

• 综述 •

# 从结构化分级到智能分级: 人工智能与 ACR-RADS 融合的研究进展

黄润楸<sup>1</sup>, 徐飞佳<sup>2</sup>, 刘育宏<sup>3</sup>, 冯钰淇<sup>4</sup>, 曾锦辉<sup>3</sup>, 梁定邦<sup>3</sup>, 周 淳<sup>4</sup>, 汪 洋<sup>1,2\*</sup>

<sup>1</sup>南方医科大学珠江医院 放射科, 广东 广州 510280; <sup>2</sup>同济大学附属第十人民医院 放射科, 上海 200072; <sup>3</sup>南方医科大学第二临床医院, <sup>4</sup>第一临床医学院, 广东 广州 510515

**[摘要]** 影像报告与数据系统(reporting and data system, RADS)是美国放射学会(American College of Radiology, ACR)旨在降低影像判读主观性、提升一致性的结构化风险分层框架。虽已广泛应用,但在真实世界实践中仍面临评分一致性受限、特征量化不足及跨中心可重复性差等方法学挑战。近年来,人工智能(artificial intelligence, AI)尤其是大语言模型(large language model, LLM)的发展为解决上述局限提供了新路径。文章系统梳理了主要 RADS 体系的结构特征及实践局限,综述了 AI 在病灶识别、风险再分层及流程规范化等方面的实证进展。现有证据显示, AI 更适宜作为 RADS 的“增强层”,以提升其客观性与重复性。文章重点探讨了 LLM 在语义理解、自动推理及质控中的潜力,并从监管视角展望了人机协作的演进方向。该趋势对推动我国影像诊断同质化、提升基层诊疗水平及构建智能化监管体系具有重要指导意义。

**[关键词]** 影像报告与数据系统; 人工智能; 结构化影像报告; 风险分层

**[中图分类号]** TP18

**[文献标志码]** A

**[文章编号]** 1007-4368(2026)07-1083-09

**doi:** 10.7655/NYDXBNSN251492

## From structured reporting to intelligent grading: research advances in the integration of artificial intelligence and ACR-RADS

HUANG Runqiu<sup>1</sup>, XU Feijia<sup>2</sup>, LIU Yuhong<sup>3</sup>, FENG Yuqi<sup>4</sup>, ZENG Jinhui<sup>3</sup>, LIANG Dingbang<sup>3</sup>, ZHOU Chun<sup>4</sup>, WANG Yang<sup>1,2\*</sup>

<sup>1</sup>Department of Radiology, Zhujiang Hospital, Southern Medical University, Guangdong 510280; <sup>2</sup>Department of Radiology, Tenth People's Hospital of Tongji University, Shanghai 200072; <sup>3</sup>Second School of Clinical Medicine, <sup>4</sup>First School of Clinical Medicine, Southern Medical University, Guangdong 510515, China

**[Abstract]** The reporting and data systems (RADS), established by the American College of Radiology (ACR), serve as structured risk-stratification frameworks designed to mitigate subjectivity and enhance consistency in radiological interpretation. Despite their widespread adoption, RADS still face methodological challenges in real-world practice, including limited inter-observer consistency, insufficient feature quantification, and suboptimal cross-center reproducibility. In recent years, the evolution of artificial intelligence (AI), particularly large language models (LLMs), has offered novel pathways to address these limitations. This paper systematically reviews the structural characteristics and practical constraints of major RADS frameworks and synthesizes empirical progress regarding AI in lesion identification, risk re-stratification, and workflow standardization. Current evidence suggests that AI is best positioned as an “augmentative layer” for RADS to bolster objectivity and reproducibility. Furthermore, this article explores the potential of LLMs in semantic understanding, automated reasoning, and quality control, while projecting the evolution of human-computer collaboration from a regulatory perspective. This trend holds significant implications for promoting the homogenization of diagnostic imaging, empowering primary healthcare services, and establishing intelligent regulatory systems in China.

**[Key words]** reporting and data system; artificial intelligence; structured imaging reporting; risk stratification

[J Nanjing Med Univ, 2026, 46(07): 1083-1091]

**[基金项目]** 广东省自然科学基金项目(2023A1515012499, 2025A1515010318); 广州市科技计划重点项目(2024B03J0781); 广东省医学装备学会科研基金项目(YZXH2025KT10); 广东省基础与应用基础研究基金(2024A1515220081)

\*通信作者(Corresponding author), E-mail: wangyang98289@smu.edu.cn (ORCID: 0000-0003-1588-3208)

传统医学影像学的重要任务之一,是将影像专家的主观经验性判断以稳定、可重复的方式表达为临床可理解且可执行的决策信息。作为该领域的权威发起者,美国放射学会(American College of Radiology, ACR)自乳腺影像标准化工作出发,逐步建立并推广影像报告与数据系统(reporting and data system, RADS)。然而,在真实世界的临床实践中,RADS常被简化为诊断分级工具,这种认知局限往往掩盖了该系统在标准化语境下的宏观意义。事实上,这一体系在降低跨学科沟通成本、提升多中心科研数据一致性方面具有不可替代的双重价值。以乳腺影像报告和数据系统(breast imaging reporting and data system, BI-RADS)为代表的体系通过统一术语与报告组织方式、明确判读要素并形成分级框架,使影像结论可与随访、活检及进一步检查等管理建议相衔接,从而在不同医师、不同机构及医患沟通之间形成高度一致的决策共识。ACR官方亦将RADS作为一组综合工具,涵盖术语词典、风险分层规则、模板与质量管理相关要素,服务于结构化报告与规范化决策链条的构建<sup>[1-4]</sup>。尽管如此,RADS在真实世界应用中仍面临可重复性与一致性方面的挑战。多项研究显示,在不同阅片者之间,部分RADS相关特征描述与分级结果的一致性可呈中等水平,且在筛查或随访场景中,这种分歧可能进一步影响管理建议的选择<sup>[5-8]</sup>。特别是在面对“灰区病例”,如BI-RADS 3与4A、肝脏影像报告和数据系统(liver imaging reporting and data system, LI-RADS 3)时,基于视觉经验的离散分级往往难以充分刻画病变的连续风险谱<sup>[9-12]</sup>。此外,跨中心应用还受到设备差异、采集协议与重建及后处理流程异质性的影响;影像定量指标在多中心研究中需要强调重复性与一致性,并依赖标准化协议与质量控制体系以降低变异,这已在定量影像生物标志物的共识框架中得到反复强调<sup>[8,13]</sup>。

