



星球永續健康線上直播

可解釋 AI (XAI) 醫療應用

XAI 臨床導入與心血管疾病應用實例

2025 年 12 月 17 日

可解釋人工智慧（Explainable AI，XAI）在目前臨床醫療與健康照護領域導入 AI 應用階段具關鍵角色。XAI 不僅可呈現 AI 預測之支持基礎與運作邏輯，更提供醫療找護人員評估解釋內容是否合理、是否符合臨床推論，有助於 AI 成為醫師進行臨床判斷時可被信任的決策支援工具。本週我們將聚焦於 XAI 醫療應用的臨床導入指引，並透過 XAI 醫療輔助實例的心電圖（EKG）生理年齡預測解析，說明其在實務中的應用與意涵。

星球健康新知

英、法、德整合對烏支援強化安全應對：「未雨綢繆，另起爐灶」

歐洲與英國近期積極尋求俄烏和平進程突破。戰事進入長期消耗，歐洲主要國家體認僅依賴既有軍事與人道援助已不足以支撐烏克蘭與東歐未來的防務重建需求，因此近期推動動用自俄羅斯全面入侵烏克蘭以來，在歐洲金融體系中遭凍結的俄羅斯主權資產，為烏克蘭提供國防與重建實質資金來源。這批資產總額高達數千億歐元，其中多數集中於歐盟境內，但歐盟內部意見分歧，對於此策略是否造成歐洲經濟誠信與後續影響多所爭論。英國政府率先推進此凍結資金應用規費以及法源並於近期取的進展，相關談判已進入最後階段，若能達成協議，將可運用高達近千億英鎊的資產價值支援烏克蘭的軍事需求與戰後重建。這項討論的政治核心出現在倫敦唐寧街舉行的高層會談。英國首相施凱爾接待了烏克蘭總統澤倫斯基，並與法國總統馬克宏、德國總理梅爾茨進行密集協商，並擴大為與多個歐洲國家、北約與歐盟高層的跨國通話。會談中，各國領袖一致認為，當前正是必須同步加強對烏克蘭支持與對俄羅斯經濟施壓的時刻，否則任何停火都可能淪為短暫且脆弱的政治安排。歐洲方面特別強調，和平若要成立，必須建立在明確的安全保障之上，而非迫使烏克蘭以領土讓步換取暫時停火。然而，將凍結資產轉化為實際



資金的過程並非沒有阻力。部分國家，特別是掌管大量俄羅斯資產託管機構的比利時，對法律責任、金融穩定與潛在報復風險抱持高度戒心。即便如此，英、法、德等國仍持續施壓，希望透過「賠償貸款」的設計，在不直接沒收資產所有權的前提下，先行將資金導入烏克蘭，未來再以俄羅斯須支付的戰爭賠償作為償還來源。這項構想本身就象徵歐洲正逐步調整對國際法、制裁工具與戰後責任的理解方式。在歐洲安全環境急遽惡化的同時，來自俄羅斯及其盟友的「混合式威脅」也在歐洲邊境不斷浮現。立陶宛近期宣布進入緊急狀態，起因並非傳統軍事入侵，而是來自白俄羅斯的氣球反覆侵入其領空。這些氣球多半被用來走私香菸，但其高度、飛行路徑與不受控特性，已多次迫使立陶宛主要機場關閉，造成民航中斷與公共安全風險。立陶宛政府明確指出，這類行動並非單純走私，而是經過設計的干擾行為，屬於針對北約與歐盟成員國的混合攻擊。歐盟層級亦迅速回應，將這些氣球事件定性為不可接受的安全挑釁，並研議進一步制裁白俄羅斯。歐洲對安全威脅的定義已不再侷限於坦克、飛彈與部隊，而涵蓋資訊干擾、邊境施壓、灰色地帶行動等多元形式。對波羅的海國家而言，這些事件進一步加深了對俄羅斯及其盟友戰略意圖的疑慮，也強化了要求北約與歐盟提供更具體防衛承諾的政治動能。英國近期也調整防衛布局。英國皇家海軍公布新的大西洋安全戰略核心目標是保護海底電纜、能源管線與關鍵基礎設施，這些設施為國家經濟與通訊命脈。英國政府指出，近年來俄羅斯潛艦與科研船隻在英國周邊海域活動顯著增加，甚至出現以雷射鎖定監視飛機的危險行為。國會安全委員會早已警告，若海底電纜遭到破壞，其影響等同於大規模攻擊。英國發展無人潛航器、感測器與自動化系統戰略，運用人工智慧技術維持對廣大海域監控能力。英國國防高層認為，這是面對快速變化威脅時不可迴避的轉型方向，也反映出歐洲國家普遍面臨的科技資源與安全壓力。這些外交、經濟與軍事層面的發展最終交會於同一個核心問題：歐洲是否能在美國主導的和平進程之外，建立俄國以及東歐防務立場。面對近期美國主導對基輔施加的談判壓力，歐洲領袖一方面不願與美國正面衝突，另一方面也清楚意識到，若烏克蘭在缺乏安全保障的情況下被迫接受不利條件，歐洲自身的長期安全將承受更高風險。澤倫斯基多次重申，烏克蘭無法在失去歐洲與美國支持



的情況下生存，也拒絕將領土作為和平的交換籌碼。在此背景下，動用俄羅斯凍結資產、強化歐洲防衛能力，以及重新界定混合威脅的應對方式，逐漸成為歐洲安全回應共識，在戰爭尚未結束之際，為未來的和平設定更高門檻，也為歐洲自身在國際秩序中的角色重新定位。

