定理，例如参见 van der Vaart (1998年)；van der Vaart 和 Wellner (1996年)）给出

$$\sqrt{n}\begin{pmatrix} \mathbb{H}_n^{uc} - H^{uc} \\ \mathbb{H}_n - H \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbb{G}^{uc} \\ \mathbb{G} \end{pmatrix} \qquad \text{in } (D[0,\tau])^2,$$

其中，$D[0,\tau]$ 是 $[0,\tau]$ 处的 Skorokhod 空间，其中 $\tau > 0$ 为某个值，使得 $(1-F(\tau))(1-G(\tau)) > 0$（例如参见 van der Vaart（1998）例 20.15；van der Vaart 和 Wellner（1996）例 3.9.19）。极限 $(\mathbb{G}^{uc}, \mathbb{G})$ 是一个紧的、均值为零的高斯过程，其协方差结构为

$$\begin{aligned} \mathrm{E}\mathbb{G}^{uc}(s)\mathbb{G}^{uc}(t) &= H^{uc}(s \wedge t) - H^{uc}(s)H^{uc}(t) \\ \mathrm{E}\mathbb{G}(s)\mathbb{G}(t) &= H(s \wedge t) - H(s)H(t) \\ \mathrm{E}\mathbb{G}^{uc}(s)\mathbb{G}(t) &= H^{uc}(s \wedge t) - H^{uc}(s)H(t). \end{aligned} \tag{12}$$

由式（5）可见，$S(t_p)$ 和 $\hat{S}_n(t_p)$ 分别通过以下映射序列依赖于这两组数据对 (H^{uc}, H) 和 $(\mathbb{H}_n^{uc}, \mathbb{H}_n)$，

$$(A,B) \mapsto \Big(A, \frac{1}{1-B}\Big) \mapsto \int_0^{t_p} \frac{\mathrm{d}A}{1-B} \mapsto \exp\Big(-\int_0^{t_p} \frac{\mathrm{d}A}{1-B}\Big). \tag{13}$$

这些映射的复合记为z_t:

$$z_{t_p}(H^{uc}, H) = S(t_p) \qquad z_{t_p}(\mathbb{H}_n^{uc}, \mathbb{H}_n) = \hat{S}(t_p). \tag{14}$$

复合映射 z_{t_p} 在 $\mathscr{D}$ 上是 Hadamard可微的：$= \{(A,B) : \int |\mathrm{d}A| \le M \ 1 - B \ge \varepsilon\}$，其中给定值为 M和$\varepsilon > 0$；对于每一对满足 $1/(1-B)$ 有界变差的 (A,B)，该结论均成立。如果 t_p 位于区间 $[0,\tau]$ 内，且 $H(\tau) < 1$，则过程对 $(\mathbb{H}_n^{uc}, \mathbb{H}_n)$ 以趋近于 1 的概率包含于 $\mathscr{D}$，且 $M \ge 1$ 和 ε 足够小（van der Vaart（1996），第 384 页）。式（13）中该复合映射的导数映射为

$$(\alpha,\beta) \mapsto -\exp\Big(-\int \frac{\mathrm{d}A}{1-B}\Big)\Big(\int \frac{\mathrm{d}\alpha}{1-B} + \int \frac{\beta \mathrm{d}A}{(1-B)^2}\Big).$$

根据泛函Delta方法，可得

$$\begin{aligned} \sqrt{n}(\hat{S}_n(.) - S(.)) \rightsquigarrow & -\exp\Big(-\int \frac{\mathrm{d}H^{uc}}{1-H}\Big)\Big(\int \frac{\mathrm{d}\mathbb{G}^{uc}}{1-H} + \int \frac{\mathbb{G}\mathrm{d}H^{uc}}{(1-H)^2}\Big) \qquad \text{in } D[0,\tau] \\ = & -S(.)\Big(\int \frac{\mathrm{d}\mathbb{G}^{uc}}{1-H} + \int \frac{\mathbb{G}\mathrm{d}H^{uc}}{(1-H)^2}\Big). \end{aligned}$$

得出结论：

$$\sqrt{n}(\hat{S}_n(t_p) - S(t_p)) \rightsquigarrow -S(t_p)\Big(\int_0^{t_p} \frac{\mathrm{d}\mathbb{G}^{uc}}{1-H} + \int_0^{t_p} \frac{\mathbb{G}\mathrm{d}H^{uc}}{(1-H)^2}\Big), \tag{15}$$

