arXiv:2611.32075v2 [stat.CO] 25 Nov 2026

coherent4cast：R语言中依赖历史的复发事件与终末事件的相干生存建模与预测

SundayA. Adetunji俄勒冈州立大学

Oluwatimileyin A. AdetunjiBowen大学

摘要

研究复发事件和终末结局的研究人员需要能够将不完整的事件历史转化为可解释的预测和不确定性估计的软件。我们介绍 coherent4cast，这是一个 R 软件包，可用于拟合依赖历史的复发事件与终末事件模型、预测未来事件负担和终末事件风险，并评估观测不完整带来的敏感度。用户提供事件历史、协变量和观测信息；该软件包在明确说明的假设下，返回拟合模型、动态预测、不确定性汇总和诊断指标。

该软件包可对与并列日期或观测窗口相容的事件历史进行积分，并支持多种用于处理受试者间异质性的脆弱性分布。共享概率模型可生成复发事件计数、终末事件概率和限制平均存活时间。在受支持的随机缺失和右删失情形下，交叉拟合靶向估计可提供影响函数诊断和渐近推断。单独的敏感性分析则可根据指定的记录限制给出界。基线二元治疗接口可在一致性、条件可交换性、正性和独立删失的假设下估计标准化治疗对比。

C++17 计算核心支持 R 接口。独立的数值检验、计算基准测试、模拟实验和可复现实例记录了准确性、性能和推断局限性。输出保留观测假设、收敛信息和数值诊断。该软件包、教程和复现材料均已公开，可支持临床及其他事件历史应用中的透明分析。

关键词：复发事件、终末事件、脆弱性、动态预测、粗化数据、部分识别 - 别识 别、R、C++。

1. 引言

复发事件预测可以表示预期事件计数、事件发生概率或存活概率。这些量彼此相关，但其单位和纳入资格规则有所不同。患者可能在死亡前反复再入院；设备单元可能

在发生不可修复故障之前经历多次维修。一个实用的软件实现必须从数据准备、估计，到预测和验证，始终保留这些区别。同时，还必须区分未发生的事件与未记录事件的区间。

Adetunji 和 AdetunjiAdetunji 和 Adetunji（2026年）的方法学框架基于共同的完整历史规律构建预后目标，并区分似然推断、缺失历史敏感性分析和预测集校准所需的假设。本文介绍其软件实现。本文的主要贡献是一条可审计的计算路径：声明的事件纳入资格和观测限制被纳入共享似然；一个保留的规律生成多个预测目标；独立的数值例程分别报告各自的假设和诊断。本文作为软件论文，与方法学稿件相辅相成，而并非声称其所有理论扩展都已作为通用例程提供。

该软件包针对三种常见的实现错误来源提供了解决方案。第一，观察到终末事件后，所有后续风险集都会发生变化。第二，粗化历史是由一组相容路径构成的；选择一条方便的代表性路径可能改变似然和预测。第三，有限样本预测覆盖率、参数置信覆盖率和数值精度是不同的性质。函数名称相同或优化器成功运行，并不意味着它们的保证可以互换。各接口明确体现了这些区别，分析人员可以据此判断哪种计算能够回答自己的问题。

版本 0.10.0 引入的核心架构集成了五项扩展：自动观测纤维、稀疏传播和大支持集界、通过主要建模API进行靶向、随时间演变的预测协变量，以及带有诊断功能的替代脆弱性分布。所有数据处理和高级方法均可在 R 中使用；似然、粗化混合、传播和核心有限支持集运算则通过独立的 C++17 库实现。我们报告了确定性数值一致性、实测计算性能以及五项案例研究。版本0.10.1 增加了显式的数据体制标签、survival 和竞争风险封装函数，以及审计导出。其因果修订增加了一个独立的治疗加权靶向引擎，用于实现方法学稿件中补充材料 S12 的内容；现有预后数值内核予以保留。

2. 相关软件

一些成熟的软件包提供了重要的复发事件与多状态工具。 **frailty-pack** 可拟合共享、联合和嵌套脆弱性模型，并包含针对纵向、复发和终末结局的动态预测（Rondeau、Mazroui 和 Gonzalez 2012；Król、Mauguen、Mazroui、Laurent、Michiels 和 Rondeau 2017）。 **mstate** 可组织竞争风险与多状态数据，并支持转移特异性估计和预测（ de Wreede 、Fiocco 和 Putter 2011）。 **reReg** 实现了用于复发事件的一般尺度变化回归框架，其中包括信息性终末事件，并且无需指定参数化脆弱性分布即可处理脆弱性（Chiou、Xu、Yan 和 Huang 2023）。 **JMbayes2**可拟合贝叶斯联合模型，将多个纵向结局与事件过程联系起来，并提供预测和验证工具（Rizopoulos、Miranda-Afonso 和 Papageorgiou 2026年）。这些软件包的科学应用范围各不相同；单一的排序会掩盖这些差异。

表1 记录了相关的已文档化功能以及**coher-ent4cast**所发挥的相应作用。此比较并未声称其他软件包必然会错误处理缺失数据或产生不一致预测。特别是，设定正确的联合模型本身已提供统一的概率律。此处的额外接口在各自声明的观测模型和状态模型范围内，将合法路径积分、有限路径靶向、记录敏感性优化以及明确的数值

软件	文档重点	与本实现的关系
frailtypack	共享、联合和嵌套脆弱性估计；动态预测（Rondeau等2012；Król 等2017）。	既有的脆弱性工作流程；当前接口还额外公开了合法路径积分和有限支持敏感度约束。
mstate	竞争风险与多状态数据准备、估计与预测（deWreede 等人2011）。	面向转移的工作流程；此处保留的计数状态编码选定的历史依赖以脆弱性为条件。
reReg	复发事件尺度变化回归以及可能具有信息性的终末事件事件（Chiou等 2023）。	不同的回归和脆弱性假设；不意味着准确度普遍适用或运行时间排序。
JMbayes2	贝叶斯纵向与事件时间联合模型、预测和验证（Rizopoulos等2026年）。	更丰富的纵向子模型；当前有限路径和敏感度计算具有不同的输入契约。
coherent4cast	自动纤维、替代脆弱性，有限路径和连续靶向，演变的协变量和稀疏预测/界。	明确区分工作模型预测、置信域和预测集。

表1：互补的软件范围。文中陈述描述的是有文档记载的功能；此处仅执行了文中指出的配套示例。未列入本表并不表示其他软件包无法实现相应扩展。

检查整合在同一个软件包中。0.10.0 版还针对声明的有限状态右删失实验提供连续时间靶向，并自动构建观测事件时间纤维。

为进行基本的执行比较，复现脚本还使用 **frailtypack** 对公开再入院数据拟合联合威布尔伽马模型，并使用**survival**(Therneau 2026) 对结肠试验拟合终末事件 Cox 模型。命令和已安装的版本均与输出一并记录。它们的似然、基线设定和目标量均不同于 **coherent4cast** 示例；因此，我们不将它们的似然值作为模型评分进行比较，也不报告无法匹配的速度比。

3. 模型和目标量

令 $i = 1, \ldots\ldots, n$ 表示独立个体，$k = 1, \ldots\ldots, K$ 表示复发事件类型。记$N_{ik}(t)$为计数过程，T_i^D为终末时间，$X_i(t)$为可预测的测量协变量，$h_i(t-)$为保留的历史摘要。给定均值为1的脆弱性变量时，工作强度为

$$\lambda_{ik}(t \mid Z_i, \mathcal{F}_{i,t-}) = Y_{ik}(t) Z_i \lambda_{0k}(t) \exp\{\beta_k^\top X_i(t) + \gamma_k^\top h_i(t-)\}, \qquad Z_i \sim F_\eta, \qquad \mathbb{E}_{F_\eta}(Z_i) = 1. \tag{1}$$

此处 F_η 是参数为 η的脆弱性分布，$Y_{ik}(t)$表示紧接t之前的纳入资格，Z_i 是方差为 ϕ的均值为1的脆弱性变量，λ_{0k} 是分段常数基线，$\mathcal{F}_{i,\,t-}$ 是完整历史。终末强度有其自身的系数。终末事件发生后，纳入资格为零；

不可重复标记在首次发生后也不再具有纳入资格。该软件包的标准历史状态由封顶的既往计数组成。封顶会改变强度特征，但不会改变负担奖励中所包含的未来事件数量。

在存活标志点 s 和时间范围 $\Delta > 0$处，目标类别为

$$\begin{aligned} \pi_D(s,\Delta) &= \mathbb{P}(T^D \le s+\Delta \mid T^D > s, \mathcal{H}_s), \\ \mu_k(s,\Delta) &= \mathbb{E}\{N_k(s+\Delta) - N_k(s) \mid T^D > s, \mathcal{H}_s\}, \\ r(s,\Delta) &= \mathbb{E}\{\min(T^D - s, \Delta) \mid T^D > s, \mathcal{H}_s\}. \end{aligned} \tag{2}$$

此处 $\mathcal{H}_s$ 表示可用记录， π_D 表示终末风险， μ_k 表示预期额外的 k 型负担， r 表示限制平均存活时间。计数在终末事件发生时停止。这些是预后期望。将某个系数解释为预防复发的效果，需要额外的因果假设和不同的估计目标。

对于一条精确路径，令 d统计所有符合条件的跳变， B 表示在这些跳变处排除脆弱性后的对数强度之和， H 表示排除脆弱性后的积分总强度。对于 F_η =伽马分布(a， a)，当 $a = 1/\phi$时，解析积分得到

$$\ell = B + \log\Gamma(a+d) - \log\Gamma(a) + a\log a - (a+d)\log(a+H). \tag{3}$$

此处 ℓ 表示该路径的边际对数似然， Γ 表示欧拉伽马函数， $a = 1/\phi$表示。条件脆弱性分布为伽马分布($a+d$， $a+H$)。该内核采用整数计数复发关系和稳定的对数计算，以避免在 $\phi = 0$附近发生消去误差；零方差模型也可显式使用。完整推导和识别条件见方法学稿件。其中包括独立个体、可预测协变量、非爆炸性、有效的观测模型以及充分的保留状态。

4. 软件架构与数值算法

图 1 展示了数据契约、数值内核和推断输出之间的分离。C++ 头文件不包含 R头文件。精简的 **Rcpp** 桥接层负责转换对象和异常（Eddelbuettel 和 François， 2011）；R 层实现公式处理、S3类、案例适配器、诊断和报告功能。这种设计支持独立的内核测试，并避免在绘图或数据导入代码中重复数值方法。

4.1. 数据验证与可预测历史

`as_trajectory()` 需要个体标识符、区间起点和终点、复发标记、终末事件指示变量以及声明的时间单位。可选的监测指示变量用于标识未观测区间。系统会检查有限时间值、按时间分组、重叠、事件编码和吸收性。复发与死亡同时发生的零长度区间不会被静默地加入微小扰动。在声明舍入精度后，粗化引擎会将日期转换为事件时间窗口，并对其可接受历史进行积分。精确拟合会拒绝不存在观测模型的间隔。当前的脆弱性拟合接口拒绝延迟进入，因为否则将忽略脆弱性的先验选择。

`build_history()` 在区间起点计算计数和距最近事件时间，此时当前区间的事件尚未发生。距最近事件时间缺失表示未记录更早发生的事件；不会将其替换为零



图1：五项已实现的扩展及其计算依赖关系。箭头表示数据或计算的传递，而非因果效应。靶向与敏感性界保留各自独立的观测实验；每个输出均附带其假设和诊断。

等待时间。模型公式不能包含区间端点或其事件结局。受试者选择必须保留完整历史记录，且文档所述的缺失值选项会移除整个受影响的受试者，而不是仅移除个别行。这些检查可以避免时间泄漏和风险集遭到破坏，但无法验证输入代码在科学上的含义。

