

• 专家面对面 •

## 人工智能驱动下的检验医学创新探索与实践

李 波<sup>1</sup>,袁 旭<sup>2#</sup>,李小强<sup>3#</sup>,牛长春<sup>4#</sup>,武永康<sup>5,6#</sup>,胡庆丰<sup>7#</sup>,杨大干<sup>8#</sup>

1. 重庆大学附属三峡医院检验科,重庆 404000; 2. 华中科技大学同济医学院附属同济医院,湖北武汉 430000;  
 3. 重庆医科大学附属儿童医院检验科,重庆 400014; 4. 重庆市人民医院检验科/重庆市临床检验中心,重庆  
 401147; 5. 四川大学华西医院实验医学科,四川成都 610000; 6. 四川金堂县第一人民医院,四川成都 610402;  
 7. 浙江省人民医院检验中心,浙江杭州 310014; 8. 浙江大学医学院附属第一医院检验科,浙江杭州 310003

**关键词:**人工智能; 大语言模型; DeepSeek; 检验医学

**DOI:**10.3969/j.issn.1673-4130.2025.17.002

**文章编号:**1673-4130(2025)17-2056-06

**中图法分类号:** R-1

**文献标志码:**A



**李波:**2025 年,医疗领域正经历着双重变革——一方面,按疾病诊断相关分组(DRGs)/按病种分值分组付费(DIP)医保支付改革已进入深化阶段,促使检验医学向精细化、标准化、高效化转型;另一方面,以大语言模型(LLM)为代表的人工智能技术加速渗透,为检验流程优化、结果互认、临床沟通等核心环节注入了新动能。在这一背景下,检验医学既面临“如何在控费与提质间平衡”的挑战,也迎来“以人工智能赋能学科升级”的机遇。

今天,我们聚焦“人工智能驱动下的检验医学创新探索与实践”,邀请到来自全国多家医疗机构的专家围绕 LLM 的落地场景、特殊人群数据精准化、实验室管理智能化、结果互认技术闭环、微生物报告优化等关键议题展开深度探讨。让我们以实践为锚点,以技术为引擎,共同绘制检验医学“智能+”时代的发展蓝图。

**1 同济医院率先推出“问问同检”实验室智能助手,实现了检验项目说明查询、报告解读等功能。能否请您具体分享在未来,LLM 在检验结果解读中将会发挥什么样的作用? 人工智能与检验人员的关系**

将如何发展? 在人工智能时代,检验医学行业应如何转型?



**袁旭:**LLM 早期主要由 OpenAI (ChatGPT)、Google (Gemini) 等国外公司主导,这些公司凭借庞大的算力、数据和算法积累,占据了市场的主导地位。然而,DeepSeek、通义千问等国产开源大模型的涌现,促进了本地化智能体的发展,LLM 从“精英封闭”走向“全民可用”,越来越多的开发者和企业基于这些开源模型构建自己的智能体(AI Agent)。

LLM 的赋能能够帮助解决千行百业中长期存在的核心需求与痛点。在检验医学行业中,尤其在检验结果解读方面,传统的检验结果解读往往依赖经验及经验总结出的规则,人工分析结果可能存在主观偏差,而知识转化为规则需要耗费大量的人力且系统间转化能力差。在面对海量的患者数据时,检验人员常常难以全面捕捉每个细节,尤其是面对复杂和多变的疾病时更容易遗漏关键线索。目前绝大部分检验科仅仅停留在向临床医师与患者提供检验结果的阶段,但面向个体的、贴合病情的解释与建议已成为临床医

