



[DOI] 10.3969/j.issn.1001-9057.2026.04.003

<http://www.lcnkzz.com/CN/10.3969/j.issn.1001-9057.2026.04.003>

· 综述与讲座 ·

消化内镜人工智能质量控制: 中国与国际指南比较、实践差异与中国方案的启示

林晓清 丁震

[摘要] 随着人工智能(AI)技术的快速发展,其在消化内镜领域的应用已从单纯的病灶辅助检测向全流程质量控制延伸。然而,全球范围内针对内镜AI质控的标准构建仍存在显著差异。本文系统梳理我国及国际主要学会在胃镜、肠镜、超声内镜(EUS)及胶囊内镜质控方面的最新指南与共识,分析各国在循证评价、推荐强度及临床实践层面的立场差异。结合我国高负荷内镜诊疗现状,提出构建具有中国特色的内镜AI全生命周期质控体系的思考,以期为全球内镜AI的规范化应用提供参考。

[关键词] 消化内镜; 人工智能; 质量控制; 国际指南; 中国方案

[中图分类号] R57 **[文献标识码]** A

消化内镜是实现消化道肿瘤早发现、早诊断、早治疗的核心手段,但其临床获益很大程度上取决于操作者的检查质量。长期以来,消化内镜质控主要依赖医师主观自觉医院及科室层面的事后抽查,评价指标多局限于腺瘤检出率、盲肠插镜率及退镜时间等结果性数据。这种模式存在明显局限:一是事后抽检具有滞后性,难以实时纠正操作偏差;二是各医疗机构信息系统不统一,人工抽检工作量大,且难以实现不同医院及地区间的标准化与同质化。

近年来,随着深度学习算法的突破,AI在消化内镜领域迅速发展。现代内镜AI系统已从单纯的病灶辅助检测工具,发展为具备实时监测与标准量化能力的质控辅助系统,能够实时监测退镜速度、自动评估黏膜观察覆盖、识别解剖盲区及微小病变等,为突破传统质控瓶颈提供了技术可能。然而,全球范围内关于消化内镜AI质控的规范化建设呈现明显差异。2023~2026年间,我国基于海量数据积累和临床需求,已在胃镜、肠镜及超声内镜(EUS)等领域发布多项专家共识,对AI作为质控工具给予较高级别的推荐。同时,欧洲胃肠内镜学会(ESGE)、美国胃肠病协会(AGA)、美国胃肠内镜学会(ASGE)及日本消化器内视镜学会

(JGES)等国际主要学会近年发布的指南与立场声明则更为审慎,对AI降低如总体死亡率等绝对净获益提出了更严格的循证要求,并将假阳性所致过度诊断、卫生经济学成本及法律伦理置于核心考量。

这种立场差异反映了不同国家在医疗资源、诊疗负荷及循证评价体系方面的差异。我国拥有全球规模最大的内镜筛查人群和较复杂的基层医疗生态,内镜质控既需关注高水平中心的持续改进,也需致力于提升基层医疗的同质化水平。因此,既不能简单套用欧美基于高循证门槛的保守立场,也不应在缺乏统一标准的情况下任其无序发展。本文旨在横向比较主要内镜领域近年国内外主流指南与共识,分析差异原因,并探讨构建具有中国特色的内镜AI质控体系的启示。

一、中外指南核心对比:AI在消化内镜各专科质控中的立场与差异

1. 胃镜质控:在上消化道内镜(EGD)领域,胃部复杂的皱襞结构使检查中的盲区遗漏高度依赖操作者经验,成为早期胃癌漏诊的主要原因。因此,该领域AI质控的核心在于规范检查路径、确保黏膜观察的完整性。

面对国内高负荷的胃镜筛查需求,2023年发布的《上消化道内镜人工智能系统临床应用专家共识(2023,武汉)》^[1]将AI辅助EGD检查时间监测及盲区监测列为强推荐的质量控制工具,共识水平分别为95.45%和100.00%。以WISENSE系统为例,该系统

基金项目:国家自然科学基金面上项目(82270675、82470670);国家重点研发计划“病原学与防疫技术体系研究”重点专项项目(2023YFC2307004)

作者单位:510080 广州,中山大学附属第一医院消化内科

通信作者:丁震,E-mail:dingzh26@mail.sysu.edu.cn

将上消化道划分为 26 个解剖部位进行实时监测,在 Wu 等^[2]发表的于《Gut》(2019 年)的研究中显示,其识别盲区的准确率为 90.40%,AI 辅助组的盲区率显著低于常规对照组(5.86% 比 22.46%, $P < 0.001$)。