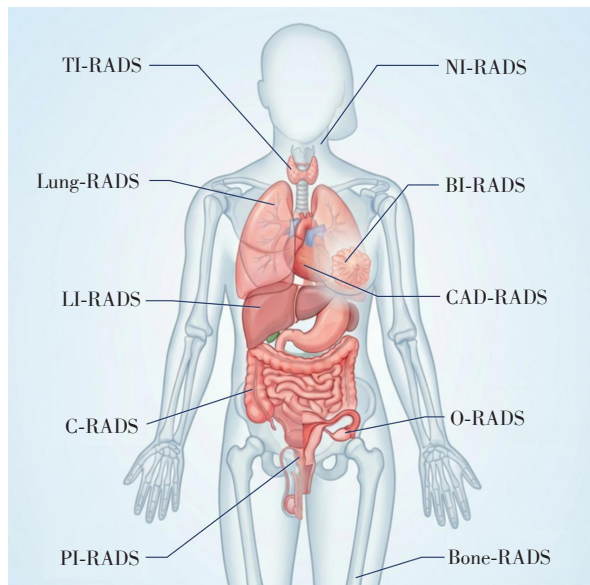
近年来,人工智能(artificial intelligence, AI)在医学影像领域快速发展,促进了病灶检出、分割与特征量化能力的提升,使影像信息表达逐步从定性及半定量描述向连续化风险评估延伸<sup>[14-16]</sup>。尤其值得关注的是,以Transformer架构为核心的大语言模型(large language model, LLM)正从语义解析与逻辑推理层面,为影像报告的质量控制、非结构化文本转化及自动化分级推理开辟了新路径<sup>[17-19]</sup>。然而,连续风险输出若缺乏统一的解释框架与管理接口,仍难以直接转化为可执行的临床路径。在此背景

下,RADS作为临床广泛接受且可解释的风险及管理框架,为AI结果嵌入结构化报告与临床决策链条提供了关键接口。虽然目前已有大量针对单一器官AI辅助RADS的应用研究,但仍缺乏从方法学演进共性、多模态融合以及大语言模型驱动的智能逻辑等宏观视角进行的系统梳理。在全球及我国大力推动影像AI规范化应用、力求通过数字化手段提升医疗同质化水平的背景下,探索AI与RADS的协同演进具有重要的理论价值与实践意义。

### 1 ACR-RADS的多模态影像分级体系:结构、共性与差异

在过去30余年的发展过程中,ACR围绕不同器官系统、疾病谱及临床应用场景,逐步构建并完善了一系列RADS。截至目前,ACR官方发布并持续更新的RADS体系已覆盖乳腺、肺、肝脏、前列腺、甲状腺、卵巢附件、头颈肿瘤治疗后随访、骨病变、结肠癌筛查及冠状动脉疾病等多个高频临床领域,形成了以风险分层与管理导向为核心的多模态结构化影像框架<sup>[20-31]</sup>(图1)。

尽管不同RADS体系在器官系统、成像技术及



Thyroid imaging reporting and data system (TI-RADS); neck imaging reporting and data system (NI-RADS); breast imaging reporting and data system (BI-RADS); liver imaging reporting and data system (LI-RADS); coronary artery disease - imaging reporting and data system (CAD-RADS); CT colonography reporting and data system (C-RADS); ovarian - adnexal reporting and data system (O-RADS); prostate imaging reporting and data system (PI-RADS).

图1 ACR影像学结构化报告与数据系统

Figure 1 ACR imaging reporting and data systems

疾病谱方面存在差异,其方法学框架仍呈现出高度一致的内在逻辑。各体系均以标准化影像术语和明确的判读要素为基础,通过规范特征定义、描述顺序与评估规则,降低自由文本报告带来的主观差异。同时,RADS普遍采用序数化风险分级,将复杂或连续的影像表现转化为临床可理解、可沟通的风险区间,而非简单的二分类判断(表1)。更重要的是,各分级通常直接对应明确的临床管理建议,使影像结论直接指向随访、进一步检查或治疗决策,从而将影像报告嵌入临床决策流程,而非作为孤立的信息输出<sup>[32-35]</sup>。在上述共性框架下,不同RADS体系的差异主要体现在3个方面。其一,风险定义的锚点不同:BI-RADS、PI-RADS与O-RADS以恶性肿瘤发生概率为核心;CAD-RADS综合解剖狭窄、斑块负荷及缺血信息,兼顾诊断与预后;LI-RADS则面向特定高危人群,强调诊断特异性。其二,成像模态与关键判据存在差异:部分体系(如TI-RADS、Bone-RADS)以形态学特征为主,而另一些体系(如PI-RADS、O-RADS MRI与LI-RADS)更依赖DWI、DCE、CEUS等功能成像参数。其三,临床应用阶段不同:Lung-RADS与C-RADS主要用于筛查,侧重风险分层与检查策略管理;NI-RADS与LR-TR则用于治疗后随访,重点在于区分复发与治疗相关改变<sup>[20-29,36-38]</sup>(表1)。因此,ACR-RADS并非单纯的诊断分级工具,而是在统一方法学原则下,围绕具

