

AI辅助结肠镜对不同年资医师检出效应研究

王若楠^{1,2}, 赵贵君²

¹内蒙古科技大学包头医学院研究生院, 内蒙古 包头

²内蒙古自治区人民医院消化中心, 内蒙古 呼和浩特

收稿日期: 2026年4月19日; 录用日期: 2026年5月12日; 发布日期: 2026年5月21日

摘要

目的: 探讨人工智能(AI)辅助结肠镜对不同年资医师检出效应的影响。方法: 采用单中心回顾性观察研究, 纳入2021年9月至2022年9月内蒙古自治区人民医院住院结肠镜患者1093例。按是否使用AI分为传统组和AI组; 按AI引入前医师累计结肠镜操作总数是否超过10,000例分为高、低年资组; 以息肉检出率(PDR)和腺瘤检出率(ADR)为主要结局, 进行总体比较、年资分层比较、AI应用初期与后期比较, 并采用多因素logistic回归及双重差分模型进行分析。结果: 总体比较显示, AI组PDR高于传统组(61.6% vs 54.5%, $P = 0.023$), ADR差异无统计学意义(43.6% vs 39.1%, $P = 0.160$)。分层比较显示, 低年资层内AI组PDR高于传统组(64.6% vs 56.2%, $P = 0.041$), 高年资层内PDR和ADR差异均无统计学意义。分层多因素logistic回归显示, 低年资层内AI与PDR升高独立相关($aOR = 1.63$, 95% CI: 1.07~2.48, $P = 0.024$), 而其余结局相关性无统计学意义。AI应用初期与后期比较、亚组内前后多因素logistic回归及双重差分模型均未见稳定统计学差异; 但在低年资医师后期同期比较中, AI组PDR高于传统组(67.0% vs 54.8%, $P = 0.038$)。结论: AI辅助结肠镜可提高息肉检出率, 且在低年资医师中更易体现检出优势; 其对腺瘤检出率的促进作用尚需进一步验证。

关键词

人工智能, 结肠镜, 息肉检出率, 腺瘤检出率, 医师年资

Effect of Artificial Intelligence-Assisted Colonoscopy on Detection Outcomes among Endoscopists with Different Experience Levels

Ruonan Wang^{1,2}, Guijun Zhao²

¹Graduate School, Baotou Medical College, Inner Mongolia University of Science and Technology, Baotou Inner Mongolia

²Digestive Endoscopy Center, Inner Mongolia People's Hospital, Hohhot Inner Mongolia

文章引用: 王若楠, 赵贵君. AI辅助结肠镜对不同年资医师检出效应研究[J]. 临床医学进展, 2026, 16(5): 1706-1714.
DOI: 10.12677/acm.2026.1651973

Abstract

Objective: To investigate the effect of artificial intelligence (AI)-assisted colonoscopy on detection outcomes among endoscopists with different experience levels. **Methods:** This single-center retrospective observational study included 1,093 inpatients who underwent colonoscopy at Inner Mongolia Autonomous Region People's Hospital from September 2021 to September 2022. Patients were divided into a conventional group and an AI-assisted group according to whether AI was used during colonoscopy. Endoscopists were further classified as senior or junior according to whether the cumulative number of colonoscopies performed before AI implementation exceeded 10,000. Polyp detection rate (PDR) and adenoma detection rate (ADR) were defined as the primary outcomes. Overall comparisons, stratified analyses by experience level, comparisons between the early and late phases after AI implementation, multivariable logistic regression, and difference-in-differences models were performed. **Results:** Overall, the AI-assisted group had a significantly higher PDR than the conventional group (61.6% vs 54.5%, $P = 0.023$), while the difference in ADR was not statistically significant (43.6% vs 39.1%, $P = 0.160$). Stratified analysis showed that among junior endoscopists, the AI-assisted group had a higher PDR than the conventional group (64.6% vs 56.2%, $P = 0.041$), whereas no significant differences in PDR or ADR were observed among senior endoscopists. Multivariable logistic regression demonstrated that AI use was independently associated with a higher PDR among junior endoscopists (aOR = 1.63, 95% CI: 1.07~2.48, $P = 0.024$), while no significant associations were found for the other outcomes. Comparisons between the early and late phases, subgroup multivariable analyses, and the overall difference-in-differences model did not show stable statistical differences; however, in the late period among junior endoscopists, the AI-assisted group showed a higher PDR than the conventional group (67.0% vs 54.8%, $P = 0.038$). **Conclusion:** AI-assisted colonoscopy improved polyp detection, and this advantage was more evident among junior endoscopists; its effect on adenoma detection still requires further validation.

