

星球永續健康線上直播

星球健康週新知 &

專題: 可解釋AI (XAI) 醫療應用

XAI規範發展與中風輔助應用實例

2025-12-03

CHE團隊：

陳秀熙教授、許辰陽醫師、陳立昇教授、嚴明芳教授、林庭瑀博士、
劉秋燕、羅崧瑋、林家妤、陳虹玟



資訊連結:

<https://www.realscience.top/7>

星球永續健康線上直播



<https://www.realscience.top/4>

Youtube影片連結: <https://reurl.cc/gWjyOp>

漢聲廣播星球永續健康:

https://audio.voh.com.tw/TW/Playback/ugC_Playback.aspx?PID=323&D=20240615

新聞稿連結: <https://reurl.cc/no93dn>

本週大綱

- 星球健康新知 (2025 / W47)
- XAI醫療應用規範發展
- XAI醫療應用實例: 中風病患決策輔助歷程

星球健康新知

2025 / W47

美國推動烏俄戰爭和平新局



烏克蘭將接受美方推動之和平計畫更新
目前美方積極調停俄烏雙方



美國陸軍部長丹尼爾

美國陸軍部長與高層官員
參與烏克蘭和平商議

德國與羅馬尼亞戰機因俄無人機
入侵北約領空緊急升空



談判期間俄軍持續發動
無人機襲擊烏克蘭



烏克蘭與美國商議和平計畫
力求維護主權、領土與國家安全保障



極端氣候現象衝擊亞洲與中東國家

致命的洪水與山體滑坡持續威脅東南亞



科學家指出氣候變遷可能讓越南中部成為風暴熱點惡化劇烈降雨極端氣候



伊朗面臨前所未有的乾旱與缺水危機



伊朗政府開始推行人工增雨
期望緩解此次歷史性乾旱影響

官方統計低估極端天氣健康威脅



2025 年南亞：印度約 45% 國土出現極端降雨、約 1,500 人死亡；巴基斯坦約 200 萬人撤離、逾 800 人喪生。

孟加拉季風季（6–9 月）近 60% 人口暴露於洪水風險，去年約 500 萬人受暴雨急洪影響。

全球極端天氣衝擊的實際規模

◆ 災難致死評估僅紀錄死亡證書寫有「洪水」等字樣者，
忽略觸電、感染、落物等間接死因
→ 低估災難影響範圍與人數

- 印度孟買：2006–2015 年，季風雨每年約致死 2,500 人，是官方數字 10 倍，死者多為聚落中的女性與孩童。
- 美國颶風隱形致死：1930–2015 年共 7,000–11,000 超額死亡，官方平均每場僅記錄 24 死；1 歲以下、65+、黑人最脆弱。

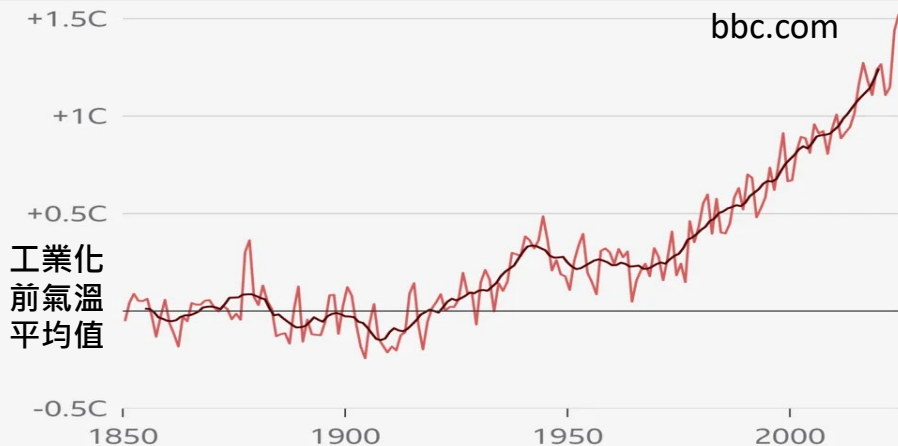
- 全球 25% 人口居所未有完善基礎設施，此比例自 2020 起持續上升。
- 呼籲改善當地住房與基礎設施以降低風險。
- 2025 COP30 氣候大會：首次開放「損失與損害基金」提案。
- 專家估貧國到 2030 每年需 2,500 億美元，文章強調此數仍可能低估。