0.10.1 接口新增前导类 `rc_trajectory`，同时保留 `surv_trajectory` 和 `data.frame` 继承。可选的 `terminal_cause` 字段声明吸收原因；它与复发标记和删失相互独立。附加的模式元数据概述已记录的内容。没有重复事件的数据集本身并不能确立单事件设计，多个事件标签也不能证明存在竞争风险结构。纳入资格、吸收和观察实验仍由分析人员声明。

可执行的 `replication/standards_mapping.R` 示例展示了选定的 CDISC ADaM 生存时间参数，以及明确定义的 OMOP 事件概念（临床数据交换标准协会 2016、2026年；观察性健康数据科学与信息学 2026年）。它们是源到模式转换，而不是经过认证的适配器。对于 ADTTE，分析人员选择分析参数，保留其持续时间和删失推导，并将 `CNSR == 0` 映射为事件指标。不得通过 `1 - CNSR`转换多个阳性删失代码。对于 OMOP，必须先明确概念集、事件段规则、人员关联、观察期和终末事件确定，记录才能成为事件历史。数据库观察并不能证明事件已被完整确定。

4.2. 自动观测纤维构造

借助 `observation = "coarsened"`，主拟合过程可以调用 `rc_coarsen_fibre()`，而无需手动展开路径。用户需提供事件标记、时间窗（L_j，U_j]、随访终点和时间单位。声明的舍入单位会将点日期转换为具有正宽度的时间窗；程序不会推断日期精度。精确的点事件保留计数测度，而区间时间则采用勒贝格积分测度。假定在本实验中，事件计数和事件标记均被完整记录。事件数量未知的未观测间隔属于另一种观测模型。

令$\mathcal{F}(o)$ 包含与观测记录 o 相容的按时间排序的标记历史。终末标记必须位于最后，且不可重复的标记须遵守其纳入资格规则。似然及其数值近似为

$$L(\theta; o) = \int_{\mathcal{F}(o)} L_c(\theta; h)\, d\mu(h), \qquad \widehat{L}(\theta; o) = \sum_{b=1}^{M} w_b L_c(\theta; h_b). \tag{4}$$

其中，θ 是参数向量，L_c 是完整历史似然，h 是合法历史，μ 是乘积积分测度，h_b 是生成历史，w_b 是其积分测度，M 是生成的求积路径数或尝试的重重要性抽样次数。在抽样过程中，被拒绝的抽样贡献为零，且 $w_b = 1/\{Mq(h_b)\}$ 对应提议密度 q。计算似然前不会对积分权重进行归一化。当提议分布的支持集覆盖观测纤维且被积函数可积时，重要性积分对似然的估计是无偏的；对数似然则不是无偏的。

同标记窗口需要特别谨慎处理。标记排列可能会重复计算同一个物理上的带标记历史。引擎会将每次后续出现分配给可用的同标记窗口中上界最早的窗口。交换两个可行分配不会改变其属于区间窗口的情况，并会产生这一规范顺序。因此，重叠的同标记窗口会作为一个并集进行积分。完成穷举后，会在窗口端点和基线不连续点处分割，并进行高斯求积。超过 `max_perms` 或`max_paths`时，将触发带种子的重要性抽样，同时保留尝试次数/拒绝次数、有效样本量和积分蒙特卡洛标准误。穷举可以消除排列误差，但不会使一般的连续时间积分在解析上精确。

4.3. 似然、得分与合法粗化

底层 `rc_loglik()`接收扩展后的标记暴露和设计矩阵，并返回受试者对数似然、解析得分和观测信息。`coherent4cast()` 构建该设计矩阵，在基线变化处拆分区间，并优化边际似然。

它会报告梯度幅值、信息特征值、优化器状态和参数边界标记。数值优化器的成功代码不足以判定拟合为常规拟合。

对于粗化观测，生成的或明确提供的合法完整路径及对数积分权重会进入 `rc_coarsen()`。给定分量对数似然 ℓ_j、得分 u_j 和观测信息 I_j，其归一化路径权重 w_j 为

$$\begin{aligned} \ell_O &= \log \sum_j \nu_j e^{\ell_j}, \qquad u_O = \sum_j w_j u_j, \qquad w_j = \frac{\nu_j e^{\ell_j}}{\sum_v \nu_v e^{\ell_v}}, \\ I_O &= \sum_j w_j I_j - \sum_j w_j (u_j - u_O)(u_j - u_O)^\top . \end{aligned} \tag{5}$$

此处 j 表示合法求积路径或重要性路径的索引，ν_j 是积分权重，O表示观测记录。对数和指数可稳定地计算混合似然。公式5 是缺失信息恒等式（Louis 1982）；仅对分量信息求平均会漏掉得分变异减项。计算似然前不会对积分权重进行归一化。对连续事件时间进行数值积分时，仍需细化观测纤维求积。

4.4. 有序传播与累积奖励

给定一个脆弱性值和一个充分有限状态后，生成器 Q 具有非负的非对角元素，且行和为零。如果 $\nu \geq \max_j(-Q_{jj})$，则均匀化给出

$$e^{Q\Delta} = e^{-\nu\Delta} \sum_{m=0}^{\infty} \frac{(\nu\Delta)^m}{m!} \, (I + Q/\nu)^m \, . \tag{6}$$

此处 Δ 是时长，m 是泊松级数索引，I 是单位矩阵。均匀化用泊松加权离散转移取代矩阵指数（Rao andTeh2013）。稠密的 `rc_propagate()` 仍作为独立的小状态参考。新的 `rc_generator_csr()` 以压缩稀疏行的形式存储非对角速率。概率行向量和积分奖励通过 `rc_propagate_sparse()` 更新，而无需构造 $S \times S$中间矩阵。对于 S 个状态、E 条存储的转移和 r 项奖励，每个保留的泊松项需要 $O(E + Sr)$ 的计算量，工作存储量则随边数和状态数线性增长。项的总数和时长分块取决于速率、时间范围和容差；这里不作无条件的线性时间声明。

该实现通过有序片段传播实数运算下的级数截断界。舍入误差会单独报告，作为诊断信息，而非区间算术认证。该实现采用时长分块和显式的几何尾部不等式；但并未声称实现了另一种 Fox–Glynn 算法（Foxand Glynn 1988）。项数和分块预算不足时，会返回结构化错误，而不会默默给出未经限定的近似结果。

生成器会按照时间顺序应用于各个基线变化。所有分段均保留同一个脆弱性分量；在每个分段独立对转移矩阵求平均会改变依赖结构。在每个分量完成有序传播后，再进行脆弱性积分。`rc_predict_paths()`还会针对明确提供的兼容地标前历史记录，额外执行一次外层混合。如果未来路径在更早的评估时刻

被拒绝，则不会仅仅因为后验混合更易于存储，就用单个脆弱性分量替代它。

4.5. 通过主要建模接口进行靶向

选择 `targeting = "finite_path"` in `coherent4cast()` 会公开由 `rc_target()`实现的全路径随机缺失（MAR）实验。令 R 表示有限路径 Y 已被记录，令 X 表示一个观测分层，令 $F(Y)$ 表示目标函数向量，$m(x) = \mathbb{E}\{F(Y) \mid X = x\}$ 以及 $\pi(x) = \mathbb{P}(R = 1 \mid X = x)$。影响函数为

$$D(O) = m(X) - \psi + \frac{R}{\pi(X)}\{F(Y) - m(X)\}, \qquad \psi = \mathbb{E}\{F(Y)\}. \tag{7}$$

其中， $O = (X, R, RY)$ 是观测单位， D 是一个向量。受试者定义折。训练折上的干扰参数估计共同输入一个指数倾斜；所得混合分布会生成每个目标。该函数返回个体EIF值、其均值和方差、记录概率估计、折分配以及得分收敛结果。双重稳健一致性和正则效率分别保留其干扰项条件与余项条件。全路径MAR并不涵盖任意的间歇性隐匿事件。

此选项 `targeting = "continuous"` 拟合一个具有显式右删失的有限状态计数过程实验。令 $V_j(t,\ h)$ 表示目标 j 在时间 t 和保留状态 h 下的剩余条件价值，并定义跳跃对比 $B_{jk}(t,\ h) = a_{jk}(t,\ h) + V_j(t,\ h+k) - V_j(t,\ h)$。其观测数据影响函数为

$$D_j(O_i) = V_j(0, X_i) - \psi_j + \sum_{k=1}^{K+1} \int_0^\tau \frac{R_i(t)}{G(t \mid X_i)} B_{jk}(t, H_i(t-))\{dN_{ik}(t) - \lambda_{ik}(t)\, dt\}. \tag{8}$$

此处， i 为独立个体的索引， j 为有限目标向量， k 为复发标记， $K + 1$为终末标记， X_i 为基线分层， $H_i(t-)$ 为充分的跳跃前状态， $R_i(t)$ 为观测风险集指示变量， G 为条件删失生存函数， N_{ik} 为计数过程， λ_{ik}为其全历史强度，a_{jk} 为该目标的即时跳跃奖励， τ 为评估时限， ψ_j 为总体目标。跳跃对比可由反向价值方程得出：一次跳跃会获得即时奖励，并改变剩余的期望价值。此实现要求历史记录从时间零起连续完整，在所声明的基线分层内满足条件独立右删失、 G为正，且保留状态充分。

受试者级折在训练受试者上估计事件和删失干扰项。在验证受试者上，共同扰动为

$$\lambda_k^\epsilon(t, h) = \widehat{\lambda}_k(t, h) \exp\left\{ \sum_j \epsilon_j \frac{\widehat{B}_{jk}(t, h)}{\widehat{G}(t \mid x)} \right\}. \tag{9}$$

$\widehat{\lambda}_k$ 和 $\widehat{G}$ 这里， λ_k 和 G 是通过折叠训练得到的干扰参数估计， x 是基线分层， B_{jk} 是当前值对比， ϵ_j 是定向系数。每次更新后，算法都会重新计算数值、对比和个体{EIF}。事件干扰参数可使用合并的工作强度或计数对数线性工作强度；在导出的实现中，删失风险率在声明的分层内为常数。时间网格细化用于检查数值插值误差。仅进行交叉拟合并不能满足正则渐近推断所需的干扰参数乘积速率、正则定向和随机等度连续性条件。

拟合对象保留一个折层混合法则。动态条件预测仅根据截至标志点的观测历史更新其折叠权重。总体目标区间采用中心化的个体层级{EIF}乘子和最大值-t 临界值；有界目标量使用logit变换，非负无界负担使用对数变换。未定义的变换结果或非正则结果会被标记。这些区间是渐近有限向量近似。总体{EIF}不会被重新用作个体历史的置信区间，仅靠变换也无法保证稀疏单元覆盖率。

当常规推断不可用时，观测奖励包络可作为另一种保守的备用方案。设 $a_j \le L_{ij} \le F_{ij} \le U_{ij} \le b_j$ 个独立同分布的单位为 $i = 1, \ldots, n$，且目标量有界： $j = 1, \ldots, m$。则

$$\left[\max\{a_j, \overline{L}_j - r_j\},\ \min\{b_j, \overline{U}_j + r_j\}\right], \qquad r_j = (b_j - a_j)\sqrt{\frac{\log(2m/\alpha)}{2n}} \tag{10}$$

