arXiv:2603.01033v1 [stat.ME] 1 Mar 2026

# 解读净生存：我们所估计的与我们以为自己所估计的

Matthew J. Smith[1]

[1]医学统计学系，伦敦卫生与热带医学学院，英国伦敦，WC1E 7HT

**摘要**

净生存率通常被定义为“假如癌症是唯一可能的死因时的生存”，该目标估计量仅对应于癌症特异性死亡率。Pohar Perme估计量旨在通过从观察到的总死亡率中去除一般人群其他原因死亡率来估计该量，但仅当癌症患者的其他原因死亡率与一般人群相同时才能实现。然而，癌症患者往往因基线健康差异和治疗所致效应而具有更高的其他原因死亡率。借助近期将总死亡率分解为四个组成部分（癌症死亡、基线健康差异、治疗所致其他原因死亡以及一般人群其他原因死亡率）的理论工作，我们表明，Pohar Perme估计量所给出的是癌症死亡、基线差异与治疗导致的死亡之和；只要任一超额来源存在，它便达不到其预期的目标估计量。依据Botta *et al*，我们提供实证证据，表明其他原因死亡的相对风险介于1.0（结直肠癌）与4.0+（头颈部癌症）之间，并且计算显示，当相对风险超过1.0时，净生存率可能大幅低估癌症特异性生存概率。至关重要的是，治疗所致其他原因死亡代表从癌症通向死亡的不可约减因果路径，无法通过更好的分层予以消除。我们建议将净生存率解读为“已去除一般人群其他原因死亡率的生存”，而不是因果反事实，并呼吁在癌症流行病学中采用更精确的表述。

**关键词：**净生存率、癌症特异性生存率、相对生存、竞争风险、因果推断、癌症流行病学

# 1 引言

癌症生存统计具有多重用途：监测人群健康、评估卫生系统、衡量治疗进展以及识别健康差异。为了在不同人群之间或不同时间进行有意义的比较，我们必须考虑与癌症无关的背景死亡率差异。净生存率正是为应对这一挑战而提出的，它提供了一种独立于其他原因死亡率差异的“纯粹”癌症生存衡量指标。

净生存率通常被解释为用于回答这样一个问题：“如果癌症是唯一可能的死因，生存会是怎样？”这一解释出现在主要的国际癌症生存研究、国家癌症报告和方法学教程中。[1] 它之所以具有吸引力，是因为它表明我们已经将癌症的效应与所有其他竞争性死因分离开来。

然而，这一解释（尽管正确陈述了预期的目标估计量）将该指标旨在估计的内容与其实际给出的结果混为一谈。Pohar Perme估计量试图通过从癌症患者观察到的总死亡率中去除一般人群其他原因死亡率，来恢复这一因果反事实。[2] 这一操作仅在特定条件下才能实现预期的目标估计量：即癌症患者经历与一般人群相同的其他原因死亡率。当该条件成立时，去除一般人群死亡率足以分离出癌症特异性死亡率。当该条件不成立时，该估计量去除的量少于所需，其结果并非“如果癌症是唯一死因时的生存”，而是仅去除其他原因死亡率中一般人群组成部分之后的生存。

当癌症患者的其他原因死亡率高于一般人群时，问题便会出现。现代癌症治疗可能导致心血管疾病、感染、继发性恶性肿瘤及其他危及生命的并发症；这些死亡在因果上可归因于癌症，却在医学编码中被列为其他原因死亡。癌症患者的基线健康状况也可能与一般人群不同，从而独立导致其他原因死亡率升高。我们认为自己估计的内容与实际估计的内容之间存在差距，这对净生存率的解释和应用具有重要影响。

Botta等人近期的研究通过将总死亡风险率分解为四个组成部分，为理解这一问题提供了理论框架。[3] 在他们的研究基础上，我们阐明 Pohar Perme估计量所提供的结果，区分可解决与不可消除的偏倚来源，展示不同癌症类型中这一问题严重程度的实证证据，并提出恰当解释净生存率的建议。

本文其余部分的结构如下。在方法部分，我们提出风险率分解框架，并引入有向无环图，以明确各组成部分的因果结构；推导相对风险作为解释失效诊断指标的作用；区分超额其他原因死亡率中可处理与不可消除的来源；并将全部四个相关生存估量目标映射到其相应的研究问题和受众。在结果部分，我们首先回顾实证证据

关于不同癌症类型间相对风险的变异，随后以一项历史性前列腺癌试验说明这四个组成部分在实践中如何运作，最后给出解析计算，量化净生存率与预期估计目标在一系列相对风险值和随访时间上的偏离。我们最后讨论对报告实践的启示，并呼吁在癌症流行病学中使用更精确的语言。

# 2 方法

## 2.1 风险分解框架

标准的相对生存框架将癌症患者的总死亡风险率分为两个组成部分：超额死亡（高于一般人群的死亡水平）和预期死亡（基于一般人群的死亡率）。净生存率的计算方法是去除预期部分，仅保留超额死亡。

Pohar Perme等人将其形式化表示为:

$$h_{\text{total}}(t) = h_{\text{excess}}(t) + h_{\text{expected}}(t) \tag{1}$$

其中， $h_{\text{expected}}(t)$表示一般人群其他原因死亡率，而 $h_{\text{excess}}(t)$表示超出这一基线的所有死亡。Pohar Perme估计量针对的是在移除 $h_{\text{expected}}(t)$这一假设情景下的生存率。

在此框架的基础上，Botta等人最近证明，超额死亡本身由三个不同的组成部分构成，从而形成四组分分解 [3]:

$$h_{\text{total}}(t) = h_A(t) + h_B(t) + h_C(t) + h_D(t) \tag{2}$$

其中:

- $h_A(t)$ 是癌症特异性死亡风险（死于肿瘤进展、转移）
- $h_B(t)$ 表示癌症患者与普通人群之间基线其他原因死亡差异，这种差异在诊断时已存在，且与癌症或其治疗无关
- $h_C(t)$ 表示治疗诱发的其他原因死亡（死于心血管疾病、感染或癌症治疗导致的其他疾病）
- $h_D(t)$ 是一般人群其他原因死亡风险（相当于上方的 $h_{\text{预期}}$ ）

In t在其框架下，Pohar Perme估计量直接给出超额死亡风险 :

$$h_{\text{excess}}(t) = h_A(t) + h_B(t) + h_C(t) \tag{3}$$

这一分解阐明，只有当两者$h_B$（$t$）$= 0$ 和 $h_C$（$t$）$= 0$时，净生存率才等于癌症特异性生存率（仅$h_A$），也就是说，癌症患者的其他原因死亡率与一般人群相同时。

这一分解还阐明了净生存目标估计量与Pohar Perme估计量实际给出的结果之间的区别。净生存目标估计量的正式定义是“如果癌症是唯一死因时的生存”（仅组分A），而Pohar Perme估计量则通过移除D组分来针对组分A；但正如公式3所示，只要B或C不为零，这一操作就会得到A+B+C。估计量执行了正确的操作；问题在于，在现实条件下，该操作不足以达到预期的目标估计量。