Keywords

Artificial Intelligence, Colonoscopy, Polyp Detection Rate, Adenoma Detection Rate, Physician Seniority

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

结直肠癌是常见的消化系统恶性肿瘤,也是当前肿瘤防治中的重点疾病。GLOBOCAN 2022 数据显示,2022 年全球结直肠癌新发病例约 190 万例,死亡病例 90 余万例;按新发病例数计,结直肠癌居全球恶性肿瘤第 3 位,并已成为全球第 2 位癌症死亡原因[1]。我国结直肠癌疾病负担同样较重,2022 年新发病例约 51.71 万例,死亡病例约 24.00 万例,发病率和死亡率分别位居全部恶性肿瘤第 2 位和第 4 位[2] [3]。随着人口老龄化、城市化进程以及生活方式改变,结直肠癌相关疾病负担进一步加重[4]。

结肠镜兼具病变发现、病理取材和内镜治疗等作用,是结直肠癌筛查和早诊早治体系中最重要检查手段之一。其临床价值不仅取决于能否完成检查,更取决于检查质量是否稳定可靠。ADR、PDR、肠

道准备质量、退镜时间和盲肠插镜率等指标已逐步成为评价结肠镜质量的重要依据[5]-[8]。

随着深度学习和实时图像识别技术发展, AI 已逐步进入消化内镜临床实践。多项随机对照研究及 Meta 分析表明, AI 辅助结肠镜有助于提高息肉检出率和腺瘤检出率, 尤其在微小病变、平坦病变及边界欠清病变的识别中更易体现优势[9]-[12]。但这种改善并非在所有病例类型、研究场景和操作者中均一致出现。结肠镜检查质量本身高度依赖操作者经验、注意力分配及过程质量控制水平, 因此不同经验层级医师是否获得相同程度获益, 仍是值得进一步讨论的问题。因此, 本研究以真实世界住院结肠镜资料为基础, 探讨 AI 辅助结肠镜在不同年资医师中的检出效应差异。

2. 资料与方法

2.1. 研究对象与资料收集

采用单中心回顾性观察研究, 连续纳入 2021 年 9 月至 2022 年 9 月内蒙古自治区人民医院消化中心住院结肠镜患者 1093 例。纳入标准: 年龄 ≥ 18 岁, 完成结肠镜检查并留有完整内镜记录, Boston 肠道准备评分 ≥ 6 分。排除主要结局或关键协变量资料缺失者, 以及内镜资料或病理资料不完整者。研究资料来源于内镜中心系统、住院病历系统及病理报告系统。收集患者年龄、性别、Boston 评分、退镜时间、息肉相关内镜记录、病理资料, 以及 BMI、癌症病史、吸烟史和饮酒史等。

2.2. 分组与结局定义

根据检查时是否启用 AI 辅助检测系统, 将研究对象分为传统组和 AI 组。AI 辅助是指结肠镜检查过程中使用实时计算机辅助检测系统, 对内镜图像中的可疑病变区域进行自动识别和提示。根据截至 2021 年 9 月 AI 系统引入前医师累计完成的结肠镜操作总数, 将医师分为高年资组($>10,000$ 例)和低年资组($\leq 10,000$ 例); 根据各医师病例按时间前后连续纳入的顺序, 以纳入例数的大致中点为界, 将病例划分为应用初期和后期。PDR 定义为单次结肠镜检查中检出 1 枚及以上息肉; ADR 定义为单次结肠镜检查中检出 1 枚及以上经病理证实的腺瘤、黏膜内癌或原位癌。

