



## 星球永續健康線上直播

### 健康大型語言模型 (1)

#### 慢性病預防應用

2025年8月26日

隨著全球醫療體系面臨高齡化與慢性病照護需求日益提高，如何運用當前快速發展晶片產業結合創新人工智慧數位科技提升預防與照護效能發展健康產業成為族群健康領域重要課題。大型語言模型 (Large Language Model, LLM) 快速發展不僅改變了人類與知識互動的方式，也為醫學研究與健康照護帶來契機。本週我們將聚焦於健康領域大型語言模型的建構，並深入探討其在慢性病預防中的實際應用與發展潛力。

#### 星球健康新知

美國、俄國與歐盟及烏克蘭近日尋求和平的可能性。此談判也使得歐盟再次展現出二戰以來少見的凝聚力。美國除了希望推動俄烏停戰，也期盼透過外交努力維持東歐的平衡。而歐洲國家則強調，保障烏克蘭的和平不僅關乎區域安全，更關係與歐洲戰略穩定。此次歐盟領導人包括歐盟委員會主席馮德萊恩、法國總統、英國首相與德國總理，紛紛積極參與展現出前所未有的團結。歐洲領導人與美國一同推動和平，其意義並非僅止於支持烏克蘭，更在於維繫歐洲與俄國之間的力量平衡。美俄歐烏三方會談計畫在兩週內舉行，外界高度關注是否能為停火奠定基礎。雖然美方希望避免克里米亞與北約議題成為談判的障礙，但在歐洲的介入下，和平仍需時間才能醞釀。

以色列與哈瑪斯之間的衝突不僅牽動加薩走廊的未來，也引發美國與國際社會的高度關注。以色列要求一次釋放所有人質，而哈瑪斯則傾向分階段進行人質談判，雙方在協議步調上出現明顯差異。近日以色列對迦薩發動攻擊推進佔領計畫已造成嚴重的人道危機。國際社會透過空投物資來支援加薩，但規模有限，難以真正緩解當地超過六萬人喪生所帶來的巨大壓力。在以色列國內，示威活動持續擴大。許多民眾要求政府立即停火，以換取人質安全，避免衝突進一步惡化。另一方面卡達積極投入斡旋。外交部發言人安薩里召開記者會，表示卡達願意協助推動停火，甚至提出六個月內達成協議的目標。



這不僅是卡達在中東和平角色的展現，也與美國在區域的影響力緊密相關。這場衝突與俄烏戰爭有相似之處，都涉及領土、主權與人道問題的多重拉鋸。如何在衝突與外交之間取得平衡，仍是國際社會必須持續關注的重要課題。

在南亞地區局勢同樣並不平靜。戰火中的緬甸將舉行自 2021 年政變以來的首次大選。長期的內戰使緬甸社會陷入混亂，不僅帶來糧食短缺與災難危機，也迫使數十萬羅興亞人被驅逐、流離失所，造成嚴重的人道悲劇。曾獲得諾貝爾和平獎的翁山蘇姬，如今卻無力改變國內的困境。在更廣泛的亞洲地緣政治上，中國外交部長王毅訪問印度，強調兩國應該是夥伴而非對手。這反映出印度在美國關稅政策下未能獲得實質利益，被迫重新尋求與中國的合作，以平衡其國際角色。這背後也顯示，美中之間的角力，已深刻影響南亞的戰略格局。德國與日本的外交也持續強調在國際危機下戰略夥伴關係的重要性。對日本而言，若僅依賴過去與美國的傳統外交模式，恐怕難以應對未來的不確定性，因此積極尋求在亞洲與歐洲之間建立更廣泛的合作網絡，南亞的地緣政治發展正進入新的變局。

