

DOI: 10.3969/j.issn.2096-045X.2025.06.019

· 调查研究 ·

## 医学教育数字化转型背景下医学生大模型使用现状调查与分析

张伟 欧阳思远 彭飞翔 赵志强\*

(首都医科大学信息化处,北京 100069)

**【摘要】目的** 调查医学生在医学教育数字化转型中对大模型的使用现状,为优化教学资源提供数据支持。**方法** 通过问卷调查的方法,调查某医学院校160名本科3~5年级医学生,从使用覆盖率、学习/科研/临床场景效果、培训需求等7个维度开展调研。**结果** 大模型普及率超95.40%,医学生中周均使用10次以上占比32.00%,DeepSeek等国产模型使用率领先。学习场景中信息检索(88.75%高效果)和文本翻译(86.25%高效果)表现突出,知识点掌握(43.75%低效果)待提升;科研场景文献分析效率显著(85.63%高效果),实验设计辅助有限;病历书写表现突出(40.62%),成效显著,呈现出明显的“单峰偏态分布”特征。病例讨论的负面评价占比最高,无效果与效果微小的评价合计达39.50%,差错改错的负面评价占62.50%。75.00%学生需求使用培训,聚焦实际操作与场景应用。**结论** 大模型已成为医学生学习科研的重要工具,但在知识内化与临床思维培养中存在局限,需结合院校需求优化功能并开展针对性培训。

**【关键词】** 大模型; 医学教育; 医学生培养; 科研辅助; 临床实践

**【中图分类号】** R-1; G465

**【文献标识码】** A

### Survey and analysis of the current status of large model usage by medical students in the context of digitalization of medical education

Zhang Wei, Ouyang Siyuan, Peng Feixiang, Zhao Zhiqiang\*

(Information Technology Office, Capital Medical University, Beijing 100069, China)

**【Abstract】 Objective** To investigate the current status of medical students' use of large models in the digital transformation of medical education, and to provide data support for optimizing teaching resources. **Methods** A questionnaire survey was conducted using a 5-point Likert scale. A total of 160 medical undergraduates in grades 3 to 5 from a medical college were surveyed, covering 7 dimensions including usage coverage, effectiveness in learning/research/clinical scenarios, and training needs. **Results** The penetration rate of large models exceeds 95.40%. Among medical students, 32.00% use them more than 10 times per week, with domestic models such as DeepSeek taking the lead in usage rate. In learning scenarios, information retrieval (88.75% high effectiveness) and text translation (86.25% high effectiveness) perform prominently, while knowledge point mastery (43.75% low effectiveness) needs improvement. In scientific research scenarios, literature analysis shows significant efficiency (85.63% high effectiveness), but the assistance in experimental design is limited. Medical record writing stands out with remarkable results, taking the lead with a 40.62% data performance and showing an obvious unimodal skewed distribution "characteristic". "Case discussion" has the highest proportion of negative evaluations, with the combined proportion of "no effect" and "slightly effective" reaching 39.50%. For "error correction", the proportion of negative evaluations is 62.50%. 75.00% of students demand training, focusing on practical operations and scenario applications. **Conclusion** Large models have become important tools for medical students' learning and research, but they have limitations in knowledge internalization and clinical thinking cultivation. It is necessary to optimize functions and carry out targeted training based on institutional needs.

**【Keywords】** large models; medical education; medical student cultivation; scientific research assistance; clinical practice

**第一作者** 张伟, 硕士, 高级实验师, 研究方向: 信息系统与信息技术。Email: zhangwei@ccmu.edu.cn

**\*通信作者** 赵志强, 博士, 研究员, 研究方向: 信息系统与信息技术。Email: zhaozhiqiang@ccmu.edu.cn

近年来,国家大力推动教育信息化建设,相继出台《教育信息化2.0行动计划》<sup>[1]</sup>、《“十四五”卫生健康人才发展规划》<sup>[2]</sup>等政策文件,明确提出要加快信息技术与教育教学的深度融合,推进医学教育创新发展。将数字技术深度融入医学教育的各个环节,推动教育理念、教学模式、管理体系和服务形态的全面革新。在医学教育数字化转型进程加速、大模型技术广泛应用的背景下,开展医学生使用大模型调查,旨在全面了解医学生对大模型的接触程度、使用频率、使用场景、使用体验及存在的问题等。本研究通过调查医学生在医学教育数字化转型中对大模型的使用现状,为医学院校优化教学资源配、改进教学方法、推动大模型在医学教育中的有效应用提供数据支持,助力提升医学教育质量,培养适应时代需求的高素质医学人才。

