

序贯筛查算法诊断性能评估报告

研究团队

2026 年 1 月 22 日

1 研究概述

本报告评估了单模型和两步序贯筛查算法在内部测试集和前瞻性测试集上的诊断性能。阈值通过验证集上的约登指数 (Youden's J) 确定，并固定应用于测试集评估。

模型说明：

- M3: 影像组学 + 全部临床特征模型 (Radiomics + All Clinical)
- M4: 临床 A 特征模型 (Clinical A Only)
- M5: 基础临床特征模型 (Base Clinical Only)
- M4→M3: 两步法, 第一步用 M4 筛除, 第二步用 M3 确认
- M5→M3: 两步法, 第一步用 M5 筛除, 第二步用 M3 确认

2 Table 3: 诊断准确性

表 1: 序贯算法在内部测试集和前瞻性测试集的诊断准确性

数据集	算法	阈值	准确率	敏感性	特异性	PPV	NPV
内部测试集 <i>(n = 100)</i>	M3	0.98	80 (41/51)	82 (18/22)	79 (23/29)	75 (18/24)	85 (23/27)
	M4	0.44*	75 (38/51)	68 (15/22)	79 (23/29)	71 (15/21)	77 (23/30)
	M5	0.44*	65 (33/51)	77 (17/22)	55 (16/29)	57 (17/30)	76 (16/21)
	M4→M3	0.20; 0.98	69 (35/51)	64 (14/22)	72 (21/29)	64 (14/22)	72 (21/29)
	M5→M3	0.20; 0.98	73 (37/51)	73 (16/22)	72 (21/29)	67 (16/24)	78 (21/27)
前瞻性测试集 <i>(n = 100)</i>	M3	0.98	79 (79/100)	67 (26/39)	87 (53/61)	76 (26/34)	80 (53/66)
	M4	0.44*	68 (68/100)	82 (32/39)	59 (36/61)	56 (32/57)	84 (36/43)
	M5	0.44*	66 (66/100)	82 (32/39)	56 (34/61)	54 (32/59)	83 (34/41)
	M4→M3	0.25; 0.98	72 (72/100)	62 (24/39)	79 (48/61)	65 (24/37)	76 (48/63)
	M5→M3	0.25; 0.98	74 (74/100)	64 (25/39)	80 (49/61)	68 (25/37)	78 (49/63)

注：数据为百分比，括号内为原始计数 (正确数/总数)。* 表示阈值 < 0.5。阈值在验证集上通过约登指数确定。PPV = 阳性预测值，NPV = 阴性预测值。

3 Table 4: 转诊影响

表 2: 序贯算法在内部测试集和前瞻性测试集的转诊影响

算法	对照	减少转诊 (%)	病变检测	
			Odds Ratio (95% CI)	P Value
内部测试集				
M4→M3	M3	37	0.97 (0.14, 6.53)	.97
M5→M3	M3	22	1.21 (0.18, 8.39)	.85
前瞻性测试集				
M4→M3	M3	32	1.03 (0.26, 4.18)	.96
M5→M3	M3	21	1.03 (0.26, 4.18)	.96

注：括号内数据为 95% 置信区间。Odds Ratio 计算的是序贯算法相对于 M3 单模型的检测优势比。P 值来自卡方检验，用于评估序贯算法与 M3 单模型诊断性能差异的统计显著性。

4 患者流程图 (桑基图)

