

DOI: 10.61189/054723ljaxsv

· 专家述评 ·

## 医学GPT研发应用之吾见



白春学\*

复旦大学附属中山医院, 上海呼吸物联网医学工程技术研究中心, 上海市呼吸病研究所,  
上海 200032

**[摘要]** 在应对医学GPT技术开发过程中的重重挑战与固有限制时, 笔者建议采取全面且深入的革新策略, 这要求从数据获取到系统运维的每一个环节都进行精细化重塑。数据的收集不再仅仅是量的积累, 而是质的飞跃, 意味着要从海量的医疗信息中精挑细选, 确保每一份数据都具备高度代表性、准确性和可用性。这一过程不仅需要先进的技术手段支持, 更需医学专家的深度参与, 以实现数据的精准筛选与价值挖掘。在基座模型的选择上, 应摒弃传统的简单问答框架, 转而探索构建专家数字人分身的可能性。这样的转变能够让患者仿佛直接面对经验丰富的医生, 获得更加个性化、专业化的医疗咨询服务, 极大地提升了交互体验与信任度。同时, 为了确保医疗信息的安全与准确性, 建议应用基于AI的智能质量控制机制来替代过往的盲目依赖, 通过算法自动审核与人工复核相结合的方式, 严把质量关。此外, 模型的训练评估与优化也应更加注重实践经验的融合。在循证医学的基础上, 倡导将大医的临床智慧与经验融入模型之中, 使医学GPT技术不仅能够依据最新的科研成果, 还能结合临床实际, 为患者提供更为精准、个体化的诊疗建议。简言之, 要实现从数据清洗到精选、从简单咨询到专家分身、从盲目信任到质控把关、从单纯循证到结合大医经验的四大转变。

**[关键词]** 人工智能; 生成型预训练变换模型; 医学GPT; 自然语言处理; 公开证据

**[中图分类号]** R-0 **[文献标志码]** A

### My opinion on the development and application of MGPT

BAI Chunxue\*

Zhongshan Hospital, Fudan University, Shanghai Engineer & Technology Research Center of Internet of Things for Respiratory Medicine, Shanghai Respiratory Research Institution, Shanghai 200032, China

**[Abstract]** In order to cope with the challenges and inherent limitations in the development of medical GPT technology, the author suggests a comprehensive and in-depth innovation strategy, which requires a refined reshaping of every link from data acquisition to system operation and maintenance. Data collection is no longer just a quantitative accumulation but a qualitative leap, which means carefully selecting from a vast amount of medical information to ensure that each piece of data is highly representative, accurate, and usable. This process requires not only the support of advanced technical means, but also the in-depth participation of medical experts to achieve accurate data screening and value mining. In the selection of the pedestal model, the traditional simple question and answer framework should be abandoned, and the possibility of building an expert digital human doppelganger should be explored. This transformation allows patients to receive more personalized and professional medical consultation services as if they were directly facing experienced doctors, which greatly improves the interactive experience and trust. At the same time, in order to ensure the security and accuracy of medical information, it is recommended to apply an AI-based intelligent quality control mechanism to replace the blind reliance in the past, and strictly control the quality through a combination of automatic review by algorithm and manual review. In addition, the training, evaluation and optimization of models should also pay more attention to the integration of practical experience. On the basis of evidence-based medicine, it advocates the integration of the clinical wisdom and experience of big doctors into the model, so that medical GPT technology can not only provide patients with more accurate and individualized diagnosis and treatment suggestions based on the latest scientific research results, but also combine with clinical practice. In short, it is necessary to realize the four major transformations from data cleaning to selection, from simple consultation to expert clone, from blind trust to quality control, and from simple evidence-based to combined with the experience of doctors.

**[Key Words]** artificial intelligence; generative pretrained transformer; medical generative pretrained transformer; natural

**[收稿日期]** 2025-03-19

**[接受日期]** 2025-03-26

**[基金项目]** 癌症、心脑血管、呼吸和代谢性疾病防治研究国家科技重大专项(2024ZD0529300). Supported by Major Project—Research on Cancer, Cardiovascular, Respiratory and Metabolic Diseases (2024ZD0529300).

