

AI新藥開發的 典範轉移與 未來融合生態系

從數據驅動的標靶發掘到自主驅動實驗室

標靶發掘、新藥設計、臨床驗證，以及法規

目錄

☒1☒ AI新藥開發的序幕與典範轉移

- 1.1 研發典範轉移：克服死亡谷(Valley of Death)
- 1.2 生命科學的人工智慧：基礎模型 (Foundation Models)
- 1.3 多重組學整合與新型生物標記物發現

☒2☒ AI主導的新藥設計與物質發掘

- 2.1 3D蛋白質結構預測與基於結構的藥物設計 (SBDD)
- 2.2 基於深度學習的候選物質生成 (De Novo Design)
- 2.3 in silico ADMET 及毒性預測系統

☒3☒ 次世代模式與智慧藥物傳輸系統

- 3.1 生物藥物與多重標靶治療藥物設計 (ADC)
- 3.2 天然物新藥資料庫的AI解讀與混合式發掘
- 3.3 奈米醫學及智慧標靶遞送系統 (Nanomedicine)

☒4☒ 臨床前與臨床驗證及精準醫療的革新

- 4.1 臨床前研究創新：動物實驗替代及生物學建模
- 4.2 數位雙生 (Digital Twins) 與次世代臨床試驗的現代化
- 4.3 精準腫瘤學 (Precision Oncology) 與病患客製化治療

☒5☒ 法規科學、數據倫理與未來融合生態系

- 5.1 AI新藥開發的監管框架與品質管理治理
- 5.2 資料倫理、演算法公平性與聯邦學習生態系統
- 5.3 製藥產業的未來：自主驅動實驗室 (Self-driving Labs)

第 1 部

AI新藥開發的序幕與典範轉移

1.1

研發典範轉移：克服死亡谷(Valley of Death)

2022年2月，特發性肺纖維化治療藥物候選物質ISM001-055首次被投藥於人體。這種肺部逐漸硬化的疾病缺乏適當的治療方法，大部分患者在確診後3年到5年內便會離世。隨著肺部組織變成僵硬的疤痕，呼吸會變得越來越困難，但連確切的病因都還沒有被查明，因此正如其名，被加上了「特發性」一詞。到那時為止，世界上能延緩這種疾病進展的藥物只有兩種，而且這兩種藥物都只能在一定程度上延緩進展，無法完全阻止。製造出這種物質的是總部位於香港的Insilico Medicine。研究團隊利用標靶發掘平台PandaOmics，找出了調節引發纖維化之信號傳導過程的酵素TNIK作為新的標靶，並利用化學分子生成引擎Chemistry42，設計出了完全契合該酵素的分子結構。

這是專案啟動不到30個月就發生的事情。如果是傳統的方式，光是尋找標靶並篩選出一個候選物質，就會是需要花費4年到6年的階段。這個後來獲得Lentosertib正式名稱的小分子，就這樣在電腦螢幕中誕生，並走進了人類的血管之中。在那之後的3年，這個物質成為了再次吸引整個業界目光之消息的主角。

長期以來，製造新藥就像是隔著山谷的角力。山谷的這一邊，數萬個在實驗室裡備受期待的候選物質排成一列；而在山谷的那一邊，則放著幾顆實際處方給患者的藥物。在跨越這段距離時掉落的候選物質佔了壓倒性的多數，因此業界一直將這個區間稱為「死亡之谷」。傳統的新藥開發從尋找標靶的那一刻起，到獲得監管機構的批准為止，平均需要花費10年到15年的時間，而加上所有失敗的候選物質，總成本已超過26億美元。

這個金額不僅包含了一個成功新藥的開發費用，還包含了無數個一起被嘗試卻失敗之候選物質的沉沒成本。儘管如此，能夠跨越山谷並進入市場的物質，還不到最初候選群的10%。剩下的物質因為療效不足，或是暴露出意想不到的毒性，又或者在體內被分解和排出的方式與預期不同，而在中途倒下。

山谷會變得如此之深，是有原因的。新藥會先在老鼠或猴子等動物身上進行試驗。然而，動物的身體和人類的身體對同一種物質的接受方式並不相同。在老鼠肝臟中能順利被分解的成分，在人類的肝臟中可能會變成毒性物質；在猴子身上確認到的效果，在人類身上可能連一半都顯現不出來。甚至就算是同為人類，也會根據年齡、基因以及已經患有的其他疾病，對同一種藥物產生不同的反應。這種試圖將動物實驗結果轉移應用到人類身上的轉譯研究過程本身，就帶有與生俱來的極限。

如果用小學生也能懂的比喻來解釋，這就類似於把對小狗有效的訓練方法直接用在小貓身上的情況。雖然外表看起來相似，但因為物種不同，體內的反應也不一樣，所以許多通過了動物試驗的候選物質，在以人類為對象的首次臨床試驗階段就直接失敗了。業界為了解釋這種失敗，會使用「物種間障礙」這個詞，意思是對老鼠有效，但對人類無效。

這種失敗的重複在統計數據上也表現得非常明顯。獲得一種新藥批准所需的費用，大約每9年就會翻倍一次，但相對於相同的投資額，獲得批准的新藥數量反而減少了。研究人員將這種趨勢命名為Eroom's Law。這是將表示半導體性能持續提升的Moore's Law的字母倒過來拼寫而成的詞。如果說Moore's Law是技術發展的象徵，那麼Eroom's Law就是製藥產業半個多世紀以來無法擺脫的生產力倒退的象徵。

專家們指出，其原因在於臨床試驗的規模越來越大，監管門檻也變得更高，而且容易開發藥物的疾病大部分已經被有效的藥物佔據，剩下的疾病就越難以攻克。也有人指出，新出現的疾病標靶往往結構複雜，剩下的都是難以處理的東西。隨著投入的資金越來越多，而成果卻越來越少的狀況長期持續，製藥公司們便無法輕易投入風險負擔大的罕見疾病或難治性腫瘤的研究之中。

而成為扭轉這種趨勢之關鍵的，就是數據。基因組和蛋白質資訊、記錄每一個細胞反應的單細胞數據、患者的電子健康紀錄和影像資料，甚至每天湧現的數千篇醫學論文，人類大腦無法全部讀完的資訊已經累積了下來。讓人去一一閱讀這些資訊並將它們相互連結起來，這在規模上原本就是不可能的。過去的新藥開發在很大程度上依賴於研究人員的直覺和大量篩選，也就是隨機測試數萬個化合物，並尋找偶然碰上的物質的方式。

人工智慧正在學習這些龐大的數據，並將研究方式本身轉向提前預測疾病的原因和藥物的反應。隨著隨機搜尋標靶的探索轉變為以數據為基礎的預測，主導新藥開發第一階段的試錯比例已明顯減少。

將這兩條管線並排對比，差異會更加鮮明。在傳統的管線中，需要投入10年到15年的時間和超過20億美元的資金，並經過化學家和生物學家一步步手工確認的順序驗證，而在那終點，十之八九會以失敗告終。在人工智慧引領的管線中，這個時間縮短到不到一半，初期探索投入的成本大幅降低，而且在多個階段同時運作的結構上，生成式人工智慧會持續完善候選物質本身。

這不僅僅是節省了時間和金錢，而是連從一開始就篩選出失敗可能性很高之候選物質的切入方式本身都發生了改變。與其在事後付出高昂的代價來確認失敗，不如將重心轉移到便宜且快速地提前確認上。

傳統的新藥開發流程遵循靶點發現、化合物篩選、優化、臨床前試驗、臨床第一期、第二期、第三期依序排列的順序。前一階段必須完全結束才能進入下一階段，所以如果任何一個階段出現延誤，隨後的所有階段都會一起延後。引入人工智能的流程大大改變了這個順序。尋找靶點的工作、設計候選物質的工作，以及預測該物質在體內如何運動的工作，現在可以在多台計算機上同時進行。

生成型人工智能創造出數千種分子結構方案的同時，其他研究團隊已經在模擬成功機率較高的候選物質的毒性。原本一個接一個排隊等待的工作，現在變成了多個地方同時進行的工作。這種平行結構是將開發時間縮短至原來一半以下的實際動力。

效果用數字表現出來。傳統上尋找新疾病靶點並選出相應的第一種化合物需要多年時間，但採用人工智能平台的生命科學公司已經將這段時間縮短到12至18個月。實際需要合成進行測試的化合物數量也從數萬個大幅減少到數百個。全球知名管理顧問公司預計，如果將人工智能引入新藥開發的整個過程，製藥行業每年將產生600億至1100億美元的經濟價值。這個價值的大部分來自於減少候選物質在晚期被淘汰時所造成的沉沒成本，其餘部分來自於開發時間縮短，使藥物能在專利保護期內銷售更長時間。

只看Insilico Medicine一家公司，從靶點發現到臨床前候選確定需要13至18個月，所需費用約為260萬美元。如果用傳統方式，同樣的步驟需要2年半到4年，花費數

千萬美元。僅看2025年這一年，投入人工智能新藥開發公司的風險投資資金到第三季度已達27億美元。資本率先搭上了這波浪潮。

這種節省效果改變了整個行業的計算方式。過去，對於患者數量只有數千人的罕見疾病，很難找到投入超過10年的時間和超過20億美元資金的理由。由於失敗的可能性很高，很少有製藥公司願意為這麼少的患者承擔那樣的風險。但隨著初期發現成本降低到數百萬美元水平，以及失敗可以在更早的階段篩選出來，那些之前被忽視的罕見疾病、小兒癌症和抗生素耐藥菌治療藥物的研究又重新獲得了資金。

這股流動之所以重要，不僅是因為統計數據或股價。像特發性肺纖維化這樣患者數量少、在新藥開發優先順序中被排擠的疾病，現在因為初期開發成本降低而重新出現在研究人員的辦公桌上。隨著跨越山谷的橋樑變得更寬，通過這座橋樑抵達患者的治療方案種類也會隨之增加。

生成型人工智能是指不僅僅是找到一個正確答案，而是根據條件自己創造新結果的人工智能。就像被要求畫畫時創作新圖畫、被要求寫句子時創作新句子一樣，在新藥開發中則是創作與靶點相匹配的新分子結構。基礎模型不是只學習某一種特定疾病的模型，而是指廣泛學習基因、蛋白質、化合物和論文等多方面數據的基礎模型，可以用於多個問題。

一旦創建了一個模型，就可以將其用於尋找靶點、分子設計、毒性預測等多種任務，這比每次都從頭創建新模型的效率要高得多。

這兩種技術實際設計新藥的方式用鑰匙的故事來解釋就容易理解。以前的虛擬篩選是在已存在的數百萬個化合物庫中一個接一個地測試是否與靶點相符。相反，最近的生成型人工智能會查看鎖，也就是靶點蛋白的三維結構，然後為那個孔洞精確設計並製作一把新鑰匙。研究人員使用生成對抗網絡、擴散模型、圖神經網絡等技術，在計算機中提前模擬藥物的結合力、選擇性，以及在體內被吸收、分解和排出的過程。

這種從頭開始設計化合物的方式被稱為從頭設計。新藥候選發現已經從尋找現有物質的探索領域進入到創造新物質的創新領域。因為即使是世界上不存在的分子，只要恰好與靶點相符，就可以成為候選物質，所以從一開始可能找到的候選物質的範圍就大大拓寬了。

有一個早期案例展示了這種設計方式的速度有多快。英國人工智能新藥開發公司 Exscientia 與日本的住友大日本製藥合作，設計了強迫症治療候選物質 DSP-1181。該物質是為血清素受體的一種類型設計的，研究開始後不到一年就完成了，而用傳統方式的話平均需要4年半。

2020年1月，DSP-1181在日本進入了首次人體臨床試驗。這是計算機設計開始12個月後發生的事，當時製藥業界對這個消息反應很大。不過這種物質之後開發被中止，這也是一個遲到的例子，說明快速通過山谷初期並不能保證安全到達另一邊。