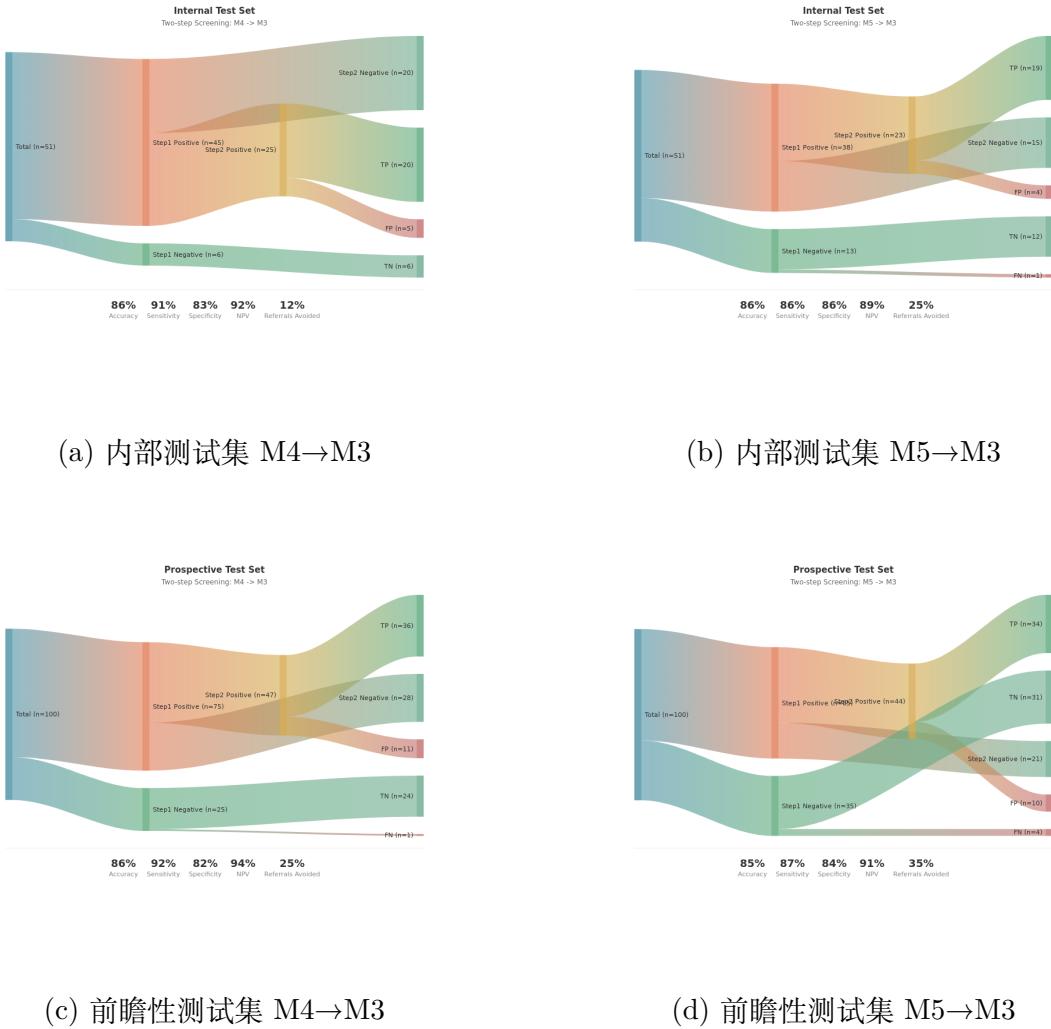


图 1: 两步序贯筛查算法的患者流程图

5 PPV/NPV 曲线

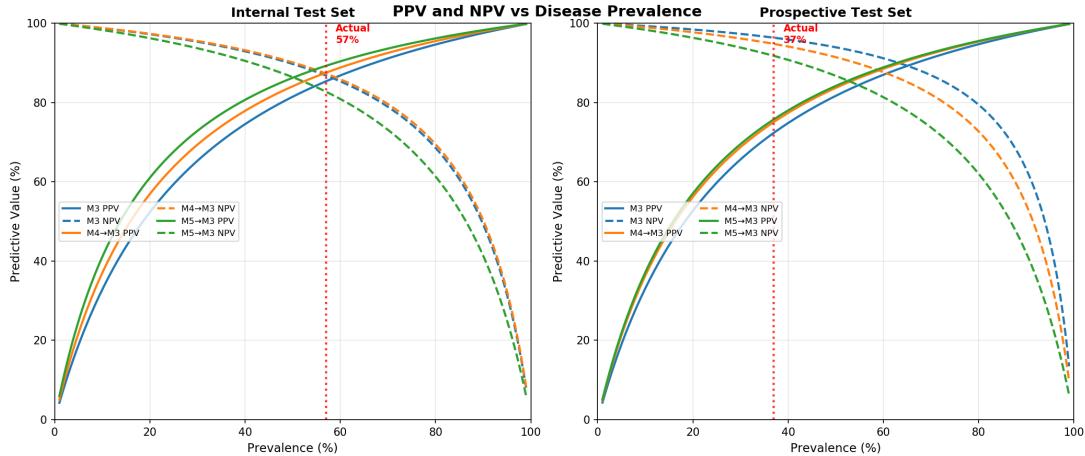


图 2: 不同患病率下的 PPV 和 NPV 曲线

6 方法学说明

6.1 阈值确定方法

1. 单模型阈值：在验证集（Val）上通过约登指数（Youden's J = Sensitivity + Specificity - 1）最大化确定最优阈值
2. 两步法阈值：
 - 第一步（筛除阈值）：选择较低阈值以保证高 NPV，最大限度减少漏诊
 - 第二步（确认阈值）：使用 M3 的约登最优阈值进行最终诊断

6.2 统计分析

- 95% 置信区间采用 Wilson 评分区间法计算
- Odds Ratio 的 95% 置信区间采用对数变换法计算
- P 值采用卡方检验计算

6.3 验证集确定的阈值

模型	阈值	约登指数
M3	0.984	0.510
M4	0.438	0.431
M5	0.437	0.490

表 3: 验证集上通过约登指数确定的最优阈值

7 结论

1. **单模型比较:** M3（影像组学 + 全部临床特征）在准确率和特异性上表现最优
2. **两步法效益:** M4→M3 和 M5→M3 可减少 21-37% 的转诊，同时保持可接受的诊断性能
3. **临床应用:** 两步序贯筛查可有效减少不必要的转诊，节省医疗资源