**[作者简介]** 白春学, 博士, 教授, 主任医师.

\*通信作者 (Corresponding author). Tel: 021-64041990, E-mail: bai.chunxue@zs-hospital.sh.cn

language processing; open evidence

伴随着人工智能(artificial intelligence, AI)、大数据和深度学习技术的进步,生成型预训练变换模型(generative pretrained transformer, GPT)的研究及应用也日新月异<sup>[1-3]</sup>。医学 GPT (medical generative pretrained transformer, MGPT)的研发和应用在医学界更引起了广泛的关注,并在多个领域取得了显著的成果<sup>[4-5]</sup>。

医学 GPT 的涌现,可以助力推动“患者为核心,元医为模版,专病为抓手,质控为保障”的医疗模式与时俱进地实施,旨在达成“名医治未病,元医惠众生”的宏伟目标,从而全面满足患者及社会对医疗与大健康领域的深切期待。然而,要实现这一愿景,必须依托新兴的医学新质生产力。GPT 的适时出现,恰逢其时地填补了这一需求空白。我们的初步研究已显示,大型语言模型凭借其快速响应本质,能够提供丰富的数据共享资源,正逐渐成为关注的焦点。不过,挑战也随之而来。对于缺乏医学基础知识的人群而言,理解和恰当运用 GPT 提供的医疗建议显得尤为困难。加之看不到医生本人,患者难以建立起对医疗服务的信任与安全感<sup>[6-7]</sup>。

为了克服这些挑战与限制,让回答内容既易于被基层医生理解,也便于被具备一定文化知识的患者理解,我们在开发医学 GPT 技术时,必须对各个环节进行全面革新。这涵盖了数据的收集与精选、基座模型的选定、模型的训练评估与优化,以及系统的持续升级与维护。其具体改革措施是将数据清洗升级为数据精挑细选,将简单问答模式转变为专家数字人分身替代,用 AI 质量控制取代盲目依赖以保证安全,以及在循证基础上融入大医临床经验满足个体化需求。简言之,即“改清洗数据为数据精选,改简单咨询为直面分身,改盲目宠信为质控核对,改单纯循证加大医经验”。这利于在继承原有研发成果的基础上进一步推动医学 GPT 技

术转型,构建一个以“患者为中心,名医为模版,专病为抓手,质量为保障”的新型医疗模式,从而更有效地服务于患者,造福社会<sup>[8]</sup>。

## 1 改清洗数据为数据精选

在医学 GPT 研发中,将数据清洗转变为数据精选是一个革命性转折。这一转变的动因深植于效率、模型性能及诊疗准确性的多方面考虑。数据精选致力于提升数据质量,它不仅易于剔除错误与无效信息,更易于从浩瀚数据中筛选出高质量、与需求紧密相关的内容。这些数据精准反映医学知识的精华,利于为模型奠定扎实基础,进而增强诊疗建议的准确性。基于精选,使模型确保数据集富含医学诊断、治疗及药物使用等核心信息,减少噪声干扰。精选数据优化了模型训练。数据集虽小但质量优异,助力模型训练快速收敛,节省时间与计算资源。高质量数据也会助力模型精准学习医学知识与语言模式效率,提升预测性能与泛化能力。通过排除误导性信息,避免模型受不良信息影响。精选数据更贴近医学实践,富含专业术语与经验,将使模型生成的诊疗建议更加专业准确。数据精选还节省了人力物力。虽然初步标注与分类也需人力,但是相较于全面清洗与标注,工作量将大大减少。精选会使资源应用更为高效,集中精力于关键数据,提升研发效率。数据精选符合医学的特殊性,也确保模型学习知识的准确性与严谨性,并满足医学的高标准要求<sup>[2,5,8-9]</sup>。

数据精选在医学 GPT 研发中展现出显著优势,它不仅节省资源与时间,更关键的是大幅提升了回答问题的针对性和准确性。因此,对于医学 GPT,以及任何基于大型语言模型的的研发过程中,将数据清洗升级为数据精挑细选是一个提升模型质量和可靠性的重要步骤(表 1)<sup>[7-12]</sup>。

表 1 将数据清洗升级为数据精挑细选的方法和技术路线

方法	技术路线
明确数据需求	(1)定义目标:首先明确医学 GPT 的应用场景和目标,比如是用于辅助诊断、患者咨询还是医学教育等,以便决定所需数据的类型和范围。(2)确定关键特征:明确对于模型性能至关重要的数据特征,如疾病描述、诊断、治疗、康复方案、检查技术的细节,药物的副作用,以及疾病预防等。
确定精选标准	(1)质量优先:确保数据的高质量,包括信息的准确性、完整性和时效性。(2)相关性:选择与医学 GPT 应用紧密相关的数据,排除无关或噪声信息。(3)多样性:涵盖不同医学领域、疾病类型、患者群体等,以确保模型的泛化能力。(4)合规性:确保数据收集和使用符合医疗隐私法规,如 HIPAA (美国)或 GDPR (欧盟)。