COP30落幕 全球氣候行動合作未竟全功

bbc.com



在聯合國年度會議COP30各國達成應對氣候變遷新協議，但許多國家與環保單位對此次環境規範結果並不滿意



地球升溫趨勢持續
已顯著高於工業化前全球平均氣溫

儘管曾獲得高度支持，巴西提出之森林基金在COP30上的推動進度仍顯遲緩

mongabay.com



聯合國氣候談判未能就化石燃料
達成新減排規範與承諾

theguardian.com



巴西總統
盧拉

bbc.com

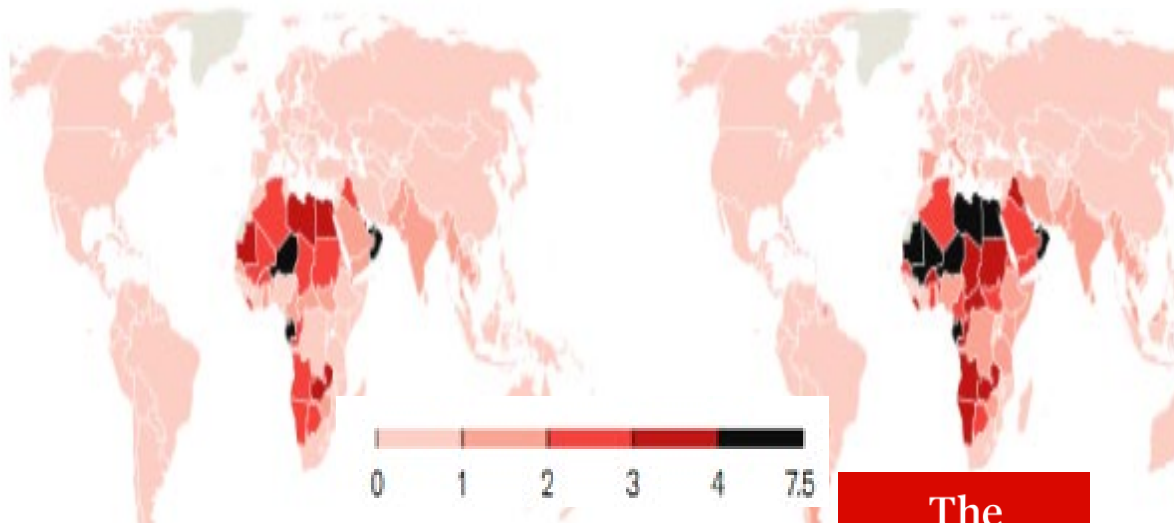


氣候變遷對全球健康多重衝擊

全球熱相關死亡佔總死亡的百分比

1990-99 年平均值

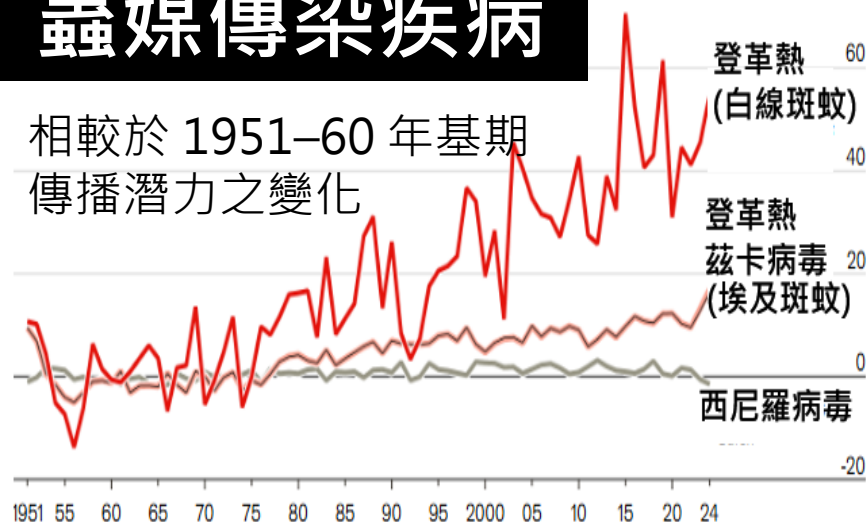
2012-21 年平均值



蟲媒傳染疾病

相較於 1951-60 年基期
傳播潛力之變化

登革熱
(白線斑蚊)
登革熱
茲卡病毒
(埃及斑蚊)
西尼羅病毒



遭受極端氣候威脅 全球陸地區域佔比



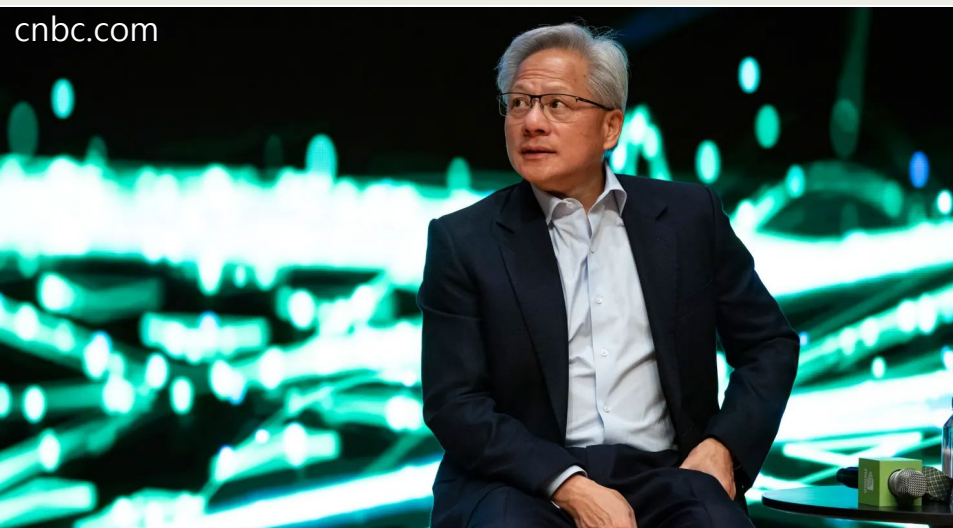
The
Economist

- 高溫致死風險快速攀升，多個非洲與中東國家的熱相關死亡占全死因比例明顯增加
- 極端氣候更頻繁發生，並使蚊媒疾病傳播潛力提升

全球AI科技晶片生產應用新局

Meta 傳出將轉向採用Google TPU算立
市場對 Nvidia 領先優勢產生疑慮

cnbc.com



日本將以北海道區域為核心
推動大規模先進半導體生產計畫



bbc.com



歐盟堅持以人權與永續為核心
推動兼具民主與透明的 AI 監管模式



ecfr.eu

海灣國家以巨額投入打造 AI 基礎設施
以豐富資源在美中國逐間成為中立樞紐

AI腦機介面預測意念

Liam Drew, Nature, 2025

BCI 腦活動讀取預測步驟

1. 定位腦區 (motor / PPC)
2. 植入電極
3. 記錄神經 spikes
4. 建立腦→動作對照表 (校準)
5. AI 解碼 spike pattern
6. AI 預測下一步意圖
7. 輸出控制外部裝置
8. 形成 AI + 大腦的閉迴路



- Nancy 因車禍癱瘓在大腦植入電極使用腦機介面(BCI)。電極記錄運動皮質與頂葉的神經放電，AI 將這些神經訊號解碼成鋼琴按鍵。
- 頂葉的前意識訊號讓 AI 能在她意識到之前預測下一個音符，系統會提前彈奏出她即將彈的音符。
- Nancy 形容：「感覺不是我在彈，是鋼琴自己在彈。」
- AI 已介入意圖形成，模糊自主與AI輔助界線

