arXiv:2302.04992v1 [stat.AP] 10 Feb 2023

心血管疾病一级预防的最佳风险评估安排

Francesca GasperoniMRC生物统计学部 剑桥大学 英国剑桥
francesca.gasperoni@mrc-bsu.cam.ac.uk

Christopher H. JacksonMRC生物统计学部 剑桥大学 英国剑桥
chris.jackson@mrc-bsu.cam.ac.uk

Angela M.Wood公共卫生与初级保健系 剑桥大学 英国剑桥
amw79@medschl.cam.ac.uk

Michael J. Sweeting健康科学系 莱斯特大学 莱斯特, 英国
michael.sweeting@leicester.ac.uk

Paul J.NewcombeMRC生物统计学部 剑桥大学 英国剑桥
paul.newcombe@mrc-bsu.cam.ac.uk

David Stevens
利物浦心血管科学中心 利物浦大学 利物浦心脏与胸科医院 心血管与代谢医学系 生命历程与医学科学研究所 利物浦大学 英国利物浦
david.stevens@liverpool.ac.uk

Jessica K. BarrettMRC生物统计学部 剑桥大学 英国剑桥
jessica.barrett@mrc-bsu.cam.ac.uk

2023年2月13日

摘要

在本研究中，我们引入了一种由获益和成本构成的个体化、年龄特异性净获益函数，用于推荐心血管疾病预防的最佳风险评估时机。我们扩展了两阶段地标模型，以估计患者特异性心血管疾病风险特征，并对时变协变量进行调整。我们将该模型应用于临床实践研究数据链中的数据，该数据包含来自英国的初级保健电子健康记录。我们发现，对于风险较低的人群，可以建议采用5年或更长的最佳风险评估间隔。对于风险较高的人群，则需要采用时变风险因素，以区分更为频繁的评估安排。

1 引言

世界卫生组织指出，心血管疾病（CVD）是全球发病率和死亡率居首位的疾病，2016年有1790万人死于心血管疾病（占全球死亡人数的31%，世界卫生组织 [2017]）。他汀类药物和其他降脂药物被公认为最常见的心血管疾病一级预防策略，适用于

心血管疾病 [Reiner， 2013] 英国国家卫生与临床优化研究所（NICE [2014]）指南建议，为10年内患心血管疾病风险达到或超过10%的人提供20 mg阿托伐他汀。建议使用QRISK2评估工具计算10年心血管疾病风险 [Hippisley-Cox 等人，2008] 男性和女性均从40岁起每5年评估一次。然而，对于最佳风险评估策略，目前尚无普遍共识 [Pylypchuk 等人，2018年，Lalor 等人，2012年，Arnett 等人，2019年，Piepoli 等人，2016]。尤其是确定最佳的心血管疾病风险评估频率仍是一个尚待解决的问题，正如 Piepoli 等人所指出的 [2016]：“[建议偶尔重复进行心血管疾病风险评估], , 例如每5年一次，但目前没有数据可以指导这一间隔。”

在预防医学中，风险评估最佳时机的问题至关重要，并且在癌症筛查领域得到了广泛研究 [Bibbins-Domingo 等，2016年，Shieh 等，2017年，Ito 等， 2019], 但在心血管疾病风险评估领域的研究则少得多 [Selvarajah 等，2013年，Lindbohm 等， 2019]。最佳风险评估计划通常通过最大化效用函数来确定 [Rizopoulos 等，2016年，Sweeting，2017]或通过最小化成本函数来确定 [Bebu 和 Lachin， 2018]。第三种方法是采用净获益函数，其定义为获益与成本之差 [Gray 等， 2011]。这些函数针对特定疾病量身定制，因为其由被认为能够区分该特定疾病的指标构成。构建这些函数时评估的要素可能包括：获得的质量调整生命年（QALY）、预期增加的寿命年、风险评估相关成本、预期风险评估次数，以及在未确诊情况下未被发现而度过的时间。此外，最佳风险评估计划可能取决于所关注疾病的阶段 , 例如 Bebu 和 Lachin 的研究 [2018], ; 也可能取决于罹患该疾病的具体风险水平 , 例如 Lindbohm 等人的研究[2019]。为应对这一问题的动态特性，研究者对多状态模型 [Bebu 和 Lachin，2018年，Lindbohm 等，2019] 以及联合模型 [Rizopoulos 等，2016年，Sweeting， 2017] 进行了研究。但只有少数作者针对下一次筛查提出了个体化建议 [Rizopoulos 等，2016年，Sweeting，2017年，Bebu 和 Lachin， 2018]。

在本研究中，我们提出一种个体化且针对特定年龄的监测计划，旨在优化他汀类药物启用相关获益与成本之间的平衡。我们的建议基于大规模电子健康记录（EHR）数据所提供的证据。鉴于数据规模庞大且结构复杂（即重复测量稀疏且存在缺失值），联合模型和多状态模型在计算上难以实现。因此，我们采用 Van Houwelingen 和 Putter [2011] 描述的标志点框架，并在每个特定标志年龄下，最大化个体层面的净获益（NB）函数。净获益函数的构成要素包括：10年时间范围内获得的无心血管疾病生存年数（作为获益）；预期就诊次数、心血管疾病事件相关成本以及他汀类药物服用成本（作为成本）。Rapsomaniki等人 [2012] 在不同背景下提出了将无心血管疾病生存年数和他汀类药物成本用于定义净获益函数的构想（他们提出将NB作为比较不同风险预测模型的替代指标）。所提出的NB函数的一个关键要素，是每个人在每个标志点的他汀类药物启动时间。为了估计他汀类药物启动时间，我们需要为每个人在每个标志点定义动态心血管疾病风险概况。风险概况通过扩展Paige等人 [2018]提出的两阶段标志点方法进行估计。具体而言，我们扩展了第一阶段：通过线性混合效应模型及随机截距和斜率，不仅提供最佳线性无偏当前预测，还提供心血管疾病发病时变风险因素的最佳线性无偏未来预测。利用未来预测，可以制定信息更充分的风险评估策略，相较于仅基于当前风险因素的策略更具优势。

本文结构如下。第 2 节介绍用于启发研究的特征数据集。第 3 节介绍所提出的模型和方法。第 4 节分别展示男性和女性的研究结果。最后，第 5 节进行讨论。

2 启发研究的数据

用于启发研究的数据集源自临床实践研究数据链（CPRD），该数据库涵盖约 6.9% 的英国人口，并且在人口年龄和性别方面具有代表性 [Herrett 等人， 2015]。该数据集与医院事件统计（HES）中的二级医疗住院记录以及国家统计局（ONS）的全国死亡记录相关联 [Herrett 等人， 2015]。关联后的数据集包含 2,610, 264 名患者和 39,189, 729 条测量记录（即体重指数、高密度脂蛋白胆固醇、收缩压、吸烟状态和总胆固醇）。

我们排除既往患有心血管疾病或在研究入组前接受过他汀类药物治疗的人群。此外，我们还排除在研究入组日期至研究结束日期之间，体质指数（体重指数）、收缩压、总胆固醇、高密度脂蛋白（高密度脂蛋白）胆固醇或吸烟状态均无测量记录的个体。

我们将以下连续变量纳入危险因素：体重指数（体质指数）、收缩压（收缩压）、总胆固醇、高密度脂蛋白胆固醇；以及以下二元变量：吸烟状态（当前吸烟者或非当前吸烟者）、他汀类药物

消费指数、降压药物指数，以及糖尿病、肾脏疾病、抑郁症、偏头痛、严重精神疾病、类风湿关节炎和心房颤动的诊断。我们还将Townsend剥夺指数作为具有20个等级的分类变量纳入。这些危险因素之所以被选中，是因为它们属于QRISK2 [Hippisley-Cox等人， 2008] 和QRISK3 [Hippisley-Cox等人， 2017] 风险评分。CVD定义为以下任一情况：急性心肌梗死、卒中、心绞痛或短暂性脑缺血发作，与QRISK3心血管疾病风险评分中采用的定义一致 [Hippisley-Cox等， 2017]。

研究纳入了来自406家全科诊所的共1,971,002名个体（914,951名男性和1,056,051名女性）。我们随机将2/3的全科诊所（270家诊所，共1,774,220名个体）分配至推导队列数据集，将1/3的全科诊所（136家诊所，共836,044名个体）分配至验证队列数据集。有关风险因素、结局定义和队列选择的更多细节见补充材料第1节。

3 模型与方法

我们引入并优化了一个净获益（NB）函数，同时考虑与风险评估和他汀类药物启用相关的获益和成本。NB函数被定义为一个年龄和个体特异性函数，其优化结果是在每个年龄下确定用于一级预防心血管疾病的个体化风险评估计划。我们将所关注的年龄定义为标志性年龄， $L_a = \{40, 45, 50, \ldots, 80\}$。这一选择模拟了NICE目前推荐的访视计划 [2014]。在每个标志性年龄，我们选取推导集中尚未被诊断患有心血管疾病、仍然存活且尚未接受他汀类药物治疗的人群，将其定义为地标队列。每个标志性年龄代表NB评估的时间起点，而 $L_a + 10$ 代表时间范围（以年为尺度），即我们考虑的潜在心血管疾病风险评估频率从每年一次到每10年一次。

模型定义中的一个关键点是他汀类药物启用，假定其发生在5年心血管疾病风险超过5%阈值后安排的首次风险评估时。遵循 van Staa 等人 [2013], 的做法，由于缺乏随访数据，我们评估5年心血管疾病风险，而非10年心血管疾病风险。跨越5%阈值的预期时间取决于界标年龄 和个体特征，记作$t^*_{i,\ L_a}$。他汀类药物启用有助于延长无心血管疾病生存年 [Ferket 等人 2012] ，此前一项他汀类药物试验荟萃分析报告称，他汀类药物已被证实可将心血管疾病事件风险降低约 20% [流行病学研究组 2005]。随后，我们通过风险比 θ对这一效应进行建模，并将其设为0。8。关于无心血管疾病生存年的定义，详见第3.2节。

由于男性的心血管疾病发病率显著高于女性，因此推导集中的所有分析均分别针对男性和女性进行。

3.1 净获益

对于每个人 i 在 L_a，我们将最佳风险评估策略定义为在一组感兴趣的 F风险评估策略（$\boldsymbol{\tau}^f \in \{\boldsymbol{\tau}^1, \boldsymbol{\tau}^2, .., \boldsymbol{\tau}^F\}$）中，使净获益函数达到最大值的策略，即 $\boldsymbol{\tau}_i^{opt}{}_{,L_a}$，如公式（1）所示。

$$\boldsymbol{\tau}^{opt}_{i,L_a} = \underset{f \in \{1,..,F\}}{\arg\max}\ NB_{i,L_a}(\boldsymbol{\tau}^f). \tag{1}$$

风险评估计划 $\boldsymbol{\tau}^f$ 是一个由就诊时间 $\boldsymbol{\tau}^f = \{\tau_1^f, \tau_2^f, .., \tau_V^f\}$ 组成的向量，其特征由 f 表示，即两次就诊之间的间隔（例如， $f = 1$ 表示年度评估）。请注意，所有就诊时间均预先固定，并由统一的时间间隔定义；τ_1^f 始终等于起始时间 L_a，而 $_f\tau_V^f \leq L_a + 10$。预计在标志年龄 L_a，个体 i 的5年心血管疾病风险将高于5%的风险评估计划记为$t^*_{i,L_a}{}^{\tau_{k^*_{i,L_a}}}$。为避免符号过于繁杂，本节余下部分省略上标 f。

$NB_{i,\ L_a}(\boldsymbol{\tau})$以货币形式定义，即将健康结局转换为成本尺度，并在必要时扣除医疗服务使用的实际成本。健康结局以预期质量调整的无心血管疾病生存年（质量调整生命年）衡量，最长时间为 10 年；将其乘以决策者愿意为一年的完全健康支付的金额λ ，即可转换为预期成本。我们假设 λ 在每年 20,000 英镑至 30,000 英镑之间 [NICE， 2014]。预期成本由与个人预期无心血管疾病生存年（最长 10 年）相关的所有成本构成，包括在$\tau_{k^*_{i,\ L_a}}$ 之后服用他汀类药物的年费用，以及风险评估就诊的预期成本。

首先，$NB_{i,\ L_a}(\boldsymbol{\tau})$ 在公式（2）中定义。

$$NB_{i,L_a}(\boldsymbol{\tau}) = QALY(\boldsymbol{\tau}) \cdot \lambda - cost(\boldsymbol{\tau}). \tag{2}$$

$QALY(\boldsymbol{\tau})$ 基于以下要素:

- $EFLY_{NS}(\tau_{k^*_{i,\ L_a}})$：他汀类药物起始治疗前未发生心血管疾病（CVD）的时间，或未使用他汀类药物时的无事件生存年（EFLY）。该时间可通过计算从时间起点到 $\tau_{k^*_{i,\ L_a}}$期间，在未启用他汀类药物的情况下不发生心血管疾病的概率积分得出。
- $EFLY_S(\tau_{k^*_{i,\ L_a}})$：服用他汀类药物后未发生心血管疾病的时间，即服用他汀类药物时的无事件生存年（EFLY）。这段时间可通过对$\tau_{k^*_i}$之间服用他汀类药物后未发生心血管疾病的概率进行积分来计算以及时间范围。

我们假设 $EFLY_{NS}(\tau_{k^*_{i,\ L_a}})$的效用为1（完全健康），而 $EFLY_S(\tau_{k^*_{i,\ L_a}})$则伴有 u_s的效用损失。他汀类药物被视为风险极低的药物，其效用降低幅度为0至0.003，这是考虑到药片负担而设定的 [Kong 和 Zhang, 2018]。这意味着 $u_s \in [0.997, 1]$。扩展定义请参见公式（3）中的 $QALY(\boldsymbol{\tau})$。

$$QALY(\boldsymbol{\tau}) = EFLY_{NS}(\tau_{k^*_{i,L_a}}) + u_s \cdot EFLY_S(\tau_{k^*_{i,L_a}}). \tag{3}$$

预先设定的风险评估策略 $\boldsymbol{\tau}$ 相关的预期成本c_s [由他汀类药物年费用]，（英镑/年） $EFLY_S$乘以用药年数$\tau_{k^*_{i,\ L_a}}$，以及预期就诊费用组成。预期就诊费用定义为单次就诊费用 c_ν [（英镑/次就诊）]乘以预期就诊次数 $\mathbb{E}_{\boldsymbol{\tau}}[N_i]$:

$$cost(\boldsymbol{\tau}) = c_s \cdot EFLY_S(\tau_{k^*_{i,L_a}}) + c_\nu \cdot \mathbb{E}_{\boldsymbol{\tau}}[N_i]. \tag{4}$$

每生命年他汀类药物的成本 c_s 为每年 4.3 至 321.2 英镑，假设每日服用 20 mg 阿托伐他汀 [联合处方委员会 2020]。假设单次就诊费用 c_ν 为 18.39 英镑/次就诊[Kypridemos 等人 2018]。为估算 $\mathbb{E}_{\boldsymbol{\tau}}[N_i]$，我们假设心血管疾病风险评估会持续进行至时间点$\tau_{k^*_i}$（即启用他汀类药物后不再进行就诊）。 $\mathbb{E}_{\boldsymbol{\tau}}[N_i]$ 估算示例见补充材料第 2.2 节。

结合公式（2）至（4），我们能够定义与特定风险评估计划 $\boldsymbol{\tau}$ 相关的标志年龄 L_a 下第 i 位个体的净获益函数。

为计算公式（2），我们估计 $EFLY_{NS}(\tau_{k^*_{i,\ L_a}})$、 $EFLY_S(\tau_{k^*_{i,\ L_a}})$、 $EFLY_{NS}$和 $\tau_{k^*_{i,\ L_a}}$。无事件生存年通过两阶段标志分析方法进行估计，其中关注事件为 L_a至 $L_a + 10$ 之间的心血管疾病诊断（见第3.2节）。鉴于 $\tau_{k^*_{i,\ L_a}}$ 被定义为 $t^*_{i,\ L_a}$ 之后的首次就诊，问题便简化为预测 $t^*_{i,\ L_a}$。我们在第3.3节中提供了扩展的两阶段地标模型，以估计 $t^*_{i,\ L_a}$。我们在第4节中设定 λ、 u_s、 c_s、 c_ν 的取值，并进行敏感性分析，以评估分析结果对这些参数变异的稳健性（见补充材料第5节）。

可以识别出方程（1）的两种特殊情况。第一种情况是，预计某人的5年心血管疾病风险在关注的时间窗口 $[L_a, L_a + 10]$ 内不会超过5%阈值，这意味着 $\tau_{k^*_{i,\ L_a}} \geq L_a + 10$。在这种情况下，方程（1）由与预期就诊次数相关的项主导。因此，最佳风险评估策略是预期就诊次数最低的策略 $\mathbb{E}_{\boldsymbol{\tau}}[N_i]$。第二种情况是，在 L_a 时预测5年心血管疾病风险高于5%阈值，这意味着 $\tau_{k^*_{i,\ L_a}} = L_a$。我们不考虑评估这种情况，因为这些人应已在 L_a开始使用他汀类药物。

