

# 人工智慧/機器學習技術之醫療器材軟體 上市前審查要求及臨床試驗相關規範

## ～臺大醫院倫理中心研究倫理研習會紀要（二）

### Pre-market Review Requirements and Clinical Trial Regulations for Artificial Intelligence / Machine Learning-Based Medical Device Software

#### ～Minutes of the Research Ethics Workshop at the Ethics Center of National Taiwan University Hospital(II)

講者 / 衛生福利部醫療器材及化粧品組 葉旭輝

整理 / 臺大醫院倫理中心 王劭慈

## 貳、人工智慧 / 機器學習技術之醫療器材軟體上市前審查要求及臨床試驗 相關規範：衛生福利部醫療器材及化粧品組葉旭輝科長

### 一、 醫療器材全生命週期管理架構

葉科長首先說明，臺灣醫療器材管理以《醫療器材管理法》為母法，AI 醫療器材相關指引屬行政規則層級。食藥署在產品全生命週期各階段均有對應服務：上市前包含屬性判定、GLP 臨床前測試規範、TFDA 與 IRB 雙軌臨床試驗審查及 GCP 查核；上市後則涵蓋不良事件（ADR）通報、各國警訊收集、安全監視及 UDI 來源流向管理等。

### 二、 AI 醫材查驗登記審查三大要項

依《醫療器材分類分級管理辦法》，目前核定的 AI 機器學習醫療器材均屬第二等級，查驗登記審查要項分為產品技術特點、臨床前測試與臨床證據三大類，且各項內容必須與中文說明書的宣稱一致，這是貫穿整個審查過程的核心原則。

#### (一) 產品技術特點：

AI 醫材的技術資料重點在於演算法架構，而非一般醫材著重的材質與生物相容性。申請者須說明演算法設計，包含演算法名稱、開發平台、函式庫及模型層數與權重等結構細節。

訓練資料的說明同樣重要。資料集須分為建立模組、訓練、驗證評估三組且不得重複，每組均需交代資料來源、納入排除標準、病患分布，且案件數量須達到統計檢定力。資料的產製方式也須明確描述，包含使用何種設備（如 CT、MRI、超音波或病理玻片影像）、儀器型號及取像參數——例如胸部 X 光的 PA 位與 AP 位對 AI 而言是截然不同的輸入，若未事先定義清楚，判讀結果可能大相逕庭。基準真相（Ground Truth）的建立方式亦須交代，例如以活體組織切片或臨床醫師判讀為標準；若採醫師判讀，建議由二至三名醫師聯合完成，並說明結果不一致時的處理程序。

產品形式目前分為電腦輔助偵測（CADe）、電腦輔助診斷（CADx）及電腦輔助篩檢（Computer Aided Triage）三種，各有對應的輸出宣稱要求。搭配特定硬體（如 CT）時，亦須詳述影像取得參數與切片厚度，規格不一致可能直接影響判讀結果。目前核准產品多為閉鎖式（locked）演算法，但食藥署已建立「軟體預定變更控制」制度，允許業者在事先申報變更控制計畫的前提下進行迭代更版。

## (二) 臨床前測試：

臨床前測試涵蓋軟體確效（參考《醫療器材軟體確效指引》及 IEC 62304）、網路安全評估（若具連網功能），以及功能性驗證。功能性驗證的核心是確認產品宣稱的輸出結果，如疾病風險高低、是否建議轉診、評分等級、靈敏度與特异性等，與中文說明書的內容相符。葉科長也指出，若試驗完成後發現說明書宣稱與結果不完全吻合，在可行範圍內可回頭調整說明書宣告。

## (三) 臨床證據：獨立效能評估與臨床性能評估：

臨床證據可採兩種路徑。獨立性能評估（Standalone Performance Assessment）以既有標記資料直接測試 AI 輸出，不需新增醫事人員判讀，適用於已有類似品上市的情境。若新產品與已核准類似品以相同資料集評估，且靈敏度或偽陽性率表現不劣於類似品，則可免除臨床性能評估。臨床性能評估則可進一步比較有無 AI 輔助時醫師判讀表現的差異，例如是否提升診斷率或縮短診斷時間，可使用前瞻性或回溯性資料，適用於全新產品或需證明臨床效益的情境。

撰寫研究計畫時，研究族群與納排條件均須與中文說明書一致，若使用非國內資料，須收納部分國內資料確認適用性，若採多中心或聯邦式學習，須說明資料整合方式。判讀方法原則上應與臨床實務一致，若有所不同，須能以資料說服參與審查的臨床醫師，確認該設計能有效驗證產品的臨床效能。效能評估指標的 cutoff 值可依實際

應用情境討論，例如某些 AI 產品的整體效能或許不如醫師，但若能在傷患分類場景發揮有效篩選作用，仍有其臨床價值。

### 三、 臨床試驗管理：無顯著風險的三種態樣

依《醫療器材管理法》第 37 條，臨床試驗原則上須申請食藥署核准。但屬「無顯著風險」者，只需 IRB 核准即可執行，無需另行申請主管機關核准，可有效簡化行政程序。葉科長指出，目前 AI 醫材試驗大多屬此類。三種態樣分別為：

- (一) 態樣一：已取得許可證且試驗未超出核准範圍。
- (二) 態樣二：以合法取得的受試者檢體或資料作為診斷試驗客體，且結果不作為臨床診斷依據——例如以常規方式收集的影像或血液檢體供實驗性 AI 分析，或引用健保資料庫資料。
- (三) 態樣三：未具游離輻射、置於受試者體表或無須接觸體表的器材（如心電圖儀等），且試驗結果不作為臨床診斷依據。

### 四、 案例探討與審查視角的啟示

葉科長以一款氣胸偵測系統為例說明實際審查要點。該系統利用 AI 分析胸部 X 光影像偵測氣胸特徵，結果以警訊方式呈現，僅適用 18 歲以上成人，須由臨床醫事人員確認。審查重點包含訓練族群的特徵描述（如是否曾手術、使用的 CT 廠牌等）、軟體確效文件的完整性，以及效能評估的設計，含影像來源格式（CR 與 DR）、參與驗證的醫事人員數量與資格。該案例亦以美國影像與臺灣影像進行種族差異比對，確認兩者判讀結果差距不顯著，從而說明訓練模型可一體適用。

演講最末，葉科長以 2017 年搞笑諾貝爾獎得主法丁先生的論文〈貓咪的流變學〉為引，說明審查標準的本質。該論文以物理學的底波拉數來判斷物質狀態：觀測時間內有形變就是液體，沒有形變就是固體，關鍵在於觀測時間的長短，喜馬拉雅山從地質時間尺度看也是流體，只是人類壽命太短無從觀察。以此為喻，葉科長說明臨床試驗是否可被接受，端視同儕審查的時代性共識而定，沒有絕對標準。食藥署扮演申請者與審查委員之間的溝通橋梁，遇到科學標準隨時代演進而位移的情形，有時須透過補充資料等方式共同協調解決。葉科長也坦言，試驗目的各有不同，某些疾病的資料收集本就困難，這些都還有探討的空間。

（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）