图 1 将这四个组成部分表示为一个有向无环图，明确展示了构成每个组分的因果路径，以及B和C组分为何在净生存率操作后仍然存在。净生存率的标准描述（即“如果癌症是唯一死因时的生存”）意味着关闭所有通向其他原因死亡的路径，只留下组分A。在实践中，Pohar Perme估计量只移除了D组分，仅关闭了从年龄／性别／日历时间通向其他原因死亡的路径。B和C组分所对应的箭头仍然有效，因此净生存率的估计值反映的是A ＋ B ＋C，而不是仅有A。只有在B和C均为零这一特殊情况下，传统解释才成立。

关键在于，我们可以通过考察癌症患者的其他原因死亡率，在经验上判断B和C组分是否为零。如果我们只单独考察其他原因死亡（排除组分A），癌症患者的其他原因死亡总风险率为：

$$h_{\text{other-cause}}(t) = h_B(t) + h_C(t) + h_D(t) \tag{4}$$

当 $h_B(t) = 0$ 和 $h_C(t) = 0$时，这等于一般人群其他原因死亡率 $h_D(t)$。当B或C组分存在时，癌症患者的其他原因死亡率会高于一般人群的死亡率。这种超额死亡率构成了经验性标志，表明对净生存率的标准解释已不再成立。

## 2.2 相对风险的作用

其他原因死亡率的相对风险（RR）量化了癌症患者的其他原因死亡率相对于一般人群的升高程度。它自然地由风险分解得出，即观察到的其他原因总死亡率与一般人群预期死亡率之比：

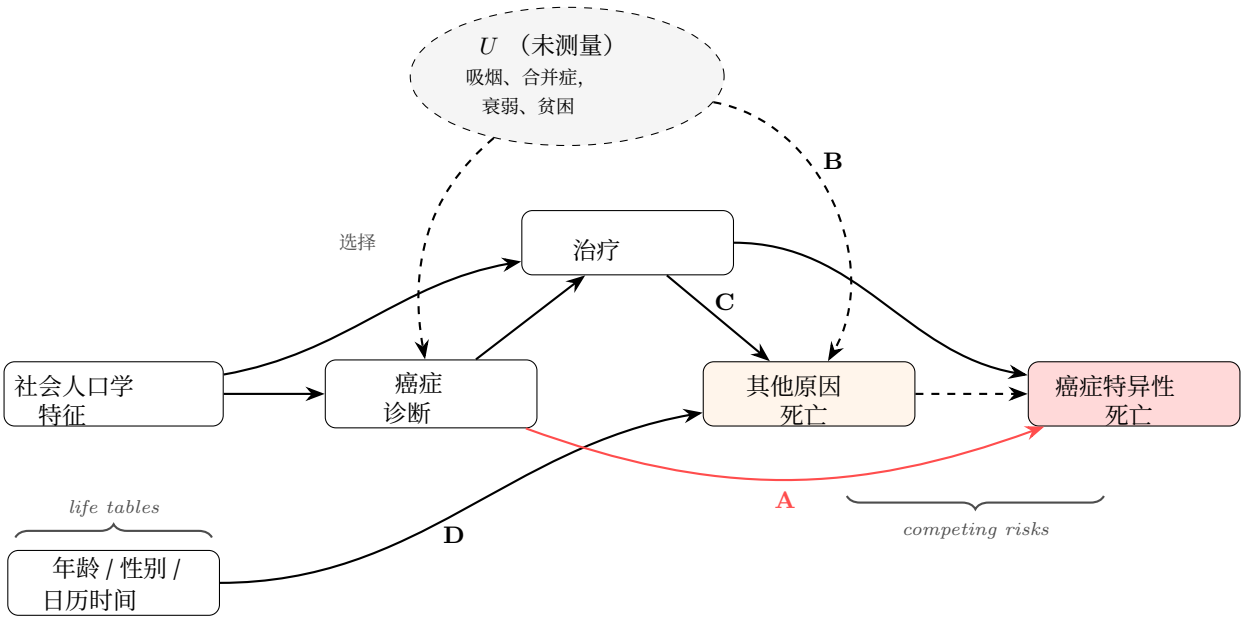


图1：四组分死亡风险分解的有向无环图。实线箭头表示直接因果效应；虚线箭头表示涉及未测量混杂的路径（$U$）。四个主要节点位于同一水平轴上。**组分A**（$h_A$，红色）：由肿瘤进展或转移导致的癌症特异性死亡（癌症诊断 →癌症特异性死亡）。**组成部分 B**（$h_B$）：由癌症患者与一般人群共有的未测量因素 $U$（吸烟、合并症、衰弱、贫困）所致的超额其他原因死亡率；这是一条原则上可通过条件化 $U$ 的代理变量、借助改进的生命表分层来处理的后门路径。**组分C**（$h_C$）：沿因果链癌症诊断 → 治疗 →其他原因死亡发生的治疗所致其他原因死亡；其不可消除，因为任何分层都无法移除因果路径上的节点。治疗也可能直接影响癌症特异性死亡。**D组分**（$h_D$）：由年龄、性别和日历时间决定的背景其他死因死亡率。从其他原因死亡指向癌症特异性死亡的箭头反映了竞争风险结构：因其他原因死亡的患者不再面临癌症特异性死亡的风险，因此其他原因死亡排除了癌症特异性死亡。下方的大括号表示这两种结局是竞争事件。Pohar Perme估计量通过匹配生命表去除D组分，但保留A、B和C。因此，仅当 $h_B = h_C = 0$时，净生存率才等于癌症特异性生存率。

$$\mathrm{RR}(t) = \frac{h_B(t) + h_C(t) + h_D(t)}{h_D(t)} \tag{5}$$

简化后为：

$$\mathrm{RR}(t) = 1 + \frac{h_B(t) + h_C(t)}{h_D(t)} \tag{6}$$

这揭　示了相对风险与净生存率解释之间的直接联系　　：

- **当相对风险 = 1.0：** 分子等于分母，这意味着 $h_B(t) + h_C(t) = 0$。B和C组分均为零。癌症患者的其他原因死亡率与一般人群相同。净生存率等于癌症特异性生存率，标准解释是正确的。

- **当相对风险>1.0：** 分子超过分母，这意味着 $h_B(t) + h_C(t) > 0$。癌症患者

的其他原因死亡率高于一般人群。净生存率包含这部分超额死亡率（A +B + C），不等同于癌症特异性生存率（仅A）。传统解释不再成立。

由于 Pohar Perme估计量保留了B和C组分带来的额外风险率，因此得出的净生存率会低于疾病特异性生存率。相对风险的大小直接表明超额其他原因死亡率的程度，因此也反映了净生存率偏离癌症特异性生存率的幅度。

Botta等人证明，相对风险在不同癌症类型之间差异显著，从结直肠癌的约1.0到头颈部癌症的4.0以上不等。[3] 这种差异表明，解释问题并非普遍存在，而是对某些癌症类型的影响远大于其他类型。此外，由于相对风险量化了B + C两者的综合影响程度，因此它提供了一个简单的经验指标，可用于判断净生存率何时偏离癌症特异性生存率。然而，单凭相对风险无法区分超额其他原因死亡率是源于基线健康差异（组成部分 B）、治疗所致效应（组分C），还是两者兼有。正如我们接下来将讨论的，这一区分十分重要：原则上，可以通过更完善的分层来处理组成部分 B，但组分C则无法如此处理。