体临床决策需求构建的结构化评估与管理支持体系。正是这种共性框架与场景特异性并存的结构特征,使其适配高度异质化的临床应用需求。

2 ACR-RADS在真实世界应用中的局限

尽管ACR提出的各类RADS体系在多个疾病领域推动了结构化报告与风险分层的规范化实践,但在真实世界临床环境中,其以规则与语义描述为核心的设计范式仍存在若干共性局限,主要体现在评分一致性、跨中心可重复性、特征量化能力、复杂临床场景适配性以及分级结果与临床管理之间的闭环落实等方面。

首先,RADS的分级结果在实际应用中仍不可避免地受到阅片者经验的影响,观察者间一致性存在客观差异。尽管各体系均建立了标准化术语与分级规则,但多项关键影像特征仍需由医师将影像表现映射为离散等级,尤其在中间等级或边界病例中,不同阅片者对形态边界、强化模式或弥散受限程度的理解差异,可能导致分级不一致,进而影响随访或干预策略的选择。相关研究亦表明,在某些具体任务(如密度分级、终末评估类别或关键征象判读)上,观察者变异仍然存在<sup>[12,39-43]</sup>。其次,RADS在跨中心应用中的可重复性在很大程度上依赖成像设备及扫描与重建协议的一致性。以定量影像生物标志物联盟(quantitative imaging biomarkers al-

表1 ACR影像报告与数据系统概览:靶向疾病、评分框架与成像模态

Table 1 Overview of ACR imaging reporting and data systems: target conditions, scoring frameworks, and imaging modalities

| RADS      | Pathologic condition                      | Score or category                                   | Score assignment          | Modality                  |
|-----------|-------------------------------------------|-----------------------------------------------------|---------------------------|---------------------------|
| BI-RADS   | Breast cancer                             | 0-6(4A-4C)                                          | Patient based             | Mammography, MRI, US, DBT |
| C-RADS    | Colon cancer                              | Colonic findings(C0-4); extracolonic findings(E0-4) | Lesion based              | CT colonography           |
| CAD-RADS  | Coronary artery disease                   | 0-5, N                                              | Patient based             | CT angiography            |
| LI-RADS   | Liver cancer(HCC)                         | 1-5, TIV, NC, M, TR                                 | CT/MR/CEUS: lesion based; | CT, MRI, CEUS, US         |
| Lung-RADS | Lung cancer(screening)                    |                                                     | US: patient based         |                           |
|           |                                           | 0-4(4A、4B、4X), S, C                                 | Patient based             | Low-dose CT               |
| NI-RADS   | Head and neck cancers (diagnostic system) | 0-4(2A、2B)                                          | Patient based             | PET, CT, MRI              |
| O-RADS    | Ovarian-adnexal masses                    | NA                                                  | Lesion based              | US, MRI                   |
| PI-RADS   | Prostate cancer                           | 1-5                                                 | Lesion based              | Multiparametric MRI       |
| TI-RADS   | Thyroid cancer(incidental lesions)        | 1-5                                                 | Lesion based              | US                        |
| Bone-RADS | Neoplastic bone lesions(potential)        | 0-4                                                 | Lesion based              | X-ray, CT, MRI, PET       |

DBT: digital breast tomosynthesis; CEUS: contrast material-enhanced ultrasound; HCC: hepatocellular carcinoma; NA: not applicable; N: non-diagnostic/inadequate study; S: clinically significant or potentially clinically significant findings (non-lung cancer); TIV: tumor in vein; TR: treated observation; M: malignant but not HCC.



liance, QIBA)为代表的共识框架指出,定量影像指标的重复性与一致性有赖于标准化采集、严格质量控制及一致的数据处理流程;在缺乏统一质控与计量学体系的情况下,多中心间的系统性变异将被放大,从而限制RADS在真实世界研究与临床推广中的稳定应用<sup>[13,44]</sup>。此外,现有RADS体系中,仍有相当一部分关键判读要素停留在描述性或半定量层面,尚未形成可机器读取、可跨平台稳定复现的明确量化阈值或连续指标。这类判据在不同设备、不同中心之间的一致性受到限制,不仅影响RADS体系的可重复应用,也客观上增加了其与自动化分析方法及数据驱动模型进行深度融合的难度<sup>[45-47]</sup>。第四,在复杂或非典型临床情境中,RADS分级体系在刻画部分病例的风险状态时存在不确定性区间。在多病灶并存、术后或放疗后改变,以及炎症与肿瘤影像表现交织等情况下,既有分级规则往往难以充分刻画病变的连续风险谱。以LI-RADS为例,LI-RADS 3与LI-RADS 4等不确定类别在临床实践中较为常见,其后续进展风险具有显著异质性<sup>[9-12]</sup>。最后,RADS分级与临床管理建议之间的闭环在真实世界实践中并不总能得到充分落实。尽管各体系通常配套明确的随访或处理路径,其实际执行仍可能受到医疗资源配置、跨科室协作流程、信息系统追踪能力及患者依从性等多重因素的制约,从而导致影像分层结果与实际诊疗行为之间出现偏离<sup>[48-49]</sup>。

### 3 图像识别和分割驱动的AI-RADS融合:从规则分级到可计算风险的量化证据

RADS分级体系的主要挑战并非来自分级框架本身,而在于连续影像证据被离散化过程中引入的不稳定性与可重复性不足。近5年来,深度学习为影像特征的客观量化与稳定映射提供了可计算路径,使AI成为推动RADS向更加客观与可重复方向演进的重要支撑。从机制上看,RADS通过分级实现临床决策衔接,而深度学习能够学习稳定表征并输出连续风险估计,二者在信息处理逻辑上的契合,为AI以受控方式嵌入而非替代RADS框架提供了依据。

