

高血压房颤患者的乳酸白蛋白比值与 28 天死亡率：一项回顾性队列研究

2025年9月19日 18:31

Background

乳酸白蛋白比值（LAR）已成为反映代谢压力和营养状况的综合生物标志物。本研究旨在评估乳酸与白蛋白比值与心房颤动（房颤）高血压患者 28 天死亡率之间的关系。

Methods

我们利用 MIMIC-IV v3.1 数据库开展了一项回顾性队列研究。根据预先设定的标准对患者进行筛选，最终确定了由 1087 名符合条件的患者组成的队列。入住重症监护室后 28 天内的死亡率是主要终点。统计分析包括 LASSO 回归和多变量 Cox 回归、接收器操作特征 (ROC) 曲线和 Kaplan-Meier 生存曲线分析

Results

28 天的总死亡率为 22.8% (n = 248)。与幸存者相比，非幸存者的 LAR 值明显更高 (0.74 对 0.52, $P < 0.001$)。多变量分析表明，LAR 是 28 天死亡率的独立预测因子 (HR 1.03, 95% CI 1.01-1.06, $p < 0.05$)，即使在调整了多种临床混杂因素后也是如此。ROC 分析证实，与其他生物标志物相比，LAR 的预测能力更强 (AUC 0.661)。KM 生存分析显示，高 LAR 组和低 LAR 组的死亡率存在显著差异 (HR 2.55, 95% CI 1.97-3.30, $p < 0.05$)

Conclusions

LAR 是高血压房颤患者短期死亡率的独立预测指标。

Introduction

- 1.虽然乳酸、白蛋白及其比率已在普通重症监护病房人群中进行过研究，但之前还没有研究专门评估过房颤重症高血压患者的 LAR
- 2.本研究旨在评估高血压房颤患者 LAR 与 28 天死亡率之间的关系
- 3.我们假设 LAR 的升高与短期死亡率的增加有独立的关联，并将提供超越传统临床和实验室指标的增量预后价值

Methods

Data sources

这项研究利用了来自 MIMIC-IV（3.1 版）的匿名患者数据

Study population

通过国际疾病分类代码从 MIMIC-IV 数据库中识别出被诊断为房颤的患者：ICD-9 代码 427.31 和 ICD-10 代码 I48、I480、I481、I4811、I4819、I482、I4820、I4821、I489 和 I4891。这些 ICD 编码包括阵发性、持续性、永久性和未指定的房颤亚型，可确保广泛捕捉 ICU 入院时的房颤表现。最初共检索到 78,887 名房颤患者，其中 27,345 人至少入住过一次 ICU。首先根据患者入住 ICU 时心房颤动的编码来确定患者。这种方法确保了将在重症监护室住院期间患有房颤的患者纳入其中；但是，MIMIC-IV 数据库并没有显示房颤是入院时的主要诊断还是入院前的合并症，因此我们的结果应在此背景下进行解释。随后根据预先确定的纳入和排除标准进行了筛选。以下患者被排除在外：(1) 首次入院时年龄小于 18 岁的患者；(2) 多次入住 ICU 的患者，为避免重复偏倚，仅保留首次入住 ICU 的患者；(3) 入住 ICU 时间小于 24 小时的患者；(4) 诊断为终末期肾病、肝硬化或恶性肿瘤的患者；(5) 未分类为高血压的患者；(6) 入住 ICU 后 24 小时内未测量血清乳酸或白蛋白的患者。应用这些标准后，最终纳入了 1087 名患者进行分析（图 1）。最常见的排除原因是血清乳酸和/或血清白蛋白水平均未在 ICU 入院的头 24 小时内获得，因此无法计算 LAR。在 MIMIC-IV 中，这种遗漏往往发生在病情较轻或诊断方法不同的患者身上。因此，我们的最终分析样本可能会过度代表病情更严重、实验室检查更全面的患者。

Data extraction

人口统计学变量包括年龄、性别和种族。生命体征数据包括心率、收缩压和舒张压、平均动脉压和呼吸频率。实验室数值也包括在内。评估的具体合并症包括心血管疾病和慢性疾病。此外，还收集了有关治疗干预和临床管理的信息，包括机械通气、持续肾脏替代疗法（CRRT）、内镜逆行胆管造影术（ERCP）的使用情况，以及血管加压药、 β -受体阻滞剂、他汀类药物、华法林、阿司匹林和非维生素 K 口服抗凝剂（NOACs）的使用情况。所有数据均通过 PostgreSQL（版本 17）结合 Navicat Premium（版本 16.3.2）执行的结构化查询语言（SQL）查询提取，用于数据库管理和界面导航。