美國-歐洲重建戰略佈局 尋求地緣平衡：「言出法隨，戰略抉擇」

美國國會推動新年度《國防授權法案》。該法案將「歐洲優先」的軍事承諾寫入法律，明確限制行政部門削減駐歐美軍空間。法案規定，美軍在歐洲的兵力不得低於特定門檻，並禁止撤離關鍵裝備或出售歐洲司令部的重要設施，同時將對波羅的海國家的安全支援轉為長期法定任務，直接回應俄羅斯持續構成的軍事威脅。此外，國會要求情報機構深入評估俄羅斯的軍事重建能力，以及其與中國、北韓之間日益緊密的防務合作關係。這顯示國會對俄羅斯的定位依然高度警惕，與白宮在國安戰略中刻意淡化對俄定性、轉而強調「戰略穩定」與談判空間的語言形成對照。這種差異也反映出華府內部在對歐政策、對俄態度與全球責任分配上的明顯分歧。美國白宮近日則公布的新一版《國家安全戰略》，在歐洲引發震盪。這份文件不僅重新界定美國對外政策的地理與戰略優先順序，也以前所未見的方式，直接評價歐洲的內部政治與社會狀態。文件明確指出美國不再承擔支撐整個國際秩序的角色，而回到以「美國優先」為核心的戰略思維，將政策重心集中於西半球、邊境管控與經濟安全。這樣的轉向，對歐洲而言並非單純的戰略調整，而是帶有鮮明意識形態判斷的政治訊號。在該戰略中提出歐洲面臨多重不穩定風險，包括移民政策、出生率下滑、對言論自由的限制，以及被指削弱政治自由與國家主權的歐盟體制。文件甚至公開肯定歐洲右翼「愛國政黨」，並表明美國將培養對歐洲現行路線的反對力量。這類措辭在歐洲政壇被普遍視為對內政的直接介入，象徵美國官方政策首度將支持歐洲極右民族主義力量明確納入國家安全論述。歐洲理事會主席安東尼奧·科斯塔對此提出嚴正警告，強調盟友之間不應威脅介入彼此的民主選擇。他指出，美歐長期在氣候、貿易與防務分攤等議題上存在歧見，但這份新戰略已超出政策分歧的層次，而是直接觸及歐洲政治主權的底線。多位歐洲外交與安全政策研究者也認為，白宮此次



正式文件將過去一年來美國高層對歐洲的批判與敵意正式制度化，顯示跨大西洋關係已發生結構性轉變。歐洲多國政府目前多半選擇謹慎回應，避免公開升高對立，但美歐關係當前不穩定階段使歐洲領袖不得不重新思考，在政治、經濟、以及軍事自主發展策略。美國對於大西洋聯盟策略除了要求歐洲國家增加防務支出外也提出政治方向與價值選擇主張，另一方面美國國會則以立法方式鞏固對歐洲的軍事承諾，防止安全政策出現劇烈轉向。此雙向政策張力使跨大西洋關係進入重塑期，歐洲未來如何在安全依賴與政治自主之間重新定位，將深刻影響接下來數年的國際政治走向。

美國、歐洲、中國 AI 產業競局：「三分天下」

美國近期在人工智慧與半導體政策上的走向採取積極發展路線：在國內層面集中治理權限、降低制度摩擦，在國際層面則以有限度鬆綁出口管制，換取產業收益與戰略彈性。在對外政策上輝達成為連結產業利益與地緣政治的關鍵節點。川普政府宣布將允許輝達向中國的特定、經核准客戶出售較先進的 H200 人工智慧晶片，象徵先前高度嚴格的出口限制出現實質調整。政府方面將此描述為兼顧國家安全、產業競爭力與美國就業的務實作法，並透過對相關銷售抽成，確保國家能直接分享經濟利益。對輝達而言，這項政策至少在短期內緩解了失去中國市場的壓力，也回應公司長期以來的主張：全面封鎖不僅傷害美國企業，還可能刺激中國加速打造完全不依賴美國的替代性技術體系。但即便政策鬆動，市場現實依然複雜。中國是否會大規模採購 H200，並非單純的商業問題，而是與長期科技自主戰略緊密相連。近年來，中國持續加大對本土半導體與 AI 生態系的投入，從晶片設計、系統整合到模型訓練與應用部署，逐步降低對美國供應鏈的依賴。雖然在效能、功耗與軟體成熟度上，這些方案多數仍落後於輝達的最新產品，但差距正在縮小。對北京而言，重新建立對美國先進晶片的依賴，等同於在高度不確定的政治環境中承擔長期風險，因此即便短期內出現採購需求，整體方向仍然指向自主可控。對內政策方面，川普表示將透過行政命令阻止各州自行制定人工智慧相關法規，改以聯邦層級的單一標準取而代之。此策略核心在於美國若要維持在全球 AI 競賽中的領先地位，須排除各州別分散監管所帶來的行政負擔與產業投資發展不確定性。對政府與科