在乳腺影像领域,AI与BI-RADS的融合是目前证据相对成熟的代表性应用场景之一。AI以增强层形式嵌入既有BI-RADS workflow,首先对病灶区域进行自动定位或分割,并在此基础上结合病灶及其周围组织的影像特征,形成连续风险评分与分级参考。阅片医师结合形态学特征作出最终裁决。其主要作用集中于阈值附近(BI-RADS 3/4A)病例,通

过提供一致的量化参照,改善灰区病例的分流稳定性。需要强调的是,相关研究的性能表现与具体任务设定密切相关:在良恶性二分类或特定风险评分任务中,深度学习模型通常可获得较高的判别性能;而在完整的BI-RADS多等级分层任务中,不同研究之间仍存在一定差异。针对临床分歧较集中的中间分层任务,多中心研究显示,在区分BI-RADS 3/4A病灶时,深度学习模型的曲线下面积(aneau under the curver, AUC)约为0.74。对于初级阅片者,AI辅助可将其AUC由0.55~0.58提升至0.74~0.77,并在保持较高敏感度的同时,使特异度提高约10%~20%,而对资深阅片者的增益相对有限<sup>[50]</sup>。同样,在前列腺多参数MRI场景中,AI首先对前列腺及可疑病灶进行自动定位与勾画,在此基础上结合多参数MRI信息,对病灶进行定量表征并形成患者级临床显著癌症风险评估。该风险评估结果并不直接替代PI-RADS分级,而是作为辅助依据嵌入既有流程,用于提示是否需要双读或进一步复核,最终分级仍由医师完成,从而在不改变分级框架的前提下,提高分流决策的一致性与效率。大规模外部验证研究报告指出,该自动化系统在识别临床显著前列腺癌方面达到患者级AUC约0.93(95% CI: 0.91~0.95);在多读者研究中,AI辅助可将总体诊断准确率由80%提升至86%。此外,在真实世界的前瞻性实施中,AI辅助流程与约32%的放射科工作量下降相关,提示其在提升诊断一致性的同时,亦具有潜在的流程优化价值<sup>[51-52]</sup>。类似的量化证据亦见于其他RADS体系。在卵巢附件影像中,基于超声的深度学习模型已被用于良恶性风险评估,并在一定程度上与O-RADS风险分层形成互补,不同研究报道的AUC存在差异,但部分研究或综述显示其判别性能可达约0.90以上<sup>[53]</sup>。在肝脏影像中,针对LI-RADS灰区病灶,已有研究基于多期动态增强MRI的时相信息构建定量评估流程,通过对不同增强时相进行配准与联合表征,在病灶及其周围区域提取时间-信号变化特征,并据此形成LR-3与LR-4/5的风险判别依据,其在边界病例中的区分性能AUC约为0.95。与前述乳腺及前列腺场景主要基于单次或有限序列形态学表征不同,该类方法的特点在于引入时间维度的动态增强信息,以补充灰区病灶在血流动力学层面的判别依据<sup>[54]</sup>。在以筛查与随访为主的Lung-RADS场景中,相关研究主要聚焦于结节的自动检出与漏诊控制。既往研究在公开数据集上总结,结节自动检测流程的灵敏度为86%~

98%,并提示,其作为第二读者应用时,可在保持可接受特异度的前提下,改善结节检出表现与流程一致性。从方法流程上看,该类系统通常先对全肺区域进行自动定位与候选区域生成,随后通过逐级筛选与假阳性抑制,输出需关注的结节提示列表,由医师在此基础上完成最终确认与随访分层。与上述分级增强不同,Lung-RADS场景中的核心作用在于前端检出与分流筛查,即通过降低漏检率来稳定随访入口,而非直接参与病灶分级本身<sup>[55]</sup>。对于Bone-RADS所涉及的偶发骨病灶评估,目前相关证据主要来源于骨转移检出等任务的AI研究与系统综述,其结果提示AI可能在一定程度上提升检出效率与读者表现,但不同模态与任务间性能异质性较大,仍需更多多中心外部验证,并与分级框架进行严格对齐<sup>[56]</sup>。

综合现有证据可见,在多个代表性RADS应用场景中,AI在特定任务上已展现出稳定且良好的AUC表现,并在灵敏度、特异度、阅片一致性或流程效率等方面带来可检测的改进。这表明,AI与RADS的融合正由概念验证阶段迈入具有实证基础的早期应用阶段,为进一步系统评估其泛化能力、可解释性及监管路径奠定了基础。

#### 4 LLM在RADS体系中的方法学角色与应用边界

RADS体系的关键并不止于影像要素提取,而在于如何将这些发现转化为符合规范的语义表述并对接分级规则。既往“影像发现-报告文本-RADS分级”之间缺乏可计算的语义接口。基于Transformer架构的LLM通过自注意力机制实现对长文本语境的整体建模,可将非结构化影像描述解析为标准化判读要素,并在既有RADS框架内完成分级推理,更符合RADS对综合征象判读的设计逻辑。相较传统规则方法或浅层自然语言处理模型,Transformer在复杂语义建模方面具有明显优势,为影像报告的结构化表达与分级推理提供了可靠支撑<sup>[17-18,57]</sup>。

在此基础上,Transformer与RADS的结合可理解作为一种“语义处理与规则约束相结合”的工作方式:模型将非结构化影像文本转化为可计算的语义表示,而RADS提供明确的分级与管理框架,对模型输出进行规范化限定,使Transformer以语义层“计算工具”的形式服务于既有体系,而非替代专家判断。现有研究显示,该模式具有应用前景明确且风险边界相对清晰的特征。例如,在基于结构化输入的乳腺超声研究中,ChatGPT-4的BI-RADS分级准