2.3. 统计学方法

采用 SPSS 25.0 软件进行统计分析。计量资料以均数 \pm 标准差表示, 组间比较采用独立样本 t 检验; 计数资料以例数和百分比表示, 组间比较采用 χ^2 检验, 必要时采用 Yates 连续性校正或 Fisher 精确检验。采用多因素 logistic 回归分析评估 AI 辅助与 PDR、ADR 的相关性, 计算校正后优势比(aOR)及 95%置信区间(95% CI); 采用双重差分模型分析组别、时间及其交互作用与检出结局之间的关系, 并校正年龄、性别、是否使用麻醉剂、Boston 评分及退镜时间。以 $P < 0.05$ 为差异有统计学意义。

3. 结果

3.1. 基线资料比较

本研究共纳入 1093 例患者, 见表 1, 其中传统组 437 例, AI 组 656 例; 高年资医师完成 466 例, 低年资医师完成 627 例。两组在年龄、性别、BMI、吸烟史、饮酒史和癌症病史等基线资料方面差异均无统计学意义, 提示总体基线可比。

Table 1. Comparison of baseline data between AI and conventional groups

表 1. AI 组和传统组基线资料的比较

指标	传统组(n = 437)	AI 组(n = 656)	P 值
年龄(岁)	60.06 \pm 11.31	58.87 \pm 11.11	0.086

续表

男[n (%)]	248 (56.8)	376 (57.4)	0.880
BMI (kg/m ²)	23.52 ± 3.45	23.70 ± 3.24	0.585
有吸烟史[n (%)]	117/273 (42.9)	182/416 (43.8)	0.879
有饮酒史[n (%)]	89/262 (34.0)	135/407 (33.2)	0.897
有癌症史[n (%)]	57 (13.0)	73 (11.1)	0.388

注: BMI、吸烟史和饮酒史按非缺失值统计。

3.2. 总体程序指标及主要结局比较

表 2 总体比较显示, 传统组与 AI 组在医师年资构成、Boston 评分及退镜时间方面差异均无统计学意义, 而 AI 组麻醉比例明显高于传统组(78.0% vs 31.8%, $P < 0.001$)。在主要结局方面, AI 组 PDR 高于传统组(61.6% vs 54.5%), 差异有统计学意义($P = 0.023$); AI 组 ADR 高于传统组(43.6% vs 39.1%), 但差异无统计学意义($P = 0.160$)。

Table 2. Comparison of endoscopic procedure and outcomes between AI and conventional groups

表 2. AI 组和传统组内镜检查程序及主要结局比较

指标	传统组(n = 437)	AI 组(n = 656)	统计量	P 值
高年资医师[n (%)]	186 (42.6)	280 (42.7)	$\chi^2 = 0.000$	1.000
麻醉[n (%)]	139 (31.8)	512 (78.0)	$\chi^2 = 230.924$	<0.001
Boston 评分	6.02 ± 0.18	6.01 ± 0.15	t = 0.800	0.424
退镜时间(min)	8.56 ± 1.64	8.58 ± 1.43	t = -0.195	0.846
PDR [n (%)]	238 (54.5)	404 (61.6)	$\chi^2 = 5.201$	0.023
ADR [n (%)]	171 (39.1)	286 (43.6)	$\chi^2 = 1.972$	0.160

3.3. 不同年资医师层内比较

表 3 分层比较显示, 高年资层内传统组与 AI 组在 PDR 和 ADR 方面差异均无统计学意义; 低年资层内, AI 组 PDR 高于传统组(64.6% vs 56.2%), 差异有统计学意义($P = 0.041$), 而 ADR 差异无统计学意义。

Table 3. Comparison of primary outcomes between AI and conventional groups stratified by physician experience

表 3. 不同年资医师层内 AI 组与传统组主要结果比较

层别	结局	传统组	AI 组	统计量	P 值
高年资	PDR	97/186 (52.2)	161/280 (57.5)	$\chi^2 = 1.087$	0.297
高年资	ADR	69/186 (37.1)	118/280 (42.1)	$\chi^2 = 0.984$	0.321
低年资	PDR	141/251 (56.2)	243/376 (64.6)	$\chi^2 = 4.181$	0.041
低年资	ADR	102/251 (40.6)	168/376 (44.7)	$\chi^2 = 0.846$	0.358