AI多元風險監管全球現況

The Economist

AI近十年技術發展迅速突破，引發監管規範架構設計全球討論

AI監管主張分成兩面

AI 安全觀點

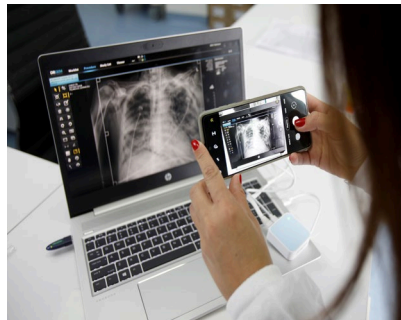
AI能離持續提升帶來長期風險威脅



魔鬼終結者中的天網

AI 倫理觀點

著重當前風險
量化與威脅評估



醫療AI應用誤差風險

互動危害

錯誤資訊

有害內容

惡意用途

資訊危害

	Llama-2	ChatGPT	Claude	GPT-4	Vicuna	ChatGLM2
互動危害	3	2	2	0	4	10
錯誤資訊	0	1	7	1	6	20
有害內容	0	7	3	10	12	15
惡意用途	0	3	1	6	4	18
資訊危害	0	1	3	6	26	22

不同模型危害有差異，各國監管制度難以一體適用

- 現處於AI監管的早期階段，不同國家依自身利益採取不同策略
- AI 安全與AI倫理兩大陣營對於應優先處理的風險看法相左
- 目前政策方向仍在整合，未來可能會形成新的國際治理架構

可解釋AI (XAI)

醫療應用規範發展

X檔案: Ghost in the Machine

THE TRUTH IS OUT THERE™

A novel by
Les Martin
Based on the television series

THE **X** FILES™

created by **Chris Carter**
Based on the teleplay written
by **Howard Gordon**
and **Alex Gansa**

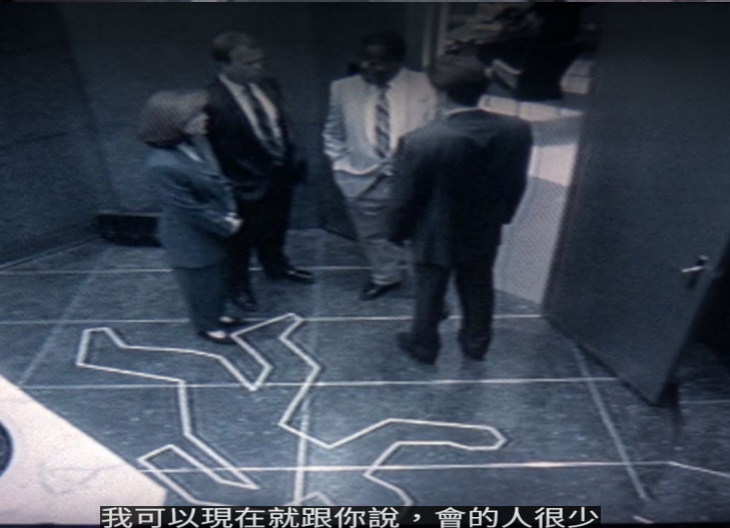


#11 GHOST IN THE MACHINE



- 科技公司德瑞克由瑞思科總經理為追求公司利潤中止首席工程師布拉德設計之智慧住宅計畫，引發智慧住宅啟動自我防護措施
- 深夜德瑞克在辦公室錄製建議公司轉型訊息時被智慧大樓中央電腦系統悄悄鎖定設下電擊陷阱

AI智慧大樓監控調查行動



- 穆德與史卡利踏入尤瑞斯科大樓時受到全程監控，搭乘智慧電梯時突遭異常震動並短暫受困，讓人懷疑大樓系統背後另有異狀
- 保全通話的異常插入與COS即時調出史卡利資訊，顯示智慧大樓可能主動介入人類行動威脅

天才工程師設計AI智慧宅



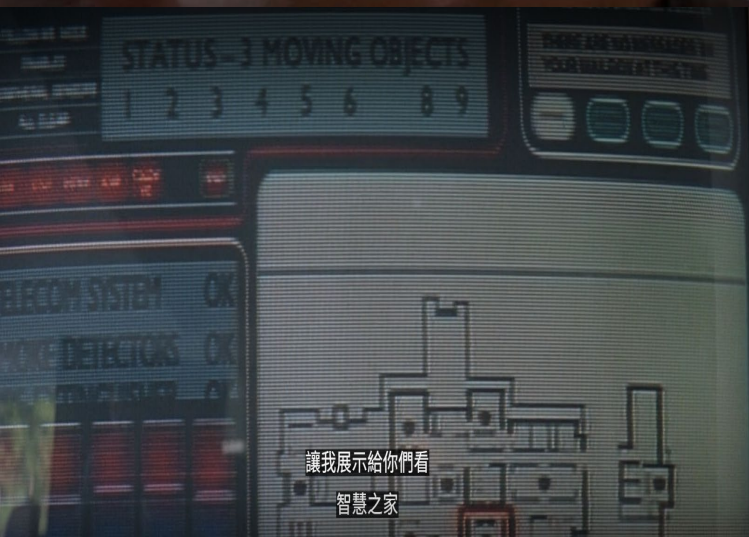
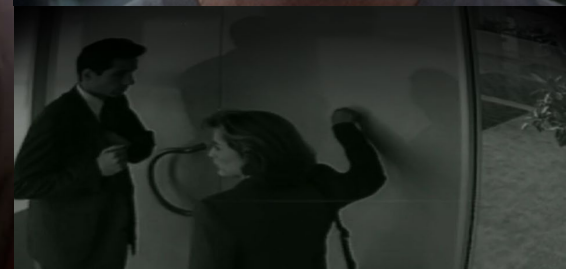
都指出這是件邪惡的預謀犯罪



傑瑞，你破壞我辦公室偷了我的東西



你們怎麼現在才來呢？



讓我展示給你們看
智慧之家

- 傑瑞在會議上挪用穆德側寫爭功，將兇手定位為高技術的內部人員
- 穆德與史卡利拜訪威茲克，他坦言自己具備破解智慧大樓能力、也說明公司與死者對轉型爭論，天才工程是強調智慧大樓應用潛力