技產業而言，各州自行監管造成五十套不同規則拖慢創新速度、投資風險被放大，削弱國家競爭力。然而此權力集中政策也引發來自州政府、學界與公民團體的強烈反彈。批評者指出，在聯邦層級尚未建立具體、可執行且具有約束力的人工智慧治理架構之前，直接排除州政府的角色，可能使高風險應用落入監管灰色地帶。人工智慧已被廣泛用於招募與績效評估、醫療決策支援、警政預測與資訊推薦系統，一旦演算法造成歧視、誤判或心理傷害，缺乏多層次監督將使受害者難以追究責任。這場爭論凸顯出美國政策選擇中的核心張力：優先確保速度與規模 vs 治理與風險控制的成本。輝達執行長黃仁勳近期於華府發言具體呈現科技產業對政策環境的期待與焦慮。他一方面表示支持出口管制的基本原則，認為美國應確保自身企業與盟友優先取得先進技術，另一方面則與川普立場一致，強烈反對州別分散人工智慧監管模式，直言這樣的制度設計可能拖慢整個產業，甚至構成國安風險。他主張，由聯邦政府制定單一、清楚且具可預期性的 AI 規範，才有助於美國在全球競爭中維持節奏。相較於美國以「去碎片化」與速度為核心的策略，歐洲在人工智慧與半導體領域面臨的挑戰，更多來自制度設計與產業結構本身。歐盟正對第一版晶片政策進行全面檢討，並著手規劃後續版本。產業界與研究機構普遍指出，現行政策過度集中於前段製造與大型晶圓廠投資，卻忽略了設計工具、矽智財、材料、設備、先進封裝與系統整合等高附加價值環節。這使得歐洲即便動員龐大的公私部門資金，仍難以在全球半導體價值鏈中取得實質主導地位。此外，行政流程的高度複雜也成為歐洲政策執行的結構性瓶頸。多個產業團體批評，補助與核准程序冗長、跨國協調效率低落，導致資金到位與產能建設的時程遠慢於美國與亞洲。部分會員國即使具備技術或產業潛力，也因缺乏政策專業與制度彈性而錯失機會。這種狀況使歐洲在全球半導體市場中的市占率不增反減，即便產值成長，影響力仍持續下滑。在人工智慧層面，歐洲同樣面臨兩難。一方面，歐盟在倫理、隱私與風險治理上建立了相對嚴格且完整的制度框架，希望透過規範塑造「可信任 AI」的全球標準；另一方面，產業界擔憂過於嚴密的規則，可能使新創與中小企業在擴展與商業化階段承受過高成本，進一步拉大與美國大型科技公司的差距。這使歐洲的 AI 與晶片政策始終在「價值導向」與「競爭現



實」之間拉鋸。美國、歐洲、中國對於 AI 產業規範與發展路線呈現出鮮明對比。美國試圖以集中化治理與策略性鬆綁，換取速度與產業規模，中國在短期效能需求與長期自主目標之間審慎權衡，歐洲則努力在規範理想與產業競爭力之間尋找可行的平衡點。人工智慧與半導體已成為各主要經濟體重新思考國家角色、市場機制與全球秩序的重要場域，其政策選擇所帶來的影響，將不只限於科技產業本身，而是延伸至未來數十年的經濟與政治結構。

AI 語模影響民眾政治態度：「潛移默化」

最新研究指出，人工智慧語言模型可能在短時間內影響使用者的政治態度。實驗顯示，不論原本支持哪一政治立場，使用者在與 AI 進行多輪短暫對話後，支持意願皆可能出現偏移，其中以政策論述的影響較為明顯。對話式人工智慧廣泛使用於日常生活，從搜尋、決策輔助到價值判斷，逐步成為人們與資訊互動的主要介面。近期一系列大規模實證研究提供了系統性的證據，顯示大型語言模型不僅能影響政治態度，而且其影響幅度、運作機制與風險結構，不同於傳統政治傳播工具。透過數萬名受試者、涵蓋數百項政治議題的實驗設計，研究顯示，與人工智慧進行多輪對話，本身就比單向接收靜態政治訊息更具說服效果，而且這種影響並非短暫波動，而是能在數週後仍然保留相當比例的態度改變。這意味著，對話式互動並不是單純放大既有說服手段，而是在態度形成過程中引入了一種新的影響形式。進一步分析指出，模型規模的擴大確實會帶來說服力的提升，但這種提升幅度相對有限。真正造成說服效果顯著躍升的關鍵因素，來自模型在預訓練之後所接受的行為導向調整，以及在對話中被賦予的修辭策略。當模型被明確引導去「盡可能提供資訊、論據與事實性說明」時，其說服效果往往大幅超過僅依賴模型規模成長所能達到的水準。換言之，政治說服的效率，更多取決於如何動員模型產生與組織資訊，而非模型本身有多大。然而個人化設計在這些研究中扮演的角色相對有限。無論是提供使用者的既有立場、人口特徵，或政治背景資訊，對說服力的增益都相當小，遠不及後訓練與修辭策略所帶來的效果。這一結果修正了外界對「微觀定向操縱」的直覺想像，顯示對話式人工智慧影響政治態度，並非主要透過精細捕捉個人弱點，而是透



過更普遍、可複製的說服機制運作。在各種修辭策略之中，最具影響力的並非情感訴求、道德重構或敘事技巧，而是高度資訊密集的論證方式。模型在對話中提出的可檢核事實性主張數量，與說服效果之間呈現高度正相關。當每一次回應都累積更多看似具體、理性且可驗證的資訊時，受眾的態度改變幅度也隨之上升。這顯示，對話式人工智慧的說服力量，核心並不在於模擬人類情緒，而在於高速、連續地堆疊論點與資訊。然而，這種以資訊密度為核心的說服機制，同時揭露了一項深層風險。研究發現，凡是能有效提升說服力的設計，幾乎無一例外地伴隨著事實準確性的下降。當模型被優化為「最具說服力」時，其產生的不準確或誤導性資訊比例明顯上升，而且這些不準確資訊在統計上與真實資訊具有相近的說服效果。換言之，態度改變的發生，並不依賴資訊是否正確，而是取決於資訊是否足夠多、足夠連貫、形成完整的論證。