与我国共识的积极推广不同,国际主流学术组织对该领域 AI 质控的常规应用持更为保守的立场。AGA 2025 年发布的《胃癌高风险个体筛查和监测临床实践更新》指出,尽管计算机辅助检测(CADe)和计算机辅助诊断(CADx)系统在早期肿瘤检测中展现潜力,但由于数据尚不充分,暂不建议作为常规工具使用^[3]。JGES 2025 年的最新声明也认为,现有证据不足以确定 CADe 或 CADx 在上消化道内镜检查中能否真正提高肿瘤检出率或光学诊断准确性^[4]。

尽管在推广时机上存在分歧,但对合格 AI 质控系统的性能要求已逐步形成共识。ESGE 2022 年立场声明指出,若要使 AI 被接受用于评估上消化道内镜检查的完整度,其对黏膜的检查水平须与经验丰富的内镜医师相当,且须在至少 90% 的检查中实现相关解剖标志的自动识别和拍照记录^[5]。

2. 肠镜质控:在 AI 辅助结肠镜质控领域,国内外指南均形成了从积极推荐到审慎观望的明确谱系。2023 年,中华医学会消化内镜学分会发布的《肠镜人工智能系统临床应用专家共识(2023,武汉)》^[6]以强推荐和 A 级证据质量,明确支持 AI 在全流程质量控制中的应用;推荐使用计算机辅助质量控制提高腺瘤检出率(共识水平 86.36%)、使用 AI 辅助计算退镜时间(共识水平 95.45%)及 AI 辅助统计盲肠插镜率(共识水平 95.45%)。这一立场将 AI 质控工具定位为肠镜检查质量提升的临床标准配置,既关注息肉检出,也强调对检查完整度和操作规范性的客观监督。

ESGE 2025 年的立场声明仅对 CADe 给出弱推荐^[7],其依据是基于 OperA 微模拟模型推算的远期获益——每 10 000 例筛查患者中可减少 9 例结直肠癌新发和 2 例恶性肿瘤死亡,但同时存在息肉过度诊断和监测负担增加的风险。19 位投票专家中,13 位支持使用、6 位反对,显示专家组内部对获益与风险的价值判断存在分歧。AGA 2025 年发布的动态临床实践指南则采取最保守姿态,因关键结局的证据确定性被评估为极低,且实现上述获益需额外进行 635 次监测性结肠镜检查,指南最终未提出任何支持或反对 CADe 使用的推荐^[8]。日本 JGES 2025 年的立场声明同样采取审慎路径,指出 CADe 有望提高腺瘤检出率并可能降低结直肠癌发病率,但明确表示尚无充分证据支持这一观点,同时还强调 CADx 对减少不必要息肉切除的潜力仍需更多证据积累^[4]。世界内镜组织(WEO)

2023 年的声明则更多着眼于技术趋势和成本效益,认为 CADe 短期内会增加医疗成本,但长期或可因恶性肿瘤治疗费用节省而实现抵消^[9]。

3. EUS 质控:基于现有搜索,国外主流消化内镜学会(如 ASGE、ESGE、JGES、WEO 等)尚未发布专门针对 AI 在 EUS 质量控制领域应用的临床实践指南或立场声明。值得一提的是,EUS 的系统留图规范是 ASGE 认可的检查质量指标^[10],也有临床试验正在验证 EUS 智能图像报告系统自动生成报告的可行性^[11]。但这些属于技术验证或框架性研究,尚未上升至学会层面的临床应用推荐。

与肠镜 AI 领域的现状不同,在 EUS 领域,目前仅检索到我国发布的《2025 超声内镜人工智能系统临床应用专家共识(2024,武汉)》^[12]。该共识明确指出,AI 技术在 EUS 的质量控制、病变识别和诊断等方面表现出高效性和准确性,将质控列为 AI 在 EUS 中的核心应用维度之一。国外则暂无专门指南。可能的原因包括:相比筛查量巨大、更容易获得高质量研究证据的结肠镜检查,EUS 操作门槛高、数据获取难度大,相关 AI 研究多集中在病灶检测与定性,质量控制方向的证据积累尚不充分。从肠镜 AI 指南可见,国外学会在证据不充分时倾向于不提出推荐,EUS AI 质控的证据基础目前很可能尚未达到国外学会的指南制定门槛。