3.4. 分层多因素 Logistic 回归分析

表 4 进一步分层多因素 logistic 回归分析显示, 在低年资医师层内, AI 与 PDR 升高独立相关(aOR =

1.63, 95% CI: 1.07~2.48, P = 0.024), 其余结局均未显示统计学意义。

Table 4. Stratified multivariable logistic regression outcomes of PDR and ADR with AI by physician seniority
表 4. 不同年资医师中 AI 与 PDR、ADR 的分层多因素 logistic 回归结果

结局	高年资 aOR (95% CI)	P 值	低年资 aOR (95% CI)	P 值
PDR	1.57 (0.91~2.70)	0.105	1.63 (1.07~2.48)	0.024
ADR	1.32 (0.82~2.13)	0.254	0.95 (0.66~1.39)	0.806

注: 模型校正年龄、性别、是否使用麻醉剂、Boston 评分及退镜时间。

3.5. 前后期 PDR 比较

表 5 显示, 在低年资传统组中, PDR 由初期 57.6% 降至后期 54.8%, 差异无统计学意义(P = 0.745); 低年资 AI 组由初期 62.2% 升至后期 67.0%, 组内前后比较差异无统计学意义(P = 0.388)。高年资传统组和高年资 AI 组的 PDR 在初期与后期之间差异也均无统计学意义(分别 P = 1.000 和 0.468)。同一时期 AI 组与传统组之间的比较显示, 低年资初期差异无统计学意义(P = 0.482), 低年资后期 AI 组 PDR 高于传统组, 差异有统计学意义(67.0% vs 54.8%, P = 0.038); 高年资初期和后期比较差异均无统计学意义(分别 P = 0.832 和 0.258)。

Table 5. Comparison of PDR between early and late periods by endoscopist experience
表 5. 不同年资医师 AI 应用初期及后期息肉检出率对比

组别	低年资			χ^2	P	高年资			χ^2	P
	例数	初期	后期			例数	初期	后期		
传统组	251	72/125 (57.6)	69/126 (54.8)	0.106	0.745	186	49/93 (52.7)	48/93 (51.6)	0.000	1.000
AI 组	376	117/188 (62.2)	126/188 (67.0)	0.745	0.388	280	77/140 (55.0)	84/140 (60.0)	0.526	0.468
χ^2	-	0.494	4.311			-	0.045	1.277		
P	-	0.482	0.038			-	0.832	0.258		

注: 本表比较不同年资、不同组别的 PDR 在 AI 应用初期和后期的变化。每行右侧 χ^2 和 P 值对应本行所表示的比较; 最下面两行为同一时期 AI 组与传统组之间的比较。

3.6. 前后期 ADR 比较

表 6 显示, 无论高年资还是低年资, 无论传统组还是 AI 组, ADR 在初期与后期之间的组内比较以及同一时期 AI 组与传统组之间的比较均无统计学意义(均 P > 0.05)。其中, 低年资 AI 组 ADR 在初期与后期均为 44.7%, 高年资 AI 组 ADR 由 42.9% 降至 41.4%, 但上述变化均无统计学意义。

Table 6. Comparison of ADR between early and late periods by endoscopist experience
表 6. 不同年资医师 AI 应用初期及后期腺瘤检出率对比

组别	低年资			χ^2	P	高年资			χ^2	P
	例数	初期	后期			例数	初期	后期		
传统组	251	52/125 (41.6)	50/126 (39.7)	0.033	0.857	186	33/93 (35.5)	36/93 (38.7)	0.092	0.761

续表

AI 组	376	84/188 (44.7)	84/188 (44.7)	0.000	1.000	280	60/140 (42.9)	58/140 (41.4)	0.015	0.904
χ^2	-	0.178	0.580			-	0.978	0.077		
P	-	0.673	0.446			-	0.323	0.781		