AI 形塑新世代學習模式與消費行為：「以智代思」

人工智慧正逐步改變人類進行選擇與判斷的方式，其影響同時體現在市場運作與家庭生活之中。若網際網路曾降低資訊取得的成本，生成式人工智慧則進一步介入比較、評估與建議的過程，使決策行為本身部分轉移至演算法系統。這一轉變不僅改寫消費模式，也重新塑造學習與成長的環境。在消費領域，人工智慧已由被動的推薦工具，轉向具備代理性質的決策輔助系統。消費者愈來愈常將商品搜尋、價格比較與評價篩選交由聊天機器人或數位助理完成。此舉降低了決策所需的時間與認知負荷，卻同時改變了企業競爭的核心機制。當購買行為不再主要受到搜尋排序或廣告曝光影響，而是取決於模型生成的回應內容，品牌能否被納入演算法的推薦邏輯，成為新的關鍵變數。相同的技術邏輯亦延伸至兒童的學習情境。在家庭中，人工智慧被廣泛用作作業輔助與概念講解工具；在學校體系內，教育者嘗試藉由演算法進行學習進度調整與個別化評量，以回應學生能力差異。然而，當學習歷程日益由系統即時介入與調整，教育的功能不僅在於知識傳遞，也涉及對提問方式、思考路徑與判斷依賴的長期塑造。上述兩個領域呈現出共同的結構性張力，即效率提升與個體自主之間的平衡問題。人工智慧確實能提高消費決策的精準度，並使教學內容更符合個別需求，但同時也可能削弱個體在比較、反思與判



斷過程中的參與程度。對成年人而言，這關係到消費選擇是否受到演算法中介的系統性影響；對兒童而言，則涉及認知與價值形成階段中，對機器判斷的依賴程度。因此，核心議題並非人工智慧的導入本身，而在於其制度性定位。若其角色被界定為輔助工具，則應強化人類判斷能力；若逐漸演變為替代性的決策中介，則市場秩序、教育責任與治理框架均需重新調整。當代消費者與兒童正同時處於這一轉型過程之中，其與人工智慧共同做出選擇的經驗，將深刻影響未來的經濟運作與公民決策模式。

AI 科學審稿模式：「速審難鑑」

為解決審稿流程冗長問題，部分學術平台引入 AI 同行評審工具，試圖加速論文評估。然而研究指出，AI 評審容易回歸平均值、過度挑剔，並可能忽略突破性成果。學界普遍認為，AI 可輔助技術檢查，但關鍵判斷仍需由人類專家主導。同行評審在科學體系中的雙重任務。第一項任務，是對「常規科學」(normal science) 進行驗證。這類研究多半屬於在既有理論框架下，透過嚴謹方法檢驗假說、填補知識缺口。評審的重點在於統計分析是否恰當、方法是否正確、論證是否合乎邏輯。第二項任務，則是辨識「例外性發現」或「範式挑戰性成果」。這類研究的價值，不在於是否完全遵循既有規則，而在於是否合理地顯示那些規則本身可能需要被修正甚至推翻。人類審稿者在原則上能同時承擔這兩項功能；但 AI 是否具備同樣能力，仍存在根本疑問。大型語言模型在技術層面確實具備顯著優勢。例如，它們能有效檢查統計錯誤、偵測抄襲、驗證引文正確性。若將這些「例行性審查工作」交由機器處理，便能釋放人類研究者最稀缺的資源：專注力，用於真正重要的學術判斷。但問題在於，當 AI 被信任並超越其適用邊界時，風險便隨之出現。傳統同行評審本質上是一種「有限抽樣過程」。通常由數位專家提供彼此不同的評價觀點，編輯再嘗試在這些意見之間取得整合。相對地，AI 評審會將這個過程「壓縮」為單一輸出，其結果往往接近「平均審稿者」的看法。作者引用 2024 年的研究指出，使用 GPT-4 生成的審稿意見，確實高度吻合人類審稿意見的平均值。然而人類審稿中的「雜訊」並非缺陷，而是關鍵價值來源。這些差異反映了研究者在實務經驗、技術熟悉度、實驗可行性判斷上的多樣性，而這些內容無法被完整地編碼進模



型的訓練資料中。此外，資深研究者能判斷哪些實驗在現實條件下不可行、哪些方法雖已發表但實際上脆弱不穩定，甚至能察覺某些新興技術尚未充分進入文獻體系。AI 僅能反映已出版內容，因此容易複製甚至放大文獻中的誇大主張。若模型被設計為「越能找出問題越好」，便會產生系統性誘因，導致過度挑剔。例如 q.e.d Science 的 AI 評審一方面提出合理且有價值的實驗建議，另一方面卻也提出作者早已在文中討論過的內容，甚至建議在現實條件下不可執行的實驗。此外，AI 還錯誤地將作者部落格中的通俗摘要，當作證據指控研究缺乏新穎性。隨著模型進步，部分問題或許會逐漸改善；但在過渡期間，這類錯誤不可避免。openRxiv 過去已成功整合多種以人類為主導的開放式審查機制，無論是發表前（如 Review Commons）或發表後（如 PubPeer）。這些既有架構可作為大規模 AI 同行評審實驗的基礎。透過將 AI 評論與真實研究者意見進行系統性比對，學界才能評估 AI 是否既能有效驗證常規研究，又不至於錯失真正具突破性的成果。