其中， $\mathbb{G}^{uc}$ 和 $\mathbb{G}$ 为前文所述的高斯过程。接下来，我们将考虑这一极限分布。首先，定义过程

$$\mathbb{M}_{t_p}(s) \;=\; -S(t_p)\Big(\mathbb{G}^{uc}(s)+\int_0^s \mathbb{G}\,\mathrm{d}\Lambda\Big) \;=\; -S(t_p)\mathbb{M}^{\star}(s),$$

其中 $\mathbb{M}^{\star}(s) = \mathbb{G}^{uc}(s)+\int_{s_0}\mathbb{G}\,\mathrm{d}\Lambda$。过程 $\mathbb{M}^{\star}$ 是一个零均值高斯鞅，其协方差函数为

$$\mathrm{E}\mathbb{M}^{\star}(s)\mathbb{M}^{\star}(t) = \int_0^{s\wedge t}(1-H)\mathrm{d}\Lambda$$

（参见 van der Vaart（1996），例 3.9.19）。因此， $\mathbb{M}_{t_p}(.) = -S(t_p)\mathbb{M}^{\star}(.)$ 是一个零均值高斯鞅，其协方差函数等于

$$\mathrm{E}\mathbb{M}_{t_p}(s)\mathbb{M}_{t_p}(t) = S(t_p)^2\int_0^{s\wedge t}(1-H)\mathrm{d}\Lambda.$$

式（15）中的极限变量可表示为 $\int_{[0,t_p]}(1-H)^{-1}\mathrm{d}\mathbb{M}_{t_p}$。利用鞅论可以证明，$\int_{[0,t_p]}(1-H)^{-1}\mathrm{d}\mathbb{M}_{t_p}$ 是一个具有独立增量的零均值高斯过程，其方差为

$$S(t_p)^2\int_0^{t_p}\frac{\mathrm{d}\Lambda(s)}{(1-F(s))(1-G(s))} \;=\; S(t_p)^2\int_0^{t_p}\frac{\mathrm{d}\Lambda(s)}{1-H(s)}.$$

综合这些结果可得

$$\sqrt{n}(\hat{S}_n(t_p)-S(t_p)) \rightsquigarrow \mathcal{N}\Big(0, S(t_p)^2\int_0^{t_p}\frac{\mathrm{d}\Lambda(s)}{1-H(s)}\Big).$$

第二项的渐近分布： $\sqrt{n}(S(\hat{}t_{p,n})-S(t_p))$

文献中已有样本分位数渐近分布的相关结果。通过应用 Delta方法，可得 $\sqrt{n}(S(\hat{}t_{p,n})-S(t_p))$ 的渐近分布。这里给出了证明，以便后续确定两项之和的渐近分布时使用其中的部分内容。

注意，$\hat{}t_{p,n}$ 和 t_p 分别是子分布 $\mathbb{H}_n^{uc}$ 和 H^{uc} 的第 p 个样本分位数。对于固定值 p，将累积（子）分布函数映射为其第 p 个分位数的映射，在分布函数域 Φ 处是Hadamard可微的；这些分布函数在 $\Phi^{-1}(p)$ 处可微，且其导数 $\phi(\Phi^{-1}(p))$ 严格为正（van der Vaart (1996)，例3.9.21）。由于 $\sqrt{n}(\mathbb{H}_n^{uc}-H^{uc})$ 在 $D(\mathbb{R})$ 中收敛于高斯过程，泛函Delta方法得到

$$\sqrt{n}(\hat{t}_{p,n}-t_p) \;=\; \sqrt{n}(\mathbb{H}_n^{uc,-1}(p)-H^{uc,-1}(p)) \;\rightsquigarrow\; -\frac{\mathbb{G}^{uc}(t_p)}{h^{uc}(t_p)}, \tag{16}$$

当 $n\to\infty$ 趋于无穷时，这是一个均值为零、方差为

$$\frac{p(1-p)}{(h^{uc}(t_p))^2}.$$

在生存函数 $S = 1-F$ 关于 t_p 可微的假设下，德尔塔方法（以及链式法则）得到，其中 A_{t_p} 定义为 $f(t_p)/h^{uc}(t_p)$，

$$\sqrt{n}(S(\hat{t}_{p,n})-S(t_p)) \;\rightsquigarrow\; \frac{f(t_p)}{h^{uc}(t_p)}\mathbb{G}^{uc}(t_p) = A_{t_p}\mathbb{G}^{uc}(t_p), \tag{17}$$