## 1 对象与方法

### 1.1 研究对象

便利抽样选取首都医科大学2022—2024级医学专业本科生共发出问卷185份,收回问卷160份,有效回收率为86.48%。本次调研内容覆盖校内学习、临床训练及科研实践三大核心场景,不涉及日常生活等事项。考虑到医学生培养体系的特殊性,大一学生尚未接触科研实践与临床训练,因此样本选取聚焦于本科3~5年级学生,具体涵盖基础医学、临床医学、护理学、预防医学、生物医学工程、康复医学等多个专业领域,以确保调研样本与研究场景的适配性。

### 1.2 研究方法

本研究采用匿名问卷调查法收集数据,所有研究对象均知情同意。通过文献查阅和主流的大模型应用,结合医学院校实际情况设计调查问卷,内容包括:①基本情况信息(如性别、专业、学院、年级);②医学生大模型使用情况调研分为大模型使用覆盖率<sup>[3]</sup>(如日均/周均使用频率,区分学习、科研、临床实践等场景)、常用大模型(如GPT-4、文心一言、DeepSeek、星火大模型、豆包、

通义千问等)、大模型知识(常用功能、使用场景等)<sup>[4]</sup>、学习场景(提升学习效率方面,如课程辅助、病例分析、知识点查询、知识点掌握速度)、科研场景(提升科研辅助方面,如文献综述生成、数据统计分析、文献阅读、论文撰写效率等)、临床实践(学生下临床方面,如考试备考、学术交流、病历书写、差错改错等)、培训需求(如纳入教学平台、开设培训课程)等。共7个维度28个条目,均采用Likert 5级计分法,采用正向计分,用行为频次测量,单位为每周次数依次为:1分=从未使用、2分=1~2次、3分=3~5次、4分=6~10次、5分=10次以上,得分越高,说明医学生对大模型使用越频繁,对大模型的发展持乐观态度<sup>[5]</sup>。数据预处理阶段,通过计算各题项均值、标准差揭示态度倾向;经Cronbach's  $\alpha$ 系数检验,量表信度 $\alpha=0.87$ ,表明内部一致性良好,确保测量结果的可靠性与稳定性。

### 1.3 统计学方法

采用SPSS 22.0软件进行数据处理,计数资料用例数(百分比)[ $n(\%)$ ]表示,计量资料采用Likert 5级计分法,正向计分,用行为频次测量。采用单因素方差分析,拒绝原假设( $H_0$ )。以 $P<0.05$ 为差异有统计学意义。

## 2 结果

### 2.1 基本情况

本次调研共回收有效数据160份,其中男生65名(占比40.62%)、女生95名(占比59.38%),该比例与医学院校传统性别分布特征相符。各专业样本分布覆盖医学院校主要学科方向,具备一定调研代表性(表1)。

### 2.2 大模型使用覆盖率

在接受调查的学生中,各专业每周使用频率分布特点如下。

(1)使用频率呈现专业分化。基础医学专业学生使用大模型最为频繁,超60.00%学生周均使用达6次以上,10次以上高频率使用者占比32.00%,居各专业之首;临床医学、护理学、预防

表1 基本信息占比情况 (n=160, %)

专业	男生	女生	合计
基础医学	9.38	21.88	31.25
临床医学	5.00	10.00	15.00
护理学	2.50	10.00	12.50
预防医学	5.00	6.25	11.25
生物医学工程	11.25	6.25	17.50
康复医学	7.50	5.00	12.50
总计	40.62	59.38	100.00

医学、康复医学专业学生使用频率相对集中在6~10次区间,其中临床医学(41.67%)、护理学(40%)、康复医学(48.28%)在该区间占比显著,体现医学实践类专业对大模型辅助学习的高需求;生物医学工程专业则以3~5次的中等频率使用为主(42.86%),契合其交叉学科属性与科研实践特点。

(2)中频使用呈主流趋势。3~5次和6~10次是各专业大模型使用的最集中频率区间,表明多数学生已将大模型融入校内学习、临床训练及科研实践等场景,形成规律性使用习惯,凸显其在辅助专业学习中的常态化价值。

(3)零使用现象趋近于无。大模型在学生群体中的普及率极高,除基础医学(2.00%)和护理学(5.00%)存在少量未使用学生外,其余专业从未使用比例近乎为零,印证大模型在专业学习场景中的高渗透与强刚需特征<sup>[6]</sup>(图1)。

在大模型使用上:DeepSeek占比最高,达到25.00%,通义千问次之(占比18.62%),文心一言占比17.44%,豆包占比16.28%,讯飞星火占比11.05%,GPT-4和智谱清言占比相同,均为5.81%。统计可以看出DeepSeek异军突起,国产大模型正在快速发展(此处为多选题)。