解讀結構的人工智慧也隨之發展。能夠以原子為單位預測蛋白質三維結構的 AlphaFold 系列，甚至找出了過去被認為難以作為標靶的蛋白質隱藏結合部位，大幅提升了分子對接模擬的速度。開發 AlphaFold 的 Demis Hassabis 與 John Jumper 也因此貢獻獲得了 2024 年諾貝爾化學獎。隨著半個多世紀以來一直是生物學難題的蛋白質結構預測問題被人工智慧解開，新藥設計的第一步 -- 掌握標靶結構所需的時間，從幾個月縮短到了幾天。Google 推出的治療藥物專用語言模型 Tx-LLM，在單一模型中共同學習分子結構、標靶蛋白質與臨床文獻，在 66 項新藥開發評估任務中，有 22 項創下了最高水準的效能。

2025 年 4 月，Google 推出了將該模型進一步發展的開源模型 TxGemma，在同樣的 66 項任務中，有 50 項取得了與既有最高水準相當或超越的成績。處理整個基因組的基礎模型 Evo 也同時解讀單細胞反應與蛋白質間的交互作用，正在將尋找標靶的過程自動化。

像這樣在電腦中設計出來的分子，現在幾乎不需要經過人手，就能製造成實際的物質。在機器手臂搬運試劑並調節反應條件的自動化實驗室裡，日以繼夜地嘗試著人工智慧所提出的數百種合成路徑。測量合成物質效果的結果會立刻反映在下一次的設計中，形成了設計、合成與測量在沒有人類介入的情況下不斷重複的封閉迴路系統。由標準化的機器人代替實驗，減少了因人而異的手法偏差，也連帶減少了在相同條件下重複實驗時結果出現分歧的問題。

研究人員親自拿著試劑瓶熬夜做實驗的景象，正轉變為人類只需設定目標，機器就會在夜裡代替進行數百次實驗的景象。這種自動化研究所將平行處理的範圍，從電腦螢幕外延伸到了實際的實驗桌上。由於即使在人類熟睡時實驗也不會停止，實驗

室原本每天只有八小時左右的實際運作時間，也因此大幅增加了。

然而，在電腦螢幕上看似完美的設計圖，若要變成實際的建築物，仍然需要漫長的施工期。即使是透過電腦模擬修飾的虛擬化合物，進入活體體內後，仍可能出現意料之外的脫靶反應、肝毒性或心臟毒性，因此必須經過體外細胞實驗、動物實驗以及利用類器官進行的臨床前驗證。

儘管人工智慧將尋找候選物質的時間從幾年縮短到幾個月，但在人體內確認安全性與療效的臨床第一期、第二期與第三期試驗，因為攸關患者的生命，是絕對不能省略的。這項驗證短則數個月，長則需要數年。這意味著我們雖然可以縮窄溪谷的寬度，卻無法消除溪谷本身。

臨床前驗證主要是一個確認五個項目的過程。藥物在體內如何被吸收、分佈到哪些器官、在肝臟如何被代謝、如何排出體外，以及是否會引起意料之外的毒性，這五者統稱為吸收、分佈、代謝、排泄與毒性。過去必須透過動物實驗與細胞實驗一一確認這五個項目，但現在人工智慧模型光看分子結構，就能先給出相當接近的預測值。

雖然這項預測無法完全取代實際的實驗，但光是能在放上實驗桌前，就先篩選掉一開始成功可能性就很低的候選物質，就足以同時減少研究人員所花費的時間以及動用的實驗動物數量。

候選物質在溪谷中失敗的方式也各有不同。有些物質在培養皿的細胞中能很好地附著在標靶上，但在人體內卻無法到達目標組織就消散了。有些物質在肝臟中分解得太快，藥效還沒出現就被排出體外。還有一些物質原本只想針對單一標靶，卻連帶影響了結構相似的其他蛋白質，引發了意想不到的副作用。

人工智慧針對這些失敗類型分別建立了對應的預測模型，讓人們從初期設計階段，就能估算候選物質在哪個階段失敗的可能性較大。過去在實驗室裡需要花費好幾週才會顯露出來的這些弱點，現在有越來越多情況是在描繪分子結構的階段就能事先察覺了。

2023 年出現了一個讓人深刻體會到這項事實的慘痛案例。Sumitomo Pharma 與 Otsuka Pharmaceutical 共同開發的思覺失調症治療候選物質 Ulotaront，是透過

PsychoGenics 基於人工智慧的行為分析平台所發掘出來的物質。與現有的思覺失調症藥物不同，它因為針對多巴胺受體以外的其他神經傳導物質受體，有望減少副作用而備受期待。然而在 2023 年 7 月，名為 Diamond 1 與 Diamond 2 的兩項臨床第三期試驗中，Ulotaront 並未能比安慰劑更明顯地改善症狀。

安全性本身雖然沒有問題，卻在必須證明藥效的最後一關失敗了。也有分析指出，由於與 COVID-19 流行期重疊，服用安慰劑的患者群症狀甚至比往年有顯著改善，這使得結果的解讀變得更加困難。這場失敗顯示，無論人工智慧預測得多麼精準，也無法完美控制個別患者的反應與安慰劑效應等臨床現場的變數。

為了解決在臨床階段發生的這種不確定性，業界所關注的解決方案是數位分身與適應性臨床試驗設計。數位分身是一項將單一患者的遺傳學與病理學特徵，原封不動地轉移到電腦中，創造出虛擬雙胞胎的技術。在實際投藥之前，先對這個虛擬患者套用各種劑量與治療情境，藉此模擬腫瘤的異質性或藥物抗藥性會發生什麼樣的變化。

這種方法將患者進行細緻的劃分，向預測會有良好反應的人優先測試藥物，使患者分層變得更加精確。利用基於真實世界數據建立的虛擬對照組，來同時減少臨床試驗的規模、時間與倫理負擔的嘗試也正在增加。不需往返醫院即可參與的去中心化臨床試驗，同樣在這種趨勢下不斷增加。

臨床試驗延遲的常見原因在於未能及時招募到參與者。甚至有統計指出，超過 80% 的臨床試驗會因為患者招募延遲而導致進度落後。人工智慧透過掃描電子健康紀錄與基因體數據，提前找出符合條件的潛在參與者，並從試驗設計階段就提出容易招募的標準，正以此方式來減少這個瓶頸。

得益於這些需求，專注於臨床試驗的人工智慧市場規模，預計將從 2024 年的 28 億美元成長到 2032 年的 540 億美元以上。就像拓寬跨越山谷的橋樑一樣，能及時將人們聚集到橋樑上，同樣左右著新藥開發的速度。

而在 2025 年，出現了第一個實際跨越這座山谷的案例。在 2022 年進入臨床一期的 ISM001-055（這種物質現在被命名為 Lentosertib），其臨床 2a 期結果於同年 6 月 3 日發表在國際期刊《Nature Medicine》上。結果並非透過企業新聞稿，而是刊登在經過同儕審查的學術期刊上，這意味著它通過了獨立專家的驗證。這項以 71 名

特發性肺纖維化患者為對象、為期 12 週的試驗，採用了醫生和患者都不知道誰真正服藥的雙盲方法，再加上以假藥作為對照組的安慰劑對照方法。Lentosertib 在所有劑量下都是安全的，且隨著劑量增加，肺功能出現了明顯的改善。

每天接受 60 毫克劑量的患者群，其肺活量指標平均增加了 98.4 毫升，而接受安慰劑的患者群則反而減少了 20.3 毫升。這個結果也在同年的美國胸腔醫學會學術大會上發表。這是在電腦中誕生的標靶與分子，在實際患者體內轉化為肉眼可見之效果的業界首個公開驗證案例。

這不僅僅是 Lentosertib 單一的故事。截至 2026 年初，進入臨床階段的人工智慧驅動新藥候選藥物已達 173 款，與 2023 年底的 24 款相比，在不到三年的時間裡成長了七倍多。根據多項產業分析，人工智慧找出的物質在臨床一期的通過率達到 80% 至 90%，遠遠超過傳統方式的 40% 至 65%。臨床一期的目的不在於確認療效，而是確認安全性。由於人工智慧從設計階段就過濾掉預期有毒性的結構，並提前模擬了在體內的吸收與排出過程，因此似乎只有那些原本就容易跨過安全性門檻的候選藥物，才會進入臨床試驗階段。

不過，一旦進入臨床二期，這個差距就會縮小。直到目前為止，完成二期試驗的十款人工智慧發掘物質中，有四款獲得成功，記錄約為 40%，這與業界平均的 30% 至 40% 幾乎處於相同水準。這似乎是因為從臨床二期開始，比起藥物設計的精確度，因人而異的遺傳背景、生活習慣、伴隨疾病等在設計階段難以控制的變數所佔的比例變得更大了。截至 2026 年夏天，尚未有獲得監管機構最終批准的人工智慧設計新藥。

這似乎意味著，雖然山谷的入口明顯變寬了，但山谷盡頭的懸崖依然陡峭。

制度也在背後推動著這股潮流。美國食品藥物管理局在 2022 年 12 月通過了 FDA Modernization Act 2.0，廢除了數十年來規定新藥批准申請必須包含動物實驗的義務條款。2025 年 4 月，更提出了具體的路線圖，從單株抗體開始，逐步以類器官和基於人工智慧的模擬來取代動物實驗。隨後的 FDA Modernization Act 3.0 在 2025 年 12 月通過了參議院，並在 2026 年 7 月 20 日通過了眾議院。

由於法案需要參議院對眾議院修改後的版本再次表決才能提交給總統，因此截至 2026 年 8 月，尚未完成簽署。一旦該法案生效，新藥批准規定中隨處可見的「動物

實驗」字眼將會消失，取而代之的將是使用人類細胞製成、能在指甲大小的裝置上重現肺或肝等器官功能的「器官晶片」與電腦模擬等非動物試驗。

大型製藥公司也爭相加入這股潮流。開發出 AlphaFold 的 Google DeepMind 旗下子公司 Isomorphic Labs，分別與 Eli Lilly 和 Novartis 簽署了價值數十億美元的新藥共同開發協議，正在針對多種疾病標靶共同進行分子設計。Sanofi 則提出了將人工智慧結合到整體研發決策過程中的目標，並重新調整了本身的組織結構。

同一時期，包括 Novo Nordisk 和 AstraZeneca 在內的其他大型製藥公司，也紛紛設立自己的人工智慧部門或與專業企業攜手，進行了類似的重組。除了人工智慧新藥開發專業企業獨自設計新藥的模式之外，現有大型製藥公司將同樣的工具移植到自家研發組織內的趨勢，也正在同步擴大。

資本市場的反應也很明顯。Insilico Medicine 於 2025 年底在香港證券交易所上市，籌集了 22.77 億港元（約合 2.92 億美元），創下當年香港生技公司最大規模的首次公開募股紀錄。公開申購的競爭率超過了 1427 比 1，包括 Eli Lilly、Tencent 和 Temasek 在內的 15 家機構列名為基石投資者。在韓國國內，類似的動向也在加快。基於人工智慧的新藥開發企業 Galux，在 2026 年 2 月獲得了由 Yuanta Investment、Korea Development Bank、Korea Investment & Securities、Mirae Asset Securities 等機構參與的 420 億韓元規模的 B 輪投資。

Breeze Bio、First Biotherapeutics 和 Elysigen 等企業也陸續籌得超過 300 億元的投資，為企業首次公開發行奠定基礎。Oncobix 正在用自主開發的生成型人工智慧平台設計低分子標靶治療藥，Arontier 則與 Celltrion 攜手發掘以新型蛋白質結構為基礎的候選藥物。這些企業大多仍停留在臨床初期階段，但從投資規模和參與機構的組成來看，國內資本對這一領域的關注已與以往截然不同，這可視為一個信號。

業界將 2026 年和 2027 年視為真正的試金石。迄今為止是驗證速度和初期通過率的時期，而今後幾年將是領先管線在臨床三期這最後的關卡前接受考驗的時期。根據這份成績單，製藥公司今後 10 年的研發經費分配將發生重大變化。這對韓國新興企業來說也是一個即將到來的試金石。

Lentoprise 能否在下一階段臨床三期中再現相同的效果，以及該結果是否能導致監管部門的最終批准，將成為衡量整個人工智慧新藥開發信任度的標準。死亡之谷並

未消失，但穿過它的橋樑已剛剛完成了第一次車輛通行。

韓國政府也在跟進這一趨勢，完善支援體系。保健福祉部和科學技術信息通信部以韓國製藥生物協會為主管機構，從2024年4月至2028年12月投入348億元實施K-Melody項目。這一仿照歐洲Melody Project的事業，未將各企業分別積累的新藥候選物質數據匯集到一處，而是採用聯邦學習方式，讓數據保留在各機構，只移動人工智慧模型進行學習。