## 2.3 组成部分 B和C的区别

虽然B和C组分都代表癌症患者中的超额其他原因死亡率，但两者在性质和影响方面存在根本差异：

**组成部分 B（基线差异）：** 这些差异在诊断时就已存在，反映了患者选择。与一般人群相比，癌症患者的吸烟率、合并症负担、社会经济地位或衰弱程度可能有所不同。重要的是，组成部分 B可通过方法学改进加以解决。根据这些特征对生命表进行更详细的分层，可以减少或消除 B。目前正在开展相关工作，以制定考虑贫困、合并症及其他因素的分层生命表。[4]

**组分C（治疗引起的死亡）：** 这些死亡源于一条因果路径：癌症诊断 → 治疗 →其他原因死亡。接受蒽环类化疗后因心力衰竭去世的患者，经历的就是治疗引起的死亡。关键在于，组分C无法通过方法学改进消除。无论如何对生命表进行分层，都无法消除 C，因为若没有癌症，这些死亡就不会发生。尽管医学编码将其归为其他原因，但从因果关系上说，这些死亡可归因于癌症。

这一区分对解释至关重要。当净生存率中包含大量组分C时，这并非只是可以通过改进方法来解决的统计假象。相反，它反映了一个根本现实：癌症治疗能够挽救患者免于死于癌症，但有时也会导致患者死于其他原因。这些治疗所致

的死亡也是癌症负担的一部分。

事实上，由于若非癌症诊断及其治疗，组分C所致的死亡便不会发生，因此这些死亡在因果上可归因于癌症，也应纳入对癌症死亡负担的完整核算。这一观察促使我们区分疾病特异性生存率（仅组分A）与疾病可归因生存率（组分A ＋ C），并将在下一节对此展开论述。

## 2.4 将估计目标对应到研究问题和受众

在癌症流行病学中，有四种相关的估计目标，各自旨在回答不同受众所关注的不同研究问题。了解某项分析得出的估计目标，以及它是否符合预期的研究问题，对于恰当解释结果至关重要。

**疾病特异性生存率（仅组分A）**反映肿瘤生长、进展和转移所致的死亡（即癌症直接导致的死亡）。临床试验研究者和肿瘤科医生在评估某种治疗能否阻止癌症进展时，主要需要这一目标估计量。死于转移性疾病的患者属于疾病特异性死亡；死于治疗诱发的心力衰竭的患者则不属于此类，即使癌症导致了该治疗。在随机对照试验中，疾病特异性生存终点旨在估计组分A，以回答这一问题：“这种治疗能否预防癌症进展导致的死亡？”这是监管批准主要关注的目标估计量。

**疾病归因生存率（组分A ＋ C）**涵盖所有可因果归因于罹患癌症的死亡，包括直接癌症死亡和治疗所致的其他原因死亡。患者因蒽环类药物诱发的心肌病而死亡，即属于疾病归因死亡，尽管死亡证明上记录的死因是心血管原因。若非癌症诊断及其治疗，这些死亡本不会发生；它们是癌症护理的真实代价。对于提出以下问题的政策制定者、卫生经济学家和公共卫生研究人员而言，这一目标估计量最为相关：“癌症给我们的人群带来了多大的完整死亡负担？”成本效益分析需要疾病归因生存率，因为必须将治疗所致死亡计入并与任何生存获益相权衡。

表1 展示了在不同组分分配情景下，净生存率与这两个临床目标之间的关系。

表1：不同情景下净生存率所估计的内容

| 情景 | 净生存率等于 | 与目标估计量的关系 |
|---|---|---|
| 无（B=0，C=0） | A | 等于疾病特异性生存率（理想情况） |
| 仅基线（B≠0，C=0） | A +B | 低估疾病特异性生存率；差值等于组成部分 B |
| 仅治疗（B=0，C≠0） | A +C | 低估疾病特异性生存率；等于疾病归因 |
| 两者均存在（B≠0、C≠0） | A + B +C | 低估疾病特异性生存率（差距 =B +C）；低估——计疾病归因生存率（差距 = B） |

**死因特异性生存率（组成部分A ＋ 部分C）**估计的是按死亡证明编码的癌症死亡

，其中通常包括组分A以及组分C的一部分，具体取决于医生在死亡时如何归因治疗并发症。在死因数据可靠且编码一致的情况下，该目标估计量与国家癌症监测项目最为相关。它所回答的问题是：“有多少患者死亡时将癌症列为根本死因？”在编码实践随时间保持稳定的单一国家或地区内，死因特异性生存率往往比净生存率更易于解释，因为它无需与人群生命表进行比较。[5]

**净生存率（组成部分A ＋ B ＋ C）**是Pohar Perme估计量从总观察死亡率中去除一般人群其他原因死亡率后所得到的量，同时保留基线健康差异和治疗导致的死亡。该目标估计量专门为国际比较以及死因数据不可靠或各国死因编码不一致的情形而开发。它所回答的问题是：“在背景死亡率不同的人群之间，癌症生存如何比较？”CONCORD和EUROCARE等主要国际研究恰当地将净生存率用于这一目的，因为即使各国之间的死因编码存在差异，它仍能实现有效比较。[6] 然而，正是这一被保留的超额部分（即B和C组分）使得净生存率无法回答本意针对疾病特异性生存或疾病归因生存率的问题。就这一目的而言（即去除背景死亡率差异，以便在预期寿命不同的人群之间对生存进行有效比较），A+B+C是恰当且有意设定的量。Pohar Perme估计量正确地提供了国际比较所需的结果。问题不在于该指标，而在于其标签：将净生存率描述为“若癌症是唯一死因时的生存”，引入的是仅属于组分A的解释；而当相对风险>1时，A+B+C并不支持这一解释。

表2 总结了全部四种估计目标及其组成部分组合、主要受众和适用场景。净生存率在解读上的核心问题不在于测量指标本身，而在于将其用于预期适用范围之外。净生存率常被用来回答有关疾病特异性生存率（组分A）或癌症总负担（组分A ＋ C）的问题，但它实际估计的是A ＋ B ＋C。当组分B占比很大时，这种差异最为显著；但即使组分C不可消除，也意味着在治疗会导致其他原因死亡的情况下，净生存率不可能等同于癌症特异性生存率。

图 2 提供了一棵决策树，将这一目标估计量映射操作化，供应用研究人员使用。从观察到的总死亡率出发，研究人员依次回答四个问题：是否已去除一般人群其他原因死亡率；其他原因死亡相对风险是否约为1.0；是否可获得可靠的死因数据；以及生命表是否已充分分层以降低组成部分 B。每个分支通向上述四个估计目标之一，并附有关于适当用途的简要说明。该决策树明确指出，对于死因编码不可靠的国际比较，净生存率是正确选择，但当

表2：将生存估计目标映射到研究问题和受众

| 目标估计量 | 组成部分 | 解释 | 主要应用场景 |
|---|---|---|---|
| 疾病特异性生存-生存 | A | 仅当疾病本身导致死亡时的生存（肿瘤进展、转移） | 临床试验、治疗疗效评估 |
| 疾病归因生存 | A + C | 计入所有因果归因于癌症的死亡后的生存——（疾病 +治疗所致） | 疾病负担研究、成本—效果分析 |
| 特定原因生存 | sur- A +（某些 C） | 基于死亡证明上将癌症编码为根本死因的生存死亡证明上的原因 | 当原因死因数据可靠 |
| 净生存率 | A +B +C | 剔除一般人群其他原因死亡后的生存率 | 国际比较，不可靠的——可进行的死因编码 |

