健康資料與生成式AI研究倫理（二）

～臺大醫院倫理中心研究倫理研習會紀要

講者 / 國立陽明交通大學生物醫學資訊研究所 吳俊穎
整理 / 臺大醫院倫理中心 王劭慈

2024 年 12 月 13 日，臺大醫院於兒醫大樓 B1 講堂舉辦「智慧醫療與資料科學研究倫理與法律」研究倫理研習會。邀請多位資料治理與法制領域之專家學者，深入剖析「健康資料應用爭議」及「生成式 AI 倫理」兩大主題。在前一期季刊（2025 年 6 月第 6 期），我們已將「健保資料庫研究利用爭議與全民健康保險資料管理條例草案評析」進行摘要整理，本期將針對「人工智慧風險評估架構：由歐盟人工智慧法案談起」、「生成式 AI 應用與挑戰」二項議題進行整理，業經講者確認內容，提供給讀者參考。

參、人工智慧風險評估架構：由歐盟人工智慧法案談起：國立陽明交通大學生物醫學資訊研究所吳俊穎教授

這場演講由吳俊穎教授主講，吳教授在開場時先簡單介紹了本次內容的架構。整場演講將圍繞以下幾個核心主題展開：人工智慧與傳統統計方法的差異、人工智慧可能帶來的偏差與風險來源、歐盟人工智慧法案（AI Act）的風險分層與法規要求、智慧醫療在法律責任上的挑戰與責任歸屬問題。

吳教授強調，這些議題彼此緊密相關，從技術到倫理、從資料到法律，將完整呈現 AI 在醫療應用中面臨的機會與風險。

一、從傳統統計到人工智慧：兩種分析思維的落差

吳俊穎教授首先談到，傳統統計在高維度資料分析上有明顯限制。當病例數有限、資料維度卻非常龐大時，統計方法往往無法勝任，例如醫療影像、病理報告、自由文字的臨床紀錄、甚至穿戴裝置蒐集的生理訊號。

人工智慧模型的出現，讓分析可以同時處理數萬甚至數十萬個特徵，並自動完

成特徵篩選（feature extraction），找出關鍵變項，建立預測模型。AI 不受限於線性假設，能發現複雜非線性的關係，這是傳統統計難以達成的。

但這種優勢也同時帶來挑戰，AI 模型常被批評為「黑箱」，缺乏統計模型常見的假設檢驗，難以解釋模型為何做出某個判斷。這使得「可解釋的 AI」（Explainable AI）成為近年重要議題，不僅要求模型結構與結果需可被專業人員理解，連資料來源、訓練過程、以及預測邏輯都必須保持透明，確保後續能進行偏差修正與責任追溯。

二、AI 偏差的來源與挑戰

演講中特別強調，AI 偏差（bias）可能來自三個層面：

1. 資料本身的偏差：
   - 醫院資料庫可能存在族群分布不均，例如男性樣本多於女性，導致模型在女性族群的表現較差。
   - 常見的還有資料不完整（missing data），如有些病人沒做特定檢查，這種缺漏本身就帶有選擇性。
2. 模型的限制：

   不同的演算法在分類、生成或回歸任務上各有優缺點，若選擇不當，會放大原始資料的偏差。
3. 使用與解釋過程的偏差：

   從定義研究問題、選擇資料來源，到實際部署 AI 系統，每個環節都可能引入新的不確定性。

吳教授提到，有些研究用生成式 AI 來補齊不平衡或缺漏的資料，但若原始資料異質性過高，生成的資料不一定能真正反映真實情況，甚至可能放大誤差。他舉了一個例子：若模型大多數影像來自女性廚師，系統可能會自動把「穿紅衣做料理的人」預測為女性。這種因資料分布不均導致的「AI artifact」，正是歐盟 AI 法案高度關注的風險來源。

三、歐盟人工智慧法案的風險分層

2024 年 5 月 21 日，歐盟正式通過 AI Act，針對 AI 應用風險建立分層架構，分為四級：

1. 不可接受風險

- 完全禁止的 AI，例如利用潛意識操縱、社會信用評分、大規模即時人臉監控等。
- 法律明確列出三個判斷要件：主觀意圖（惡意操縱）、客觀手段（欺騙或操控）、實際危害（對身體、心理或財產造成重大損害），三者缺一不可。