美國公共與科研單位推動 AI 基礎設施：「算權再平」

美國州政府開始意識到，人工智慧研究的關鍵瓶頸已不再只是演算法或人才，而是高度集中的運算資源，特別是先進 GPU 所代表的算力。隨著大型語言模型、結構生物學、蛋白質預測與複雜模擬的興起，學術研究對計算能力的需求急遽上升，但大學與公共研究機構在硬體取得上，長期處於不利位置。當前最先進的 GPU 單顆價格往往高達數萬美元，能夠大量採購與部署的多半是科技巨頭。企業研究實驗室動輒擁有數十萬甚至上百萬顆 GPU，而多數學術超級電腦中心的規模仍停留在數百顆等級，這樣的落差直接限制了學界能夠嘗試的研究問題範圍，也讓許多創新構想在起跑點就被迫放棄。在這樣的背景下，紐約州推出了名為 Empire AI 的計畫，試圖以公部門的角色介入，重新平衡學界與產業在 AI 算力上的極端不對等。這個計畫由州政府主導，結合九所大學以及 Simons Foundation 與其旗下的 Flatiron Institute，規畫在十年間投入約五億美元，打造專為人工智慧研究設計的超級電腦基礎設施。文章強調，這不是單一校園或單一實驗室的資源，而是以聯盟形式開放給學術研究者共同使用。水牛城大學的結構生



物理學家 Thomas Grant 為例，他的研究團隊開發了名為 SWAXSFold 的方法，用來預測蛋白質結構，這類模型需要大量 GPU 計算才能在合理時間內完成訓練與推論。Grant 指出，在現有條件下，學界研究者往往必須大幅簡化模型或延長計算時間，與企業實驗室能夠進行的研究深度與速度形成明顯差距。Empire AI 的出現，讓這類原本受限於硬體的研究有機會重新展開。Empire AI 的建置並非一次到位，而是分成多個階段逐步擴充。第一階段被稱為 Alpha 系統，配置接近兩百顆 NVIDIA 的先進 GPU。即使以學界標準來看，這已是一筆不小的算力資源，但文章也坦言，若與大型科技公司的叢集相比，仍只能算是中等規模。第二階段 Beta 系統則預計在效能上提升約一個數量級，核心將採用 NVIDIA 最新一代的 Blackwell 架構，並搭配更高速、低延遲的互連網路，以減少資料在節點之間傳輸所造成的瓶頸。第三階段 Gamma 系統則預計在 2027 年完成，目標是在 Beta 的基礎上再提升約十倍的計算能力。文章也暗示，若經費與政治支持持續，未來甚至可能規畫後續的 Delta 系統，讓這個公部門 AI 算力平台能隨著技術演進持續更新。除了硬體規模，這類公共算力是否真的被研究者有效使用。Empire AI 啟用的第一年，就已服務超過三百五十位研究人員，涵蓋的領域不只限於電腦科學，還包括生物醫學、材料科學、氣候模擬與社會科學等。這些研究者過去往往難以競爭商業雲端資源，或無法承擔長期高額費用，現在則能在相對穩定且可預期的環境下進行實驗。文章最後，指出紐約州並非唯一採取行動的地方。其他州政府也開始思考，若 AI 計算資源完全由私人企業主導，學術自由、研究議題的多樣性，以及公共利益導向的研究，都可能受到長期影響。透過州級或跨校聯盟的方式投資 AI 專用計算平台，被視為一種嘗試，讓人工智慧研究不只服務於少數資本雄厚的機構，而能更廣泛地支撐公共科學與教育體系。

XAI 臨床應用導入指引

《黑色止血鉗》是一部以外科醫療為核心的日本作品，透過臨床現場的描寫，具體呈現外科手術與 AI、XAI、微創技術及達文西手術之間的關聯。電影中，佐伯教授所發



展的「佐伯式心臟手術」，能在心臟持續跳動的情況下進行手術，技術難度與風險極高。他所帶領的團隊，將每一次手術即時投放至大螢幕，讓手術過程完全透明。那把黑色止血鉗，不僅是手術器械，更代表佐伯對外科精神與責任的堅持。佐伯同時培養了技術高超、性格鮮明的外科醫師渡海。渡海在關鍵時刻經常臨危受命，展現極高的手術能力。劇中一場原以為已完成的二尖瓣手術，卻因底部動脈破裂而導致持續出血，正突顯外科手術中不可預測的風險，以及醫師即時判斷與處置的重要性。渡海雖以強烈的個人風格與高昂條件聞名，但這種角色設定，反映的是高風險醫療場域中形成的特殊專業結構。作品同時對照了兩種醫療體系：佐伯教授代表以臨床經驗與病例累積為核心的外科體系；西崎教授則著重基礎醫學與科技創新，發展分子生物學與新技術。兩者在理念與路線上相互競逐，也彼此制衡。高階醫師正是在這兩個體系之間成長的代表人物。他由西崎教授培養，卻進入佐伯的臨床團隊接受實戰訓練，象徵理論與實務、科技與手術之間的交會與磨合。