*Note:* 组成部分定义如下：A = 癌症特异性死亡风险（肿瘤进展、转移所致死亡）；B = 诊断时癌症患者与一般人群之间的基线其他原因死亡差异；C = 治疗诱发的其他原因死亡（由癌症治疗引起的心血管疾病、感染或其他疾病所致死亡）；D = 一般人群其他原因死亡风险。净生存率通过从总观察死亡率中去除D组分计算得出。

相对风险大幅超过1.0时，研究人员应谨慎解读，并应在组分C可能较大时考虑疾病归因生存率。

# 3 实证证据与分析结果

## 3.1 不同癌症类型之间相对风险的变化

基于人群的癌症登记数据所提供的实证证据表明，不同癌症类型的其他原因死亡相对风险存在显著差异，这取决于癌症部位、诊断时年龄、性别和诊断后时间。[3] Botta等人使用应用于EUROCARE-6数据的混合治愈模型，估计了三种癌症类型的RR（表3）。

表3：按癌症类型、年龄组和性别划分的其他原因死亡相对风险估计值（Botta等人，2025年）

| 癌症类型 | Sex | 诊断时年龄 | | |
|---|---|---|---|---|
| | | 40–59 | 60–69 | 70–79 |
| 头颈部 | Male | 4.0 | 2.6 | 1.6 |
| | 女性 | 4.5 | 2.9 | 1.8 |
| 乳腺 | 女性 | 1.3 | 1.4 | 1.2 |
| 结直肠 | Male | ≈1.0（所有年龄组） | | |
| | 女性 | ≈1.0，在所有年龄组中 | | |

这些差异具有重要意义。对于结直肠癌，所有年龄和性别组的相对风险都接近1.0，这意味着净生存率与癌症特异性生存率非常接近（即B和C组分都很小）。对于头颈癌，相对风险显著升高，尤其是在年轻患者中（40–59岁女性的相对风险最高达4.5），这反映了吸烟、酒精和HPV感染等共同风险因素对其他慢性疾病的强烈影响。乳腺癌则属于中间情况，相对风险略有升高（1.2–1.4），这种升高主要在老年患者和较长随访期间才具有临床意义。

至关重要的是，相对风险不仅因癌症类型而异，也因年龄和诊断后时间而异。这意味着解释问题并非静态不变，而是会随时间恶化，尤其是对于癌症类型和患者亚组

相对风险升高者。值得注意的是，对于头颈部等癌症，年轻患者面临的解释问题比老年患者更为严重，因为他们的相对风险更高，尽管其绝对背景死亡率（$h_D$）更低。

重要的是，相对风险升高的来源因癌症类型而异。对于头颈癌，主要贡献可能来自组成部分 B（吸烟和酒精等共同风险因素会在基线时升高其他原因死亡率，且这一影响独立于治疗）。对于采用具有心脏毒性的治疗方案的癌症，组分C可能有显著贡献。由于相对风险无法区分这两种来源，而且二者的影响截然不同（组成部分 B 可通过改进分层来解决，而组分C则不能），因此，了解超额其他原因死亡率的构成与了解其大小同样重要。以下经验案例展示了这两个组分如何同时发挥作用。

## 3.2 经验分解：前列腺癌试验示例

为说明组成部分 B 和组分C在实际中的作用，我们考察了一项随机试验的数据。该试验于 1960 年至 1975 年间开展，比较大剂量雌激素治疗（5.0 mg 己烯雌酚）与安慰剂对晚期前列腺癌男性患者的疗效。[7] 这项试验为将总其他原因死亡率经验性地分解为各组成部分提供了理想契机，原因包括：(1) 治疗分配采用随机方式，能够清晰归因治疗效应；(2) 试验开展年代和患者特征便于与同期生命表进行匹配；(3) 雌激素的心血管毒性为治疗诱发的其他原因死亡提供了明确示例。

该试验纳入了125名分配至大剂量雌激素组的患者和127名分配至安慰剂组的患者，中位随访时间约为36个月。我们使用人类死亡率数据库中按年龄和日历年份匹配的生命表，计算了D组分（一般人群预期死亡率）。通过将安慰剂组和雌激素组观察到的其他原因死亡率与这些预期死亡率进行比较，估算了B和C组分。

表4 展示了分解结果。D组分为13.5%，代表该年龄男性在试验随访期间预期的其他原因死亡率。安慰剂组的其他原因死亡率为45.7%，因此B组分为32.1个百分点，相对风险为3.4（45.7%/13.5%）。这一显著的基线超额死亡反映了晚期前列腺癌男性总体健康状况较差；也就是说，许多人原本就患有心血管疾病、吸烟量大，或有该人群中常见的其他合并症。

大剂量雌激素使其他原因死亡率升至52.8%，比安慰剂组增加了7.1个百分点的组分C，并使相对风险升至3.9（52.8%/13.5%）。这种治疗所致的超额死亡主要由心血管原因造成：雌激素组中有40.0%死于心血管原因，而安慰剂组为28.3%（差异：11.7个百分点）。大剂量雌激素的这种心血管毒性早已为人所

docu历史记载，并导致雌激素治疗被弃用于前列腺癌 r.

组分A（前列腺癌死亡）占安慰剂组的29.1%，占雌激素组的21.6%。然而，正如我们下文将讨论的，雌激素治疗下前列腺癌死亡率表面上降低了7.5个百分点，这并不代表治疗获益。

表4：中位随访时间时死亡率的分解：大剂量雌激素与安慰剂

| 组别 | N | 前列腺癌死亡（A） | 其他原因死亡 | % 其他原因原因 | 相对 Risk |
|---|---|---|---|---|---|
| 一般人群（D） | — | — | — | 13.5% | 1.0 |
| 安慰剂（D +B） | 127 | 37（29.1%） | 58 | 45.7% | 3.4 |
| 大剂量雌激素（D + B + C） | 125 | 27（21.6%） | 66 | 52.8% | 3.9 |
| *Component B (baseline differences): 32.1 percentage points* | | | | | |
| *Component C (treatment-induced): 7.1 percentage points* | | | | | |

这个示例展示了几个关键要点。首先，两组的相对风险都远高于 1.0（安慰剂组为 3.4，雌激素组为 3.9），这证实在这种情境下，将净生存率按标准方式解释为癌症特异性生存率并不成立。即使在随机试验中，组成部分 B 仍然很大，代表基线时存在 32.1 个百分点的超额其他原因死亡率。其次，组成部分 C 确实存在，且具有重要的临床意义：即使这些治疗导致的死亡被编码为其他原因，因果上仍可归因于癌症；它们也构成癌症死亡负担的一部分，无法通过更精细的生命表分层消除。第三，将 Pohar Perme估计量应用于这项试验，得到的将是组分A + B + C，而非其预期的目标估计量（仅组分A）。B 和 C 合计为 39.2 个百分点，因此估计量实际给出的结果与其设计目标之间差距很大，远超任何可能的治疗效应幅度。