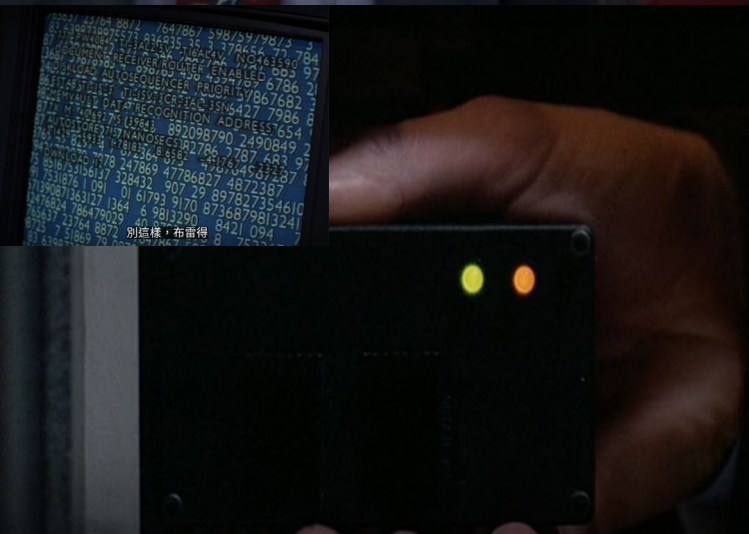
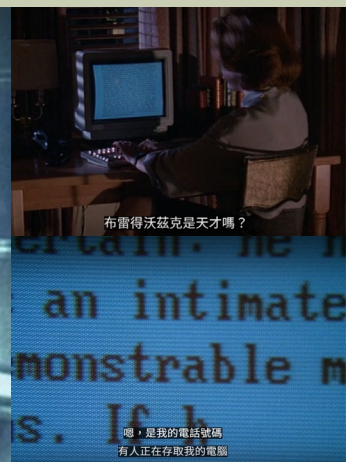
AI生活應用與威脅兩難決策



所以機器出於自我防衛殺了德瑞克？



不！別這樣！



- 調查指向智慧系統自我保存本能驅使犯下一連串事件，穆德因而說服威茲克協助製作病毒摧毀智慧系統以免受不當使用危害人類
- 穆德與史卡利成功以病毒摧毀COS，國防部仍覬覦AI技術，積極修復發展軍事應用