Grouping and endpoint events

根据 28 天的院内生存结果，将患者分为两组：幸存者组（n = 839）和非幸存者组（n = 248）

Management of missing data

- 1.为尽量减少数据不完整可能造成的偏差，分析中排除了缺失值超过 15% 的变量
- 2.对于缺失率在 5-15% 之间的变量，则采用多重插补，通过最合适的估算数据集来估计和替换缺失条目

Statistical analysis

- 1.连续变量以均数 \pm 标准差（means \pm SDs）或中位数（四分位距）[medians (IQRs)]表示，分类变量以计数（百分比）[%]表示
- 2.基线比较中，对于连续变量，根据其分布情况采用独立样本 t 检验（independent t tests）或单因素方差分析（one-way ANOVA）；对于分类变量，则采用皮尔逊卡方检验（Pearson's chi-square test）或费雪精确检验（Fisher's exact test）

3.变量选择

第一阶段：变量初筛

(1)单因素Cox回归:

分别检验每一个变量与生存结局（这里是28天内是否死亡）之间的关系。只有在单因素分析中有统计学意义（通常 $p < 0.1$ 或 $p < 0.05$ ）的变量，才被认为是潜在的预测因子，有资格进入下一步。

(2)LASSO回归:

这是一种更现代的机器学习方法。它可以在处理大量变量时，通过施加一个“惩罚项”，自动将那些贡献不大的变量的系数压缩至零，从而实现变量筛选和防止模型过拟合。同时使用两种方法可以使筛选出的变量集更加稳健。

第二阶段：确定独立危险因素

(3)多因素Cox回归:

将第一阶段筛选出的所有关键变量同时放入一个回归模型中。这么做的目的是，在控制了其他所有变量的影响之后，看某一个变量是否仍然与死亡风险显著相关。如果相关，那么它就是一个“独立危险因素”。例如，我们要看LAR是否是独立危险因素，就要在模型里同时校正年龄、性别、SOFA评分等，如果校正后LAR的p值依然小于0.05，我们才能说它独立于这些因素之外，对死亡风险有预测价值。

4. 校正模型

建立不同层次的校正模型是为了检验结论的稳健性 (Robustness)

模型1: 这是一个基础模型，只校正了最常见、最基本的人口学混杂因素 (年龄和性别)

模型2: 这是一个完全校正模型，纳入了大量临床上已知或潜在的影响预后的因素。作者把他们能想到的重要混杂因素几乎都放进去了

5. 验证LAR的预测能力

通过受试者工作特征 (ROC) 曲线来评估乳酸、白蛋白、LAR和SOFA评分对死亡率预测的区分能力，并使用约登指数 (Youden index) 来确定LAR的最佳阈值

6. 补充和验证分析

(1) 使用Kaplan-Meier生存曲线来比较高LAR组和低LAR组的生存情况，并用对数秩检验 (log-rank test) 来量化两组间的差异

(2) 使用限制性立方样条 (RCS) 拟合LAR与死亡风险之间潜在的非线性关系，并辅以阈值效应分析来探寻拐点

(3) 在临床分层的队列中进行亚组分析，以验证研究结果

Results

Baseline characteristics

一堆的基线差异信息被显示出来

这些基线显示出来有显著性差异的变量被纳入单因素cox回归进行进一步分析

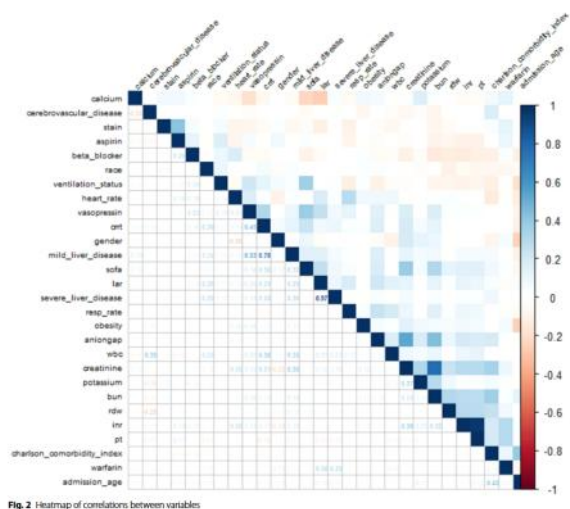
The LAR is an independent risk factor

1. 对基线特征中显示出显著统计学差异 ($P < 0.05$) 的变量进行了单变量 Cox 回归分析——分析结果显示，未经调整的 LAR 与 28 天死亡率之间存在密切联系 (HR 1.79, 95% CI 1.58-2.03, $p < 0.001$)