3.2 用于心血管疾病风险预测的两阶段地标法

本节将介绍我们如何应用Paige等人提出的两阶段标志点模型 [2018]，以估计在开始使用他汀类药物之前未被诊断为心血管疾病的概率。

在每个标志年龄 $L_a \in \{40, 45, .., 80\}$，我们对推导数据集中所有属于地标队列的个体拟合带有随机截距和斜率的线性混合效应模型（LMEM）。我们关注的结局是心血管疾病的时变风险因素。设 $smoke_{ij}$, SBP_{ij}, $TCHOL_{ij}$, HDL_{ij}, BMI_{ij}, BPM_{ij}, $statin_{ij}$和 age_{ij}分别表示个体 i, $i \in \{1, .., N_{L_a}\}$的吸烟状态、收缩压、总胆固醇、高密度脂蛋白胆固醇、体重指数、降压药物使用史、他汀类药物处方和年龄的重复测量值，记录于第 j, $j \in \{1, .., J_i\}$ 次访视，其中 N_{L_a}为标志队列规模。为获得最精确的估计值

模型回归参数，我们不仅纳入标志年龄之前测得的既往测量值，也纳入未来测量值 [Paige等人， 2018]。LMEM定义如下：

$$\begin{aligned} smoke_{ij} &= \beta_{10} + \beta_{11} age_{ij} + u_{10i} + u_{11i} age_{ij} + \varepsilon_{ij} \\ HDL_{ij} &= \beta_{20} + \beta_{21} age_{ij} + u_{20i} + u_{21i} age_{ij} + \varepsilon_{ij} \\ SBP_{ij} &= \beta_{30} + \beta_{31} age_{ij} + \beta_{32} BPM_{ij} + u_{30i} + u_{31i} age_{ij} + \varepsilon_{ij} \\ TCHOL_{ij} &= \beta_{40} + \beta_{41} age_{ij} + \beta_{42} statin_{ij} + u_{40i} + u_{41i} age_{ij} + \varepsilon_{ij} \\ BMI_{ij} &= \beta_{50} + \beta_{51} age_{ij} + u_{50i} + u_{51i} age_{ij} + \varepsilon_{ij} \end{aligned} \tag{5}$$

其中， $\begin{pmatrix} 0i \\ \mathbf{u}_{1i} \end{pmatrix} \sim MVN_{\mathbf{u}(\mathbf{0}, \Sigma)}$ 和 Σ为满矩阵； $\boldsymbol{\varepsilon}_{ij} \sim MVN(\mathbf{0}, \boldsymbol{\sigma}_e I)$ 和 I为单位矩阵。其中， $\boldsymbol{\beta}_0$表示每个危险因素的固定截距， $\boldsymbol{\beta}_1$ 表示每个危险因素的固定斜率。 β_{32} 表示在测量时正在使用降压药物的受试者的收缩压水平调整因子。 β_{42} 是表示他汀类药物处方对总胆固醇影响的回归参数。$\mathbf{u}_{0i}$ 和$\mathbf{u}_{1i}$ 分别是各危险因素特异的随机截距和随机斜率向量，并且在危险因素之间相关。最后， $\boldsymbol{\varepsilon}_{ij}$ 表示各危险因素不相关的残差。

我们的模型假设所有危险因素共同服从多元正态分布。对于体质指数、收缩压、总胆固醇和高密度脂蛋白胆固醇而言，这一假设是合理的；但对于作为二分类变量的吸烟状态而言，则不太合理。然而，即使不满足多元正态性，基于多元正态分布的推断通常仍可能是合理的 [Schafer， 1997]。

我们通过为地标队列中每个人 i 在时间 L_a （记为 SBP_{iL_a}, $TCHOL_{iL_a}$, $smoke_{iL_a}$, HDL_{iL_a}和 BMI_{iL_a}）预测当前危险因素值，利用$\widehat{\quad}$最佳线性无偏预测值（BLUPs），完成两阶段地标法的第一阶段。值得注意的是，计算最佳线性无偏预测值时，我们只利用既往观测，因为时变心血管疾病危险因素的预测不应依赖未来信息。因此，最佳线性无偏预测值的预测方式与预测一名新个体时相同，即我们仅观测到该个体截至标志年龄的信息。

在地标法的第二阶段，我们在每个标志年龄拟合一个Cox比例风险模型。关注的事件是未来10年内确诊心血管疾病所需的时间（在时间范围$L_a + 10$ 之后确诊心血管疾病的人将被删失）。在时间 L_a纳入Cox比例风险模型的危险因素分为两类：时间固定或时变。时间固定危险因素包括糖尿病、降压药物、肾脏疾病、抑郁症、偏头痛、严重精神疾病、类风湿关节炎、房颤诊断和Townsend贫困指数。假设这些危险因素在标志年龄 L_a 已知，且从该标志年龄起保持恒定。我们将个体 i 的这些危险因素值记为$\mathbf{x}_{i,fixed}$。时变危险因素包括体质指数、收缩压、总胆固醇、高密度脂蛋白胆固醇和吸烟状态。在时间 L_a 纳入Cox模型的值为第一阶段所得的最佳线性无偏预测值。我们将个体 i 的这些值记为$\mathbf{x}_{i,BLUP}(L_a)$。我们假设时间 L_a 的风险率如公式（6）所示。

$$\lambda^{NS}(t; \mathbf{x}_i(L_a), L_a) = \lambda_0^{NS}(t; L_a) \cdot \exp\left\{ \mathbf{x}_{i,fixed}^T \boldsymbol{\beta}_{fixed}(L_a) + \mathbf{x}_{i,BLUP}^T(L_a) \boldsymbol{\beta}_{BLUP}(L_a) \right\}. \tag{6}$$

根据公式（6），我们能够估计某人在未服用他汀类药物的情况下，到时间 t 时未被诊断为心血管疾病的概率，$S^{NS}(t; \mathbf{x}_i(L_a)\ L_a)$，即 $\Lambda_0(t;\ L_a) \exp\{\mathbf{x}_i(L_a)^T \boldsymbol{\beta}(L_a)\}$，其中 $\Lambda_0(t;\ L_a)$ 为累积基线风险，$\mathbf{x}_i(L_a)$ 是某人 i 在 L_a 时测量的所有风险因素向量。根据上一节给出的定义，我们可以将 $EFLY_{NS}(\tau_{k_{i,\ L_a}^*})$ 写成公式（7）的形式。

$$EFLY_{NS}(\tau_{k_{i,L_a}^*}) = \int_{L_a}^{\tau_{k_{i,L_a}^*}} S^{NS}(t; \mathbf{x}_i(L_a), L_a)\, dt. \tag{7}$$

类似地， $EFLY_S(\tau_{k_{i,\ L_a}^*})$ 在公式（8）中给出。

$$EFLY_S(\tau_{k_{i,L_a}^*}) = \int_{\tau_{k_{i,L_a}^*}}^{L_a+10} S^S(t; \mathbf{x}_i(L_a), L_a)\, dt \tag{8}$$

其中， $S^S(t;\mathbf{x}_i(L_a)\ L_a)$ 表示倘用他汀类药物后未被诊断为心血管疾病的概率，且等于$S^{NS}(\tau_{k_i^*}; \mathbf{x}_i(L_a))$ $\times \begin{pmatrix} S^{NS} \\ S^{NS}\ (t;\mathbf{x}_i(L_a))\,(\tau_{k_i^*};\underline{\mathbf{x}_i(L_a)}) \end{pmatrix}$。关于 $S^S(t;\ \mathbf{x}_i(L_a)\ L_a)$ 推导的完整细节，请参见补充材料第2.1节。

3.3 扩展两阶段模型以预测5%交叉时间

我们介绍了对两阶段地标模型的扩展，该扩展用于提供5年心血管疾病个体化风险概况，以预测t^*_{i,L_a}，并以个人在标志年龄 L_a时的病史为条件。

首先，我们将标志年龄 L_a （$\mathcal{P}_{L_a} = \{L_a, L_a+1, L_a+2, .., L_a+10\}$）下的预测时间集 定义为一组时间点，即在当前标志年龄之后，我们希望在这些时间点估计5年心血管疾病风险。整个地标队列中的成员不可能在1、2、3.....10年后仍然全部存活，因此，预测感兴趣时间点的风险时，不应将已死亡或已被诊断患有心血管疾病的人纳入计算 $s, s \in \mathcal{P}_{L_a}$。因此，有必要创建一个亚队列，仅纳入在每个预测时间点仍然存活且尚未被诊断患有心血管疾病的人 $s, s \in \mathcal{P}_{L_a}$。使用第3.2节中对地标队列成员拟合的线性混合效应模型（5），我们能够为地标亚队列中的每个人 i，在预测时间集 $\mathcal{P}_{L_a}$中的每个时间点 s，计算出 $\widehat{SBP}_{is}, \widehat{TCHOL}_{is}, \widehat{smoke}_{is}, \widehat{HDL}_{is}$ 和 $\widehat{BMI}_{is}$，作为其最佳线性无偏预测值（BLUPs）。

给定在每个时间点 $s \in \mathcal{P}_{L_a}$ 计算的最佳线性无偏预测值，我们在每个时间点 s，针对地标亚队列拟合 Cox比例风险模型。我们关注的是5年心血管疾病风险预测，因此，所有发生时间晚于 $s+5$、 $s \in \mathcal{P}_{L_a}$ 的事件均视为在时间 $s+5$ 删失。我们使用在每个预测时间 s、$s \in \mathcal{P}_{L_a}$ 拟合的 Cox比例风险模型，见公式（9）。该公式与公式（6）相同，区别在于：（i）起始时间 s（上一节为 L_a，此处为 $s \in \mathcal{P}_{L_a}$）；（ii）收缩压、总胆固醇、高密度脂蛋白、体质指数和吸烟的最佳线性无偏预测值（此处不仅在 L_a 进行评估，还在 $\mathcal{P}_{L_a}$ 的每个时间点进行评估）；（iii）公式（6）中分析的队列是地标队列（由 N_{L_a} 名个体组成），而公式（9）中则是地标亚队列（由 $N_{L_a,\, s}$ 名个体组成）；（iv）关注的时间窗口分别为10年和5年。

$$\lambda(t; \mathbf{x}_i(s), s) = \lambda_0(t; s) \cdot \exp\left\{\mathbf{x}^T_{i,fixed}\boldsymbol{\beta}_{fixed}(s) + \mathbf{x}^T_{i,BLUP}(s)\boldsymbol{\beta}_{BLUP}(s)\right\}$$
$$i \in \{1, .., N_{L_a,s}\} \quad s \in \mathcal{P}_{L_a} \quad s \le t \le s+5. \tag{9}$$

给定累积风险函数的 Nelson-Aalen估计量 $\hat{\Lambda}(t;\ s)$、 $s \le t \le s+5$，以及根据公式（9）得到的估计回归参数 $\hat{\boldsymbol{\beta}}(s)$、 $s \in \mathcal{P}_{L_a}$，并将$\mathbf{x}_i(s)$定义为个体 i 在时间s 测得的所有协变量向量，以及在时间 s估计的最佳线性无偏预测值，我们可以如下估计个体 i 的5年心血管疾病风险 $\hat{r}_i(s+5;\mathbf{x}_i(s)\ s)$:

$$\hat{r}(s+5; \mathbf{x}_i(s), s) = 1 - \exp\{-\hat{\Lambda}_0(s+5; s) \cdot \exp\{\mathbf{x}^T_i(s)\hat{\boldsymbol{\beta}}(s)\}\} \quad i = \{1, .., N_{L_a,s}\}, \quad s \in \mathcal{P}_{L_a}. \tag{10}$$

在每个标志年龄 L_a，我们可以计算地标队列中每个个体 i的5年心血管疾病风险向量。该向量的各元素为在每个时间 $s \in \mathcal{P}_{L_a}$估计的5年心血管疾病风险 $\hat{r}_i(s+5;\mathbf{x}_i(s)\ s)$。

最后，我们通过在首次估计出5年心血管疾病风险高于5%的年份与前一年之间进行线性插值，预测个体 i 的5年心血管疾病风险超过5%阈值的时间t^*_{i,L_a}。请注意，对于预测时间集中的任意 s，个体的风险都可能不会超过该阈值。

我们使用动态一致性指数 [Harrell Jr 等，1996 年，Van Houwelingen 和 Putter， 2011] 以及动态布里尔分数 [Graf 等，1999 年，Van Houwelingen 和 Putter， 2011]对所有预测模型进行验证。验证使用验证数据集进行，以避免过拟合。有关更多详细信息，请参阅补充材料第 4 节。

4 结果

图1报告了男性（顶行）和女性（底行）所选地标队列的规模。颜色表示不同标志年龄下5年心血管疾病风险的分类，即式（10）中的 $\hat{r}(L_a+5;\ \mathbf{x}_i(L_a)\ L_a)$。分类如下： $> 5\%$ 及以上为极高；处于区间（3.75%,5%]）内为高；处于区间（2.5% 3.75%]）内为中高；处于区间（1.25% 2.5%]）内为中低； $\le 1.25\%$ 以下为低。

请注意，女性和男性的最大标志队列规模均记录在标志年龄45岁时。这并非异常，因为人们可以在40岁之后加入研究（见第2节）。此外，随着标志年龄增加，我们观察到极高风险人群的比例上升，而低风险人群的比例下降。但请注意，在每个标志年龄计算出的亚队列规模只会减少。根据第3.1节末尾所述的考虑，我们在风险评估策略评估中排除了极高风险人群。

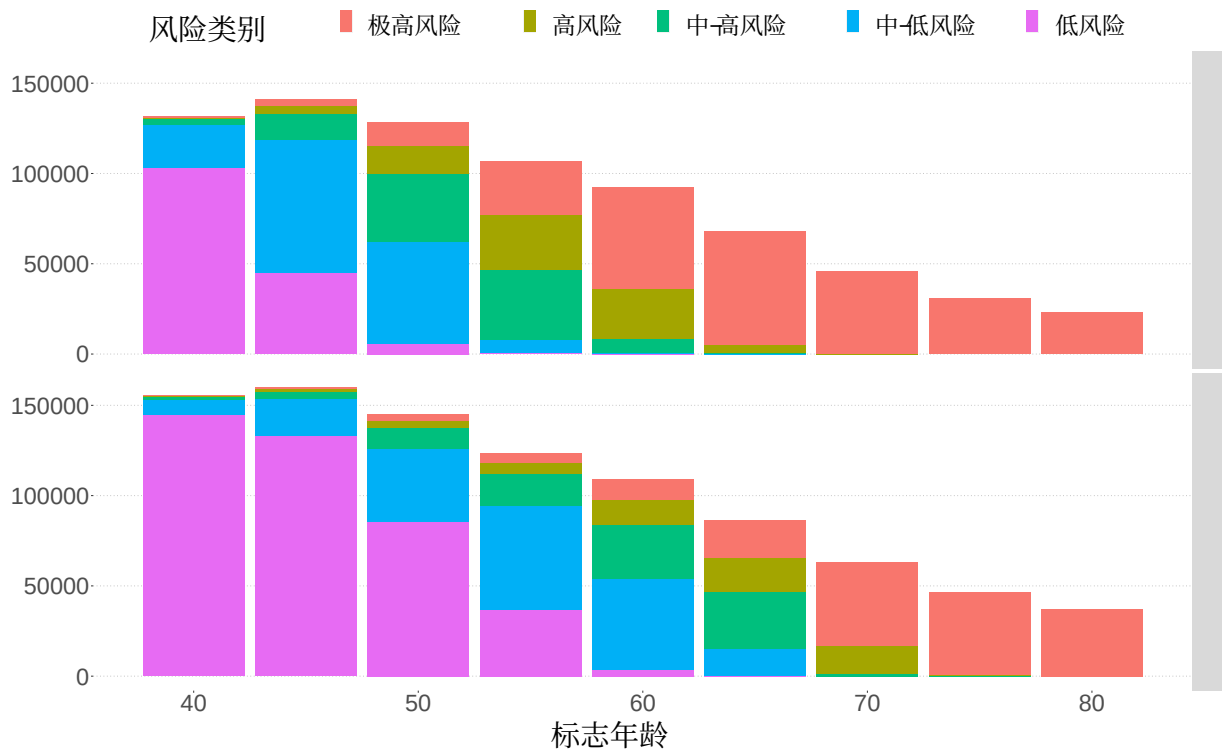

图1：推导集中各标志年龄下男性（上排）和女性（下排）各标志队列的参与者人数。每种颜色代表标志年龄时估算的5年心血管疾病风险。本文电子版中的此图为彩色。

4.1 最佳风险评估策略

在本小节中，我们介绍根据公式（1）得出的、适用于每个标志年龄所有个体的最佳风险评估策略。公式（1）的参数设置如下：$\lambda = 25,000$英镑/年，$u_s = 0.997$、$c_s = 150$英镑/年，以及$c_\nu = 18.39$英镑/次就诊。我们对$F = 10$种不同的风险评估计划$\boldsymbol{\tau}^f$、$f \in \{1, ..., 10\}$进行了公式（1）的评估。每5年进行一次风险评估的计划（$f = 5$）符合NICE指南的建议。