**作者简介:**李波,男,主任技师,主要从事实验室诊断技术、实验室精益管理及实验室数字化和智慧化研究。<sup>#</sup> **共同第一作者简介:**袁旭,男,主管技师,主要从事实验室信息化建设工作、系统运维管理、检验流程设计、大数据分析等工作。<sup>#</sup> **共同第一作者简介:**李小强,男,主任技师、副教授,主要从事检验信息化、儿童代谢性分子标志物等相关研究。<sup>#</sup> **共同第一作者简介:**牛长春,男,主任技师,主要从事智慧检验及检验结果互认策略等研究。<sup>#</sup> **共同第一作者简介:**武永康,男,教授,主要从事智慧检验及检验结果互认策略等研究。<sup>#</sup> **共同第一作者简介:**胡庆丰,男,副主任技师,主要从事微生物学检验、检验信息系统建设、生物安全管理相关研究。<sup>#</sup> **共同第一作者简介:**杨大干,男,主任技师,主要从事临床基础检验、实验室管理、数据科学研究。

师与患者的迫切需求。而 LLM 的应用使这一需求的实现成为可能。

我们畅想,未来在 DeepSeek、通义千问等 LLM 的基础上,融入检验医学行业知识库与检验知识图谱训练而来的检验医学专业 LLM 能够快速、准确地解读检验项目的临床意义。随着多模态 LLM 技术的不断成熟,该技术有望结合患者的主诉、检验及检查结果、诊断、用药、治疗、手术记录等与时间序列融合的多模态患者数据,实现检验结果数据的个性化精准解读,为临床医师提供更精准、更丰富的诊断支持和决策依据,推动检验结果解读模式从经验依赖向数据与知识双驱动的转变。LLM 技术的引入,无疑让检验医学这一临床的重要“眼睛”看得更加透彻和清晰,可显著提升检验医学服务临床的价值。

目前,检验科通过自动审核与自动化流水线的结合实现了敏捷实验室的建设,在未来,人工智能通过智慧化的结果解读及检验审核智能体、检验工作流智能体将进一步赋能实验室自动化。人工智能技术的迅速发展引发检验人员对失去工作的担忧,但实际上,它的核心价值在于它能够赋能有效运用其技术的人,使他们更加高效地发挥主观能动性。以检验报告的自动审核为例,这项功能的出现曾让部分检验人员担心检验医学行业将消失,但实际情况是通过自动审核的过滤,它让检验人员将关注点聚焦在异常结果报告的审核上,缩短了样本周转时间,提高了报告质量。LLM 的出现进一步优化了检验流程的自动化水平并开拓了以往难以实现的个性化检验结果解读功能。人工智能技术的深度运用,使结合患者病情、病史、治疗方案等信息的个性化检验解读报告成为可能,这将使检验医学行业从单纯提供结果数据的服务模式,转型为提供深入的检验结果解释与咨询服务。此外,检验人员在推动人工智能技术落地医学检验领域的过程中,还需关注并着力解决技术应用中的诸多现实挑战,包括人工智能系统可能产生的机器欺骗与机器幻觉问题,以及模型本地化部署所面临的高成本投入及医疗数据安全保障等关键难题。

因此,检验医学行业未来建设的核心方向,将是充分利用人工智能技术,构建智能化的检验结果解读系统,持续优化检验流程,实现检验服务由单纯提供数据向深入解读数据的转型。通过提高检验结果解

读的深度和效率,人工智能将真正成为检验医学行业的得力助手。在人工智能技术加速发展的新时代,检验人员应把握发展机遇,积极转型,掌握与 LLM 高效沟通的技巧,提早适应检验医学领域人机协作的新范式。