從製藥公司的角度來看，失敗的實驗記錄是敏感資訊，它們擔心會洩露給競爭對手。這種方式打開了在不外洩數據情況下共同學習的道路。主導項目的Kim Hwa-jong主任在一次採訪中表示，必須積累更多失敗案例，人工智慧才能提前篩選出下一次失敗。換句話說，失敗也是數據。

食品藥物安全處也在2024年6月發佈了藥物開發中人工智慧應用指南，為新藥審批過程中如何處理人工智慧生成的資料提供了方向。Daewong Pharmaceutical作為共同研究機構參與保健福祉部的K-AI新藥開發臨床前和臨床模型開發項目，決定在4年內將其在抗癌藥和代謝性疾病治療藥研究中積累的臨床前數據用於人工智慧模型訓練。

該公司基於包含8億種化合物分子模型的自有數據庫，另行開發並運營了用於尋找新藥候選物質的平台Daisy。不僅是新興創業企業，擁有悠久歷史的製藥公司也開始手握相同的工具。

儘管如此，人們仍可以問為什麼人工智慧設計的新藥中還沒有達到監管批准的案例。即使是最早進入臨床試驗的DSP-1181，也直到2020年才首次給予人類使用，Lentoprise也是到2022年才開始進行首次臨床試驗。如果我們按照從臨床一期到三期、再到監管審查所需的時間來計算，即使是目前領先的候選物質，最早也要到2027年，最遲也要到2028年才能出現結果。

人工智慧設計的新藥並非異常地頻繁失敗，而是充分的時間還未流逝至足以導致最終批准的程度。無論計算機運行多快，也無法縮短患者身體對藥物反應所需的數年時間。未來幾年內出現的臨床三期結果將實際驗證這一判斷是否正確。

1.2

生命科學的人工智慧：基礎模型 (Foundation Models)

2025年2月，在美國加州的非營利研究機構Arc Institute的電腦前，研究人員正在觀看大量基因序列不停地流入模型中，這些序列來自人類、水稻、大腸桿菌和古菌，不分種類。屏幕上閃過的不是熟悉的基因名稱，而是只由腺嘌呤、鳥嘌呤、胞嘧啶和胸腺嘧啶四個字母組成的序列，但其中混雜著從人類心臟病到微生物抗生素耐藥性的各種生命現象的規律。

他們正在製作的程序不是針對某一特定疾病、某一特定生物的計算器。直到幾年前，生物學研究室使用的人工智慧還停留在專注於單一基因、單一靶蛋白的狹隘工具上，因此不分種類地用一個模型給出答案的這個嘗試本身就是一種突破。無論是人類遺傳病、植物的乾旱抗性還是工業微生物的酶設計，無論問題是什麼，目標是讓同一個模型給出答案。

對於製藥的人來說，基因組一直以來只是被用作參考資料，沒有人想過可以像預測下一句話一樣完整地計算和處理它。幾週後向世界公開的這個模型的名字是Evo 2，其出現明確改變了人工智能在製藥和生命科學領域走過的方向。

只會做好泡菜湯的廚師和掌握切菜、火候控制、調味等基本功的廚師面對新訂單時的結果是不同的。前者做泡菜湯雖然出神入化，但收到義大利麵訂單就無計可施。掌握基本功的後者只要看到冰箱裡的食材就能立即做出新菜，無論是湯還是炒菜。即使來了陌生食材，後者也不會慌張，而是靈活運用基本功做出可口的味道。所謂基礎模型的人工智慧更接近後者。以前的人工智慧程序像前者一樣，如果被訓練去給特定蛋白質和特定化合物的結合力評分，就只能做那一個計算。

同樣的程序如果被問到某種化合物是否會對肝臟造成毒性，就無法給出答案，必須從頭開始訓練不同的程序。基礎模型先從大量數據中學習生命現象的基本文法，然後針對新提出的問題，無論是靶點預測、新藥候選物質設計還是臨床試驗結果估計，都能靈活運用實力。正是因為這種差異，製藥公司現在正在改變研究組織本身，

不再是每個部門、每個項目都編寫新程序，而是由多個部門共享一個具備基本功的基礎模型。

新藥開發現場若要讓這個原理發揮作用，需要跨越一道障礙。醫院的電子健康紀錄和基因組數據庫、製藥公司的化合物庫、學術期刊上散落的論文長期以來被困在不同的服務器、不同的格式、不同的所有權下。某大學醫院收集的肺癌患者基因組資料只在那家醫院內流通，某製藥公司通過實驗確認的化合物毒性數據不會流出公司外。

監管機構在新藥批准審查過程中積累的不良事件報告除了少數公開數據庫外，大多也只存放在機構內部的文件夾中。當數據這樣被分割時，再出色的算法也只能根據眼前狹隘倉庫中的信息進行學習。從2023年左右開始，製藥業界、學術界和大科技公司同時投入的工作就是把這些分散的儲存庫整合在一起。如果將疾病信息、患者基因組和臨床數據一起放入一個模型中進行訓練，模型就能掌握通用生物學推理能力，即使只看到一個肺癌患者的基因變異，也能同時回想起其他癌症類型、其他藥物和其他臨床試驗結果中類似的模式。

這種通用推理能力正是區分前一代只計算一種疾病、一個靶點的程序和基礎模型的決定性差異。

觀察生命現象的數據也有多個層次。基因組就像設計身體的完整藍圖，從出生時起幾乎不變，而轉錄體顯示在這個藍圖中此刻正在實際使用哪些基因。蛋白質體顯示按照這些展開的藍圖實際製造出來的部件，即蛋白質清單，而代謝體則顯示這些部件運作時產生的最終產物。就像看同一棟建築物，只看設計圖無法知道現在哪幾層的燈亮著一樣，僅從基因組信息也無法完全了解特定細胞或特定患者身上現在發生著什麼。

基礎模型之所以同時學習基因組、轉錄體、蛋白質體和代謝體，是因為必須一起觀看設計圖、照明狀態、部件清單和排放物才能正確診斷整棟建築物的狀態。如果再加上患者的診療記錄和影像資料，模型就能同時推理分子水平的變化與實際症狀如何關聯。

這次整合中最關鍵的洞察是發現可以像閱讀人類句子一樣的方式來閱讀蛋白質序列和分子結構。蛋白質是由二十種氨基酸像鏈一樣連接起來的結構，這與二十六個字

母排列成單詞和句子的方式在框架上並無不同。化學家長期以來只用圖畫表示的分子結構也可以轉換為用原子和鍵作為一個符號序列來表示。本來是為了翻譯人類句子和預測下一個單詞而設計的神經網絡結構 -- 轉換器，即使不輸入句子而是輸入氨基酸序列或分子符號序列，也能同樣運作。在重複訓練預測下一個氨基酸或原子的過程中，模型開始自動學習生命現象的文法，即使研究人員沒有特別教授規則，它也自動識別出了折疊方式或反應性等化學性質。

學會這種文法的模型，能夠全新創造出過去不存在的蛋白質序列，或是從頭設計出完全契合標靶蛋白質凹陷結合部位的分子。隨著原本處理文字的生成式人工智慧的力量，原封不動地轉移到生物學與化學的語言上，新藥開發已經從在現有候選藥物中挑選，跨越到了創造世上未曾有過的候選藥物的階段。

基礎模型的建立順序也有其邏輯。首先，它會經歷預先訓練階段，面對龐大的數據，在沒有特定目的的情況下，學習生命現象本身的規則。在這個階段，模型就像在小學和國中廣泛學習國語、數學與科學基礎的學生一樣，雖然還沒有決定特定的出路，但已經累積了無論遇到什麼問題都能著手處理的基礎。讓這樣打好基礎的模型，經歷針對特定任務重新訓練的微調階段後，才會具備專門用於設計新藥候選物質或預測臨床試驗結果等具體任務的實力。

TxGemma就是如實遵循這個順序的案例，它是建立在Google的通用語言模型Gemma 2透過預先訓練所打下的基礎之上，再加入與治療藥物相關的數據進行微調而成的。基礎打得越廣泛、越穩固，建立在上面的微調效果也就越大，這一點再次證明了數據整合為何是決定基礎模型成敗的關鍵。

如此建立起來的基礎模型結構，很像一棟三層樓的建築。在最底層，化合物資料庫、基因體、轉錄體（顯示細胞內基因實際被讀取多少的數據），以及數百萬篇臨床文獻，就像基礎建材一樣堆疊著。在修整這些建材以建立骨架的中間層，則放置了理解句子的的大型語言模型，以及專門讀取蛋白質序列與分子結構的基礎模型。在中間層修整好的骨架會往上送到最高層，完成生成新藥候選物質、為標靶排定優先順序、最佳化臨床試驗設計等實際成果。

舉例來說，如果中間層的模型讀取了堆疊在底層的肺癌患者基因體與相關論文，挑選出一個在肺癌中經常出現異常的蛋白質，最高層就會立刻描繪出幾種能附著在該

蛋白質上的新藥候選物質。如果底層建材不佳，中間層再怎麼精巧也蓋不出東西；如果中間層垮了，底層就算堆滿了再好的建材，力量也無法傳遞到最高層。唯有三層同時堅固，才能產出在新藥開發實務中有用的成果，這項事實正由2025年與2026年間大量出現的基礎模型一一證明。

清楚展現連接中間層與最高層力量的案例，是Google推出的治療藥物專用大型語言模型TxLLM。TxLLM是將分子結構、標靶蛋白質、細胞株、臨床試驗文字等不同格式的數據，一起放入同一個模型中進行訓練的成果。儘管分子以原子符號表示，蛋白質以胺基酸序列表示，臨床試驗紀錄則以人類使用的句子表示，各自的表達方式都不同，但TxLLM將這些全部融入到同一個共通的表達空間中，使其能夠相互比較與連結。研究團隊使用了從發掘標靶到預測臨床試驗結果等涵蓋新藥開發整個過程的66項評估任務，來測試TxLLM的實力。

評估項目中，混合了從估算特定化合物能存活到臨床試驗第三期的機率的問題，到找出同時投予兩種藥物時可能出現的交互作用的問題。TxLLM在其中22項任務裡，創下了比當時出現過的任何專業模型都要好的成績。一個模型能同時做好分子設計與臨床預測這兩件相距甚遠的工作，這項事實被視為基礎模型能夠走出論文，在實際開發實務中獲得應用的證據。

新藥開發團隊實際面臨的困難，不在於標靶候選數量不足，而是數量過多。在人類基因體中，光是製造蛋白質的基因就超過2萬個，其中有潛力成為新藥標靶的候選者，只要搜尋論文、專利與臨床紀錄，就會以數千個的數量湧現。由於研究經費與人力有限，製藥公司只能從中挑選幾十個，最多幾百個來優先深入研究。過去，這種篩選工作依賴少數幾位經驗豐富的科學家的直覺與論文搜尋來進行，結果導致研究總是集中在已經流行標靶的傾斜現象一再發生。

像TxLLM或TxGemma這類的模型，會一次瀏覽標靶、疾病與現有藥物的成敗紀錄，為特定標靶能成功進入臨床試驗的機率打分數，因此也能將人類容易漏掉的候選者連同根據一起往前拉升。與過去排定標靶優先順序的工作需要花費幾週到幾個月不同，使用基礎模型的現在，幾天內就能產出排行榜草案。

TxLLM所點燃的可能性，在2025年3月Google公開TxGemma時又向前邁進了一步。TxGemma是以Google的輕量級語言模型Gemma 2為骨架，利用超過700萬筆與治

療藥物相關的訓練案例重新訓練而成的模型，分為20億、90億與270億個參數規模的三種大小推出。參數就像是模型在學習過程中會自行調整數值的內部旋鈕，其數量越多，通常就越能做出精巧的判斷。除了預測專用的模型之外，還一併推出了以對話形式運作的TxGemma-Chat版本，讓研究人員可以像聊天一樣提出問題，模型則會提出根據來回答的運作方式也變得可能。

最大的270億參數版本在與前一個TxLLM相同的66個任務中有64個達到相當或更好的成績，其中45個刷新了既有紀錄。TxGemma與TxLLM的決定性差異在於公開性。與TxLLM作為Google內部研究成果而未曾公開不同，TxGemma將模型權重通過Hugging Face和Google Cloud的Vertex AI完全公開，為大學研究室乃至資金不夠充裕的生物創業公司打開了取得同等水準基礎模型的大門。