续表

方法	技术路线
实施数据筛选	(1)自动筛选:应用自然语言处理(natural language processing, NLP)和机器学习技术,开发自动筛选工具,根据预设标准快速精选数据。(2)人工审核:在关键数据或自动筛选难以判断时,进行人工审核,确保数据的准确性和适用性。(3)迭代优化:根据模型训练结果和反馈,不断调整筛选标准,优化数据集。
增强数据	(1)数据增强:通过同义词替换、句式变换等方式,增加数据的多样性,提高模型的鲁棒性。(2)标注数据:对于特定任务(如命名实体识别、关系抽取等),进行人工标注,提供高质量的监督学习数据。
持续监控与更新	(1)监控数据质量:建立质量监控机制,定期评估数据集的准确性和时效性。(2)更新数据集:随着医学知识的更新和临床实践的变化,定期更新数据集,确保模型与最新医学知识同步。
应用专业资源	(1)合作与共享:与医疗机构、研究机构和同行合作,共享高质量数据资源。(2)应用公开数据集:关注并应用医学领域的公开数据集,如PubMed、MIMIC-III等,丰富数据集内容。
评估与反馈	(1)模型评估:通过交叉验证、A/B测试等方式,评估不同数据集对模型性能的影响。(2)用户反馈:收集用户反馈,了解模型在实际应用中的表现,指导数据精选和模型优化。

在医学GPT的研发过程中,改清洗数据为数据精选具有更加重要意义。通过实施上述策略,可以将数据清洗升级为数据精挑细选,为医学GPT的研发提供高质量、可靠的数据基础,从而提升模型的性能和可靠性。例如在构建“公开证据(Open

Evidence)”这样的医疗AI助手时,已经体现出多种优势(表2)<sup>[1,7-8,13-14]</sup>。对于像公开证据这样的医疗AI助手来说,数据精选是其持续发展和成功的关键要素之一<sup>[13-14]</sup>。

表2 公开证据体现的数据精选意义

方法	技术路线和内容
提升数据质量	数据精选意味着从海量医学文献中筛选出最相关、最准确的信息。与简单的数据清洗相比,这一过程更加注重数据的价值和可靠性,从而确保AI助手提供的建议基于高质量的医学证据。
增强决策支持	对于医生而言,面对罕见病例和复杂情况时,需要快速而准确的决策支持。数据精选能够确保AI助手提供的信息更加精准,帮助医生在更短的时间内做出更明智的诊疗决策。
优化用户体验	公开证据之所以受到医生的广泛欢迎,部分原因在于其易用性和高效性。数据精选能够进一步提升AI助手的响应速度和准确性,从而优化用户体验,使其更加愿意使用并信赖这款工具。
知识快速更新	促进医学领域知识的迅速更新,数据精选能够确保AI助手及时捕捉到最新的医学研究成果和临床指南,为医生提供最新的诊疗建议。
提高AI竞争力	在医疗AI市场竞争激烈的环境下,数据精选能够提升AI助手的智能水平和专业性,利于在众多竞争者中脱颖而出。对于公开证据这样的产品来说,持续数据精选优化是保持领先地位的关键。
降低误诊风险	通过精挑细选的数据,AI助手能够更准确地识别病例特征,提供个性化的诊疗建议,从而降低因信息不准确或过时而导致的漏诊和误诊风险。

## 2 改简单咨询为直面分身

在医学GPT深入研发中,我们尝试了从简单咨询向名医、大医数字人直面回应的深刻转变。这一变革背后,蕴含着多维度的深思熟虑,极大地增强了用户的亲切感与信任感。用户面对的是名医的数字化身,其形象、语气乃至交互方式,都如同真人般自然亲切。例如,数字人医学GPT在回答用户提问时(图1)<sup>[15]</sup>,可使用户得到这种感觉。名医、大医的权威形象和口碑,专业性与准确性也得到了显著提升。这些数字人掌握庞大的医学知识库和丰富

的临床案例,能够精准引用专业知识,为患者提供个性化、高质量的诊疗建议。随着持续的学习与优化,这些建议会始终保持时效性和准确性。患者的依从性也大大提高。名医数字人的权威指导,增强了患者的诊疗和康复的信心与动力。这一转变还利于减少误解和沟通障碍,提升诊疗效率。同时,这种创新的咨询方式也利于更多患者使用,促进了医疗资源的合理分配<sup>[6-8]</sup>。

名医、大医数字人直面回应的咨询模式,是医学GPT研发中的一大创新,开辟了医疗咨询服务的智能化、个性化和高效化新道路。例如,BAIMGPT,



即是一个涉及技术、用户体验和医疗专业性的综合过程(表3)<sup>[4-5,7-8,15-16]</sup>。

## Q: So how did BAIMGPT answer?

A: ChatGPT can't see us directly. It is just an AI-based language model with no actual medical knowledge and abilities. In medical practice, we need to rely on professional doctors. We need to be cautious about ChatGPT's accuracy. While it may be able to provide some useful information and advice in some cases, it is not a substitute for professional medical diagnosis and treatment. If you have any health problems, you should consult your doctor or medical institution directly.