表1列出了女性和男性在关键年龄40岁时净获益评估的结果。我们发现，对于高风险人群，通常更倾向于采用更频繁的评估安排；而对于低风险或中低风险人群，每十年评估一次似乎较为合适。在标志年龄时被归类为低风险，是推荐采用10年风险评估策略的良好依据。然而，对于标志年龄时5年心血管疾病风险较高的个体，由于我们的预测模型利用了这些个体当前及未来预测危险因素具体数值所提供的额外信息，因此可能会给出多种风险评估建议。由表1可见，两组队列中的大多数人都被归类为低风险（女性占93.09%，男性占78.75%），而被标记为高风险的女性仅485人（0.31%），男性仅977人（0.75%）。在这一标志年龄，几乎整个女性队列（99.56%）以及96.25%的男性队列，每10年进行一次检查是最佳安排。聚焦于高风险人群，我们发现最常推荐的风险评估策略是每1、2、3或4年一次（男性中这一趋势比女性更明显）。不过，也有少数高风险人群在关键年龄40岁时每10年进行一次风险评估即可（女性104人，男性8人）。这是因为模型预测部分个体的5年心血管疾病风险曲线呈平坦趋势。补充图3展示了在$L_a = 40$被标记为高风险的女性的5年心血管疾病风险曲线。

此外，在L_a被归类为低风险的人，通常预计未来10年内不会开始使用他汀类药物。事实上，查看补充表1和补充表2可以发现，在标志年龄40岁时被标记为低风险的144,416名女性中，有143,864人的5年心血管疾病风险预计不会超过5%阈值；在标志年龄40岁时被标记为低风险的102,989名男性中，有100,361人的5年心血管疾病风险预计不会超过5%阈值。

女性和男性在所有标志性年龄下的结果概览见图2和图3。随着标志年龄增加，最常见的最佳风险评估策略会从每10年一次转变为每年一次，适用于

表1：女性和男性在标志年龄40岁时的最佳风险评估策略。女性中被归类为极高风险的有359人（0.23%），总人数为155,497人；男性中被归类为极高风险的有761人（0.58%），总人数为131,548人。

	f	高风险	中高风险	中低风险	低风险	
Women	1	76	5	0	0	81 (0.05%)
	2	88	51	4	0	143 (0.09%)
	3	67	65	12	0	144 (0.09%)
	4	5	32	27	0	64 (0.04%)
	5	76	102	0	0	178 (0.11%)
	6	69	6	0	0	75 (0.05%)
	10	104	1315	8618	144416	154453 (99.56%)
	总计	485 (0.31%)	1576 (1.02%)	8661 (5.58%)	144416 (93.09%)	
Men	1	165	3	0	0	168（0.13%）
	2	288	99	3	0	390（0.3%）
	3	252	199	10	0	461（0.35%）
	4	156	363	261	0	780（0.6%）
	5	58	766	583	0	1407（1.08%）
	6	40	637	775	0	1452（1.11%）
	7	10	236	0	0	246 (0.19%)
	10	8	767	22119	102989	125883 (96.25%)
	总计	977 (0.75%)	3070 (2.35%)	23751(18.16%)	102989(78.75%)	

两种性别。不过请注意，对于标志年龄为65岁且该年龄时5年心血管疾病风险较高的女性，最常见的最佳风险评估计划是每1至3年一次。男性也有类似情况，从标志年龄55岁开始（图3）。男性和女性之间的心血管疾病风险存在差异。图2和图3所报告的数字详见补充材料第3节。

4.2 敏感性分析：探索NB参数的影响

我们针对净效益优化的NB参数 λ、u_s、c_s c_v进行了敏感性分析。总体而言，我们发现结果对参数选择具有稳健性，观察到的预期变化较小。具体来说，λ 增加时，每10年一次的频率对较少人群而言是最佳选择，而中间频率对更大比例的人群而言成为最佳选择。对于与他汀类药物相关的效用 u_s，也可以观察到类似现象：他汀类药物对生活质量的影响越小，每10年进行一次风险评估就越不受青睐。相反，他汀类药物的价格越高，c_s，人们就越倾向于选择与较低频率就诊相关的风险评估策略。

这项研究的敏感性分析完整结果见补充材料第 5 节。 l.

5 讨论

本文介绍了一种新颖的统计方法，用于解决识别心血管疾病风险预防最佳风险评估策略这一多方面问题。全球各地均推荐了不同的心血管疾病风险预防策略，例如生活习惯/饮食调整和他汀类药物处方，以及不同的风险评估计划 [Lalor 等人，2012年，NICE，2014年，Pylypchuk 等人， 2018]。本研究聚焦于他汀类药物起始治疗，因为他汀类药物已被证实是最常见的心血管疾病预防方法 [Reiner， 2013] ；我们还聚焦于英国 NICE 指南 [NICE，2014]。

本研究的创新之处有两点：首先，我们扩展了两阶段地标模型 [Paige等人， 2018] ，以估算5年心血管疾病风险超过5%阈值的确切时间；其次，我们定义了一个净获益函数，用于区分不同的就诊方案，从而评估不同标志年龄下每个人的最佳心血管疾病风险评估计划。

两阶段地标模型的扩展包括：根据一系列关注的预测时间定义一系列标志点亚队列；估计最佳线性无偏预测值；并在每个关注的预测时间，基于固定协变量和最佳线性无偏预测值拟合Cox模型。

净获益函数基于获益（即无心血管疾病生存年数）与成本（即生活质量下降、就诊费用和他汀类药物购买费用）之间的差值，并设计为一种界标型、个体特异性的函数

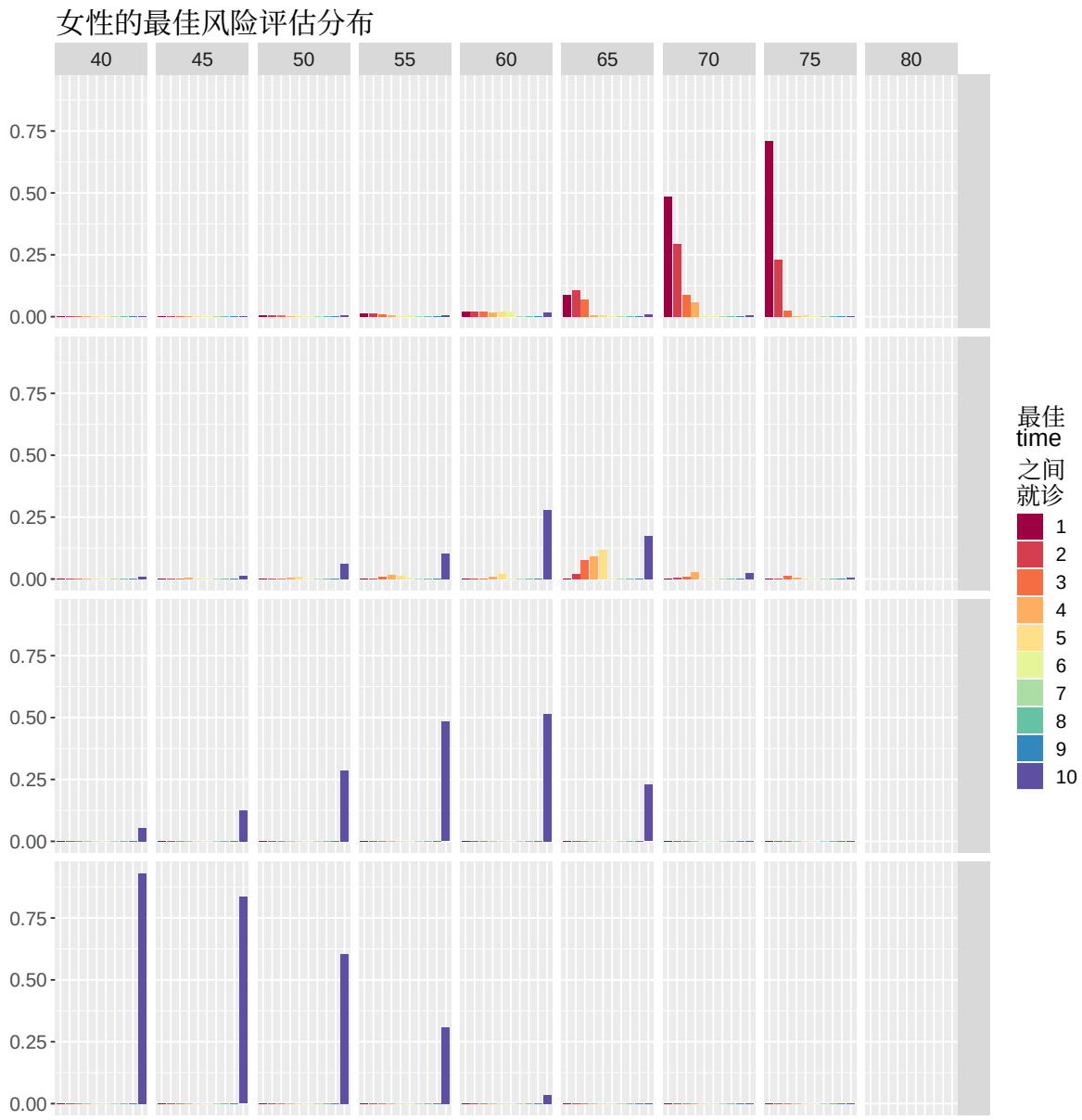

图2：女性在每个标志年龄所对应的最佳风险评估计划的比例。该图以彩色显示于本文的电子版中。

的风险评估计划 $\boldsymbol{\tau}^f$。第 i 位个体（$\boldsymbol{\tau}_i^{opt}$，$L_a$）的最佳心血管疾病风险评估计划，是与最高净获益值相关联的计划。

我们将所提出的模型应用于一个电子健康记录数据集。该数据集通过将 CPRD 数据与 HES 的二级医疗住院记录及 ONS 的死亡记录进行关联而获得。根据我们的研究结果，预计只有队列中的一部分人会超过 5%阈值，且这一人群的比例会随年龄增长而增加。由于女性的心血管疾病发病率低于男性，尤其是在较年轻时，因此，对男性和女性均从 40 岁开始每 5 年评估一次心血管疾病风险，可能不是最佳策略。利用我们的方法，我们能够针对每个人在每个界标年龄推荐最佳风险评估计划。对于 5 年风险低于 3.75% 的较低风险类别，我们发现每 10 年评估一次心血管疾病风险是最常见的最佳选择；而对于大多数界标

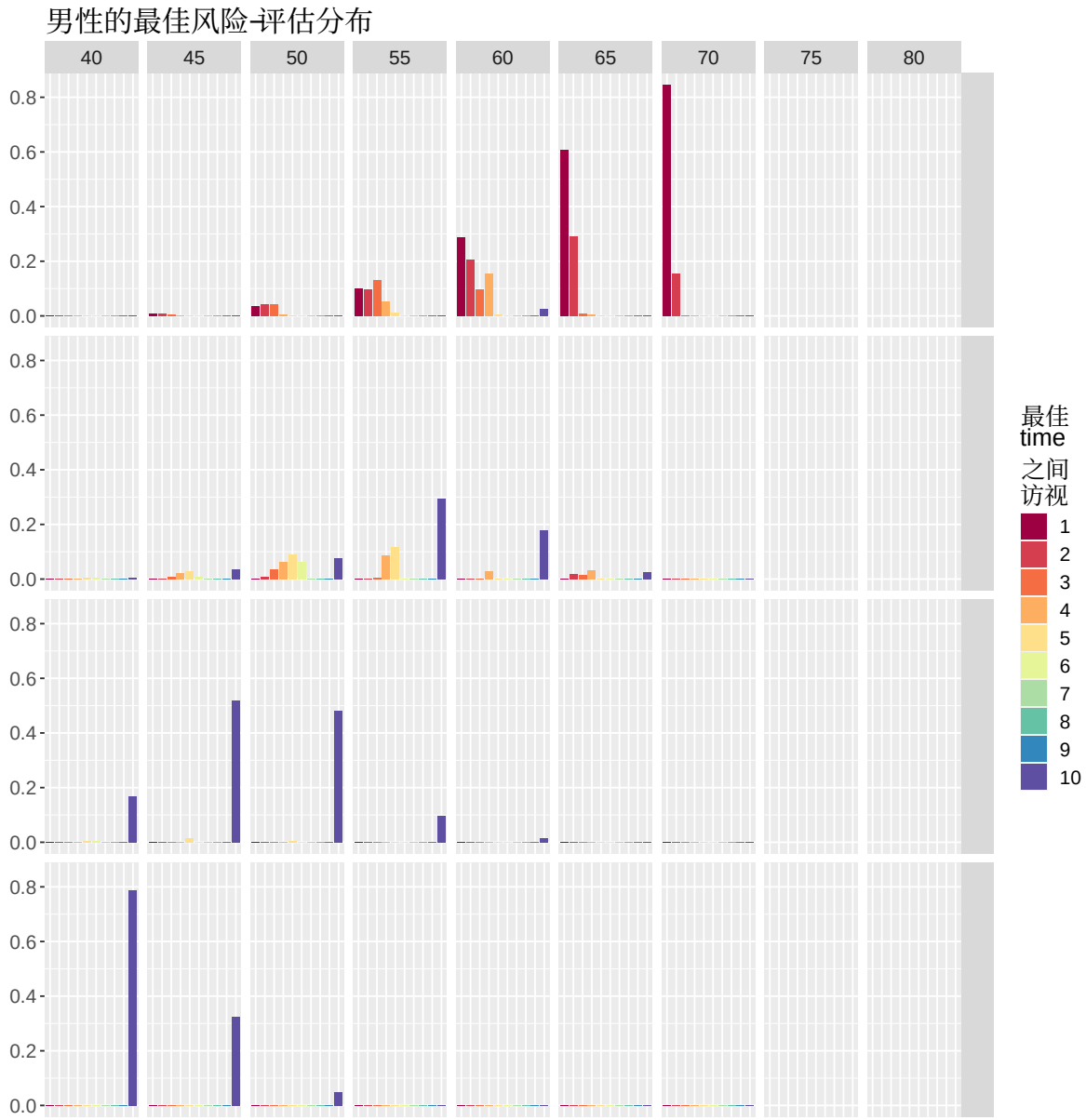

图3：男性在各标志年龄对应的最佳风险评估计划比例。此图的彩色版本见本文电子版。

风险较高的队列。 请注意，几乎所有 75 岁以上的女性和 70 岁以上的男性都被归类为极高风险 .
这与以下事实一致： 年龄是心血管疾病诊断最重要的危险因素。

为了研究这一复杂问题，我们不得不作出一些假设。这些假设可能构成本研究的局限性，但也指出了进一步研究的方向。例如，我们假设每位开始接受他汀类药物治疗的人都会完全依从，尽管他汀类药物依从性不足是一个众所周知的问题 [Simpson Jr 和 Mendys,2010]。我们模型的另一项假设是，在确定越过阈值的时间以及进行净获益计算时，均将死亡作为删失处理。这一选择符合 NICE 指南 [NICE, 2014]。第三，我们假设时变心血管疾病危险因素呈线性趋势，这可能不适合用于预测未来长达 10 年的情况。最后，我们定义了一个较为通用的净获益函数，用于为一般人群确定最佳风险评估计划。然而，该净获益函数无法处理被标记为极高

风险的人群（特定标志年龄时的5年心血管疾病风险高于 5%），因此需要针对这一人群的心血管疾病风险管理另行提出建议。

未来的研究将进一步探讨这些局限性。可以通过提供经调整的 θ，甚至提供随时间变化的 θ，来调整他汀类药物依从性不足带来的影响。通过拟合更灵活的混合效应模型，可以改进内生时变变量背后的线性假设，不过这可能需要比临床实践研究数据链数据集现有数据更完整、频率更高的测量结果。为了处理死亡竞争风险，并将罹患心血管疾病期间的生存时间纳入考量，可以定义竞争风险模型或多状态模型来评估心血管疾病特异性风险。还可以设计更复杂的净获益函数，同时考虑心血管疾病和死亡。我们的健康结局仅包括经他汀类药物生活质量调整后的最长10年无事件生存年；我们假设，NICE采用的每获得一个质量调整生命年的成本适用于这些受限结局。还可以针对老年人群设计另一种净获益函数，因为他们都被归类为极高风险。在这种情况下，风险评估策略可以建议采用的测量类型 （即血液检查、收缩压等），而不是风险评估计划。

致谢

本研究基于从英国药品和保健品管理局获许可获取的临床实践研究数据链数据（方案编号 162RMn2）。这些数据由患者提供，并由英国国家医疗服务体系在提供护理和支持的过程中收集。本研究中的阐释和结论仅代表作者本人观点。F.G. 和 J.K.B. 的经费来自英国医学研究委员会，项目编号为 MRC_MC_UU_00002/5。C.J. 的经费来自英国医学研究委员会，项目编号为 MRC_MC_UU_00002/11。M.J.S. 的经费来自英国医学研究委员会、英国心脏基金会，以及英国国家卫生研究院血液与移植研究中心（NIHR BTRU）DonorHealth andGenomics 项目（NIHR BTRU-2014-10024）。A.M.W. 获英国心脏基金会—图灵心血管数据科学奖及欧盟创新药物倡议（BigData@Heart）资助。本研究亦获艾伦·图灵研究所／英国心脏基金会（BHF）资助（资助编号 SP/18/3/33801）。心血管流行病学研究组的基础经费来自英国医学研究委员会（MR/L003120/1）、英国心脏基金会（RG/13/13/30194 和 RG/18/13/33946），以及英国国家卫生研究院剑桥生物医学研究中心（BRC-1215-20014）。为实现开放获取，作者已对本次投稿产生的任何作者接受稿版本采用知识共享署名（CC BY）许可。