### 2.3 学习场景

在调研中学习场景主要是医学生在学校学习基础知识时使用大模型在提升学习效率方面,如课程辅助、知识点查询、知识点掌握速度方面的辅助作用,采用“1分=无效果,2分=效果微小,3分=有一定效果,4分=效果显著,5分=效果很明显”的五级评估体系。研究结果表明:通过数据

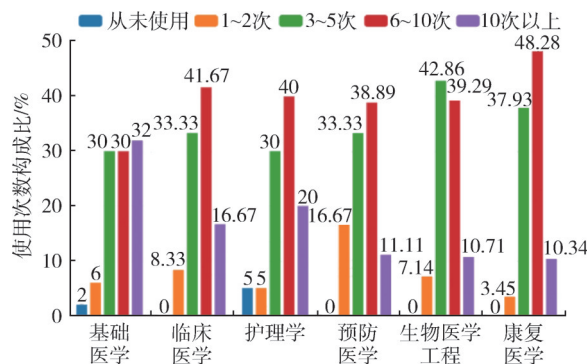


图1 各专业使用次数构成比

图分析可知,各学习场景使用效果呈现显著差异:“信息检索”效果最为突出,“效果显著”与“效果很明显”合计占比88.75%，“无效果”占比仅0.62%;“文本翻译”紧随其后,两项高效果等级合计占比86.25%,用户认可度高。“课程辅助”和“语言润色”效果处于中等水平,集中在“有一定效果”和“效果显著”等级,但部分用户反馈效果微小,仍有优化空间。“完成作业”和“知识点掌握速度”效果较差,“无效果”与“效果微小”合计占比分别达36.67%和43.75%,正面评价显著低于其他场景。综上所述:医学生在基础学习阶段的核心需求呈现“分层化”特征,不同场景的需求强度、性质与紧迫性直接影响大模型的应用效果,大模型在医学生基础学习中展现出“工具性辅助”的强优势,在“信息检索”和“文本翻译”中的高效率,源于其海量数据覆盖与自然语言理解的核心优势。医学领域知识更新快、文献基数大(如PubMed收录超3000万篇文献),大模型可通过预训练的医学知识库(如整合textbooks、指南、文献摘要)实现快速信息聚合,但在“思维培养”(如作业完成、知识点掌握)场景中效果受限<sup>[7-8]</sup>,暴露了大模型在医学专业深度与认知思维训练上的短板。未来可聚焦医学专业特性,优化模型在临床思维训练、个性化学习路径规划等方面的功能,以提升深层次学习辅助效果(表2)。

### 2.4 科研场景

科研场景是指医学生在本科期间参与科研,培养科研素养的能力,主要表现在文献检索与阅读分析能力,实验设计与操作能力,数据收集能

表2 在校学习场景数据 (n=160,%)

在校学习场景	无效果	效果微小	有一定效果	效果显著	效果很明显
信息检索	0.62	1.88	8.75	57.50	31.25
课程辅助	5.00	12.50	43.15	26.25	12.50
语言润色	6.67	33.33	40.00	20.00	20.00
完成作业	10.00	26.00	26.00	26.67	10.00
文本翻译	0.00	1.28	12.20	42.50	43.75
知识点掌握速度	18.75	25.25	25.25	18.75	12.50

力和数据分析能力。在调研中发现：“文献检索与阅读分析能力”中“效果显著”(48.75%)与“效果很明显”(36.88%)，合计占比85.63%，“无效果”与“效果微小”占比为0%。“实验设计与操作能力”中：“有一定效果”占比最高(46.25%)，“效果显著”(26.25%)次之，“效果微小”占比12.5%。“数据分析能力”中“效果显著”(38.71%)、“有一定效果”(29.03%)与“效果很明显”(25.81%)占比分布均衡，无明显短板。从中展现出大模型在文献筛选、关键词检索及核心观点提炼上效率突出。本科生科研入门的首要障碍在于海量文献的筛选与核心观点提炼，需高效解决“如何精准定位关键文献、快速抓取研究脉络”的问题。调研中“文献检索与阅读分析能力”85.63%的正面反馈(“效果显著”+“效果很明显”)，直接印证了大模型对这一核心痛点的精准适配，通过关键词检索优化、文献核心观点提炼等功能，填补了本科生文献处理经验不足的短板。人工智能已成为本科生科研入门的重要辅助工具<sup>[9]</sup>。但在实验方案参考、流程优化等方面提供基础支持，创新性设计(如变量控制、样本量计算)上辅助作用有限。实验设计需兼顾规范性与创新性(如变量控制、样本量计算)，本科生在实验创新性设计上缺乏经验，在数据分析中需高效工具辅助，这两类需求推动了大模型在基础实验流程参考、数据逻辑整理等场景的应用，但也因创新性需求的高阶性，导致实验设计环节的效果反馈呈现“中等偏上但非顶尖”的特征(“有一定效果”占比46.25%)，大模型精准适配了本科科研的入门级痛点，其在信息处理环节不可替代但创造性不足。综上所述，大模型已成为本科生科研能力培