**2 特殊人群(如儿童、孕产期女性)生物参考区间的缺失是临床痛点,能否请您结合儿童检验的工作经验,说明人工智能在特殊人群参考区间设立及验证等工作方面具备哪些优势?**



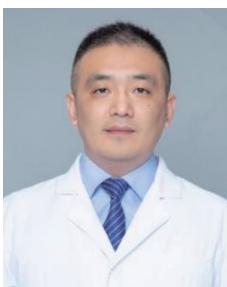
李小强:检验结果的临床诊断价值和有效性取决于合理的医学决定水平和生物参考区间。生物参考区间是指从参考下限到参考上线的区间,通常取参考值分布的中间 95% 区间。目前的来源主要包括卫生行业标准(如 WS/T404、WS/T405、WS/T779、WS/T780 等)、试剂说明书、权威教材、通过研究建立的参考区间或转移法获取的参考区间。建立方法包括国际标准 CLSIEP28-A3c、卫生行业标准 WS/T402 及其他文件规定的直接法、间接法等,这些方法均是以抽样为基本方法。但是抽样的参考个体必须考虑包括年龄、性别、周期节律等多种因素,同时可能因为检测系统不同、方法学差异、环境条件变化而使得所引用的生物参考区间不合适。传统的方法,在人群定义、样本和数据收集、方法选择及数据验证等各个环节均存在诸多问题,目前鲜见对特殊人群(如新生儿、儿童、孕产期女性、老年人群体等)和特殊检验项目的生物参考区间设立和验证,可供选择借鉴和转移使用的数据很少,因此对该类人群的疾病诊疗和健康判断极为不利。

随着 DeepSeek 及类似语言大模型在医院信息系统的部署和植入,相关技术在特殊人群的生物参考区间的获取和细分中可能会发挥重要作用;语言大模型的优势在于通过真实世界检测数据获取样本,并可以有效避免直接抽样法所需的大量样本、人群筛查、分组及差异化统计、临床验证等工作,其具有更强的操作性。语言大模型可以根据特殊人群的需求,设定明确的人群特征和入组条件,包含年龄(如新生儿、儿童、不同年龄段的老人)、性别、种族、地域、孕期(孕前/早/中/晚期)、昼夜节律、生理周期、饮食状况等,

甚至不同基础状态(如肾病、糖尿病、人类免疫缺陷病毒感染等)、特殊生理状态(如运动员、不同海拔地区)等;合理排除可能的混杂因素(如生活习惯、用药、合并症),按照统计学要求设定基础分组和样本量,并不断扩充样本量和加大模型范围;协助样本采集流程标准化或排除信息不清晰的样本;对统计结果做数据库间比较和验证。

语言大模型的介入有助于实现检测结果及生物参考区间的动态化管理和结果验证,并能对各种特殊场景做出针对性和个性化研究,适配不同实验室所用的检测系统和环境,从而让特殊人群和特殊项目的生物参考区间更加精准和具备更高的诊断价值。

### 3 实验室质量管理是检验结果可靠的基石,能够请您分享 DeepSeek 在实验室质量管理方面具有哪些优势? 具体有哪些应用方面?



**牛长春:**在临床实验室质量管理实践中,DeepSeek 的优势在于逻辑推理、决策支持等多种应用技能,输出结果是以问答的方式呈现。在临床实验室质量管理实践中,目前常用的机器学习的方式有决策树、随机森林、支持向量机、卷积神经网络等,这些机器学习算法的输出结果是精确的计算结果,已在疾病诊疗模型建立等研究方面应用广泛。DeepSeek 与以上机器学习的算法不一样,其优势在于具备逻辑推理、决策支持等多种应用技能,输出结果是以问答的方式呈现。DeepSeek 虽然包含一些人工智能计算方法,但是不能够进行具体的、复杂的、精确的数据清洗等操作,也不能对人工智能计算方法进行参数优化,这是其局限性。

DeepSeek 在临床实验室质量管理中的实施监测、方案制订、风险评估等方面具有优势。应用要点首先是建立本地化知识库,DeepSeek 本身的知识更新是截至 2024 年 7 月,虽然其包含了人工智能、计算机科学、自然科学等知识,但其缺乏结合实践的知识,因此需要建立本地化的知识库。本地化知识库的建立需要使用 maxKB、ragflow、Dify 等工具,知识库的输入项为标准操作规程、行业标准等文件,实验室的体系文件[质量手册、程序文件及标准操作规程(SOP)文件等],以及性能验证、不符合项改进等质量

控制过程相关记录等。并且知识库的输入项最理想的是结构化、标准化的输入知识内容,使用标准化的命名规则和标签系统,通过结构化数据提高检索和回答效率,避免机械照搬通用知识和简单翻译,需要通过清洗去除冗余或重复的信息。