补充材料

补充材料可在出版商网站上的文章在线版本中获取。本研究基于从英国药品和保健品管理局获得许可的数据，数据来自临床实践研究数据链（CPRD）（协议编号162RMn2）。本研究使用了患者提供、英国国家医疗服务体系在提供护理和支持过程中收集的数据。代码可在 https://github.com/fgaspe04/CPRD/ 公开获取。

参考文献

Donna K Arnett、Roger S Blumenthal、Michelle A Albert、Andrew B Buroker、Zachary D Goldberger、Ellen J Hahn、Cheryl Dennison Himmelfarb、Amit Khera、Donald Lloyd-Jones、J William McEvoy 等。2019年ACC/AHA心血管疾病一级预防指南：美国心脏病学会/美国心脏协会临床实践指南工作组报告。《美国心脏病学会杂志》，74(10)：e177–e232，2019年。

Ionut Bebu 和 John M Lachin。用于评估疾病进展的最佳筛查方案：以糖尿病视网膜病变为例。《生物统计学》，19(1):1–13，2018年。

Kirsten Bibbins-Domingo、David C Grossman、Susan J Curry、Karina W Davidson、John W Epling、Francisco AR García、Matthew W Gillman、Diane M Harper、Alex R Kemper、Alex H Krist 等。结直肠癌筛查：美国预防服务工作组推荐声明。《美国医学会杂志》，315(23):2564–2575，2016年。

Bart S Ferket、Bob JH van Kempen、Jan Heeringa、Sandra Spronk、Kirsten E Fleischmann、Rogier LG Nijhuis、Albert Hofman、Ewout W Steyerberg 和 MG Myriam Hunink。针对无症状个体的他汀类药物治疗终身获益个体化预测：一项建模研究。*PLoSMed*，9(12):e1001361，2012年。

Erika Graf、Claudia Schmoor、Willi Sauerbrei 和 Martin Schumacher。生存数据预后分类方案的评估与比较。《医学统计学》，18(17-18):2529–2545，1999 年。

Alastair M Gray、Philip M Clarke、Jane L Wolstenholme 和 Sarah Wordsworth。《医疗保健成本效益分析的应用方法》，第3卷。牛津大学出版社，2011年。

Frank E Harrell Jr、Kerry L Lee 和 Daniel B Mark。多变量预后模型：模型开发、评估假设与充分性，以及测量和减少误差中的问题。《医学统计学》，15(4):361–387，1996年。Emily Herrett、Arlene M Gallagher、Krishnan Bhaskaran、Harriet Forbes、Rohini Mathur、Tjeerd Van Staa 和 Liam Smeeth。数据资源概况：临床实践研究数据链（CPRD）。《国际流行病学杂志》，44 (3):827–836，2015年。Julia Hippisley-Cox、Carol Coupland、Yana Vinogradova、John Robson、Rubin Minhas、Aziz Sheikh 和 Peter Brindle。预测英格兰和威尔士的心血管风险：QRISK2的前瞻性推导与验证。《英国医学杂志》，336(7659): 1475–1482，2008年。Julia Hippisley-Cox、Carol Coupland 和 Peter Brindle。开发并验证QRISK3风险预测算法，以估算未来心血管疾病风险：前瞻性队列研究。《英国医学杂志》，357:j2099，2017年。Kazuto Ito、Ryo Oki、Yoshitaka Sekine、Seiji Arai、Yoshiyuki Miyazawa、Yasuhiro Shibata、Kazuhiro Suzuki 和 Isao Kurosawa。前列腺癌筛查：个体化筛查的历史、证据、争议与未来展望。《国际泌尿学杂志》，26(10):956–970，2019年。联合处方委员会。《英国国家处方集》（在线版）。https://bnf.nice.org.uk/medicinal-forms/ atorvastatin.html，2020年。访问日期：2022年6月20日。Nan Kong 和 Shengfan Zhang。《疾病预防与治疗中的决策分析与优化》。Wiley Online Library，2018年。Chris Kypridemos、Brendan Collins、Philip McHale、Helen Bromley、Paula Parvulescu、Simon Capewell 和 Martin O'Flaherty。英国国家医疗服务体系（NHS）健康检查心血管疾病预防项目未来的成本效益与公平性：基于利物浦（英国）数据的微观模拟建模。《公共科学图书馆·医学》，15(5)，2018年。Erin Lalor、Andrew Boyden、Dominique Cadilhac、Stephen Colagiur、Jennifer Doust、Dianne Fraser、Mark Harris、Nancy Huang、David Johnson、Greg Johnson 等。绝对心血管疾病风险管理指南。2012年。Joni V Lindbohm、Pyry N Sipilä、Nina J Mars、Jaana Pentti、Sara Ahmadi-Abhari、Eric J Brunner、Martin J Shipley、Archana Singh-Manoux、Adam G Tabak 和 Mika Kivimäki。心血管疾病预防的5年筛查间隔与风险类别特异性筛查间隔：一项队列研究。《柳叶刀·公共卫生》，4(4):e189–e199，2019年。NICE。国家临床指南中心。英国国家卫生与临床优化研究所（NICE）临床指南CG181：血脂调节、心血管风险评估及血脂调节，用于心血管疾病的一级和二级预防。https://www.nice.org.uk/guidance/cg181，2014年。2016年更新。访问日期：2022年6月20日。Ellie Paige、Jessica Barrett、David Stevens、Ruth H Keogh、Michael J Sweeting、Irwin Nazareth、Irene Petersen 和 Angela M Wood。地标模型：优化电子健康记录中风险因素重复测量的使用，以预测未来疾病风险。《美国流行病学杂志》，187(7):1530–1538，2018年。Massimo F Piepoli、Arno W Hoes、Stefan Agewall、Christian Albus、Carlos Brotons、Alberico L Catapano、Marie- Therese Cooney、Ugo Corra、Bernard Cosyns、Christi Deaton 等。2016年欧洲心血管疾病临床预防指南：欧洲心脏病学会及其他学会心血管疾病临床预防联合工作组第六次报告（由10个学会代表及受邀专家组成），并特别得到欧洲心血管预防与康复协会（EACPR）的支持。《欧洲心脏杂志》，37(29):2315–2381，2016年。Romana Pylypchuk、Sue Wells、Andrew Kerr、Katrina Poppe、Tania Riddell、Matire Harwood、Dan Exeter、Suneela Mehta、Corina Grey、Billy P Wu 等。新西兰40万名初级保健患者的心血管疾病风险预测方程：一项推导与验证研究。《柳叶刀》，391(10133):1897–1907，2018年。Eleni Rapsomaniki、Ian R White、Angela M Wood 和 Simon G Thompson。量化替代预后模型净获益的框架。《医学统计学》，31(2):114–130，2012年。Željko Reiner。他汀类药物在心血管疾病一级预防中的应用。《自然综述：心脏病学》，10(8):453，2013年。Dimitris Rizopoulos、Jeremy MG Taylor、Joost Van Rosmalen、Ewout W Steyerberg 和 Johanna JM Takkenberg。利用纵向数据与生存数据联合模型制定生物标志物个性化筛查间隔。《生物统计学》，17(1):149–164，2016年。Joseph L Schafer。《不完整多变量数据分析》。CRC出版社，1997年。

Sharmini Selvarajah、Jamaiyah Haniff、Gurpreet Kaur、Tee Guat Hiong、Adam Bujang、Kee Chee Cheong 和 Michiel L Bots。识别发展中国家心血管疾病预防的有效筛查策略：利用心血管风险评估和风险降低工具提出政策建议。*BMC*《心血管疾病》，13(1):1–10，2013年。Yiwey Shieh、Martin Eklund、Lisa Madlensky、Sarah D Sawyer、Carlie K Thompson、Allison Stover Fiscalini、Elad Ziv、Laura J van't Veer、Laura J Esserman、Jeffrey A Tice 等。精准医学时代的乳腺癌筛查：基于人群试验中的基于风险的筛查。《美国国家癌症研究所杂志》，109(5):djw290，2017年。Ross J Simpson Jr 和 Philip Mendys。他汀类药物治疗患者的依从性和持续性对临床结局的影响：系统综述。《临床脂质学杂志》，4(6):462–471，2010年。Michael J Sweeting。利用纵向数据和生存数据联合模型的预测结果，为腹主动脉瘤筛查项目确定最佳干预时间提供依据。《生物计量学杂志》，59(6):1247–1260，2017年。流行病学研究组。降胆固醇治疗的疗效和安全性：对14项他汀类药物随机试验中90,056名参与者数据的前瞻性荟萃分析。《柳叶刀》，366(9493):1267–1278，2005年。Hans Van Houwelingen 和 Hein Putter。临床生存分析中的动态预测。CRC出版社，2011年。Tjeerd-Pieter van Staa、Liam Smeeth、Edmond SW Ng、Ben Goldacre 和 Martin Gulliford。心血管风险评估的效率：是否为正确的患者提供了他汀类药物治疗？《心脏》，99(21):1597–1602，2013年。WHO。WHO心血管疾病（CVD）报告。https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)，2017年。访问日期：2022年6月20日。

arXiv:2302.04992v1 [stat.AP] 10 Feb 2023

《心血管疾病一级预防的最优风险评估安排》的补充信息

Francesca GasperoniMRC生物统计学部
剑桥大学 英国剑桥
francesca.gasperoni@mrc-bsu.cam.ac.uk

Christopher H. JacksonMRC生物统计学部 剑桥大学 英国剑桥
chris.jackson@
mrc-bsu.cam.ac.uk

Angela M.Wood公共卫生与初级保健系 剑桥大学 英国剑桥
amw79@medschl.cam.ac.uk

Michael J. Sweeting
健康科学系
莱斯特大学
英国莱斯特
michael.sweeting@leicester.ac.uk

Paul J. Newcombe
MRC生物统计学部
剑桥大学
英国剑桥
paul.newcombe@mrc-bsu.cam.ac.uk

David Stevens
利物浦心血管科学中心
利物浦大学
利物浦心脏与胸科医院
心血管与代谢医学系
生命历程与医学科学研究所
利物浦大学
英国利物浦
david.stevens@liverpool.ac.uk

Jessica K. BarrettMRC生物统计学部 剑桥大学 英国剑桥
jessica.barrett
@mrc-bsu.cam.ac.uk

2023年2月13日

1 队列选择、风险因素和结局定义

本节　　我们报告队列选择以及纳入拟议模型的变量详情　　.

图1展示了队列选择流程。最终推导数据集包含1,774,220人，分布于英国的270家全科诊所；验证数据集则包含836,044人，分布于136家全科诊所。

按照Xu等人 [2021], 的做法，我们将每个人的研究入组日期定义为以下四个日期中最晚的一个：在全科诊所注册满6个月之日；该个人年满30岁之日；该诊所数据达到标准之日 [Tate等人， 2017]；或2004年4月1日，即质量与结果框架（QOF）推出之日 [英国国家医疗服务体系， 2011]。我们将每个人的研究结束 日期定义为以下日期中最早的一个：在该诊所注销登记之日；该个人死亡之日；

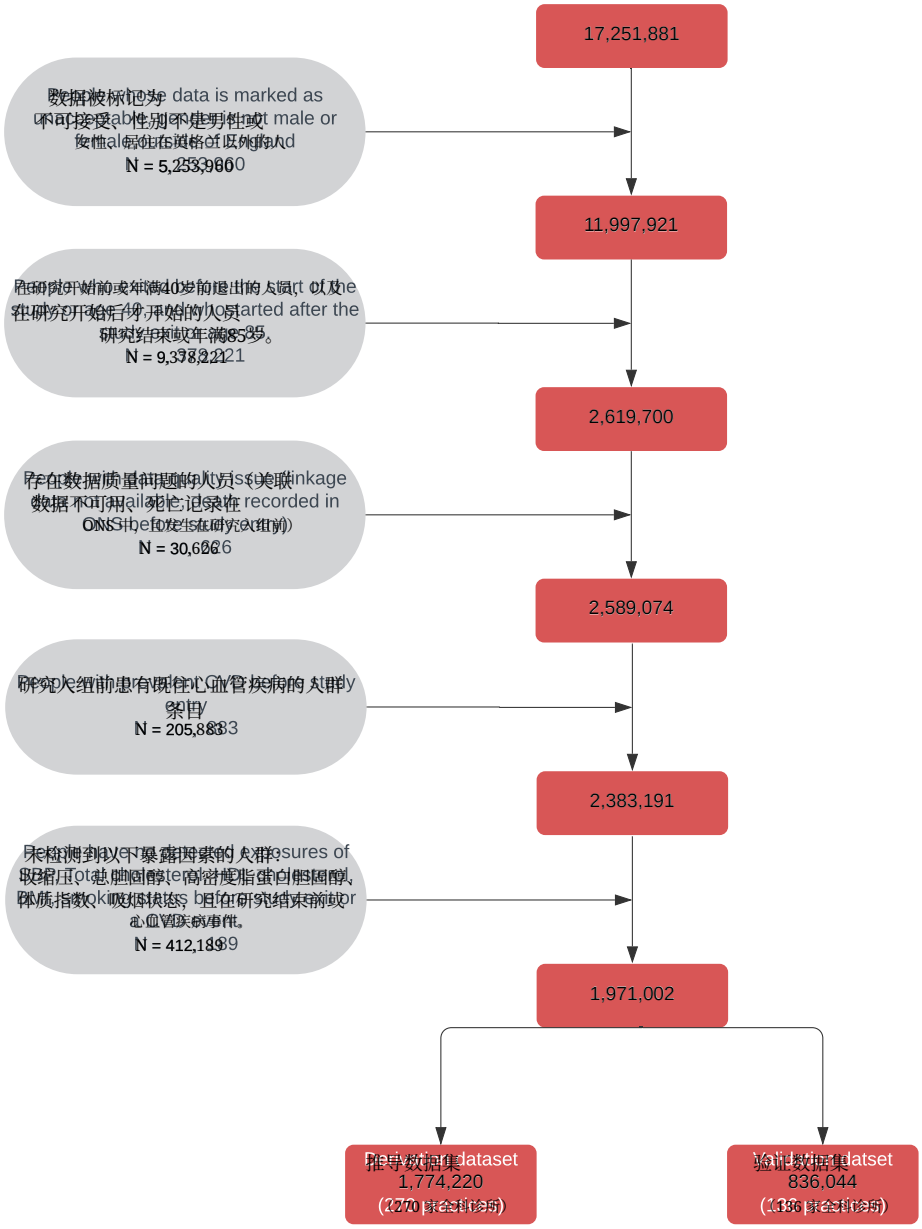

图1： 所分析队列的筛选流程图。

该个人年满95岁之日；该诊所在CPRD中的最后联系日期；或行政管理结束日期（2017年11月）。

Read编码（用于识别CPRD中的结局）和《国际疾病分类》第十次修订版编码（用于识别医院事件统计中的主要或次要诊断字段，以及国家统计局的基础或次要死因字段）见Xu等人的网络附录1、网络表1和表2 [2021]。

既往糖尿病、肾脏疾病、抑郁症、偏头痛、严重精神疾病、类风湿关节炎和心房颤动的诊断均根据 CPRD Read编码确定。降压药物（是/否）根据 CPRD 处方信息确定，并定义为首次开具处方的日期。他汀类药物起始治疗定义为首次 CPRD 处方的日期（CPRD 处方的代码列表见 Xu 等人 [2021] 的网络附录2、网络表3）。

最后，Townsend剥夺指数的范围为 1 至 20。该指数共有 1979 个缺失值（占整个队列的 0.08%），通过对每个界标取平均值进行插补。
我们将以下生物学上不合理的值设为缺失：体质指数 > 80 kg/m^2；收缩压 > 250 mmHg 或 < 60 mmHg；总胆固醇水平 > 20 mmol/L 或 < 1.75 mmol/L；高密度脂蛋白胆固醇水平 > 3.1 mmol/L 或 < 0.3 mmol/L。
此外，如果我们在标志年龄60岁之前或之后进行分析，则考虑两组不同的风险因素。在所有标志性年龄，我们都纳入体质指数、高密度脂蛋白、收缩压、总胆固醇和吸烟的最佳线性无偏预测值。在所有标志性年龄，我们都纳入降压药物、Townsend剥夺指数、既往糖尿病诊断、抑郁症、偏头痛和严重精神疾病。既往肾脏疾病、类风湿关节炎和心房颤动诊断仅在标志年龄60岁之后纳入。之所以作出这一选择，是因为这些特定疾病在较年轻年龄段极为罕见，且主稿件公式12中的Cox模型估计无法实施。