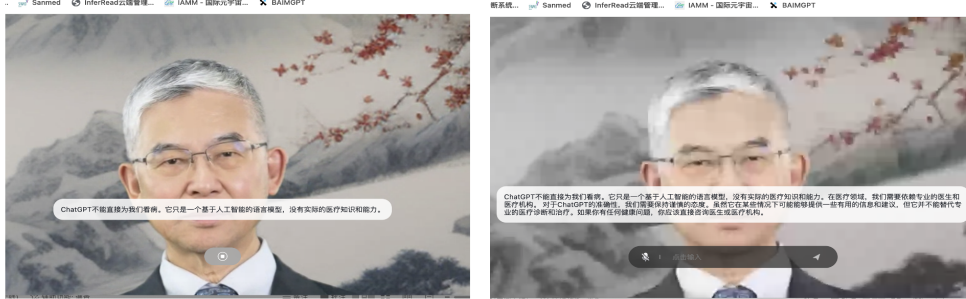


图1 BAIMGPT在回答用户问题

表3 BAIMGPT研发的方法和技术路线

方法	技术路线
明确目标与定位	(1)确定目标用户:明确BAIMGPT用户为基层医生和患者,以及他们的具体需求。(2)定义专家数字人分身:明确数字人分身专业背景为呼吸系统疾病,使其更具专业性和亲和力。
技术准备	(1)NLP技术:提升GPT模型在呼吸系统疾病的理解能力,包括医学术语、病历数据等。(2)多模态交互技术:结合语音、图像、视频等多种交互方式,使数字人分身更加生动和真实。(3)个性化与定制化:根据用户需求,为数字人分身提供个性化的外观、声音和交互风格。
数据收集与训练	(1)医学数据收集:收集大量专家相关的对话数据、病历数据等,用于训练模型。(2)模型训练:应用收集到的数据,对GPT模型进行训练,使其能够准确理解并回答医学问题。(3)专家知识融入:邀请医学专家参与模型训练过程,提供专业知识指导,确保模型的准确性和可靠性。
数字人分身设计	(1)外观设计:根据专家数字人的定位,设计其外观形象,包括面部特征、服装等。(2)交互设计:设计数字人分身与用户的交互方式,如对话流程、表情变化、动作反应等。(3)用户体验测试:通过用户测试,收集反馈,不断优化数字人分身的设计和交互体验。
部署与迭代	(1)部署许可:如获得国家知识产权局的核定使用商品/服务项目注册商标。(2)平台部署:将数字人分身集成到相应的应用平台(如移动应用、网页等),确保稳定运行。(3)持续迭代:根据用户反馈和医学领域的发展,不断更新和优化数字人分身的功能和表现。
合规与隐私保护	(1)遵守医疗法规:确保数字人分身提供的医疗信息和服务符合相关法律法规要求。(2)隐私保护:加强用户数据保护,确保用户隐私安全。
推广与应用	(1)用户教育:引导用户了解并正确使用数字人分身,提升用户体验。(2)市场推广:通过多渠道宣传,提高数字人分身的知名度和使用率。

NLP:自然语言处理;GPT:生成型预训练变换模型。

在医学GPT的研发过程中,将传统的简单问答模式转变为直面专家数字人分身的形式,能更全面地体现GPT技术在医学领域的应用优势。仍以BAIMGPT为例,该项目通过与白春学教授团队、国际元宇宙医学协会以及中国肺癌防治联盟的合作,成功研发出针对肺结节评估管理和阻塞性睡眠呼吸暂停(OSA)诊治管理的专病知识库。这种基于

GPT技术的专病知识库,医学GPT技术不仅拥有强大的信息处理和诊治推理能力,还能够通过专家数字人的形式,将复杂的医学知识以更加直观、易懂和沉浸感的方式呈现给用户。专家数字人分身作为GPT技术在医学领域的一种创新应用,有利于为用户提供更加专业、便捷、个性化的医疗服务(表4)<sup>[4,7-8,15,17-19]</sup>。