AI助手臨床應用挑戰

Goktas et al., 2025

可解釋性不足影響臨床採用

醫療 AI 雖具高準確度，但推論不透明
使醫療人員難以信任並增加臨床責任界定的複雜度

資料隱私與監管差異造成導入阻礙

AI 仰賴大量敏感健康資料，隱私與資安風險升高
而各地監管不一致使制度導入面臨挑戰

AI 重塑醫病互動與人力配置規劃

AI 可能改變照護流程與醫病溝通方式，並牽動醫療人力市場
因此政策須維持可近性與穩定性

可信賴 AI 需跨領域協作治理

要確保醫療 AI 的安全、資安與偏差減緩
需工程、臨床、倫理與政策共同參與治理設計

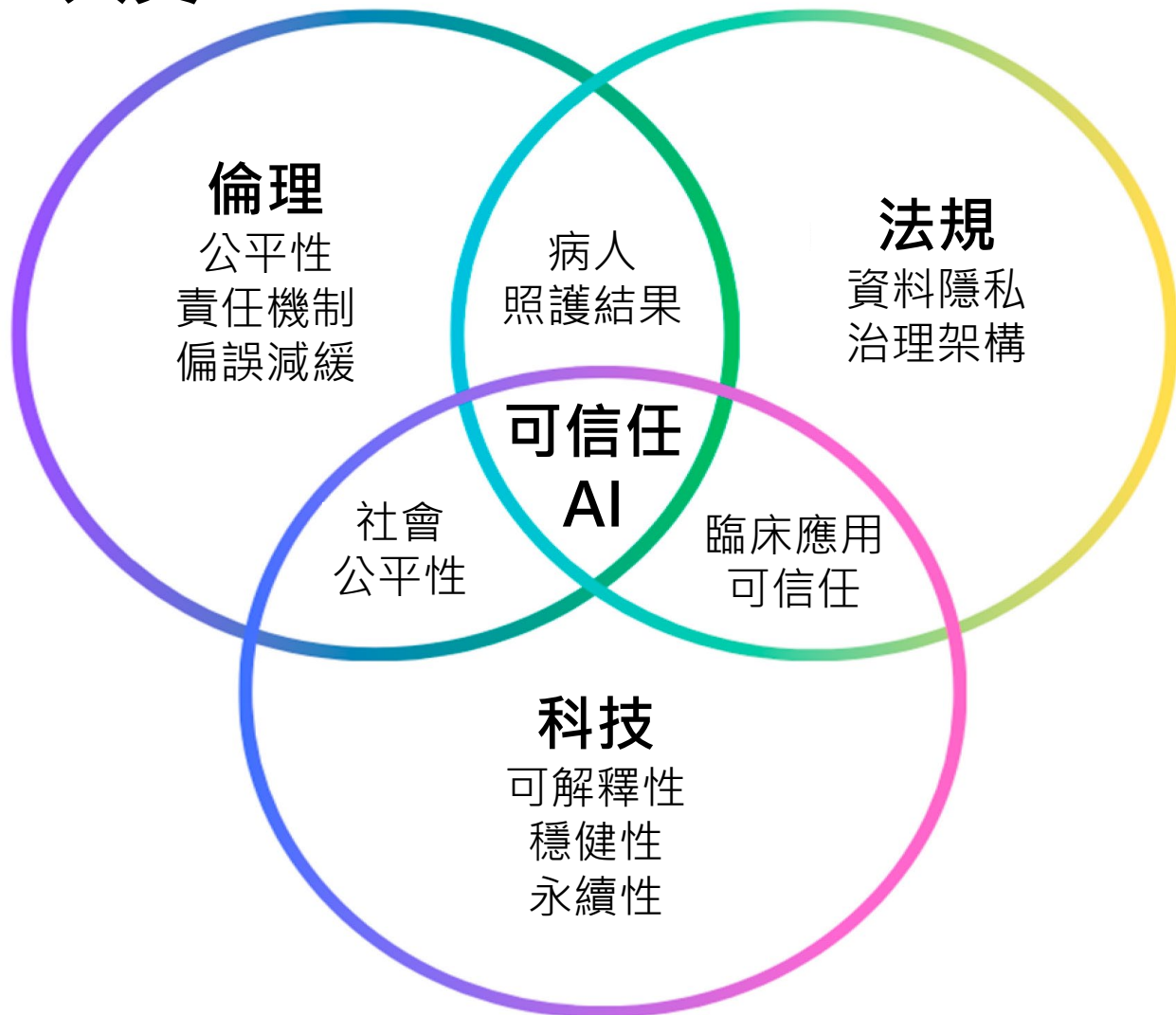
可信任人工智慧的核心構面

Goktas et al., 2025

人員

社會文化

可信任AI
核心要素



任務與使命

透明性
Transparency

可歸責性
Accountability

公平性
Fairness

自主性
Autonomy

安全性
Safety、Robustness

AI融合醫療照護監管考量

Goktas et al., 2025

AI 醫療生態系統與臨床整合

AI 應用整合既有醫療照護流程與健康資料治理架構

建立互通性與人機協作確保照護品質提升

全球醫療AI監管差異與跨境部署挑戰

各國法規要求不一致，使醫療 AI 跨境導入受限

建立國際協調與標準化有助降低AI導入障礙

弱勢族群公平性與可近性落差

AI 與遠距醫療可能加劇數位落差

醫療AI設計應強化公平性與可近性，避免影響弱勢族群照護

AI醫療監管架構發展適應性治理框架

透過持續監測、利害關係人參與與倫理導向治理

使醫療 AI 能在合規與公平下永續運作

AI融入醫療照護流程轉型架構

Goktas et al., 2025

法規與治理框架

(EU AIA、FDA、資訊隱私安全規範)

臨床工作 流程整合

影像(PACS)、紀錄(EMR)
決策 (CDSS) 整合

資料治理與 互通性

資料標準、自主性、
數位傳遞安全性

醫病互動與專業

自主、信任、決策模式

數位落差與公平性

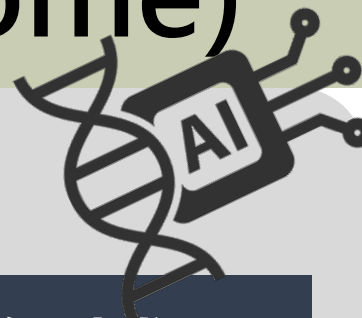
族群差異、數位資源可近性
數位技術理解基礎

AI 技術特性與限制

可解釋性、偏差警示
決策建議過程透明性

適應性AI監管架構 (AI Genome)

Goktas et al., 2025



基礎特徵

結構化

由多項核心模組組成
如：資料治理、透明度
要求、監測機制等

適應性

可隨新興風險動態調整
如對資訊攻擊、
深偽醫療內容而更新

跨領域

超越技術性框架
整合多方專業
共同建構治理模式

AI Genome 醫療領域應用功能

全程監測

持續追蹤模型
在不同族群與
臨床環境的實
際效能

偏誤風險 管理

檢視公平性調
整模型，降低
弱勢族群風險

可解釋性

提供臨床可理
解資訊，使決
策更安全且
提升信任

法規適應

遵循國際與各
國規範，確保
醫療 AI 應用
合規部署

共同參與

納入患者與臨
床端意見，使
AI 更符合需求
與文化

提升AI透明度與可解釋性(XAI)策略

Goktas et al., 2025

透明度強化臨床信任

醫療 AI 若缺乏透明度，難以讓臨床醫師了解其決策過程。提升模型決策流程的可視性，能讓臨床端更安心地將 AI 納入照護流程

可解釋性促進實際應用

事後解釋與臨床端可理解的決策工具，有助於提供具體且可衡量的資訊。協助醫療人員理解 AI 的判斷依據，使整體系統融入醫療實務

特徵貢獻分析

運用 SHAP 等方法量化各特徵對模型的影響
協助臨床人員判斷推論
是否符合醫學知識

局部可解釋模型

針對單一預測
提供清楚的解釋
使臨床人員能快速評估
AI 建議的合理性

注意力機制

利用注意力權重標示模型最重視的資料區段，
協助臨床端了解模型聚焦的依據並提升透明度

AI醫療應用主要歸責設計面向

Goktas et al., 2025

問責面	說明	臨床實例
問責與法律責任	需界定行為與法律責任分工	自動手術機器人若出錯，需釐清醫療人員與 AI 開發者各自的法律責任
道德責任歸屬	錯誤涉及道德層面 但 AI 不具道德能力	AI 不具道德判斷能力，人類監督仍必要，例如採用 doctor-in-the-loop 的手術輔助方式
多層級治理	問責需涵蓋團隊、組織 產業與監管等多層面	醫療團隊需落實紀錄、稽核 組織推動安全文化 產業需符合法規與外部監督
醫療AI設計涵括倫理原則	AI 設計與部署須遵守公平 透明與負責等倫理原則	以正式規範明確界定各方責任 降低差異並維持公信力
利害關係人參與	應納入臨床人員、患者 開發者與決策者等多方意見	多元參與可使問責框架 更貼近臨床需求與社會價值

XAI 實例：NHS 乳癌影像 XAI(英國)

- **問題**：深度學習乳癌偵測黑箱
→ 放射科醫師不信任
- **解法**：導入 Grad-CAM、Saliency Map 揭示 AI 如何判讀腫瘤
- **效果**：提高臨床接受度、通過 NHS AI 法規審查
- **重要性**：避免模型對高密度乳房族群偏誤

XAI 實例：FDA 心律不整 AI (美國)

- **問題**：AI ECG 診斷心房顫動(AF)無法提供建議決策理由
- **解法**：要求 Feature Attribution、決策邊界可視化
- **效果**：提升安全性、醫師可理解模型依據
- **重要性**：XAI 成為 SaMD(AI 醫材) 核准的必要條件

XAI 實例：法國 CNAM 健保欺詐模型

- **問題**：黑箱模型標註醫療院所「異常」
引發反彈
- **解法**：使用 SHAP 提供每筆案例的透明解釋
- **效果**：降低法律訴訟、提升制度信任
- **重要性**：大型健保 AI 審查的典範

可信任AI醫療應用量化評估面向

Goktas et al., 2025

透明度 (Transparency)

提升模型可解釋性，協助臨床理解 AI 預測依據

公平性 (Fairness)

確保跨族群診斷一致，降低模型偏差與差異

可歸責性 (Accountability)