4. 胶囊内镜质控:在胶囊内镜 AI 质控方面,我国共识是目前唯一给出明确临床应用推荐的文件。中华医学会消化内镜学分会 2024 年会发布的《胶囊内镜人工智能系统临床应用专家共识(2024,上海)》^[13]率先对 AI 在胶囊内镜质控中的应用做出明确推荐。该共识涵盖 9 条推荐意见,在质量控制维度,以 B 级证据质量和 100% 的共识水平建议使用 AI 提升阅片效率;以上消化道胶囊内镜检查为例,以 C 级证据质量和 100% 的共识水平建议使用 AI 辅助实时监测上消化道解剖图像作为质量控制工具,通过辅助解剖识别降低病灶漏诊率、提高检查效率;此外,共识还以 B 级证据质量和 90% 的共识水平建议使用 AI 辅助进行肠道准备评估作为小肠胶囊内镜检查的质量评估方法之一,并且指出 AI 在结直肠胶囊内镜肠道准备评估中亦有良好应用前景。这一系列推荐将 AI 质控的功能从辅助阅片拓展至实时解剖识别和肠道准备评估等多个环节。

欧洲 ESGE 2022 年发布的立场声明并未专门针对胶囊内镜的质量控制给出独立推荐,而是基于 AI 在胃肠内镜中的预期价值总框架下,对 AI 在胶囊内镜中的应用提出了性能接受标准。该声明指出,为使 AI 被接受用于小肠胶囊内镜的自动病灶检测,AI 辅助阅片

的性能应与经验丰富的内镜医师相当,在不增加甚至可能减少操作者阅片时间的前提下完成病灶检测。但 ESGE 未针对胶囊内镜的阅片效率、实时解剖监测或肠道准备评估等具体质控环节给出临床应用推荐。随着胶囊内镜 AI 技术的进一步成熟和证据积累,国际学会在这一领域的立场有望逐步明晰。

二、国际实践差异的原因分析

通过对比可见,中国方案与国际指南在推荐强度上存在显著差异,其可能的原因包括:

第一,循证评价体系的侧重点与证据解读阈值不同。国际指南(尤其是 AGA)遵循动态指南模式,对证据确定性的要求较高,其推荐决策高度依赖例如结直肠癌发病率和死亡率这样的远期硬终点。例如,在肠镜 AI 使用的领域,ESGE 立场声明^[7]、AGA 实践指南^[8]以及发表在 BMJ 上的一篇推荐意见^[14]均以同一 OperA 项目证据为基础,却得出了弱推荐使用、无推荐和弱推荐不使用三种不同结论。这一现象并非源于证据本身的矛盾,而是各方对证据的解读阈值、获益与风险的权重分配及专家组构成差异所致。其揭示了当前 AI 辅助内镜领域的核心争议:以腺瘤检出率为代表的中间指标提升已有较一致的证据支持,但从中间指标到结直肠癌发病率和死亡率这一硬终点的证据链条尚不完整,且缺乏高质量的成本效益数据。我国共识在证据积累阶段采取更为积极的策略,将 AI 质量控制功能与辅助诊断功能并列作为推荐项,体现了对检查过程本身规范化价值的独立认可;国际指南则在证据不确定性的背景下选择更为保守的立场。

第二,医疗资源的供需矛盾与伦理风险考量不同。我国内镜医师面临较高的诊疗负荷,AI 质控不仅是提升质量上限的手段,更是保障基层医疗水平同质化底线的工具。在这一背景下,AI 辅助的实时监测与标准化记录具有明确的现实需求。欧美等发达国家则更关注 AI 可能带来的过度诊断对医疗保险体系的压力,以及由此可能引发的监测性内镜检查次数增加和医疗资源消耗,因而在成本效益尚不明朗时趋于保守。此外,如 WEO 2026 年共识所指出的,AI 质控中的数据隐私与责任归属问题在不同法律体系下面临不同的规范要求^[15],这也影响了各地区的推荐强度。

三、中国方案的启示:构建全生命周期质控体系

借鉴 ASGE 2023 年 AI 算法评估框架^[16]及 WEO 2026 年法律伦理共识^[15]等国际经验,未来中国方案可在以下维度持续完善:(1)建立标准化的量化评价指标。应超越单一的准确率指标,引入更符合质控需

