

晚期肝纤维化序贯筛查算法

性能分析报告

自动生成分析

2026 年 1 月 22 日

1 引言

本报告呈现了两种两步序贯筛查算法的性能比较分析：

- **M4→M3**: Clinical-A（临床参数）→ Fusion-Net（最优配置）
- **M5→M3**: Clinical-Net（基础临床深度学习）→ Fusion-Net

两种方法均与 M3 单模型（Fusion-Net）进行比较。

2 诊断准确性指标

表 1: 表 1：内部测试集和前瞻性测试集中不同方法评估肝纤维化的诊断准确性指标

场景和纤维化检测	阈值	准确率	敏感性	特异性	PPV	NPV
内部测试集 (n = 51)						
Fusion-Net (M3)	0.09	84 (43/51)	91 (20/22)	79 (23/29)	77 (20/26)	92 (23/25)
Two-step#1 (M4→M3)	0.09; 0.10	84 (43/51)	86 (19/22)	83 (24/29)	79 (19/24)	89 (24/27)
Two-step#2 (M5→M3)	0.42; 0.10	84 (43/51)	77 (17/22)	90 (26/29)	85 (17/20)	84 (26/31)
前瞻性测试集 (n = 100)						
Fusion-Net (M3)	0.05	85 (85/100)	95 (37/39)	79 (48/61)	74 (37/50)	96 (48/50)
Two-step#1 (M4→M3)	0.19; 0.04	86 (86/100)	92 (36/39)	82 (50/61)	77 (36/47)	94 (50/53)
Two-step#2 (M5→M3)	0.40; 0.04	85 (85/100)	87 (34/39)	84 (51/61)	77 (34/44)	91 (51/56)

注：除非另有说明，数据为百分比，括号内为患者数量。M3 (Fusion-Net) 是完整多模态融合模型，M4 (Clinical-A) 是临床参数模型，M5 (Clinical-Net) 是基础临床深度学习模型。**加粗**表示最优两步配置。

3 转诊影响分析

表 2: 表 2：内部测试集和前瞻性测试集中序贯算法的转诊影响

	Step 1	Step 1	避免转诊	
算法	阴性排除	阳性进入	比例 (%)	Δ NPV
内部测试集 ($n = 51$)				
M4→M3	8	43	16%	−3.1%
M5→M3	20	31	39%	−8.1%
前瞻性测试集 ($n = 100$)				
M4→M3	25	75	25%	−1.7%
M5→M3	35	65	35%	−4.9%