以至少 $\mathbb{E}(F_{ij})$ 的概率共同覆盖 $1 - \alpha$。其中 F_{ij} 是完整目标奖励， L_{ij}， U_{ij} 是其可观测包络， a_j， b_j 是其已知支撑端点，横线表示样本均值， α 是族错误率。Hoeffding不等式（Hoeffding 1963）对每个 $2m$ 单侧偏差给出的界为 $\alpha/(2m)$；并集界与 $\mathbb{E}L_{ij} \le \mathbb{E}F_{ij} \le \mathbb{E}U_{ij}$ 共同推出该结论。终末风险以观测到的死亡作为下界包络，并将时间范围前的删失计入上界包络。限制生存以观测随访作为下界包络；若存在删失，则以时间范围作为上界包络。这是一个粗化外集，而不是精确的敏感性界，也不能替代有效影响函数置信区间。它允许信息性删失；未设上限的负担返回 $[0, \infty)$。

对于这一无条件有限样本陈述，整个预先指定的分析过程中都应使用 request `interval = "projection"`。在近似高斯区间与投影区间之间自动进行依赖数据的切换，既不能使合并后的程序继承无条件有限样本覆盖率，也不能使所选分支继承选择条件覆盖率。相同的限定也适用于自动零事件回退方案。

4.6. 基线治疗效应

`rc_causal()` 实现了 Adetunji 和 Adetunji在补充材料 S12 中提出的(2026年)静态基线二元治疗扩展。令 $A \in \{0,1\}$ 表示治疗（在该补充材料中记作 W），X 表示基线混杂因素， Y^a 表示分配为 a时的潜在完整历史，而 F_j 表示复发计数、终末事件风险或限制生存奖励。干预均值为

$$\psi_j(a) = \mathbb{E}\{F_j(Y^a)\} = \mathbb{E}\{V_j^a(0, X)\}, \qquad a \in \{0, 1\}. \tag{11}$$

此处， j 表示结局坐标， $V_j^a(0, X)$ 是其在基线时针对相应治疗组的反向价值。识别要求满足一致性和无干扰、给定 X 条件下的治疗可交换性、正治疗概率，以及条件独立的右删失和正删失生存概率。该软件包会检查输入支持集和估计重叠性，但无法验证可交换性。两个治疗组均基于相同的 X 经验分布进行标准化。

对于独立受试者 i，估计的治疗组梯度遵循 Rytgaard、Gerds 和 van derLaan（2022）提出的连续时间靶向原则：

$$D_j^a(O_i) = V_j^a(0, X_i) - \psi_j(a) + \frac{\mathbf{1}(A_i = a)}{e_a(X_i)} \sum_{k=1}^{K+1} \int_0^\tau \frac{R_i(t) B_{jk}^a(t, H_i(t-))}{G_a(t \mid X_i)} \{dN_{ik}(t) - \lambda_k^a(t, H_i(t-))\, dt\}. \tag{12}$$

这里 O_i 是观测历史，$e_a(X_i) = \mathbb{P}(A_i = a \mid X_i)$是治疗臂概率，$R_i(t)$是观测风险集指示变量，$G_a$ 是治疗臂特异的删失生存函数，k 是事件标记，N_{ik} 是其计数过程，$H_i(t-)$是保留状态，λ_k^a 是其可纳入强度，B_{jk}^a 是反向跳跃对比，τ 是固定时间范围。每位受试者在两个治疗臂中都会贡献基线值项；事件鞅项仅使用该受试者实际观测到的治疗臂。因此，在某个治疗臂中没有记录的受试者不会从该治疗臂的标准化总体中被剔除。

治疗与删失加权更新为

$$\lambda_k^{a,\epsilon}(t, h) = \widehat{\lambda}_k^a(t, h) \exp\left\{ \sum_j \epsilon_j \frac{\widehat{B}_{jk}^a(t, h)}{\widehat{e}_a(x) \widehat{G}_a(t \mid x)} \right\}. \tag{13}$$

这里 a 标识治疗臂，h 是基线部分为 x的保留历史，j 是目标坐标，k 是事件标记，ϵ_j 是其波动系数；上标符号表示折内训练的辅助量或当前的反向对比。正乘子可保留不可纳入强度中的零值以及终端吸收。事件、治疗和删失辅助量均在每位受试者所在的折之外进行训练。合并验证评分决定靶向更新；更新后的分布本身并不被描述为完全由训练折拟合得到。

基线治疗对比保留了治疗臂间协方差：

$$\widehat{\Delta}_j = \widehat{\psi}_j(1) - \widehat{\psi}_j(0), \qquad \widehat{D}_{\Delta,j,i} = \widehat{D}_j^1(O_i) - \widehat{D}_j^0(O_i), \qquad \widehat{\Sigma}_\Delta = \frac{\widehat{\mathrm{Cov}}(\widehat{D}_{\Delta,i})}{n}. \tag{14}$$

其中，n 为独立个体数，i 为受试者，j 为目标，Δ_j 为治疗臂 1 与治疗臂 ^0 之差的效应，Σ_Δ 为其估计量的协方差矩阵。`rc_causal_contrast()`还会返回比值和对数比值，使用 delta 方法梯度以及比值在对数尺度上的正值区间限。期望计数之比是负担比，而非风险比。对较小的联合协方差矩阵使用高斯乘子，可为规则坐标计算同时渐近区间，而无需分配一个“抽样次数 × 队列规模”的矩阵。

导出的实现最多使用 30 个预先指定的离散基线分层、6 个复发标记和 128 个存活状态。默认事件干扰参数采用合并计数，或对保留计数进行对数线性建模；删失风险在每个治疗臂和基线分层内保持恒定。样本外事件率数组可随声明的时间网格变化。治疗概率可以根据设计已知，也可以作为按折和分层划分的估计值提供。不使用自动倾向性截断。保留计数上限用于概括状态依赖性；它不会截断累积的复发奖励。

增广估计函数的一致性要求事件强度正确，或治疗概率与删失生存函数的乘积正确。仅倾向得分正确并不能弥补删失模型设定错误。效率和正态推断还要求满足补充材料 S12 中的乘积速率条件和正则性条件。零事件、未收敛或

零方差坐标会保留其估计值和诊断标记，但不会获得常规置信限。预后性记录敏感度区间和观测包络区间不能替代因果治疗界限。时变治疗、未测量治疗混杂敏感性以及粗化因果历史需要额外的估计量，不在此入口的范围内。

4.7. 大支持集敏感性优化

`rc_bounds()` 采用记录类别分布、已知记录比例 ρ、完整路径目标向量以及可选的粗化纤维。已记录与缺失质量满足带有敏感性参数 $\kappa \geq 1$的边际记录优势比约束。条件目标的形式为 $c^\top q / d^\top q$。Charnes–Cooper变换设定 $t = (d^\top q)^{-1}$ 和 $v = tq$，因此 $d^\top v = 1$，目标函数则变为 $c^\top v$ (Charnes和Cooper 1962)。求解器会检查原始可行性和对偶一致性，并返回端点取值分布。可选的已记录计数会产生同时二项置信约束和外置信投影。有效影响函数得分不能替代这些已记录分布约束。条件分母必须为正，声明的支持集必须有限，且记录比例必须已知。

对于大支持集，`rc_bounds_sparse()` 通过利用现有的记录优势比约束，避免构建稠密的Charnes–Cooper单纯形表。在记录纤维内，线性目标可以集中于其最大化原子。双资源排序分配可求解由此产生的已记录/未记录质量问题。参数化分式迭代（Dinkelbach1967）可提供可行端点分布和可重构的对偶证书。每次迭代的计算量为 $O(J + M \log M)$，内存用量为 $O(J + M)$，其中 J是完整原子数，M是记录类别数。这种精确的结构化约简支持任意划分纤维，以及已实现的已记录质量置信盒；但它并非可处理其他任意跨纤维约束的通用求解器。条件分母必须有一个严格大于零的下界，才能构造分式证书。当分母下界为零时，仍可使用明确标注的支持集包络回退方案。

4.8. 未来协变量演化

未来演化通过 `covariate_evolution` 进行选择。默认情况下，`"freeze"` 在整个时间范围内使用标志点取值。不带未来时间表的选项 `"locf"` 与冻结一致；使用 `newdata_future` 时，每次声明的更新都会向前应用，并持续保持到下一次更新。这类时间表属于预测情景，而不是事后导入的未来测量值。生成器片段会在其更新时间处分割，并在完整的有序传播过程中保留同一次脆弱性抽样。

对于`"joint_submodel"`，`rc_covariate_model()` 使用最大似然或限制性最大似然估计高斯线性随机系数模型，遵循 Laird和Ware (1982的纵向研究框架。预测需要 `covariate_history` 以及截至标志点的实际测量时间。重复事件区间行和潜在求积时间不视为额外的生物标志物测量。对于一个协变量，条件随机效应分布为

$$
\begin{aligned}
W_i &= X_i \beta + Z_i b_i + e_i, \quad b_i \sim N(0, D), \quad e_i \sim N(0, \sigma^2 I), \\
b_i \mid W_i &\sim N\Big(D Z_i^\top V_i^{-1} (W_i - X_i \beta),\, D - D Z_i^\top V_i^{-1} Z_i D \Big), \quad V_i = Z_i D Z_i^\top + \sigma^2 I.
\end{aligned}
\tag{15}
$$

此处 W_i 是受试者 i的既往测量向量， X_i 而 Z_i 分别为固定效应和随机效应设计矩阵， β 为固定系数向量， b_i 为随机截距及可选斜率，D 为其协方差， e_i 为测量噪声， σ^2 为其方差， I 为单位矩阵， V_i 为边际测量协方差。这些 X_i、 Z_i 是局部混合模型矩阵，并非第 3 节中的脆弱性或强度记号。高斯条件化给出了所示分布。对抽取的未来潜在协变量曲线进行传播并取平均；报告的蒙特卡洛误差是以估计的子模型参数为条件计算的。协变量子模型相互独立、未来事件不会产生反馈且演化模型设定正确，这些仍是所需假设。这并非联合估计的测量误差生存模型，也不会自动纳入纵向参数不确定性。

4.9. 替代脆弱性分布与诊断

选择 `frailty = "gamma"`、 `"lognormal"` 或 `"discrete"` 会改变混合分布，同时保留均值为一的尺度。伽马积分仍可解析求解。对数正态积分使用高斯–埃尔米特节点，结合 $\log Z \sim N(-v/2,\ v)$ 和 $v = \log(1+\phi)$，其中 Z表示脆弱性， ϕ 表示其方差。离散分布选项会估计两个至八个质量点及其权重。使用不受约束的权重参数 a_ℓ 和对数位置参数 v_ℓ 时，它采用

$$w_\ell = \frac{e^{a_\ell}}{\sum_{m=1}^{L} e^{a_m}}, \qquad z_\ell = \frac{e^{v_\ell}}{\sum_{m=1}^{L} w_m e^{v_m}}, \qquad \mathbb{P}(Z = z_\ell) = w_\ell. \tag{16}$$

这里 L 是选定的点数， ℓ， m 为这些点的索引， w_ℓ 是一个概率， z_ℓ 是一个正的脆弱性值，而 Z 是随机脆弱性。该构造施加了 $\sum_\ell w_\ell z_\ell = 1$。估计质量点混合分布的思路源自 Heckman 和 Singer (1984)，但固定有限数量的点并不等同于不受约束的非参数分布估计量。权重和归一化位置的导数会纳入似然得分和观测信息；Louis校正会同时对路径不确定性和脆弱性不确定性积分。点彼此接近或权重趋于零时，可能产生非正则拟合；不能仅因优化器停止运行，就将其视为正则拟合。对数正态拟合会逐步加倍求积节点，并通过热启动优化检查个体对数似然和平均得分的变化。若达到声明的最大值仍未收敛，则标记为积分尚未解决。这是一项数值细化诊断，并非严格的求积误差界。