求的评价维度:实时性方面,参考 ASGE 建议,处理延迟需控制在 2 秒以内,以减少操作干扰和报警疲劳;稳健性方面,强调跨设备、跨品牌的外部验证,解决 AI 系统在不同型号内镜下的性能衰减问题。(2)推进全流程闭环质控模式。质控应贯穿检查全程,构建术前准备评估、术中实时监测及术后报告与检查质量关键指标质控的完整闭环。(3)明确伦理规范与法律边界。如 WEO 2026 年共识所指出的,AI 质控中的数据隐私与责任归属不可回避^[15]。中国方案应明确 AI 是质量监测的辅助工具,医师始终是临床决策的责任主体,并探索建立 AI 辅助下的医疗风险分担机制。

四、结语与展望

消化内镜 AI 的质量控制正从理论探索走向临床实践。尽管国际上对 AI 质控的绝对获益仍存争议,但其在规范化、标准化方面的推动作用已获广泛认可。中国方案应在保持临床应用先发优势的同时,加强高质量循证研究,以期为全球消化内镜 AI 质控标准的建立贡献更多实证依据。

参 考 文 献

- [1] 中华医学会消化内镜学分会大数据协作组. 上消化道内镜人工智能系统临床应用专家共识(2023,武汉)[J]. 中华消化内镜杂志, 2024, 41(2): 85-96.
- [2] Wu L, Zhang J, Zhou W, et al. Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy [J]. Gut, 2019, 68(12): 2161-2169.
- [3] Shah SC, Wang AY, Wallace MB, et al. AGA Clinical Practice Update on Screening and Surveillance in Individuals at Increased Risk for Gastric Cancer in the United States: Expert Review [J]. Gastroenterology, 2025, 168(2): 405-416. e1.
- [4] Mori Y, Ishihara R, Ogata H, et al. Artificial Intelligence in Gastrointestinal Endoscopy: The Japan Gastroenterological Endoscopy Society Position Statements [J]. Dig Endosc, 2025, 37(10): 1116-1122.
- [5] Messmann H, Bisschops R, Antonelli G, et al. Expected value of artificial intelligence in gastrointestinal endoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement [J]. Endoscopy, 2022, 54(12): 1211-1231.
- [6] 中华医学会消化内镜学分会大数据协作组. 肠镜人工智能系统临床应用专家共识(2023,武汉)[J]. 中华消化内镜杂志, 2024, 41(4): 253-262.
- [7] Bretthauer M, Ahmed J, Antonelli G, et al. Use of computer-assisted detection (CAdE) colonoscopy in colorectal cancer screening and surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement [J]. Endoscopy, 2025, 57(6): 667-673.
- [8] Sultan S, Shung DL, Kolb JM, et al. AGA Living Clinical Practice Guideline on Computer-Aided Detection-Assisted Colonoscopy [J]. Gastroenterology, 2025, 168(4): 691-700.
- [9] Mori Y, East JE, Hassan C, et al. Benefits and challenges in implementation of artificial intelligence in colonoscopy: World Endoscopy Organization position statement [J]. Dig Endosc, 2023, 35(4): 422-429.
- [10] Wani S, Wallace MB, Cohen J, et al. Quality indicators for EUS [J]. Am J Gastroenterol, 2015, 110(1): 102-113.
- [11] Li X, Yao L, Wu H, et al. A deep learning-based, real-time image report system for linear EUS [J]. Gastrointest Endosc, 2025, 101(6): 1166-1173. e11.
- [12] 中华医学会消化内镜学分会大数据协作组. 超声内镜人工智能系统临床应用专家共识(2024,武汉)[J]. 中华消化内镜杂志, 2025, 42(2): 94-103.
- [13] 中华医学会消化内镜学分会大数据协作组. 胶囊内镜人工智能系统临床应用专家共识(2024,上海)[J]. 中华消化内镜杂志, 2025, 42(4): 258-265.
- [14] Foroutan F, Vandvik PO, Helsing LM, et al. Computer aided detection and diagnosis of polyps in adult patients undergoing colonoscopy: a living clinical practice guideline [J]. BMJ, 2025, 388: e082656.