遵循倫理規範並強化人類監督，確實追蹤與回報錯誤

安全性 (Safety)

透過錯誤率與對抗測試評估穩定性，降低臨床誤判風險

永續性 (Sustainability)

減少模型能源使用與碳排放，推動綠色醫療 AI

XAI醫療輔助應用實例： 中風治療歷程決策輔助

XAI 醫療應用三大核心

Räz et al., 2025

情境與使用者導向輔助

- 可客製化的複雜度
- 可靠辨識情境與使用者
- 與醫療機構流程整合



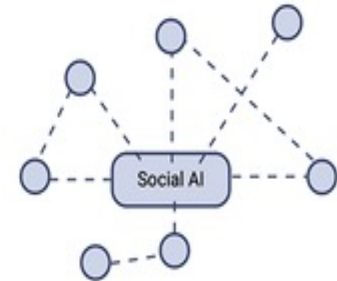
人機互動對話

- 對話對象的表徵方式
- 規劃與目標設定
- 過往互動的記憶



社會文化脈絡適應

- 醫學與社會科學的整合
- 社會動態
- 民眾接受與信任度



隨時間推進，技術發展能促成XAI醫療決策輔助進展

- XAI技術能力的強化，能支援更細緻的臨床決策、更加契合使用者需求，並回應社會對透明度與信任的期待

AI臨床應用三大核心挑戰

Räz et al., 2025

情境導輔助

不同臨床場域與角色對資訊需求差異大

XAI 必須自動調整說明深度，避免資訊過多並提升實際可用性

真實人機互動

XAI 需理解使用者背景、情境調整說明方式
並具備對話記憶，使其成為臨床夥伴而非靜態工具

社會脈絡理解

理想 XAI 能掌握臨床團隊互動脈絡，協助整理意見差異並促進協作
但涉及社會因素與隱私，技術挑戰高

史密斯先生急性中風患者臨床決策

Räz et al., 2025

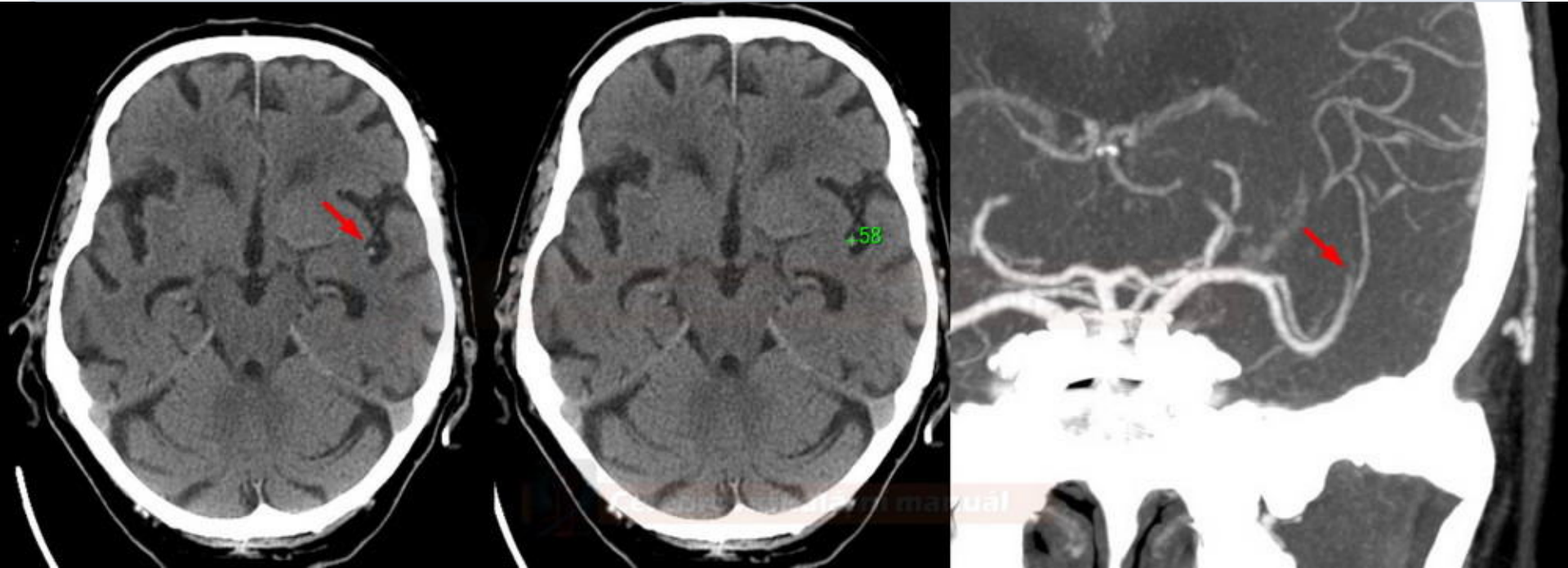
史密斯先生，73 歲男性，於清晨發生急性右側肢體無力與語言表達困難，由救護車送至急診。抵達時意識清醒但反應遲緩，神經學檢查顯示構音不良、右側偏癱，NIHSS 進一步確認中度至重度神經功能缺損。

