

State of the Art in Data Science Research Ethics and Law (中)

～臺大醫院倫理中心研究倫理研習會紀要

整理 / 臺大醫院倫理中心 王劭慈、戴君芳

2024 年 11 月 1 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」研習會，邀請多位在醫療倫理及法律領域的重量級專家學者，就「資料法制的回顧與前瞻」、「健康資料研究國際趨勢」、「資料研究倫理與公民參與義務」、「健康臺灣引領大數據應用之契機與挑戰」、「生成式 AI 在醫療機構之隱私權保護及倫理議題」等議題進行研討。在前一期季刊 (2024 年 12 月第 4 期)，我們已摘要前二項講題之精彩內容，本期將針對「健康臺灣引領大數據應用之契機與挑戰」、「生成式 AI 在醫療機構之隱私權保護及倫理議題」二項議題進行整理，業經講者確認內容，提供給讀者參考。

肆、健康臺灣引領大數據應用之契機與挑戰：演講者國家衛生研究院群體健康科學研究所特聘研究員兼所長 邱弘毅教授

一、現今公共衛生概況

臺灣精準健康方案的推動與挑戰

臺灣目前正處於人口結構劇變的時代，面臨高齡化與少子化的雙重壓力，醫療資源與健保體系的負擔日益加重。根據國發會的數據，臺灣將於 2025 年進入超高齡社會，老年人口將突破 20%，2050 年達到最高峰，在未來 20 年內攀升至 40% 以上，其中一半以上將是 80 歲以上的長者。這意味著醫療資源的需求將大幅提升，如何有效運用現有資源，推動「精準健康」與「精準醫療」成為當務之急。

政府在政策上逐步推動健康大數據的建構與應用，試圖透過資料整合來改善醫療服務效率及促進產業創新。目前現有的資料多分散在不同單位與平台，格式標準化成為資料整合的難題。例如健保資料庫雖然龐大，但資料主要來自財務給付申

報，缺少常見穿戴裝置紀錄的生活習慣、基因與環境等更全面的數據，必須透過其他資料庫加以補強。未來的理想目標是將所有健康數據上傳至雲端，並透過 AI 模型進行風險預測分析。舉例來說，醫師可透過 AI 運算提供病人未來 5 到 10 年的健康風險預測，並視覺化呈現，讓病人更容易理解與配合治療。這樣的模式雖然會增加醫師的工作量，但健保署已開始考慮調整健保給付制度，在修法後將健康促進納入支付範疇，鼓勵醫師進行預防醫療。

二、精準健康方案及數據推動現況

資料整合的必要性與現況及產業應用

面對臺灣高齡化與少子化帶來的衝擊，政府近年來積極推動「健康臺灣」計畫，並透過「精準健康方案」來應對未來醫療資源短缺的困境。現今的資料收集來源多元，從各醫院、國家單位到私人機構，都在不同層面收集健康數據。然而這些資料格式不一致，導致整合難度高，限制了資料的有效應用與價值發揮。

為了解決這一問題，臺灣逐步推動資料格式標準化，建構統一的數據平台。NBCD 平台就將全臺醫學中心的人體生物資料庫進行格式一致化，讓不同機構之間的資料能夠快速交換與整合。當研究者提出研究案後，如臺大或北榮等機構有興趣參加即可快速上傳資料，透過自動轉換功能，迅速完成跨機構的資料串聯與研究合作。目前包括 FIRE 標準與癌症資料 MCO 格式等技術框架，已逐漸成形並投入應用。

在資料整合的基礎上，臺灣也逐步展開產學合作，推動健康與醫療開發。例如羅氏與國衛院合作，使用臺灣的 2000 個檢體進行基因定序，研發新的藥物反應標靶。西方人與東方人在基因組上的差異，使得本土資料的重要性不容忽視，這些資料對於藥物開發與臨床應用具有關鍵價值。

為了滿足不同需求，未來數據平台將提供主題式資料庫與 Raw Data 服務。研究能力較強的機構可進行在地分析，而資源較有限的學校或企業，則能透過套餐式資料庫取得完整數據集，迅速展開研究。

法規挑戰與治理架構

臺灣在推動精準健康方案時，仍面臨法規限制與治理架構的不確定性。尤其健

保資料庫的合規應用常受到質疑，反對方認為這些數據的使用目的與健保設立初衷相違背。另外資料治理機制也需明確化：究竟由行政院或部會層級負責？如何建立資料共享的單一窗口？這些都是需再釐清的問題。

國際趨勢與臺灣的下一步

歐美國家已明確設定資料應用的長遠目標。例如歐盟推動 Single Market for Data 計畫，強化資料經濟效益，並逐步提升數據應用的專業人才數量。美國則透過明確的十年願景，從基礎建設到資料治理，逐步邁向資料驅動的決策模式。