确率达96.87%;在良恶性预测任务中,其AUC为0.82(95% CI:0.79~0.85),总体准确率80.63%,灵敏度90.56%,特异度73.51%<sup>[58]</sup>。这些结果提示,在输入要素明确、语义范围受控的条件下,基于Transformer的语言模型可获得稳定且可量化的性能表现。然而,在更贴近真实世界的跨语言、多中心验证研究中,直接基于乳腺影像自由文本报告进行BI-RADS赋值时,GPT-4与人类专家之间的一致性仅为中等水平(Gwet AC1=0.52)。其中,18.1%的分级差异可能引发临床管理路径改变,另有10.6%可能对管理决策产生不利影响<sup>[4]</sup>。上述结果提示,在缺乏明确结构化输入与外部规则约束的条件下,Transformer的语义推断不确定性仍可能在分级与管理层面被放大。相比之下,Transformer在RADS相关流程中的辅助性环节,尤其是在报告质量控制与差错拦截方面,显示出更具可行性的安全性与效率优势。一项影像报告差错检测研究显示,基于Transformer的GPT-4对人为植入错误的检出率为82.7%(124/150),与不同资历放射科医师的平均水平相当;其单份报告处理时间为(3.5±0.5)s,显著快于最快人类读者的(25.1±20.1)s,且单份报告纠错成本约为0.03美元,明显低于成本效率最高的放射科医师(约0.42美元)<sup>[59]</sup>。此类任务侧重语义一致性与逻辑完整性的校验,与Transformer在全局语义建模与模式识别方面的优势高度契合。因此,从近期应用角度看,Transformer驱动 RADS 更适合作为结构化信息抽取、分级一致性校正及差错控制模块嵌入既有工作流程,在不改变分级框架的前提下提升报告的可重复性与执行一致性。中长期发展应侧重于将Transformer与指南版本、机构本地流程及可追溯证据体系进行系统对齐,通过指南检索增强、字段级校验及分级过程的可审计化设计,将潜在的一致性不足与管理漂移风险转化为可监管、可验证且可持续迭代的临床级工具。

#### 5 RADS体系的演进原则、证据边界与AI的辅助作用

在当前由ACR推动并持续维护的体系中,RADS主要覆盖若干高频临床领域,其核心目标在于通过标准化术语、结构化报告与分级评估,提升医患及跨学科沟通的一致性,并将影像结论与后续管理决策有效衔接。但不同体系及具体条目之间的证据层级与稳定性仍存在差异,难以将其简单概括为“均已被充分证实”<sup>[20-30]</sup>。从这一角度看,尚未



被ACR正式纳入RADS体系的成像方式,并不意味着其缺乏潜在价值,而是提示截至目前,其在真实世界应用中的稳定性、可重复性或特异性尚不足以支撑跨中心一致的标准化分级与管理决策。这种相对审慎的纳入策略,有助于避免在日常临床实践中对证据等级尚不充分的成像手段赋予过高的决策权重。随着免疫治疗、靶向治疗及局部消融等策略的广泛应用,传统以“是否恶性”为核心的分级逻辑正逐步延伸至治疗反应、复发风险及疾病动态演变评估,这一趋势已在部分体系中有体现,例如围绕治疗后评估与随访管理构建的相关模块<sup>[20, 26]</sup>。另一方面,随着多模态影像、功能成像及影像组学证据的不断积累,部分疾病领域可能逐步形成具有RADS特征的结构化评估框架,但其是否最终被ACR正式吸纳,仍取决于相关证据能否达到与现有RADS相当的循证强度与跨中心一致性水平。

在这一过程中,AI主要体现为促进现有体系演进的技术力量。通过对大规模真实世界数据的持续分析,AI有望评估不同影像征象与成像策略在多中心、多设备条件下的判别稳定性,从而为是否建立新的RADS子体系或优化既有判据条目提供量化依据<sup>[60]</sup>。在明确的验证流程、版本治理与持续监测机制支持下,AI在理论上可缩短证据积累与共识更新之间的时间滞后,使部分条目更接近证据驱动的渐进式优化;但这一过程仍依赖制度与监管层面的协同,难以由算法本身独立完成<sup>[61-62]</sup>。至于通用人工智能(artificial general intelligence, AGI)的潜在影响,更适合作为概念层面的讨论。即便模型能够整合影像、临床信息与最新证据形成逻辑自洽的风险评估,医学决策仍不可避免地涉及价值判断、风险权衡、患者偏好及伦理责任等难以完全形式化的因素。因此,人类专家对关键决策的最终审查与责任承担,仍是临床决策支持系统不可替代的核心环节。国内相关研究亦普遍指出,临床AI的安全、有效与可持续应用,有赖于覆盖研发、部署与使用全流程的治理框架与持续监测机制,并需在制度层面明确人类监督与责任边界,以确保技术应用始终处于可控与可追责的范畴之内<sup>[60-66]</sup>。

## 6 总结与展望

RADS体系的未来演进更可能建立在证据充分与应用可控的基础之上,通过AI技术辅助与人机协同决策,实现有序拓展与持续优化。在这一过程中,AI技术,包括AGI相关方法,或将在影像特征量

化、证据整合及流程一致性提升等方面不断拓展RADS的技术边界与应用深度。然而,在可预见阶段,其核心功能仍应定位为增强既有分级与管理框架,而非替代临床判断或责任主体。唯有在可解释性、可审计性与临床安全性得到充分保障的前提下,将AI以受控方式嵌入RADS体系,才能使其在真实世界临床环境中持续发挥规范化风险分层与管理支持的价值。