一个均值为零、方差为 $^2p(1-p)/(h^{uc}(t_p))^2$ 的正态分布。

为准备证明的下一部分，定义这些映射的复合如下：

$$v_s(\mu_1,\mu_2) \;=\; S(\mu_1^{-1}(H^{uc}(s))), \tag{18}$$

对于 μ_1 单调非降函数 $[0,1]\to\mathbb{R}^+$。反函数 μ_1^{-1}定义为 $\mu_1^{-1}(s)=\inf\{u:\mu_1(u)\geq s\}$。根据此定义，

$$v_{t_p}(H^{uc},H)=S(H^{uc,-1}(H^{uc}(t_p)))=S(t_p),$$

$$v_{t_p}(\mathbb{H}_n^{uc},\mathbb{H}_n)=S(\mathbb{H}_n^{uc,-1}(H^{uc}(t_p)))=S(\hat{t}_{p,n}).$$

请注意，v_s 中的第二个参数并未使用，但为了简化证明其余部分的记号，仍将其纳入定义中。

第三项的收敛：$\sqrt{n}(\hat{S}_n(\hat{t}_{p,n})-\overline{S}(\hat{t}_{p,n})-(\hat{S}_n(t_p)-S(t_p)))$

仍需证明随机变量序列 $\sqrt{n}(\hat{S}_n(\hat{t}_{p,n})-S(\hat{t}_{p,n})-(\hat{S}_n(t_p)-S(t_p)))$ 依概率收敛至零。此前已证明，$\sqrt{n}(\hat{S}_n(.)-S(.))$ 在 $D[0,\tau]$ 中依分布收敛于一个具有连续样本路径的零均值高斯过程（假设 $t\to\Lambda(t)$ 连续）。根据这一连续性以及 $\hat{t}_{p,n}$ 依概率收敛于 t_p 这一事实，可知前述显示中的项依概率收敛至零。

合并式 (11) 中三项的收敛性

(11) 中的第三项依概率收敛于零。因此，$\sqrt{n}(\hat{S}_n(\hat{t}_{p,\,n})-S(t_p))$ 的渐近分布由 (11) 中第一项和第二项之和决定。前文分别讨论了这些项的渐近分布，现在需要将它们结合起来。

定义 $\phi_t(x,y)=z_t(x,y)+v_t(x,y)$，其中 z_t和 v_t如 (14) 和 (18) 中所定义。那么，$\phi_{t_p}(H^{uc},H)=z^{t_p}(H^{uc},H)+v^{t_p}(H^{uc},H)=2S(t_p)$ 且 $\phi_{t_p}(\mathbb{H}_n^{uc},\mathbb{H}_n)=\hat{S}_n(t_p)+S(\hat{t}_{p,n})$。这意味着

$$\sqrt{n}(\hat{S}_n(t_p)-S(t_p))+\sqrt{n}(S(\hat{t}_{p,n})-S(t_p))=\sqrt{n}(\phi_{t_p}(\mathbb{H}_n^{uc},\mathbb{H}_n)-\phi_{t_p}(H^{uc},H)).$$

映射 ϕ_{t_p} 是 Hadamard可微的，因为映射 z^{t_p} 和 v^{t_p} 都具有这一性质。导数 ϕ_{t_p} 等于z^{t_p} 和 v^{t_p} 的导数之和，根据泛函Delta方法

$$\begin{aligned}
&\sqrt{n}(\hat{S}_n(t_p)-S(t_p))+\sqrt{n}(S(\hat{t}_{p,n})-S(t_p))\\
&\rightsquigarrow -S(t_p)\Big(\int_0^{t_p}\frac{d\mathbb{G}^{uc}}{1-H}+\int_0^{t_p}\frac{\mathbb{G}dH^{uc}}{(1-H)^2}\Big)+\frac{f(t_p)}{h^{uc}(t_p)}\mathbb{G}^{uc}(t_p)\qquad \text{in } D[0,\tau]\\
&=\int_0^{t_p}\frac{d\mathbb{M}_{t_p}}{1-H}+A_{t_p}\mathbb{G}^{uc}(t_p)=-S(t_p)\int_0^{t_p}\frac{d\mathbb{M}^\star}{1-H}+A_{t_p}\mathbb{G}^{uc}(t_p),
\end{aligned}\tag{19}$$