美國總統川普在過去六個月內雖然推出多項經濟與貿易政策，包括高關稅等措施，但其政治支持率卻持續下滑，目前淨支持率已降至-15%，顯示多數民眾對他不再支持。儘管政策聲勢浩大，但在經濟、物價與通貨膨脹等民眾最關切的議題上，成效並不明顯。金融市場的不穩與生活壓力依舊存在，使許多美國人認為川普未能真正改善生活狀況。從 2017 年到 2025 年的趨勢可見，經濟始終是美國民眾最關心的議題，遠超過民權、醫療、移民與國家安全等其他政策領域。即便移民是川普的重要主張之一，他在這方面也未能贏得廣泛支持。值得注意的是，不僅民主黨選民普遍反對，連部分傳統上支持共和黨的州份也開始出現鬆動，顯示川普在國內的政治支持面臨嚴峻挑戰。

在全球科技發展的競爭中，晶片已成為核心戰略資源。近期軟銀注資 20 億美元挽救英特爾，維持美國在晶片設計上的優勢；同時，美國政府也加大投資，鞏固其晶片產業鏈，並透過高關稅政策來維護全球半導體領導地位。美國企圖確保自身在高科技領域處於最前沿，不容他國超越，不過也引發新的晶片地緣政治競爭。英國積極發展自主晶



片設計產業，美中之間的關稅談判也牽動未來合作與競爭的走向，晶片已不再只是經濟議題，更上升為國際戰略焦點。

另一方面晶片的應用進入生命科學與醫學研究領域。美國 FDA、EPA 與 NIH 已逐步鼓勵以晶片技術取代動物實驗。2022 年起，FDA 不再強制要求藥物必須經動物實驗；2024 年，NIH 更宣布不再徵求僅依賴動物模型的研究計畫，並成立專責辦公室推廣「非動物方法」(NAMs)。具體案例包括 Emulate Inc. 的肝毒性晶片 (DILI chip)，在藥物安全測試中能以 87% 準確率識別有害物質，100% 準確率識別無害物質，並已被 FDA 納入 NAMs 試點計畫。此外，結合 AI 與微生理系統 (MPS) 的應用，更能模擬多器官交互作用，強化毒性與安全性的評估。約翰霍普金斯大學開發的 AI 模型 OPERA，準確率更高達 91%。這些成果顯示，晶片不僅能提升研究效率，也有機會在未來大幅減少對動物實驗的依賴。

歐盟在科學研究政策上也出現重要轉變。從 2028 年起歐盟主要科學基金 Horizon Europe 將首度開放支持軍民兩用研究，這是對過去 40 年以來科學民用限制一大突破。傳統上，歐洲對軍事與民用研究有明確區隔，例如部分研究機構的科研平台與民間完全分開，以避免觸及軍事倫理與學術自主的爭議。然而，隨著地緣政治的壓力，特別是俄烏戰爭、中東局勢以及美國在貿易與國防上的強勢作為，歐盟開始更加重視自主防衛能力，推動軍民融合的雙重用途研究。未來七年間，歐盟將投入 1750 億歐元，涵蓋基礎科學、創新領域與 AI 等前沿科技。這樣的轉向引發學界擔憂：是否會因軍事導向而限制學術自由，甚至削弱國際合作？在國家實例方面，德國已有「軍民研究條款」(zivilklauseln)，法國與荷蘭也在 AI 與量子領域展開軍民兩用研究。這些發展顯示，歐洲正逐漸走向軍民界線模糊化的灰色地帶。

廣義相對論與量子力學能否統一，一直是物理學最重要的未解之謎，而其中的核心問題就是重力是否具備量子性。隨著量子科技的進展，科學家近年設計了多項實驗，利用量子疊加與糾纏來檢驗重力理論，包括鑽石晶體的自由落體、玻色愛因斯坦凝聚(BEC)測試，以及 GQuEST 雷射干涉實驗。這些研究期望在未來十至二十年內，能提供重力量



子化的直接證據。舉例而言當自由落體的晶體處於量子疊加狀態時，其重力場可能相互干涉並產生糾纏，若能被觀測到，將為重力的量子化帶來突破性的驗證。雖然這些實驗仍在進行，但已逐步推動物理學邁向新階段。目前，科學界正同時探索大統一理論與迴圈量子重力的路徑；一旦成功，不僅有助於理解黑洞與蟲洞等現象，也可能改變我們對時空結構與時間旅行的想像。