### [参考文献]

- [1] YADAV A, SINGH A, MADAAN S, et al. International academy of cytology Yokohama system for reporting breast cytology and the ACR breast imaging reporting and data system (BIRADS): are they concordant? [J]. Iran J Pathol, 2024, 19(4): 400-407
- [2] SPAK D A, PLAXCO J S, SANTIAGO L, et al. BI-RADS® fifth edition: a summary of changes [J]. Diagn Interv Imaging, 2017, 98(3): 179-190
- [3] MERCADO C L. BI-RADS update [J]. Radiol Clin North Am, 2014, 52(3): 481-487
- [4] COZZI A, PINKER K, HIDBER A, et al. BI-RADS category assignments by GPT-3.5, GPT-4, and google bard: a multilanguage study [J]. Radiology, 2024, 311(1): e232133
- [5] YOON S H, KIM Y J, DOH K, et al. Interobserver variability in Lung CT Screening Reporting and Data System categorisation in subsolid nodule-enriched lung cancer screening CTs [J]. Eur Radiol, 2021, 31(9): 7184-7191
- [6] KIRI L, ABDOLELL M, COSTA A F, et al. US LI-RADS visualization score: interobserver variability and association with cause of liver disease, sex, and body mass index [J]. J L' association Can Des Radiol, 2022, 73(1): 68-74
- [7] LEE A Y, WISNER D J, AMINOLOLAMA-SHAKERI S, et al. Inter-reader variability in the use of BI-RADS descriptors for suspicious findings on diagnostic mammography: a multi-institution study of 10 academic radiologists [J]. Acad Radiol, 2017, 24(1): 60-66
- [8] 姜 艳, 项霞青, 周 婉. 美国放射学会甲状腺影像报告和数据系统分类在甲状腺结节中的观察者一致性研究 [J]. 医疗装备, 2022, 35(21): 114-117  
JIANG Y, XIANG L Q, ZHOU W. Observer consistency of American college of radiology thyroid imaging Reporting and data system classification in thyroid nodules [J]. Medical Equipment, 2022, 35(21): 114-117
- [9] SEVIM S, DICLE O, GEZER N S, et al. How high is the inter-observer reproducibility in the LIRADS reporting system? [J]. Pol J Radiol, 2019, 84: e464-e469

- [10] ABDEL RAZEK A A K, EL-SEROUGY L G, SALEH G A, et al. Reproducibility of LI-RADS treatment response algorithm for hepatocellular carcinoma after locoregional therapy [J]. *Diagn Interv Imaging*, 2020, 101 (9): 547-553
- [11] MULÉ S, RNOT M, GHOSN M, et al. Automated CT LI-RADS v2018 scoring of liver observations using machine learning: a multivendor, multicentre retrospective study [J]. *JHEP Rep*, 2023, 5 (10): 100857
- [12] GALGANO S J, SMITH E N. LI-RADS version 2018 for MRI and CT: interreader agreement in real-world practice [J]. *Radiology*, 2023, 307 (5): e231212
- [13] SHUKLA-DAVE A, OBUCHOWSKI N A, CHENEVERT T L, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials [J]. *J Magn Reson Imaging*, 2019, 49 (7): e101-e121
- [14] WANG C D, SHAO J, HE Y C, et al. Data-driven risk stratification and precision management of pulmonary nodules detected on chest computed tomography [J]. *Nat Med*, 2024, 30 (11): 3184-3195
- [15] YANG D D, LEE L K, TSUI J M G, et al. AI-derived tumor volume from multiparametric MRI and outcomes in localized prostate cancer [J]. *Radiology*, 2024, 313 (1): e240041
- [16] 陈敏. 智能赋影, 融合创新: 人工智能时代下医学影像学科的发展与展望 [J]. *中国医学影像学杂志*, 2026, 34 (1): 1-2  
CHEN M. Intelligent imaging empowerment and integrative innovation: development and prospects of the medical imaging discipline in the era of artificial intelligence [J]. *Chinese Journal of Medical Imaging*, 2026, 34 (1): 1-2
- [17] YAN B, PEI M T, ZHAO M, et al. Prior guided transformer for accurate radiology reports generation [J]. *IEEE J Biomed Health Inform*, 2022, 26 (11): 5631-5640
- [18] BHAYANA R. Chatbots and large language models in radiology: a practical primer for clinical and research applications [J]. *Radiology*, 2024, 310 (1): e232756
- [19] 刘再毅, 赵致禾, 石镇维. 医学影像大模型: 肿瘤精准诊疗的范式革新 [J]. *协和医学杂志*, 2025, 16 (4): 805-811  
LIU Z Y, ZHAO Z H, SHI Z W. Large medical imaging models: a paradigm innovation for precision diagnosis and treatment of tumors [J]. *Medical Journal of Peking Union Medical College Hospital*, 2025, 16 (4): 805-811
- [20] AIKEN A H, RATH T J, ANZAI Y, et al. ACR neck imaging reporting and data systems (NI-RADS): a white paper of the ACR NI-RADS committee [J]. *J Am Coll Radiol*, 2018, 15 (8): 1097-1108
- [21] REINHOLD C, ROCKALL A, SADOWSKI E A, et al. Ovarian-adnexal reporting lexicon for MRI: a white paper of the ACR ovarian-adnexal reporting and data systems MRI committee [J]. *J Am Coll Radiol*, 2021, 18 (5): 713-729
- [22] HOANG J K, MIDDLETON W D, TESSLER F N. Update on ACR TI-RADS: successes, challenges, and future directions, from the AJR special series on radiology reporting and data systems [J]. *AJR Am J Roentgenol*, 2021, 216 (3): 570-578
- [23] TESSLER F N, MIDDLETON W D, GRANT E G, et al. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee [J]. *J Am Coll Radiol*, 2017, 14 (5): 587-595
- [24] WEINREB J C, BARENTSZ J O, CHOYKE P L, et al. PI-RADS prostate imaging - reporting and data system: 2015, version 2 [J]. *Eur Urol*, 2016, 69 (1): 16-40
- [25] AGRAWAL R, AHUJA J, STRANGE C D, et al. Lung-RADS v2022 update [J]. *Radiol Clin North Am*, 2025, 63 (4): 507-516
- [26] SCHIMA W, KOPF H, EISENHUBER E. LI-RADS made easy [J]. *Rofo*, 2023, 195 (6): 486-494
- [27] TESSLER F N, MIDDLETON W D, GRANT E G. Thyroid imaging reporting and data system (TI-RADS): a user's guide [J]. *Radiology*, 2018, 287 (1): 29-36
- [28] GHASEMI A, AHLAWAT S. Bone reporting and data system (bone-RADS) and other proposed practice guidelines for reporting bone tumors [J]. *Rofo*, 2024, 196 (11): 1134-1142
- [29] CURY R C, LEIPSIC J, ABBARA S, et al. CAD-RADS™ 2.0-2022 coronary artery disease-reporting and data system an expert consensus document of the society of cardiovascular computed tomography (SCCT), the American college of cardiology (ACC), the American college of radiology (ACR) and the North America society of cardiovascular imaging (NASCI) [J]. *Radiol Cardiothorac Imag*, 2022, 4 (5): e220183
- [30] YEE J, DACHMAN A, KIM D H, et al. CT colonography reporting and data system (C-RADS): version 2023 update [J]. *Radiology*, 2024, 310 (1): e232007
- [31] 乔敏, 冯尚勇, 沈德娟, 等. 中国甲状腺影像报告和数据分析系统对甲状腺结节良恶性的鉴别诊断 [J]. *中国医学影像学杂志*, 2021, 29 (11): 1070-1075  
QIAO M, FENG S Y, SHEN D J, et al. Differential diagnosis of benign and malignant thyroid nodules using the Chinese Thyroid Imaging Reporting and Data System [J]. *Chinese Journal of Medical Imaging*, 2021, 29 (11): 1070-1075
- [32] KAHN C E Jr, LANGLOTZ C P, BURNSIDE E S, et al.