2. 为减少潜在的过拟合，对单变量 Cox 模型中确定为重要的变量采用了 LASSO 回归

3. 在进行 LASSO 分析之前，通过皮尔逊相关矩阵和多重共线性诊断对共线性进行了评估，大多数协变量显示出可接受的共线性水平，方差膨胀因子 (VIF) 值小于 5，但也有少数协变量超过了这一阈值

4. 经过 LASSO 回归，保留了 17 个系数不为零的变量进行进一步分析



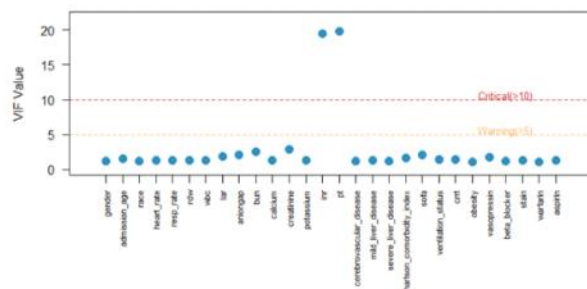


Fig. 3 VIF plot for multicollinearity checking. VIF: variance inflation factor

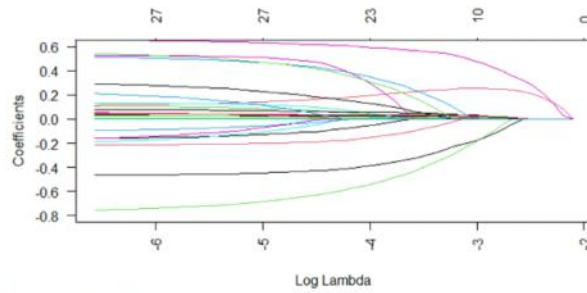


Fig. 4 Characterization of the changes in the variable coefficients

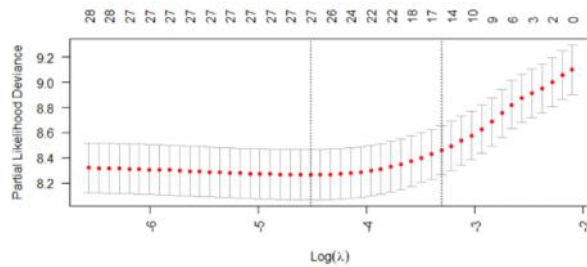


Fig. 5 Cross-validation selection process for optimal values of parameter λ in LASSO regression models

5.单变量分析和 LASSO 分析中 P 均小于 0.05 的变量随后被纳入多变量 Cox 回归模型

表 3 总结了多变量分析的结果。在调整了潜在的混杂因素（包括年龄、性别、心率、呼吸频率、白细胞计数、LAR、阴离子间隙、BUN、血钾、PT、脑血管疾病、CCI、SOFA 评分、机械通气、血管加压素的使用以及 β 受体阻滞剂、他汀类药物和华法林的使用）后，LAR 仍然是一个独立的风险因素（模型 II: HR 1.03, 95% CI 1.01-1.06, $P < 0.05$ ）