2. 高風險

- 涉及公共安全、醫療器材、關鍵基礎建設、司法與移民等領域的 AI 系統。
- 凡屬高風險類別，必須完成嚴格的取證、資料治理、透明度與安全性要求，並列入歐盟官方登錄。

3. 有限風險

   包括大型生成式 AI、開源模型等，必須滿足透明度與資訊揭露義務，如公開訓練資料來源、告知使用者系統能力與限制，並建立對抗性測試與安全機制。

4. 極低風險

   例如一般聊天機器人或遊戲應用，歐盟僅鼓勵業界自律，無強制規範。

吳教授特別提醒，高風險分類下的醫療器材 AI，無論用於診斷、治療或病歷管理，都需要符合歐盟醫療器材法（MDR，IVDR）規範，這與美國 FDA 對醫療軟體分級管理的做法相呼應。

四、智慧醫療中的法律責任

AI 在醫療中的使用，讓責任歸屬變得複雜：

1. 產品責任：若 AI 系統本身有缺陷，製造商需負責。但根據台灣《消費者保護法》第 7 條，舉證責任可能會轉移，由製造商證明其無過失。
2. 醫師責任：若醫師在使用 AI 後仍做出錯誤診斷，可能被認定為醫療疏失，就可能產生醫療糾紛。
3. 醫療機構責任：醫院作為 AI 導入者，若系統維護或管理不足，也可能承擔部分責任。

吳教授分享一個有趣的研究：

- 醫師單獨診斷的正確率約 74%，
- AI 單獨診斷可達 92%，

- 但醫師與 AI 共同診斷時，錯誤率反而上升到 24%。

這帶出一個關鍵問題：當 AI 與醫師共同參與診斷時，錯誤究竟該由誰負責？實務上可能出現三種情境：

1. AI 正確、醫師改錯，屬於醫療疏失，責任在醫師；
2. AI 出錯、醫師未發現，則涉及產品責任與醫療責任交疊；
3. 醫院作為醫療契約主體，理論上應承擔一定責任，但在台灣醫療訴訟裡，醫師個人被告的風險仍遠高於醫院。

根據《消費者保護法》第 7 條，AI 產品的製造商必須舉證自己無過失，然而在實務上，醫師仍可能因診斷錯誤而成為主要被告。這形成一種道德風險：隨著 AI 正確率愈來愈高，醫師可能出現「反正 AI 做決策，我就照蓋章」的心態，導致病患安全反而被忽視。責任的分配與法律制度的調整，將是智慧醫療發展中最需要回應的議題之一。

在臺灣，醫療糾紛仍以刑事訴訟為主，醫師被告的風險高於醫院，未來若沒有新的立法調整，AI 導入後的責任分擔仍將充滿爭議。

五、結語： AI 時代的適應與治理

吳教授最後引用達爾文的名言：能存活下來的物種，並非最強或最聰明的，而是最能適應環境的。

AI 正在重塑醫療生態，從資料治理、風險分層到法律責任，都需要新的制度與倫理規範。未來的關鍵不僅在於技術進步，更在於如何在創新與安全之間取得平衡，讓 AI 真正成為可信任的醫療夥伴。

肆、中場綜合討論

中場綜合討論時間，與會者首先關注 AI 在醫療應用下是否會增加犯錯風險。吳俊穎教授指出，並非犯錯的機會增加，而是 AI 提升了錯誤被發現的機率。過去醫療錯誤不易被檢視，如今 AI 診斷準確度更高，可能成為新的標準，讓醫療教育與法律責任必須同步調整，以因應社會對錯誤率降低的期待。

在法規部分，有人詢問美國與歐盟資料規範的差異。回應指出，美國只要經專家認定完成去識別化，資料便可自由使用；歐盟限制較嚴，台灣若要跟進，需先建立明確的去識別化標準與審查機制。

針對 AI 風險分級與「動態同意」的關係，吳俊穎教授認為，動態同意的要求

過於繁複，實務上難以落實。重點應放在資訊提供的充分性與受試者理解，而非簽署次數。

最後，主持人提到，目前 IRB 審查已推動在同意書首頁提供試驗關鍵資訊（Key Information），讓受試者更快掌握試驗及研究重點，提升知情同意的實質意義。

（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）