零均值高斯分布，其中 $A^{t_p}=f(t_p)/h^{uc}(t_p)$ 和 $\mathbb{M}_{t_p},\mathbb{M}^\star$，且 $\mathbb{G}^{uc}$如上所定义。这是通过结合这两项的渐近理论结果得出的。极限分布的方差等于

$$\begin{aligned}
&\text{var}\left(-S(t_p)\int_0^{t_p}\frac{d\mathbb{M}^\star}{1-H}+A_{t_p}\mathbb{G}^{uc}(t_p)\right)\\
&=\text{var}\left(S(t_p)\int_0^{t_p}\frac{d\mathbb{M}^\star}{1-H}\right)+\text{var}\left(A_{t_p}\mathbb{G}^{uc}(t_p)\right)-2\text{cov}\left(S(t_p)\int_0^{t_p}\frac{d\mathbb{M}^\star}{1-H},A_{t_p}\mathbb{G}^{uc}(t_p)\right)\\
&=S(t_p)^2\int_0^{t_p}\frac{d\Lambda}{1-H}+A_{t_p}^2p(1-p)-2S(t_p)A_{t_p}\text{cov}\left(\int_0^{t_p}\frac{d\mathbb{M}^\star}{1-H},\mathbb{G}^{uc}(t_p)\right),
\end{aligned}$$

with

$$\text{cov}\left(\int_0^{t_p}\frac{d\mathbb{M}^\star}{1-H},\mathbb{G}^{uc}(t_p)\right)=(1-p)\Lambda(t_p)+\int_0^{t_p}\frac{(H^{uc}(s)-H^{uc}(t_p)H(s))d\Lambda(s)}{1-H(s)}.\tag{20}$$

协方差的这个表达式由此得出

$$\begin{aligned}
\text{cov}\left(\int_0^{t_p}\frac{d\mathbb{M}^\star}{1-H},\mathbb{G}^{uc}(t_p)\right)&=\text{cov}\left(\int_0^{t_p}\frac{d\left(\mathbb{G}^{uc}(s)+\int_0^s\mathbb{G}\,d\Lambda\right)}{1-H(s)},\mathbb{G}^{uc}(t_p)\right)\\
&=\text{cov}\left(\int_0^{t_p}\frac{d\mathbb{G}^{uc}(s)}{1-H(s)},\mathbb{G}^{uc}(t_p)\right)+\text{cov}\left(\int_0^{t_p}\frac{\mathbb{G}(s)d\Lambda(s)}{1-H(s)},\mathbb{G}^{uc}(t_p)\right)
\end{aligned}$$

with

$$\operatorname{cov}\left(\int_0^{t_p}\frac{\mathbb{G}(s)\mathrm{d}\Lambda(s)}{1-H(s)},\mathbb{G}^{uc}(t_p)\right)=\int_0^{t_p}\frac{\mathrm{E}\mathbb{G}^{uc}(t_p)\mathbb{G}(s)\mathrm{d}\Lambda(s)}{1-H(s)}=\int_0^{t_p}\frac{(H^{uc}(s)-H^{uc}(t_p)H(s))\mathrm{d}\Lambda(s)}{1-H(s)}$$

并且对于 $0=s0,\ldots,sK_\varepsilon=tp$，在区间 $[0,t_p]$上，宽度为 $\varepsilon>0$ 的网格

$$\begin{aligned}\operatorname{cov}\left(\int_0^{t_p}\frac{\mathrm{d}\mathbb{G}^{uc}(s)}{1-H(s)},\mathbb{G}^{uc}(t_p)\right)&=\int_0^{t_p}\frac{\mathrm{E}\mathbb{G}^{uc}(t_p)\mathrm{d}\mathbb{G}^{uc}(s)}{1-H(s)}\\&=\lim_{\varepsilon\downarrow 0}\sum_{i=1}^{K_\varepsilon}\frac{\mathrm{E}\mathbb{G}^{uc}(t_p)(\mathbb{G}^{uc}(s_i+\varepsilon)-\mathbb{G}^{uc}(s_i))}{1-H(s_i)}\\&=\lim_{\varepsilon\downarrow 0}\sum_{i=1}^{K_\varepsilon}\frac{(H^{uc}(s_i+\varepsilon)-H^{uc}(s_i))(1-H^{uc}(t_p))}{1-H(s_i)}\\&=(1-H^{uc}(t_p))\int_{[0,t_p]}\frac{\mathrm{d}H^{uc}(s)}{1-H(s)}=(1-p)\Lambda(t_p).\end{aligned}$$