如果不进行这种分解就考察净生存效应，会产生误导。大剂量雌激素似乎使前列腺癌死亡（组分A）从 29.1% 降至 21.6%，风险差约为−7.5 个百分点，提示治疗有益。然而，这种表面上的保护作用是因为患者在死于前列腺癌之前，已因雌激素而死于心血管疾病；这是组成部分 C 通过竞争风险掩盖治疗危害的典型例子。[8] 雌激素对前列腺癌进展的直接效应（仅组分A）可能为零，甚至有害，但组成部分 C 导致的大量死亡掩盖了这一点。[8]

**通过四种目标估计量解读试验**

前列腺癌试验具体说明了目标估计量的选择如何改变对治疗获益的表面结论（表2）。

**疾病特异性生存率（仅组分A）**将其他原因死亡视为非信息性删失。在这一目标估计量下，雌激素似乎有益：雌激素组中有21.6%的人死于前列腺癌，而安慰剂组为

29.1%，相差7.5个百分点，提示治疗疗效具有实际意义。报告这一目标估计量的临床试验研究者会得出结论：雌激素可减少癌症进展。

**疾病归因生存率（组分A ＋ C）**将治疗所致的其他原因死亡重新计入癌症负担。在雌激素组中，C组分所致的治疗所致心血管死亡率为7.1个百分点，使疾病归因死亡率总计约为28.7%（21.6%+ 7.1%）。与安慰剂组的29.1%相比，表面上的治疗获益几乎完全消失。评估癌症治疗全部负担的卫生经济学家会正确地得出结论：雌激素并未带来有意义的死亡率优势。

**死因特异性生存率（组分A ＋ 部分C）**取决于医生如何编码归因于雌激素治疗的心血管死亡。如果将心血管死亡编码为与癌症无关（这是当时的标准做法），那么这一目标估计量就近似于仅组分A，再次误导性地显示治疗获益。如果部分治疗所致死亡被编码为与癌症相关，情况会有所改善，但结果并不一致。

**净生存率（组分A ＋ B ＋ C）**仅剔除一般人群风险（D组分），同时在两组中保留基线健康差异（组分B）。由于组分B在随机化后于两组间保持平衡，且两组共有，因此在这一试验内比较中，净生存率会呈现与疾病特异性生存率相似的模式（再次看似有利于雌激素），但较大的组分B会使两组的绝对生存率均降低32.1个百分点。因此，净生存率会低估两组的生存率，同时仍无法揭示表面上的治疗获益完全是由组分C死亡取代组分A死亡所致。

这说明了一个更广泛的问题：净生存率的预期适用场景是死因数据不可靠时进行国际比较。在随访完整且治疗分配已知的随机试验中，使用净生存率而非疾病归因生存率没有方法学依据；这样做还会掩盖试验旨在检测的治疗危害。

这个例子说明，选择目标估计量并非只是技术偏好，而是会决定实质性结论：只有疾病归因生存率才能准确反映雌激素治疗是以癌症死亡换取心血管死亡，因此并未给患者带来净生存获益。

## 3.3 分析结果：净生存率与预期估计目标的偏离

为系统量化B和C组分如何影响净生存率与其预期估计目标之间的偏差，我们根据风险分解以解析方式推导出生存概率。该方法可得出精确结果，因为净生存率与癌症特异性生存率之间的偏差是分解框架在数学上的结果，而非需要通过模拟

个体。

我们采用了以下参数设定：

- 癌症特异性风险率：威布尔分布，形状参数为 1.5，尺度参数为 5.3，对应 5 年癌症特异性生存率为 40%

- 一般人群其他死因风险率：每年恒定为 $h_D = 0.025$ ，近似于英国生命表中 70 岁男性的死亡率（即代表前列腺癌、结直肠癌和头颈癌人群的年龄分布）

- 相对风险情景：1.0、1.5、2.0、3.0、4.0

- 组分分配：三种情景：（1）所有超额部分均分配给组成部分 B（纯基线差异）；（2）所有超额部分均分配给组分C（纯治疗效应）；（3）B 和 C 平均分配（混合）

对于每种情景和相对风险值，我们推导出三个量。疾病特异性生存率：

$$S_A(t) = \exp\left[-(t/5.3)^{1.5}\right] \tag{7}$$

疾病归因生存率：

$$S_{A+C}(t) = \exp\left[-(t/5.3)^{1.5} - h_C \cdot t\right] \tag{8}$$

净生存率：

$$S_{\text{net}}(t) = \exp\left[-(t/5.3)^{1.5} - (\text{RR}-1) \cdot h_D \cdot t\right] \tag{9}$$

其中， $h_C = \text{frac}_C \times (\text{相对风险} - 1) \times h_D$ and $\text{frac}_C$ 表示分配给组分C的超额风险比例。所有计算均使用 R 4.3.0 版本完成。

图 3 显示了三种组分分配方案下不同相对风险情景中的 5 年生存率估计值。当相对风险 $= 1.0$ 时，三个估计目标完全一致，这证实了在标准解释所假设的理想条件下，净生存率等于癌症特异性生存率。随着相对风险增加，净生存率与两个目标均出现偏离：当相对风险 $= 2.0$ 时，净生存率在 5 年时低估癌症特异性生存率 4.7 个百分点；当相对风险 $= 4.0$ 时，这一差距达到 12.5 个百分点。

**情景1（纯基线差异）：**当所有超额其他原因死亡率都代表基线健康差异（组成部分 B）时，净生存率会低估疾病特异性生存率和疾病归因生存率。在相对风险为 $= 2.0$ 时，净生存率与癌症特异性生存率之间的差距在 5 年时为 4.7 个百分点；在相对风险为 $= 4.0$ 时，这一差距达到 12.5 个百分点。通过更好的分层可以消除这种偏倚。如果能够在生命表中为癌症患者找到完全匹配的适当比较组，组成部分 B 就会消失，净生存率也将等于癌症特异性生存率。

**情景2（纯处理效应）：**当所有超额死亡都代表治疗引起的死亡（组分C）时，净生存率会低估疾病特异性生存率，但能够准确估计疾病归因生存率。这是预期之中的：组分C代表通过治疗路径因果归因于癌症的死亡，而疾病归因生存率能够正确反映这些死亡。对于给定的相对风险，净生存率与癌症特异性生存率之间的差距与情景1完全相同，因为超额风险率总量相同。关键的是，这一差距无法消除。任何分层都无法消除这一差距，因为这些死亡是由癌症治疗所致。

**情景3（混合）：**在现实的混合情景中，对于给定的相对风险，净生存率低估癌症特异性生存率的幅度与情景1和情景2相同（因为总超额风险率保持不变），但此时还会低估疾病归因生存率，低估幅度为组成部分 B 所占的部分。因此，净生存率无法准确估计其中任何一个目标。

图 4 展示了各分配情景下三个固定相对风险值（1.0、2.0、4.0）对应的生存曲线随时间的变化。这一视角揭示，净生存率与癌症特异性生存率之间的差距并非在整个随访期间保持不变。从诊断时起，差距逐渐扩大，在约4–5年时达到最大值（此时已有40%的癌症特异性死亡发生，超额其他原因死亡率的累积影响也达到最大），随后到10年时逐渐缩小，因为所有组的生存概率都趋近于零。在相对风险为 $= 4.0$ 时，最大差距超过12个百分点，恰好集中在临床和政策层面最为相关的中期时间窗内。相对风险为 $= 1.0$ 的面板证实，当癌症患者没有超额其他原因死亡率时，三种估计目标完全一致。