試圖以一種語言讀取整個基因體的努力在Evo 2中達到了頂峰。由Arc Institute與NVIDIA、Stanford University、加州大學舊金山分校和加州大學柏克萊分校共同開發的這個模型擁有400億個參數，能夠一次最多讀取100萬個鹼基（組成基因的化學字母單位）長度的序列。用於學習的數據來自超過10萬種生物（人類、植物、細菌和古菌），共計超過9兆個鹼基。Evo 2於2025年2月先以預印本形式公開，隨後經過驗證後正式發表在國際學術期刊Nature上。

該模型展示了它能夠完成的任務範圍，從預先識別導致人類疾病的遺傳變異，到設計新的微生物酶，甚至到從無到有地創造原本不存在的遺傳序列。Arc Institute解釋說，該模型不僅可用於人類遺傳病研究，也可用於農業育種、工業微生物開發和新材料探索。Evo 2之所以被視為象徵性案例，在於一個模型同時處理診斷和設計兩個不同的任務，展示了遺傳體數據整合如何為新藥開發賦予力量。

Google DeepMind也在2025年6月25日推出了名為AlphaGenome的類似方向的模型。AlphaGenome接收100萬個鹼基長度的DNA序列作為輸入，一次預測數千種調控信號，包括基因表達、染色質可及性（表示DNA有多容易被打開讀取的指標）、組蛋白修飾和連接位點等。判斷人類基因體中不直接製造蛋白質的區域（即所謂的非編碼區）中發生的突變是否實際導致疾病，這一直是困擾遺傳學家們的難題，AlphaGenome在24個評估項目中的22個以及26個變異預測項目中的24個上超越了現有模型的成績。

但是該模型並未設計為輸入個人完整基因體來診斷疾病，同時在讀取相距超過10萬個鹼基的調控序列對基因的影響方面也存在準確度不足的局限。Google DeepMind在2026年1月將該模型的源代碼和權重為非營利研究目的免費開放，因此全球遺傳學實驗室得以使用同一工具比較不同疾病的調控變異，建立了共同基礎。

隨著分散在各個醫院的患者基因體測序結果在這個共同模型上匯聚，像罕見疾病診斷這樣的單靠單一醫院數據無法解答的問題也開始有了突破口。

在蛋白質領域，美國初創公司EvolutionaryScale開發的ESM3實現了類似的飛躍。如果說迄今廣泛使用的AlphaFold系列模型專注於根據給定的胺基酸序列預測三維結構，那麼ESM3則更進一步，從預測跨越到設計 - - 即使從結構或功能的角度反過來給定條件，它也能生成符合這些條件的新序列。ESM3同時學習了胺基酸序列、三維結構以及蛋白質的實際功能，具備了只要給定其中任何一個作為條件，就能自動補全其他兩個的能力。

該模型的能力在2025年1月作為正式論文發表在國際學術期刊Science上時得到了驗證。研究人員只告訴ESM3螢光蛋白質應該具備的條件，然後讓它設計其餘的序列，結果生成的蛋白質在序列上與自然界中的任何螢光蛋白質都明顯不同，卻能發出明亮的光。研究人員評價說，該模型通過計算達到了自然界在5億年進化過程中可能達到的境界。ESM3的權重和代碼在學術研究目的下通過GitHub公開，並為一般用戶提供了免費使用的應用編程接口（API），使得只讀過論文的研究生也能夠直接嘗試設計新蛋白質。

但是，商業規模應用需要單獨的合同，因此應該將進入門檻降低到什麼程度的問題仍在討論中。

試圖完整地模擬單個細胞反應的努力在2025年夏季由Arc Institute舉辦的Virtual Cell Challenge中留下了明顯的里程碑。這場競賽旨在確定最準確預測細胞在經歷基因修飾或藥物刺激時反應的模型，來自114個國家的超過5000名研究人員報名參加，超過1200支隊伍提交了結果。這場獎金為10萬美元的競賽基於Arc Institute公開的名為Virtual Cell Atlas的數據存儲庫進行，該存儲庫包含了來自多個實驗室的超過6億個細胞的觀察和刺激反應數據。

這份資料中，不僅包含了生技公司 Tahoe Therapeutics 提供的大規模藥物反應資料集，還有 Arc Institute 讓人工智慧代理程式搜尋公開論文，自動收集而來的單細胞資料。比賽結束後，Arc Institute 推出了一個名為 STATE 的模型，該模型無需實驗，就能預先計算出在給予特定細胞株特定刺激時，基因表現會如何改變。這場比賽的成績單證實了一項事實：不侷限於單一疾病或單一患者的資料，而是將多個實驗室、多種細胞株與多種刺激條件的資料匯集一處時，細胞層級的通用預測能力才得以誕生。

Arc Institute 預告將在 2026 年再次舉辦這場比賽。這意味著，當散落各處的細胞實驗資料匯集一處時，那些單一實驗室無法回答的問題，便開始出現了解答的線索。

到了 2026 年，這股趨勢的規模已超越了個別模型之間的競爭，擴大到重塑整個產業基礎設施的程度。OpenAI 在 2026 年 4 月 16 日推出了生命科學研究專用模型 GPT-Rosalind。這個以對揭示 DNA 雙螺旋結構留下決定性線索的科學家 Rosalind Franklin 命名之模型，在同年 6 月再次調整了效能，與前一版本相比，在減少 31% 的權杖（處理句子的最小單位）消耗量的同時，還提升了醫藥化學與基因體分析的實力。不過，這個模型並沒有以任何人都能立即使用的形式釋出，而是先將存取權限賦予了 Amgen、Moderna、Thermo Fisher Scientific 和 Allen Institute 等經過驗證的機構。

名列這份名單的 Amgen 與 Moderna 皆是已經自行累積了龐大臨床資料的大型製藥公司，這一點顯示出 OpenAI 的模型並未排擠現有資料，而是以疊加於其上的方式生根發芽的可能性。同時，OpenAI 為了防止生命科學技術被用於錯誤的目的，也啟動了一項名為 Rosalind Biodefense 的獨立計畫。連生命科學專用模型都掛上公司名稱投入研發，這項事實本身就顯示出這類基礎模型已不再僅限於學術界與少數生技公司的領域。

Amazon Web Services (AWS) 也在同年的 4 月 15 日於生命科學研討會上展示了 Amazon Bio Discovery。這項服務具備了人工智慧助理的形式，將 40 多種生物學基礎模型匯集於同一個畫面中，當研究人員以自然語言提出請求時，便會挑選合適的模型，甚至代為安排實驗設計。Amazon Web Services (AWS) 表示，負責合成與驗證的合作實驗室名單中，Twist Bioscience 與 Ginkgo Bioworks 已經名列其中，A-Alpha Bio 也即將加入。這項服務還具備一項功能，只要研究人員以自然語言寫

入實驗步驟，就能將其轉換為由多個模型與分析階段組成的工作流程，因此即使是不懂得編寫程式碼的生物學家，也能親手設計複雜的管道。

Memorial Sloan Kettering Cancer Center 利用這項服務設計了 30 萬種抗體候選分子，接著將其中的 10 萬種送到 Twist Bioscience 完成了實際的合成與驗證，讓原本需要花費數個月的工作縮短到了幾週。NVIDIA 早在 2025 年 4 月就已將處理小分子的生成式模型 MolMIM 加入了 BioNeMo，並在 2026 年 1 月於同一平台上全新添加了 RNA 結構預測模型與合成可能性驗證模型。同月 12 日，NVIDIA 並排發表了兩項合作案：一項是與 Eli Lilly 建立共同創新實驗室，共同解決新藥研發的難題；另一項則是與 Thermo Fisher Scientific 合作，將人工智慧結合到實驗設備中，藉此投入了將基礎模型延伸至電腦之外的實驗室現場的任務。

這種將預測與驗證直接串聯起來的循環結構，顯示出基礎模型已超越了論文中的效能指標，真正在實際的新藥管道中生根發芽。

要從頭開始訓練這類模型，需要動用數千台圖形處理器連續運作數週的計算量，光是其電費與設備費用，就遠遠超過了一般生技新創公司一年的預算。結果，從底層開始建立新模型的工作，明顯傾向於集中在 NVIDIA、Google、Amazon、OpenAI 等具備自家資料中心的少數企業，以及少數獲得政府支援的研究所身上。

規模較小的生技公司或大學實驗室，越來越常選擇下載這些企業所公開的模型權重，並使用自身特殊的資料來進行微調。TxGemma、ESM3 與 AlphaGenome 公開權重，也可視為是為了稍微縮小這種差距的嘗試。

目前為止出現的基礎模型，各自擅長的領域皆有所不同。Evo 2 與 AlphaGenome 善於解讀基因體序列，ESM3 善於處理蛋白質結構，TxGemma 則善於結合文字與臨床資料。將一名患者的基因體、細胞反應、診療紀錄與影像資料全數解讀，並整合為單一模型的工作，目前仍在進行中。像 Amazon Bio Discovery 這樣將多個模型匯集於同一個畫面，讓人能在需要時挑選使用的做法，可以看作是達到完全整合的過渡階段。

不同的公司、不同的研究所推出的模型，目前尚未建立起將它們串聯在一起的標準，這是該領域面臨的顯著課題。儘管如此，看到短短兩三年間，原本只解讀單一基因體的模型，已朝著涵蓋細胞反應與臨床紀錄的方向快速擴展的趨勢，便可推測出

下一個整合階段已經離我們不遠了。

基礎模型的價值，比起演算法程式碼本身，更取決於該模型所吸收的資料廣度，以及透過實驗來確認這些資料的速度。即使套用相同的 Transformer 架構，吸收了 10 萬種生物基因組的模型，與只吸收了一家公司化合物資料的模型，會展現出完全不同的實力。因此，現在的問題已經不是哪一種演算法比較聰明，而是哪一個組織站在更廣泛、更深厚的資料基礎之上。

廣泛掌握資料，並且具備將這些資料透過實驗回饋的循環結構的組織，在下一階段的新藥開發競爭中，佔據優勢的可能性看來很大。掌握了能夠跨越疾病、基因組和臨床資料的推論基礎，而不僅僅是猜中單一標靶的計算機的組織，將會把生成新藥候選物質、判斷標靶優先順序以及設計臨床試驗這三項實務工作掌握在手中。如果在 2022 年，Insilico Medicine 在 30 個月內將一個候選物質推進到臨床 1 期試驗被視為一項破例的成果，那麼現在，在多種疾病上同時重現那樣的速度，則是擺在具備基礎模型的組織面前的下一個課題。

就像只會煮一種 Kimchijjigae 的廚師，要變成具備各項基本功的廚師需要長時間的修練一樣，生命科學基礎模型要轉化為病床邊的實際處方，也仍然需要收集資料並透過實驗進行驗證的修練時間。

Gwagi Jeongtongbu 在 2025 年 10 月，選出了由醫療人工智慧企業 Lunit 所帶領的聯盟，作為開發處理醫學科學領域基礎模型的主管機構。這裡有 Kakao Healthcare 和 SK Biopharm 等企業、KAIST 和 Seoul Daehakgyo 研究團隊，加上 9 家醫院，總共 23 個機構參與，並且計畫將從分子和蛋白質到臨床論文等不同層次的醫學科學資料整合在一起，在 2026 年 9 月前完成模型。Bogeon Bokjibu 也將 Hanguk Jeyak Baio Hyeophoe 設為總管機構，與 31 個機構一起，正在建立連結臨床前階段到臨床階段的 K-AI 新藥開發模型。

由 Seoul Daehakgyo Byeongwon、Samsung Seoul Byeongwon 以及 Hanguk Saengmyeong Gonghak Yeonguwon 分別帶領的這項工作，採用了一種聯合學習的方式：不將散落於各家醫院不同地方的患者資料轉移到同一個伺服器，而是只把在各家醫院內部學習的結果收集起來，合併成一個模型。這算是一場韓國式的實驗，試圖用不轉移資料只收集知識的方法，來解決資料被困在不同伺服器裡的古老問