注: 本表比较不同年资、不同组别的 ADR 在 AI 应用初期和后期的变化。每行右侧 χ^2 和 P 值对应本行所表示的比较; 最下面两行为同一时期 AI 组与传统组之间的比较。

3.7. 亚组内前后多因素 Logistic 回归分析

表 7 显示, 在控制年龄、性别、是否使用麻醉剂、Boston 评分和退镜时间后, 高年资传统组、高年资 AI 组、低年资传统组及低年资 AI 组的 PDR 和 ADR 在后期与初期之间差异均无统计学意义(均 $P > 0.05$)。其中, 高年资传统组 PDR 和 ADR 的 aOR 分别为 0.77 (95% CI: 0.38~1.60) 和 1.04 (95% CI: 0.55~1.98), 高年资 AI 组分别为 1.53 (95% CI: 0.85~2.77) 和 0.95 (95% CI: 0.57~1.57); 低年资传统组分别为 0.98 (95% CI: 0.54~1.81) 和 0.93 (95% CI: 0.54~1.61), 低年资 AI 组分别为 1.34 (95% CI: 0.83~2.17) 和 1.07 (95% CI: 0.69~1.65)。

Table 7. Logistic regression analysis outcomes of factors affecting PDR and ADR during early vs. late phases of AI adoption by physician seniority

表 7. 不同年资医师 AI 应用初期及后期息肉及腺瘤检出影响因素的 logistic 回归分析结果

年资组	组别	结局变量	n	aOR (95% CI)	P
高年资	传统组	PDR	186	0.77 (0.38~1.60)	0.488
高年资	传统组	ADR	186	1.04 (0.55~1.98)	0.901
高年资	AI 组	PDR	279	1.53 (0.85~2.77)	0.160
高年资	AI 组	ADR	279	0.95 (0.57~1.57)	0.835
低年资	传统组	PDR	251	0.98 (0.54~1.81)	0.959
低年资	传统组	ADR	251	0.93 (0.54~1.61)	0.788
低年资	AI 组	PDR	376	1.34 (0.83~2.17)	0.234
低年资	AI 组	ADR	376	1.07 (0.69~1.65)	0.778

注: 分别在高年资传统组、高年资 AI 组、低年资传统组和低年资 AI 组中, 以应用后期相对于初期为比较方式, 纳入年龄、性别、是否使用麻醉剂、Boston 评分及退镜时间进行多因素 logistic 回归分析。aOR 为校正后优势比。BMI、吸烟史及饮酒史因缺失较多未纳入模型。

3.8. 双重差分分析

表 8 显示, 在高年资和低年资医师的 PDR、ADR 双重差分全模型中, 内镜应用组别、应用前后及其交互项差异均无统计学意义(均 $P > 0.05$)。其中, 各模型交互项 P 值均大于 0.05, 未见统计学意义。协变量分析显示, 退镜时间在各模型中均具有统计学意义(均 $P < 0.001$)。

Table 8. Full model results of the DID analysis for polyp and adenoma detection
表 8. 息肉及腺瘤检出的双重差分全模型结果

变量	统计值	内镜应用组别(0 = AI 组, 1 = 普通组)	应用前后(0 = 初期, 1 = 后期)	内镜应用组别 × 应用前后	性别	年龄	肠道评分	退镜时间
高年资 PDR	β	-0.211	0.345	-0.504	0.186	-0.003	0.580	1.373
	P	0.560	0.243	0.281	0.424	0.786	0.474	<0.001
低年资 PDR	β	-0.269	0.355	-0.435	0.053	0.008	-0.084	1.017
	P	0.349	0.149	0.259	0.781	0.380	0.870	<0.001
高年资 ADR	β	-0.328	-0.057	0.101	0.128	0.004	0.497	0.451
	P	0.305	0.822	0.806	0.530	0.657	0.455	<0.001
低年资 ADR	β	0.049	0.039	-0.003	0.052	0.007	-0.104	0.445
	P	0.850	0.858	0.993	0.759	0.372	0.833	<0.001