2 增量净获益函数详情

2.1 他汀类药物起始治疗前后无事件生命年的推导

我们将限制性无事件生命年（EFLY）作为获益进行评估，即考察心血管疾病诊断的时间，T，并将其限制在可观察时间窗 $[L_a, L_a + 10]$内。
为定义限制性无事件生命年，$min(T,\ L_a + 10)$，我们假设心血管疾病发生时间可表示为 $T = T^{NS} + T^S$，其中 T^{NS}是他汀类药物起始治疗前经过的无事件时间，T^S是他汀类药物起始治疗后经过的无事件时间。心血管疾病发生时间的分布定义 T应反映他汀类药物使用可延长无心血管疾病预期寿命这一事实 [Ferket 等人，2012]。为表示无事件生命年的这一增益，我们通过风险比 θ量化他汀类药物的影响。本文根据既往他汀类药物试验的荟萃分析 [流行病学研究组，2005]，设定 $\theta = 0.8$。
T定义中的不连续点将未使用他汀类药物时至心血管疾病发病的时间 T^{NS}与使用他汀类药物时的时间 T^S区分开来，定义为$\tau_{k^*_{i,\ L_a}}$，即预定访视中首次发生在$t^*_{i,\ L_a}$之后的访视；$_{L_a}$是标志年龄为 L_a时，个人 i 的5年心血管疾病风险超过5%阈值的预测时间。$\tau_{k^*_{i,\ L_a}}$ 具有界标和个人特异性，并取决于正在评估的风险评估策略 $\boldsymbol{\tau}$。

根据时间 L_a 已知的个人协变量$\tau_{k^*_i}$，我们可以定义心血管疾病发病的风险率，如式（1）所示。

$$\begin{aligned}\lambda(t; \mathbf{x}_i(L_a), L_a, \tau_{k^*_{i,L_a}}) &= \lambda^{NS}(t; \mathbf{x}_i(L_a), L_a) \cdot \mathbb{1}\{t \le \tau_{k^*_{i,L_a}}\} + \lambda^{S}(t; \mathbf{x}_i(L_a), L_a) \cdot \mathbb{1}\{t > \tau_{k^*_{i,L_a}}\} \\ &= \lambda_0(t; L_a) \cdot \exp\{\mathbf{x}_i(L_a)^T \boldsymbol{\beta}(L_a)\} \cdot \left[\mathbb{1}\{t \le \tau_{k^*_{i,L_a}}\} + \theta \cdot \mathbb{1}\{t > \tau_{k^*_{i,L_a}}\}\right], \quad t \ge L_a. \quad (1)\end{aligned}$$

我们可以按照式（2）中的说明，计算累积风险函数 $\Lambda(t;\ \mathbf{x}_i(L_a)\ L_a,\ \tau_{k^*_{i,\ L_a}})$).

$$\begin{aligned}\Lambda(t; \mathbf{x}_i(L_a), L_a, \tau_{k^*_{i,L_a}}) &= \int_{L_a}^{t} \lambda_0(u) \cdot \exp\{\mathbf{x}_i(L_a)^T \boldsymbol{\beta}(L_a)\} \cdot \left[\mathbb{1}\{u \le \tau_{k^*_{i,L_a}}\} + \theta \cdot \mathbb{1}\{u > \tau_{k^*_{i,L_a}}\}\right] du \\ &= \begin{cases} \Lambda_0(t; L_a) \exp\{\mathbf{x}_i(L_a)^T \boldsymbol{\beta}(L_a)\}, & t \le \tau_{k^*_{i,L_a}} \\ \Lambda_0(\tau_{k^*_{i,L_a}}) \exp\{\mathbf{x}_i(L_a)^T \boldsymbol{\beta}(L_a)\} + (\Lambda_0(t; L_a) - \Lambda_0(\tau_{k^*_{i,L_a}})) \cdot \theta \cdot \exp\{\mathbf{x}_i(L_a)^T \boldsymbol{\beta}(L_a)\}, & t > \tau_{k^*_{i,L_a}} \end{cases} \\ & \quad (2)\end{aligned}$$

将 $\Lambda^{NS}(t;\ \mathbf{x}_i(L_a)\ L_a)$定义为 $\Lambda_0(t;\ L_a) \exp\{\mathbf{x}_i(L_a)^T \boldsymbol{\beta}(L_a)\}$，我们可以将式（2）改写如下：

$$\Lambda(t; \mathbf{x}_i(L_a), L_a, \tau_{k^*_{i,L_a}}) = \begin{cases} \Lambda^{NS}(t; \mathbf{x}_i(L_a), L_a), & t \le \tau_{k^*_{i,L_a}} \\ \Lambda^{NS}(\tau_{k^*_{i,L_a}}; \mathbf{x}_i(L_a), L_a) + (\Lambda^{NS}(t; \mathbf{x}_i(L_a), L_a) - \Lambda^{NS}(\tau_{k^*_{i,L_a}}; \mathbf{x}_i(L_a), L_a)) \cdot \theta, & t > \tau_{k^*_{i,L_a}} \end{cases} \quad (3)$$

最后，我们在式（4）中使用生存函数 $S(t;\ \mathbf{x}_i(L_a)\ \tau_{k^*_{i,\ L_a}})$。

$$S(t; \mathbf{x}_i(L_a), L_a, \tau_{k^*_{i,L_a}}) = \exp\{-\Lambda(t; \mathbf{x}_i(L_a), \tau_{k^*_{i,L_a}}, L_a)\} = \begin{cases} S^{NS}(t; \mathbf{x}_i(L_a), L_a), & t \le \tau_{k^*_{i,L_a}} \\ S^{S}(t; \mathbf{x}_i(L_a), L_a), & t > \tau_{k^*_{i,L_a}} \end{cases} \quad (4)$$

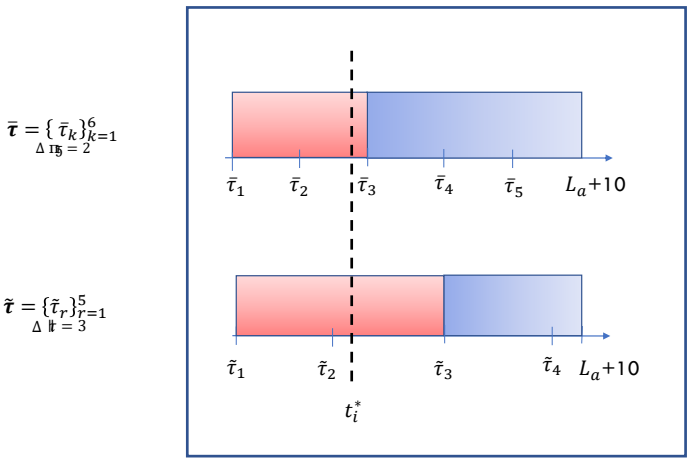

图2： 示例图，说明在给定 t_i^* 和两种特定风险评估策略的情况下，计算 k_i^*、$\tau_{k_i^*}$ 的过程。图的上半部分考虑风险评估计划 $\bar{\boldsymbol{\tau}}$（每2年进行一次访视），下半部分则考虑另一种风险评估计划 $\tilde{\boldsymbol{\tau}}$（每3年进行一次访视）。预计此人将在 t_i^* 达到5%阈值，这意味着两种情况下此人都将在第三次访视时（$k_i^* = k_i^* = 3$）开始服用他汀类药物（$\mathbb{E}_{\bar{\boldsymbol{\tau}}}[N_i] = \mathbb{E}_{\tilde{\boldsymbol{\tau}}}[N_i] = 3$）。然而，$\bar{\tau}_3$ 和 $\tilde{\tau}_3$ 并不相同，$\bar{\tau}_3 < \tilde{\tau}_3$，这意味着如果采用 $\bar{\boldsymbol{\tau}}$（每2年进行一次风险评估），服用他汀类药物的时间（蓝色方框）会更长。红色方框表示未服用他汀类药物的时间。

其中：

$$S^S(t;\mathbf{x}_i(L_a),L_a) = S^{NS}(\tau_{k^*_{i,L_a}};\mathbf{x}_i(L_a),L_a)\cdot\left(\frac{S^{NS}(t;\mathbf{x}_i(L_a),L_a)}{S^{NS}(\tau_{k^*_{i,L_a}};\mathbf{x}_i(L_a),L_a)}\right)^{\theta} \tag{5}$$

请注意，如果某人从未被开具他汀类药物，则我们有 $S(t;\mathbf{x}_i(L_a)\ L_a) = S^{NS}(t;\mathbf{x}_i(L_a)\ L_a)$.
C以 $\tau_{k^*_{i,\ L_a}}, \tau_{k^*_{i,\ L_a}} \le L_a + 10$ 为条件，我们便能够定义公式（6）中的受限无事件生命年 .

$$\begin{aligned} EFLY &= EFLY_{NS}(\tau_{k^*_{i,L_a}}) + EFLY_S(\tau_{k^*_{i,L_a}}) \\ &= \int_{L_a}^{\tau_{k^*_{i,L_a}}} S^{NS}(t;\mathbf{x}_i(L_a),L_a)\,dt + \int_{\tau_{k^*_{i,L_a}}}^{L_a+10} S^{S}(t;\mathbf{x}_i(L_a),L_a)\,dt. \end{aligned} \tag{6}$$

2.2 预期风险评估次数

预期风险评估次数 $\mathbb{E}_{\boldsymbol{\tau}}[N_i]$ 是与特定风险评估策略相关的成本的一部分$\boldsymbol{\tau}$。我们假设，对某人的心血管疾病风险评估持续进行至时间 $\tau_{k^*_{i,\ L_a}}$（即，他汀类药物启用后不再就诊）。

在图2中，我们给出一个示例，说明如何针对一名5年心血管疾病风险超过5%阈值的普通人，在 t_i^* 时（黑色虚线）计算两种不同风险评估计划（上排为 $\bar{\boldsymbol{\tau}}$，下排为 $\tilde{\boldsymbol{\tau}}$）下的预期就诊次数。按照这两种风险评估计划，此人 i 应从第三次就诊开始服用他汀类药物，这意味着 $k_i^* = k_i^* = 3$ 和 $\mathbb{E}_{\bar{\boldsymbol{\tau}}}[N_i] = \mathbb{E}_{\tilde{\boldsymbol{\tau}}}[N_i] = 3$。

如果某人的风险从未超过5%阈值，则其将永远不会开始服用他汀类药物，包括基线访视在内的预期就诊次数为 $\mathbb{E}_{\boldsymbol{\tau}}[N_i] = 1 + 10/\Delta\tau$，其中 $\Delta\tau$是根据访视计划 $\boldsymbol{\tau}$确定的两次访视之间的时间。根据图2，某人的5年心血管疾病风险从未超过5%阈值时，其预期就诊次数根据 $\bar{\boldsymbol{\tau}}$为6次，根据 $\tilde{\boldsymbol{\tau}}$为4.33次。

3 最佳风险评估计划：无心血管疾病生存年数

本节报告表1中针对不同标志年龄女性的建议风险评估策略，以及表2中针对男性的建议策略。在每张表中，我们分别根据标志年龄（列中的数值）以及在各标志年龄估算的5年心血管疾病风险类别（行中的数值）描述结果。表中每个数字对应的百分比均以标志队列为基准计算。例如，如果我们关注标志年龄为40岁的高风险女性（表1中从左数第三列），

可以看到，建议其中76名女性（155497人的0.05%）每年进行一次风险评估。此外，我们还报告了属于各风险类别的人数总计，以及其中预计未来10年内5年心血管疾病风险不会超过5%阈值的人数。这两个数字列在各风险类别区块的底部。例如，如果我们关注标志年龄为40岁的高风险女性（表1中从左数第三列），可以看到485人被标记为高风险，其中50人（10.31%）预计不会超过5%阈值。两张表的底部列出了标志队列规模，以及预计未来10年内5年心血管疾病风险不会超过5%阈值的人数总计。查看这几行可以发现，标志队列规模随标志年龄增加而减少。预计5年心血管疾病风险不会超过5%阈值的人数总计也随标志年龄增加而减少。这一行应与上面一行结合阅读；上面一行报告的是被标记为极高风险的人数总计（请注意，这些人数随标志年龄增加而增加）。女性和男性均呈现这些观察结果。

表1中的女性与表2中的男性之间，相关差异在于预计5年心血管疾病风险不会越过阈值的人数开始减少的时间（最后一行），以及被标记为极高风险的人群比例开始上升的时间（第一行）。事实上，女性中这一人数从标志年龄65岁开始减少（从34.79%降至8.18%），而男性则从标志年龄55岁开始减少（从43.01%降至5.6%）。同样，女性中标志年龄为65岁的地标队列里，近四分之一（24.62%）被标记为极高风险；男性达到相近比例的时间则早至标志年龄55岁（28.1%）。

表***1***：不同标志年龄（列）的最佳心血管疾病风险评估频率，按基线风险类别（行）分层。本表针对女性，关注的结局为*10*年心血管疾病。最后两行报告了标志队列规模，以及预计未来*10*年内*5*年心血管疾病风险不会超过*5%*阈值的人数总计。

风险类别	最佳	40	45	50	55	60	65	70	75	80
极高	-	359 (0.23%)	1119（0.7%）	3586（2.47%）	5068（4.12%）	11953（10.95%）	21274（24.62%）	46926（74.18%）	46361 (99.09%)	37178 (100%)
High	1	76 (0.05%)	405 (0.25%)	876 (0.6%)	1498 (1.22%)	1995（1.83%）	5795（6.71%）	7928（12.53%）	303（0.65%）	-
	2	88（0.06%）	465（0.29%）	969（0.67%）	1746（1.42%）	1963（1.8%）	7044（8.15%）	4798（7.58%）	98（0.21%）	-
	3	67（0.04%）	96（0.06%）	585（0.4%）	1004（0.82%）	2000（1.83%）	4434（5.13%）	1431 (2.26%)	11 (0.02%)	-
	4	5 (0%)	144 (0.09%)	163 (0.11%)	876 (0.71%)	1759（1.61%）	463（0.54%）	946（1.5%）	-	-
	5	76（0.05%）	69（0.04%）	387（0.27%）	217（0.18%）	2108（1.93%）	351（0.41%）	-	2（0%）	-
	6	69（0.04%）	128（0.08%）	263（0.18%）	491（0.4%）	1871（1.71%）	10（0.01%）	-	-	-
	7	-	7（0%）	17（0.01%）	99（0.08%）	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-
	9	-	-	-	-	-	-	-	-	-
	10	104（0.07%）	84（0.05%）	1000（0.69%）	522（0.42%）	1632（1.49%）	617 (0.71%)	113 (0.18%)	-	-
	总计	485	1398	4260	6453	13328	18714	15216	414	0
	从未交叉	50 (10.31%)	8 (0.57%)	438（10.28%）	82（1.27%）	3（0.02%）	123（0.66%）	2（0.01%）	-	-
Med-high	1	5（0%）	6 (0%)	1 (0%)	10 (0.01%)	-	-	-	-	-
	2	51 (0.03%)	235 (0.15%)	118 (0.08%)	123（0.1%）	58（0.05%）	1432（1.66%）	111（0.18%）	1（0%）	-
	3	65（0.04%）	448 (0.28%)	210 (0.14%)	950 (0.77%)	342 (0.31%)	4915 (5.69%)	151 (0.24%)	6 (0.01%)	-
	4	32 (0.02%)	762 (0.48%)	529 (0.37%)	1827 (1.48%)	772 (0.71%)	5868（6.79%）	445（0.7%）	3（0.01%）	-
	5	102（0.07%）	543（0.34%）	1178（0.81%）	1651 (1.34%)	1864 (1.71%)	7755 (8.98%)	-	-	-
	6	6 (0%)	382 (0.24%)	433 (0.3%)	807（0.66%）	2（0%）	-	-	-	-
	7	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-
	9	-	-	-	-	-	-	-	-	-
	10	1315（0.85%）	1911（1.19%）	8943（6.17%）	12354（10.03%）	27238（24.95%）	11430（13.23%）	408（0.64%）	2（0%）	-
	总计	1576	4287	11412	17722	30276	31400	1115	12	0
	从未交叉	821（52.09%）	627（14.63%）	4870（42.67%）	7729（43.61%）	1138（3.76%）	1437（4.58%）	2 (0.18%)	-	-
Med-Low	1	-	-	-	-	-	-	-	-	-
	2	4 (0%)	-	-	-	-	-	-	-	-
	3	12 (0.01%)	52 (0.03%)	1 (0%)	35 (0.03%)	-	14 (0.02%)	-	-	-
	4	27 (0.02%)	237 (0.15%)	20 (0.01%)	177 (0.14%)	12 (0.01%)	123 (0.14%)	-	-	-
	5	-	174 (0.11%)	12 (0.01%)	132 (0.11%)	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-
	7	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-
	9	-	-	-	-	-	-	-	-	-
	10	8618 (5.54%)	20071 (12.55%)	40276（27.79%）	56987（46.28%）	50104（45.89%）	14861（17.2%）	-	-	-
	总计	8661	20534	40309	57331	50116	14998	0	0	0
	从未交叉	7692 (88.81%)	16253 (79.15%)	32028 (79.46%)	54396 (94.88%)	33458 (66.76%)	5493 (36.62%)	-	-	-
Low	1	-	-	-	-	-	-	-	-	-
	2	-	-	-	-	-	-	-	-	-
	3	-	-	-	-	-	-	-	-	-
	4	-	-	-	-	-	-	-	-	-
	5	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-
	7	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-
	9	-	-	-	-	-	-	-	-	-
	10	144416（92.87%）	132637（82.91%）	85340（58.89%）	36561（29.69%）	3500（3.21%）	17（0.02%）	-	-	-
	总计	144416	132637	85340	36561	3500	17	0	0	0
	从未交叉	143864 (99.62%)	131791 (99.36%)	81119 (95.05%)	36540 (99.94%)	3384（96.69%）	17（100%）	-	-	-
	总计	155497	159975	144907	123135	109173	86403	63257	46787	37178
	从未交叉	152427（98.03%）	148679（92.94%）	118455 (81.75%)	98747 (80.19%)	37983 (34.79%)	7070 (8.18%)	4 (0.01%)	-	-