## 3.4 分析结果的含义

从这些结果可以得出三点认识。第一，即使相对风险升高幅度不大，也会产生具有临床意义的差异：当相对风险为 $= 1.5$时，5年时的差距约为2.4个百分点，这与癌症监测项目通常作为进展报告的生存改善相当。第二，仅凭相对风险无法揭示超额其他原因死亡率是来自组成部分 B、组分C，还是两者兼有，因此无法仅凭净生存率判断该差距是可解决的还是不可消除的。第三，即使理想分层的生命表消除了全部组成部分 B，在存在组分C时也无法恢复疾病特异性生存率，因为治疗导致的死亡是癌症的真实因果后果，而非方法学伪象。

# 4 讨论

## 4.1 主要发现

本分析阐明了净生存率所估计的内容，以及它在何种条件下对应于癌症特异性生存率。我们采用将总死亡率分解为四个组成部分的理论框架，区分了超额其他原因死亡率的两个来源：基线健康差异（组成部分 B）和治疗导致的死亡（组分C）。只有前者可通过方法学改进加以解决；后者代表一条经由治疗效应从癌症通向死亡的不可消除的因果路径。分析结果表明，差异的幅度可能相当大。当相对风险为2.0（即对许多癌症类型而言较为保守的估计）时，净生存率在5年时将疾病特异性生存率低估4.7个百分点；当相对风险为 $= 4.0$时，差距达到12.5个百分点。这些并非微小的技术性差异，而是对解读生存趋势、国际比较和治疗进展具有实际意义。

前列腺癌试验对此作了具体说明：7.1个百分点的组分C足以完全掩盖治疗毫无获益这一事实，而只有通过疾病归因生存率，才能得出正确结论（即雌激素以癌症死亡换取心血管死亡）。因此，当前标准解释在一个明确且可以纠正的方面具有误导性。“假设癌症是唯一可能的死因时的生存率”这一说法准确指明了预期的目标估计量（即仅组分A），但只有在癌症患者的其他原因死亡率与一般人群相同这一严格条件下，Pohar Perme估计量才能实现这一目标。解决办法不是放弃净生存率，而是替换这一标签，因为它一直以来都夸大了估计量实际能做到的事情。

## 4.2 对癌症生存研究的启示

我们的发现对癌症生存统计的报告和解释方式有多方面的启示。

在措辞和解释方面，我们应停止将净生存率描述为“假设癌症是唯一死因时的生存率”。这种表述准确指明了预期的目标估计量，但当相对风险超过1.0时，Pohar Perme估计量并不能实现这一目标。我们建议改为：“净生存率表示剔除了一般人群其他原因死亡率后的生存率。”这一表述准确描述了统计操作，同时避免了缺乏依据的因果主张。

就目标估计量的选择而言，净生存率在某些用途上仍然很有价值，尤其适用于国际比较，因为死因数据可能不可靠或编码不一致。当死因数据可靠时，死因特异性生存率可能更易于解读，因为它承认治疗导致的其他原因死亡在因果上可归因于癌症。

关于报告　　实践中，在可行的情况下，研究人员应报告多项指标：净生存率（用于

与现有文献和国际比较保持连续性）、死因特异性生存率（便于解释），以及其他原因的标准化死亡比（以评估超额幅度）。这一组合将使读者既能了解癌症死亡负担，也能了解治疗所致其他原因死亡负担。

解释问题的程度因癌症类型而异。[3] 对于结直肠癌（相对风险 $\approx 1.0$），净生存率与癌症特异性生存率大致相当。对于头颈癌（相对风险最高可达4.0– 4.5），两者可能存在显著差异；根据我们的参数假设，5年时差异可能超过12个百分点。研究人员在解读特定癌症类型或治疗时期的净生存率时，应考虑B和C组分可能产生的影响程度。

CONCORD和EUROCARE等项目报告了不同国家和时期的净生存率，因此应在报告中阐明其解读方式。[6] 为帮助读者判断相对风险何时显著偏离1.0，这些项目可以考虑在报告净生存率的同时，报告其他原因导致的标准化死亡率比。

## 4.3 与现有文献的关系

以往研究已指出净生存率可能存在解读问题。Belot等人的教程提到，净生存率所依据的风险框架对独立性作出了假设，但没有强调相对风险超过1.0所带来的影响。[9] Dickman及其同事讨论过净生存率与死因特异性生存率之间的区别，但没有采用正式的风险率分解框架。[10]

Botta等人提供了关于超额其他原因死亡率的重要实证证据，并提出了四组分解框架。[3] 我们的研究在此基础上，明确将分解与净生存率的解释联系起来，区分可解决与不可消除的偏倚来源，并通过分析量化二者的差异程度。

我们强调，我们批评的并非净生存率这一指标本身，而是对它的解释。净生存率具有重要用途，今后也将继续发挥价值。我们的目标是澄清它所估计的内容，以便对其作出恰当解释并加以运用。

## 4.4 局限性

我们的分析结果仅用于说明问题，并非面面俱到。我们采用了简化的威布尔分布来描述癌症特异性风险，并以常数近似一般人群其他原因死亡率，而没有重现真实癌症登记数据的全部复杂性。我们考察的情景（纯B、纯C、混合）是经过简化的；实际上，B和C组分之间的分配可能因癌症类型、治疗时期和患者特征而异。

我们依赖已发表的　而是采用 Botta等人的实证证据，而非自行分析

登记数据。[3] 前列腺癌试验的分解确实构成了原创性实证分析，但该分析依据的是一项历史试验，其患者群体和治疗时期均具有特定性，因此限制了研究结果的普遍适用性。虽然登记证据清楚地表明，相对风险会因癌症类型而异，且往往大于 1.0，但仍有必要开展更深入的实证研究，以了解 B和C组分随时间变化的特性，以及它们在不同患者亚组中的差异。

我们的框架假设存在竞争风险结构，因此可能无法充分涵盖半竞争风险情形，也无法充分捕捉癌症与其他原因死亡率之间的时变关系。我们未处理时期与队列分析、年龄-时期-队列模型或死因编码随时间变化等相关方法学问题。为便于说明，我们的分析结果假设其他原因风险率恒定；在时变风险率下，净生存率与其组成部分目标估计量之间的概念关系仍普遍成立，因为这一关系可直接由风险率分解框架推导得出。

# 5 结论

净生存率是比较癌症生存的一项有价值的工具，但其标准解释需要修正。传统说法认为，净生存率代表“假如癌症是唯一可能的死因时的生存率”；这一说法正确指出了预期的目标估计量，但只有当癌症患者的其他原因死亡率与一般人群相同时，Pohar Perme估计量才能实现这一目标，而实证证据表明，这一条件往往并不成立。若该条件不成立，估计量给出的是 $A + B + C$，而非单独的组分A，因此这一名称并未准确反映实际估计的内容。

当癌症患者因基线健康差异或治疗所致效应而面临较高的其他原因死亡率时，净生存率会低估癌症特异性生存概率，换言之，会高估癌症死亡负担。治疗诱发组成部分代表一条不可消除的因果路径，无法通过更好的分层加以排除；即使医学编码将这些死亡归为其他原因，也可以认为它们由癌症所致。