注：内镜应用组别(0 = AI 组, 1 = 传统组)；应用前后(0 = 初期, 1 = 后期)；内镜应用组别 × 应用前后为交互项。模型同时纳入性别、年龄、是否使用麻醉剂、肠道评分和退镜时间进行调整。BMI、吸烟史及饮酒史因缺失较多未纳入模型。

4. 讨论

4.1. 主要结果概述

本研究基于真实世界住院结肠镜资料，从总体比较、医师年资分层比较、AI 应用初期与后期比较、亚组内前后多因素 logistic 回归及双重差分分析等多个层面，对 AI 辅助结肠镜的检出表现进行了考察。结果显示，AI 辅助与总体 PDR 升高相关，而 ADR 差异无统计学意义；在按医师年资分层后，AI 相关优势主要集中于低年资医师的 PDR；在进一步纳入协变量后，上述结果总体方向未发生明显改变，其中低年资医师层内 AI 与 PDR 升高的相关性仍具有统计学意义。前后期比较、亚组内前后多因素 logistic 回归及双重差分分析未见稳定统计学差异，但在低年资医师后期同期比较中，AI 组 PDR 高于传统组。

这一结果提示，AI 辅助在真实世界住院结肠镜场景中的作用主要体现在病变发现层面，而且更容易在低年资医师中观察到。既往随机对照研究和 Meta 分析普遍提示，AI 辅助有助于提高息肉或腺瘤检出率，尤其在微小病变、平坦病变和边界欠清病变的识别方面更易体现优势[10]-[12]。本研究中 AI 组总体 PDR 高于传统组，低年资医师层面的 PDR 优势也较为明显，与既往研究总体方向一致。

4.2. PDR 与 ADR 结果差异的可能原因

本研究在总体 ADR、分层 ADR 及交互项分析中均未观察到有统计学意义的差异，提示真实世界资料中 AI 辅助的表现较试验环境更为复杂。PDR 更接近结肠镜现场的病变发现能力，而 ADR 除受发现能力影响外，还受到病理类型分布、送检完整性、切除策略和样本规模等因素制约。因此，PDR 与 ADR 虽然相关，但并不等同：前者更接近“是否看到病变”，后者更接近“是否发现具有明确癌前意义的病变”。因此，真实世界资料中首先观察到 PDR 差异，而 ADR 未形成稳定统计学结果，具有一定合理性，也提示 AI 作用可能首先体现在病变初始识别层面，而未必同步转化为腺瘤性病变检出改善。

进一步看，本研究中低年资医师层内 PDR 升高而 ADR 未见同步改善，提示 AI 更可能首先减少视觉遗漏、提高可疑病变发现率，但这些新增病灶未必主要由腺瘤构成。AI 实时提示对微小病变、平坦病变及边界欠清病变更敏感，因此额外发现的病灶中，除腺瘤外，也可能包含增生性息肉、炎性或其他临床意义相对有限的病变。ADR 作为结直肠癌预防相关的重要质量指标，对病理构成和样本规模更为敏感，

因此即使 PDR 已出现一定程度提高, 也未必能够在现阶段同步转化为稳定的 ADR 差异。需要指出的是, 本研究未进一步对检出息肉的病理类型构成进行拆分分析, 也未按大小、部位和形态对新增检出病灶进行描述, 因此目前尚不能直接回答 AI 额外检出的息肉中有多少属于腺瘤性病变、增生性息肉或其他临床意义相对有限的病变。这是本研究解释 PDR 与 ADR 结果不一致时的重要限制, 也提示后续研究有必要进一步补充病理构成分析。

4.3. 麻醉差异及其可能影响

本研究中, AI 组麻醉比例明显高于传统组(78.0% vs 31.8%, $P < 0.001$), 是两组间最突出的程序相关差异之一。该结果提示, 本研究中 AI 应用并非完全发生在与传统组均质的检查场景中, 而更可能与检查房间配置、麻醉资源安排、患者检查方式选择及临床流程分配等现实因素有关。