`rc_frailty_diagnose()` 拟合零方差原假设，并构建经干扰参数调整的单侧方差得分。仅当正则原假设设定正确且标量方差方向可识别时，才使用其半点质量／半χ^2_1 参考分布；这与 Self 和 Liang (1987)的边界原理一致。原假设下的干扰参数拟合本身必须收敛，具有满秩正定信息矩阵、较小的平均得分，且不存在数值上的系数边界接触；否则，诊断仍会保留其统计量，但不报告 p 值。非单位个案权重采用基于导数的干扰参数校正，并要求权重固定或外生，且满足有效的加权得分中心极限定理。替代模型拟合使用相同的权重。伽马分布、对数正态分布和估计的离散替代模型之间的比较属于非嵌套或奇异比较。接口会报告这些模型的似然、参数个数、AIC 和正则性标记；但不会为这些比较指定普通的卡方似然比参考分布。AIC 是描述性指标，尤其对于非正则混合模型而言；加权目标函数值未必具有普通 AIC 的风险解释。验证补充材料提供了两个加权分支。

4.10. 校准与稀疏风险不确定性

`rc_regularize()` 在符合条件的精确历史伽马似然上，加上观测信息对数行列式的一半。更新过程始终处于正信息域内，使用解析迹得分，并通过中心差分估计惩罚曲率。这是一种局部正则化；对于此非线性脆弱性模型族，不主张其会自动实现一阶 Firth 偏差消除。

`rc_band()` 使用给定的夹心协方差抽取高斯参数扰动，计算非线性预测，并在 logit 或恒等尺度上构建同时最大值-t 置信带。这是一种频率学派近似，并非字面意义上的后验分布。无效抽样会导致明确的失败，不会被截断或替换。

加权保形反演由 `rc_conformal()`提供。它保留候选项权重并处理得分并列，遵循Tibshirani、Barber、Candès 和 Ramdas (2019年) 的加权可交换性表述。已知权重、估计权重和外部认证的权重界具有不同的输出标签。仅凭估计权重无法保证精确覆盖率。该例程对给定的候选网格进行反演；它不是通用的连续结果集搜索引擎。 `rc_pivot()`而是针对具有已知离散参数和恒定基线的独立未删失终末时间，反演模型特定的精确枢轴量。

`rc_cornish_fisher()` 则明确实现渐近分位数展开。它们的假设并未被刻意隐藏在单一的通用置信论证之后。

5. 使用coherent4cast

源代码包使用 C++17 工具链和 Rcpp 安装。建议安装的程序包支持公开数据示例、图和长篇教程。三个合成数据集规模较小，并随程序包一同发布。以下完整示例使用月度时间和明确的评估起点：

```
R> library("coherent4cast")R> data("sim_perinatal")
R> fit <- coherent4cast(~ x, sim_perinatal, event_types = c("1", "2"),
+                   time_unit = "months")R> glance(fit)
R> one <- sim_perinatal[sim_perinatal$id == 1, ]
R> pred <- predict(fit, one, landmark = 0,
+                  horizon = c(0.25, 0.5, 1), type = "all")
R> pred[, c("unit", "horizon", "target", "estimate")]
```

表2 由此工作流程生成。三个数值域使用相同的数据契约，但声明的单位和生成率不同。域名称仅用于说明，并不表示模拟协变量是临床测量的暴露或经过验证的工程失效模式。

拟合对象支持标准的 S3 方法，包括 `print()`、 `summary()`、 `coef()`、`vcov()` 和 `logLik()`。额外的 `tidy()`、 `glance()` 和 `augment()`接口用于整理系数、模型摘要和预测。预测结果包含目标名称、单位和数值诊断；图会保留这些标签。审计元数据记录调用、被省略的个体、数据指纹和假设。数据指纹用于完整性检查，并非密码学隐私保证。

时间范围	终末风险	生存	负担 1	负担 2	存活时间
0.25	0.0547	0.9453	0.1478	0.0729	0.2431
0.50	0.1062	0.8938	0.2917	0.1436	0.4729
1.00	0.2000	0.8000	0.5660	0.2774	0.8958

表 2：在评估起点对第一位合成围产期受试者的预测。负担为预期计数；限制生存以月为单位。

0.10.1 封装函数在不改变数值核心的前提下，提供无脆弱性、精确观测特例。 `rc_surv()` 接受单侧预测变量公式（其中包含已命名的时间/状态列），或接受受支持的 `Surv(time, status)` 响应。二元状态中，0 表示删失，1 表示事件。该模型采用分段指数比例风险模型，并使用完整似然；它不是 Cox 部分似然估计。显式的`frailty="none"` 别名会在主接口中将脆弱性方差固定为零。起止数据需要个体标识符，并要求从原点开始连续观测；这些封装函数目前不支持延迟进入。

`rc_competing()` 拟合声明的、相互排斥的吸收原因对应的无脆弱性原因别风险。它将每个拟合原因风险合并到一个稀疏生成矩阵中，返回`cif_<cause>`、 `survival`、全因 `risk` 和 `rmst`。因此，累积发生率并非通过将每个原因单独的生存概率补数视为其概率来获得。对于没有观测事件的声明原因，程序会明确报错，而不会为其提供常规拟合区间。联合夹心协方差使用对齐的个体得分；不会自动生成条件累积发生率置信带。竞争风险预测支持固定协变量或显式未来计划，不支持潜在纵向子模型。

对于包含声明列的数据框 `d` 和 `dc` ，调用方式如下：

```
R> fit_s <- rc_surv(~ x, d, time = "time", status = "status",
+                   unit = "id", time_unit = "years")
R> fit_c <- rc_competing(~ x, dc, time = "time", status = "cause",
+                        unit = "id", causes = c("1", "2"),
+                        censor_code = 0, time_unit = "years")
R> predict(fit_c, landmark = 0, horizon = 1, type = "all")
```

这里， `dc$cause` 将 0 用于删失，并将声明的标签用于吸收。预测表会增加 `rc_forecast` 类别，同时保留 `coherent4cast_prediction`；预测通过现有的 `predict()`通用函数请求。如果分析人员在标志点明确提供符合条件的无事件档案，则允许使用仅协变量预测档案；这些临时映射区间并非观测随访。

竞争对象 `validate()`方法用于评估某个声明原因的累积发生率。较早发生的竞争事件被视为已观测到的非病例，并在其事件时间计算删失权重；这些事件不会被重新编码为失访。默认设置仍要求边际独立删失。外部提供的条件删失概率必须符合文档中说明的事件时间与时间范围对齐要求。指标仍是留出点估计，而非自动生成的 CIF 置信带。

可执行文件 `replication/standard_regimes.R` 使用闭式概率检验一个双原因示例，并将单事件封装器应用于 survival 的 lung 数据。

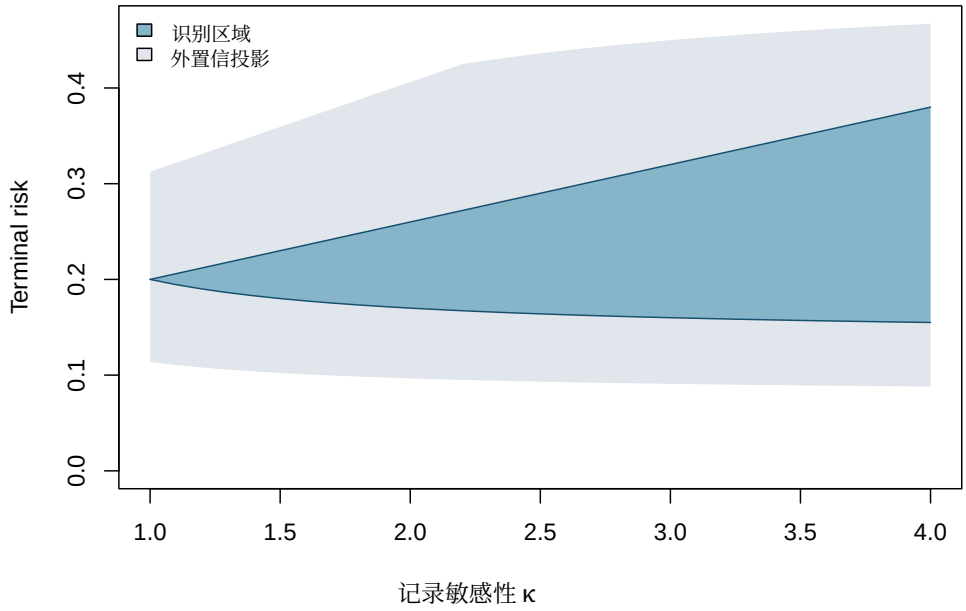


图2：三原子示例中的识别界限和同时外置信限。已知记录比例为 0.7。更宽的敏感性类别允许纳入更多完整概率律；抽样不确定性则单独进行投影。

专项补充说明报告数值一致性、留出数据指标和确切的数据编码选择；这些示例并不构成临床验证或因果验证方面的额外主张。
有限支撑敏感性计算的输入契约有所不同：

```
R> count <- c(50, 30, 20)
R> sensitivity <- rc_bounds(count / sum(count), c(0, 0, 1),
+     kappa = 2, recording_fraction = 0.7, counts = count)
R> c(sensitivity$lower, sensitivity$upper)
R> c(sensitivity$outer$lower, sensitivity$outer$upper)
```

第一个区间是在经验记录律下的识别区域；第二个区间则进一步纳入了抽样不确定性。图 2 在不改变观测计数的情况下改变敏感性参数。区间宽度描述的是本次明确界定的实验，不应被解释为不确定性的生物学分解。

5.1. 从事件记录到领域特定报告

可复用的对象是事件历史，而非疾病或行业标签。分析人员在拟合前定义独立单位、时间起点、复发标记、吸收终点、观测过程和预测时域。 `predict(..., type = "all")` 随后基于同一个拟合律给出终末风险、survival、预期复发计数和受限存活时间。预期计数并非发生概率。表3 将这些输出与报告联系起来；配套的跨领域指南提供公式、准确的字段名称、输入项和示例。

配套的复现存档还包含 `replication/research_metrics.R`，必须明确指定其来源。其 `rc_research_report()` 会添加删失加权平均预-

问题	原生输出或接口	最小报告背景
未来终点和负担	`predict()`: `risk`、`survival`、`burden_*`、`rmst`; `estimate`	单位、标志点、时间范围、事件纳入资格、时间单位和协变量演化规则。
总体推断	目标 `predict()`: `std_error`, `lower`, `upper`, `interval`, `inference_status`	总体目标与个体预测；正则性、干扰项、删失与选择限定条件。
记录敏感性	`rc_bounds()` 或 `rc_bounds_sparse()`: 端点界和可选的外投影投影	声明的支持集/纤维，记录比例、敏感度参数、分母和类别计数。
预测误差和排序	`validate()`: `Brier`, `AUROC`	独立评估单位，固定时间范围、观测事件以及删失权重。
校准	`CITL`, `calibration_slope`, `ICI`	病例/对照计数，校准拟合状态和局部平滑约定。
阈值决策	`sensitivity`, `specificity`, `PPV`, `NPV`, `net_benefit`	预设阈值，加权混淆表、删失模型以及相关的决策后果。
数值可靠性	拟合收敛/信息量；序列、混合压缩和协变量蒙特卡罗诊断	数值误差将单独报告与置信区间和预测误差分开。
外部报告计算	配套报告函数或分析员代码	成本/效用系数，精确率–召回率约定，自助法范围或模拟预言机（如适用）。