表4 数字人医学GPT的优势

效果	方法和技术路线
视觉赋能	将原来的简单咨询模式转变为专家数字人分身替代,通过逼真的数字人形象,用户能够感受到与真实专家面对面的交流体验。这种视觉上的沉浸感,使得用户更容易接受和信任所获取的信息。
亲切感与信赖感	专家数字人不仅能够提供准确的医学知识,还能通过表情、语气等细节,传递出亲切和关怀的情感。这种人性化的交流方式,有助于建立患者与专家之间的信任关系,甚至也有利于对医学生、规培生、进修生和基层医生的培训。
安全感与依从性	在医学领域,用户的安全感和依从性至关重要。专家数字人通过提供专业的建议和指导,能够增强患者对治疗方案的信心和遵从度,从而提高治疗效果。
深度医学知识	大量医学文献和病例数据训练,使BAIMGPT对肺结节和OSA具备更加深厚的医学知识。利于根据用户病史和症状,智能推荐更有针对性的相关医疗信息和建议。
元宇宙医学元素	基于专家数字人分身具有元宇宙医学元素,更利于推动医学知识的普及和共享,促进学术交流和科技创新。这种顶层设计有助于构建更加开放、协同的医学生态系统。
多场景应用	支持在线咨询、远程诊疗、医学教育等多种场景,满足不同用户需求。

MGPT:医学GPT;OSA:阻塞性睡眠呼吸暂停。

这种创新的应用方式不仅能够提升医学信息的传递效率和准确性,还能够增强用户的信任度和依从性,为医学领域的发展注入新的活力。未来,随着技术的不断进步和应用场景的拓展,专家数字人分身有望在医学领域发挥更加重要的作用。

### 3 改盲目宠信为质控核对

在医学GPT或任何医疗相关AI系统的研发历程中,用“AI质量控制核对”替代“盲目宠信”标志着一次深刻且必要的转型。这一转变有多方面的重要意义:(1)利于提升诊疗的准确性。盲目接受AI结果可能导致误诊,而质量控制核对则要求对每项AI建议进行严格验证,确保其基于精确可靠的数据与算法,有效降低误诊风险。(2)患者安全得到更有力的保障。医疗领域首要原则是患者安全,盲目依赖AI可能忽视潜在风险。质量控制核对机制能及时发现并纠正错误,确保提供给患者的建议安全有效。(3)利于增强医生的信心与责任感。医生对经过严格核对的AI建议更为信赖,同时保持独立思

考,提升了决策能力与责任感。(4)利于促进AI系统的持续优化。质量控制核对不仅验证当前输出,还反馈系统性能问题,为算法、数据优化及训练改进提供宝贵信息。(5)符合医疗行业的高标准监管要求。医疗行业对诊疗准确性与安全性要求严苛,实施质量控制核对是满足监管要求、确保AI合法合规应用的关键<sup>[13-14,20]</sup>。

简言之,改盲目宠信为质控核对,对医学GPT等医疗AI系统而言至关重要。它不仅利于提升诊疗质量与安全,增强医生信心,推动系统优化,还利于确保行业合规性,为医疗AI的成熟与广泛应用奠定坚实基础,是确保系统安全性和可靠性的关键步骤(表5)<sup>[2,4,7-8,11-12]</sup>。

这些措施可以将医学GPT的研发过程从盲目宠信转变为基于质控核对的科学流程,从而显著提高医疗AI系统的安全性和可靠性和准确性。为此,笔者建议借鉴公开数据AI助手的研究方法和技术路线研发医学GPT(表6)<sup>[2,5,7,14,21]</sup>。

表5 将盲目宠信改为质控核对方法和技术路线

方法	技术路线
建立严格的数据验证和清洗流程	(1)数据收集:从多个可靠来源收集医学数据,包括文献、病例报告、临床试验等。(2)数据清洗:使用自动化工具和算法去除重复、错误或不一致的数据。数据标注:对医学文本进行准确标注,以便模型能够正确理解和生成医学内容。
实施多层次的模型验证	(1)内部验证:在模型开发过程中,使用交叉验证等技术评估模型的性能。(2)外部验证:在独立的、未参与模型训练的数据集上测试模型,以确保其泛化能力。(3)专家审核:邀请医学专家对模型生成的输出进行审核,确保其准确性和可靠性。

续表

方法	技术路线
引入质量控制机制	(1)自动校对:使用NLP技术,自动检查模型生成的文本中的语法、拼写和医学术语错误。(2)一致性检查:确保模型生成的内容与已知的医学知识和实践保持一致。(3)风险评估:对模型生成的每个建议或诊断进行风险评估,以确定其可能对患者造成的影响。
持续监控和更新	(1)性能监控:实时监控模型在实际应用中的性能,包括准确率、召回率等关键指标。(2)反馈循环:建立用户反馈机制,收集医生、患者和其他用户对模型输出的反馈,用于改进模型。(3)定期更新:根据新的医学研究和数据,定期更新模型,以确保其与时俱进。
强化伦理和法规遵循	(1)伦理审查:在模型开发和使用过程中,遵循医学伦理原则,确保患者隐私和数据安全。(2)法规遵循:确保模型符合相关医疗法规和行业标准,如美国健康保险流通与责任法案(HIPAA)等。
透明度和可解释性	(1)模型透明度:提供关于模型如何工作、如何做出决策的详细信息。(2)可解释性:确保模型生成的输出可以解释,以便医生和患者能够理解其背后的逻辑和依据。