图3详细展示了女性的风险曲线，$\hat{r}_i(s+5; \mathbf{x}_i(s),\ s)$ $s \in \{40, \ldots, 50\}$，其$L_a = 40$的5年心血管疾病风险在 $L_a = 40$ 被归类为高风险。显而易见，中位数$\tau_{k^*_{i,\ 40}}$（黑色

表*2*：不同标志性年龄（列）下，按基线风险类别（行）分层的最佳心血管疾病风险评估频率。本表针对男性，关注的结局为*10*年心血管疾病。表格底部两行报告了标志队列规模，以及未来*10*年内*5*年心血管疾病风险预计不会跨越*5%*阈值的人数总计。

风险类别	最佳	40	45	50	55	60	65	70	75	80
极高	-	761 (0.58%)	4077（2.89%）	13792（10.73%）	30009（28.1%）	56178（60.93%）	63144（93.15%）	45486（99.94%）	30978 (100%)	23189 (100%)
High	1	165 (0.13%)	1385 (0.98%)	3956 (3.08%)	7634 (7.15%)	10305 (11.18%)	2816 (4.15%)	22 (0.05%)	-	-
	2	288 (0.22%)	1388 (0.98%)	4891 (3.8%)	7441 (6.97%)	7357 (7.98%)	1350 (1.99%)	4 (0.01%)	-	-
	3	252 (0.19%)	606 (0.43%)	4893 (3.81%)	9968 (9.33%)	3500 (3.8%)	32 (0.05%)	-	-	-
	4	156 (0.12%)	331 (0.23%)	669 (0.52%)	4094 (3.83%)	5584 (6.06%)	24 (0.04%)	-	-	-
	5	58 (0.04%)	386 (0.27%)	161 (0.13%)	968 (0.91%)	211 (0.23%)	-	-	-	-
	6	40 (0.03%)	82 (0.06%)	205 (0.16%)	9（0.01%）	-	-	-	-	-
	7	10（0.01%）	105（0.07%）	85（0.07%）	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-
	9	-	-	-	-	-	-	-	-	-
	10	8（0.01%）	222（0.16%）	98 (0.08%)	117 (0.11%)	909 (0.99%)	2 (0%)	-	-	-
	总计	977	4505	14958	30231	27866	4224	26	0	0
	从未交叉	1（0.1%）	139（3.09%）	70（0.47%）	12（0.04%）	61（0.22%）	-	-	-	-
Med-high	1	3（0%）	24（0.02%）	-	1（0%）	-	-	-	-	-
	2	99（0.08%）	228（0.16%）	792（0.62%）	60（0.06%）	9（0.01%）	81（0.12%）	-	-	-
	3	199（0.15%）	1211（0.86%）	4038（3.14%）	418（0.39%）	76 (0.08%)	70 (0.1%)	-	-	-
	4	363 (0.28%)	2820 (2%)	7097 (5.52%)	6711 (6.28%)	991（1.07%）	144（0.21%）	-	-	-
	5	766（0.58%）	3935（2.78%）	10357（8.06%）	9012（8.44%）	-	-	-	-	-
	6	637 (0.48%)	962 (0.68%)	7175 (5.58%)	-	-	-	-	-	-
	7	236 (0.18%)	150 (0.11%)	50 (0.04%)	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-
	9	-	-	-	-	-	-	-	-	-
	10	767（0.58%）	4954（3.51%）	8638（6.72%）	22640（21.2%）	6485（7.03%）	125（0.18%）	-	-	-
	总计	3070	14284	38147	38842	7561	420	0	0	0
	从未交叉	116 (3.78%)	1839 (12.87%)	2287 (6%)	2258 (5.81%)	937 (12.39%)	-	-	-	-
Medium-low	1	-	-	-	-	-	-	-	-	-
	2	3 (0%)	-	-	-	-	-	-	-	-
	3	10 (0.01%)	-	13 (0.01%)	1 (0%)	-	-	-	-	-
	4	261 (0.2%)	132 (0.09%)	221 (0.17%)	8 (0.01%)	-	-	-	-	-
	5	583 (0.44%)	2191 (1.55%)	501 (0.39%)	1（0%）	-	-	-	-	-
	6	775（0.59%）	237（0.17%）	38（0.03%）	-	-	-	-	-	-
	7	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-
	9	-	-	-	-	-	-	-	-	-
	10	22119（16.81%）	71349（50.49%）	55336 (43.05%)	7547 (7.07%)	592 (0.64%)	2 (0%)	-	-	-
	总计	23751	73909	56109	7557	592	2	0	0	0
	从未交叉	14996（63.14%）	59809（80.92%）	47403（84.48%）	3593（47.55%）	507（85.64%）	-	-	-	-
Low	1	-	-	-	-	-	-	-	-	-
	2	-	-	-	-	-	-	-	-	-
	3	-	-	-	-	-	-	-	-	-
	4	-	-	-	-	-	-	-	-	-
	5	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-
	7	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-
	9	-	-	-	-	-	-	-	-	-
	10	102989（78.29%）	44529 (31.51%)	5541 (4.31%)	149 (0.14%)	3 (0%)	-	-	-	-
	总计	102989	44529	5541	149	3	0	0	0	0
	从未交叉	100361 (97.45%)	44111 (99.06%)	5533 (99.86%)	114 (76.51%)	3 (100%)	-	-	-	-
	总计	131548	141304	128547	106788	92200	67790	45512	30978	23189
	从未交叉	115474 (87.78%)	105898 (74.94%)	55293 (43.01%)	5977 (5.6%)	1508 (1.64%)	-	-	-	-

实线）按照最佳频率建议而增加。事实上，预计5年心血管疾病风险将在较晚时超过5%阈值的人，更有可能被建议采用较低频率的风险评估策略。此外，值得注意的是，基线时估算的5年心血管疾病风险与风险评估建议相关，但风险曲线所提供的信息对于风险评估建议至关重要。事实上，我们观察到，不同的风险特征趋势（更陡或更平缓）可能会导致高心血管疾病风险人群得到截然相反的风险评估建议。

3.1 界标队列的描述性统计

本节报告各地标队列的描述性特征，并按标志年龄时的5年心血管疾病风险分类进行分层。女性和男性的描述性统计数据分别列于表3和表4。我们观察到，使用降压药物以及确诊抑郁症、糖尿病、偏头痛、肾脏疾病、类风湿关节炎、严重精神疾病和系统性红斑狼疮的人群风险较高。收缩压、总胆固醇（TCHOL）和体质指数较高，也与较高风险相关。汤森德20指数和吸烟同样与较高风险相关。

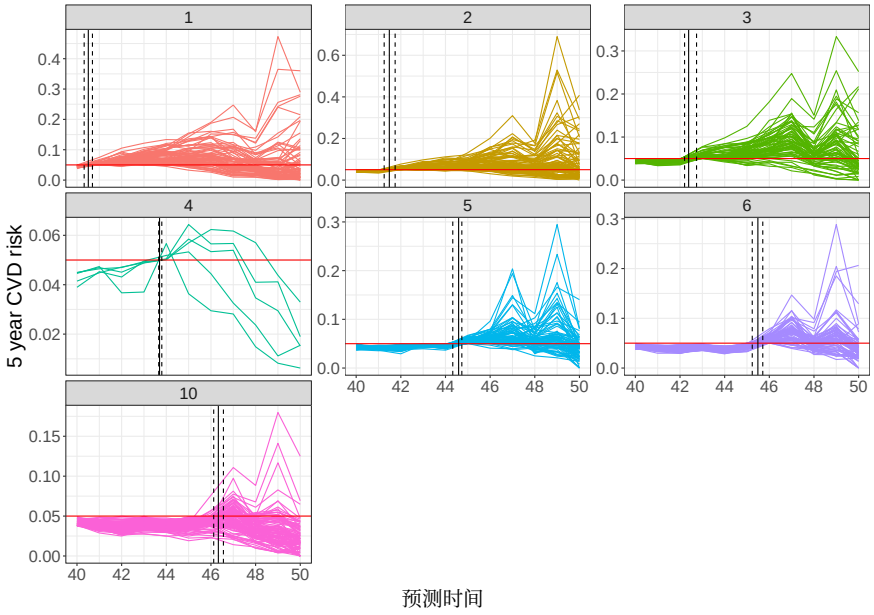

图3： $L_a = 40$时女性的风险曲线表示。每个面板代表一种最优风险评估频率。每条线表示特定个体的估计风险曲线。黑色实线表示超过5%阈值（以红色水平线表示）的预期时间中位数。虚线表示$t_{k_i^*,\ 40}$ 分布的第一和第三四分位数。

4 验证：C指数和布里尔分数

4.1 验证

我们验证了用于估计在启用他汀类药物前未被诊断为心血管疾病的概率的两阶段标志点模型，该模型在主稿件的第3.2节中进行了描述。估计的c指数以点的形式显示于图4中，而布里尔分数则显示于图5中。

在图4中，我们发现整体区分度良好（整体C指数为0.77，以黑色实线表示）。

其次，我们验证了主稿第 3.3 节所述的扩展的两阶段标志点方法，并在图 6 中报告了估计的 C 指数，在图 7 中报告了估计的布里尔分数。

请注意，用于估计 $t_{i,\ L_a}^*$ 的扩展的两阶段标志点方法通常具有良好的区分能力和预测准确性（布里尔分数越低，模型的预测准确性越高）。该模型对女性的表现优于男性（在每个时间点 s，女性的 C 指数均高于男性的 C 指数，而布里尔分数则相反）。此外，模型性能往往会随着标志年龄的增加而下降。在较高的标志年龄（75 岁、80 岁）下，可以观察到较低的 C 指数和较高的标准差，且预测时间越晚$s \in \{83, 84, 85\}$，情况越明显，分别为 $s \in \{88, 89, 90\}$。这可能是因为较高标志年龄下的平均随访时间较短，并且在标志年龄为 75 岁时，83 岁以后仍被观察到的人数极少；在标志年龄为 80 岁时，88 岁以后仍被观察到的人数也极少。

我们观察到，图4、图5、图6和图7呈现出相似的趋势：男性地标队列的区分度和预测准确性较低，并且随着地标年龄的增加，模型性能逐渐下降。

5 敏感性分析：探索NB参数的影响

我们通过改变以下关键参数，开展了三项敏感性分析：

- $\lambda \in [20,000;\ 30,000]$ 英镑/年，而 $u_s = 0.997$、 $c_s = 150$ 英镑/年，以及 $c_\nu = 18.39$ 英镑/次就诊。结果报告于图8的A面板中。
- $u_s \in [0.997;\ 1]$, 而 $\lambda = 25,\ 000$英镑/年， $c_s = 150$ 英镑/年，且 $c_\nu = 18.39$ 英镑/次就诊。结果见图8的B面板。

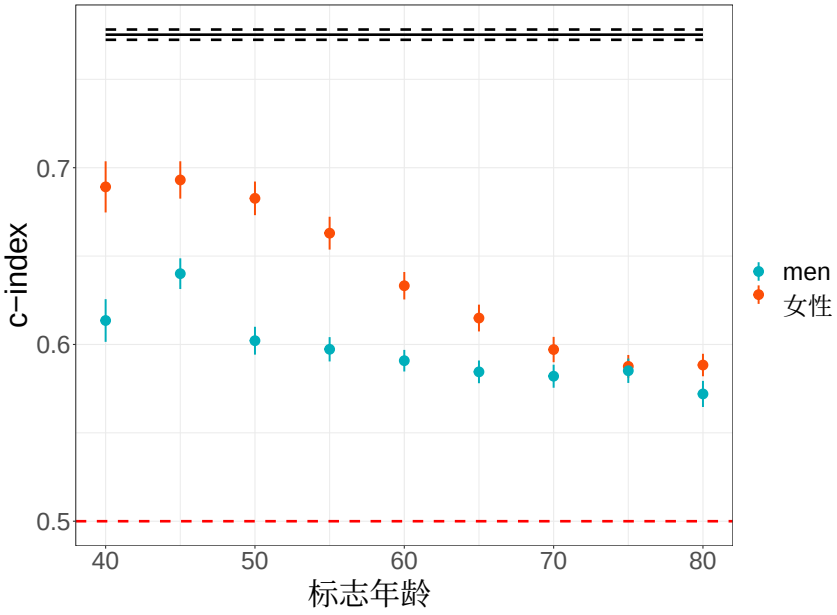

图4： 不同起始时间 t下，男性（蓝点）和女性（红点）第二个标志模型的估计 C 指数。每个点代表针对特定 $s = L_a \in \{40, 45, .., 80\}$ 和 $w = 10$计算出的 C 指数，因为我们关注的是10年心血管疾病风险的区分准确度。 $L_a = 40$处的点代表使用 $s = 40$ 和 $w = 10$估计的 C 指数。黑色实线表示不同标志年龄和性别的总体 C 指数（虚线表示95%置信区间）。位于0.5处的红色虚线表示 C 指数的最低合理值。

- $c_s \in [4;\ 320]$ 英镑/年，而 $\lambda = 25,\ 000$英镑/年，$u_s = 0.997$，且 $c_\nu = 18.39$ 英镑/次就诊。结果见图8的C面板。

总体而言，我们观察到结果对参数选择具有稳健性。

在图8的A面板中，我们将 λ 的值从20,000英镑/年调整至30,000英镑/年。随着 λ 增加，对于更少的人群而言，10年频率是最佳选择，而中间频率（如4–7年）则逐渐成为更大比例人群的最佳选择。每1年、2年或3年进行一次访视，对队列中固定比例的人群而言都是最佳选择。 λ 增加可以解读为人们更愿意付费以增加无心血管疾病生存年，因此，与每10年进行一次风险评估相比，人们往往更倾向于选择中间频率。

我们在 u_s 中（图8的B面板）观察到了类似的情况。我们将效用因子从0.997（这意味着生活质量下降0.003）调整到1（这意味着服用他汀类药物对生活质量完全没有影响）。我们注意到，对于固定的人数，1至4年一次的风险评估策略是最佳选择。当他汀类药物对生活质量的影响较小（u_s 趋近于1）时，每10年一次的方案较不受青睐，而中等频率（每4至8年一次）则更受青睐。如果服用他汀类药物没有任何负担（$u_s = 1$），那么最佳策略是立即开始使用他汀类药物。

相比之下，他汀类药物的价格越高（ c_s），越倾向于选择与较低就诊频率相关的风险评估策略（图8的C面板）。这是意料之中的，因为成本越高，他汀类药物使用的净获益就越低。

我们还考察了 c_ν 在15英镑/次就诊至1000英镑/次就诊之间变化的情况（未展示结果）。尽管考察的范围很广，最佳日程安排的比例在 c_ν的所有取值下均保持不变。预期会出现这一结果，因为在稿件公式（2）中，该净获益项与预期无事件生存年相关项的量级并不可比。

从图8纵轴所示的范围也可以立即看出，整个队列中占比最大的一部分（>70%）建议每10年评估一次。这是因为我们考虑的是堆叠地标队列，而规模最大的地标队列是在 $L_a = 40$、 $L_a = 45$收集的，这些队列由更年轻、健康状况更好的人组成。

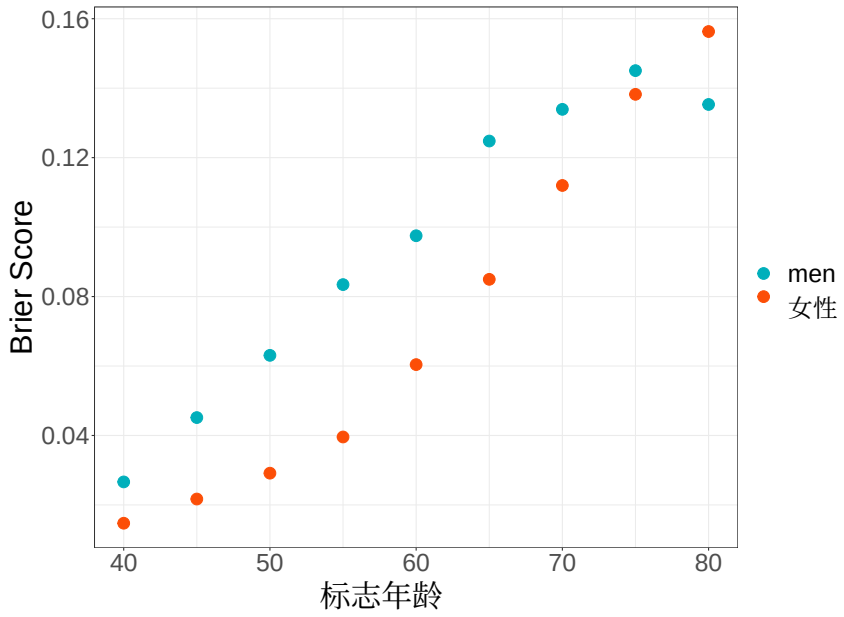

图 5： 布里尔分数的估计值，$BS_s(w)$，其中 $s = L_a \in \{40, 45, .., 80\}$ 和 $w = 10$。每个 $BS_{L_a}(10)$ 都以彩色圆点表示（蓝点代表男性，红点代表女性）。