- Toward best practices in radiology reporting[J]. *Radiology*, 2009, 252(3): 852-856
- [33] KNIGHTLY D V, PLACE J N. Best practices in computerized tomography[J]. *Radiol Manage*, 2010, 32(1): 16-23
- [34] GRANATA V, DE MUZIO F, CUTOLO C, et al. Structured reporting in radiological settings: pitfalls and perspectives[J]. *J Pers Med*, 2022, 12(8): 1344
- [35] 张紫欣, 吕志彬, 王宇新, 等. LI-RADS 2018 版 MRI 辅助征象在肝癌分类及预后评估中的应用[J]. *临床放射学杂志*, 2024, 43(5): 858-861
- ZHANG Z X, LV Z B, WANG Y X, et al. Application of LI-RADS 2018 MRI ancillary features in hepatocellular carcinoma classification and prognostic evaluation[J]. *Journal of Clinical Radiology*, 2024, 43(5): 858-861
- [36] AN J Y, UNSDORFER K M L, WEINREB J C. BI-RADS, C-RADS, CAD-RADS, LI-RADS, lung-RADS, NI-RADS, O-RADS, PI-RADS, TI-RADS: reporting and data systems[J]. *RadioGraphics*, 2019, 39(5): 1435-1436
- [37] 唐雅伦, 李 瑞, 高 磊, 等. 人工智能辅助诊断系统与 Lung-RADS 对不同临床特征肺结节的良恶性预测效能[J]. *分子影像学杂志*, 2025, 48(6): 668-677
- TANG Y L, LI R, GAO L, et al. Predictive performance of an artificial intelligence - assisted diagnostic system and Lung - RADS for malignancy of pulmonary nodules with different clinical characteristics[J]. *Journal of Molecular Imaging*, 2025, 48(6): 668-677
- [38] 刘丹丹, 韦 平, 高 彦, 等. 联合 Kaiser 评分、DWI 及 MRS 对乳腺 BI-RADS3-5 类病变的诊断效能[J]. *临床放射学杂志*, 2025, 44(12): 2290-2294
- LIU D, WEI P, GAO Y, et al. Diagnostic performance of combined Kaiser Score, diffusion-weighted imaging, and magnetic resonance spectroscopy for breast BI-RADS category 3-5 lesions[J]. *Journal of Clinical Radiology*, 2025, 44(12): 2290-2294
- [39] VAN RIEL S J, JACOBS C, SCHOLTEN E T, et al. Observer variability for Lung - RADS categorisation of lung cancer screening CTs: impact on patient management[J]. *Eur Radiol*, 2019, 29(2): 924-931
- [40] BERG W A, CAMPASSI C, LANGENBERG P, et al. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment[J]. *AJR Am J Roentgenol*, 2000, 174(6): 1769-1777
- [41] ALIKHASSI A, ESMAILI GOURABI H, BAIKPOUR M. Comparison of inter - and intra - observer variability of breast density assessments using the fourth and fifth editions of Breast Imaging Reporting and Data System[J]. *Eur J Radiol Open*, 2018, 5: 67-72
- [42] BECKER A S, BARTH B K, MARQUEZ P H, et al. Increased interreader agreement in diagnosis of hepatocellular carcinoma using an adapted LI-RADS algorithm[J]. *Eur J Radiol*, 2017, 86: 33-40
- [43] CHOI H H, KIM S, SHUM D J, et al. Assessing adherence to US LI-RADS follow-up recommendations in vulnerable patients undergoing hepatocellular carcinoma surveillance[J]. *Radiol Imaging Cancer*, 2024, 6(1): e230118
- [44] BOSS M A, MALYARENKO D, PARTRIDGE S, et al. The QIBA profile for diffusion-weighted MRI: apparent diffusion coefficient as a quantitative imaging biomarker[J]. *Radiology*, 2024, 313: e233055
- [45] RAUNIG D L, MCSHANE L M, PENNELLO G, et al. Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment[J]. *Stat Methods Med Res*, 2015, 24(1): 27-67
- [46] HAGIWARA A, FUJITA S, OHNO Y, et al. Variability and standardization of quantitative imaging: monoparametric to multiparametric quantification, radiomics, and artificial intelligence[J]. *Invest Radiol*, 2020, 55(9): 601-616
- [47] 程 枫, 艾慧俊, 冯娅琴, 等. 描述性标准化超声报告对甲状腺乳头状癌诊断的影响[J]. *现代实用医学*, 2018, 30(11): 1507-1509
- CHENG F, AI H J, FENG Y Q, et al. Impact of descriptive standardized ultrasound reporting on the diagnosis of papillary thyroid carcinoma[J]. *Modern Practical Medicine*, 2018, 30(11): 1507-1509
- [48] LIN Y N, FU M Z, DING R W, et al. Patient adherence to lung CT screening reporting & data system-recommended screening intervals in the United States: a systematic review and meta-analysis[J]. *J Thorac Oncol*, 2022, 17(1): 38-55
- [49] LACSON R, WANG A J, COCHON L, et al. Factors associated with optimal follow-up in women with BI-RADS 3 breast findings[J]. *J Am Coll Radiol*, 2020, 17(4): 469-474
- [50] LIN X H, LIAO T T, YANG Y T, et al. Value of deep learning model for predicting breast imaging reporting and data system 3 and 4A lesions on mammography[J]. *Quant Imaging Med Surg*, 2025, 15(5): 4047-4058
- [51] WU H C, LIU F, YANG Q S, et al. Automated MRI system for clinically significant prostate cancer detection development validation and real-world implementation[J]. *Nat Commun*, 2025, 16(1): 11583
- [52] XING Z Y, CHEN J, PAN L, et al. Enhanced detection of prostate cancer lesions on biparametric MRI using artificial intelligence: a multicenter, fully-crossed, multi-reader multi-case trial[J]. *Acad Radiol*, 2025, 32(10): 5954-5963
- [53] CHEN H, YANG B W, QIAN L, et al. Deep learning pre-