Table 3 Multivariate Cox analysis of risk factors

Variables	p	HR (95%CI)	p	HR (95%CI)
Model I			Model II	
Admission age	<0.001	1.03 (1.02–1.05)	<0.001	1.03 (1.02–1.05)
Gender				
Female		1.00 (Reference)		1.00 (Reference)
Male	0.129	0.81 (0.62–1.06)	0.129	0.81 (0.62–1.06)
Heart rate	0.027	1.01 (1.01–1.02)	0.027	1.01 (1.01–1.02)
Resp rate	0.067	1.03 (1.00–1.07)	0.067	1.03 (1.00–1.07)
WBC	0.075	1.17 (0.98–1.39)	0.075	1.17 (0.98–1.39)
LAR	0.002	1.03 (1.01–1.06)	0.002	1.03 (1.01–1.06)
Aniongap	0.343	1.02 (0.98–1.07)	0.343	1.02 (0.98–1.07)
BUN	0.472	1.00 (1.00–1.01)	0.472	1.00 (1.00–1.01)
Potassium	0.019	1.31 (1.05–1.63)	0.019	1.31 (1.05–1.63)
PT	0.007	1.01 (1.01–1.02)	0.007	1.01 (1.01–1.02)
Charlson Comorbidity Index	0.020	1.09 (1.01–1.16)	0.020	1.09 (1.01–1.16)
SOFA	0.062	1.04 (1.00–1.09)	0.062	1.04 (1.00–1.09)
Cerebrovascular disease				
No		1.00 (Reference)		1.00 (Reference)
Yes	<0.001	1.74 (1.27–2.37)	<0.001	1.74 (1.27–2.37)
Ventilation status				
No		1.00 (Reference)		1.00 (Reference)
Yes	<0.001	1.85 (1.33–2.58)	<0.001	1.85 (1.33–2.58)
Vasopressin				
No		1.00 (Reference)		1.00 (Reference)
Yes	<0.001	1.93 (1.38–2.71)	<0.001	1.93 (1.38–2.71)
Beta blocker				
No		1.00 (Reference)		1.00 (Reference)
Yes	<0.001	0.58 (0.45–0.76)	<0.001	0.58 (0.45–0.76)
Statin				
No		1.00 (Reference)		1.00 (Reference)
Yes	0.051	0.75 (0.57–1.00)	0.051	0.75 (0.57–1.00)
Warfarin				
No		1.00 (Reference)		1.00 (Reference)
Yes	<0.001	0.46 (0.32–0.65)	<0.001	0.46 (0.32–0.65)

Model I was adjusted for age and gender

Model II was adjusted for Model I plus heart rate, respiratory rate, white blood cell count, lactate-to-albumin ratio, anion gap, blood urea nitrogen, potassium, prothrombin time, cerebrovascular disease, Charlson comorbidity index, Sequential Organ Failure Assessment score, ventilation status, vasopressin, beta blocker, statin, and warfarin

6.RCS 的 Cox 比例危险模型

为了进一步描述 LAR 与死亡风险之间的关系，采用了包含 RCS 的 Cox 比例危险模型（图 6）。该模型显示出明显的线性关系（总体 $p < 0.001$ ），并有非线性证据（ $p = 0.003$ ）。样条曲线显示了危害随剂量的增加而增加，当 LAR 临界值超过约 0.56 时，可观察到有临床意义的风险加速。这些发现强调了 LAR 作为一种连续性生物标志物的预后价值，在这种生物标志物中，高于生理标准（ $LAR > 0.56$ ）的增量与死亡风险的增加独立相关，而不依赖于预定义的临界值

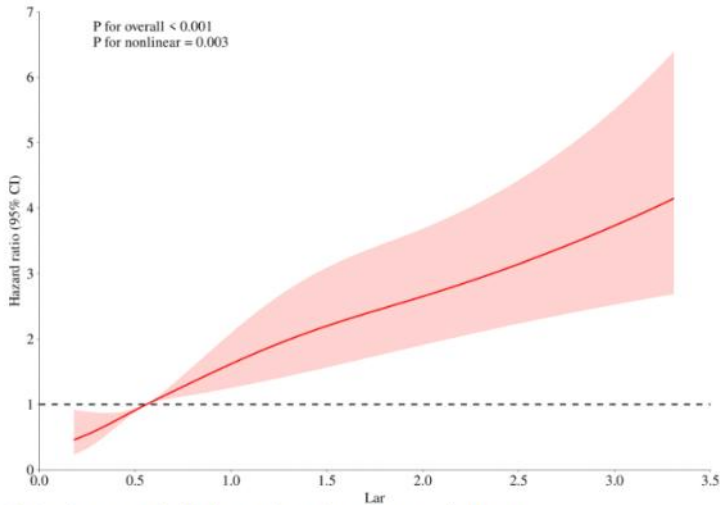


Fig. 6 Restricted cubic spline curve analysis for the LAR and mortality. LAR: lactate-to-albumin ratio

ROC curve analysis and Kaplan–Meier curves

单个指标的ROC曲线: 评估的是这个指标自身的、独立的预测能力。它回答的问题是：“如果我只看LAR这一个数值，我能多准确地猜出病人会不会在28天内死亡？”
模型的ROC曲线: 评估的是由多个变量组合而成的整个模型的综合预测能力。它回答的问题是：“当我综合考虑了LAR、年龄、SOFA评分等多个因素后，我的整体预测能力有多强？”

作者在这里构建ROC曲线，并没有使用一个特定的回归模型（如Cox或逻辑回归），而是直接使用了每个生物标记物本身的原始数值来对患者的最终结局（28天内死亡/存活）进行预测