相較之下，臺灣雖已有資料平台與標準化的初步進展，但在基礎建設與治理規劃上仍存在落差。2024 年生技產業策略諮議委員會重點希望發展結合民間力量及各大公私立醫學中心成立健康數據服務公司，整合來自各種來源的數據，確保不同健康數據之間的互操作性，並加速去連結後的數據應用於產業，上述目標若無法突破目前法規與治理障礙，臺灣將難以實現真正大數據的實證決策，亦難以與國際接軌，推動大數據產業的全面發展。

三、臺灣健康大數據現況及應用實例

現有數據平台的應用與挑戰

目前臺灣擁有多個健康數據來源，包括健保資料庫、疾病診斷數據、實驗室數據以及醫學影像資料等。健保署已開放四大癌症（肺癌、大腸癌、乳癌、肝癌）的主題式資料庫，並計劃於明年新增心血管疾病及糖尿病的主題式資料，這些資料庫均諮詢相關疾病專家，將重要資料項目放入，將可包含更高質量的實驗室數據。

數據應用仍面臨的問題，包含醫學影像的數據格式在不同醫院間存在差異，影響整合與解讀，需持續突破。個人穿戴式裝置雖然能提供大量數據，但因用戶行為不一致（如忘記充電或佩戴），導致數據中斷，需開發統計模型以補足遺漏資料，確保數據連續性。

衛福體系外之部會資料的整合

除了醫療數據，外部資料如環境污染、社會經濟背景、自殺風險等，也對健康研究至關重要。目前政府正努力與環保署、法務部、教育部等單位合作，將空污監測數據、犯罪及自殺相關資料及教育部體育署掌握 18 歲以下學童健康資料串接到

健康大數據平台。舉例來說，針對急診病摘資料萃取臨床端方具有的家暴與自殺高風險個案資料，透過數據串聯及分析，以進行更有效識別與預防措施。

在社會弱勢群體的支持方面，目前有超過十萬高風險家庭納入定期關懷訪視。但傳統訪查需耗費大量人力與時間，未能有效應對突發事件。此時利用 AI 與數據分析，優先識別最急迫的個案就成為未來發展的重點。例如預測高風險自殺個案的模型，能提前 20 天識別潛在風險，幫助社工優先介入協助。

臺灣現正推動健康大數據平台計畫，整合健保、癌症次世代基因定序、罕見疾病資料等。每年由國科會支持的醫院項目，也收集了約 2000 例電子健康紀錄、實驗室數據和基因數據，這些資料將進一步串接到平台中，提供研究者使用。國衛院運用糖尿病風險預測模型，成功識別出 10% 的隱性糖尿病患者，避免延誤診斷導致併發症進一步惡化。未來研究人員計劃加入基因分數，提升風險預測的精準度，進一步推動精準醫療的發展。

四、健康大數據平台合規使用之挑戰

健保資料的合規使用與挑戰

過去十年，臺灣透過健保資料庫進行大規模研究，累積超過一萬篇學術論文。隨著個資法的規範不足與法律挑戰，現行的資料使用制度正面臨壓力。憲法法庭要求，未來健保資料的使用需滿足三個要件：建立獨立監督機制、明確法規框架，以及提供資料退出機制。退出機制的設計必須兼顧實務需求與研究連續性，否則將影響數據的完整性與研究效益。為解決此問題，政府與學界建議應設立公正的第三方監督機構，並明確規範資料的二次利用範圍與限制。建立社會信任與透明機制也至關重要，讓民眾了解資料使用的公益價值，並消除隱私洩露的疑慮。未來若能結合完善的治理架構與法規配套，臺灣在健康大數據的應用上將更具國際競爭力。

資料治理與跨部會合作

目前臺灣資料整合面臨兩大挑戰：跨部會協作困難與產業應用受限。各部門擁有不同資料庫，但資料串接受限於法規與管理體制。健保資料僅限於學研機構申請，產業界難以直接參與。未來應建立中介機構，負責資料的媒合、監控及公益回饋，促進資料共享與有效利用。

同時資料提供者與使用者之間需建立透明的信任機制，讓民眾了解資料的公共利益，而非過度擔憂個資洩漏。目前學研機構對於資料管理已十分謹慎，並無資料外洩案例發生，這也證明現有制度在技術層面具備安全性。

資料格式標準化與人工智慧發展

資料的標準化是推動大數據應用的基礎。未來政府應加速採用國際標準，如美國的 CDI 流程及 FHIR 格式，確保數據在正確的時間、渠道與格式下流通，實現「Right Format, Right Channel, Right Time」。

人工智慧的風險評估是未來發展的重點。隨著 AI 技術快速進展，歐盟已率先制定風險評估框架，確保 AI 應用在醫療領域的安全性與有效性。臺灣也應加速跟進，推動 AI 技術在醫療數據中的合理應用，實現精準健康與精準醫療的目標。(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)