- diction of ovarian malignancy at US compared with O-RADS and expert assessment[J]. *Radiology*, 2022, 304(1): 106-113
- [54] WU Y N, WHITE G M, CORNELIUS T, et al. Deep learning LI-RADS grading system based on contrast enhanced multiphase MRI for differentiation between LR-3 and LR-4/LR-5 liver tumors[J]. *Ann Transl Med*, 2020, 8(11): 701
- [55] HENDRIX W, HENDRIX N, SCHOLTEN E T, et al. Deep learning for the detection of benign and malignant pulmonary nodules in non-screening chest CT scans[J]. *Commun Med*, 2023, 3(1): 156
- [56] DONG X Y, CHEN G L, ZHU Y P, et al. Artificial intelligence in skeletal metastasis imaging[J]. *Comput Struct Biotechnol J*, 2024, 23: 157-164
- [57] HUANG R Q, MENG X L, ZHANG X X, et al. Artificial intelligence - driven change redefining radiology through interdisciplinary innovation[J]. *Interdiscip Med*, 2025, 3(1): e20240063
- [58] SU M J, LIANG X, ZENG X T, et al. Using a large language model for breast imaging reporting and data system classification and malignancy prediction to enhance breast ultrasound diagnosis: retrospective study[J]. *JMIR Med Inform*, 2025, 13: e70924
- [59] GERTZ R J, DRATSCH T, BUNCK A C, et al. Potential of GPT-4 for detecting errors in radiology reports: implications for reporting accuracy[J]. *Radiology*, 2024, 311(1): e232714
- [60] WELLS B J, NGUYEN H M, MCWILLIAMS A, et al. A practical framework for appropriate implementation and review of artificial intelligence(FAIR-AD) in healthcare[J]. *NPJ Digit Med*, 2025, 8(1): 514
- [61] YOU J G, HERNANDEZ-BOUSSARD T, PFEFFER M A, et al. Clinical trials informed framework for real world clinical implementation and deployment of artificial intelligence applications[J]. *NPJ Digit Med*, 2025, 8(1): 107
- [62] 唐蛰雨,李绍钦,贾中芝. 基于CT图像的深度学习在主动脉夹层中的应用进展[J]. *南京医科大学学报(自然科学版)*, 2024, 44(8): 1174-1178
- TANG H Y, LI S M, JIA Z Z. Application progress of deep learning based on CT images in aortic dissection[J]. *Journal of Nanjing Medical University(Natural Sciences)*, 2024, 44(8): 1174-1178
- [63] AL KUWAITI A, NAZER K, AL-REEDY A, et al. A review of the role of artificial intelligence in healthcare[J]. *J Pers Med*, 2023, 13(6): 951
- [64] YANG Y, CUI Y U, WANG Y T, et al. Interpretation of the WHO's "ethics and governance of artificial intelligence for health: guidance on large multi-modal models" and its implications for China[J]. *Chin J Prev Med*, 2025, 59(6): 960-969
- [65] 汪 洋,马 玉,刘 颖. 人工智能背景下医疗器械创新监管研究与启示[J]. *中国设备工程*, 2025, (23): 272-274
- WANG Y, MA Y, LIU Y. Research and implications of innovative regulation of medical devices in the context of artificial intelligence [J]. *China Plant Engineering*, 2025, (23): 272-274
- [66] 许琳琳,黄德群. 人工智能应用于医疗领域的法律法规及技术挑战[J]. *临床医学工程*, 2025, 32(11): 1177-1182
- XU L L, HUANG D Q. Legal regulations and technical challenges of artificial intelligence applications in the medical field[J]. *Clinical Medical Engineering*, 2025, 32(11): 1177-1182
- (收稿: 2025-12-30; 修回: 2026-01-17; 录用: 2026-03-02)
- (本文编辑:戴玉娟)