因此，极限分布是均值为零、方差为 σ^2 的高斯分布，如式（10）所示。

证明：Kaplan-Meier估计量， $\tilde{S}_n(\hat{t}_{p,n})$

Kaplan-Meier估计量的渐近分布可以用与推导Breslow估计量相似的方法得到。同样，$\sqrt{n}(\tilde{S}_n(\hat{t}_{p,n})-S(t_p))$可写成三个项之和，如式（11）所示。对于分解中的第一项，在固定点t_p 处，Kaplan-Meier估计量的渐近分布与在t_p处求得的Breslow估计量的渐近分布完全相同。式（11）中第二项的渐近分布也与定理的第一部分完全相同，因为它不取决于S的估计量选择。证明的其余部分与Breslow估计量的证明完全相同。■

附录 B：定理1的证明

在本附录中，我们再次考虑双臂试验。试验在日历时间中的 $\hat{t}_{p,n+m}$ 时停止，即一旦p100%的患者经历了事件。这个随机时刻取决于两组中的观测值。下面给出定理1的证明。该证明将遵循附录A中证明所采用的相同思路。不过，还存在一些额外的复杂性。

定理1

设$S_{A,n}$ 和$S_{A,n}$ 分别为S_A的Breslow估计量和Kaplan-Meier估计量，它们基于在时间点$\hat{t}_{p,n+m}$之前进入研究的n 名A 组患者的观测值确定，其中$\hat{t}_{p,n+m}$是经验分布$\mathbb{H}^{uc,\star}{}_{mix,n+m}$的$p$分位数。令 $\delta>0$ 为预先指定的，$\hat{t}^{\delta}_{p,n+m}=\hat{t}_{p,n+m}-\delta$及其极限$t_p^\delta=t_p-\delta$。Breslow估计量和Kaplan-Meier估计量的渐近分布为

$$\sqrt{n}(\hat{S}_{A,n}(\hat{t}^{\delta}_{p,n+m})-S_A(t_p^\delta))\rightsquigarrow\mathcal{N}(0,\sigma^2_{A,\delta}),$$

$$\sqrt{n}(\tilde{S}_{A,n}(\hat{t}^{\delta}_{p,n+m})-S_A(t_p^\delta))\rightsquigarrow\mathcal{N}(0,\sigma^2_{A,\delta}),$$

（其中 $\rightsquigarrow$ 表示依分布收敛）作为 $n,m\to\infty$ 和

$$\begin{aligned}\sigma^2_{A,\delta}=S_A(t_p^\delta)^2\int_0^{t_p^\delta}\frac{\mathrm{d}\Lambda_A(s)}{1-H_{A,t_p}(s)}+\frac{q_Af_A(t_p^\delta)^2}{(h^{uc,\star}_{mix,t_p}(t_p))^2}p(1-p)\\-2S_A(t_p^\delta)\frac{q_A\sqrt{q_A}f_A(t_p^\delta)}{h^{uc,\star}_{mix,t_p}(t_p)}\left((1-H^{uc,\star}_{A,t_p}(t_p))\Lambda_A(t_p^\delta)+\int_0^{t_p^\delta}\frac{(H^{uc}_{A,t_p}(s)-H^{uc,\star}_{A,t_p}(t_p)H_{A,t_p}(s))\mathrm{d}\Lambda_A(s)}{1-H_{A,t_p}(s)}\right).\end{aligned}$$

证明：Breslow估计量， $\hat{S}_{A,n}(\hat{t}_{p,n+m})$

首先，序列 $\sqrt{n}(\hat{S}_{A,n}(\hat{t}^{\delta}_{p,n+m}) - S_A(t^{\delta}_p))$ 被分解为三项之和：第一项涉及在固定点 t^{δ}_p处对S_A 的估计，第二项涉及通过$S_A(\hat{t}^{\delta}_{p,n+m})$ 对时间$S_A(t^{\delta}_p)$的估计，第三项依概率收敛于零：

$$\begin{aligned}\sqrt{n}(\hat{S}_{A,n}(\hat{t}^{\delta}_{p,n+m}) - S_A(t^{\delta}_p)) = & \sqrt{n}(\hat{S}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p)) + \sqrt{n}(S_A(\hat{t}^{\delta}_{p,n+m}) - S_A(t^{\delta}_p)) \\ & + \sqrt{n}((\hat{S}_{A,n}(\hat{t}^{\delta}_{p,n+m}) - S_A(\hat{t}^{\delta}_{p,n+m})) - (\hat{S}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p))). \end{aligned} \tag{21}$$

