健康資料與生成式AI研究倫理（四）

～臺大醫院倫理中心研究倫理研習會紀要

講者 / 臺大醫院倫理中心 蔡甫昌
整理 / 王劭慈

伍、生成式醫療 AI 之研究倫理考量：臺大醫院倫理中心蔡甫昌主任

生成式人工智慧（GAI）迅速進入醫療現場，成為臨床與科學研究日常中的重要工具。臺大醫院倫理中心主任蔡甫昌教授於本場講題中，以實務角度談AI應用的倫理風險與治理挑戰，亦回應與會者對「生成式AI是否會取代醫療人力」的焦慮。演講以多起國際案例與比喻帶出嚴肅議題，也從AI的演算法透明度、責任歸屬、資料偏見到可解釋性等面向，具體指出倫理審查的未來挑戰。

一、從「iPhone 時代」看生成式 AI 普及速度

蔡教授以輝達執行長黃仁勳談話為引，形容生成式AI已如智慧手機般普及，內建於日常生活與專業決策中。他分享Apple如何以「裝置端運算」、「私有雲環境」與「資料使用承諾」三層架構面對AI治安挑戰，也指出iOS與ChatGPT整合後，用戶可邊洗碗邊向AI提問、要求摘要、詢問醫學內容，甚至請其解析演講PPT。

在這樣的速度下，「AI無所不在」已成新常態。蔡教授以動漫中的閃電俠與量子電腦為例，說明生成式AI帶來的效率革命。無論是陪伴型機器人、手術輔助機器、家庭小教師，甚至外骨骼裝置，都已陸續投入醫療現場。這些科技讓人類執行原本難以達成的行動，未來在智慧醫療、尖端醫療與精準健康中都將是關鍵基礎。

二、生成式 AI 在醫療應用的倫理議題

面對這波變革，生成式 AI 的優勢與潛在衝擊成為討論焦點。蔡教授指出，GAI 具備通用性、模態整合能力、不需標註資料的自學能力與高效的模型訓練能力，使其成為科研與臨床中的強大助手。

但同時，AI 亦帶來以下六大倫理挑戰：

1. **隱私：** GAI 可能生成與原始資料相似內容，衍生再識別風險。AI 學會的資訊難以刪除，「被遺忘權」在生成式模型中難以落實。蔡教授幽默地說：「要 AI 忘記，它可能會回你：這個產品無法處理這個問題」，凸顯技術限制與法律保障之間的落差。
2. **同意：**病人是否知情其資料被用於 AI 模型訓練？模型若產出新的內容，是否屬於病人所有？是否應取得再同意？系統本身亦難以溯源判斷資料出處。
3. **二次使用：** GAI 透過「記憶」資料間關聯進行推理，若用戶希望退出模型訓練資料，執行的實際可行性與範圍仍不明確。
4. **偏見：**來自特定文化資料的偏見、模型設計缺陷可能造成錯誤決策與差別待遇。潛藏歧視尤其難以在開發階段被發現或修正。
5. **資料真實性：** AI 並非法律主體，若誤導診斷，醫師仍須負責。蔡教授指出，當 AI 表現優於人時，醫師面對採信與否都將進退兩難；即使懷疑 AI，也需舉證合理裁量。舉產檢為例，說明臨床判斷背後的倫理選擇與風險承擔，未來 AI 參與愈深，這類責任劃分問題將更複雜。生成式 AI 也讓「醫療深偽」(medical deepfake) 成為風險焦點。虛構病歷若結合真實語音或影像，將可能誤入臨床系統、誤導專業判斷，甚至引發詐欺、誹謗等法律與公共衛生問題。蔡教授也提醒，倫理與審查機制也須更新，「現在要問的，不只是你有沒有看病歷，而是你有沒有檢查 AI」。
6. **可解釋性：**傳統 AI 即有「黑箱問題」，生成式 AI 更難說明輸出理由。當 AI 建議與醫師直覺不同，若採信出錯，是誰負責？若拒絕 AI 建議而出錯，又如何舉證合理性？這些將是法律與倫理的重要爭點。

三、制度與法規：從歐盟到臺灣，走向分級管理

面對 AI 風險，國際法規已開始分類管理。蔡教授指出，歐盟《人工智慧法案》(AI Act) 將 AI 系統依風險程度分為四級：

- **不可接受風險 (Unacceptable Risk) ：**如即時人臉辨識、社會評分等用途，因涉及基本人權侵害，被列為全面禁止。
- **高風險系統 (High Risk) ：**應用於教育、醫療、司法、基礎建設等領域，須接受嚴格檢驗與合規機制。

- **有限風險（Limited Risk）**：如聊天機器人或生成式 AI，需告知使用者正在與 AI 互動，保障資訊透明。
- **極小風險（Minimal Risk）**：如過濾垃圾郵件、自動分類文件等，可自由使用，無須額外規範。

此一風險分級邏輯，未來可能影響台灣在倫理審查、研究規範乃至行政應用上的參考依據。蔡教授也提醒，AI 風險管理不該只停留在總則式宣示，而需落實在審查、使用到部署過程中的實際分級與責任機制。

臺灣則可參考《醫用軟體分類分級參考指引》，將 AI 工具視為醫療器材進行分級與審查。蔡教授指出，未來醫療 AI 若欲「落地」應用，勢必會被視為醫療器材或技術，須接受衛福部食藥署管理。這類 AI 工具只要具備資料收集、儲存、分析、轉換、顯示等功能，均在規範之列。即便是生成病歷草稿，亦屬醫療文書，依《醫師法》仍應由醫師親自簽署並負法律責任，不能因為 AI 參與而模糊責任歸屬。

進一步進入臨床使用階段時，蔡教授提醒應審慎思考幾個關鍵問題：

- 醫師與病人是否都知情 AI 已參與診斷或溝通？是否已妥善解釋使用情境？
- 此類應用是否會進一步加劇健康不平等（health disparity）？其價格與普及程度是否會限縮部分病人使用？
- 訓練資料是否具代表性、多樣性？偏倚風險如何控制？
- 有無足夠資安與濫用防護機制？ GAI 若遭惡意使用，是否會擴大資訊戰或商業競爭下的資料操控？

蔡教授舉例：「將來戴著耳機行醫，AI 隨時提供你應該怎麼回答病人。病人可能會說：耳機借我，我自己聽就好了。」這段幽默舉例背後其實是嚴肅提醒： AI 能否輔助醫療判斷並非問題，關鍵是病人是否充分知情、有無選擇權，資訊是否足夠透明。

GAI 若參與診斷，是否應建立記錄機制？可否標示 AI 參與比例與角色？這些機制都關係到未來的倫理審查、法律追溯與病患信任。GAI 或可輔助知情同意流程，讓病人更清楚理解資訊來源與風險，但前提是制度與技術配套要跟得上。

四、結語：建立可持續的監理框架，回應醫療 AI 挑戰

生成式 AI 已深刻改變醫療場域的運作樣貌，不僅能提升效率、強化溝通，也為醫療平等帶來新契機。幻覺風險與運作不透明等挑戰，仍需制度機制持續監督與

修正。蔡教授也提到既然 GAI 被定位為醫用軟體，其審查與使用規範應比照醫療裝置嚴謹，從信賴性、隱私保障到問責與資料治理，都須有與時俱進的對應標準。

醫療 AI 的治理，不能僅止於研發與應用，而是需要一套能長期運作、持續追蹤與評估的倫理與法規框架。

(感謝講師授權記錄演講內容與季刊刊載，著作權與智慧財產權歸屬講師本人)