应用方向主要包括以下几个方面:(1)识别差错、不符合项等。结合临床实验室质量管理知识库与 DeepSeek 的强大推理能力可形成智能体,智能体如同经验丰富的专家,能够依据标准的要求,对实验室运行中的偏离、差错进行识别,查找原因,并提出改进的建议。智能体可以通过 DeepSeek 将出现的差错、不符合项结合本身和输入的知识库,并利用鱼骨图等工具分析出关键原因,制订纠正措施及持续优化的措施。(2)分析趋势。基于 DeepSeek 在处理临床实验室质量管理各个要素中非线性关系的优势,可利用 DeepSeek 对检测结果、质量指标参数时间序列大数据进行分析,结合仪器维护、质量管理措施实施等要素,分析质量变化的趋势,识别潜在的变化模式及影响因素,为提高临床实验室质量提供科学依据。(3)策划新方案。利用 DeepSeek 本地质量管理知识库,根据输入年度的内审、管理评审及风险分析评估结果等输入项,建立下一步质量控制的工作计划,如人员培训计划等。(4)辅助计算、分析质量控制数据并形成报告。DeepSeek 以 python 语言为基础,包含了 python 处理数据的 pandas、numpy 等数据处理库,以及相关性分析、回归分析(包括岭回归、Lasso 回归等人工智能计算方法)、假设检验、分类与聚类(包括支持向量机、决策树等人工智能计算方法)、因子分析等统计学工具,其能够运用这些数据处理库和统计学工具对质量控制数据进行分析,用于失控数据识别、失控规则识别、仪器间比对数据分析等,并可以进一步依据知识库、SOP 文件等,结合输入的质量控制数据形成质量控制报告。(5)提高风险预警效率。本地化部署的 DeepSeek 可以通过对聊天记录、不符合改进过程、内部审核、外部审核相关记录等文本数据进行语义、情境等方面的分析来判断相关质量管理环节潜在的风险。例如,如果发现有用户高频次咨询仪器结果异常提示的相关问题,可能提示此仪器存在故障风险。(6)通过联网获得最新理论及进展。利用 Page Assist 等插件,可实现本地化部署 DeepSeek 的联网功能,从

而可以获取质量管理相关文献的最新进展，再结合本地知识库，进一步应用到实践中。

**4 检验结果互认是国家层面的重要战略目标，能否请您结合四川大学华西医院的经验，说明人工智能如何赋能检验结果互认工作？具体解决了哪些问题？发挥了哪些作用？**



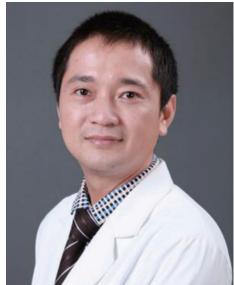
武永康：人工智能具备感知、学习、推理、决策和自然语言处理等类人能力。从某种意义上讲，人工智能可以视为一个拥有强大知识、不断进步且不会退休的工作人员，能够辅助或替代人类完成重复性、高精度或高风险工作。

检验结果互认是国家层面的重要目标，《关于进一步推进医疗机构检查检验结果互认的指导意见》要求到 2025 年底实现医联体内全部项目互认，到 2027 年底实现省域内医疗机构间互认项目超过 300 项。这一目标可促进管理部门和医疗机构加强科研创新。而检验结果互认的核心目标为检验结果能够被不同医疗机构的医生有效使用。因此，有两个关键问题需要解决：一是不同医疗机构间的检验数据对接；二是检验结果的可比性评估。检验结果需要具有可比性才能为基于 SaaS 构建的检验医学知识库提供有效的数据基础，解读报告的结论才能可靠、可信，而人工智能在前述两方面均可发挥重要作用。（1）人工智能在检验数据对接中的作用。目前，医疗机构的信息化水平参差不齐。部分医疗机构仅能为患者提供纸质检验报告，另有部分医疗机构则已实现检验报告的网络共享。对于后者，统一的数据格式标准帮助数据互联互通的实现，但要在全国范围内形成医疗“一张网”的应用格局仍需长期努力。但人工智能技术可以帮助解决未联网医疗机构数据对接的问题。例如，患者可通过微信或互联网医院上传纸质检验报告的图片，人工智能引擎能够自动识别报告中的检验项目、方法、数据和参考范围，并与全国统一项目代码对接，实现数据标准化。随着数据的不断积累与校正，人工智能将推动不同医疗机构间数据对接的持续迭代升级，助力实现检验结果互认这一目标。（2）人工智能在检验结果可比性评估中的作用。检验结果互认的基础是数据可比性。传统方法需通过仪器比对（如 EP9-A2