推导项 $\sqrt{n}(\hat{S}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p))^2 \int$ 的渐近分布时，主要困难在于进行期中分析时，$\hat{t}_{p,n+m}$ 是随机的，并且是所有观测值的函数。为处理这一问题，取 $\varepsilon > 0$，使得 $0 < \varepsilon < \delta$。$t_p - \varepsilon$ 的值是固定且非随机的，并且（暂时）视为新的“研究结束”时点。此时，C_A 的分布变为 $G^{t_p-\varepsilon}$。可能出现 $G_{Acc}(t_p - \varepsilon) < 1$ 的情况，这意味着期中分析时可能并非所有患者都已进入试验，从而使样本量变为随机。可以通过重新定义 G_{Acc} 来避免这一问题，使其在 $t_p - \varepsilon$ 处有一个点质量，大小为 $1 - G_{Acc}(t_p - \varepsilon)$，并假设 $G^{t_p-\varepsilon}$ 在 0 处也有一个点质量，大小同样为 $1 - G_{Acc}(t_p - \varepsilon)$。这意味着可能有一定比例的患者在试验停止时进入研究，其随访时间为 0。通过这一重新定义，样本量再次固定为 n，但可能有一定比例的患者无法为生存曲线提供任何信息（随访时间为 0）。据此，可沿用附录A证明中的相同思路，推导 $\hat{S}_{A,n}(\hat{t}^{\delta}_p)$ 的渐近分布。由此可得，在 $t_p - \varepsilon$ 处删失时，式（21）中第一项的渐近分布为渐近正态分布，均值为零，方差为 $S_A(t^{\delta}_p)^2 \int_0^{t_p} (1 - H_{A,t_p-\varepsilon})^{-1} d\Lambda_A$。对新的终点 $t_p + \varepsilon$ 也可采用相同推理，得到渐近正态分布，其均值为零，方差为 $S_A(t^{\delta}_p)^2 \int_0^{t_p} (1 - H_{A,t_p+\varepsilon})^{-1} d\Lambda_A$。接下来请注意，$1 - H_{A,t_p+\varepsilon}(s) = S_A(s) G_{Acc}(t_p + \varepsilon - s)$ 随 ε 增大，而渐近方差随 ε 增大而减小。这符合预期，因为随访时间越长，信息越多，渐近方差也越小。对于 $\hat{S}^{t_p-\varepsilon}_{A,n}$ 和 $\hat{S}^{t_p+\varepsilon}_{A,n}$，它们分别是基于在 $t_p - \varepsilon$ 和 $t_p + \varepsilon$ 处删失的数据计算得到的 Breslow 估计量，并且对于 $t > 0$

$$\begin{aligned}
&\lim_{n\to\infty} P(\sqrt{n}(\hat{S}^{t_p-\varepsilon}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p)) \le t) \\
&= \lim_{n,m\to\infty} P(\sqrt{n}(\hat{S}^{t_p-\varepsilon}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p)) \le t, t_p - \varepsilon \le \hat{t}_{p,n+m} \le t_p + \varepsilon) \\
&\le \liminf_{n,m\to\infty} P(\sqrt{n}(\hat{S}^{\hat{t}_{p,n+m}}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p)) \le t, t_p - \varepsilon \le \hat{t}_{p,n+m} \le t_p + \varepsilon) \\
&\le \limsup_{n,m\to\infty} P(\sqrt{n}(\hat{S}^{\hat{t}_{p,n+m}}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p)) \le t, t_p - \varepsilon \le \hat{t}_{p,n+m} \le t_p + \varepsilon) \\
&\le \lim_{n,m\to\infty} P(\sqrt{n}(\hat{S}^{t_p+\varepsilon}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p)) \le t, t_p - \varepsilon \le \hat{t}_{p,n+m} \le t_p + \varepsilon) \\
&= \lim_{n\to\infty} P(\sqrt{n}(\hat{S}^{t_p+\varepsilon}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p)) \le t),
\end{aligned}$$