題。

也出現了基礎模型提出的預測在實驗台上立刻被推翻的案例。瑞士巴塞爾大學藥學部的 Markus Lill 教授研究團隊，對包含 AlphaFold 3 在內的多個結構預測模型進行了惡作劇般的試驗。他們將藥物附著的結合部位的胺基酸整個替換掉，或是反轉化合物的電荷，讓藥物在實際上無法附著在蛋白質上，然後讓模型再次解答相同的問題。

從 2025 年 10 月刊登在國際學術期刊《Nature Communications》上的結果來看，在 285 個複合體中，有超過 40% 的情況，模型給出了和以前一模一樣的答案，就好像結合部位沒有改變一樣。研究團隊針對這個結果解釋說，這並不是因為模型理解了為什麼藥物會附著在蛋白質上，而只是把在學習資料中看過的圖像原封不動地重複出來而已。

學習資料和計算資源集中在少數企業的趨勢，看來是一個無法僅靠公開權重就能完全解決的問題。因為下載權重來進行微調的一方，與一開始就收集龐大數量的鹼基序列和臨床紀錄，從頭開始訓練模型的一方，其起跑線本身就有所不同。由於很少有企業會把訓練模型時所使用的原始資料的組成和來源也一起公開，因此外部的研究人員也很難驗證該模型實際上學到了什麼。如果沒有像韓國的聯合學習方式那樣，由政府籌集預算來整合醫院和大學的資料，若是公共部門不挺身而出，這種傾斜的現象在未來似乎會變得更加嚴重。

1.3

多重組學整合與新型生物標記物發現

在日本多家醫院進行的臨床二期TRIUMPH試驗，是以人類表皮生長因子受體2型，也就是HER2基因大幅增加的轉移性大腸癌患者為對象。這些患者先透過一項名為GOZILA的循環腫瘤DNA篩檢研究進行了篩選，這是一種不需再次切取組織，而是透過一瓶血液來找出標靶基因是否擴增的液體活檢方式。這等於是為難以再次切取腫瘤組織的患者，開啟了標靶治療的大門。

研究團隊將接受事前篩檢的143名患者，以及實際參與試驗的30名患者的組織玻片，輸入到人工智慧病理判讀程式中。這個程式是韓國企業Lunit所製作名為SCOPE HER2與SCOPE IO的軟體，它能在一張玻片影像中，同時計算出癌細胞被HER2染劑染得有多深，以及腫瘤周圍密集聚集了多少免疫細胞與纖維細胞。與病理醫師的判定相比，這個程式的準確度為86.7%，而在挑選出染色強度最深的第3級患者方面，則達到了100%的吻合。

超過半數癌細胞被染成該第3級的患者群，與未能達到的患者群相比，無惡化存活期為4.4個月對上1.4個月，而整體存活期則拉開到了16.5個月對上4.1個月。這項探索性分析結果，已刊登在2025年的國際學術期刊《臨床腫瘤學精準醫學》（JCO Precision Oncology）上。

若要正確解讀這項結果，必須先從「標靶」與「生物標記」這兩個詞彙開始解釋。標靶是指成為引發疾病原因的蛋白質或基因。您可以將其想像成鎖頭。而藥物就是為了完全契合該鎖頭所打造的鑰匙，鑰匙必須插入鎖孔轉動，才能關上引發疾病的大門。與此不同，生物標記則是一種告訴我們身體裡面現在正在發生什麼事的訊號。

就像平交道的號誌燈或汽車儀表板的警告燈一樣，只要測量一滴血或一個組織樣本中所含特定物質的量，就能預先推測出這個身體處於什麼狀態、哪種藥物會有效，以及疾病未來會如何發展。在TRIUMPH試驗中，人工智慧判讀出的HER2染色強度與免疫細胞密度，正是扮演了該號誌燈角色的生物標記。

人類病理醫師看著同一張玻片卻做出不同判定的情況並不罕見。用目測來區分HER2染色強度的2+與3+之間的界線相當模糊，在這種模糊的案例中，往往必須另外重新接受螢光原位雜交法檢查。人工智慧判讀程式將人類眼睛容易錯漏的這個界線，轉換為染色強度與細胞比例的連續數字，大幅減少了因判定者不同而產生的結果偏差。當然，這個分數並不代表能取代病理醫師。它比較像是病理醫師做出最終判斷前，先幫忙篩選出容易錯漏的微細訊號的輔助工具。

癌症或失智症這類退化性腦部疾病，很少是因為單一基因故障而引起的。數百、數千個基因與蛋白質相互交織，只有在特定的組合下，疾病才會顯現於外。只靠一次組織檢查或一次基因檢查，是無法解開這團錯綜複雜的線球的。因此，研究人員開始將從不同層面觀察身體的各種技術匯聚在一起。基因體學、轉錄體學與蛋白質體學便是如此。基因體學是利用次世代定序設備，將細胞內的DNA序列整個讀取出來。這可以看作是蓋建築物時所使用的設計圖。設計圖一旦畫好，就不太會改變。

轉錄體學則相當於看著該設計圖，細胞每天所發出的工作指示書。它利用名為RNA測序的技術，測量目前有哪些基因被開啟了多少。即使是擁有相同設計圖的細胞，肝細胞與神經細胞的工作方式之所以不同，原因就在於此。蛋白質體學則是更進一步，實際用質譜儀一一確認在現場搬磚頭、釘釘子的工人，也就是蛋白質本身。就算是相同的蛋白質，也會根據是否附加上磷酸基等微小標籤而有完全不同的活動方式，像這樣連轉譯後修飾都進行觀察的精準蛋白質體分析，能更緊密地告訴我們藥物在訊息傳遞路徑中實際能介入的位置。

由於藥物通常不是直接附著在基因上，而是附著在蛋白質上來發揮作用，因此確切知道這些工人有幾名、彼此以什麼樣的形態牽著手，在新藥開發中至關重要。問題在於，這三種數據的測量方式、單位與規模都各不相同，人類要將它們並列比較是非常困難的。

若在此加上代謝體學與表觀遺傳體學，圖像就會變得更加細緻。代謝體學是測量細胞實際吃了什麼、排出了什麼，像是葡萄糖或胺基酸等小分子濃度的作業；而表觀遺傳體學則是在保持DNA序列本身不變的情況下，觀察為特定基因貼上化學標籤，扮演著開啟或關閉該基因之開關角色的變化。

雖然已知人類基因大約有兩萬個左右，但這些基因經過轉錄體、蛋白質體、代謝體階段後所製造出的組合數量，要比這多出好幾個位數。人類要親手比對這龐大的組合，從一開始就幾乎是不可能的，而這正是將多體學數據整個吞下並找出模式的人工智慧發揮龐大作用之處。

就在這個環節，人工智慧介入了。像是變分自編碼器或Transformer這類深度學習模型，會將基因體、轉錄體、蛋白質體數據一口氣全丟進去，找出人類肉眼看不見的潛在規則。它不是將這三種數據勉強平均或並列排開，而是在模型內部建立新的座標空間，將每一位患者或每一個細胞移轉為該空間中的一個點。觀察這些被移轉的點的距離與配置，就能揭露出傳統統計手法無法捕捉到的交織關係，例如一個基因與一個蛋白質只有在特定條件下才會一起引發問題的關係。

市場調查機構預測，多重組學分析市場從2025年的28億1千萬美元增長到2026年的約32億美元，並且預計到2035年為止每年將增長超過15%。

單細胞轉錄體分析(scRNA-seq)的出現是尋找靶標工作進一步精細化的契機。過去，人們會將整塊組織磨碎，並籠統地測量其中包含的所有細胞的基因表達。這就像拿一張混雜了多個人的照片，然後畫出平均臉孔一樣。實際上，隱藏在其中的少數癌幹細胞或已經疲憊的免疫細胞都被掩埋在統計數據中而消失了。

單細胞分析則不同。它為每一個細胞貼上標籤，像逐一進行訪談一樣分別測量基因表達。在腫瘤組織這樣的正常細胞、免疫細胞和突變細胞混雜在一起的地方，除了這種方法，沒有辦法了解每個細胞的不同情況。

研究人員通過這樣逐一讀取數萬個細胞，追溯從癌症前期的細胞開始，經過細分的各種癌細胞亞群，直到力竭無法發揮其作用的免疫細胞的時間流程。在2025年至2026年期間，Illumina通過其Billion Cell Atlas項目，一年內就積累了達20拍字節的單細胞轉錄體數據，並且最終目標是建立包含50億個細胞的數據庫。

學習了這樣積累的數據的單細胞基礎模型可以做到給特定細胞命名說明其種類，甚至在給予特定刺激時提前估計細胞將如何反應。

單細胞數據訓練的人工智能也開始與文獻基礎語言模型協作。2026年公開的ELISA框架將理解單細胞基因表達的scGPT等模型與基於BioBERT的文獻檢索功能整合為

一體。當研究人員發現某個特定細胞類型中特別容易啟動的基因時，它會立即同時展示相關的論文證據。通過這樣在一個屏幕上實驗數據和文獻知識相互契合，就可以直觀確認新的生物標誌物候選物，並立即驗證其有效性。

但是，單單知道每個細胞的身份並不意味著故事的結束。因為細胞根據其所處位置的不同而執行不同的功能。空間轉錄體學(Spatial transcriptomics)不粉碎組織，而是將其保持原樣，並繪製每個位置啟動了哪些基因以及啟動程度如何的地圖。這就像一張不僅列出居民名單，而是標明幾棟幾號住著誰的地圖。10x Genomics的Visium或同時貼附多個熒光標記的多路熒光原位雜交(MERFISH)等技術以細胞單位的精確度繪製了這張地圖。

這個領域在2025年至2026年期間迅速改變。10x Genomics在2025年8月推出了Xenium Protein，為在同一細胞、同一組織切片中同時測量RNA和蛋白質開闢了道路。2026年4月，它發布了一款名為Atera的新設備。相比現有的Xenium，這款新設備的面積大四倍，使用兩倍數量的載玻片，能以細胞單位測量整個基因組，同時利用圖形處理器將分析時間從小時級別縮短到分鐘級別。

用該設備獲得的數據中發現了新的位置相關生物標誌物候選物，例如在低氧腫瘤組織中聚集的ODAM和在侵襲性強的分裂細胞中聚集的C21orf58。

人工智能將單細胞數據和空間數據疊加在一起，計算腫瘤細胞與周邊結締組織細胞、免疫細胞之間的實際物理距離和相互傳遞的信號。在2025年發表的一項研究中，聯合學習空間轉錄體和空間蛋白質體數據的圖神經網絡預測免疫檢查點抑制劑反應的精確度(以ROC-AUC指標表示)超過了0.90，並發現細胞外基質厚重包圍的腫瘤周圍區域在對藥物無效的患者中特別突出。

Triumph試驗中Lunit程序讀出的免疫細胞密度方式也不例外。Lunit在2025年12月針對291名膽道癌患者和309張切片進行的研究中，HER2判讀結果與病理醫生共識的符合率為83.5%，相關論文已在國際學術期刊上發表。在2026年美國臨床腫瘤學會學術大會上發表了五項研究，在美國癌症研究學會學術大會上則發表了六項研究，包括與Ajou University及Agilent合作進行的非小細胞肺癌研究。

腫瘤微環境中的巨噬細胞也是人工智能關注的對象。巨噬細胞的性質並非固定，既可以朝著幫助腫瘤的方向改變，也可以朝著與腫瘤對抗的方向改變。一項空間轉錄

體研究揭示了CXCL16和SPP1這兩種信號物質調節巨噬細胞的這種方向轉變，並報告稱已經用於其他用途的Sirolimus和Trichostatin A這兩種藥物可能成為靶向這條信號途徑的候選物。

採用這種藥物重新定位的方法 -- 即不從頭設計新化合物，而是重新使用已確認安全性的藥物治療其他疾病 -- 大大縮短了開發時間。

迄今為止檢視的數據都是實驗室設備從細胞和組織中直接提取的數字。但是，關於人類疾病的知識中相當一部分不是以數字形式，而是以論文、專利和臨床試驗報告中的文字形式存在。每年產生的數百萬篇這樣的文獻對於人類來說全部閱讀和整理是不可能的。因此應運而生了BioBERT和PubMedBERT這樣的語言模型，它們是用大量生物醫學論文預先訓練的。