对各种生物标记物预测房颤高血压患者 28 天死亡率的判别性能进行了 ROC 曲线分析（图 7）。在所评估的标记物中，LAR 的曲线下面积（AUC）最高（0.661；95% CI 0.622-0.700）

在最佳尤登指数临界值（ $LAR = 0.605$ ）时，灵敏度和特异度分别为 63.3% 和 62.3%，表明在检测高危个体和排除低危个体方面表现均衡
相比之下，单纯乳酸盐（临界值 = 1.75 mmol/L）的 AUC 稍低，灵敏度（63.7%）相似，但特异性（57.3%）降低。白蛋白的临界值为 3.35 g/dL，灵敏度最高（74.2%），但特

异性最低 (43.4%)，而 SOFA 评分 (临界值 = 7.5) 的特异性 (70.1%) 高于灵敏度 (50.8%)。这些发现突出表明，相对于单个成分或临床评分，LAR 的整体判别能力更强

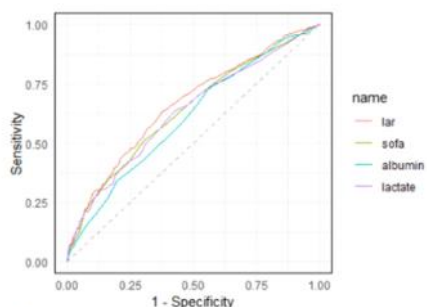


Fig. 7 ROC curves. ROC: receiver operating characteristic; LAR: lactate-to-albumin ratio

在临床相关阈值 0.605 时对患者进行二分，进一步证实了 LAR 的分层效用。卡普兰-梅耶生存分析显示，高 LAR 组和低 LAR 组的 28 天死亡率存在显著差异 (HR 2.55, 95% CI 1.97-3.30)

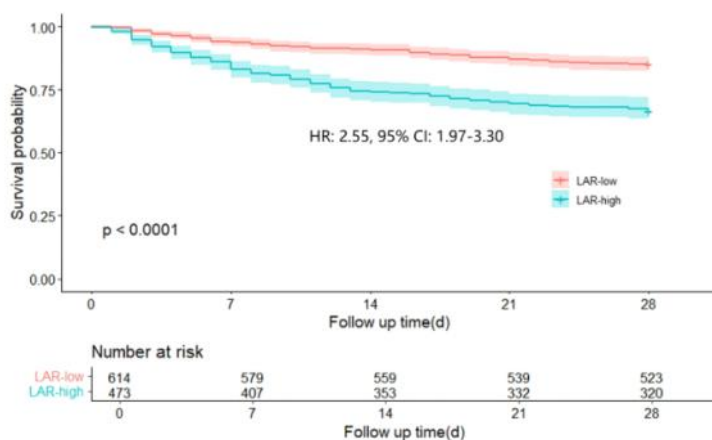
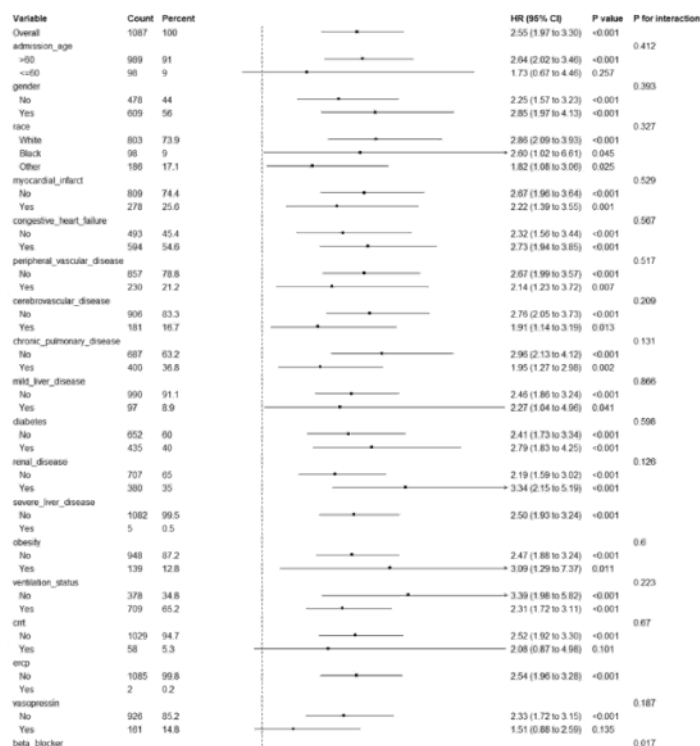