注： Δ NPV 表示与 M3 单模型相比的 NPV 变化。M4→M3 在保持更高 NPV 的同时实现有效转诊节省。

4 序贯筛查流程

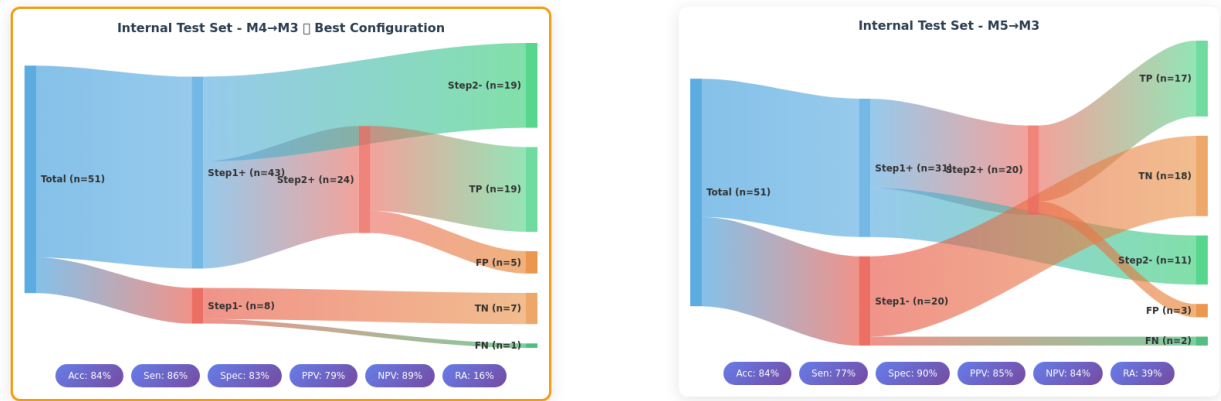


图 1: 内部测试集的序贯筛查算法流程图。左：M4→M3（排除 8 人，16%）；右：M5→M3（排除 20 人，39%）

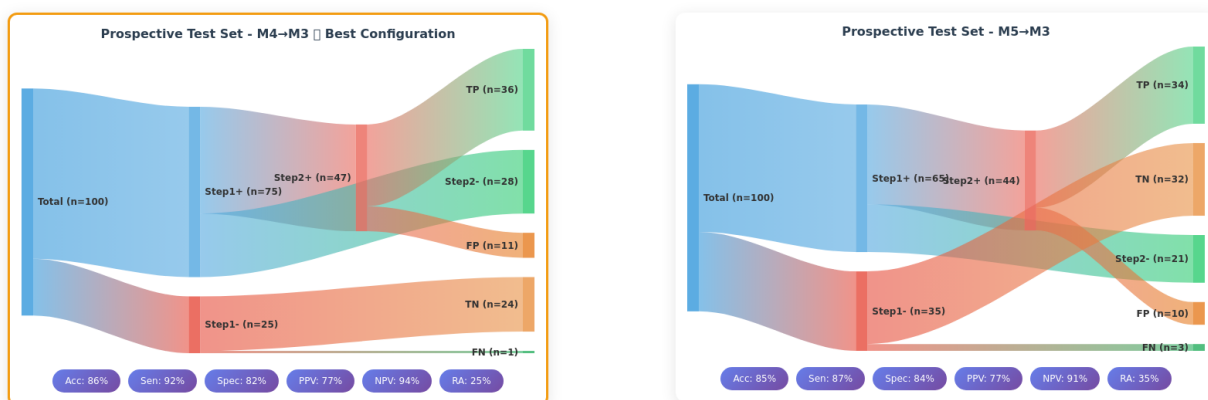


图 2: 前瞻性测试集的序贯筛查算法流程图。左: M4→M3 (排除 25 人, 25%); 右: M5→M3 (排除 35 人, 35%)

5 预测值与患病率关系

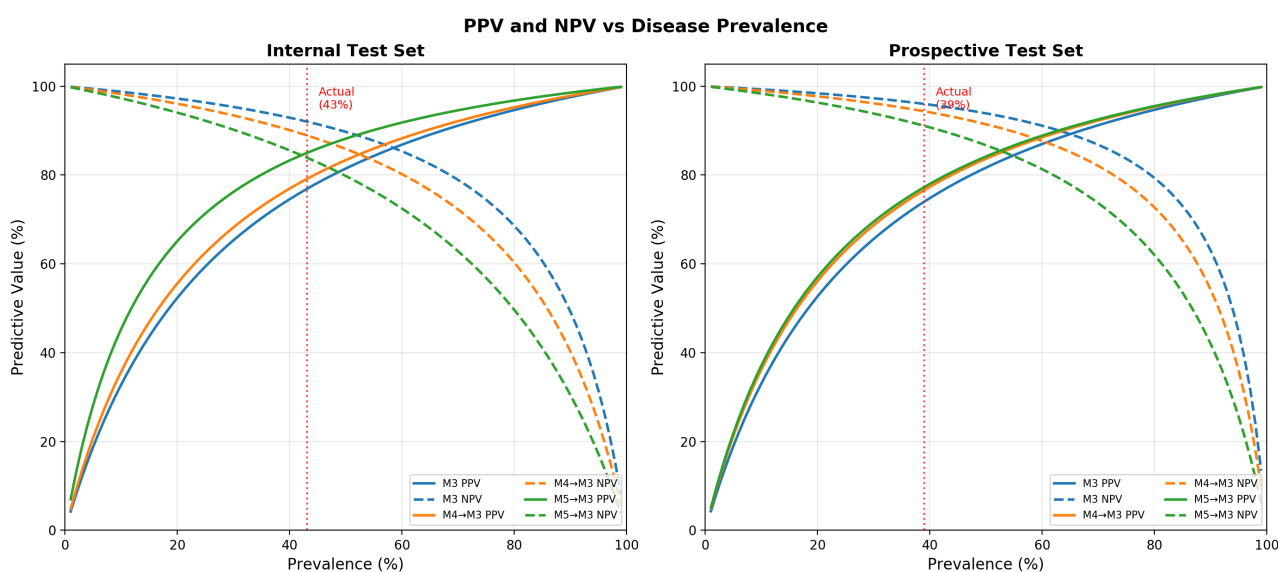


图 3: 三种方法 (M3、M4→M3、M5→M3) 的 PPV 和 NPV 随患病率变化曲线。实线为 PPV, 虚线为 NPV。

6 主要发现

6.1 M4→M3 为最优配置

内部测试集:

- 准确率: 84% (与 M3 持平)
- NPV: 89% (仅下降 3%)
- 避免转诊: 16%

前瞻性测试集：

- 准确率：86%（优于 M3 的 85%）↑
- NPV：94%（仅下降 2%）
- 避免转诊：25%

6.2 M5→M3 作为备选

M5→M3 虽然避免转诊率更高（内部 39%，前瞻 35%），但 NPV 下降较多（5-8%），可能增加漏诊风险。

7 结论

1. **M4→M3 是推荐的最优配置：**在保持高诊断准确性的同时，实现有效的资源节省。
2. **NPV 保持良好：**M4→M3 的 NPV 下降仅 2-3%，临床可接受。
3. **权衡考虑：**若资源极度有限，可考虑 M5→M3 以获得更高的转诊节省率，但需接受更高的漏诊风险。