由于 ε 与渐近方差之间的关系。对于 $t < 0$，同样的推理适用，只是将不等号反向。现在令 ε 递减至零，从而得出

$$\sqrt{n}\left(\hat{S}^{\hat{t}_{p,n+m}}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p)\right) \rightsquigarrow \mathcal{N}\left(0, S_A(t^{\delta}_p)^2 \int_0^{t^{\delta}_p} (1 - H_{A,t_p})^{-1} d\Lambda_A\right).$$

对于式（21）中第二项渐近分布的推导，可以沿用附录A中的证明思路。矩 $\hat{t}_{p,n+m}$ 是合并样本的经验分布函数 $\mathbb{H}^{uc}_{,\star\, mix}$的分位数，样本量为$n + m$。对于$q_A = n/(n + m)$，记

$$\begin{aligned}\sqrt{n}(S_A(\hat{t}^{\delta}_{p,n+m}) - S_A(t^{\delta}_p)) = & -\sqrt{q_A}\sqrt{n+m}(\hat{t}^{\delta}_{p,n+m} - t^{\delta}_p) f_A(t^{\delta}_p) + o_P(1) \\ = & -\sqrt{q_A}\sqrt{n+m}(\hat{t}_{p,n+m} - t_p) f_A(t^{\delta}_p) + o_P(1)\end{aligned}$$

其中，第一个等式直接由一次泰勒展开（或德尔塔方法）得出。最后一项渐近正态，均值为零，方差等于$\sigma^2_{A,\delta}$中的第二项，正如附录A的证明中所示。因此：

$$\sqrt{n}(S_A(\hat{t}^{\delta}_{p,n+m}) - S_A(t^{\delta}_p)) \rightsquigarrow \sqrt{q_A}\frac{f_A(t^{\delta}_p)}{h^{uc,\star}_{mix,t_p}(t_p)}\mathbb{G}^{uc,\star}_{mix}(t_p),$$

其中 $\mathbb{G}^{uc,\ \star\text{混合分布}}(t\quad p)$ 是均值为零、方差为 $p(1-p)$ 的高斯过程（同前）。

(21) 中的第三项依概率收敛于零，这一点已在附录A中证明。

现在，我们需要再次将这三项的渐近分布合并起来。由于第二项的推导基于混合分布，因此还需要一些额外的论证。我们已经看到，与附录A中的证明类似，

$$\sqrt{n}(\hat{S}_{A,n}(t^{\delta}_p) - S_A(t^{\delta}_p)) + \sqrt{n}(S_A(\hat{t}^{\delta}_{p,n+m}) - S_A(t^{\delta}_p)) \rightsquigarrow -S_A(t^{\delta}_p)\int_0^{t^{\delta}_p}\frac{\mathrm{d}\mathbb{M}^{\star}}{1-H_{A,t_p}} + \sqrt{q_A}\frac{f_A(t^{\delta}_p)}{h^{uc,\star}_{mix,t_p}(t_p)}\mathbb{G}^{uc,\star}_{mix}(t_p),$$

一个均值为零的高斯分布（见附录A中的证明），其中 $\mathbb{M}^{\star}_{t_p}$ 和 $\mathbb{G}^{uc,\ \star\text{混合分布}}$ 如前所定义。该分布的方差为

$$\begin{aligned}
&\mathrm{var}\left(-S_A(t^{\delta}_p)\int_0^{t^{\delta}_p}\frac{\mathrm{d}\mathbb{M}^{\star}}{1-H_{A,t_p}} + \sqrt{q_A}\frac{f_A(t^{\delta}_p)}{h^{uc,\star}_{mix,t_p}(t_p)}\mathbb{G}^{uc,\star}_{mix}(t_p)\right)\\
&= \mathrm{var}\left(S_A(t^{\delta}_p)\int_0^{t^{\delta}_p}\frac{\mathrm{d}\mathbb{M}^{\star}}{1-H_{A,t_p}}\right) + \mathrm{var}\left(\frac{\sqrt{q_A}f_A(t^{\delta}_p)}{h^{uc,\star}_{mix,t_p}(t_p)}\mathbb{G}^{uc,\star}_{mix}(t_p)\right)\\
&\quad - 2\mathrm{cov}\left(S_A(t^{\delta}_p)\int_0^{t^{\delta}_p}\frac{\mathrm{d}\mathbb{M}^{\star}}{1-H_{A,t_p}}, \frac{\sqrt{q_A}f_A(t^{\delta}_p)}{h^{uc,\star}_{mix,t_p}(t_p)}\mathbb{G}^{uc,\star}_{mix}(t_p)\right)\\
&= S_A(t^{\delta}_p)^2\int_0^{t^{\delta}_p}\frac{\mathrm{d}\Lambda_A}{1-H_{A,t_p}} + q_A\Big(\frac{f_A(t^{\delta}_p)}{h^{uc,\star}_{mix,t_p}(t_p)}\Big)^2 p(1-p)\\
&\quad - 2\sqrt{q_A}\frac{S_A(t^{\delta}_p)f_A(t^{\delta}_p)}{h^{uc,\star}_{mix,t_p}(t_p)}\,\mathrm{cov}\left(\int_0^{t^{\delta}_p}\frac{\mathrm{d}\mathbb{M}^{\star}}{1-H_{A,t_p}}, \mathbb{G}^{uc,\star}_{mix}(t_p)\right),
\end{aligned}$$