表3：当前软件包的报告契约。原生验证包含十项指标。额外的报告计算不会仅仅因为出现在报告中，就成为软件包导出项或产生实证收益。

决策、阈值曲线、比较器净获益以及可选的固定错误分类损失；`rc_value_burden()` 使用提供的成本，将预期计数转换为加性估值。这些函数是配套分析代码，并非软件包导出项。可选的独立验证单位自助法将训练好的预测器固定不变，保留未定义的抽样结果，并在不满足其规定的报告要求时不提供区间。它不提供同时置信覆盖率或精确置信覆盖率。除非分析人员提供并传播适当的联合预测抽样结果，否则估值报告不涉及联合不确定性。

在医院研究中，一个人在死亡前可能反复入院或出现临床预先规定的事件标记；报告给出预期事件计数、存活时间和经过校准的时间范围风险。在试验和政策研究中，同样的输出可用于预后分层、随访规划或评估预先规定的预测规则；仅凭随机化研究的起点，并不能将这些预测转变为治疗效应估计。在经济学或银行业中，符合条件的账户在预先规定的吸收性违约发生前可能会反复出现拖欠事件；在工程领域，运行中的资产可能在发生不可修复故障前反复发生故障。这些是对数据契约的映射，并非外部银行或工业验证。成本、违约损失率、生活质量权重和干预效应需要额外输入及适当的研究设计。

验证区分概率准确性、排序能力和决策后果。删失加权Brier评分遵循生存预测误差原理（Graf、Schmoor、Sauerbrei 和 Schumacher 1999年）；时间范围AUROC考察截至该时间范围发生事件的病例与仍未发生事件的受试者之间的区分能力（Heagerty、Lumley 和 Pepe 2000 年）。此处实现的 ICI 是相对于局部校准曲线的加权绝对差异（ Austin 和 Steyerberg 2019年）；净获益是按阈值加权的真阳性与假阳性对比（Vickers 和 Elkin 2006年）。这些指标没有任何一个能单独证明因果获益。原生 `validate()`返回点估计及已定义/状态标记，而不返回置信区间。验证集重抽样可以量化在给定拟合预测条件下的不确定性；模型开发的不确定性则需要重复完整的训练流程。按受试者整体划分数据必须避免历史信息泄漏，并且在投入实际使用前，必须在预期应用场景中重复评估。

6. 计算验证与基准测试

验证分为三个层次。代数单元测试将导数与独立数值差分进行比较，并将传播结果与闭式解进行比较。跨语言测试样例将新例程与 Version 5 研究实现中未更改的数值输出进行比较。工作流测试检查受试者分离、观测限制、概率守恒、排除未来信息以及对未定义指标的处理。表 4报告实际差异；微小数值差分可以验证实现一致性，但不能证明在任意规律下的统计覆盖率。

交叉拟合测试样例包含 800 名受试者、13 条合法路径和四个目标函数。LP 测试样例包括条件目标、粗化纤维和置信投影。这些是确定性验证示例，并非新的蒙特卡洛覆盖率研究。方法学档案中较早的大规模模拟结果仍作为参考证据保留，不会重新标记为新的软件包运行结果。表4–5 和五项案例研究保留了测得的前代版本 0.9.1 结果；新的 0.10.0 验证结果将在下文单独报告。配套脚本会在性能测量前记录源代码哈希值。

比较	最大绝对差
似然得分与数值导数的比较	2.37e-12
观测信息与得分导数的比较	3.45e-13
Louis信息与得分导数	2.81e-13
概率与解析传播	7.77e-16
奖励与解析积分	4.22e-15
有序半群一致性	1.67e-16
有限路径目标与 Python 预言机的比较	4.44e-16
个体EIF与 Python 预言机的比较	2.72e-15
条件 LP 端点与 Python 预言机	1.11e-16
外投影与 Python 预言机	3.89e-15
观测信息惩罚得分与预言机	6.68e-15

表 4：已执行的数值检验。每一行均针对其所述的测试样例和比较，不构成普遍的准确性或覆盖率保证。

受试者	历史（秒）	分数（秒）	核心拟合（秒）	峰值 GB
1,000	<0.01	<0.01	0.030	0.171
10,000	<0.01	0.030	0.320	0.156
50,000	<0.01	0.150	1.660	0.234
250,000	0.030	0.720	8.110	0.771

表5：每个样本量下五次重复运行的中位耗时（秒）。历史准备、似然与得分评估以及四参数核心拟合是彼此独立的操作。内存为新进程的峰值工作集，包括准备过程。低于计时器分辨率的耗时以界限值报告。这些是在所述主机上测得的工作负载数据；并非所有高级数据转换或任意模型规模的计时结果。

性能测试方案固定了四种样本量，并对每种样本量重复五次。分别测量历史构建、似然与得分评估以及底层优化的耗时。较大的工作负载会在不同的单元标识符下重复使用固定的合成历史数据模板；这些是计算工作负载，并非相互独立的蒙特卡洛队列。基准模型的系数维度固定且较小，因此不应将其耗时结果外推至历史数据设计规模不断增大的情况。另有一项高层拟合包含 R 数据准备。测量所用机器配备 AMD Ryzen 9 3900X、12 个物理核心和 24 个逻辑处理器，运行 Windows 11；计时的数值工作负载使用单线程。表5 报告中位数，图3 展示扩展情况。进程内存包含数据准备部分，其测量与模型对象大小分开进行。

方案规定的两秒和45秒目标，是针对相应的实测操作进行评估，而不是针对为方便起见而缩短的内核。相同模型的独立 R 实现提供了数值参考。配套软件包拟合仅作为可执行示例报告，因为其统计设定存在差异。5、15和30节点的伽马求积计算量化了对积分分辨率的敏感度；在一个剖面上的一致并不能证明统一的求积误差界。对1000名受试者执行的高层拟合，其耗时中位数为1.66秒。在本方案下未对更大规模的高层拟合进行计时，因此完整的250,000名受试者工作流程是否达到45秒目标仍未得到验证。不得用小型核心拟合和进程内存测量来替代这一缺失的比较。5节点伽马求积与30节点数值参考之间的差异

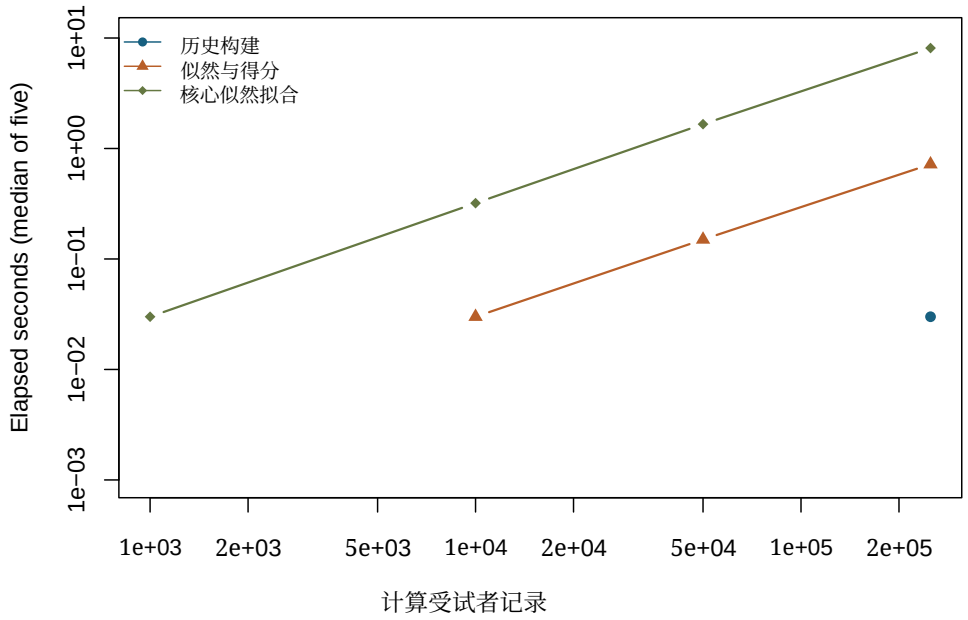


图3：冻结基准测试协议下测得的耗时。折线连接指定工作负载下五次运行的中位数；这些折线并非复杂度证明，也不是对不匹配统计模型的比较。

在测试坐标中最大为 1.48×10^{-5}，超过了登记的 10^{-5} 容差。15个节点的最大差异为 7.93×10^{-10}。默认值仍为30个节点；这一特定剖面的比较并不能证明统一的积分界。

6.1. 五项扩展的验证

表6 区分了实现一致性、数值积分诊断和经验推断验证。自动纤维生成器保留了全部929名结肠参与者和五个同日配对，生成的路径数量与专用前代适配器相同，均为7,447条。我们还将两个重叠的同标记窗口与直接并集积分进行核对，以检测排列重复计数。在一个刻意构造的大阶纤维中，进行了200次相互独立的重要性积分，每次包含2,000次尝试抽样；经验标准差为0.1513，报告的平均蒙特卡洛标准误为0.1502。该诊断正确揭示了该预算下显著的积分噪声；它并不是统计目标量的置信覆盖率。

大支持集求解器与80个相互独立的SciPy/HiGHS Charnes–Cooper问题结果一致。其中三个是可行箱体，而旧的稠密求解器将其判定为不可行；新的约简结果与独立判定程序一致。稀疏规模检查达到了50,000个原子和50,000个状态。这些确定性实例证明了系统能够在所声明的规模下执行，而不代表其在所有图上的通用速度。在这台机器上，包含50,000个状态的原生计算使用了估计4.40 MB的工作空间和15个泊松项。记录运行中，包含50,000个原子的端到端R界限加投影调用耗时0.06秒；编译和无关的数据准备不计入其中。

预定协变量预测与有序强度闭式计算的结果一致，差异为1.75×10^{-15}。在一个高斯未来协变量示例中，使用了 500 次共同随机效应抽样，

模块	记录的比较或诊断	结果
自动纤维	929 个结肠受试者；与专用方法相比的对数似然适配器	3.14×10^{-13}
Louis信息	生成路径与专用适配器的比较	1.14×10^{-13}
稀疏界	80个独立的HiGHS问题；最大端点差异	1.72×10^{-15}
大支持集界	50,000 个原子；分数证书间隙	4.28×10^{-18}
CSR 传播	50,000 个状态；概率级数界	3.74×10^{-14}
连续靶向	个体EIF 与未更改的 V5预言机的比较	7.77×10^{-16}
未来调度	有序预测与解析预言机	1.75×10^{-15}
对数正态脆弱性	40节点对数似然与自适应积分	1.98×10^{-13}

表6：新增的 0.10.0 版本验证。差异衡量数值一致性；级数界涉及实数算术截断。它们不同于经验置信覆盖率。

机制	受试者	原始Wald	变换高斯	数值成功率
普通	250	0.9314 (0.0036)	0.9510 (0.0031)	5000/5000
普通	1000	0.9460 (0.0032)	0.9516 (0.0030)	5000/5000
稀疏	250	0.5582 (0.0070)	0.5942 (0.0069)	4773/5000
稀疏	1000	0.8612 (0.0049)	0.9254 (0.0037)	4972/5000

表 7：已注册的 0.10.0 验证中的同时覆盖率。括号内为二项分布蒙特卡洛标准误。原始比较先于公共零事件正则性保护机制进行；每个单元格中的全部 5,000 个队列均予以保留。

在包含 8、16 和 32 个区间的网格上，终末风险估计分别为 0.17273、0.17339 和 0.17373。对应的条件蒙特卡洛标准误分别为 0.00088、0.00089 和 0.00090。网格细化和蒙特卡洛积分是两种不同的误差来源；二者都不是拟合模型的置信区间。使用 40 个对数正态求积节点时，固定参数对数似然与自适应数值积分所得结果相差 1.98×10^{-13}。