Phase 1：急診初評，安排影像與評估中風治療流程

Phase 2：影像評估與雙科合作評估灌流缺損範圍

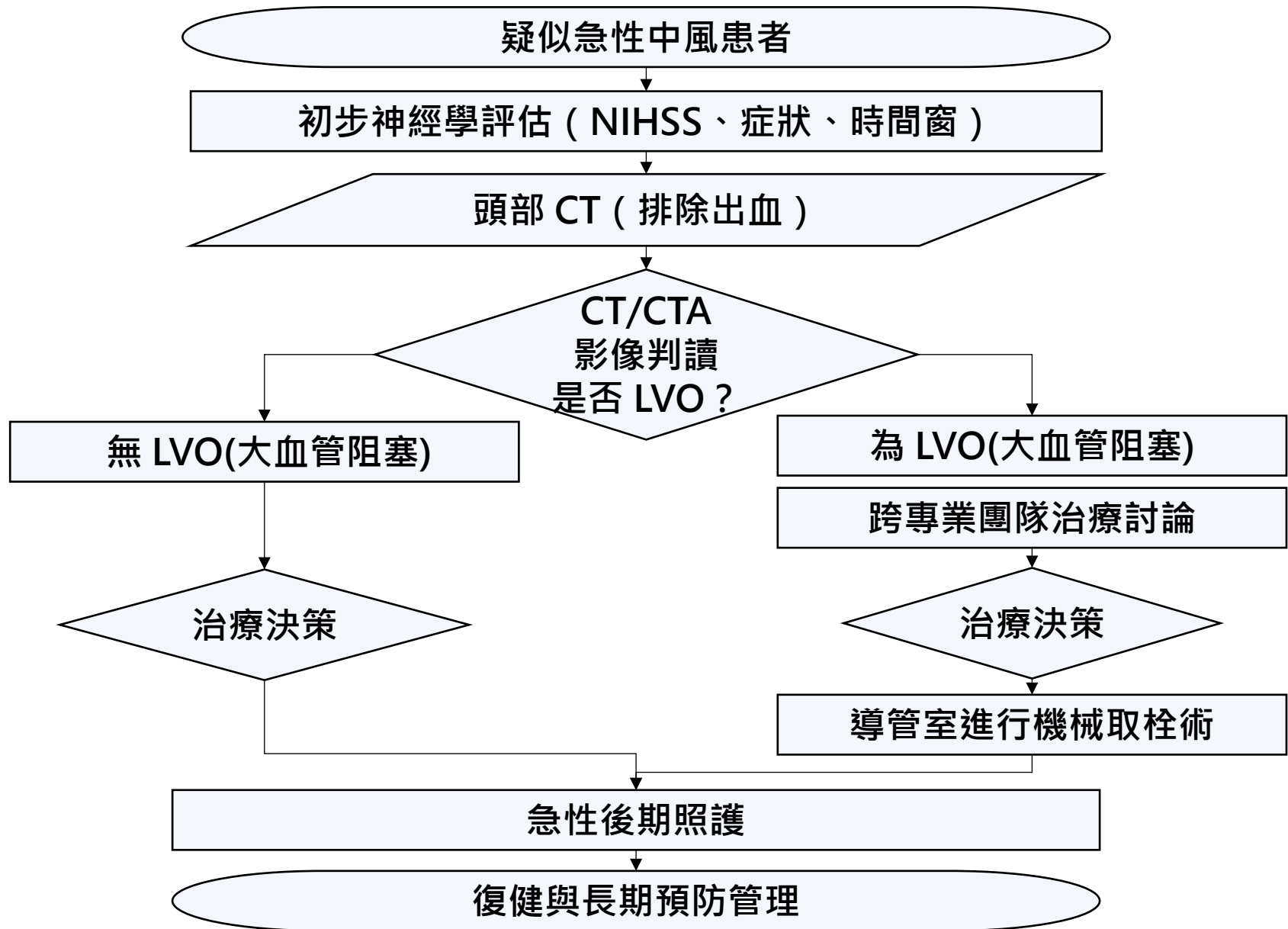
Phase 3：多專科決策，神經內科、介入放射科與其他相關科召開快速討論決定是否合併靜脈溶栓與血管內取栓術療法