這類模型只要讀取一個句子，就能自動辨識出其中出現了哪些基因、哪些疾病、哪些藥物名稱，甚至還能萃取出它們彼此之間的關係。一個名為 reguloGPT 的模型更進一步，能從單個句子中，將基因是抑制還是促進其他基因表現等調控關係，轉換成整理好的表格。

像這樣將從文獻中萃取出的關係，與實驗室獲得的基因體、轉錄體、蛋白質體數據匯集在一起，就會形成巨大的知識圖譜。基因與蛋白質、疾病與藥物各自成為節點，而它們之間的關係則成為連接節點的線，這是一張如蜘蛛網般密集的網狀地圖。研究人員在這張地圖上，運用圖神經網路或知識圖譜嵌入等技術，在如網般交織的連結網中，找出特別多線條匯聚的基因是什麼。就連尚未在論文中報導過的藥物與標靶之間的關係，也能在這張圖譜上預先推敲出來。

研究人員將這種圖譜分析稱為疾病網路分析。就像不直接結交朋友，也能透過朋友的朋友介紹一樣，圖神經網路即使發現基因 A 與基因 C 從未在同一篇論文中同時出現過，也能僅憑兩者都經常與基因 B 有所關聯這項事實，找出兩者之間隱藏的關聯。2025年發表的知識引導型圖譜學習技術 KGDRP，在藥物反應預測與標靶預測兩項任務中，即便是針對幾乎沒有既有已知數據的新疾病，也將預測效能提升了12%。

這意味著，即使在文獻或臨床數據依然貧乏的罕見疾病中，人工智慧也能從頭開始提出標靶候選者，並有望大幅減少新藥開發初期階段的反覆試錯。

Insilico Medicine 的 PandaOmics 是及早將此方式商業化的平台。它將基因表現與蛋白質數據、基於文獻的知識圖譜結合在一起，在肝細胞癌與特發性肺纖維化等多種疾病中，為新標靶排定優先順序，並已透過實際實驗進行驗證。2025年，他們精進了綜合考量可靠度、商業潛力、成藥性以及作用機轉明確程度的評分系統，全新推出了 TargetBench 1.0，以及涵蓋腫瘤學、代謝疾病、免疫疾病、纖維化疾病、神經疾病等38種疾病的架構 TargetPro。

同年，Insilico Medicine 為了尋找胃癌與食道癌標靶，與世界級的癌症中心攜手合作；2026年，針對 PandaOmics 挑選出的腦部疾病標靶，他們與中國醫療系統集團簽訂了規模達1億6,500萬美元的第二次合作協議。

這個平台明顯展現成果的案例是 rentosertib（開發代號 ISM001-055）。

PandaOmics 在瀏覽關於纖維化與老化的文獻及體學數據時，找出了名為 TNIK 的蛋白質，而這個蛋白質在此之前並非特發性肺纖維化治療藥物標靶所關注的候選者。在中國22家醫院招募71名患者，進行為期12週、比較安慰劑與三種劑量的臨床2a期 GENESIS-IPF 試驗中，每天接受60毫克投藥的患者組別，其用力肺活量（FVC）平均增加了98.4毫升，而安慰劑組別則減少了20.3毫升。

2025年6月3日刊登於 Nature Medicine 的這項結果，被視為業界首次證明由人工智慧從頭到尾參與並找出標靶與藥物，實際上能在人體內產生療效的案例，隨後也推進到了第3期試驗。

尋找標靶的方法還包含合成致死率（Synthetic lethality）分析。想像飛機上裝有兩具引擎，就很容易理解了。即使一具引擎停止運轉，仍可靠另一具繼續飛行，但如果兩具都停止，飛機就會墜毀。細胞內也有許多這種成對的安全裝置，因此即使一個基因損壞，另一個基因也會接下該工作，讓細胞存活下來。

然而，癌細胞通常是在已有一個基因損壞的情況下生長出來的，因此如果用藥物將其剩餘配對的基因也阻斷，正常細胞會安然無恙，但癌細胞就會死亡。人工智慧透過共同學習泛癌多體學數據以及將基因逐一關閉的模擬，大規模地找出這樣的配對。

尋找這種配對的人工智慧模型，通常會使用被稱為基因依賴性圖譜的大規模公開資料庫作為學習材料。美國 Broad Institute 建立的 Cancer Dependency Map 即是代表性例子，這是一份在數百種癌細胞株中，將基因逐一關閉，藉由實驗記錄下當哪

個基因消失時該細胞株就會死亡的資料。在此疊加上各細胞株的基因體、轉錄體與蛋白質體資訊，就能在統計上挑選出唯獨在具有特定基因突變的細胞株中特別容易死亡的配對。

在 BRCA 基因損壞的癌細胞中，只要阻斷名為 PARP 的酵素，就能專挑癌細胞將其殺死，這項事實首次證明了這種方法的潛力。其後雖然推進至臨床的配對並不多，但在2025年與2026年之間，新的候選者接連出現。研究確認了在失去 MTAP 基因的癌細胞中，只要阻斷名為 PRMT5 的酵素，癌細胞就會死亡的關係，於是 IDEAYA Biosciences 在2025年9月提交了名為 IDE892 的 PRMT5 抑制劑的臨床試驗計畫，並於同年第4季針對 MTAP 缺失的肺癌患者開始了第1期試驗。

Bristol Myers Squibb的BMS-986504進入了第二期臨床試驗，招募了60名MTAP缺陷型胰腺癌患者。在胰管腺癌研究中，研究人員結合使用單細胞空間轉錄體學、蛋白質體學、競爭性內源RNA網絡分析，以及基於轉換器的深度學習模型SEFormer，建立了一套分析管道。這套管道發現了由TNFRSF10A基因編碼的死亡受體TRAIL-R1這個新靶點，並透過計算機模擬確認了已被批准用於腎癌治療的替西羅利姆斯可以結合到這個受體上。

已發現的生物標記物在實際臨床應用中的表現，在黑色素瘤治療藥物 Pembrolizumab 的大規模臨床試驗KEYNOTE-002和KEYNOTE-006中得到了驗證。研究人員從575名患者的CT影像中提取了表示腫瘤大小和形狀變化的25個特徵，然後使用隨機森林演算法篩選出了其中與大小相關的2個、與形狀變化相關的2個，共4個影像生物標記物。

這四個組合在治療開始後6個月時點能夠相當準確地預測患者的總體生存情況。無需進行血液檢查或組織取樣，僅使用已有的CT影像就可以提取生物標記物，這說明該技術多麼容易融入醫院實踐。

疾病網絡或多元組學模型在電腦屏幕上篩選出的靶點候選物本身並不是藥物。研究人員必須透過使用細胞系、類器官或實驗動物進行體內實驗，確認抑制該靶點後疾病是否真正得到改善。只有通過這一階段的候選物才能進入臨床試驗設計。對於已在進行中的大規模臨床試驗(如TRIUMPH試驗或KEYNOTE試驗)，探索性分析方法可以利用人工智能重新檢查其組織切片和影像資料，這樣無需從頭設計新的臨床試驗，就可

以快速驗證生物標記物候選,大大節省時間和費用。

經過驗證的生物標記物必須獲得伴隨診斷認證後,才能最終在醫院診療室中用於實際患者。

經過驗證的生物標記物本身也正在改變臨床試驗的設計方式。過去,研究人員隨機分組患者,只有在試驗結束後才能確定哪種治療效果更好。現在,適應性臨床試驗設計越來越普遍,可以在試驗進行期間實時監測生物標記物水平,對反應不佳的患者群體進行轉換治療,對反應良好的患者群體繼續原有治療。

像TRIUMPH試驗一樣,事先根據循環腫瘤DNA篩選患者,然後才開始本次試驗的方式,也是這一趨勢的一部分。透過減少患者的耽擱,同時縮少所需患者人數,更快速地得出結論,生物標記物的發現似乎擁有改變新藥開發速度本身的力量。

為了使這樣發現的生物標記物真正轉化為臨床處方,必須通過稱為伴隨診斷(CDx)的關卡。這是一種檢測工具,用於將特定藥物的使用與具有特定生物標記物的患者配對並批准。美國食品和藥物管理局批准的伴隨診斷檢測從2017年的34項、2021年的44項增加到2025年3月的188項,相關生物標記物增加到45種。在2026年5月,Natera公司的Signatera伴隨診斷(用循環腫瘤DNA測定膀胱癌患者的復發風險)和Guardant Health公司的Guardant360 Liquid伴隨診斷(涵蓋乳腺癌、結腸癌和非小細胞肺癌)先後獲得了批准。

除此之外,美國食品和藥物管理局還在考慮將基於次世代測序和聚合酶鏈式反應的檢測等已確認安全性的抗癌伴隨診斷方法,從目前嚴格的上市前許可改為更輕便的程序,意見徵集至2026年1月。然而,新藥與伴隨診斷在同一時間點同時批准的案例仍然很少見,在實際臨床實踐中經常會出現藥物先發布而相應檢測隨後發布,或相反情況。

在韓國,也有越來越多企業跟隨這一趨勢。Arontier公司建立了一個連貫的平台,涵蓋了新藥探索的全過程:使用多元組學數據尋找新靶點和候選物質的REMEDY、設計小分子化合物的DRUG DESIGNER,以及透過病理影像劃分腫瘤微環境的PATH。在2026年美國臨床腫瘤學會年會上,除了Lunit以外,S.T.Cube、TiumBio、ImmuOnsia、OnconicTherapeutics等國內企業也計劃發表人工智能生物標記物和合成致死率基礎抗癌藥物的臨床數據。

儘管取得了如此多的成就,人工智能基礎的靶點發現要在醫院日常工作中紮根,仍需克服若干障礙。首先遇到的問題是用於模型訓練的數據本身的偏差。目前積累的大多數基因組和臨床數據庫偏向於西方國家或特定發達國家人口,將相同的模型直接應用於其他種族和地區患者時,預測往往會出現偏差。數據偏差必然導致模型輸出結果也會出現偏差。這可能導致擴大健康差距的風險。

另一個障礙是深度學習模型為什麼給出某些答案,人類難以逐一追蹤。多層堆疊的神經網絡和混合多個模型的方法具有出色的預測性能,但很難透視其內部,通常被稱為黑盒。在患者生命攸關的處方前,以及在新藥批准決定權掌握在監管機構手中的關鍵時刻,這種不透明性是難以接受的。

因此,研究人員正在努力採用SHAP(Shapley加法解釋)或LIME(局部可解釋模型無關解釋)等可解釋人工智能技術,將模型的判斷是源自哪個基因變異或哪條代謝途徑轉換為人類可以理解的語言。

法律與制度的空白依然存在。在將散落於各個醫院與各個國家的患者數據收集起來進行訓練的過程中,保護個人隱私是最敏感的問題。為了遵守歐盟的《一般資料保護規則》(GDPR)或美國的《健康保險便利和責任法案》(HIPAA),不將原始數據移至外部,而僅在各醫院內部訓練模型,隨後只交換結果值的聯邦學習技術正逐漸普及,但是各國不同的數據標準與加密成本依然是絆腳石。

關於人工智慧自行找出的標靶或設計的生物標記,應該承認誰為發明人,以及其權利究竟屬於開發演算法的公司,還是提供數據的醫院,這類爭議正在世界各地發生。美國食品藥物管理局與歐洲藥品管理局為了因應這些問題,正在不斷修改法規文件,以闡明模型的透明度與品質管理,以及上市後的安全性責任歸屬。

業界預測,在2025年與2026年之間,從這類人工智慧找出的標靶出發的藥物,將會出現十件以上的臨床結果,而且隨著這些成績的累積,結合體學與文獻數據來尋找標靶的這種方法的可靠度,也預期會隨之提高。

從2025年與2026年間展現出的趨勢來看,多體學與人工智慧的結合,似乎已經度過炫耀計算能力的實驗室展示階段,正在轉移到能實際影響患者生存期間的臨床現場。Triumph 試驗的 HER2 患者群、服用 Lentosertib 的肺纖維化患者群,以及針對 MTAP 缺失癌症的新抑制劑,全都是在將異質的體學數據統合為一,並將文獻中的

知識編織成知識圖譜之後才獲得的結果。

將來自不同醫院、不同國家的數據統一標準進行連接，並將黑盒子模型的判斷依據轉換為人類可以理解的語言，這項工作極有可能決定未來幾年的成敗。從患者的一張組織切片、一滴血中發出的信號，與地球另一端研究人員的知識圖譜相連結的當下趨勢，似乎正在重新建構新藥開發這項古老工作的根基。