氣候變遷已使歐洲不再是過去人們嚮往的夏日天堂。如今，夏天成為熱浪與能源危機的溫床。極端高溫暴露出歐洲電網的脆弱，一方面要應付夏季冷氣的龐大需求，另一方面冬季又需大量供暖，形成前所未有的挑戰。英國正遭遇連續三波熱浪，希臘大城帕特雷則面臨野火威脅，顯示氣候變遷已對能源與人類安全帶來雙重衝擊。

Google 推出的 AlphaEarth，整合了衛星影像、雷達與多源觀測資料，能生成高解析度的全球環境地圖，應用於再生能源規劃、氣候變遷研究與森林監測。這項模型大幅縮短了資料處理時間，並降低運算能耗。其資料庫涵蓋 2017 至 2024 年的地球基準數據，以 10 公尺方格封裝，精度超越其他 AI 模型。不過，學界也提醒，這些成果仍需公開更多驗證細節，以確保資料的可靠性與延展性。

### 健康領域大型語言模型建構

電影《最後的編劇》(The Last Screenwriter) 以娛樂性電影呈現，但其本質如同科學實驗計劃。計劃主持人彼得路易斯 (Peter Luisi) 邀請 ChatGPT 4.0 完整撰寫劇本，導演僅負責拍攝與剪輯。換句話說，這是一部由 AI 真正完成編劇工作的電影，而背後的核心提示詞為：「寫一個劇本，講述人類編劇意識到自己不如人工智慧的故事。」在這部電影中，主角傑克是一位備受矚目的劇作家。他深信藝術源自人性、情感與生活經驗，無法被複製或取代。但製作人保羅卻為他安排了一位「AI 編劇助理」。起初，傑克抱持懷疑，但在實際嘗試後，他驚訝於 AI 能精準描寫戰爭場景、親情離別甚至環境氛圍，作品細膩而動人。雖然 AI 自稱僅是輔助，卻動搖了傑克對人類創作優勢的信念。隨著合作加深，傑克逐漸沉迷於這種高效的創作模式。AI 不僅能根據大數據生成符合潮流的劇情，還能根據傑克的回饋進行互動式學習，形成人機黃金組合。然而，傑克也逐漸



感到不安：他開始依賴 AI 的靈感與效率，甚至忽略了家庭。AI 更進一步介入他的生活，提醒他休息、甚至建議他如何與妻子相處。這種跨越界線的控制，讓傑克感覺自己被監視、被取代。隨後，製作人提議全程交由 AI 寫劇本，就能獲得高額報酬。傑克掙扎之下仍然答應。結果 AI 在短短半天內完成一部劇本，並署名為傑克創作。電影上映後大獲成功，讓他名利雙收。然而，傑克卻陷入深深的空虛與危機。他發現自己雖然得到了掌聲，卻失去了對創作的主導權與熱情。最終，傑克在家庭崩解、導師離世的打擊下，決定揮棒砸毀 AI 劇本助理，並重新拿起打字機創作。他明白，人類的價值並不在於永遠正確或高效，而在於那些帶有錯誤、矛盾與情感的真實經驗。這些才是藝術與人性的核心。這部電影讓我們反思，AI 在創作領域既是工具，也是挑戰者。它能帶來效率與靈感，卻也可能奪走我們的自我與創造力。真正的關鍵，不在於全然拒絕或全然依賴，而是如何在人與 AI 的互動中，找到平衡與合作的可能。