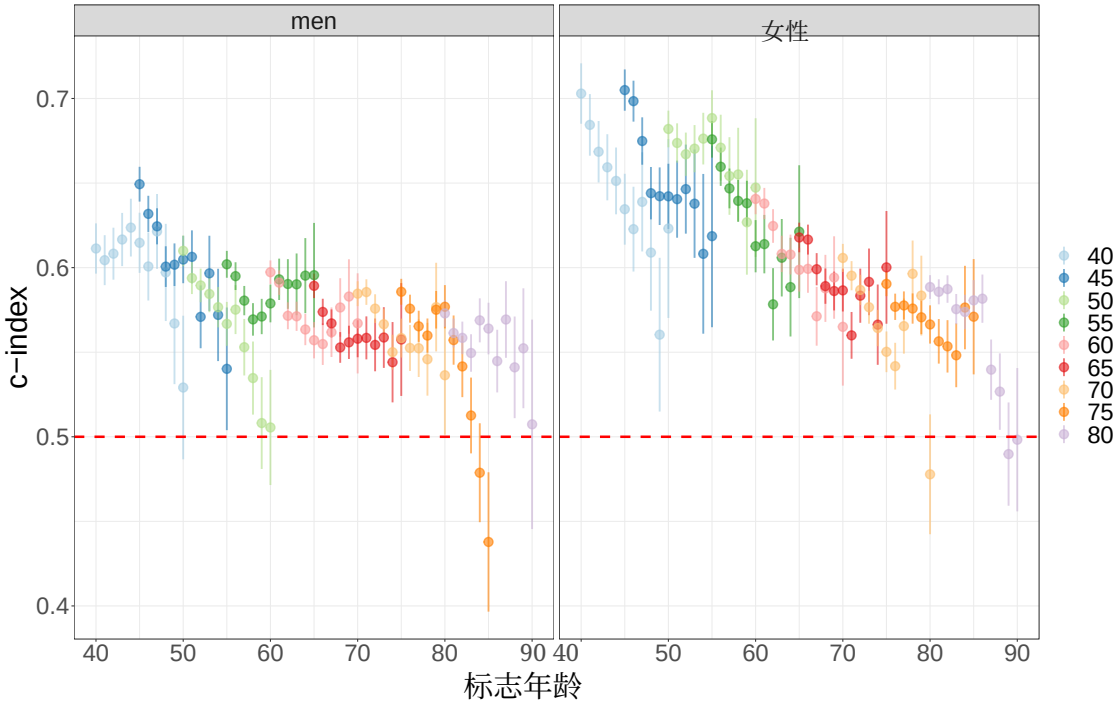

图 6： 不同起始时间 s下，女性（左图）和男性（右图）的首个标志模型的估计 C 指数。每个点代表针对特定 $s \in \mathcal{P}_{L_a}$ 和 $w = 5$计算得出的 C 指数，因为我们关注的是 5 年心血管疾病风险的区分准确度。我们为每个标志集合指定一种颜色。所有浅蓝色的点均对应于 $L_a = 40$，从最左侧的第一个点开始依次为$s \in \{40, 41, 42, .., 50\}$。0.5 处的红色虚线表示 C 指数的最低合理值。在年龄较大时，最新的时间窗口（即男性橙色点对应的 83-88、84-89、85-90，以及男性紫色点对应的 88-93、89-94、90-95）记录到低于 0.5 的数值。

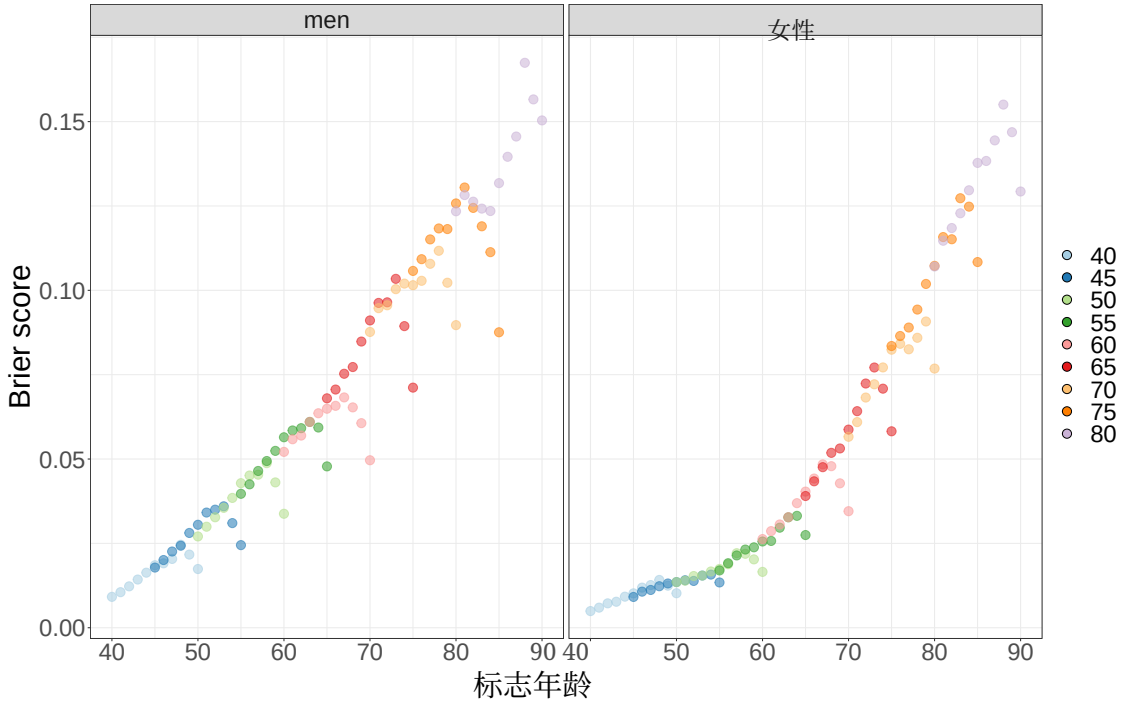

图7： 布里尔分数的估计值， $BS_s(w)$，其中 $s \in \mathcal{P}_{L_a}$ 和 $w = 5$。每个 $BS_s(5)$均以彩色圆点表示（每个标志年龄对应一种特定颜色）。女性对应的 BS 显示在左图中，而男性对应的 BS 显示在右图中。

参考文献

Bart S Ferket、Bob JH van Kempen、Jan Heeringa、Sandra Spronk、Kirsten E Fleischmann、Rogier LG Nijhuis、Albert Hofman、Ewout W Steyerberg 和 MG Myriam Hunink。无症状个体接受他汀类药物治疗的终身获益个体化预测：一项建模研究。*PLoSMed*，9(12):e1001361，2012年。英国国家医疗服务体系。NHS数字化QOF。https://digital.nhs.uk/data-and-information/publications/statistical/quality-and-outcomes-framework-achievement-prevalence-and-exceptions-data/quality-and-outcomes-framework-2010-11，2011年。访问日期：2022年6月20日。A Rosemary Tate、Sheena Dungey、Simon Glew、Natalia Beloff、Rachael Williams 和 Tim Williams。英国糖尿病记录质量：全科医生记录临床数据的编码方法如何影响发病率估计？一项使用CPRD数据库的横断面研究。*BMJ Open*，7(1)，2017年。流行病学研究组。降胆固醇治疗的疗效与安全性：对14项他汀类药物随机试验中90 056名参与者数据进行的前瞻性荟萃分析。《柳叶刀》，366(9493):1267–1278，2005年。Zhe Xu、Matthew Arnold、David Stevens、Stephen Kaptoge、Lisa Pennells、Michael J Sweeting、Jessica Barrett、Emanuele Di Angelantonio 和 Angela M Wood。将未来开始他汀类药物治疗纳入考虑的心血管疾病风险预测。《美国流行病学杂志》，2021年。

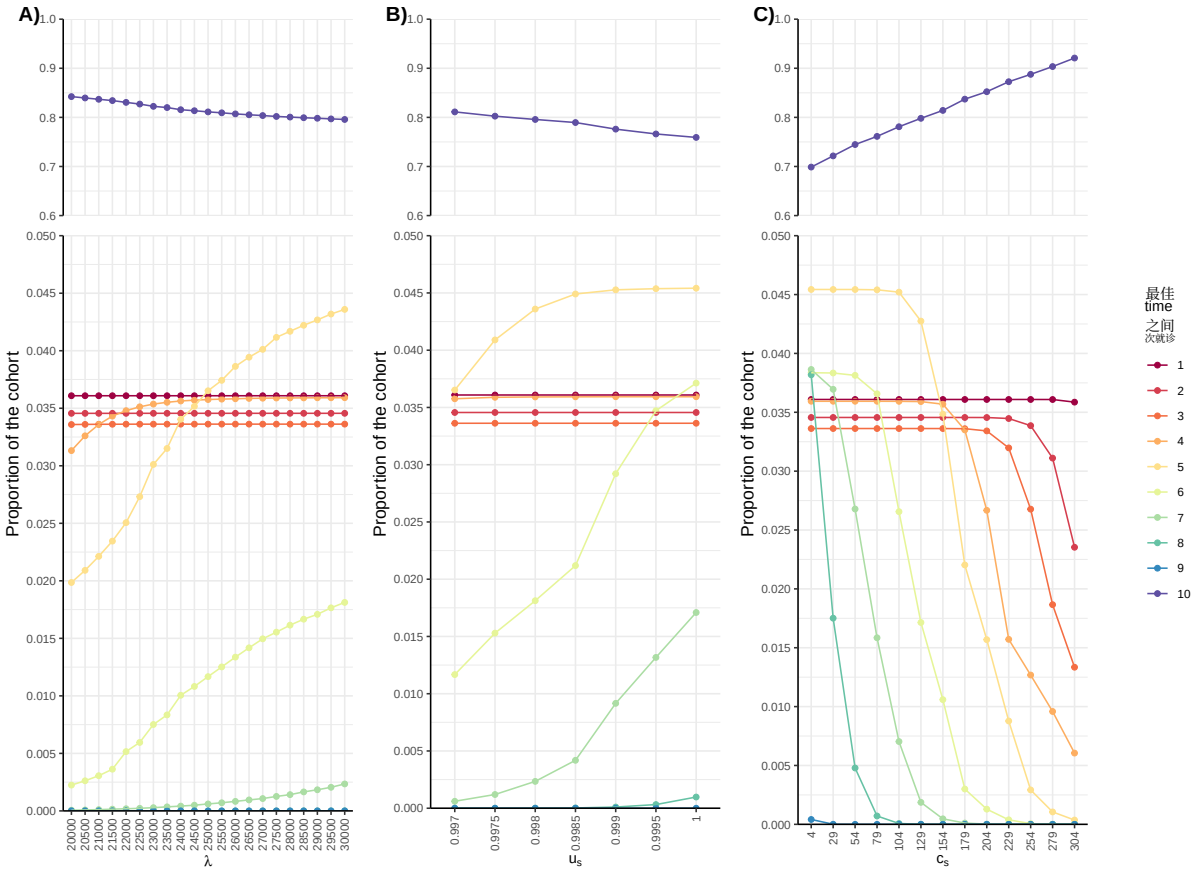

图8： λ（面板A）、c_s（面板B）、u_s（面板C）的敏感性分析。所有面板的y轴均表示队列的比例。面板A的x轴表示 λ 的取值，范围为 $[20,000;\ 30,000]$ 英镑/年。$u_s = 0.997$，$c_s = 150$ 英镑/年，以及 $c_\nu = 18.39$ 英镑/次就诊。面板B的x轴表示 u_s 的取值，范围为 $[0.997;\ 1]$。$\lambda = 25,000$英镑/年，$c_s = 150$ 英镑/年，以及 $c_\nu = 18.39$ 英镑/次就诊。面板C的x轴表示 c_s 的取值，范围为 $[4;\ 320]$ 英镑/年。$\lambda = 25,000$英镑/年，$u_s = 0.997$，以及 $c_\nu = 18.39$ 英镑/次就诊。本文的电子版中，此图以彩色显示。