标准）进行评估，但耗时费力。人工智能技术可借助国家或省市级临床检验中心的质量控制数据库，结合 ISO15189 及 CAP 认可标准，自动评估医疗机构检验项目的质量控制情况。检验结果若通过考核，即可采信。此外，还可以依托临床医师的实际应用场景，借助人工智能技术构建检验结果认可体系。当检验结果无法直接互认时，医生可结合临床判断决定检验结果是否采信。人工智能技术可将相关数据（如医疗机构信息、检验项目符合情况）传输至数据质控比对中心，随着数据的积累和更新，可形成检验结果认可体系的评价依据。这种“大众点评”模式能够逐步提升检验结果的认可度。若不同医疗机构的检测结果的差异有规律可循，人工智能可通过数据转换实现互认，这一方法已应用于临床中<sup>[1]</sup>；若差异无规律，则表明检测仪器可能存在故障，需及时停止相关检测项目以避免误诊。

总之，人工智能通过构建“数据对接-质量评估-临床应用-迭代更新”的技术闭环，正在重塑检验互认和报告解读的模式<sup>[2]</sup>。人工智能不仅能破解传统检验互认的难题，还将推动检验医学进入智能协作新时代，最终实现“一次检测，全国认可”的智慧医疗愿景。

**5 微生物检测周期长、耐药性复杂是临床的难点问题，能否请您分享人工智能在“血培养四级报告模式”、缩短临床决策时间、提升抗菌药物使用的规范性等方面如何发挥作用？**



胡庆丰：在传染性疾病的诊断与治疗过程中，微生物检测发挥着至关重要的作用。然而，传统的微生物检测流程存在诸多挑战，包括检测周期长、微生物耐药机制复杂多变及结果解读困难等问题，这些问题与临床实践中降低住院时间、推进循证医学的目标形成了显著的矛盾。在血流感染等重症感染情境下，早期准确的诊断和合理使用抗菌药物显得尤为关键，这能够显著降低患者的死亡率。因此，微生物分级报告制度应运而生，通过分阶段反馈检测结果，可显著提升临床决策的效率，确保了医疗质量与患者安全。

在临床微生物学检验中，血培养三级报告模式遵循“多级递进”原则：一级报告阶段，血培养呈现阳性

后,检验人员需将阳性样本接种至平板并进行革兰染色涂片,随后将涂片结果作为紧急信息传递给临床部门;二级报告阶段,检验人员将血培养瓶内容物直接涂布于药敏平板,附上相应的药敏试纸,经过 18~24 h,检验人员初步解读药敏结果并及时通报给临床;三级报告阶段,检验人员依据标准化的菌液浓度执行鉴定和药敏试验,最终将详尽的报告、结果评估及建议反馈至临床。该三级报告模式为临床诊疗提供了早期、迅速、精确的用药指导,然而,在实际临床应用中,二级报告阶段进行直接药敏所需时间较长,同时由于临床医师在感染性疾病的诊断与治疗能力上存在差异,很多医师难以充分利用分级报告所提供的信息来优化抗菌药物的使用策略。

近年来,质谱技术、信息技术及人工智能技术等新兴技术的迅猛发展与广泛运用,极大地促进了血培养分级报告流程的优化。优化后具体流程包括:一级报告阶段,检验人员在血培养阳性结果出现后立即进行接种,并出具涂片报告;二级报告阶段,在培养 4~6 h 后,检验人员通过菌苔进行质谱鉴定、酶型检测和药敏试验;三级报告阶段,仪器完成药敏测试后,药敏结果被传输至微生物信息系统,系统进行智能化审核,并出具仪器药敏的初步报告;四级报告阶段,检验人员整合完整的药敏数据与临床结果解读,出具最终报告。