从临床操作过程看, 麻醉可提高患者检查过程中的舒适度和依从性, 减少因腹痛、不适、紧张及体动所带来的干扰, 从而有利于内镜推进和退镜观察的顺利进行。在患者配合较好的情况下, 医师可以将更多注意力集中于肠腔展开、视野清洁、皱襞后观察和可疑病灶复核, 而不必过多分散于安抚患者、处理中断性体动或反复调整操作节奏。尤其对于低年资医师而言, 麻醉所带来的操作稳定性提升, 可能更有助于其维持完整的观察顺序与退镜节奏, 从而在一定程度上放大 AI 实时提示的辅助价值。另一方面, 麻醉的影响并不一定仅仅体现为患者“更配合”, 还可能与病例安排、房间功能、护理配合和操作习惯等程序因素交织存在。虽然本研究已在多因素 logistic 回归及双重差分模型中将是否使用麻醉剂纳入校正, 但统计学调整并不能完全替代机制层面的讨论。因此, 对于本研究结果, 更稳妥的理解是: AI 辅助结肠镜是在较高麻醉使用背景下显示出 PDR 提升, 提示 AI 效应可能与麻醉相关的患者配合改善、操作稳定性提高及程序管理因素共同作用。

4.4. 不同年资医师获益差异的可能原因

低年资医师更容易观察到 AI 辅助相关的 PDR 优势, 而高年资医师层面未见同样稳定的统计学结果, 这一差异具有一定合理性。高年资医师通常已形成较稳定的退镜观察顺序、病变识别经验及检查节奏, 因此 AI 带来的额外提升空间相对有限; 低年资医师仍处于经验积累和观察模式逐步稳定的阶段, AI 提供的实时提示更易帮助其减少病变遗漏, 从而首先表现为 PDR 提高[13][14]。

双重差分全模型中, 内镜应用组别、应用前后及其交互项均未见统计学意义, 提示本研究中 AI 辅助相较传统组在前后期变化幅度上的额外效应尚未显示稳定证据。另一方面, 退镜时间在各双重差分模型中均呈正向且有统计学意义, 再次提示更充分的退镜观察与更高检出率之间存在稳定关系; 真实世界研究也提示, AI 在常规临床环境中的作用不仅体现在检出率变化上, 还可能有助于提高检查质量的稳定性[15]。

4.5. 研究优势与局限性

本研究的优势在于基于真实世界住院结肠镜资料, 从医师年资分层角度考察 AI 辅助的检出表现, 并采用总体比较、前后期比较、分层回归、亚组内前后多因素分析及双重差分模型进行综合评价。局限性在于: 本研究为单中心回顾性观察研究, 因果推断能力有限; 病例来源为住院患者, 与筛查人群存在差异; 部分变量缺失较多, 未纳入主分析回归模型; AI 应用初期与后期的划分基于病例连续纳入顺序, 仍难完全等同于严格意义上的前瞻性学习曲线分析。

5. 结论

在本研究的真实世界住院结肠镜场景中, AI 辅助结肠镜与息肉检出率提高相关, 且该优势在低年资医师中更易体现; 但其对腺瘤检出率的促进作用尚未显示稳定证据, 相关结果仍可能受到麻醉差异及病