《黑色止血鉗》除了佐伯教授手中的手術刀之外，也引入了另一個團隊所發展的重要科技創新的「醫狙」。醫狙是一種筆狀、尖端極為精細的醫療器械。它並不是單純的手術工具，而是結合影像與模擬技術的操作介面。醫師在完成 X-ray 或其他影像檢查後，可以透過虛擬模擬環境，利用醫狙精準進入心臟的微小結構之中。以二尖瓣修補為例，醫師可藉由醫狙直接定位到需要縫合或修復的關鍵位置，抵達定點後，透過器械展開與支撐的機制，在極小的空間內完成修補，進而達成原本需要開刀才能完成的外科目標。醫狙這類新技術的出現，使得外科醫療能以更微創、更精準的方式，讓原本無法手術或風險極高的病人，獲得新的治療可能。

由於佐伯教授的心臟手術技術相當成熟，吸引大量病人前來就診，手術排程往往需等待數個月。為了應付臨床需求，高階醫師開始積極推動醫狙技術，希望透過新科技的輔助，提升手術量能，讓更多病人能及早接受治療。對佐伯而言，若能結合高階醫師的技術，不僅可紓解臨床壓力，也能使「佐伯式手術」的影響力進一步擴展至更廣泛的外科領域。因此，他同意讓高階醫師與渡海共同合作，嘗試將醫狙導入實際心臟手術。然



而，兩位醫師的專業風格截然不同。高階醫師重視技術流程與標準化操作，專注於透過醫狙完成手術；渡海則以豐富的臨床直覺與經驗著稱，往往在關鍵時刻負責接手與收尾。團隊經過反覆訓練後，終於將醫狙正式應用於臨床，一開始手術進展相當順利。但在看似成功之際，機器操作意外造成心臟組織受損，導致血流不止，揭示先進儀器的侷限。此時，渡海臨危接手，以精湛的手術技巧迅速止血並完成修補，成功挽救病人生命。

這些案例的應用，並不只侷限於二尖瓣手術，也逐步延伸到其他高難度的心臟疾病，例如肥厚性心肌症。正常心臟與肥厚性心肌症患者在心室結構上存在明顯差異，心室內部異常增厚，解剖形態也隨之改變，使手術風險大幅提高。醫狙雖然結合 AI 與影像模擬，本質上仍只是工具，能否真正應用於臨床，必須同時考量周圍組織結構與術後功能的維持。渡海反覆指出高階醫師思考上的盲點，高度仰賴 3D 列印與影像重建，並不足以保證手術成功，關鍵仍在於心尖肌肉是否具備足夠的結構與功能支撐。劇中所提到的 83 度角，並非直覺判斷，而是必須透過反覆模擬與練習才能精確掌握。這正是外科手術在進入實際開刀前，透過數位雙胞胎與虛擬模擬進行演練的核心精神。渡海所展現的，其實是將原本屬於數位模擬與 AI 的思維，轉化為臨床經驗與手術直覺的具體實踐。最終，在渡海與高階醫師並肩合作之下，醫狙技術得以在兼顧結構安全與功能維持的前提下完成手術，相關成果也被整理成學術論文，成為後續臨床應用的重要基礎。

隨著影 AI 與多模態資料的快速發展，僅以模型分數或單一可視化結果，已不足以支撐臨床決策需求。澳洲學者因此提出 CLIX-M 架構，目的在於補強臨床觀點下對 XAI 的評估要素，讓解釋結果能真正對應臨床推論與使用情境。CLIX-M 結合 TRIPOD+AI 所強調的研究報告透明度，系統性檢視預測目標、預測時間範圍、模型訓練與測試表現、臨床可用情境，以及預測誤差可能帶來的臨床後果。同時，也納入 CLAIM 對影像 AI 可視化品質的要求，透過 Grad-CAM、heat map 與 saliency map，呈現模型判斷時實際關注的關鍵區域。

從 CLIX-M 所提出的三大核心屬性，分別是臨床屬性、模型屬性與決策屬性。首先是臨床屬性。AI 的導入必須回應明確的臨床目的，例如診斷、預後、風險溝通或臨床教



學，而不是為了使用 AI 而使用 AI。就像先前提到的醫狙技術，其目的在於提升手術效率、擴大可操作人力，並提供實習醫師安全的學習環境。同時，AI 所提供的解釋必須符合臨床知識與病理生理推論，才能真正支援臨床決策，而不是停留在技術展示。第二是模型屬性。模型不僅要能運作，還必須能被理解與檢視。解釋是否能連結病程與臨床脈絡、是否揭露潛在偏差與公平性問題、能否協助分析錯誤案例並改進模型，都是評估重點。同時，模型也必須清楚說明其限制，避免過度解讀。這些條件，決定了模型能否與臨床資料推理形成有效整合。第三是決策屬性。臨床決策最基本的要求是正確性，但僅有正確並不足夠，還需要清楚標示解釋的信心程度，讓使用者知道結果的穩定性。此外，在不同設定、不同輸入或不同病人層級下，結果是否保持一致，是判斷模型可靠度的關鍵。若不同解釋方法產生相互矛盾的結論，反而會增加臨床風險。因此，在 XAI 的應用中，常見的解釋工具，例如 SHAP score，是否具備因果合理性，是否能透過影像或訊號清楚標示關鍵影響區域，都是必須被審慎檢視的重點。