人工智能技术借助自然语言处理(NLP)和逻辑推理引擎,人工智能知识库能够解析复杂的临床问题,并提供精确的建议。在整合微生物检验与治疗的相关指南、导则、标准及实时监测数据的基础上可构建结构化的人工智能知识库,其通过分析患者的病史及病原体的耐药性,结合患者的个体特征可生成定制化的治疗方案,从而为临床医师提供精确的用药建议。

一级报告阶段,检验人员对阳性血培养瓶进行接种及涂片,对涂片进行镜检发布细菌形态的初步报告。涂片结果根据病原体的形态特征(球形/杆状、革兰阳性/阴性),结合系统通过大数据分析技术,提供的科室、重症监护病房(ICU)及非 ICU 病区的年度分离率及耐药率数据,可预测可能的病原体种类,为经验性抗感染治疗提供了参考依据。

二级报告阶段,检验人员对转种 4~6 h 后生长出

的菌苔进行质谱分析、药敏试验及酶型检测。质谱技术能够迅速鉴定出细菌种类,并结合细菌学知识库、耐药性趋势,进一步推荐经验性用药方案。人工智能技术结合实时微生物监测数据,能够预测病原体的耐药性,辅助临床医师提前筛选出更有效的抗菌药物,并预测其对特定抗菌药物的敏感性,从而为临床医师提供更为精确的用药指导。

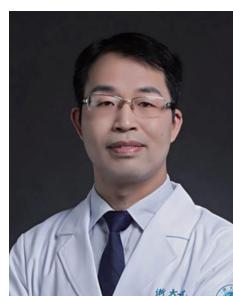
三级报告阶段,仪器完成药敏试验后,药敏结果将被传输至微生物信息系统。该系统将对药敏结果执行智能化审核,并依据既定规则对合格的药敏结果进行初步报告。此外,基于已有的药敏结果与患者的基础健康资料,人工智能系统将提出相应的用药建议。人工智能技术结合了仪器药敏结果及患者个体情况,为临床医师提供已知药敏结果的药物用药的参考方案。

四级报告阶段,检验人员整合详尽的药敏数据与临床解读,形成最终的综合报告。该报告是综合了罕见耐药表型和药敏试验结果的全面的药敏数据。人工智能技术对患者的个体特征(例如基因信息、病史等)及微生物的耐药模式进行深入分析,提出个体化的用药建议,生成个体化治疗方案。这种个体化治疗方案不仅能够提升治疗效果,还能减少抗菌药物的过度使用,有效遏制细菌耐药性的扩散。

在临床沟通的视角下,微生物分级报告与人工智能技术的结合具有极其重要的意义。首先,人工智能技术通过迅速识别病原体和预测耐药性,协助临床医师在第一时间启动精准治疗,避免因等待完整报告而延误病情。其次,人工智能技术提供的初步药敏结果和用药建议能够协助临床医师在治疗过程中及时调整抗菌药物的使用,提高治疗的精准性。最后,人工智能技术驱动的个体化治疗方案为临床医师提供了全面且精准的指导,最大限度地提升治疗疗效并减少耐药风险。

综上所述,微生物分级报告制度与人工智能技术的应用相结合,不仅缩短了临床决策时间,提升了抗菌药物使用的规范性,还有效控制了细菌耐药性的蔓延,保障了医疗质量与患者安全。这种分级报告模式为临床医师与微生物检验人员之间的沟通搭建了桥梁,使双方能够更高效地协作,共同为患者的健康保驾护航。