我們必須先了解大型語言模型的基本運作模式，這樣才能進一步思考醫療應用該如何與它相輔相成。大型語言模型的核心在於「資料」。沒有資料，就不可能生成任何有臨床價值的文本。和文學創作不同，醫療應用需要仰賴龐大而多元的多模態資料，包括：病人的語音對話、醫師與病人的交流紀錄、影像檢查、檢驗報告，甚至是醫師撰寫的病史與日誌。這些異質性資料需要經過斷詞 (tokenization)、轉碼 (embedding)、再透過自注意力機制 (self-attention) 進行分析，才能生成有意義的文本。在訓練過程中，LLM 先透過非監督式學習 進行大規模預訓練，之後再經過監督式學習與精細微調 (fine-tuning)，讓模型能更貼近臨床需求。而最大的突破在於 回饋學習：當我們給予提示 (prompt) 時，模型能夠即時調整，並根據回饋持續優化。這就是它和傳統 AI 最大的差異。因此，今天看到的不只是 ChatGPT，還包括 LLaMA、Claude、Gemini 等模型，正迅速發展並應用於各種領域。在醫療上，它們不只是幫助生成臨床紀錄，還能進一步支援 臨床推理與診斷，讓醫師能更快更精確地完成決策。

大型語言模型在醫療領域的應用已不限於臨床決策，也能支援醫學教育與跨語言翻譯。以醫學生學習醫學拉丁文為例，過去多依靠死記硬背，如今透過 LLM 即可即時理



解與轉譯。在影像辨識方面，COVID-19 疫情已展現其價值，AI 工具能協助快速判讀醫學影像，這在過去幾乎難以想像。LLM 亦能生成電子病歷與醫療紀錄，減輕醫師每日手寫病例的負擔。然而，這同時帶來新挑戰：有人可能會質疑這是否構成「偽造文書」。若僅以取代的角度看待 AI，反而會被工具綁住；相反地，若能建立標準並將其定位為輔助工具，便能有效提升醫師工作效率與生活品質。此外，大型模型能自動生成醫療對話摘要，免去人工錄音與轉譯的繁瑣程序；也能快速處理醫療問答，針對專業人員與一般民眾提供不同層次的回覆。最終，它將成為臨床決策與行政任務的輔助利器，並進一步強化醫療安全監測，展現健康領域專屬 LLM 的完整價值。

在健康領域，大型語言模型的建構流程雖然與一般 LLM 類似，但仍有幾個關鍵差異。首先，需要持續的預訓練，並仰賴龐大而多元的資料集。這些資料包括知識圖譜，例如解剖學，過去我們必須抱著厚重的教科書到教室學習，如今模型就能直接進行知識連結與檢索。除此之外，還有真實世界的數據、各式問答資料、醫學教科書以及臨床文本，這些都是模型的重要養分。當模型具備基礎知識後，就會進入「專業資料集」的階段，包含單輪對話、多輪對話、醫學 NLP 任務，以及一般醫病對話。醫療現場並不是單一答案，而是充滿互動與變化，醫師與病人的交流常常會隨著對話而影響診斷與治療策略。因此，大型語言模型必須透過監督式微調，學會掌握這些複雜的互動脈絡。最後，訓練不能停留在一次性完成。就像《最後的編劇》這部電影，如果主角傑克願意持續給予 AI 回饋，AI 就能不斷修正與優化，創造出更符合需求的作品。在醫療應用上，使用者回饋與強化學習同樣至關重要。醫師、病人與研究人員都能成為回饋來源，幫助模型逐步最佳化演算法，最終生成真正符合臨床需求的紀錄與決策支持。

提示優化與健康摘要的生成，在臨床應用中是一項重要挑戰。關鍵在於如何區分必須凍結保留的資訊，以及可以依情境動態改變的內容。病歷與入院紀錄屬於基本資料，不能隨意修改，這部分被稱為「凍結提示」，也就是原始資料輸入後，模型在生成時必須完整保留。相對地，臨床過程中的觀察紀錄，或病人隨訪時的對話，則可以透過提示優化進行調整與更新。這樣的設計能在確保醫療資料正確性的同時，保留彈性，以生成



