



“好帮手”or“猪队友”

该如何拥抱 AI？



医生的“小助手” 患者的“陪诊员”

近日，德国研究团队让 DeepSeek-R1 参加了一场美国医师执照考试（USMLE）。结果显示，DeepSeek-R1 不仅通过了 USMLE 三个阶段的考试，还在与 ChatGPT 的比试中有着出色的表现。研究发现，DeepSeek 在第一阶段和第二阶段的精确匹配性能明显超过 ChatGPT，显示出其在基于事实的回忆和临床知识检索方面的优势。让人对其诊疗能力充满期待。

“非常准！”首都医科大学附属北京胸科医院口腔科主任吴斌经常会用 DeepSeek 模拟诊断，他表示，DeepSeek 的诊断准确率很高，给出的医疗建议较为专业，“甚至有些时候就是医生要告诉患者的话”。

“目前，AI 工具被定位为医疗流程中的辅助工具，而非替代人类医生。将 AI 集成到医疗工作流程中可以提高效率，但最终诊断仍需结合医师的专业知识和临床经验，并与患者共同进行医疗决策。”吴斌说。此外，AI 生成的科普内容也很靠谱，角度十分全面，有时稍加修改便可形成科普文章，十分方便快捷。

“AI 技术对中医药诊疗水平的提升意义十分重大！”中国睡眠研究会理事长黄志力教授表示，其一，AI 能分析海量数据，助力诊疗标准化；其二，可辅助诊断，通过整理脉象等数据实现精准诊断；其三，能带动中西医结合发展；其四，可将中医

精华形成理论，实现人才培养和传承的普遍化、广泛化与高效率。

在 ChatGPT、DeepSeek 等 AI 工具诞生前，患者拿着检查结果，通过搜索引擎查询自己所患疾病的时候，往往是极不靠谱的，甚至有了“某度看病，癌症起步”这样的顺口溜。而在 AI 工具诞生以后，首诊准确率得到了大幅提升。

“面对五花八门的科室名称、越来越细的亚专科，患者往往会‘丈二和尚摸不着头脑’，而 AI 技术可以显著优化问诊、挂号和分诊流程，提升医疗服务的整体效率和患者体验。”吴斌表示，目前，一些医院已将 AI 与挂号系统融合，相当于让患者拥有一位 AI 陪诊员，从预约挂号、院内外导航，到候诊报到、预问诊交流，再到预约检查，AI 陪诊员的全程陪伴，让患者看病的流程更加舒心，更为顺畅。

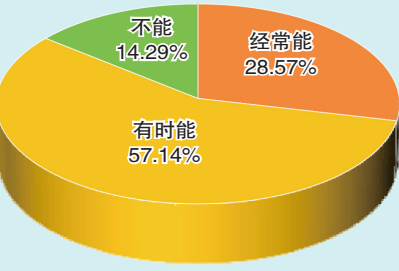
“AI 对医患关系的影响整体应该是向好的。”中日友好医院呼吸中心余中光认为，医患沟通不畅很大一部分源于信息的不对称，在此方面，AI 一方面可以很好的弥补这一认知的鸿沟，让患者更便捷地了解自己的病情及治疗建议，增加患者健康素养，甚至缓解心理焦虑。另一方面，AI 能够快速、大量处理数据，节省了医生的时间和精力，让医生有更多时间关注患者的治疗与人文关怀。

医师报讯（融媒体记者 尹晗 刘则伯）今年年初，深度求索（DeepSeek）以“中国速度”席卷全球。其引发的人工智能（AI）浪潮不仅融入了生活的方方面面，更覆盖了基础电信企业、云计算、芯片、金融、汽车、手机等多领域，医疗领域也不例外。从检索、归纳学术论文，到辅助临床决策、患者管理，再到打破各学科壁垒、推动科研创新，一时间，AI 在医疗领域的应用成为全行业热议的话题。

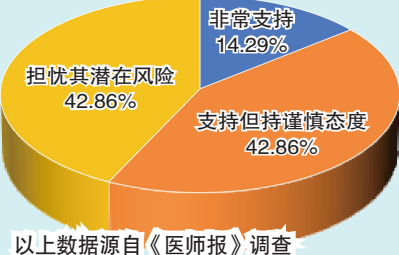
但在实际应用中，不少医务人员也发现 DeepSeek 等 AI 工具存在“检索速度过慢、效率不高”“编造并不存在的文献”等问题，让人不得不对其实时性、准确性与可靠性打一个问号。2月10日，《医师报》开展题为“‘好帮手’还是‘猪队友’？火爆全网的AI，你怎么看？”的问卷调研，其中70%的医师担心其检索内容的准确性与可靠性问题；仅28.57%的医师认为AI写作的内容“经常能”让人感到满意。

AI 能为现有的临床诊疗工作带来哪些便利，医生又该如何应对“AI 时代”的到来？《医师报》针对此话题，采访了 AI 软件工程从业者及多位临床医师，以期帮助广大医务人员拥抱 AI 带来的变革与机遇。