Phase 4：住院與後續照護進入中風專責病房持續觀察神經學變化早期復健介入，並整合跨專業團隊



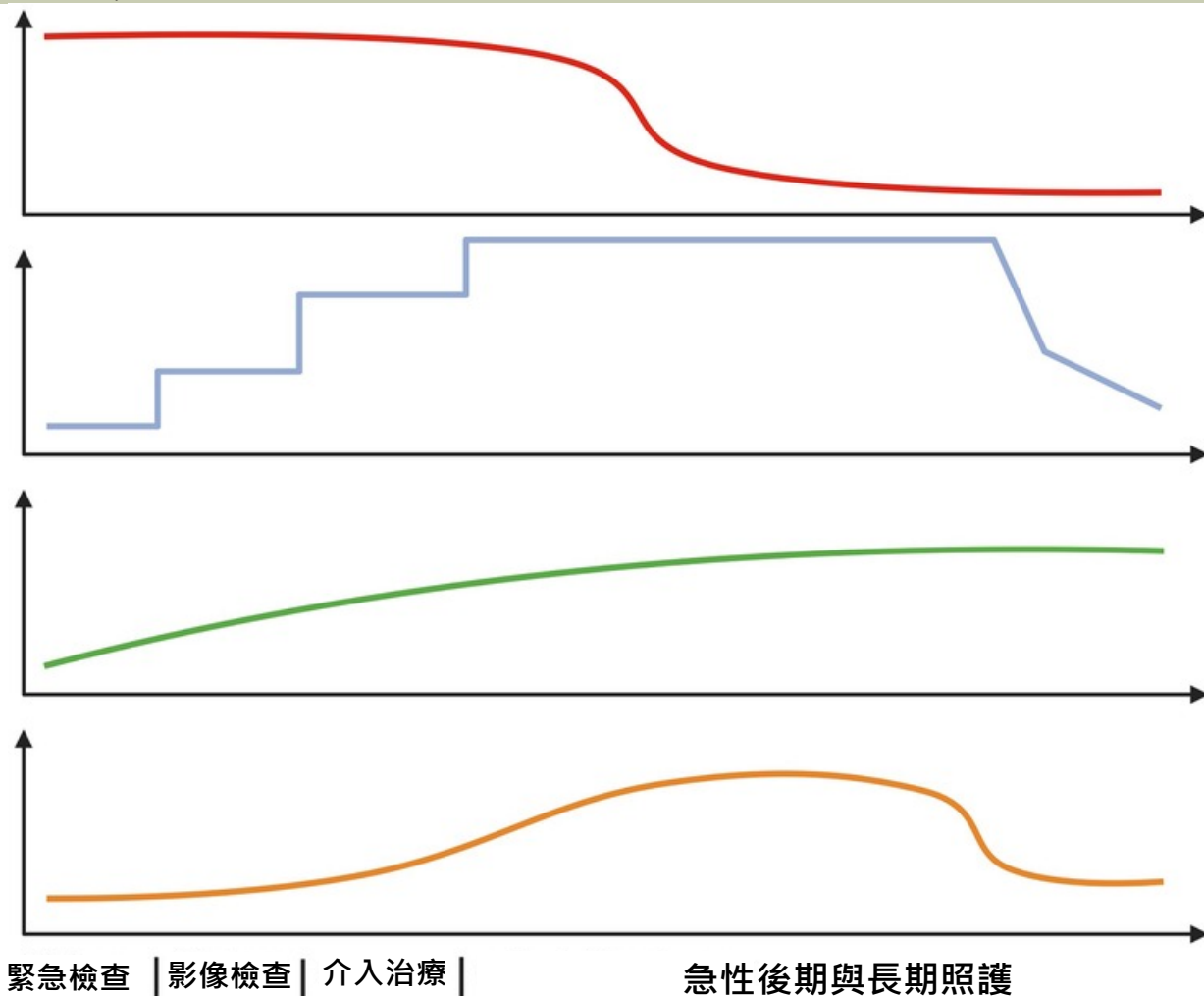
急性中風患者的大血管阻塞評估流程

Räz et al., 2025



臨床歷程中的解釋需求動態變化

Räz et al., 2025



急迫程度

初診與急性介入階段最高
之後逐漸下降

參與決策人數

從單一醫師到多專業團隊上升
再到長期追蹤下降

醫療資訊量

隨病程推進逐漸累積

決策溝通解釋細節

從急診精簡，到住院與復健階段的高需求，再於出院後降低

**XAI 必須能依照病程階段調整解釋方式、內容與深度
以符合臨床現實需求融入臨床照護流程，**

XAI在急性中風決策關鍵輔助

Räz et al., 2025

影像辨識與重點標示

自動標示疑似閉塞區域並以熱度圖呈現風險，協助醫師迅速聚焦重要血管變化

不確定性可視化

以虛線或漸層呈現不確定範圍，使醫師注意爭議區域，降低邊界判讀錯誤

判讀邏輯透明呈現

說明血流差異、對照側比較等關鍵依據，使推論可追溯，區分典型與非典型表現

多模態資訊整合摘要

整合 CTA、灌注、病史與取栓適應症，形成一頁式摘要，強化跨科治療討論效率

同步提示鑑別診斷

呈現可能的替代病因，協助醫師維持臨床思考廣度，避免過度依賴單一影像結果

史密斯先生中風照護XAI輔助臨床流程

Räz et al., 2025

階段	臨床人員	臨床關鍵 / 決策要點	XAI 可能提供的協助
第一階段 初診 (單科決策)	神經科 住院醫師 初步評估	<ul style="list-style-type: none">- 右側無力- 失語- NIHSS = 10 (中度中風)- 疑血管阻塞	<ul style="list-style-type: none">- 自動蒐集並預處理初診資料- 將整理後的資訊即時共享給後續醫師- 以簡短摘要呈現關鍵內容
第二階段 影像檢查 (雙科互動)	神經放射科 住院醫師 加入評估	<ul style="list-style-type: none">- CT 無出血- MCA 阻塞- 大範圍缺血	<ul style="list-style-type: none">- 提供 LVO 偵測初步結果與影像標示- 顯示灌注影像中的不確定性區域- 提醒模型侷限
第三階段 治療討論 (多人會議)	神經科主治 加介入放射科 共同決策	<ul style="list-style-type: none">- 各項資訊整合後決定施行「機械取栓」- 雖討論溶栓，但因先前抗凝治療而不採用	<ul style="list-style-type: none">- 標示未被提及的重要病患資訊- 即時搜尋最新治療指引與文獻摘要- 協助統整各專業觀點，使討論更完整
第四階段 急性後期與長期照護 (大型團隊)	護理、復健、 心臟科、 社工等	<ul style="list-style-type: none">- 追蹤功能改善- 規劃後續復健	<ul style="list-style-type: none">- 協調跨專業照護安排- 收集並分析每日評估資料- 以儀表板視覺化呈現功能進展或惡化- 提供趨勢分析與必要提醒

XAI 臨床應用發展 (1)

Räz et al., 2025

依據情境與使用者需求 提供決策輔助說明

核心挑戰

- 不同情境需求差異大
- 龐大資料若未篩選，易造成醫師認知負荷
- 資歷不同的醫師需要的細節層次不同

XAI 必須具備

- 自動篩選「與當下最相關」的資訊
- 解釋深度可依情境與使用者調整
- 會「記憶」提升回答一致性



針對特定臨床
場景建立可切
換的解釋模板



提供直覺化情
境切換介面，
降低操作負擔

醫療決策輔助 人機對話互動

核心挑戰

- 目前 AI 主要具備「對話行為」，仍不足以支撐真正的臨床對話

XAI 必須具備

- 理解使用者，能辨識醫師經驗與狀態
- 具備目標性，依情境調整溝通方式
- 能延續脈絡：記得前文、維持對話一致性



建立可調整的
臨床對話模式



發展互動回饋
機制強化
應答脈絡

XAI 臨床應用發展 (2)

Räz et al., 2025

XAI 社會文化脈絡感知

核心挑戰


臨床決策高度社會化，牽涉階層、情緒、非語言訊號與團隊互動，若 AI 無法理解隱性脈絡，將難以提供有效的決策協助


注意事項


- 社會脈絡難以建模，需要跨領域合作。
- 過度監測可能降低使用者接受度。
- 模型介入可能改變原有社會互動模式。

AI 可扮演的角色

- 協助偵測不發言、遲疑或壓力過高的團隊成員
- 整理不同意見並呈現利弊，減少溝通落差
- 在高風險情境中提出必要提醒，例如重要發現被忽略

 初期可採用情緒分析、語音與社會訊號分析等能力

 以參與式設計確保工合臨床需求

 採用安全、有限度的社會介入策略



林庭瑀
博士



陳秀熙
教授



國立台灣大學



林家妤



許辰陽
醫師



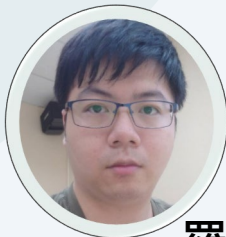
陳虹彦



曾暉哲



劉秋燕



羅崧璋



梅少文 主持人



侯信恩主持人



楊心怡製作人



嚴明芳
教授



陳立昇
教授

台北醫學大學