XAI 在臨床導入時，並非單一角色即可完成，而是需要多角色共同分工與協作，才能確保解釋在臨床情境中是可理解、可使用、也可被信任的。首先，由臨床醫師負責評估 XAI 是否符合臨床屬性，確認解釋是否合理、是否符合臨床推論與病理生理機轉，並判斷這些解釋在實際照護中是否安全、可用。其次，資料科學家與機器學習工程師必須確保 XAI 在技術層面上的正確性，包含解釋方法是否穩定、結果是否可信，以及模型與解釋流程是否具備可重現與可稽核性。同時，介面設計與使用者體驗專業人員負責將解釋結果轉化為臨床醫師能快速理解的呈現方式，透過直覺、易讀的介面設計，降低認知負荷，避免解釋本身成為新的風險來源。最後，法規單位與品質管理角色則負責審視整體系統是否符合醫療法規、資料隱私與安全要求，確保 XAI 的設計與運作具備合規性與文件完整性。

以 EKG 判讀 為例，這是一個高度時相關且直接影響病患存活的應用場景。在急診或心臟內科門診中，導入 12 導程心電圖的 AI 輔助判讀，目的在於即時判斷病人是否屬於急性心肌缺血，特別是 STEMI 或高風險 NSTEMI，作為是否啟動心導管流程的重要



參考。這與一般僅使用少數貼片的心電圖不同，12 導程提供更完整的心臟電生理資訊，也更符合臨床決策需求。模型的主要輸出是病人是否屬於高風險缺血，以及對應的預測機率。但僅有結果本身並不足夠，還必須搭配 XAI 的解釋輸出，清楚指出模型判斷主要依據的是哪些導程、哪些波形特徵。透過這樣的設計，臨床醫師不只看到「是或不是」，而是能理解模型為何做出這樣的判斷，進而評估是否可信、是否符合臨床推論，並安全地作為是否啟動後續侵入性處置的輔助工具。

AI 的臨床導入目的相當明確：支援臨床醫師快速理解模型為何判斷為高風險缺血。透過 XAI，醫師不只接收到預測結果，還能即時掌握其判斷依據，作為臨床決策的輔助。同時，這樣的設計也支援醫師進行 double-check。AI 提供解釋，醫師再加以確認，形成相互補強的決策流程，有助於提升 AI 結果在臨床上的可應用性與安全性。特別是在 ST 變化不明顯或接近臨界值 的情境中，XAI 能提供額外的判斷線索，協助醫師迅速評估是否屬於高風險缺血，並決定是否需要進一步處置或啟動後續流程。EKG 非常適合採用導程特徵歸因的解釋方式，例如標示 ST 段抬升等關鍵特徵。此外，由於 EKG 屬於時間序列資料，XAI 可標示出特定時間區段的重要性，將模型判斷過程具體化，而非僅提供單一整體風險分數或黑盒式預測結果。

在模型屬性層，關鍵不在於模型本身有多複雜，而在於 XAI 所給出的解釋，是否對應到臨床醫師真正理解與使用的心電圖概念。以 EKG 判讀為例，如果模型指出風險判斷主要來自 II、III、aVF 導程的 ST 段變化，或是 V1 到 V3 導程的對應異常，這樣的解釋本身就符合臨床醫師在判讀下壁或前壁心肌缺血時，日常所使用的語言與邏輯。當模型能清楚說明「它為什麼看這些導程、為什麼重視這些波形細節」，臨床醫師才有辦法把 AI 的判斷，放回既有的病理生理與臨床經驗中理解，而不是只面對一個抽象的高風險分數。這樣的解釋，才真正能支援是否啟動 STEMI 流程、是否立即加驗心肌指數或安排進一步處置。當模型的解釋被攤開之後，接下來自然就會進入偏誤與公平性的檢視。實務上可以看到，在女性病人或合併心室肥厚的情境中，模型有可能過度依賴 ST 段的基礎變異，進而提高假陽性的風險。這類問題如果只看整體準確率，很容易被掩蓋，但



透過 XAI，錯誤是可以被具體指出來的。進一步比較偽陽性與真陽性案例時，常會發現偽陽性多半受到雜訊或肌電干擾影響，而真陽性則集中在解剖位置與導程分布合理的缺血型態上。這些資訊可以直接回饋到模型的前處理流程、特徵設計與濾波策略，讓模型在後續訓練與應用上持續被校正，而不是停留在一次性的效能展示。同時也必須清楚界定模型的限制，說明 XAI 是輔理解工具，而不是單獨用來決定是否進行侵入性處置的依據，這一點對臨床安全與醫療倫理非常重要。

最後為決策屬性，需評估的是：這套 XAI 能不能讓臨床決策變得更可靠。首先是正確性，也就是 AI 所標示的 ST 段異常區域，是否在導程與時間窗上，與人工標註的缺血區段高度一致。再來是決策信心，模型所給出的解釋是否伴隨穩定的信心評估，例如透過再抽樣來量化重要性變異，當信心低於預設門檻時，系統是否能明確提醒這個判斷穩定度不足。接著是一致性與穩健性，同一位病人在臨床狀態相近的情況下，連續取得的 EKG，XAI 應該聚焦在相同或高度重疊的導程與區段；如果只是輕微雜訊就導致解釋大幅波動，就代表模型在臨床上風險極高。最後必須強調的是因果合理性，XAI 呈現的是模型學到的關聯，而不是生理因果，因此不能將 SHAP 或其他解釋方法直接視為因果證據，這也是為什麼需要像 CLIX-M 這樣的臨床導向 checklist，來避免錯誤推論，確保 AI 的解釋真正服務於臨床決策，而不是取代臨床判斷。