养的有效辅助手段，尤其在信息检索与数据处理环节价值突出。建议在科研训练中结合人工智能工具优势，重点提升文献分析与数据处理效率，同时通过导师指导弥补实验设计等创造性环节的不足，实现“工具辅助”与“思维训练”的协同发展<sup>[10]</sup>(表3)。

表3 科研场景数据 (n=160,%)

科研场景	无效果	效果微小	有一定效果	效果显著	效果很明显
文献检索与阅读分析能力	0.00	0.00	14.37	48.75	36.88
实验设计与操作能力	0.00	12.50	46.25	26.25	15.00
数据收集能力和数据分析能力	0.00	5.45	29.03	38.71	25.81

## 2.5 临床实践场景

临床实践是医学生学习生涯中最重要的一环，是理论知识向临床能力的转化枢纽，临床实践是医学教育中“知行合一”的关键环节，在此次调研中主要调研临床知识储备、病历书写，差错改错、病例讨论等方面。

在统计数据中发现：“临床知识”呈现“中间高、两头低”分布，效果中等占比最高(34.59%有一定效果)。“病历书写”效果显著一枝独秀(40.62%)，“单峰偏态分布”特征显著，但效果两极分化明显。核心源于其需求的强标准化、高重复性特征。临床工作中，病历书写是每日高频任务，且需遵循固定格式规范，属于“结构化信息填充”需求，大模型的模板调用、语义提取能力可直接匹配此类需求，形成“需求明确，工具适配，效果突出”的正向循环。“病例讨论”负面评价占比高(“无效果”+“效果微小”)。“差错改错”负面评价占比超60%(“无效果”+“效果微小”)，大模型缺乏实际互动中的实时反馈、灵活推理能力，无法匹配此类需求的核心诉求，导致效果低下。在临床实践中展现出大模型在基础医学知识检索与解读中表现稳定，但在复杂临床情境的知识迁移上存在局限，在结构化病历模板填充(如主诉、查体记录)中效率突出，但在临床思维逻辑推导、鉴别诊断策略建议等方面能力薄弱，难以替代实际互动中的知识碰撞<sup>[11-12]</sup>，大模型的核心功效特

征可概括为“结构化任务强、复杂推理弱”,其能力边界直接映射到临床场景的效果差异中(表4)。

表4 临床实践场景数据 (n=160,%)

临床场景	无效果	效果微小	有一定效果	效果显著	效果很明显
临床知识	6.29	15.72	34.59	22.05	21.38
病历书写	9.38	21.88	18.75	40.62	9.38
病例讨论	15.88	23.12	21.88	34.38	3.75
差错改错	28.12	34.38	21.88	10.62	5.00

## 2.6 培训需求

在本科生培养调研中发现,将大模型作为辅助学习工具的必要性已得到广泛认可。具体来看,75.00%的学生明确希望获得大模型相关的使用培训,10.00%的学生期待开设系统性的大模型课程,另有15.00%的学生对大模型背后的理论知识表现出浓厚兴趣。

在培训形式上,学生更倾向于通过手册、讲座等实用方式开展学习;内容聚焦高度贴合实际需求,主要包括:大模型的实际操作技巧、科学提问方法、精准关键词定位策略、降低“幻觉”输出的实用技巧,以及大模型在学习场景中的常见应用。此外,学生对知识库体系构建、私有智能体搭建等进阶内容也展现出强烈的学习需求<sup>[14]</sup>。

## 3 结语

本研究通过调查160名医学生在本科教育阶段使用大模型的情况,重点分析学生在大模型应用的体验和在不同学习场景的需求,没有涉及大模型在教育领域的应用。通过本次研究发现:大模型普及率高但使用频率呈专业分化特征,应用效果因场景而异,“工具性辅助”优势与“思维培养”短板并存;学习场景中信息处理类效果突出,认知训练类效果受限;科研场景中信息处理环节不可替代,创造性环节辅助不足;临床场景中结构化任务强,复杂推理弱,学生培训需求明确,注重实用性。创新性地突破“泛化讨论”局限,将研究场景细化为“校内学习、科研实践、