单独注册的目标估计实验在四个单元格中各生成了 5,000 个队列：样本量 250 和 1,000 分别与普通和稀疏事件率交叉组合。历史记录由不含潜在脆弱性的依赖历史的有限状态规律生成；拟合的事件干扰项采用合并估计，删失干扰项则按基线分层进行估计。四个同时靶向目标包括两个复发负担、终末风险和限制生存。全部 20,000 个生成队列均保留在分母中，包括未通过分数或正方差检验的队列。未发生任何执行异常。

变换区间在普通单元格中接近名义覆盖率，但未能修复稀疏事件失效（表7 和图 4）。在 250 个稀疏事件受试者中，同时覆盖率为 0.5942，而不是 0.95；在 1,000 个受试者中则为 0.9254。因此，该模拟支持一种有条件的正则情形接口，而非普遍的稀疏单元有效性。当某个目标存在观测到的零事件边界时，公共 API 现在会不提供正则高斯极限。该诊断是在保留原始比较之外新增的，并未事后将其作为覆盖率改善进行评分。

公共回退方案在删失情况下使用观测到的奖励下界和上界包络。对于支撑集已知的目标，霍夫丁不等式与并集界可给出同时的保守外部置信集；这一论证要求单位独立同

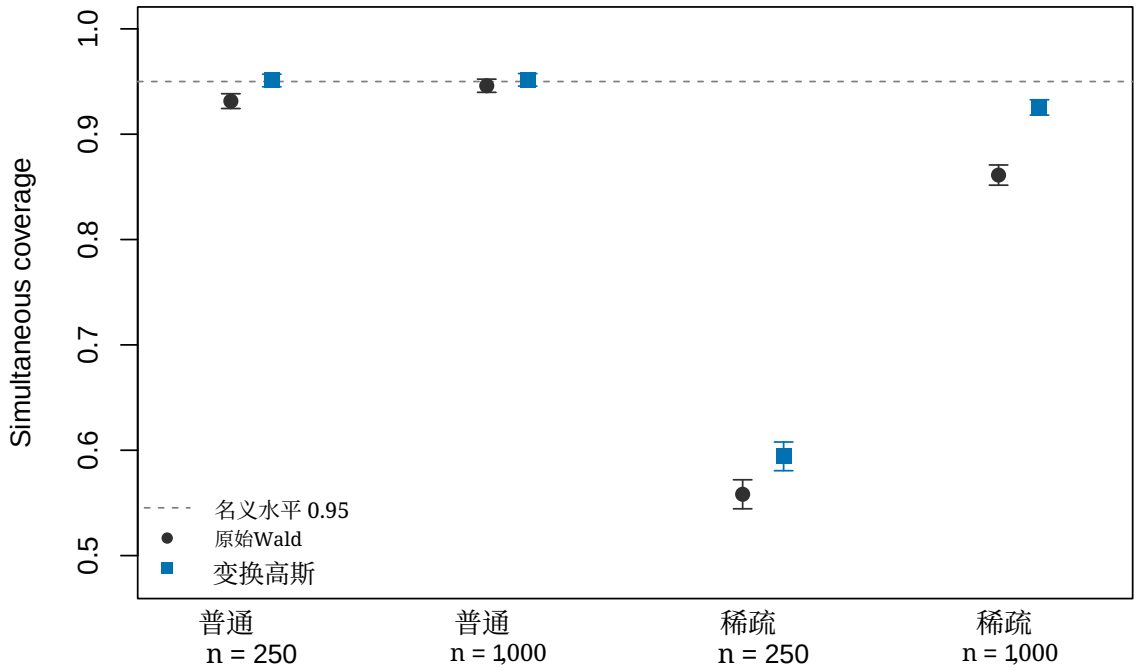


图 4：注册的同时覆盖率比较。误差线表示 1.96 倍蒙特卡洛标准误；虚线表示 0.95 的名义水平。稀疏失败仍然可见。

分布，但不要求独立删失。终末事件指标和限制生存具有此类界。未设上限的负担采用无信息外部支持 $[0, \infty)$，而非臆造有限上界。输出标记为`bounded_coarsening_projection`；它不同于精确的有限支持敏感性求解器，也无法实现点识别。在零事件边界处，明确请求的高斯方法仍不可用。若要获得所述有限样本保证，请统一请求 `interval ="projection"` 。基于数据在近似高斯区间和投影区间之间自动切换，既没有仅使用投影区间程序的无条件保证，也无法从这一论证中获得选择条件保证。数值积分容差、蒙特卡洛标准误和置信覆盖率仍是彼此独立的诊断指标。另行注册的包络实验在完全观测、独立删失和信息性删失三种情况下，各生成了 5,000 个队列。受界坐标的联合覆盖率分别为 0.9992、1.0000 和 1.0000。第一个估计值的蒙特卡洛标准误为 0.00040；两个完全覆盖的单元格的二项式蒙特卡洛下限为 0.99926，这并非确凿无疑的证据。风险区间宽度的中位数分别为 0.2960、0.6380 和 0.6480，限制生存区间宽度的中位数分别为 0.2406、0.4291 和 0.4329。这些保守集合以精度换取更弱的观测假设。15,000 个队列的检查不能取代或事后修正独立开展的 20,000 个队列高斯实验。

Case	受试者	复发事件	终末事件	脆弱性方差	状态
围产期（合成的）	160	335	59	0.485	已收敛
肿瘤学（合成的）	160	248	68	0.398	已收敛
设备（合成的）	160	398	41	0.618	已收敛
再入院	403	458	109	0.751	已收敛
结肠试验	929	468	452	0.693	已收敛

表8：全数据病例拟合。事件计数是源事件的计数，而不是区间行数或求积路径数。所有图表和验证统计量均保留所声明的观测模型。

Case	时间单位	Risk	负担 1	负担 2	存活时间
围产期（合成的）	月	0.177	0.610	0.261	0.907
肿瘤学（合成的）	年	0.231	0.525	0.208	0.880
设备（合成的）	运行小时数	0.090	0.615	0.253	0.955
再入院	年	0.087	0.316	–	0.951
结肠试验	年	0.136	0.143	–	0.929

表9：持续一个所述时间单位时的留出预测均值。负担为预期计数；存活时间以注明的单位计量。破折号表示该病例中不存在此事件类型，而非估计值为零。

7. 可复现的案例研究

每个案例均记录数据来源、估计目标、预测摘要、留出验证及局限性。表8汇总了拟合样本和诊断结果。真实数据适配器从各自的规范源软件包中获取记录。软件源代码发行版仅包含合成记录和测试夹具；汇总案例结果已足以支持稿件。图5展示了生成的预测摘要，并沿用各项分析中的时间单位和纳入资格。

7.1. 合成围产期病史

围产期示例包含两个复发标记和一个月度尺度上的吸收终点。其生成律具有共享伽马脆弱性和封顶计数反馈。应用层面的解释是，在明确的随访窗口内预期增加的发病事件数和终末风险；而数据本身仍为模拟数据，并非产妇记录。预测表依据同一拟合律报告所有目标类型。留出布里尔评分和判别能力展示了验证接口；表11 中的数值仅针对这一生成机制。标记有意采用通用定义，并不表示不同疾病之间存在临床序列。

7.2. 合成肿瘤病史

肿瘤学演示以年为单位，并采用不同的事件率配置。它区分复发事件与终末风险和受限存活时间。相同的模式和函数调用适用，无需重命名数学目标量。验证是在留出受试者上进行的，而不是在训练受试者的重复记录上进行的。此示例

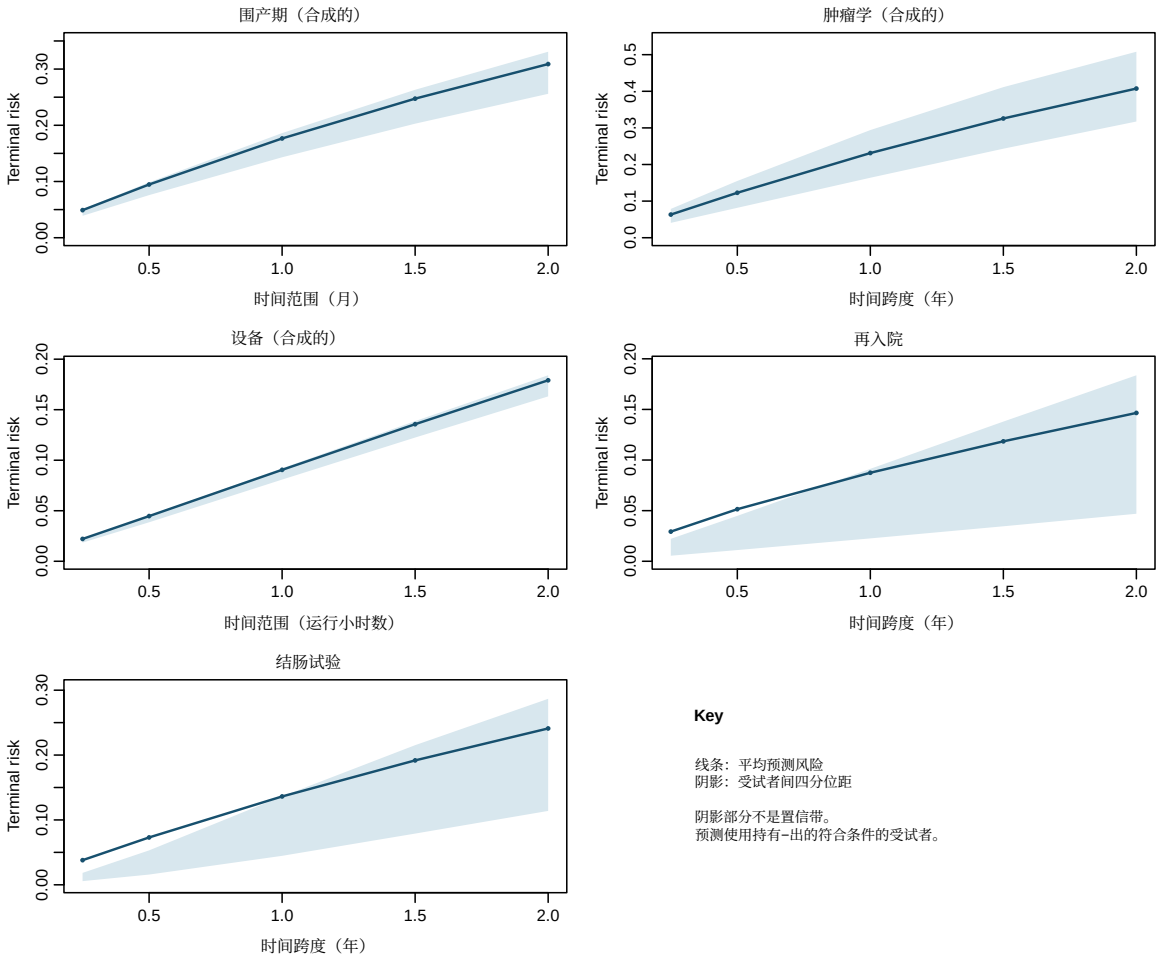


图5：五个案例研究中的普通法预测。曲线概括了在既定历史和观测规格下生成的示例。它们是预后模型输出，而非干预效应。合成案例已明确标注。

展示了如何将接口迁移到不同的时间尺度和强度配置；但这并不能证明其适用于癌症人群，也不能确定治疗效应。

7.3. 合成设备历史

机器人学演示使用终末故障发生前的运行小时数和复发性维修类型标记。预期负担是预期维修次数，限制生存则是时间范围内终末故障发生前的预期运行时间。这种解释上的变化不会改变概率计算。报告了终末故障的留出布里尔评分，并采用相同的删失约定。实际工程部署需要科学上足够完善的状态、观测模型和验证设计；通用合成标签无法满足这些要求。