我们呼吁在癌症流行病学中使用更精准的表述。净生存率应描述为“剔除了与一般人群其他原因死亡率相当的死亡后的生存率”，而不应将其描述为一种因果反事实，即除目标疾病外消除所有其他死因。研究人员在解读特定癌症类型和治疗时期的净生存率时，应考虑超额其他原因死亡率可能达到的程度，并在可行时报告补充指标。

正确解释净生存率，将有助于加强癌症生存研究，使我们的表述与实际估计的内容相一致，承认治疗引起的死亡所带来的真实负担，并帮助读者理解生存统计数据能够和不能告诉我们哪些有关癌症影响的信息。

## 利益冲突

作者声明不存在利益冲突。

## 参考文献


[1] Cronin KA、Lake AJ、Scott S等。关于癌症状况的全国年度报告，第一部分：全国癌症统计。癌症。2018;124(13):2785–2800。

[2] Perme MP、Stare J、Est`eve J。关于相对生存的估计。*Biometrics*。2012;68(1):113–120。

[3] Botta L、Capocaccia R、Bernasconi A等。癌症患者中非癌症原因所致的超额死亡：一项使用混合治愈模型的基于人群的研究。*American Journal of Epidemiology*。2025;194(11):3316– 3324。

[4] Rubio FJ、Rachet B、Giorgi R、Maringe C、Belot A。关于在生命表分层不足情况下估计超额死亡风险的模型。*Biostatistics*。2021;22(1):51–67。

[5] Lambert PC, Dickman PW, Nelson CP, Royston P. 使用相对生存模型估计癌症及其他原因导致的粗死亡概率。*Statistics in Medicine*. 2015;34(13):2092–2108.

[6]Allemani C, Matsuda T, Di Carlo V, et al. 全球监测癌症生存趋势（2000–14 年，CONCORD-3）：分析来自 71 个国家 322 个基于人群的登记处的个体记录，涵盖 37,513,025 名被诊断患有 18 种癌症之一的患者。*Lancet*. 2018;391(10125):1023–1075.

[7] 退伍军人管理局泌尿外科合作研究组。前列腺癌：治疗比较。*Journal of Urology*. 1967;98(5):516–522.

[8] Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hern´an MA. 针对存在竞争事件的失败时间情境中经典统计估计目标的因果框架。*Statistics in Medicine*. 2020;39(8):1199–1236.

[9] Belot A, Ndirangu M, Luque-Fernandez MA 等. 根据受众总结和传达生存数据：以基于人群的癌症登记数据为例，介绍不同的衡量指标。*Clinical Epidemiology*. 2019;11:53–65.

[10] Dickman PW, Sloggett A, Hills M, Hakulinen T. 相对生存回归模型。*Statistics inMedicine*. 2004;23(1):51–64.