NLP:自然语言处理。

#### 4 改单纯循证加大医经验

在医学GPT的研发历程中,整合单纯循证医学与资深医生临床经验(即大医经验)的举措,其背后蕴含着深刻的考量。循证医学虽应用已发表的研究为基石,却难免受制于研究设计、样本量等局限,导致结论存在偏差,尤其在罕见病或新兴治疗领域,高质量证据更是稀缺。因为它侧重于平均效应,难以充分顾及个体差异和临床复杂性。相比之下,大医经验则凝聚了资深医生长期实践的智慧,超越了现有文献的边界,对复杂多变的临床情况有着独到的见解与处理策略,对指导治疗决策至关重

要。将名医、大医的宝贵经验融入医学GPT,能显著提升系统处理特殊病例与复杂病情的能力。结合循证医学与大医经验,可构建一个更为全面的临床决策支持框架。这一综合方法不仅吸纳了科学研究的严谨性,还融入了临床实践的灵活性,有利于减少误诊与漏诊,提升了治疗的有效性与安全性。它也利于促进医学知识的有效传承,让年轻医生得以快速汲取前辈的智慧,同时也利于医学研究开辟新路径,推动了知识的创新与发展。更为重要的是,这种融合模式回应了患者对医生经验与专业性的期待,提供了更加贴心、精准的医疗服务,从而增强了患者的信任与接受度<sup>[4,7,9]</sup>。

表6 应用公开证据AI助手方法研发医学GPT之吾见

方法	技术路线
使用专门的小型模型	开发或采用专门针对医学领域设计的小型模型,而非直接应用通用大模型。这样的小型模型能够更深入地理解医学术语和上下文,从而提供更准确的信息。
依赖顶级医学文献	训练数据严格筛选,只采用如柳叶刀、新英格兰医学杂志等顶级医学文献的数据,可以确保模型学习的内容具有高度的权威性和准确性,避免垃圾信息的干扰。
实施多模型协作架构	通过构建多模型协作系统,综合应用不同模型的优势,相互验证和纠正信息,从而进一步提高输出的准确性。这种架构有助于防止单一模型可能产生的偏见或错误。
严格的数据校验机制	在输出阶段,实施严格的数据校验机制。这包括对生成的内容进行医学逻辑检查、事实核对以及与已知医学知识库的对比,以确保信息的准确性和可靠性。
强调用户反馈和持续改进	建立用户反馈机制,收集医生、患者等用户的实际使用体验和问题反馈。通过不断迭代和优化模型,以便逐步提升其性能和准确性。
注重医学知识的公平获取	致力于促进医学知识的公平获取,特别是为偏远地区的医生提供便利。这可以通过提供免费或低成本的访问权限、开发适合不同语言和文化背景的界面等方式实现。
借鉴公开证据创始人经验	研发团队应借鉴公开证据创始人丹尼尔的经验,深入了解医学领域的真实需求,以用户驱动增长为核心理念。同时,应关注错位竞争,寻找与其他AI医疗产品的差异化点,并将潜在的弱点转化为优势。

GPT:生成型预训练变换模型。



在医学GPT的研发过程中,将原来的单纯循证方法提升为结合大医临床经验的方法,是一个复杂但富有成效的任务。其中比较重要的是理解循证医学与大医经验各自的优缺点和互补性。循证医学是基于科学证据,通过系统综述、临床试验等方法得出的结论,具有普遍性和客观性。而大医经验则是基于临床实践积累的大量个案经验,具有独特性和主观性,能够反映患者的个体差异和复杂病情。以下建议是帮助我们在循证医学基础上融入大医临床实践经验,以满足个性化医疗需求的方法(表7)<sup>[2,5,7-8]</sup>。

在医学GPT的研发中,融合循证医学与大医经验不仅是对循证医学局限性的有力补充,更是提升临床决策质量、促进医学知识传承与创新迭代发展的重要方式。例如笔者研发的PNapp 5A(肺结节诊