脆弱性	对数似然	AIC	方差	常规拟合
伽马分布	-870.60	1775.20	1.012	yes
对数正态分布	-861.76	1757.51	1.827	yes
离散分布	-881.08	1798.16	0.929	yes

表10：在共同设计和观察方案下，基于403名受试者进行的翻修再入院敏感性分析。离散模型具有两个估计质量点。赤池信息准则仅用于描述性比较拟合的工作分布；未使用非嵌套卡方检验。

7.4. 结直肠再入院队列

`readmission` frailtypack分发的数据包含术后再入院间隔和终末死亡指标。`rc_readmission()` 保留复发事件，并在基线示例中使用首次记录的协变量值；之后的Charlson值仍作为记录字段保留，而不会被移至更早的时间点。其目标量包括预期未来入院次数、死亡风险和限制生存。再入院是医疗利用终点，而非一组彼此独立的疾病类型标记。留出集指标属于描述性内部验证。该公开数据集的观察方案和有限的协变量，无法证明观测系统之外的医疗服务也被完整记录。

新的脆弱性敏感性拟合对三个分布族均采用同一再入院设计（表10）。报告的三个拟合结果均通过了正则性诊断。在本次分析中，拟合的对数正态模型具有最小的赤池信息准则。共享的零方差得分统计量约为29.66，单侧边界参考 p 值为 2.57×10^{-8}。这表明指定的无脆弱性零假设受到了反对，但并不表示首选的混合分布为真实分布。该分析独立于前例，并保留其自身的会话和拟合对象。

对数正态比较使用的是精细拟合，而不是其初步的40节点优化器结果。从320个节点增加到640个节点后，个体对数似然的变化至多为 5.90×10^{-10}，平均得分分量的变化至多为 2.00×10^{-11}。早期的精细化未达到声明的容差，相关结果仍保留在审计表中。因此，不能将优化器收敛视为混合积分已收敛的证据。

7.5. 按日粗化的事件时间的结肠试验

survival 中的 `colon` data 提供了一项随机辅助治疗试验中的配对复发和死亡结局。该适配器保留所有参与者，并将复发视为不可重复的标记。修订后的自动生成器和专用适配器在合法区域 $d-1 < t_R < t_D \le d$ 上对同日复发和死亡进行积分，其中 d 是记录的日期。这一约定保留了两个事件，但无法恢复它们在当日内的实际发生时间。积分区间在基线不连续点处分割。该应用比较二元 `node4` 协变量与连续 `nodes`，并按照适配器元数据中的说明进行缺失值处理，同时检查求积细化。

试验的随机化起源并不意味着这些拟合得到的预后对比就是因果估计。似然和预测仍取决于基线、脆弱性、观测和纳入资格的设定。在较早标志时间点进行的预测只使用标志时间点前纤维。用于拟合模型的全随访求积路径，不能被当作在……时已经观测到的路径而重复使用

Case	符合条件	事件	布赖尔分数	受试者工作特征曲线下面积	CITL	斜率
围产期（合成的）	31	3	0.114	0.379	-0.611	-2.058
肿瘤学（合成的）	30	6	0.135	0.701	-0.196	1.811
设备（合成的）	28	2	0.075	0.557	-0.016	-1.401
再入院	65	6	0.066	0.743	0.170	0.733
结肠试验	173	21	0.069	0.871	-0.176	1.278

表11：留出数据终末风险验证。按排序后每五个受试者标识符留出一个；其余受试者用于训练模型。这些小型确定性划分用于说明工作流程，并非外部验证。尽管合成的围产期判别结果的受试者工作特征曲线下面积较低，仍予以保留。

CITL 表示总体校准。预测时长为一个规定的时间单位；标志时间点和删失假设记录在随附的指标文件中。删失权重采用边际 Kaplan–Meier估计。观测事件较少，因此判别能力和校准结果的精确性有限。

那个较早时间点。表 11以及补充输出文件报告了验证、收敛和敏感性诊断，且未将难以拟合的结果排除在其分母之外。

7.6. 基线治疗因果工作流程

配套的 `causal/run_causal.R`生成了 1,600 名合成受试者，每名受试者有一个观测到的二元基线混杂因素、两种复发事件类型、一种吸收事件，并存在右删失。基线治疗概率根据混杂因素取 0.3 或 0.7。事件强度取决于治疗、混杂因素和保留的复发次数。治疗和事件的生成过程均有明确规定；一致性和可交换性由构造保证。这是一个实现示例，而非实证干预研究。

```
R> causal_fit <- rc_causal(~x, d, treatment = "treatment",
+    horizon = 1, censor = "censored", time_unit = "years",
+    event_types = c("1", "2"),
+    spec = list(history_cap = 1L, grid = 64L), seed = 81)
R> rc_causal_contrast(causal_fit)
R> rc_causal_contrast(causal_fit, "ratio")
R> rc_audit_report(causal_fit, format = "json")
```

这里， `d` 是生成的起止数据集， `x`是观测到的混杂因素， `treatment` 是固定的基线分配。该函数在训练折内估计治疗概率。也可以通过 `propensity = 0.5` 声明已知的随机分配。终末事件会终止两个复发过程。基线混杂因素和治疗列是明确指定的输入，不会从结局标记中推断。平均得分绝对值的最大值为 1.84×10^{-9}，将时间网格从 64 加倍至 128，某个对比值的变化最多为1.36×10^{-8}。另一个独立的解析指数示例与治疗加权原生有效影响函数的差异为 4.37×10^{-13}；将治疗概率设为 1，可精确复现未改变的预后核。这些是数值检验，并非抽样覆盖率保证。

提供的余项检验在四种干扰参数配置下分别独立模拟了 100 万名受试者，共用一个生成的数据集。当事件和加权干扰项均设定错误时，观测偏倚为 0.06987，而解析乘积余项为 0.07020（蒙特卡洛标准误为 0.00044）。全正确、事件正确和权重—

目标	组 0	组 1	差异	逐点 95% 置信区间	Oracle 差异
1 型负担	0.7088	0.5659	-0.1429	[-0.2423, -0.0435]	-0.1989
2 型负担	0.3113	0.3471	0.0358	[-0.0368, 0.1084]	0.0517
终末风险	0.4078	0.3487	-0.0591	[-0.1157, -0.0024]	-0.0777
存活时间（年）	0.7922	0.8173	0.0250	[-0.0095, 0.0596]	0.0439

表12：合成基线-治疗示例。两组均按相同的
队列进行标准化。区间为影响函数正态近似。总体预言值取两个混杂因素分层的等权平均，并通过稠密矩阵独立计算。
指数分布；基线构成和结局两方面的抽样变异仍存在于
估计值中。

乘积正确这几种情况下的差异均小于一个蒙特卡洛标准误。所有配置和数值结果均予以保留；这验证了该设计下所述的单位矩阵性质，但并未声称其具有普遍的经验效率。

8. 验证与可复现性

`validate()` 在明确的时间范围评估所提供的留出风险。病例在该时间范围内发生终末事件，而对照者在此时间范围之后仍保持观察。删失的逆概率用于对布里尔评分、受试者工作特征曲线下面积和阈值表进行加权。默认的Kaplan–Meier删失估计要求边际独立删失；分析人员也可以改为提供有充分依据的条件删失概率。该函数不会自动估计此类条件模型。

总体校准通过以预测对数优势为偏置项的逻辑回归截距进行估计；单独的斜率则概括其乘数关系。综合校准指数采用对观察到的结局与预测风险进行加权平滑的方法。敏感度、特异度、阳性和阴性预测值以及净获益共用一个特定于阈值的加权表。未定义的量会保留相应标记，而不会被替换为零。这些统计量衡量的是不同属性：受试者工作特征曲线下面积不能证明校准性，布里尔评分的改善不能证明临床获益，而内部留出估计也不能替代外部验证。

配套的 `code.R`入口点可根据软件包计算结果和随附的已记录实验输出，重建稿件中的表格与图形。其默认模式会根据这些输出复现文章；单独的重新拟合脚本和实验脚本则会重新运行底层分析。它还会记录会话信息、源代码哈希值和预期输出。运行时间和内存使用量本质上取决于机器；确定性的数值输出及其容差会单独进行检查。源代码发行包包含教程、帮助页面、结构化误差和测试套件。修订检查会记录自身的版本和运行时信息，而不是继承前一项检查的结果。独立的 C++测试采用严格的编译器警告设置。随附的检查记录会标明源代码修订版本、平台、编译器和已执行的测试。

在加入因果功能之前，0.10.1 源码归档在记录的 Windows `R CMDcheck` 运行中通过了 R 4.6.2 和 R-devel 4.7.0 测试，使用 `-as-cran -no-manual`：每次运行均无错误、警告或注释。全新的 Windows二进制安装通过了 1,107 项测试断言，没有失败、警告或跳过。R/C++ 综合行覆盖率为 95.46236%（2,777

行，共 2,909 行可执行代码），未排除任何文件。六个独立的 C++17 测试样例通过了严格的编译器检查，且全部 23 个本机源代码、头文件和 Makevars 文件均与保留的 0.10.0 版本一致。这些是特定版本的计算检查；源代码行覆盖率不同于统计置信覆盖率。托管的 R-release 和 R-devel 检查也已在 Linux、macOS 和 Windows 上完成，每项任务均有 1,107 项测试断言通过，且无错误、警告或注释。四个 Linux 本机任务通过了 GCC/Clang 严格构建，分别启用和未启用 Address- Sanitizer/UndefinedBehaviorSanitizer。配套检查记录保留了所测试的源代码修订版本及日志。

前一版本0.10.1的因果修订在 Windows R-devel 下通过了全新完整的 `R CMD check -as-cran`：0 个错误、0 个警告，并出现了预期的新提交说明。全部 1,187 项测试断言、五篇长篇教程、示例以及 PDF/HTML 手册检查均已通过。另外六项 Linux、macOS 和 Windows 任务也在 R-release 和 R-devel 下通过。七个独立的原生测试样例也全部通过了四种 GCC/Clang 构建，包括启用和未启用 AddressSanitizer 与 UndefinedBehaviorSanitizer 的构建。这些检查涵盖了新增的因果核。随附的检查清单记录了提交的归档文件哈希值和经过测试的源代码提交。

复现材料提供了软件源代码、附有注释的脚本、记录的输出、随机种子和会话信息（《统计软件杂志》2026年）。源代码依据 GPL-3 许可证发布于公开代码仓库 https://github.com/setrojalab/coherent4cast。安装说明和文档可在 https://setrojalab.github.io/coherent4cast/找到。此版本提供源代码归档文件、参考手册、文章源文件和计算补充材料的链接。此次提供了重命名后的源代码归档文件供 CRAN 审核；该归档文件取代了先前使用旧软件包名称提交、目前仍待处理的版本。提供源代码归档文件并不意味着 CRAN 已予以接受。同一套材料也随期刊投稿一并提交。

对于版本固定的源代码安装，请使用：

```
remotes::install_github("setrojalab/coherent4cast", subdir = "package",
ref = "main",  dependencies = NA,  upgrade = "never")
```

源代码安装需要声明的 R 依赖项和 C++17工具链；Windows 用户需要使用与其 R 安装相适配的 Rtools 版本。安装所提供的源代码存档还会提供五篇预构建的长篇教程。

8.1. 审计导出、离线执行与研究治理

`rc_audit_report()` 从拟合对象导出列表、JSON 或文本报告，还可以选择汇总预测中的数值诊断，并对明确提供的源文件进行哈希处理。版本 3 R 序列化数据的 SHA-256 指纹用于识别所捕获的数据对象；源文件指纹用于识别文件字节。捕获时间／会话元数据与报告导出元数据有所区分。旧版 MD5 指纹仍保留标注，以实现兼容性。哈希值仅针对其所声明的表示形式建立完整性比较，并不代表数据匿名性、数字签名、不可变存储或监管批准。