韓國政府也正在推動於國家層面累積多體學數據的計畫。由疾病管理廳主導的國家整合生物大數據建置計畫，將收集100萬名國民的基因體與臨床資訊作為目標。在從2024年延續至2028年的第一階段中，決定招募罕見疾病患者、重症患者與一般參與者，共計77萬2000人。有意願參與的國民可前往全國38個招募機構的其中一處，填寫同意書與問卷，留下血液或尿液等檢體，並同時提供自己的臨床資訊。

計畫團隊說明，這樣收集來的資料，將成為尋找適合韓國人的標靶與生物標記的基礎。如此收集而來的基因體數據預計將從2026年起陸續對研究人員開放，截至2026年4月底，參與者已增加至16萬5722人。

然而，對新藥開發非常有用的重症患者數據，卻遠遠達不到目標。截至2026年4月底，重症患者在14萬人的目標中，僅收集到了1萬1274人，勉強超過了8%。罕見疾病患者也同樣在4萬7000人的目標中，僅收集到了4476人，僅佔9.5%。計畫團隊表示，其原因在於缺乏收集檢體與說服患者的研究護理師，而且醫療團隊主動參與的動機也很薄弱。

回想全世界用於基因體研究的數據中，歐洲裔人口比例超過90%的調查結果，這項將填補韓國人與其他亞洲人基因資訊的計畫，其重要性絕對不容小覷。儘管如此，如果實際罹病者的資料累積得如此緩慢，那麼糾正數據傾斜的機會，似乎也將長時間難以掌握。

發掘出的生物標記若要應用於醫院處方，必須依序跨過三個門檻。第一是分析效度，觀察檢測工具是否能準確且一致地捕捉到目標標靶。第二是臨床效度，確認該標靶能多準確地預先告知患者的治療反應。第三是臨床實用性，評估使用這項檢測後，患者的治療與結果是否確實有所改善。美國食品藥物管理局建議，在設計新藥候補的第三期臨床試驗時，就應將伴隨式診斷檢測一併納入。

因為這樣一來，藥物與檢測在同一天同時獲得批准的可能性才會增加。如果檢測準備延遲，就會發生藥物雖然先問世，卻因為沒有工具可以篩選出該使用該藥物的患者，而導致處方延遲的情況。韓國食品藥物安全處也規定，若要使用伴隨式診斷醫療器材進行臨床效能試驗，必須獲得額外的計畫批准，以遵守類似的原則。

第 2 部

AI主導的新藥設計與物質發掘

2.1

3D蛋白質結構預測與基於結構的藥物設計 (SBDD)

2020年11月30日，全球數百名結構生物學家聚集在電腦螢幕前。在第14屆國際蛋白質結構預測學術會議 - 通常被稱為CASP的這場大會中，Google DeepMind製造的人工智能AlphaFold 2的評分結果公開的那一刻。螢幕上顯示的分數遠遠超過90分。用X射線結晶學或極低溫電子顯微鏡需要花費數個月，有時甚至數年才能勉強確認的蛋白質三維形狀，一台電腦竟然預測得和實驗結果幾乎完全相同。

困擾生物學家近半個世紀的難題在那一天基本上得到了解決。在1970年代初期，氨基酸序列決定蛋白質三維結構這一原理本身早已被發現，但僅根據序列推算出實際形狀的工作卻在此後半個世紀內一直未能解決。

蛋白質是由稱為氨基酸的小部件以鏈狀長度相連的物質。在細胞中剛剛製造的鏈就像線球一樣隨意伸展，但在水中瞬間巧妙地摺疊，獲得獨特的三維形狀。如果想到摺紙就容易理解了。同樣是正方形紙張，根據摺疊的順序和角度不同，可以摺成鶴，也可以摺成船。

蛋白質也是這樣，氨基酸排列的順序（即序列）決定了摺疊方式，如果形狀有絲毫偏差，就無法發揮原本的作用，或者粘附在不該粘附的地方而引起問題。這就是為什麼像阿茲海默病和帕金森氏病這樣因蛋白質錯誤摺疊而聚集產生的疾病有很多。

製藥公司在設計新藥時首先需要了解的信息就是這種三維形狀。因為必須精確了解導致疾病的蛋白質具有什麼樣的形態，才能設計出能附著在該蛋白質上阻止或幫助其功能的小分子。這種基於靶向蛋白質三維結構來設計藥物的方法被稱為結構基礎藥物設計，簡稱SBDD。過去要獲得這種結構，需要進行X射線結晶學實驗（製作蛋白質晶體，照射X射線，讀取衍射圖樣），或者進行極低溫電子顯微鏡實驗（將樣本冷凍到極低溫，照射電子束，重疊影像），通常稱為冷凍電鏡。

兩者都需要熟練的人力、昂貴的設備，以及數個月到數年不等的時間。有時候，將蛋白質製成晶體的過程本身無法按預期進行，必須徹底改變方法方式的情況也並不罕見。

AlphaFold 2在CASP舞台上展示的預測能力根本上動搖了這種古老的工作方式。2024年5月，Google DeepMind和其子公司Isomorphic Labs公開了後續版本AlphaFold 3，更進一步。AlphaFold 3不僅限於預測單個蛋白質的形狀，還用一個神經網絡共同預測包含遺傳信息的DNA和RNA、可作為藥物候選的小分子配體、金屬離子，甚至翻譯後修飾在內的整個複合體的三維結構。

採用進化的EvoFormer模塊與擴散模型相結合的方式，從原子散亂如雲的狀態開始，逐步消除雜訊，找到接近正確配置的排列。公開時學界認為相比既有方法，相互作用預測準確度提高了超過50%。

雖然擴散模型這個詞彙陌生，但如果想起老式電視螢幕的雜訊聲，原理就容易理解了。就像螢幕上滿是雜訊的狀態中頻道逐漸調整清晰的畫面浮現一樣，AlphaFold 3從原子隨機散亂的雜訊狀態開始，一步步消除雜訊，逐漸完成接近實際分子配置的圖像。在重複這個過程數十次的過程中，最初模糊的原子雲逐漸縮小成蛋白質和配體相互嵌合的清晰三維結構。

這樣製造的預測結構逐步堆積，成為了AlphaFold蛋白質結構數據庫這一公共資產。2025年介面經過翻新的這個數據庫上傳了超過2億1400萬個蛋白質結構，可以通過一次搜索查看包括人類在內實際上所有生物物種的蛋白質。實際上，中國研究人員利用這個數據庫和AlphaFold找到了葡萄膜黑色素瘤患者出現的315個異常結構，並獲得了新治療靶點的線索。這在實驗室里解明一個蛋白質結構需要花費數年時間的時代是難以想像的規模。

能夠感受到的變化首先體現在速度上。在AlphaFold 2出現之前，新藥候選發現的第一個關口——靶向蛋白質結構確認工作需要花費數個月。自AlphaFold 2推出後，這項工作縮短到了以天為單位，節省下來的時間被直接再投資到下一階段的對接和模擬中。

AlphaFold 3最初公開時，研究人員無法直接下載代碼運行，只能將序列輸入由Google運營的AlphaFold服務器並獲取結果，這讓人感到遺憾。進入2025年，Google DeepMind擴大了範圍，允許非商業性學術研究下載推論代碼和申請獲得模型權重，同時將AlphaFold服務器每天的工作處理量增加了一倍。期間，開發替代方案的行動也很活躍。

中國ByteDance推出了完全復現AlphaFold 3的開源模型ProteinNX，哥倫比亞大學的al-Quraishi實驗室以Apache 2.0許可證公開了OpenFold3-preview，該許可證允許學術和商業用途免費使用。曾經被一家公司幾乎獨占的技術在短短幾年內擴散到了多個研究團隊共享的開源生態系統。

如果說 AlphaFold 3 是以精緻度取勝，那麼 Meta 所開發的 ESMFold 則是以速度取勝。大多數的結構預測人工智慧，都必須在多個生物物種中找出相似的蛋白質，並將它們並列進行多重序列比對後，才能計算出結構。這項工作光是在資料庫中進行搜尋，就會耗費相當多的時間。相反地，ESMFold 以預先學習過數億個蛋白質序列的大型語言模型 ESM2 為基礎，不需要多重序列比對，只要輸入一個序列，就能立刻給出結構。

對於像是在演化上幾乎沒有相似親戚的孤兒蛋白質，或是新發現的病毒蛋白質這類缺乏適當比較對象的情況，ESMFold 的優勢就會特別突出，目前已廣泛應用於大規模篩選與發掘新酵素的現場。Meta 也以這項技術為基礎，大規模預測並公開了自然界中存在但尚未被人類所知的蛋白質結構。

與 AlphaFold 共同支撐蛋白質結構預測生態系的另一個主軸，是由華盛頓大學 David Baker 教授研究團隊所開發的 RoseTTAFold。其中，最新版本的 RoseTTAFold All-Atom 就像 AlphaFold 3 一樣，會同時預測蛋白質、核酸與小分子配體混合在一起的複合體結構。Baker 教授研究團隊真正的拿手絕活，不僅止於預測結構，更在於能從白紙狀態開始，描繪出能與目標分子結合的全新蛋白質。應用擴散模型的 RFdiffusion 系列工具，便是負責這項任務。

RoseTTAFold 於 2021 年首次發表，採用的是三重神經網路結構，也就是將胺基酸序列資訊、原子間距離資訊與三維座標資訊，分別在不同的神經網路通道中同時處理，並讓彼此互相交換資訊。這款模型與 AlphaFold 2 幾乎在同一個時期問世，雖然在預測準確度上被評為稍微不及 AlphaFold，但它具有計算所需的電腦資源相對較少的實用優勢，因此在個別的研究室裡也能實際運行。

這項工具在 2025 年 12 月 3 日以 RFdiffusion3 的名稱更往前邁進了一步。這款處理到原子等級的模型，不僅是蛋白質，還能同時設計出與 DNA 或各種配體相契合的蛋白質，推論速度更比之前的版本快了近十倍。該團隊發表表示，可以在電腦中描繪

出性能接近大自然長時間淬鍊出的酵素的全新酵素，這在學術界引起了不小的迴響。這是因為，過去開發新工業用酵素需要耗費數年的週期，如今可望轉變為先在電腦中描繪出候選者，再於實驗室中進行驗證的模式。

同一個研究團隊在此之前，就已經在國際學術期刊《自然》上發表了成功設計出能附著在沒有固定形狀、難以處理的本質無序蛋白質上的結合劑的論文，同時也提出了完全只用電腦設計抗體重鏈與輕鏈全體的 RFantibody 研究成果。既然抗體已經被作為多種難治疾病治療的主力藥物，能夠在電腦螢幕上設計抗體骨架的這項事實，便具備了改變生物製藥開發策略本身的潛力。這意味著，重心正在從讀取結構的人工智慧，轉移到能隨心所欲創造結構的人工智慧。

一直苦苦等待這類技術的，正是被稱為不可標靶蛋白質的領域。沒有明顯凹陷的結合口袋而缺乏藥物能抓住的把手的蛋白質、在細胞內不斷改變形狀的本質無序蛋白質、表面寬廣平坦而單靠一個小分子絕對無法阻擋的蛋白質與蛋白質之間的結合部位，這些對象長久以來都被排除在新藥開發的候選名單之外。這是因為，僅憑一張靜態的結晶結構照片，很難將這些蛋白質作為設計對象。學習了數百萬個序列與結構資料的生成式人工智慧，正在用另一種方式跨越這道高牆。

雖然本質無序蛋白質這個名字聽起來有些陌生，但它的意思並不難懂。與大多數蛋白質會像摺紙一樣完美摺疊成固定形狀不同，這種蛋白質沒有固定的形態，會在水中持續搖晃擺動。既然沒有能抓在手裡的形狀，自然也就沒有合適的鑰匙孔可以對應，因此長久以來都停留在新藥開發者的關注之外。在 2026 年，還出現了專門收集口袋較淺的標靶來比較性能的 ShallowBench 基準測試，開始仔細篩選處理不可標靶蛋白質的人工智慧模型的實力。