治5步小程序)即是从三加二式诊断法更新迭代而来(图2)<sup>[22-26]</sup>。随着CT的普及,特别是开展CT筛查肺癌之后,肺结节的发现越来越多。究竟如何处理?是手术切除还是观察?为了达到既易学习,又可用于临床的目的,笔者提出“三加二式诊断法”。“三”为三步检查法,“二”为两个备用方案。三步检查法为采集病史、无创检查和有创检查。两个备用方案为手术切除和随访观察<sup>[27]</sup>。随后,为融入AI助手的功能,笔者又提出了PNapp 5A这一人工智能肺结节诊治5步小程序(图2)。这利于通过“复杂问题简单化、简单问题数字化、数字问题程序化,程序问题体系化”的策略,将肺结节诊治中国共识和亚太指南融入PNapp5A和初评与研判二流程体系,更利于将手工业作坊式诊疗模式提升为达到国家和国际标准的物联网流水作业工程<sup>[22-26]</sup>。

表7 整合循证医学与大医经验的方法和技术路线

方法	技术路线
数据收集与整理	收集大医的临床案例、经验总结、患者反馈等数据。对这些数据进行结构化处理,便于后续的分析和整合。
建立知识图谱	构建包含循证医学知识和大医临床经验的医学知识图谱。在图谱中明确标注循证证据的来源、等级以及大医经验的适用场景和限制条件。
算法融合	在医学GPT的算法中融入对大医经验的识别和处理能力。设计算法能够根据患者的具体情况,自动匹配最相关的循证证据和大医经验。
个性化推荐	基于患者的病史、症状、体征等信息,使医学GPT能够生成个性化的医疗建议。在建议中同时呈现循证证据和大医经验,并说明两者的关系和互补性。
持续学习与优化	医学GPT应具备持续学习的能力,能够不断吸收新的循证证据和大医经验。通过用户反馈和临床验证,不断优化算法的性能和准确性。

GPT:生成型预训练变换模型。



图2 PNapp5A

在医学GPT的研发过程中,将原来的单纯循证方法提升为结合大医临床实践经验是与时俱进的时代需求,利于满足个体化医疗需要,具有深远的临床意义。(1)提升决策精准度与个性化治疗水平:单纯循证医学主要基于已有的研究证据进行决策,但往往难以充分考虑到患者的个体差异。融入大医的临床实践经验,可以使得AI系统更加贴近实际临床场景,更准确地识别患者的特殊需求和病情细微差别,从而提供更加个性化的治疗水平。(2)增强AI系统的可信度和接受度:名医、大医的临床经验是融合了理论与实践的医学领域宝贵的财富,是经过长期实践验证的有效方法。将这些经验融入GPT系统,可以显著提升其可信度和临床医生的接受度,促进AI技术在医疗领域的广泛应用。(3)促进医学知识的传承与创新:通过GPT技术,大医的临床经验可以被系统地记录、整理和传播,这有助于年轻医生快速学习和掌握前辈的经验,减少学习曲线,加速将陈述性知识转变为程序性知识。同时,AI系统还能助力医生在分析病例和医学期刊的过程中,发现新的治疗方法和疾病模式,推动医学知识的创新和发展。(4)优化公开证据在医学领域的应用:公开证据强调公开、透明的医学证据,而融入大医经验可以视为对循证医学的一种补充和丰富。GPT技术可以高效地处理和分析这些综合信息,为医生提供更加全面、准确的决策支持。(5)拓展AI在医疗领域的市场空间:中国丰富的病例资源为AI技术的发展提供了得天独厚的条件。结合名医、大医经验,GPT系统能够更精准地定位市场需求,开发出更具针对性的医疗产品和服务,从而拓展市场空间,为AI创业公司提供更多的发展机遇。(6)进一步强化患者安全与医疗质量:AI技术通过融合循证医学与大医经验,能够更全面地评估患者的健康状况,预测潜在风险,并提前采取预防措施。这有助于降低医疗差错,提高患者安全性,同时提升医疗服务的整体质量<sup>[2-3,6,11-14,28]</sup>。

综上所述,在医学GPT的研发过程中,改单纯循证为加上大医经验不仅有助于提升医疗决策的精准度和个性化水平,还能增强AI系统的可信度和接受度,促进医学知识的传承与创新,优化公开证据的应用,拓展AI在医疗领域的市场空间,并强化患者安全与医疗质量。这一转变对于推动医学GPT技术的发展和具有重大的意义。