表*3*：女性队列各标志年龄的描述性特征。

变量	风险类别	40	45	50	55	60	65	70	75	80
AF	极高风险	0 %	0.27 %	0.47 %	0.67 %	5.53 %	4.37 %	2.8 %	3.4 %	5.63 %
	高风险	0.41 %	0.14 %	0.56 %	0.67 %	0.65 %	0.14 %	0 %	0 %	
	中高风险	0.19 %	0.26 %	0.39 %	0.63 %	0.1 %	0 %	0 %	0 %	
	中低风险	0.16 %	0.23 %	0.36 %	0.39 %	0 %	0 %			
	低风险	0.12 %	0.15 %	0.2 %	0.29 %	0 %	0 %			
BP med	极高风险	95.82 %	91.6 %	92.58 %	83.88 %	80.16 %	77.39 %	57.62 %	52.06 %	62.15 %
	高风险	87.01 %	85.62 %	81.34 %	64.71 %	62.07 %	52.84 %	1.27 %	0.48 %	
	中高风险	78.43 %	72.31 %	65.12 %	53.63 %	39.6 %	16.35 %	0 %	0 %	
	中低风险	55.01 %	44.93 %	31.13 %	22.5 %	8.73 %	1.63 %			
	低风险	7.61 %	7.47 %	3.43 %	3.38 %	0.2 %	0 %			
Depression	极高风险	83.29 %	73.19 %	82.12 %	72.97 %	64.78 %	43.67 %	28.51 %	20.35 %	19.39 %
	高风险	75.05 %	64.02 %	71.53 %	64.7 %	47.21 %	29.05 %	4.3 %	2.17 %	
	中高风险	75.19 %	63.24 %	61.72 %	49.15 %	32.24 %	18.46 %	0.63 %	0 %	
	中低风险	66.92 %	52.9 %	44.86 %	28.02 %	11.96 %	7 %			
	低风险	26.21 %	25.03 %	14.42 %	7.79 %	0.97 %	0 %			
Diabetes	极高风险	22.84 %	21.18 %	11.88 %	13.12 %	6.22 %	4.31 %	2.07 %	1.96 %	2.11 %
	高风险	15.67 %	11.52 %	4.91 %	3.47 %	1.41 %	0.37 %	0 %	0 %	
	中高风险	6.92 %	5.09 %	2.28 %	1.06 %	0.43 %	0.09 %	0 %	0 %	
	中低风险	3.54%	1.91%	0.76%	0.11%	0.07%	0 %			
	低风险	0.31 %	0.15 %	0.06 %	0.01 %	0 %	0 %			
Migraine	极高风险	62.95 %	52.37 %	41.27 %	26.2 %	20.06 %	15.24 %	7.61 %	6.7 %	5.19 %
	高风险	52.58 %	46.28 %	33.12 %	21.57 %	16.63 %	11.44 %	8.68 %	0.97 %	
	中高风险	54.57 %	41.99 %	27.26 %	18.33 %	12.51 %	8.08 %	25.83 %	0 %	
	中低风险	42.54 %	30.5 %	16.82 %	12 %	6.93 %	3.47 %			
	低风险	9.99 %	8 %	6.11 %	4.97 %	2.8 %	0 %			
Renal dis.	极高风险	3.06 %	2.59 %	3.04 %	3.18 %	5.09 %	9.32 %	6.86 %	7.86 %	12.18 %
	高风险	1.86%	1.79%	2.68%	2.76%	3.4%	3.03%	0.97 %	0 %	
	中高风险	1.46 %	2.19 %	2.22 %	2.06 %	2.13 %	1.15 %	0.09 %	0 %	
	中低风险	1.2 %	1.24 %	1.29 %	1.3 %	1.18 %	0.32 %			
	低风险	0.36 %	0.5 %	0.67 %	0.93 %	1.26 %	0 %			
RA	极高风险	2.23 %	1.52%	1.78%	2.41%	8.52%	6.64%	3.53%	2.89 %	3.03 %
	高风险	1.03 %	1.14 %	1.62 %	2.08 %	3.3 %	1.94 %	0.03 %	0.48 %	
	中高风险	0.7 %	1.52 %	1.64 %	1.81 %	1.51 %	0.43 %	0 %	0 %	
	中低风险	0.89 %	1.01 %	1.27 %	1.44 %	0.2 %	0.03 %			
	低风险	0.66 %	0.86 %	1.05 %	1.42 %	0 %	0 %			
SMI	极高风险	9.75 %	6.43 %	9.82 %	8.07 %	5.19 %	4.52 %	1.99 %	1.61 %	1.73 %
	高风险	7.01 %	4.94 %	5.92 %	3.89 %	2.54 %	0.96 %	0.02 %	0 %	
	中高风险	8.38 %	4.25 %	3.93 %	2.43 %	1.31 %	0.22 %	0 %	0 %	
	中低风险	5.05 %	2.9 %	1.72 %	0.96 %	0.32 %	0.01 %			
	低风险	0.78 %	0.81 %	0.23 %	0.2 %	0.03 %	0 %			
SLE	极高风险	1.11 %	0.89 %	0.45 %	0.41 %	0.42 %	0.44 %	0.3 %	0.32 %	0.23 %
	高风险	0 %	0.36 %	0.35 %	0.46 %	0.38 %	0.32 %	0.2 %	0.24 %	
	中高风险	0.57 %	0.44 %	0.41 %	0.37 %	0.29 %	0.21 %	0.18 %	0 %	
	中低风险	0.38 %	0.37 %	0.34 %	0.27 %	0.23 %	0.27 %			
	低风险	0.19 %	0.22 %	0.23 %	0.25 %	0.4 %	0 %			
BMI	极高风险	1.22 (1.13)	0.89 (1.04)	0.33 (0.86)	0.08 (0.77)	0.01 (0.75)	-0.14 (0.63)	-0.27 (0.53)	-0.37 (0.49)	-0.46 (0.44)
	高风险	0.88 (1.03)	0.58 (0.94)	0.16 (0.78)	-0.04 (0.68)	-0.11 (0.6)	-0.24 (0.52)	-0.45 (0.4)	-0.78 (0.46)	
	中高风险	0.54 (0.94)	0.36 (0.86)	0.04 (0.72)	-0.1 (0.62)	-0.21 (0.51)	-0.32 (0.44)	-0.74 (0.4)	-1.18 (0.52)	
	中低风险	0.23 (0.83)	0.05 (0.71)	-0.14 (0.59)	-0.21 (0.52)	-0.34 (0.41)	-0.48 (0.42)			
	低风险	-0.33 (0.48)	-0.32 (0.47)	-0.34 (0.45)	-0.39 (0.44)	-0.66 (0.39)	-0.97 (0.3)			
HDL	极高风险	-1.08 (0.44)	-0.7 (0.57)	-0.46 (0.57)	-0.32 (0.57)	-0.03 (0.61)	0.01 (0.59)	0.2 (0.59)	0.29 (0.59)	0.3 (0.58)
	高风险	-0.86 (0.48)	-0.57 (0.54)	-0.28 (0.55)	-0.14 (0.52)	0.09 (0.56)	0.19 (0.54)	0.44 (0.56)	1.77 (0.68)	
	中高风险	-0.68 (0.47)	-0.43 (0.54)	-0.16 (0.53)	-0.01 (0.49)	0.19 (0.53)	0.3 (0.52)	1.12 (0.82)	2.52 (0.28)	
	中低风险	-0.48 (0.42)	-0.26 (0.46)	-0.02 (0.47)	0.17 (0.45)	0.33 (0.51)	0.57 (0.65)			
	低风险	-0.12 (0.33)	0.01 (0.4)	0.2 (0.46)	0.46 (0.57)	0.72 (0.71)	1.91 (0.76)			
SBP	极高风险	0.19 (0.67)	0.23 (0.64)	0.04 (0.6)	0.05 (0.55)	0.15 (0.51)	0.18 (0.46)	0.2 (0.4)	0.26 (0.39)	0.33 (0.42)
	高风险	0 (0.57)	0.01 (0.58)	-0.1 (0.54)	-0.07 (0.5)	0.01 (0.43)	0.06 (0.37)	0.02 (0.3)	-0.02 (0.36)	
	中高风险	-0.24 (0.55)	-0.17 (0.53)	-0.22 (0.51)	-0.15 (0.46)	-0.07 (0.37)	-0.02 (0.33)	-0.35 (0.32)	-0.32 (0.24)	
	中低风险	-0.47 (0.47)	-0.4 (0.45)	-0.37 (0.42)	-0.26 (0.38)	-0.2 (0.34)	-0.2 (0.34)			
	低风险	-0.85 (0.33)	-0.69 (0.34)	-0.54 (0.35)	-0.44 (0.36)	-0.58 (0.33)	-0.78 (0.23)			
Smoke	极高风险	0.71 (0.24)	0.71 (0.24)	0.68 (0.26)	0.72 (0.24)	0.51 (0.3)	0.4 (0.29)	0.26 (0.23)	0.2 (0.19)	0.18 (0.19)
	高风险	0.66 (0.25)	0.68 (0.24)	0.6 (0.26)	0.61 (0.25)	0.42 (0.26)	0.29 (0.22)	0.17 (0.1)	0.08 (0.08)	
	中高风险	0.64 (0.24)	0.65 (0.24)	0.56 (0.26)	0.51 (0.25)	0.35 (0.21)	0.24（0.14）	0.07（0.08）	0.07（0.08）	
	中低风险	0.61（0.24）	0.6（0.24）	0.5 (0.24)	0.38 (0.18)	0.28 (0.15)	0.18 (0.12)			
	低风险	0.46 (0.2)	0.43 (0.2)	0.38 (0.18)	0.27 (0.16)	0.14 (0.13)	0.04 (0.06)			
TCHOL	极高风险	0.07 (0.56)	0.34 (0.67)	0.3 (0.6)	0.42 (0.6)	0.44 (0.53)	0.43 (0.51)	0.39 (0.46)	0.36 (0.46)	0.3 (0.48)
	高风险	0.01 (0.56)	0.15 (0.57)	0.22（0.51）	0.38（0.5）	0.43（0.44）	0.44（0.43）	0.44（0.39）	0.31（0.55）	
	中高风险	-0.1 (0.48)	0.09 (0.5)	0.16 (0.44)	0.33 (0.42)	0.42 (0.4)	0.43 (0.38)	0.43 (0.48)	0.12 (0.38)	
	中低风险	-0.22 (0.4)	-0.05 (0.38)	0.09 (0.38)	0.28 (0.36)	0.39 (0.35)	0.39 (0.39)			
	低风险	-0.38 (0.24)	-0.21 (0.27)	0.01 (0.31)	0.21 (0.35)	0.24 (0.42)	0.04 (0.57)			
Townsend20	极高风险	15.64 (3.66)	15.01 (4.15)	14.75 (4.16)	13.64 (4.73)	12.97 (4.8)	11.68 (5.05)	9.44 (5.19)	8.69 (5.19)	9.01 (5.2)
	高风险	15.45 (3.74)	14.35 (4.46)	13.28 (4.82)	12.71 (4.96)	11.32 (5.1)	9.22 (5.17)	5.58 (3.79)	4.28 (2.97)	
	中高风险	14.97 (4.2)	13.91 (4.69)	12.4 (5.15)	11.29 (5.31)	9.59 (5.13)	7.61 (4.41)	3.65 (2.64)	2 (1.28)	
	中低风险	14.09 (4.6)	12.93 (5.09)	11.05 (5.4)	9.18 (5.06)	6.32 (4.12)	4.24 (3.1)			
	低风险	9.02 (5.52)	8.31 (5.31)	7.04 (4.75)	5.38 (3.97)	3.1 (2.24)	2.24 (1.48)			

表*4*：男性队列各标志年龄的描述性特征。

变量	风险类别	40	45	50	55	60	65	70	75	80
AF	极高风险	1.18 %	0.78 %	1.01 %	1.4 %	2.42 %	2.35 %	3.42 %	5.45 %	8.27 %
	高风险	0.61 %	1.09 %	0.96 %	0.81 %	0.01 %	0 %	0 %		
	中高风险	1.04 %	0.73 %	0.54 %	0.63 %	0.01 %	0 %			
	中低风险	0.45 %	0.33 %	0.41 %	0.69 %	0 %	0 %			
	低风险	0.2 %	0.29 %	0.32 %	0.67 %	0 %				
BP med	极高风险	84.23 %	80.23 %	59.76 %	50.48 %	36.44 %	31.47 %	36.19 %	44.84 %	55.08 %
	高风险	78.51 %	56.12 %	30.35 %	11.88 %	5.5 %	2.37 %	0 %		
	中高风险	61.79 %	29.84 %	9.59 %	3.13 %	3.85 %	1.43 %			
	中低风险	18.26 %	4.46 %	2.22 %	2.17 %	2.53 %	0 %			
	低风险	1.13 %	0.84 %	1.08 %	0 %	0 %				
Depression	极高风险	55.98 %	44.94 %	39.69 %	31.7 %	21.12 %	14.4 %	11.64 %	10.35 %	9.81 %
	高风险	56.29 %	39.13 %	27 %	14.88 %	4.52%	2.94%	3.85%		
	中高风险	47.88%	33.09%	17.43 %	4.83 %	3.08 %	1.67 %			
	中低风险	42.5 %	15.17 %	6.74 %	3.71 %	1.86 %	0 %			
	低风险	7.2 %	5.39 %	3.83 %	2.68 %	0 %				
Diabetes	极高风险	25.23 %	4.34 %	6.81 %	4.17 %	2.92 %	2.45 %	2.94 %	3.7 %	3.82 %
	高风险	8.7 %	3.44 %	1.68 %	0.78 %	0.16 %	0.28 %	0 %		
	中高风险	5.64 %	1.79 %	0.61 %	0.29 %	0.09 %	1.19 %			
	中低风险	1.1 %	0.59 %	0.16 %	0.17 %	0.51 %	0 %			
	低风险	0.08 %	0.46 %	0.18 %	1.34 %	0 %				
Migraine	极高风险	8.67 %	17.41 %	10.62 %	7.6 %	4.75 %	3.45 %	3.16 %	2.63 %	2.27 %
	高风险	10.95 %	14.85 %	6.32 %	4.15 %	2.22 %	3.41 %	0 %		
	中高风险	9.15 %	10.21 %	4.54 %	1.57 %	1.93 %	2.38 %			
	中低风险	6.18 %	3.66 %	1.86 %	1.01 %	2.03 %	0 %			
	低风险	3.84 %	1.06 %	1.3 %	0 %	0 %				
Renal dis	极高风险	3.42 %	1.91 %	1.89 %	1.76 %	2.19 %	2.26 %	3.66 %	6.29 %	10.26 %
	高风险	1.74 %	1.95 %	1.38 %	0.86 %	0.27 %	2.72 %	0 %		
	中高风险	1.47 %	1.04 %	0.64 %	0.62 %	0.2 %	3.1 %			
	中低风险	0.7%	0.42%	0.5%	0.98%	0.34%	50 %			
	低风险	0.25 %	0.38 %	0.83 %	0 %	0 %				
RA	极高风险	0.66 %	0.69 %	0.77 %	0.79 %	1.3 %	1.09 %	1.24 %	1.38 %	1.44 %
	高风险	0.51 %	0.55 %	0.66 %	0.59 %	0.03 %	0.09 %	0 %		
	中高风险	0.55 %	0.55 %	0.5 %	0.52 %	0.04 %	0 %			
	中低风险	0.39 %	0.41 %	0.47 %	0.91 %	0 %	0 %			
	低风险	0.31 %	0.42 %	0.7 %	0.67 %	0 %				
SMI	极高风险	6.44 %	5.96 %	3.73 %	2.3 %	1.92 %	1.22 %	1.2 %	1.14 %	1.11 %
	高风险	4.09 %	5.44 %	2.17 %	1.15 %	0.14 %	0.14 %	0 %		
	中高风险	3.78 %	3.4 %	1.33 %	0.6 %	0.08 %	0.24 %			
	中低风险	2.74 %	1.16 %	0.63 %	0.81 %	0.17 %	0 %			
	低风险	1.17 %	0.49 %	0.76 %	0.67 %	0 %				
SLE	极高风险	0 %	0.1 %	0.09 %	0.09 %	0.07 %	0.08 %	0.09 %	0.09 %	0.09 %
	高风险	0.2 %	0.02 %	0.1%	0.05%	0.04%	0.14%	0 %		
	中高风险	0 %	0.09%	0.05 %	0.06 %	0.07 %	0.24 %			
	中低风险	0.05 %	0.05 %	0.04 %	0.08 %	0 %	0 %			
	低风险	0.04 %	0.02 %	0.07 %	0 %	0 %				
BMI	极高风险	0.78 (1.09)	0.45 (0.94)	0.09 (0.73)	-0.1 (0.62)	-0.22 (0.53)	-0.31 (0.49)	-0.4 (0.49)	-0.46 (0.47)	-0.56 (0.43)
	高风险	0.38 (0.91)	0.16 (0.7)	-0.14 (0.56)	-0.22 (0.44)	-0.32 (0.37)	-0.62 (0.44)	-1.15 (0.5)		
	中高风险	0.21 (0.77)	-0.05 (0.56)	-0.18 (0.44)	-0.26 (0.39)	-0.5 (0.45)	-0.89 (0.43)			
	中低风险	-0.11 (0.53)	-0.19 (0.37)	-0.25 (0.37)	-0.49 (0.46)	-0.82 (0.43)	-1.02 (0.78)			
	低风险	-0.28 (0.32)	-0.32 (0.38)	-0.52 (0.45)	-0.94 (0.43)	-1.47 (0.18)				
HDL	极高风险	-0.7 (0.51)	-0.41 (0.52)	-0.23 (0.51)	-0.07 (0.5)	0.08 (0.5)	0.2 (0.56)	0.28 (0.63)	0.28 (0.64)	0.3 (0.62)
	高风险	-0.51 (0.46)	-0.28 (0.48)	-0.07 (0.43)	0.08 (0.41)	0.25 (0.46)	0.79 (0.82)	1.91 (1.3)		
	中高风险	-0.39 (0.44)	-0.18 (0.41)	-0.01 (0.37)	0.17 (0.43)	0.63 (0.76)	1.78 (1.09)			
	中低风险	-0.22 (0.32)	-0.07 (0.3)	0.1 (0.41)	0.68 (0.81)	1.73 (1.01)	2.52 (0.98)			
	低风险	-0.07 (0.31)	0.11 (0.48)	0.66 (0.83)	2.36 (0.93)	4.04 (0.08)				
SBP	极高风险	0.17 (0.65)	0.13 (0.59)	-0.08 (0.54)	-0.08 (0.5)	-0.04 (0.44)	-0.03 (0.41)	-0.01 (0.44)	0.03 (0.45)	0.05 (0.46)
	高风险	-0.11 (0.54)	-0.15 (0.47)	-0.28 (0.41)	-0.21 (0.34)	-0.17 (0.32)	-0.41 (0.37)	-0.66 (0.42)		
	中高风险	-0.23 (0.5)	-0.33 (0.38)	-0.33 (0.33)	-0.29 (0.32)	-0.39 (0.38)	-0.59 (0.42)			
	中低风险	-0.46 (0.34)	-0.44 (0.27)	-0.41 (0.31)	-0.49 (0.37)	-0.57 (0.42)	-0.56 (0.37)			
	低风险	-0.58 (0.25)	-0.58 (0.29)	-0.62 (0.35)	-0.66 (0.37)	-0.58 (0.59)				
Smoke	极高风险	0.7 (0.22)	0.68 (0.23)	0.65 (0.25)	0.54 (0.28)	0.4 (0.26)	0.29 (0.25)	0.22 (0.23)	0.17 (0.22)	0.14 (0.21)
	高风险	0.68 (0.22)	0.64 (0.23)	0.61 (0.25)	0.42 (0.2)	0.28 (0.15)	0.15 (0.15)	0.06 (0.08)		
	中高风险	0.65 (0.21)	0.65 (0.23)	0.5 (0.19)	0.34 (0.17)	0.17 (0.16)	0.11 (0.12)			
	中低风险	0.66 (0.19)	0.54 (0.17)	0.39 (0.17)	0.2 (0.19)	0.14 (0.15)	0.04 (0.02)			
	低风险	0.52 (0.18)	0.38 (0.2)	0.18 (0.18)	0.15 (0.15)	0.13 (0.15)				
TCHOL	极高风险	0.68 (0.75)	0.59 (0.59)	0.52 (0.52)	0.45 (0.48)	0.39 (0.42)	0.3 (0.41)	0.19 (0.43)	0.08 (0.44)	0.02 (0.45)
	高风险	0.5 (0.56)	0.5 (0.48)	0.44 (0.42)	0.41 (0.36)	0.36 (0.32)	0.16 (0.46)	-0.45 (0.64)		
	中高风险	0.47 (0.51)	0.43 (0.4)	0.42 (0.33)	0.37 (0.33)	0.23 (0.44)	0.02 (0.57)			
	中低风险	0.33 (0.33)	0.37 (0.28)	0.36 (0.32)	0.21 (0.46)	0.03 (0.53)	-1.1 (0.51)			
	低风险	0.24 (0.24)	0.27 (0.33)	0.17 (0.45)	0.03 (0.61)	0.09 (0.25)				
Townsend 20	极高风险	13.97 (4.83)	12.85 (5.17)	12.81 (5.16)	11.26 (5.35)	10.34 (5.25)	8.7 (5.22)	8.39 (5.17)	8.53 (5.2)	8.73 (5.21)
	高风险	13.63 (4.75)	12.2 (5.38)	11.61 (5.34)	10.22 (5.31)	5.96 (4.02)	4.8 (3.7)	3.23 (2.41)		
	中高风险	12.9 (5.18)	12.2 (5.36)	10.92 (5.25)	6.42 (4.41)	4.98 (3.86)	4.4 (3.47)			
	中低风险	12.9 (5.09)	10.29 (5.16)	6.48 (4.44)	5.35 (4.11)	4.55 (3.61)	6.5 (0.71)			
	低风险	8.65 (5.41)	5.98 (4.81)	4.96 (4)	4.72 (4.23)	5 (1.73)				