在安装了合适的 R 环境、编译包及其开源依赖项后，模型和预测例程可离线运行。准备该运行环境和获取可选数据集是两个独立步骤。可复现的安装说明和机器检查有助于分析人员建立验证记录，但不能证明某个机构的运行环境符合要求。

ICH E9(R1) 要求试验分析与明确定义的治疗效果估计目标及其伴发事件策略保持一致（国际人用药品注册技术协调会 2019年）。预后 API 不会自动识别此类治疗效果。第 4.6 节中单独的基线治疗接口提供了估计目标和假设；但它不会替分析人员选择伴发事件策略。提交用分析仍需具备相应方案、估计目标、预先规定的分析、敏感性策略、来源可追溯性及适用性验证。模式映射、软件哈希值和离线执行并不构成监管资质认定。

平台状态依据已执行的日志报告。专项补充说明和随附的检查记录区分了已完成的 0.10.1 平台运行及消毒器运行，与继承自 0.10.0 的证据。

9. 讨论与开发范围

coherent4cast 为事件史似然和预测提供了通用接口，同时开放独立的观测实验，用于靶向、敏感性推断和预测校准。它在实践中的主要优势是可追溯性：分析人员可以检查纳入了哪些历史记录、哪些约束界定了边界、哪些受试者贡献了影响函数值，以及采用了哪种数值近似来生成预测。基准测试和示例展示了这些计算，但不会将机器精度解读为经验有效性。

五项工程需求均已通过明确的模块实现。对于声明的完整计数时序实验，自动纤维无需再手动枚举顺序。CSR 传播消除了稠密转移存储，而结构优化则针对指定的记录约束处理大支持集。有限路径靶向和连续时间靶向均已纳入主 API。未来协变量可以遵循时间表或估计得到的高斯子模型。替代脆弱性分布和边界诊断则使分布敏感性变得可检查。

这些统计条件仍然具有重要影响。自动计时纤维不会重建数量未知的隐藏事件。稀疏存储不会使不充分的保留状态具备马尔可夫性质。Logit 变换并不能保证稀疏的有限样本覆盖。高斯演化需要其自身的测量假设和反馈假设，而有限质量点拟合可能是非正则的。这些边界均附加在相关输出上，以便分析人员评估适用性，而不是从函数名称中推断出保证。

未来工作包括更丰富的删失模型和纵向干扰模型、更多观测适配器，以及经过单独验证的语言接口。成本和生活质量评估需要明确的奖励定义。静态基线治疗对比通过 `rc_causal()` 实现。时变策略、可预防比例和治疗建议，除这一静态扩展之外，还需要额外的因果估计目标和识别。新功能应保留其自身的观测契约、独立判定程序和验证证据。

将其组织为四个软件包的设想被记录为开发路线图，而非四个已发布的软件包。coherent4cast 将保留事件史似然、共同分布预测和共享数值核心；recurbounds 将专注于兼容集和敏感性推断；recurcausal 可将已实现的静态基线治疗接口扩展至有单独依据的纵向策略；recurmetrics 则将统筹验证、校准和决策报告。在 0.10.1 中，这些名称不会拆分代码仓库，也不会引入新的依赖项。

已实现的敏感性和验证接口仍保留在 coherent4cast 中；额外的报告脚本仍明确作为需显式调用的配套代码。任何后续拆分都必须保留各项约定、数值等效性、依赖方向和独立测试。下游软件包名称仍是提议；此处所述的基线治疗因果能力已在 coherent4cast 中实现。

计算细节

软件包名称和版本。 该软件以 coherent4cast 版本 0.10.2 发布；早期开发版本使用的名称是 recurcast。本次修订更改了软件包的单位矩阵和注册信息、文档及链接，同时保留数值内核。历史基准测试描述的是相同的数值方法。新代码使用`library(coherent4cast)`和 `coherent4cast()`；早期的拟合和控制函数名称仍作为兼容性别名保留。

本文所述软件版本为 0.10.2；本稿件于 2026年 11 月 24 日修订。前身基准测试和五病例输出保留其原始快照 0.9.0.9000 以及候选版本 0.9.1 的来源信息，运行环境为 R 4.4.2 UCRT 和 GCC 14.3，操作系统为 Windows 11。0.10.0 模块验证保留其原始来源信息；针对性的 0.10.1 API 检查则有单独记录。材料中包含 C++17 核心、GPL-3 源代码许可证、LaTeX 源文件、参考文献、已执行的因果分析小插曲以及生成的图表。记录的结果均标注了版本；原始实验输出与当前接口检查结果一并保留。

数据可用性与伦理

捆绑提供的三个示例数据集均为模拟数据，并附有生成参数和随机种子。两个临床示例使用 frailtypack 和 survival 软件包已发布的数据，并通过其文档说明的接口加载。因果示例另外生成了一个合成队列。另一个 survival 封装器示例使用 `survival::lung`；这是一个 API 操作示例，并非外部验证队列。研究未招募新的参与者。分析使用各来源软件包提供的公开示例数据集；本文不声称获得机构伦理认定。软件仓库材料中不包含患者层面临床记录。

经费与利益冲突

作者声明未获得任何经费，也不存在利益冲突。

参考文献

Adetunji SA, Adetunji OA（2026年）。“在观测不完整的情况下，对依赖历史的多类型复发事件和终末风险进行相干生存建模与动态预测。”未发表的方法学稿件，第 13 版，附补充材料 S12。

Austin PC, Steyerberg EW（2019年）。“综合校准指数（ICI）及相关指标：量化逻辑回归模型的校准情况。”《医学统计》，38(21)，4051–4065。 doi:10.1002/sim.8281。Charnes A, Cooper WW（1962年）。“线性分式泛函规划。”《海军研究后勤季刊》，9(3–4)，181–186。 doi:10.1002/nav.3800090303。Chiou SH, Xu G, Yan J, Huang CY（2023年）。“使用 R 软件包 reReg 对可能伴有信息性终末事件的复发事件进行回归建模。”《统计软件杂志》，105(5)，1–34。 doi:10.18637/jss.v105.i05。临床数据交换标准协会（2016年）。《ADaM 生存时间分析基础数据结构》，第 1.0 版。官方公开概览；访问日期：2026年11月23日，URL https://www.cdisc.org/standards/foundational/adam/adam-basic-data-structure-bds-time-event-tte-analyses-v1-0。临床数据交换标准协会（2026年）。《胰腺癌 ADTTE 生存和随访持续时间》。官方示例；访问日期：2026年11月23日，URLhttps://www.cdisc.org/kb/examples/pancreatic-cancer-adtte-survival-and-duration-follow-103759150。de Wreede LC, Fiocco M, Putter H（2011年）。“mstate：用于竞争风险和多状态模型分析的 R 软件包。”《统计软件杂志》，38(7)，1–30。doi:10.18637/jss.v038.i07。Dinkelbach W（1967年）。“论非线性分式规划。”《管理科学》，13(7)，492–498。 doi:10.1287/mnsc.13.7.492。Eddelbuettel D, François R（2011年）。“Rcpp：R 与 C++的无缝集成。”《统计软件杂志》，40(8)，1–18。 doi:10.18637/jss.v040.i08。Fox BL, Glynn PW（1988年）。“计算泊松概率。”《美国计算机协会通讯》，31(4)，440–445。 doi:10.1145/42404.42409。Graf E, Schmoor C, Sauerbrei W, Schumacher M（1999年）。“生存数据预后分类方案的评估与比较。”《医学统计》，18(17–18)，2529–2545。 doi:10.1002/(SICI)1097-0258(19990915/30)18:17/18<2529::AID-SIM274>3.0.CO;2-5。Heagerty PJ, Lumley T, Pepe MS（2000年）。“删失生存数据和诊断标志物的时间依赖 ROC 曲线。”《生物统计学》，56(2)，337–344。 doi:10.1111/j.0006-341X.2000.00337.x。Heckman JJ, Singer B（1984年）。“一种最大限度降低计量经济持续时间模型中分布假设影响的方法。”《计量经济学》，52(2)，271–320。 doi:10.2307/1911491。Hoeffding W（1963年）。“有界随机变量之和的概率不等式。”《美国统计学会杂志》，58(301)，13–30。 doi:10.1080/01621459.1963.10500830。

国际人用药品注册技术协调会（2019年）。E9(R1)：临床试验中估计目标和敏感性分析附录。最终指南于2019年11月20日通过，URL https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf。《统计软件杂志》（2026年）。作者须知与格式指南。访问于2026年11月22日，URL https://www.jstatsoft.org/authors。Król A、Mauguen A、Mazroui Y、Laurent A、Michiels S、Rondeau V（2017年）。“联合建模与预测教程：一种用于相关纵向结局、复发事件和终末事件的统计软件。”《统计软件杂志》，81(3)，1–52。doi:10.18637/jss.v081.i03。Laird NM、Ware JH（1982年）。“纵向数据的随机效应模型。”《生物统计学》，38(4)，963–974。doi:10.2307/2529876。Louis TA（1982年）。“使用EM算法时如何求观测信息矩阵。”《皇家统计学会杂志B辑》，44(2)，226–233。doi:10.1111/j.2517-6161.1982.tb01203.x。观察性健康数据科学与信息学组织（2026年）。OMOP通用数据模型，5.4版。特定版本规范；访问于2026年11月23日，URL https://ohdsi.github.io/CommonDataModel/cdm54.html。Rao V、Teh YW（2013年）。“马尔可夫跳跃过程及其扩展的快速MCMC抽样。”《机器学习研究杂志》，14(103)，3295–3320。URL https://www.jmlr.org/papers/v14/rao13a.html。Rizopoulos D、Miranda-Afonso P、Papageorgiou G（2026年）。JMbayes2：适用于纵向数据和事件发生时间数据的扩展联合模型。软件包文档，0.6-6版，URL https://cran.r-project.org/web/packages/JMbayes2/refman/JMbayes2.html。Rondeau V、Mazroui Y、Gonzalez JR（2012年）。“frailtypack：一种使用惩罚似然估计或参数估计的脆弱性模型分析相关生存数据的R软件包。”《统计软件杂志》，47(4)，1–28。doi:10.18637/jss.v047.i04。Rytgaard HC、Gerds TA、van der Laan MJ（2022年）。“干预特异性平均结局的连续时间目标最小损失估计。”《统计年刊》，50(5)，2469–2491。doi:10.1214/21-AOS2114。Self SG、Liang KY（1987年）。“非标准条件下最大似然估计量和似然比检验的渐近性质。”《美国统计学会杂志》，82(398)，605–610。doi:10.1080/01621459.1987.10478472。Therneau TM（2026年）。survival：用于R中生存分析的软件包。URL https://CRAN.R-project.org/package=survival。Tibshirani RJ、Barber RF、Candès EJ、Ramdas A（2019年）。“协变量偏移下的保形预测。”载于《神经信息处理系统进展》，第32卷。URL https://proceedings.neurips.cc/paper_files/paper/2019/hash/8fb21ee7a2207526da55a679f0332de2-Abstract.html。

Vickers AJ、Elkin EB（2006年）。“决策曲线分析：一种评估预测模型的新方法。”
Medical Decision Making，**26**(6)，565–574。doi:10.1177/0272989X06295361。

所属机构：
Sunday A.Adetunji 俄勒冈州立大学健康学院生物统
计学与流行病学系，美国俄勒冈州科瓦利斯 通讯作者：
sundayadetunjisa@gmail.com

Oluwatimileyin A. Adetunji Bowen大学管理与社会科学学
院经济系，尼日利亚奥孙州伊沃 电子邮箱：
eniola16.to@gmail.com