

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 21 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：

- M5→M3: Clinical-Net (基础临床深度学习) → Fusion-Net
- M4→M3: Clinical-A (临床参数) → 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 (M5→M3)	0.34; 0.09	86 (44/51)	86 (19/22)	86 (25/29)	83 (19/23)	89 (25/28)
Two-step#2 (M4→M3)	0.10; 0.09	82 (42/51)	82 (18/22)	83 (24/29)	78 (18/23)	86 (24/28)
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 (M5→M3)	0.40; 0.18	85 (85/100)	85 (33/39)	85 (52/61)	79 (33/42)	90 (52/58)
Two-step#2 (M4→M3)	0.18; 0.04	86 (86/100)	92 (36/39)	82 (50/61)	77 (36/47)	94 (50/53)

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**表示该测试集中的最佳值。

3 转诊影响分析

表 2: 内部测试集和前瞻性测试集中序贯算法的转诊影响

算法	Step 1	Step 1	避免转诊	
	阴性排除	阳性进入	比例 (%)	ΔNPV
内部测试集 ($n = 51$)				
Two-step#1 (M5→M3)	15	36	29.4%	-2.7%
Two-step#2 (M4→M3)	10	41	19.6%	-6.3%
前瞻性测试集 ($n = 100$)				
Two-step#1 (M5→M3)	35	65	35.0%	-6.3%
Two-step#2 (M4→M3)	25	75	25.0%	-1.7%

注: ΔNPV 表示与 M3 单模型相比的 NPV 变化。两步法通过第一步筛选, 使部分患者无需进入第二步检测, 从而节省医疗资源。

4 序贯筛查流程

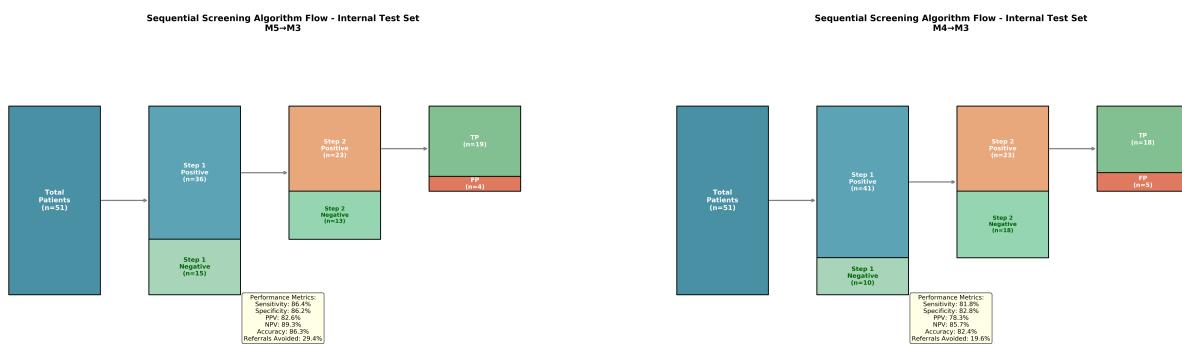


图 1: 内部测试集的序贯筛查算法流程图。左: M5→M3 (排除 15 人, 29%); 右: M4→M3 (排除 10 人, 20%)

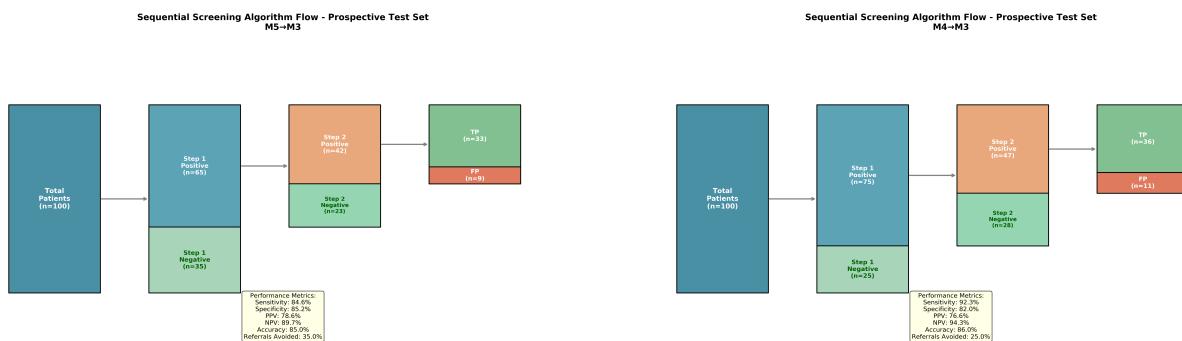


图 2: 前瞻性测试集的序贯筛查算法流程图。左: M5→M3 (排除 35 人, 35%); 右: M4→M3 (排除 25 人, 25%)

