AI 諮詢醫療資訊：什麼能信、什麼不能信

Using AI for medical information: what to trust, what not to

林協霆, MD, 內科專科醫師, 腫瘤內科專科醫師
醫療財團法人辜公亮基金會和信治癌中心醫院 腫瘤內科部 · ORCID: 0009-0002-3974-4528

發表日期：2026/05/12 · 最後更新：2026/05/12 · 審稿：林協霆 (2026/05/12) · 主題：病人使用 AI 取得醫療資訊 (Patient use of AI for medical information)

DOI: 10.5281/zenodo.20131267 · 此版本 10.5281/zenodo.20131268 ·
https://lin.hsiehting.com/posts/2026/using-ai-for-medical-info/

摘要 · ABSTRACT

ChatGPT、Claude、Gemini 在醫學知識答題已超過多數住院醫師，但仍會「自信地編造」（hallucination）、引用不存在的論文、給予過時建議。本文用腫瘤科醫師視角整理：AI 能做什麼、不能做什麼、怎麼問才不被誤導，並列出 8 個高風險情境（藥物劑量、急症判讀、保單理賠依據等）絕對要找真人專業。

AI 在醫學知識答題已超過多數住院醫師（USMLE、TWMLE 模擬考過及格），但仍會「自信地編造」——編造不存在的論文、給過時建議、誤譯關鍵字。本文以腫瘤科醫師視角整理：AI 能做什麼、不能做什麼、怎麼問才不被誤導、哪些情境絕對要找真人專業。並提供 8 個高風險情境清單（藥物劑量、急症判讀、保險理賠依據等）。

閱讀對象
本文設定讀者為使用 ChatGPT / Claude / Gemini 查醫療資訊的病友與家屬，以及第一線醫療同仁的衛教輔助。本文整理一般使用原則，不取代個別醫療決策。

AI 在醫療資訊上的能與不能

AI 適合的場景（綠燈）

場景	為什麼
解釋醫學名詞	「dMMR 是什麼意思？」 — 知識相對穩定、容易驗證
翻譯英文論文摘要	翻譯品質高，但專有名詞要核對
整理常見副作用清單	對標準藥物資訊整合度高
生成「該問醫師的問題」	幫病人組織思緒
白話解釋已知治療策略	「什麼是新輔助治療？」這類概念題
練習對話／預演就診	安全感、降低焦慮

AI 容易出錯的場景（紅燈）

場景	問題
引用論文與 DOI	hallucination 比例 10–40%，常見「看似真實但不存在」的論文
最新指引與試驗結果	訓練資料有 6–18 個月延遲；2024 後試驗常缺失
罕見癌、罕見變異	訓練資料少、容易混淆
個別藥物劑量、藥物交互作用、過敏	缺乏個人病歷與檢驗值
台灣健保給付條件	多數 AI 不熟悉台灣本土規範
保單理賠條款細節	條款逐字判讀超出 AI 能力
影像判讀（X 光、CT、病理）	通用聊天 AI 沒有專科認證、無責任歸屬

8 個「絕對找真人」的高風險情境

> 紅旗清單
> 以下情境**不要靠 AI 拍板**，直接找主治醫師、24 小時專線或急診：
>
> 1. **急性症狀**：發燒、呼吸困難、胸痛、嚴重出血、抽搐、意識變化
> 2. **化療相關副作用 ≥ CTCAE Grade 2**
> 3. **新藥劑量、停藥決策、藥物交互作用**
> 4. **過敏 / 不耐受處置**
> 5. **保單理賠資料準備**（保險公司只認診斷書、病理報告）
> 6. **重大治療決策**（要不要手術、要不要參加試驗、要不要安寧）
> 7. **遺傳諮詢**（BRCA、Lynch、CDKN2A 等檢測結果解讀）
> 8. **末期照護決定**（DNR、AD、安寧轉介）

Hallucination：AI 為什麼會「編造」？

Hallucination 是 LLM（大型語言模型）的根本特性，不是 bug。模型生成文字時不是「查資料」而是「預測下一個字最可能是什麼」。當被問到不確定的細節（例如某個試驗的精確 PFS 數字、某個 DOI），AI 會用「**最像答案的東西**」生成內容——可能完全正確、可能部分錯、可能完全虛構。

常見 hallucination 模式

類型	範例
虛構論文	「KEYNOTE-XXXX 試驗（Smith et al, NEJM 2023）」— 試驗號或作者不存在
錯誤 DOI	DOI 結構正確但 doi.org 解析 404
混淆藥名	把 olaparib（PARPi）與 osimertinib（EGFR TKI）混淆
過時數字	引用 2018 試驗，但 2022 已有更新版本
錯誤適應症	把 A 藥的適應症套到 B 藥
編造健保條件	「健保 2025 年起給付 XXX」— 實際沒有

怎麼驗證？

想驗證	工具
論文 DOI	doi.org 直接貼
試驗 NCT 編號	clinicaltrials.gov 搜尋
藥物適應症	TFDA 仿單、FDA label、Drugs@FDA
健保給付	健保署網站「藥品給付規定」
治療指引	NCCN（要會員）、ASCO、ESMO、台灣腫瘤醫學會
醫院個案經驗	主治醫師或個案管理師