with

$$\begin{aligned}
&\mathrm{cov}\left(\int_0^{t^{\delta}_p}\frac{\mathrm{d}\mathbb{M}^{\star}}{1-H_{A,t_p}}, \mathbb{G}^{uc,\star}_{mix}(t_p)\right) = \mathrm{cov}\left(\int_0^{t^{\delta}_p}\frac{\mathrm{d}\big(\mathbb{G}^{uc}_A(s) + \int_0^s \mathbb{G}_A\,\mathrm{d}\Lambda\big)}{1-H_{A,t_p}(s)}, \mathbb{G}^{uc,\star}_{mix}(t_p)\right)\\
&= \mathrm{cov}\left(\int_0^{t^{\delta}_p}\frac{\mathrm{d}\mathbb{G}^{uc}_A(s)}{1-H_{A,t_p}(s)}, \mathbb{G}^{uc,\star}_{mix}(t_p)\right) + \mathrm{cov}\left(\int_0^{t^{\delta}_p}\frac{\mathbb{G}_A(s)\mathrm{d}\Lambda(s)}{1-H_{A,t_p}(s)}, \mathbb{G}^{uc,\star}_{mix}(t_p)\right)\\
&= q_A\mathrm{cov}\left(\int_0^{t^{\delta}_p}\frac{\mathrm{d}\mathbb{G}^{uc}_A(s)}{1-H_{A,t_p}(s)}, \mathbb{G}^{uc,\star}_A(t_p)\right) + q_A\mathrm{cov}\left(\int_0^{t^{\delta}_p}\frac{\mathbb{G}_A(s)\mathrm{d}\Lambda(s)}{1-H_{A,t_p}(s)}, \mathbb{G}^{uc,\star}_A(t_p)\right),
\end{aligned}$$

其中，最后一个等式源于两组观测值之间相互独立。上一公式最后一行中的协方差等于

$$\begin{aligned}
\mathrm{cov}\left(\int_0^{t^{\delta}_p}\frac{\mathbb{G}_A(s)\mathrm{d}\Lambda_A(s)}{1-H_{A,t_p}(s)}, \mathbb{G}^{uc,\star}_A(t_p)\right) &= \int_0^{t^{\delta}_p}\frac{\mathrm{E}\mathbb{G}^{uc,\star}_A(t_p)\mathbb{G}_A(s)\mathrm{d}\Lambda_A(s)}{1-H_{A,t_p}(s)}\\
&= \int_0^{t^{\delta}_p}\frac{(H^{uc}_{A,t_p}(s) - H^{uc,\star}_{A,t_p}(t_p)H_{A,t_p}(s))\mathrm{d}\Lambda_A(s)}{1-H_{A,t_p}(s)}
\end{aligned}$$

采用与附录A类似的方法，可得

$$\operatorname{cov}\left(\int_0^{t_p^\delta} \frac{\mathrm{d}\mathbb{G}_A^{uc}(s)}{1-H_{A,t_p}(s)}, \mathbb{G}_A^{uc,\star}(t_p)\right) = (1-H_{A,t_p}^{uc,\star}(t_p))\Lambda_A(t_p^\delta).$$

证明：Kaplan-Meier估计量， $\tilde{S}_{A,n}(\hat{t}_{p,n+m})$

这可以像 Breslow 估计量那样在同一行完成。 ■