## 6 请您分享一下您的团队关于 LLM 在检验医学中的应用过程管理？介绍一下当下 DeepSeek 相较于其他人工智能模型具备哪些优势？



**杨大干：**在不久的将来，LLM 将成为实验室信息生态圈的智能基座，助力全面实现智慧检验医学。而选择合适的 LLM 是构建智慧检验医学的基石，目前可采用开源 LLM（如 DeepSeek、Llama、通义千问、ChatGLM 等）进行本地化部署和微调，其中 DeepSeek 具备出色的对话、创作和逻辑推理能力，其性能接近甚至超过 ChatGPT 最新版，其已在全国上千家医疗机构实现本地化部署，用于报告解读、辅助诊断、智能问答、质量管理、教学科研等；也可采用商业闭源模型（如 ChatGPT 系列、文心一言、Claude 等）的应用程序接口（API）进行开发，模型能力强，但存在数据安全担忧，以及成本较高、定制性受限等问题。

LLM 的应用框架涵盖需求适配与场景设计、数据工程与知识构建、模型开发与优化、多维度验证与测试、部署架构与效能优化、持续维护与价值迭代等关键阶段。首先，应剖析检验医学领域的需求，依据应用场景选择 LLM，可优先考虑本地化部署，如 DeepSeek-R1 参数（7B、70B、671B），也可采用 API 调用方案，实现高效推理和精准预测。检验结果中宜整合电子病历数据，便于提取患者信息、病史、诊断结果等信息，在基于遵守数据隐私保护、相关法律法规等要求的前提下构建检验医学领域知识库。其次，提示词、检索增强生成（RAG）、LLM 微调等技术方案宜无缝嵌入现有检验前、中、后及管理流程，以减少对现有操作的干扰。为确保人工智能的性能和可靠性，应多维度验证算法的透明度、可解释性等，以确保人工智能模型的应用能达到预期的效果。LLM 作为辅助工具，有助于临床应用指南的制订，并能够帮助医生更好地理解和使用人工智能工具，但最终决策和审核应由专业人员负责。最后，应从低风险、高价值场景入手，逐步验证、迭代优化，重点解决 LLM 的“幻觉”问题，确保输出信息的准确性和可靠性。

提示词工程 + 结构化查询语言（SQL）是所有信息系统应用 LLM 的基础，应设计精确、有效的指令来

实现文本生成、文本改写、摘要生成、简单问答、代码生成等任务。通过 SQL 获取实验室信息生态圈中的特定数据，配置和调整特定场景的提示词，与 LLM 交互获取所需的服务。RAG 旨在解决 LLM 的知识局限性，如知识截止日期、缺乏领域知识、“幻觉”生成与事实不符内容，通过 RAG 里的知识库来增强模型回答的准确性和可靠性。RAG 适用于需要精确引用的 SOP、最新指南或专业文献、报告注释、知识查询等场景。LLM 可学习特定领域的语言风格、复杂模式、决策逻辑或执行特定任务的格式，从而适用于需要模型深度理解和模仿特定行为模式的场景，如智能结果审核、个性化结果解释、质量管理分析。

选择何种技术及组合，应基于应用场景的需求复杂度、对准确性和可靠性的要求、可用数据资源的质量与数量、机构的技术能力与计算资源，以及预期的模型专业化程度来综合判断。以智能报告解读为例，首先需通过 RAG 检索相关指南和文献，然后利用经过微调的模型结合检索到的信息和原始报告数据，最终在精心设计的提示词引导下生成结构化、个性化的解读建议。LLM 的应用应遵循从简单到复杂、从辅助到逐步自动化、持续验证与迭代的原则。LLM 在检验医学全过程、全要素、全场景的应用中具有知识掌握与问答能力强、推理和特定任务处理能力强、用户偏好与共情体验好、RAG 技术增强效果好等优点。在 LLM 及 AI Agent 的支持下，将有望实现智能甚至无人的智慧检验、管理和服务。

## 参考文献

- [1] 周鑫, 刘在栓, 武宇翔, 等. 人工智能算法在不同仪器游离甲状腺素检测结果可比性中的应用[J]. 检验医学, 2025, 40(3): 299-305.
- [2] 陆小琴, 伍柯, 雷玉倩, 等. 基于大语言模型的实验室报告解读系统构建研究[J]. 医学信息学杂志, 2025, 46(3): 73-78.

（收稿日期：2025-04-12 修回日期：2025-07-28）

（本文编辑：陈玮嘉 张耀元）