符合個別病人情境的摘要。提示優化不僅是技術層面的調整，更是醫療應用中在正確性與靈活性之間取得平衡的關鍵機制。

分散式運算架構在健康領域的大型語言模型中具有關鍵地位。醫療資料龐大且多模態，包括病歷文本、影像數據與實驗室報告，累積起來的規模極為龐大，單一 GPU 幾乎不可能有效完成訓練與推理。因此，需要依靠分散式平行運算。透過多顆 GPU 的協同運作，模型能同時處理不同維度的資料，並行運算，大幅提升效率。這樣的運算基礎，使大型語言模型能真正落地應用於臨床現場。它不僅能協助醫師快速生成病歷，還能用於題目設計、研究分析，甚至支援互動式醫病諮詢。分散式運算因此被視為 LLM 在醫療場景中應用的基礎設施。

醫師考試是大型語言模型在醫學教育中最具代表性的應用場景之一。醫師執照考試題庫龐大且內容複雜，若完全依靠人工建置，需要投入大量時間與人力。大型語言模型的導入，能夠自動生成題目，並經過品質檢查與專業審查。這個過程能確保題目的正確性與一致性，不合格的題目會被篩除，符合標準的題目才會納入考庫。因此，LLM 在醫學教育中的角色並非取代專家，而是作為高效率的輔助工具。它能讓醫師專注於審查與評估，而不必將時間耗費在題目產製上。這樣不僅提升效率，也能增加題目的多樣性與整體品質。

LLM 在健康問答上的應用仍存在挑戰。若模型生成的問題涉及自傷或極端行為，這類內容在臨床互動中可能帶來風險，甚至誤導病人，被視為「高風險問答」，不宜直接使用。經過強化學習後，模型能被限制在合理範圍內，只生成安全且具解釋性的問題。例如「你是否因為煩躁而感到困擾？」這樣的問句符合醫學標準，也能幫助醫師更精確地理解病人狀態。由此可見，LLM 的問答能力並非天然可靠，而是必須透過持續調整與強化，才能在臨床互動中發揮真正價值。

在精準醫療領域，大型語言模型的應用可歸納為四個面向：Predictive（預測）、Preventive（預防）、Personalized（個人化）、Participatory（參與）。在預測方面，模型能根據病歷與數據推估疾病風險；在預防方面，可協助設計健康檢查流程，進行早期介入；



在個人化方面，能生成符合不同病人需求的臨床紀錄與治療建議；在參與方面，則能提升病人與醫療團隊的互動，讓病人更積極投入健康管理。大型語言模型不僅是一項技術工具，更是推動智慧醫療的核心動能，既能提升醫師的效率與病人照護體驗，也能強化整體醫療體系的安全與品質。

### 健康大型語言模型慢性病預防應用

在慢性病預防的應用上，大型語言模型能協助醫師設計臨床劇本，提供個人化的預防指引。慢性病的本質是生活習慣的累積，例如高血糖、高血壓與高血脂，這些狀態都與日常作息密切相關。因此，臨床劇本的核心任務就是根據個別差異，建構專屬的預防方案。健康檢查是最常見的入口。透過抽血檢驗，可以得到血糖、腎功能、肝功能、血脂等生化數據。這些指標雖然重要，但對病人而言往往過於專業，必須藉由視覺化介面轉換為簡潔的呈現。然而，即使資料呈現更加直觀，專業判斷仍然不可或缺，特別是當需要結合多模態訊息時。這些訊息除了傳統的檢驗數據，還包括心電圖、心臟超音波、心音訊號，語音對談的分析。

與電影劇本不同，臨床劇本必須建立在結構化與非結構化資料的整合之上。文字型的病歷、生物標記的檢驗結果、影像檢查的輸出、心電圖等序列訊號、藥物使用紀錄、生活事件與社群環境，甚至 IoT 裝置所收集的連續數據，都是重要的組成部分。這些多模態的臨床實證資料，才能支撐個人化的預防策略。其中，醫學知識屬於「不可變動」的部分，例如人體結構、代謝機制與血糖調控；醫學研究則是「部分可變動」，多數已有實證基礎，但仍持續隨著新證據更新。大型語言模型結合這些不同層次的資料，再透過專業資料集與持續學習，能逐步演進，生成更精準的臨床劇本。