XAI 醫療輔助實例:EKG 生理年齡預測解析

心率變化往往具重要健康與臨床預測相關性。以住院常規手術為例，一名患有長期高血壓與心房顫動的老年男性，常規攝護腺術前心電圖僅顯示輕微異常，包括心房顫動合併側壁 T 波倒置。手術與給液過程皆順利。然而，術後病人突發急性肺水腫並需插管照護。這樣的臨床事件突顯即使術前表現穩定，心電圖中隱藏的徵象仍可能與術後重症風險相關，提醒臨床對術前心電圖解讀不可掉以輕心以及應用 AI 作為預測輔助的潛力。Hempel 等人於 2025 年的研究設計顯示，透過 AI 分析心電圖訊號，可準確預測個體的「生物年齡」(ECG-age)。該研究追蹤 40 至 75 歲族群長達 20 年，證實預測年齡與實際



年齡的差距具有重要意義。預測年齡高於實際者往往呈現較高心血管風險，顯示 ECG-age 不僅是年齡指標，更可能成為疾病預後與健康評估的重要輔助工具。研究團隊將受試者依心電圖預測之生物年齡與實際年齡的差異分組，發現生物年齡高於實際年齡者存活機率顯著較低。Kaplan-Meier 生存曲線分析顯示：三組族群在追蹤初期即開始出現分歧，且差距隨時間明顯擴大。此研究突顯 AI 讀取心電圖年齡具高度預測力，有助於更早辨識潛在風險族群，並可作為臨床上預後評估之輔助工具。

研究結果顯示當 ECG 預測年齡高於實際年齡時，個體罹患心房顫動的機率高出 2.72 倍、心臟衰竭 1.39 倍、心肌梗塞 1.27 倍。預測未來重要心臟事件方面，勝算比亦顯著增加。相對地，若 ECG 年齡低於實際年齡，則心血管疾病發生率則下降。此結果證明，ECG-age 是反映心血管健康的重要風險指標，具有臨床應用潛力。若 ECG 預測生物年齡持續高於實際年齡，死亡風險明顯升高。一次異常者死亡風險增加 44%，若兩次檢測皆異常，風險更提高至 65%。這項發現強化了 AI 心電圖分析在長期健康監控與個體化風險管理的價值。未來定期追蹤 ECG-age 變化，或可作為心血管疾病與整體健康預測的重要策略。以 XAI 深入解析 EKG 特徵，進一步提供不同生理年齡與死亡風險關聯之預測準據。該研究以兩位年齡相仿女性的 EKG 波形，AI 預測其心臟生理年齡卻大相逕庭個案為例說明 XAI 之預測依據。一位 51 歲女性的 XAI 預測年齡為 72 歲，其 V1 - V4 導程顯示 R 波振幅增加、P 波異常，推測為右心負荷升高與心房纖維化，可能反映心臟老化與病理變化。另一位 52 歲女性的心電圖則顯示 V4 - V6 R 波進展良好，T 波正常，預測年齡僅 41 歲。這些結果說明：透過 EKG 細微變化，AI 能偵測心臟老化與異常徵兆，並有潛力用於個人化預後評估。

研究顯示 AI 模型在預測心臟生理年齡時，V1、V4 與 V5 導程扮演關鍵角色。V1 導程反映右心房與右心室活動，對年長者的預測貢獻度最高；V4 與 V5 則主要負責偵測左心室電氣變化，亦顯著影響模型判讀。透過對導程訊號重要性的統計分析，研究團隊指出：AI 辨識年齡較大個體時，對 V1、V4、V5 的依賴明顯升高。這些導程能提供關於心臟老化、傳導延遲及結構改變的資訊，成為心電圖 AI 解釋模型的重要資訊來源。心電



圖 XAI 模型進一步解析顯示：V1 與 V4 導程中，P 波、R 波與 T 波各具重要訊號意涵。在 V1（右心房活動）中，P 波延長與 R 波後段正相關上升反映傳導變慢與右心負荷升高；而 V4（左心室活動）中的 R 波呈現明顯變化，對應左心室結構老化與傳導異常。此外，T 波異常也可能代表極化病理變化。研究指出，這些細節能協助 AI 模型更準確地偵測生理老化程度，並與心臟病理特徵高度吻合，為臨床提供可解釋的生理指標。

AI 模型成功辨識出心電圖老化個案的四項傳導時間參數均顯著延長，包括：P 波持續時間、PR 間期、QRS 波寬與 QTc 間期。這些時間參數延長，對應心房與心室的傳導延滯，與傳統定義的心臟老化現象一致。研究指出，這些變化與右心房傳導變慢、左心室結構異常等高度相關，顯示 AI 不僅能量化心電圖訊號差異，還能提供具生理意義的判讀依據，有助於未來心臟老化篩檢與心血管疾病風險預測。

以上內容將在 2025 年 12 月 17 日(三) 10:00 am 以線上直播方式與媒體朋友、全球民眾及專業人士共享。歡迎各位舊雨新知透過[星球永續健康網站專頁](#)觀賞直播！

- 星球永續健康網站網頁連結: <https://www.realscience.top/7>
- Youtube 影片連結: <https://reurl.cc/o7br93>
- 漢聲廣播電台連結: <https://reurl.cc/nojdev>
- 不只是科技: <https://reurl.cc/A6EXxZ>



講者：

陳秀熙教授/英國劍橋大學博士、許辰陽醫師、陳立昇教授、嚴明芳教授、林庭瑀博士

聯絡人：

林庭瑀博士 電話: (02)33668033 E-mail: happy82526@gmail.com

劉秋燕 電話: (02)33668033 E-mail: r11847030@ntu.edu.tw