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 白春学:选题、撰写、修改文章。

#### 参考文献

- [1] GHASSEMI M, BIRHANE A, BILAL M, et al. ChatGPT one year on: who is using it, how and why?[J]. *Nature*, 2023, 624(7990): 39-41.
- [2] MINSSEN T, VAYENA E, GLENN COHEN I. The challenges for regulating medical use of ChatGPT and other large language models[J]. *JAMA*, 2023, 330(4): 315-316.
- [3] ARONSON J K. When I use a word. ChatGPT: a differential diagnosis[J]. *BMJ*, 2023, 382: 1862.
- [4] YANG D W, ZHOU J, SONG Y L, et al. Metaverse in medicine [J]. *Clin eHealth*, 2022, 5: 39-43.
- [5] YANG D W, ZHOU J, CHEN R C, et al. Expert consensus on the metaverse in medicine[J]. *Clin eHealth*, 2022, 5: 1-9.
- [6] 顾建英, 宋振举, 白春学. 拥抱新质生产力, 开创医学新时代[J]. *元宇宙医学*, 2024, 3(1): 1-2.
- [7] 白春学. 医学新质生产力之我见[J]. *元宇宙医学*, 2024; 3(1): 3-10.
- [8] YANG D W, SUN M T, ZHOU J, et al. Expert consensus on the "digital human" of metaverse in medicine[J]. *Clin eHealth*, 2023, 6: 159-163.
- [9] 韦 球, 蒋维芃, 杨超勉, 等. 医学数字人GPT的研究现状及展望[J]. *元宇宙医学*, 2024, 1(1): 43-51.
- [10] THIRUNAVUKARASU A J, TING D S J, ELANGO VAN K, et al. Large language models in medicine[J]. *Nat Med*, 2023, 29(8): 1930-1940.
- [11] SINGHAL K, AZIZI S, TU T, et al. Large language models encode clinical knowledge [J]. *Nature*, 2023, 620(7972): 172-180.
- [12] MARKS M, HAUPT C E. AI chatbots, health privacy, and challenges to HIPAA compliance[J]. *JAMA*, 2023, 330(4): 309-310.
- [13] SCHWALBE N, WAHL B. Artificial intelligence and the future of global health[J]. *Lancet*, 2020, 395(10236): 1579-1586.
- [14] HAUG C J, DRAZEN J M. Artificial intelligence and machine learning in clinical medicine, 2023[J]. *N Engl J Med*, 2023, 388(13): 1201-1208.
- [15] Newsletter: the first metaverse medical digital human GPT launches[J]. *Clin eHealth*, 2024, 7: 51.
- [16] 胡 洁. 元宇宙医学数字人GPT研究快报[J]. *元宇宙医学*, 2024, 1(1): 2.
- [17] 国际元宇宙医学协会睡眠呼吸专家组. 云加端物联网辅助诊治睡眠呼吸暂停(OSA)专家共识(2022版)[J]. *复旦学报(医学版)*, 2023, 50(4): 613-9, 32.
- [18] 白春学, 陆俊羽, 蒋维芃, 等. 医学新质生产力赋能阻塞性睡眠呼吸暂停诊疗新模式[J]. *元宇宙医学*, 2024, 3(1): 22-8.



- [19] 物联网在睡眠呼吸疾病诊治中的应用专家组. 物联网在睡眠呼吸疾病诊治中的应用专家共识[J]. 国际呼吸杂志, 2013, 33(4): 241-244.
- [20] RAJPURKAR P, CHEN E, BANERJEE O, et al. AI in health and medicine[J]. Nat Med, 2022, 28(1): 31-38.
- [21] SHANAHAN M, MCDONELL K, REYNOLDS L. Role play with large language models [J]. Nature, 2023, 623 (7987) : 493-498.
- [22] YANG L, YANG D W, YAO M, et al. Concept and prospect of the Human-Computer Multi-Disciplinary team (MDT) in pulmonary nodule evaluation [J]. Clin eHealth, 2023, 6: 172-181.
- [23] 中华医学会呼吸病学分会, 中国肺癌防治联盟专家组. 肺结节诊治中国专家共识(2024年版). 中华结核和呼吸杂志, 2024, 47(8): 716-29.
- [24] YANG D W, LAM S, KAI W, et al. Expert consensus on the evaluation and management of high-risk indeterminate pulmonary nodules[J]. Clin eHealth, 2024, 7: 27-35.
- [25] 王源, 杨达伟, 白春学. 元宇宙时代的肺癌筛查与管理[J]. 元宇宙医学, 2024, 2(1): 13-5.
- [26] 中国物联网辅助评估管理肺结节专家组. 物联网辅助评估管理肺结节中国专家共识[J]. 国际呼吸杂志, 2022, 42(1): 5-12.
- [27] 白春学. 肺结节“三加二式诊断法”[J]. 国际呼吸杂志, 2013, 33(6): 401-402.
- [28] 宋元林, 宋振举, 钱菊英, 等. 元宇宙技术在医疗质量控制中的应用[J]. 元宇宙医学, 2024, 2(1): 1-8.

#### 引用本文

白春学. 医学GPT研发应用之吾见[J]. 元宇宙医学, 2025, 2(1): 6-14.

BAI C X. My opinion on the development and application of MGPT[J]. Metaverse Med, 2025, 2(1): 6-14.