伍、生成式 AI 在醫療機構之隱私權保護及倫理議題：演講者陽明交通大學科技法律學院 陳鈺雄院長

醫療機構內生成式 AI 的多元應用

醫療機構內生成式 AI (Generative AI) 的應用正在迅速擴展，臨床文件處理是其最常見的應用之一，如進行病歷摘要，減輕臨床人員的行政負擔。最早，IBM Watson 就利用自然語言處理技術來分析癌症資料，這類技術如今已被廣泛應用於醫學研究與藥物開發。此外，診斷支援系統可分析大量醫療資料，輔助處理複雜病例，並透過比對廣泛的資料庫提供建議。

生成式 AI 還可協助與病患溝通，透過通訊軟體提供數位內容，讓病患更清楚了解治療過程，並增強參與感。健康監控功能則能透過解析穿戴裝置數據，提供健康預測與建議，提升預防醫學的效率。而上述臨床的應用，因不屬於治療或診斷，不一定需要經過主管機關的審查，因此較無明確之規範，是目前所面臨較大的挑戰。

使用生成式 AI 的潛在挑戰與風險

儘管應用廣泛，生成式 AI 的使用同時帶來隱私與倫理挑戰，值得各界關注。生成式 AI 在臨床決策中也可能導致風險，例如，在病歷摘要的生成過程中，如有

錯誤可能對診斷及治療結果產生影響。以 IBM Watson Health 為例，其定位於類似圖書館檢索系統進行資料檢索，而不被認為是醫療器材，但若生成的摘要不夠準確，可能間接干擾臨床判斷。有國際實務經驗顯示，部分 AI 在整理病歷時可能發生內容順序顛倒的問題，影響醫師對病患病情的掌握，進而影響臨床決策。

生成式 AI 在不同階段應用有不同風險考量，分為資料產生階段 (data collecting phase)、開發階段 (development phase) 及部署階段 (deployment phase)。

資料產生階段之隱私議題是關鍵挑戰，此階段需使用大量數據資料，完全匿名化的數據難以連結不同資料庫，因此通常採取假名化處理。匿名化包括「技術上不得再識別」以及「法律上不得再識別」兩種情形，例如健保資料庫限制在封閉環境中，僅允許查看結果而無法接觸原始個人資料，也無法將資料帶走，這樣的運作模式，配合使用情境，該環境可能可以在歐美國家的法律上被視為等同於匿名化。而法律上的匿名化常透過條件約束，限制還原數據的可能性。例如美國 HIPAA 提到的限制數據集 (Limited Dataset)，允許在特定條件簽署規範下免除病患同意。而對於醫院內生成式 AI 的應用，IRB 是否能免除受試者同意仍具爭議。目前衛福部關於免除同意的規範，多是針對學術研究用途，但這些系統不僅服務於學術研究，未來還可能涉及商業化，因此能否適用現行標準由 IRB 免除病患同意仍需進一步討論。

美國、歐盟與其他國家正在積極推動相關立法，使資料較能提供廣泛之運用。例如，歐盟的 AI 法規考量了個資保護與偏誤校正之平衡，亦即面對高風險 AI 系統時，若僅使用匿名數據無法有效矯正偏誤 (bias) 的可能時，該法規允許在滿足假名化、資安的要求下，可保留使用特種個資的豁免規定。

在模型開發階段，多仰賴廠商之算力，由於涉及病人資料，且可能涉及跨境傳輸，由於醫院機構面對實力極大之廠商巨擘，恐無法掌握其資料使用約定，宜有政府之介入，例如歐盟規範即要求廠商透明化，或美國要求廠商必須提供內部安全性資料，使雙方之關係較為平衡。

在部署階段，生成式 AI 面臨的最大挑戰是資訊安全。如日本的 Line 資料外洩事件與英國遠距醫療公司 Babylon 的病歷洩漏事件，都警示了 AI 系統的潛在風

險。因此若類此 AI 不需衛福部許可時，須遵循之規範為何?建議可朝將某些生程式 AI 應用納入醫療裝置的範圍、將舉證責任置於發展 AI 機構，或建立準核准機制三個方式。本國衛福部今年執行之三大 AI 中心之一：負責任 AI 執行中心，即是朝此方向規劃，在藥證機關之外，建立醫療場域 AI 治理之標準，進行文件化，並對不良反應事件加以檢視，確認是否 AI 應用所產生，以及定期檢視，以決定改進或下架。

展望與結論

在醫院中採用生成式人工智慧，對於改變醫療照護有著無窮的潛力。然而，它也會引起重大的隱私權和安全問題，必須透過健全的法規框架和創新的隱私權強化技術來解決。政府、開發者、部署者（醫院）等利害關係人應該意識到這些風險，並共同努力在臨床實務中保護病人。而教育訓練是推動生成式 AI 安全應用的關鍵，醫療從業人員需要深入了解 AI 風險與應對策略，確保臨床決策的正確性，即可降低 AI 生成錯誤資訊的可能性。(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)