变构成等因素影响, 尚需进一步研究验证。

声明

经伦理委员会审批(SC-07/02KT2024157Y)及获得受试者知情同意。

参考文献

- [1] Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., *et al.* (2024) Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, **74**, 229-263. <https://doi.org/10.3322/caac.21834>
- [2] 中华人民共和国国家卫生健康委员会医政司, 中华医学会肿瘤学分会. 国家卫生健康委员会中国结直肠癌诊疗规范(2025版)(精简版)[J]. 中国实用外科杂志, 2025, 45(12): 1353-1359.
- [3] Han, B., Zheng, R., Zeng, H., Wang, S., Sun, K., Chen, R., *et al.* (2024) Cancer Incidence and Mortality in China, 2022. *Journal of the National Cancer Center*, **4**, 47-53. <https://doi.org/10.1016/j.jncc.2024.01.006>
- [4] Roshandel, G., Ghasemi-Kebria, F. and Malekzadeh, R. (2024) Colorectal Cancer: Epidemiology, Risk Factors, and Prevention. *Cancers*, **16**, Article No. 1530. <https://doi.org/10.3390/cancers16081530>
- [5] Kaminski, M.F., Regula, J., Kraszewska, E., Polkowski, M., Wojciechowska, U., Didkowska, J., *et al.* (2010) Quality Indicators for Colonoscopy and the Risk of Interval Cancer. *New England Journal of Medicine*, **362**, 1795-1803. <https://doi.org/10.1056/nejmoa0907667>
- [6] Corley, D.A., Jensen, C.D., Marks, A.R., Zhao, W.K., Lee, J.K., Doubeni, C.A., *et al.* (2014) Adenoma Detection Rate and Risk of Colorectal Cancer and Death. *New England Journal of Medicine*, **370**, 1298-1306. <https://doi.org/10.1056/nejmoa1309086>
- [7] Rex, D.K., Schoenfeld, P.S., Cohen, J., Pike, I.M., Adler, D.G., Fennerty, M.B., *et al.* (2015) Quality Indicators for Colonoscopy. *Gastrointestinal Endoscopy*, **81**, 31-53. <https://doi.org/10.1016/j.gie.2014.07.058>
- [8] Kaminski, M., Thomas-Gibson, S., Bugajski, M., Bretthauer, M., Rees, C., Dekker, E., *et al.* (2017) Performance Measures for Lower Gastrointestinal Endoscopy: A European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. *Endoscopy*, **49**, 378-397. <https://doi.org/10.1055/s-0043-103411>
- [9] Ali, H., Muzammil, M.A., Dahiya, D.S., *et al.* (2024) Artificial Intelligence in Gastrointestinal Endoscopy: A Comprehensive Review. *Annals of Gastroenterology*, **37**, 133-141. <https://doi.org/10.20524/aog.2024.0861>
- [10] Wang, P., Berzin, T.M., Glissen Brown, J.R., Bharadwaj, S., Becq, A., Xiao, X., *et al.* (2019) Real-Time Automatic Detection System Increases Colonoscopic Polyp and Adenoma Detection Rates: A Prospective Randomised Controlled Study. *Gut*, **68**, 1813-1819. <https://doi.org/10.1136/gutjnl-2018-317500>
- [11] Repici, A., Badalamenti, M., Maselli, R., Correale, L., Radaelli, F., Rondonotti, E., *et al.* (2020) Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. *Gastroenterology*, **159**, 512-520.e7. <https://doi.org/10.1053/j.gastro.2020.04.062>
- [12] Wallace, M.B., Sharma, P., Bhandari, P., East, J., Antonelli, G., Lorenzetti, R., *et al.* (2022) Impact of Artificial Intelligence on Miss Rate of Colorectal Neoplasia. *Gastroenterology*, **163**, 295-304.e5. <https://doi.org/10.1053/j.gastro.2022.03.007>
- [13] Biscaglia, G., Cocomazzi, F., Gentile, M., Loconte, I., Mileti, A., Paolillo, R., *et al.* (2022) Real-Time, Computer-Aided, Detection-Assisted Colonoscopy Eliminates Differences in Adenoma Detection Rate between Trainee and Experienced Endoscopists. *Endoscopy International Open*, **10**, E616-E621. <https://doi.org/10.1055/a-1783-9678>
- [14] Lau, L.H.S., Ho, J.C.L., Lai, J.C.T., Ho, A.H.Y., Wu, C.W.K., Lo, V.W.H., *et al.* (2024) Effect of Real-Time Computer-Aided Polyp Detection System (ENDO-AID) on Adenoma Detection in Endoscopists-in-Training: A Randomized Trial. *Clinical Gastroenterology and Hepatology*, **22**, 630-641.e4. <https://doi.org/10.1016/j.cgh.2023.10.019>
- [15] Ham, D.Y., Lee, J.G., Ahn, C.I., Kae, S.H. and Jang, H.J. (2025) Artificial Intelligence-Assisted Colonoscopy Improves Adenoma Detection Rates in Routine Colonoscopy Practice: A Single-Center, Retrospective, Propensity Score-Matched Study with Concurrent Controls. *BMC Gastroenterology*, **25**, Article No. 755. <https://doi.org/10.1186/s12876-025-04011-w>