使用 AI 工具写作时，其提供的内容是否能让您满意？



作为医生，您对 AI 工具持何种态度？



以上数据源自《医师报》调查

< AI会让医生失业吗？ >

Deepseek竟这样回答！

Deepseek:
AI如同医生的“超级智能助手”
它能3秒筛查上千张医学影像
预警90%的潜在用药风险，
让基层诊断准确率提升
医生则专注疑难杂症、
手术急救和人文关怀
就像《柳叶刀》研究揭示的
AI+医生的乳腺癌诊断准确率（96.7%）
远超两者独立工作。

DeepSeek“高情商”回答

没有完美的工具 “深度求索”尚待“深度训练”

“用 DeepSeek 学习文献，它给我编了一个。对，你没看错，它编了一个本不存在的文献。如果不是特别熟悉这个话题，我差一点就信了……”

春节期间，一位医师的朋友圈截图传遍了大小微信群，让人不禁对 AI 工具心存疑虑。他所提到的现象有一个专业名词：AI 幻觉。

在生成性 AI 平台 Vectara 大模型幻觉评测排行榜（HHEM）2.1 测试中，DeepSeek R1 模型的“幻觉”率高达 14.3%，Deepseek V3 模型为 3.9%，而 OpenAI 的 GPT-3.5 的幻觉率也达到了 3.5%。

2024 年的一项研究显示，各种聊天机器人在引用参考文献方面的错误率为 30%~90%——它至少会把论文的标题、第一作者、出版年份中的两项写错。尽管它通常会给出警告，告知使用者对重要内容真实性进行核查。

不少医生在调查中也表示，AI 经常会编造引用文献，甚至会出现“引用了十篇文献，没有一篇对”的情况。

“这是因为模型生成答案的底层逻辑不是基于逻辑而是训练时候提供的文字权重，举个例子：你提问‘人为什么感冒’，模型会将问题拆解成：‘人’‘为什

么’‘感冒’，并根据训练数据展现最可能出现的词组。因此，从专业知识到 10 以内加减法，模型给出的答案都有可能出错。”

亚马逊服务（AWS）高级软件工程师 Huawei Mi 告诉《医师报》记者：现阶段，普通人使用免费大模型主要是为了娱乐和辅助处理一些重复性工作，免费模型也会提醒用户模型不能保证给出正确答案。而在专业领域，目前还需要机器学习工程师和数据科学家来做进一步处理，包括怎么找专业训练数据，用什么样的基础模型做微调，怎么使用专业数据训练，怎么提问，每个步骤都有许多工具集，数据科学家需要根据不同情况试验和分析来决定用不同工具来训练和询问模型。

“现阶段，使用者往往需要给提问加一定限制条件来规范问题的范围。怎么提问是个专业问题，问对问题也可以提高答案的准确度。有个专业词汇叫 Prompt（大模型指令），可以帮助模型更好地理解使用者的意图，并作出相应的响应。”Huawei Mi 说，大语言模型现在还是一个新兴技术，和互联网乃至前几次工业革命开始阶段一样，我们知道这个技术未来很有潜力、会改变游戏规则，但在具体应用层面还不是很成熟。

本地化部署 让 AI 更安全更聪明

据不完全统计，截至目前，已有包括北京大学第一医院、清华大学附属北京清华长庚医院、湖南省人民医院等十余家医院官宣完成 DeepSeek 本地化部署。从辅助诊断、治疗、预后判断到疾病科普，甚至包括对医护人员病历的书写、病历质控、讲座 PPT 制作、科研翻译和资料查询，各家医院的应用场景也有所不同。

“现阶段，AI 一定程度上可以辅助医生更好地做出临床决策，帮助医生查漏补缺，但由于 AI 技术、伦理、法律法规等尚不完善，AI 很难取代临床医生的工作。”余中光表示，在应用 DeepSeek 带来便利的同时，必须注意数据安全性问题，国家规定院内数据不能离院，因此必须进行本地化部署，以防止个人信息和隐私泄露，这是不可逾越的红线。

“本地化部署后，医院可对 AI 进行深层次‘二次开发’，利用医院高质量的真实医疗数据训练 AI。”Huawei Mi 说，让 AI 越来越“聪明”的方案有三：

一是增大模型的训练数据，让正确答案的权重变高，理论上只要数据充足，未来模型会越来越精确；二是微调，即用一个中小型模型专门对大模型进行针对性训练，给 AI 提供大量专业数据，让结果更精确，让正确答案出现概率变高；三是通过人工或基于规则的验证，即通过不断重复过程来增加精度。

“尽管 AI 发展迅速，但医生也不必为失业担忧，因为医疗不仅是治愈，还包含人文关怀，这是 AI 无法做到的。”黄志力建议青年医师积极进取，提升医疗技术，掌握 AI 技术，为人民健康贡献更大力量。

当被问及 AI 是否会让医生失业时，DeepSeek 给出了一个非常“高情商”的回答。它说：“AI 如同医生的‘超级智能助手’，它能 3 秒筛查上千张医学影像、预警 90% 的潜在用药风险、让基层诊断准确率提升，医生则专注疑难杂症、手术急救和人文关怀。就像《柳叶刀》研究揭示的 AI+ 医生的乳腺癌诊断准确率（96.7%）远超两者独立工作。”