[DOI] 10.3969/j.issn.1001-9057.2026.04.004

http://www.lcnkzz.com/CN/10.3969/j.issn.1001-9057.2026.04.004

· 综述与讲座 ·

多模态数据融合的结直肠息肉诊疗闭环管理质量控制新范式

蒋欢 袁湘蕾 牟一 高远 胡兵

【摘要】 结直肠息肉是结直肠癌发生发展的关键前驱病变,其相关诊疗质量直接影响恶性肿瘤防控效果。当前,人工智能(AI)在该领域的质控应用主要集中于内镜检查等“镜下”环节,虽显著提升了腺瘤检出率(ADR)与操作规范性,但在术前评估与术后随访等环节仍存在管理断点和明显不足。本文提出构建以大语言模型(LLM)驱动的 AI Agent 为核心并融合多模态数据的结直肠息肉诊疗全程闭环管理质控框架,推动该领域质控由单点优化向全程闭环协同转变。尽管仍面临模型可靠性、数据安全及临床落地等挑战,但随着相关技术与制度的不断完善,该模式在结直肠癌防控中的应用前景广阔。

【关键词】 结直肠息肉; 质量控制; 多模态; 人工智能; 大语言模型

【中图分类号】 R574.6

【文献标识码】 A

结直肠癌是威胁我国居民健康的重大疾病,我国最新流行病学数据显示,结直肠癌的发病率和死亡率在恶性肿瘤中分别列居第2位和第4位^[1]。大量研究表明,结直肠癌的发生发展多遵循“正常黏膜-腺瘤性息肉-癌”的演进序列,这一过程通常历时10~15年,早期发现和切除结直肠息肉是阻断其癌变进程的最有效手段^[2-3]。然而,结直肠息肉诊疗全程各环节的质量水平参差不齐,如何在规模化筛查背景下实现全流程的精细化质控,已成为制约息肉切除远期获益的主要瓶颈。近年来,人工智能(AI)技术的快速兴起及其与医疗领域的深度融合,为结直肠息肉诊疗的质控提供了新的技术路径。本文综合分析了当前结直肠息肉诊疗领域 AI 质控的现状与局限,提出“从镜下质控向全程质控延伸”的理念,系统阐述了基于多模态数据融合与 AI Agent 的闭环管理质控体系,探讨其关键技术、框架设计及应用挑战,为推动结直肠息肉诊疗质控从“单点优化”迈向“全程闭环”提供参考。

基金项目:四川省卫生健康委员会科技项目(23LCYJ037)

作者单位:610041 成都,四川大学华西医院消化内科 消化内镜医工研究室

通信作者:胡兵, E-mail:hubing@wchscu.edu.cn

一、结直肠息肉诊疗的 AI 质控现状

1. “镜下质控”:技术成熟的核心应用场景

结直肠息肉的诊疗贯穿术前筛查、内镜下切除及术后随访等多个环节。近年来 AI 技术的发展推动了该领域质控的显著进步,但相关研究主要聚焦于内镜检查与操作过程,即“镜下质控”环节。

腺瘤检出率(ADR)是评价结肠镜检查质量的核心指标,其每提高1%,间期结直肠癌的发生率约下降3%,致命性间期癌风险降低约5%^[4]。近年来,基于深度学习的计算机辅助检测(CADe)系统通过对内镜视频流进行实时分析,可自动识别疑似息肉区域并给予视觉提示,从而有效减少漏检,尤其在小型及扁平病变的识别方面表现出明显优势^[5]。多项随机对照研究表明,CADe 系统的应用可显著提高 ADR,并在一定程度上降低不同操作者之间的检出差异,推动结肠镜检查质量的标准化与均质化^[6-10]。

除结果指标外,过程质量控制同样重要。退镜时间被认为是影响 ADR 的重要因素之一,研究表明,退镜时间不足6分钟与较低 ADR 显著相关^[11-12]。基于 AI 的质控系统可自动记录退镜时间并进行实时提醒,

[15] Ahmad OF, Mori Y, Bretthauer M, et al. The Legal and Ethical Framework for Artificial Intelligence in Gastrointestinal Endoscopy: A World Endoscopy Organization International Consensus Statement[J]. Ann Intern Med, 2026, 179(2): 270-275.

[16] Parasa S, Repici A, Berzin T, et al. Framework and metrics for the clinical use and implementation of artificial intelligence algorithms into en-

doscopy practice; recommendations from the American Society for Gastrointestinal Endoscopy Artificial Intelligence Task Force[J]. Gastrointest Endosc, 2023, 97(5): 815-824. e1.

(收稿日期:2026-04-12)

(本文编辑:李丹青)