Fig. 8 Kaplan-Meier survival analysis curves. LAR: lactate-to-albumin ratio

亚组分析

以LAR分层，并进行亚组分析

亚组分析 当对不同亚组进行分层时，LAR 与亚组之间没有显著的交互作用



第一步：理解这张图的构成

这张图叫做森林图 (Forest Plot)，是展示亚组分析结果的标准方式。

1. 左侧 (Variable, Count, Percent):

- Variable: 这是用来划分亚组的临床特征。例如，第一个亚组是按 admission_age (入院年龄) 划分的，分成了 >60 岁和 ≤60 岁两组。
- Count 和 Percent: 显示了每个亚组里有多少病人，以及他们占总人数的百分比。

2. 中间 (HR, 95% CI, P value):

- HR (Hazard Ratio, 风险比): 这是核心指标。它表示**高LAR组**相较于**低LAR组**的死亡风险倍数。
 - Overall (第一行)**: HR是**2.55**。这说明, 在整个研究人群中, 高LAR组的死亡风险是低LAR组的2.55倍。这是研究的主要结论。
 - 各个亚组**: 下面的每一行都显示了在**该特定亚组内**计算出的HR。例如, 在>60岁的患者中, HR是**2.64**; 而在renal_disease (肾病) Yes (有)的患者中, HR高达**3.34**。
- 黑点和横线**: 黑点代表HR的点估计值, 横线代表95%置信区间(CI)。置信区间越窄, 说明结果越精确。
- P value**: 这个P值检验的是**在当前亚组内, HR是否显著大于1**。可以看到, 几乎所有亚组的P值都小于0.05 (除了极少数, 如vasopressin Yes组), 说明在绝大多数亚组中, 高LAR都是一个显著的危险因素。

3. 右侧 (P for interaction):

- 交互作用P值 (P for interaction)**: 这是回答你问题的**关键**。它检验的是:**LAR的预测效果 (即HR的大小) 在不同的亚组之间是否存在显著差异?**
- 例如, 对于gender (性别) 这个变量, 男性的HR是2.25, 女性的HR是2.85。交互作用P值 (0.393) 回答的问题是: “2.25和2.85这两个数字之间的差异, 是真实存在的 (即LAR对女性的预测性真的比男性强), 还是仅仅是由于抽样误差造成的随机波动?”

第二步: 解读 “没有显著的交互作用”

研究结果显示, 所有的P for interaction都**大于0.05**。这意味着什么?

我们以gender (性别) 为例, 它的P for interaction是0.393。

- 因为 **0.393 > 0.05**, 所以我们得出结论: 虽然从数值上看, 女性的HR (2.85) 比男性的HR (2.25) 要高, 但这种差异在统计学上**不显著**。我们不能认为LAR对女性的风险预测能力比对男性更强。正确的解读是: **LAR对男性和女性的死亡风险预测能力是相似的, 没有本质区别。**

我们可以将这个逻辑应用到图中所有的亚组划分上:

- 年龄**: LAR的预测效果在大于60岁和小于等于60岁的人群中没有显著差异 (P for interaction = 0.412)。
- 种族**: LAR的预测效果在白人、黑人和其他种族中没有显著差异 (P for interaction = 0.327)。
- 合并症**: 无论患者是否合并有心肌梗死、心力衰竭、肾病或糖尿病等, LAR的预测能力都是稳定且显著的 (所有 P for interaction > 0.05)。

特别是看beta_blocker (β 受体阻滞剂) 这一组, 它的P for interaction是0.017, 是所有交互P值中唯一一个小于0.05的。这提示LAR的预测效果在**使用和不使用** β 受体阻滞剂的患者中**可能存在差异**。具体来看, 未使用者的HR (3.61) 远高于使用者 (2.35)。这可能是一个值得未来进一步研究的发现点, 但在本文中, 作者可能基于多重检验校正等原因, 认为整体上不存在显著的交互作用。不过, 这确实是图中最值得注意的一个例外。*(通常论文会对此进行讨论)*

结论与意义

所以, 当作者说 “**LAR 与亚组之间没有显著的交互作用**” 时, 他是在用严谨的统计学语言告诉你一个非常重要的临床结论:

LAR是一个非常 “**皮实**” (Robust) 和普适的生物标志物。它的价值不会因为病人的个体差异而大打折扣。无论你是年老的还是年轻的, 男性还是女性, 无论你同时患有什么其他慢性病, 只要你的LAR升高了, 就都意味着你的短期死亡风险显著增加。