临床训练”三大医学生培养的核心领域,并进一步拆解具体应用场景,捕捉大模型在医学教育不同环节的“效果差异”,为后续技术适配与教学改革提供“场景化”数据支撑,填补了医学教育数字化转型实践研究的空白。

未来研究可扩大样本范围,涵盖不同地区、层次的医学院校,提升结论普适性;结合访谈、实验等多种方法,构建定量与定性结合的研究体系,全面剖析大模型应用痛点<sup>[15]</sup>。在实践层面,医学院校可根据本次调研结果,针对性开发大模型辅助教学资源,如搭建医学专属知识库、开设定制化培训课程;同时,联合技术企业优化模型功能,增强其在医学复杂场景的适配性,为医学教育数字化转型提供更科学的决策依据。

**利益冲突** 所有作者声明不存在利益冲突。

**作者贡献声明** 张伟:研究构思,问卷设计,论文撰写;欧阳思远、彭飞翔:数据收集与分析,方法改进;赵志强:学术指导,论文审定。

## 参考文献

- [1] 中华人民共和国教育部. 教育部关于印发《教育信息化2.0行动计划》的通知[EB/OL]. (2018-04-25) [2025-05-25]. [http://www.moe.gov.cn/srcsite/A16/s3342/201804/t20180425\\_334188.html](http://www.moe.gov.cn/srcsite/A16/s3342/201804/t20180425_334188.html).
- [2] 中华人民共和国国家卫生健康委员会. 国家卫生健康委关于印发“十四五”卫生健康人才发展规划的通知[EB/OL]. (2022-08-03) [2025-05-10]. [https://www.gov.cn/zhengce/zhengceku/2022-08/18/content\\_5705867.htm](https://www.gov.cn/zhengce/zhengceku/2022-08/18/content_5705867.htm).
- [3] 赵志君, 庄馨予. 中国人工智能高质量发展: 现状、问题与方略[J]. 改革, 2023(9): 11-20.
- [4] 黄明芳, 侯青涵, 张伟. 生成式人工智能在医学教育领域的应用现状与未来趋势[J]. 医学与社会, 2025, 38(1): 29-34, 47.
- [5] 卢宇, 余京蕾, 陈鹏鹤. 基于大模型的教学智能体构建与应用研究[J]. 中国电化教育, 2024(7): 99-108.
- [6] 吴永和, 姜元昊, 陈圆圆, 等. 大语言模型支持的多智能体: 技术路径、教育应用与未来展望[J]. 开放教育研究, 2024, 30(5): 63-75.
- [7] 杨宗凯, 王俊, 吴砥, 等. ChatGPT/生成式人工智能对教育的影响探析及应对策略[J]. 华东师范大

- 学学报(教育科学版), 2023, 41(7): 26-35.
- [8] 兰雪, 张晗, 何佳陆, 等. 教育数字化转型背景下医学生数字素养现状及对策研究[J]. 医学信息学杂志, 2024, 45(11): 99-103.
- [9] 方建锋, 王克宇, 房欲飞. 生成式人工智能对教育的颠覆性影响和应对[J]. 全球教育展望, 2024, 53(8): 17-32.
- [10] 赵晓伟, 祝智庭, 沈书生. 教育提示语工程: 构建数智时代的认识论新话语[J]. 中国远程教育, 2023, 43(11): 22-31.
- [11] 余胜泉. 数据赋能的未来教育评价[J]. 中小学数字化教学, 2021(7): 5-10.
- [12] 陈隆升. 数智化时代分布式动态学情测评范式转型与路径建构[J]. 开放教育研究, 2025, 31(1): 110-118.
- [13] 阮秦莉, 张琪, 李国春. 本科生“预防医学”混合式教学的问卷调查分析[J]. 教育教学论坛, 2025(12): 28-32.
- [14] 祝智庭, 胡姣. 教育数字化转型的理论框架[J]. 中国教育学刊, 2022(4): 41-49.
- [15] 李海伟, 王龚, 陆美晨. 教育数字化转型的路径探索与上海实践[J]. 华东师范大学学报(教育科学版), 2023, 41(3): 110-120.

(收稿日期:2025-07-07,修回日期:2025-07-23)

(本文编辑:高健)

**开放获取** 本文使用遵循知识共享署名-非商业性-禁止演绎 4.0 协议(CC BY-NC-ND 4.0), 详细信息请访问 <https://creativecommons.org/licenses/by-nc-nd/4.0/>。

**OPEN ACCESS** This article is licensed for use under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Deed (CC BY-NC-ND 4.0). For more information, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>.