5 预测值与患病率关系

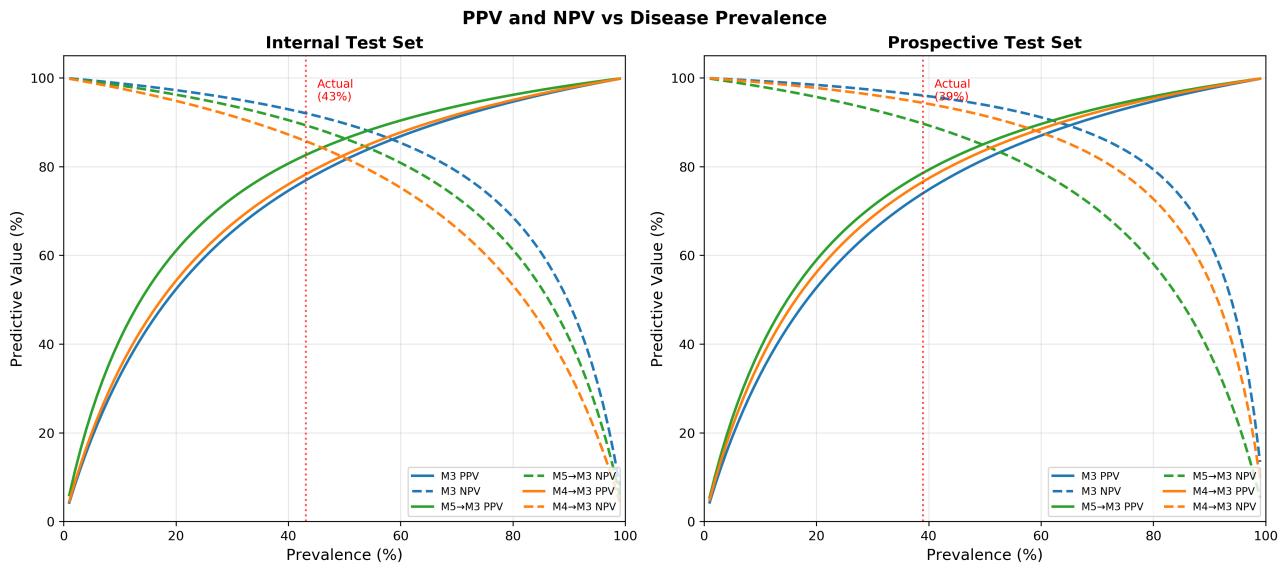


图 3: 三种方法 (M3、M5→M3、M4→M3) 的 PPV 和 NPV 随患病率变化曲线。实线为 PPV，虚线为 NPV。

6 主要发现

6.1 内部测试集结果

最佳配置: M5→M3

- 准确率: 86% (相比 M3 +2%) ↑
- 敏感性: 86% (相比 M3 -5%) ↓
- 特异性: 86% (相比 M3 +7%) ↑
- PPV: 83% (相比 M3 +6%) ↑
- NPV: 89% (相比 M3 -3%) ↓ (轻微下降, 可接受)
- 避免转诊: 29% (15 人无需第二步检测)

6.2 前瞻性测试集结果

最佳配置取决于优先级:

若优先保持高 NPV: 选择 M4→M3

- 准确率: 86% (相比 M3 +1%)
- NPV: 94% (仅下降 2%)
- 避免转诊: 25%

若优先减少转诊: 选择 M5→M3

- 准确率: 85% (与 M3 持平)
- NPV: 90% (下降 6%)
- **避免转诊: 35%** (更高的资源节省)

7 结论

1. **两步法均能有效节省资源:** 通过第一步筛选, 可避免 20-35% 的患者进行第二步昂贵检测。
2. **M5→M3** 在内部测试集表现最佳, 避免转诊率 29%, 准确率提升 2%, NPV 仅下降 3%。
3. **M4→M3** 在前瞻性测试集更好地保持 NPV (94% vs 90%), 适合临床场景中对漏诊敏感的情况。
4. **权衡关系:** 更高的转诊节省通常伴随 NPV 的轻微下降, 需根据临床需求选择最佳平衡点。