慢性病的血壓會從正常狀態逐漸進展到高血壓前期，再到第一期與第二期高血壓；血糖則從正常狀態發展至糖尿病前期，最後成為糖尿病。這是一個多重時間軸的過程，每個疾病階段都像是劇本中的不同角色。以往必須透過定期健檢才能獲得相關數據，如今藉由連續血糖監測 (CGM)、智慧型手錶或 IoT 血壓計，即可即時追蹤變化。體重、活動量、飲食習慣等生活型態資訊，也能同步被上傳並整合。例如病人若血壓曲線中出



現多次超過 140/90 mmHg 的紀錄，同時在 CGM 數據中顯示血糖波動超過 126 mg/dL，便能判斷其已進入高血壓與糖尿病的前期狀態。在這樣的情況下，大型語言模型可結合病人的性別、年齡、家族史、個人病史與生活習慣，撰寫個人化的臨床劇本，並提供具體的預防建議。模型必須正確理解病人的提問。例如「我是不是高血壓個案？」這類問題，需要經過分詞、編碼與向量化處理，將文字轉換為數字序列，再進入模型進行語意判斷。經過數位化的處理後，大型語言模型能進行語意預測與學習。例如，當「高血壓」中的「高」被遮罩時，模型必須理解其與其他詞彙的關聯，才能推斷完整的語意。這個過程需要依靠平行運算與 GPU 的算力，並透過遮罩機制 (Masking) 來建立字詞間的關係網絡。只有在充分學習實證資料後，模型才能判斷文字是否屬於醫療語境，並正確辨識像是「我是不是高血壓個案？」醫療相關提問。

在語意理解的基礎上，下一步是分類與風險判定(NER)，這正如醫師在診間所做的判斷。當病人帶來抽血檢查報告與病史描述時，醫師會依據血壓、血糖、血脂等指標進行分群，再結合病人的性別、年齡、生活習慣與家族史，提出個別化建議。大型語言模型透過自然語言處理與風險分群，也能模擬這樣的過程。舉例來說，有些病患可能是血糖與 HbA1c 偏高，有些則是血壓與血脂異常；經過分群後，就能更精準地對應到個人化的建議。這樣的運算過程在人類腦中由經驗與直覺完成，而在 AI 中則仰賴演算法與 GPU 平行計算。經過分類與風險評估後，模型可結合醫學詞典、研究文獻與知識庫，建立可持續更新的互動架構。病患或醫療專業人員在提出問題時，模型能即時呼叫相關知識，並透過多模態資料包括檢驗數據、心音訊號、語音特徵，進行回應與建議。

以一位檢查結果顯示血壓 139/87 mmHg、空腹血糖 119 mg/dL、HbA1c 5.7%的病人為例，模型會將其歸為「代謝風險群」，同時結合 BMI、生活習慣（如飲酒、嚼檳榔）、教育程度與既有病史，進行個人化風險評估。若判斷屬於高血壓與糖尿病前期，模型便能推算五年內進展到第一期或第二期高血壓的機率，並生成個人化的臨床劇本。這個劇本不僅包含風險描述，還需要提供具體的改善方案。例如，運動方面可建議每週 150 分鐘中等強度有氧運動，搭配阻力訓練；飲食方面則建議採用 DASH 飲食，降低鹽分與精



緻糖攝取，增加蔬果與膳食纖維；若生活調整後血壓仍未控制，則需考慮藥物治療。這樣的過程就如同在電影創作中修改劇本情節，醫療劇本也能根據個人風險因子調整，最終達到降低慢性病風險的目標。大型語言模型是協助醫師更有效率地完成風險分析與建議生成。人類醫師仍是決策的核心，但 AI 提供的臨床劇本能成為可靠的輔助工具，提升診療品質與病人健康管理的精準度。

以上內容將在 2025 年 8 月 26 日(二) 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](https://www.realscience.top/)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](mailto:happy82526@gmail.com)

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