5 個提問技巧（讓 AI 答得更準）

> **要求引用具體來源**
>
> 例：「依 NCCN 2026 或 UpToDate，乳癌術後 5 年內復發風險的數據是多少？請列出 DOI。」AI 在被要求引用時較不敢編造（但仍需驗證）。
>
> **指定日期範圍**
>
> 例：「以 2024 年後發表的試驗為準。」減少過時資訊。
>
> **改問「我該問醫師什麼」**
>
> 把「我該怎麼治療」換成「我要去看醫師，該問哪些問題？」這類提問風險較低、AI 表現較穩。
>
> **分段問、不要一次塞太多脈絡**
>
> 一次給 10 個檢驗值 + 5 個用藥 + 3 個症狀，AI 容易出錯。一次一個重點。
>
> **核對關鍵數字**
>
> AI 列出的 ORR、mPFS、DOI、健保條件，**全部要在原始來源驗證**。

不同類型 AI 工具的差別

工具類型	範例	是否為醫療器材	用途
通用聊天 AI	ChatGPT、Claude、Gemini	否	衛教查詢、翻譯、整理問題
臨床決策支援系統（CDSS）	UpToDate、DynaMed、BMJ Best Practice	部分通過醫療器材認證	醫師臨床決策
影像 AI	Aidoc、Lunit、Mirai	多通過 FDA / TFDA	影像輔助診斷（醫師判讀）
病理 AI	PathAI、Paige	部分通過認證	病理輔助
語音轉病歷	Abridge、Nuance DAX	工具類	醫師效率工具

病人最常接觸的是通用聊天 AI——這是「**搜尋引擎升級版**」，不是醫療器材，沒有臨床責任。把它當圖書館員，不當醫師。

OpenEvidence 與其他「醫療 AI 搜尋引擎」

近年出現專為醫療設計的 AI 工具（如 OpenEvidence、Glass Health、Consensus）：

特色	評估
引用真實論文	較不易 hallucination，但仍需驗證
醫師為主要使用者	對病人不夠直觀
訓練資料偏歐美	台灣健保、亞洲族群數據較少
多為訂閱制	一般民眾不易使用

這類工具對醫師有用，對病人仍建議透過醫師轉述。

適用對象 / 不適用對象

本文適用

- 使用 ChatGPT / Claude / Gemini 查病情的病友與家屬
- 想了解 AI 風險的醫療同仁
- 第一線衛教師資

本文不適用

- 取代專業醫療諮詢

- 醫院 / 機構 AI 系統評估（需專業驗證）

結語：AI 是助手，不是醫師

該做	不該做
用 AI 整理「該問醫師的問題」	用 AI 決定「要不要做手術」
用 AI 翻譯英文衛教	用 AI 自我診斷急症
用 AI 練習對話、降低焦慮	用 AI 計算藥物劑量
用 AI 解釋醫學名詞	用 AI 判讀影像 / 病理
驗證 AI 引用的論文與數字	直接相信沒有來源的數字

參考文獻

1. Singhal K, et al. **Large language models encode clinical knowledge.** *Nature.* 2023;620(7972):172–180. doi:10.1038/s41586-023-06291-2
2. Goodman RS, et al. **Accuracy and Reliability of Chatbot Responses to Physician Questions.** *JAMA Netw Open.* 2023;6(10):e2336483. doi:10.1001/jamanetworkopen.2023.36483
3. Ayers JW, et al. **Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum.** *JAMA Intern Med.* 2023;183(6):589–596. doi:10.1001/jamainternmed.2023.1838
4. Topol EJ. **High-performance medicine: the convergence of human and artificial intelligence.** *Nat Med.* 2019;25(1):44–56. doi:10.1038/s41591-018-0300-7
5. Walters WP, et al. **Fabrication and errors in the bibliographic citations generated by ChatGPT.** *Sci Rep.* 2023;13:14045. doi:10.1038/s41598-023-41032-5
6. US Food and Drug Administration. **Artificial Intelligence and Machine Learning in Software as a Medical Device.** fda.gov/AI-medical-devices

> 引用整理協力：Singhal Nature 2023、Ayers JAMA Intern Med 2023、Goodman JAMA Netw Open 2023、Topol Nat Med 2019、FDA AI/ML SaMD 指引 (2026/05/12)。

SOURCE　https://lin.hsiehting.com/posts/2026/using-ai-for-medical-info/

CITATION　林協霆. AI 諮詢醫療資訊：什麼能信、什麼不能信. 林協霆 · 臨床筆記. 2026/05/12. doi:10.5281/zenodo.20131267

DISCLAIMER　本文整理公開發表之臨床試驗結果與 NCCN／ASCO／ESMO 治療指引，僅供醫學新知與病人衛生教育參考，不構成個別醫療建議，亦不取代主治醫師之診療判斷。實際治療決策請與您的主治團隊面對面討論。