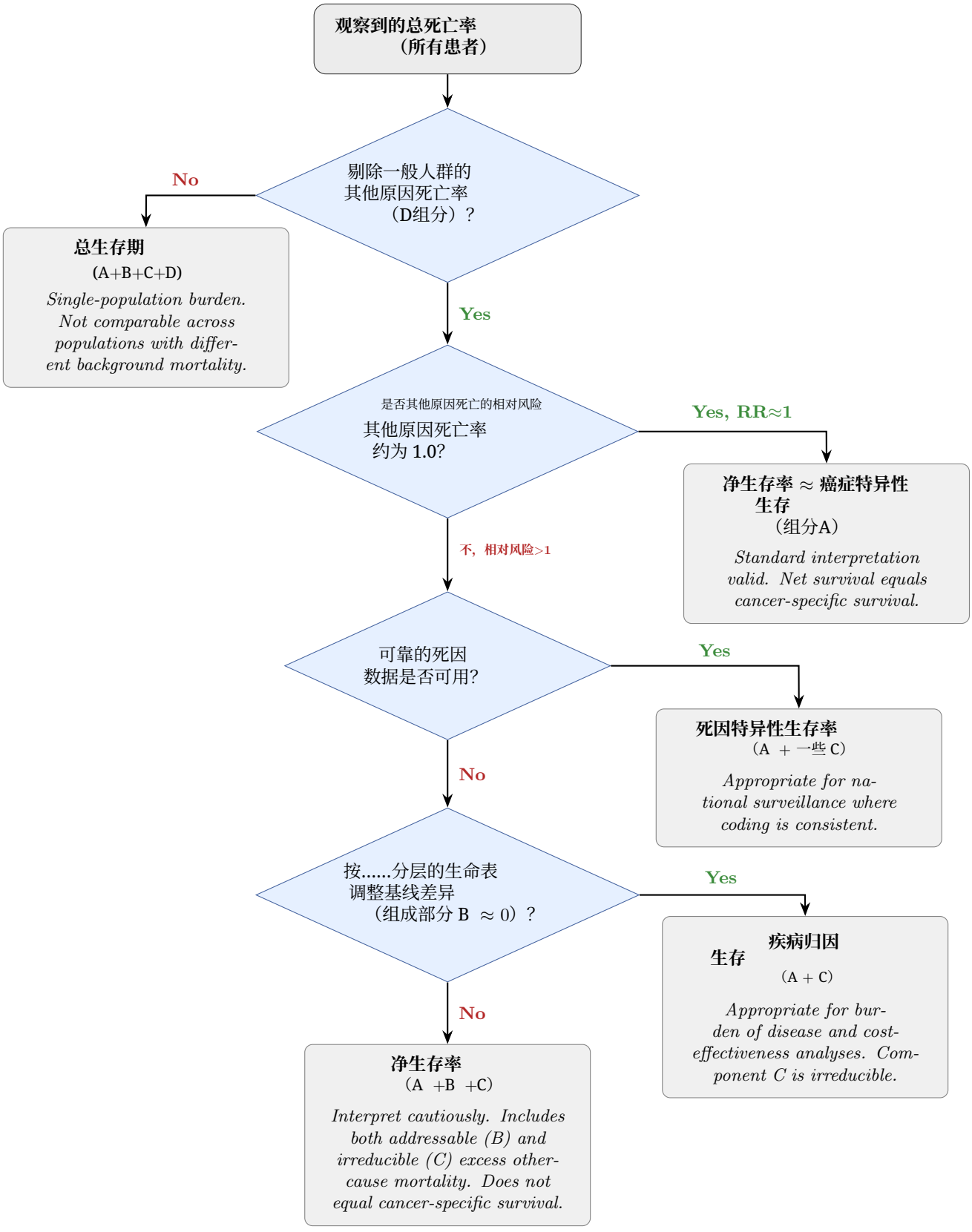


图2：用于选择适当生存估计目标的决策树。从观察到的总死亡率出发，研究者依次回答四个问题，以确定其分析所产生的目标估计量。各组成部分的字母对应方程2中的风险分解。

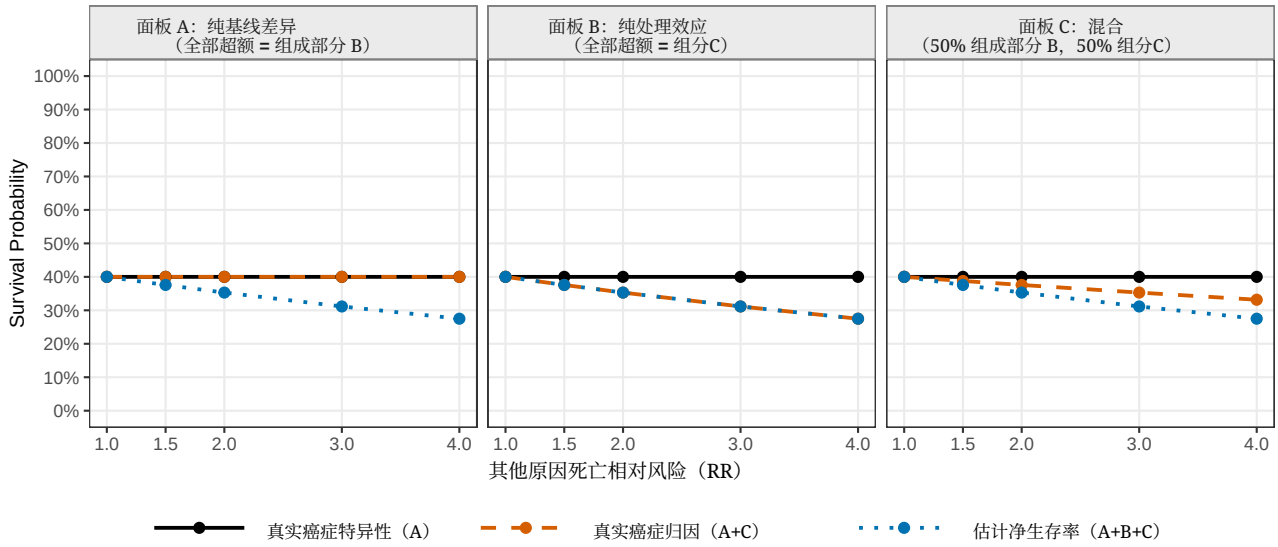


图3：在三种组分分配情景下，不同其他原因死亡相对风险对应的五年生存概率。每个面板展示五年随访时，在相对风险值为 1.0、1.5、2.0、3.0 和 4.0 时的疾病特异性生存率（组分A；黑色实线）、疾病归因生存率（组分A ＋C；橙色虚线）和估计净生存率（组分A ＋B ＋C；蓝色点线）。面板 A：所有超额其他原因死亡率均归入组分B（纯基线差异）。面板 B：所有超额死亡均归入组分C（纯治疗效应）。面板 C：在B和C之间平均分配（混合）。癌症特异性风险：威布尔分布（形状参数 1.5，尺度参数 5.3），对应五年癌症特异性生存率为 40%。一般人群其他原因风险：$h_D = 0.025$ 每年（英国生命表，年龄 70 岁）。当 RR 为 ＝ 1.0 时，三个估计目标相同。

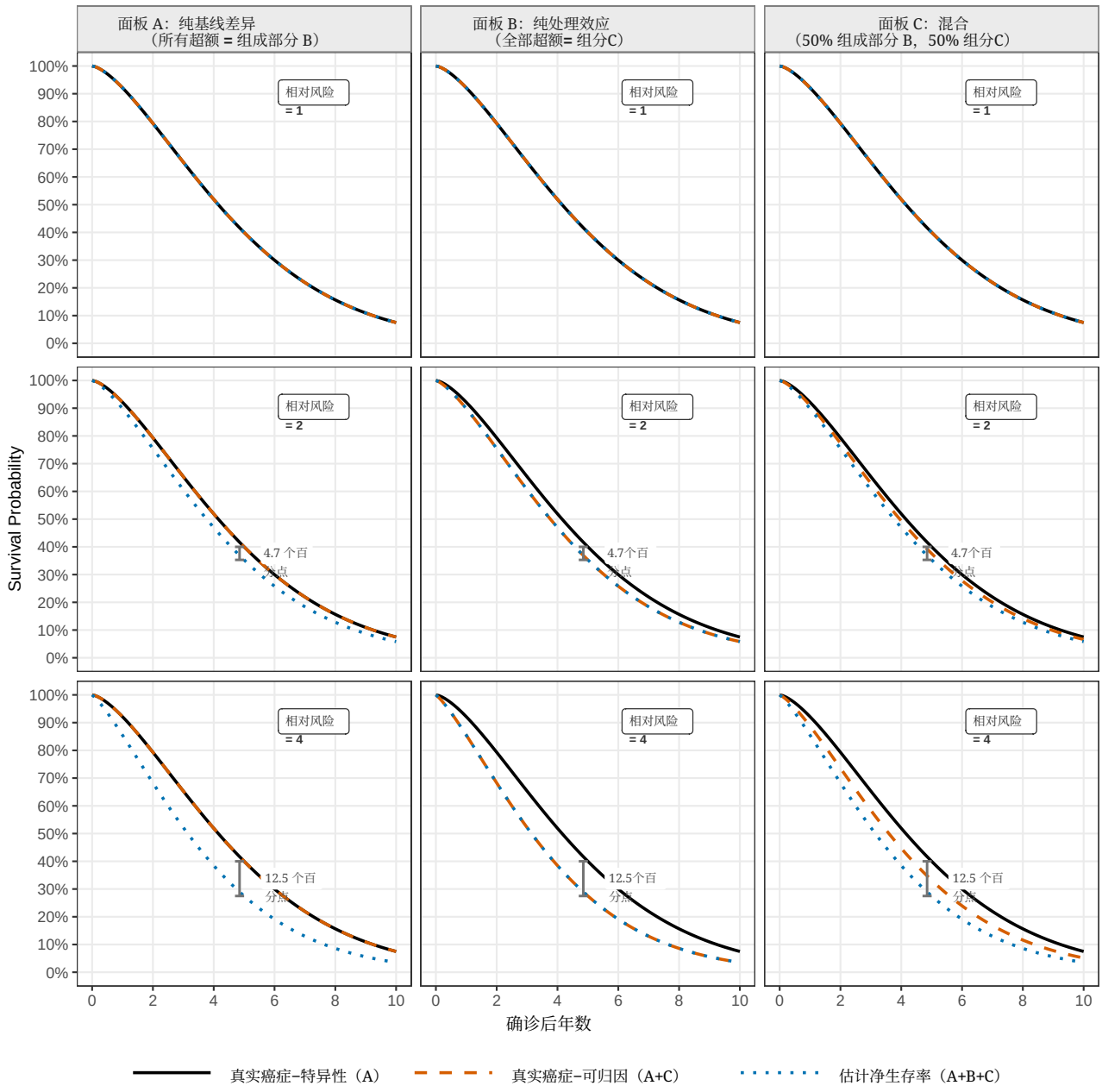


图4：按相对风险值和组分分配情景绘制的10年生存曲线。各面板按 $3 \times 3$ 网格排列：行对应相对风险值（相对风险 $= 1.0$、2.0、4.0）；列对应组分分配情景（纯基线差异、纯处理效应、混合）。每个面板内，三条线分别显示疾病特异性生存率（组分A；黑色实线）、疾病归因生存率（组成部分A + C；橙色虚线）和估计净生存率（组成部分A + B + C；蓝色点线）。五年处的括号标示净生存率与疾病特异性生存率之间的绝对差距（百分点）。癌症特异性风险：威布尔（形状参数1.5，尺度参数5.3）。一般人群其他死因风险：每年 $h_D = 0.025$ 。在相对风险 =1.0面板中，三条线重合。净生存率与癌症特异性生存率之间的差距自诊断起扩大，约在四至五年达到最大，随后随生存概率趋近于零而收窄。