蛋白質除了正門之外，還有後門。如果藥物經常鎖定的活性部位是正門，那麼異位部位就相當於遠離活性部位的蛋白質表面上的後門。只要有小分子輕輕觸碰這個後門，蛋白質整體的立體形狀就會改變，進而調節酵素的活性或訊號傳遞。像 DeepAllo 這種結合蛋白質語言模型萃取出的特徵與結合口袋形狀一起分析的工具，能夠找出僅憑一張結晶學照片無法顯現的激酶或 G 蛋白偶聯受體的異位部位。即使是正門緊閉而無從下手的蛋白質，只要利用後門，就能夠選擇性地調節其功能。

實際的成果也隨之而來。AlphaFold 找出的表皮生長因子受體，也就是 EGFR 突變的結構，以及 KRAS 蛋白質的立體形狀變化資訊，實際上已用於改進肺癌與乳癌患者使用的精準標靶治療藥物 Sotorasib 與 Erlotinib 的療效。長久以來，KRAS 是一個因為表面平滑、沒有把手而聲名狼藉的標靶蛋白質。雖然 KRAS 是與癌症相關基因的事實本身在很久以前就已經被發現，但因為找不到能抓在手裡的結合口袋，這數十年間始終無法撕下不可標靶基因的標籤。隨著結構資訊變得越來越詳細，我們也正在逐步找出攻略這類蛋白質的方法。

近年來，AlphaFold系列模型的研究更進一步，不是預測單一的正確結構，而是同時預測蛋白質可能呈現的多種形態的集成生成研究也很活躍。考慮到蛋白質在結合前後呈現不同的形狀，與其擺出一張照片，不如把多張快照並排展示，這是更貼近實際運動的嘗試。

但AlphaFold呈現的結構就像拍下某個瞬間的一張照片。實際的蛋白質不是照片中靜止的樣子，而是被體溫和水分子包圍著，不斷搖晃和扭轉的活的結構物。這跟看影片而非照片才能看到真實運動是一樣的道理。用計算重現這種運動的技術就是分子動力學，簡稱為MD模擬。

MD根據牛頓運動定律，在極短的時間間隔內持續更新每個原子的位置和速度，追蹤在從皮秒到毫秒的時間尺度內，蛋白質如何搖晃，以及配體結合時周圍形狀如何隨之改變。

如果說人工智慧基礎的分子對接快速生成多張結合的快照，那麼MD模擬就負責驗證這些快照在實際體內環境中能維持多長時間、有多穩定。配體結合時蛋白質形狀隨之變化的誘導適配現象，甚至水分子進出結合位點產生的微妙效應，MD模擬都能捕捉。僅根據對接結果判斷有前景的候選物質，在MD模擬階段幾個小時內結合就解開的情況也並不少見。經過這樣的驗證，候選物質才獲得進入下一階段的資格。

化學鍵新生或斷裂的瞬間，也就是酶催化反應或藥物以共價鍵結合到靶蛋白的瞬間，其中的電子運動即使用MD模擬也無法精確計算。要處理這樣細微的電子世界需要量子力學方程。但是如果要用量子力學計算整個蛋白質，即使是超級電腦也需要難以承受的時間。因此，只在結合實際發生的活性位點周圍進行精密的量子力學計算，而將其餘龐大的蛋白質骨架和水分子用計算成本較低的古典力學處理，這種量子

力學和分子力學的混合計算，簡稱為QM/MM技術。

這類似於建造建築物時不用相同的精密度全部施工，只在人手直接接觸的門把手採用精密部件，其餘牆壁和柱子則用標準材料建造的做法。

近來，結合量子機器學習和圖神經網絡的計算方法正在大大降低QM/MM的成本。國內浦項科技大學的研究團隊開發了結合電荷平衡原理和圖神經網絡的機器學習勢，提出了既接近量子力學精度又計算速度快得多的方法。在2026年的一次新藥開發會議上，Schrödinger公開的工作流程將基於物理的模擬和生成式人工智慧連接起來，據稱僅用4小時30分鐘就生成了超過1萬2千個化合物候選，引起了關注。

這意味著在臨床前階段，通過計算機預先篩選化合物毒性和藥動學特性的驗證的可信度正在不斷提高。

隨著多項技術的積累，人工智慧尋找新藥候選的程序大體上要經過四個階段。研究人員首先輸入導致疾病的蛋白質的胺基酸序列。接著，像AlphaFold這樣的人工智慧根據這個序列預測蛋白質的三維結構。得到結構後，分子對接程序將候選化合物逐一與該結構的結合位點匹配，篩選出吻合度高的候選。

最後，QM/MM計算和MD模擬精密地再次確認結合位點周圍的電子運動和隨時間的穩定性，進一步完善最終候選。從一個序列出發到得到驗證的候選物質的這個過程，過去需要花費幾年，現在已經壓縮到幾周或幾個月。

這個流程中第三階段的分子對接就像逐一將鑰匙與鑰匙孔匹配的工作。將靶蛋白表面的凹槽稱為鑰匙孔，候選物質相當於在這個孔中尋找匹配的鑰匙。但在新藥開發現場，區別在於需要在數百萬、數億個候選中同時測試哪把鑰匙符合一個鑰匙孔。過去是像AutoDock Vina或Glide這樣依賴物理能量函數和經驗評分的古典程式承擔這項工作。

近年來，結合深度學習、圖神經網絡和擴散模型的DiffDock、SurfDock、Chai-1、Boltz-1等人工智慧對接模型陸續出現，在學習靶蛋白表面形態和靜電特性後，能更快地找到靈活的結合姿態。查閱數億個化合物庫的虛擬篩選時間因此大大縮短。

蛋白質與蛋白質相互作用的對接情況略有不同。像ZDOCK這樣基於物理定律的傳統程式細緻地反映結合過程中發生的立體結構變化，而許多深度學習模型則假設結合

前後的形態固定，只進行靜態形狀匹配，有時會扭曲實際結合環境。不過，在蛋白質與小分子相互作用的對接中，最近的情況有所不同。

2025年發表的名為PoseX的新基準測試將訓練數據中沒有的陌生蛋白質和配體組合進行混合測試，在這裡一些人工智慧模型的表現也超過了傳統基於物理的方法。但說哪一方已經完全佔據優勢還為時過早。

2025年6月，麻省理工學院與新藥開發公司 Recursion 共同推出的 Boltz-2 為這一趨勢又添上了一筆。因為它不僅僅停留在預測結構，而是在同一個模型內預測結合強度，也就是結合親和力，是首個共摺疊模型。它被評價為能以高達一千倍的速度，提供與被稱為自由能微擾的精密物理基礎計算相當的準確度，同時吸引了學界與業界的關注。

Boltz-2 最初是以開源形式發布，但在 2026 年上半年推出的後續版本 Boltz 2.1 轉變為僅以 API 形式提供的封閉型。這是一個展示技術趨勢從開放再次轉向商業化的案例。

圍繞公開範圍的拉鋸戰，似乎是貫穿這整個領域的趨勢。在 AlphaFold 3 向學界敞開大門，且有 Protenix 與 OpenFold3-preview 等替代方案緊隨其後時，原本能預測到結合親和力的 Boltz-2 卻在後續版本中鎖上了大門。不亞於建立一個優秀的預測模型，決定將該模型向誰開放以及開放多少程度，似乎也正在成為決定這個生態系勝負的因素。

在華麗的成果背後，尚未解決的缺陷也很明顯。如果用 PoseBusters 這樣嚴格的驗證標準來審視 AlphaFold 3 的結合姿態預測結果，在基本資料集中僅有 76% 左右，即使放寬評分標準，也停留在 84% 上下。真正嚴重的問題另有其他。在深度學習模型得出的結合姿態中，原子與原子在物理上互相重疊，或者無法區分像鏡中倒影般方向相反的光學異構物等，在現實中不可能存在的結構並不罕見地混雜在其中。

面對與訓練資料相似的結合類型時，人工智慧展現出令人驚訝的準確度，但在從未見過的陌生蛋白質或完全全新的化合物面前，預測能力就會大幅下降的情況，也在多個基準測試中被反覆證實。因此，將人工智慧得出的結合姿態用物理基礎計算再修飾一次，以糾正不自然結構的程序，正逐漸成為提升準確度的實質解決方案。

正面瞄準這個問題的是 2021 年從 Google DeepMind 分拆成立的 Isomorphic Labs。2026 年 2 月，該公司發布的 IsoDDE 宣布，在與訓練資料的序列相似度不到 20% 的極其陌生配體結合案例中，展現了比 AlphaFold 3 高出約兩倍的準確度。

Isomorphic Labs 在與 Novartis、Eli Lilly、Johnson & Johnson 等大型製藥公司攜手共同開發新藥候選藥物的同時，也在腫瘤學和免疫學領域培育了自家的候選物質。

Demis Hassabis 在 2026 年 1 月的達沃斯論壇上表示，雖然原本定於 2025 年的首次臨床試驗日程被推遲到了 2026 年底，但也明確表達了為患者投藥的日子已不遠。2026 年初，Johnson & Johnson 也宣布展開了深入合作，以 AlphaFold 3 為基礎，設計能阻止蛋白質與蛋白質之間結合的新型抑制劑。

在韓國國內也有相關研究緊隨其後。KAIST 研究團隊展示了一種人工智慧，它不僅超越了預測藥物是否附著在目標蛋白質上，甚至能同時預測附著後是否實際按照預期運作。這是一項將僅觀察是否結合的現有模型標準提升了一個層次的嘗試。由 Seoul National University 化學部教授 Seok Cha-ok 於 2020 年創立的新創公司 Galux，以在 AlphaFold 出現之前就在國際蛋白質結構預測競賽中取得名列前茅的成績為基礎，利用以物理定律為骨架的人工智慧來設計蛋白質。該公司與 LG Chem 進行了共同開發抗癌蛋白質的合作，期間募集的投資金額達到了 680 億韓元。

同一時期在韓國成立的 Simplex 也值得關注。由在全球製藥公司 GSK、BMS 和 Amgen 開發新藥研究平台超過 20 年的代表 Cho Sung-jin 所創立，這家公司主打能同時顯示預測結果產生原因的可解釋人工智慧，其優勢在於能在進入臨床階段之前，提前篩選出失敗可能性很高的候選藥物。

直接確認結構的實驗技術也與人工智慧攜手合作。處理超低溫電子顯微鏡影像的軟體 RELION 或 cryoSPARC 中引入了人工智慧，將挑選、分類粒子並重新組裝成三維立體的過程自動化，把過去需要耗費數週的處理時間縮短到了幾個小時。在形狀模糊且難以解析的密度圖面前，AlphaFold 得出的預測結構會作為參考地圖，協助拼湊模型的過程。這等於是轉變成了一種結構，由實驗設備產生的模糊原始資料與人工智慧預先繪製好的接近標準答案的地圖，彼此互相核對。

以 Pfizer、Novartis、AstraZeneca 為首的大型製藥公司，將 AlphaFold 與 RoseTTAFold 引入公司內部的新藥開發管道，也與這種趨勢息息相關。結構生物學部門與運行預測模型的運算團隊共用同一間辦公室的景象，現在也不再陌生。實驗與計算交替互相驗證的循環結構，正持續縮短掌握單一個目標蛋白質所需的時間。

在所有這些預測的基礎中，深度學習特有的不透明性依然存在。神經網路為何挑選特定的原子排列作為標準答案，研究人員要一行一行地追蹤並解釋其判斷的內部邏輯仍然不容易。由於難以追溯單一個結合親和力分數是出自何種物理依據的這種特性，計算結果正被當作需要透過實驗來縮小範圍的候選清單，而不是最終答案。

整個流程的官方里程碑是 2024 年諾貝爾化學獎。瑞典皇家科學院將該年諾貝爾化學獎的一半授予大衛·貝克，另一半共同授予德米斯·哈薩比斯和約翰·茱珀。貝克因設計出自然界不存在的新蛋白質而獲得認可，哈薩比斯和茱珀因只根據胺基酸序列預測蛋白質結構而獲得認可。獎金 1100 萬瑞典克朗流向人工智能公司首席執行官、研究人員和大學教授，而非單一實驗室研究人員，這個事實本身就是不尋常的。

可以說，科學界本身將通過計算機計算解決化學和生物學長期難題的成就認定為最高權威，隨後整個製藥業界加快了將人工智能基礎結構預測從邊緣實驗轉變為新藥開發標準程序的進程。曾經長期被視為計算機科學領域實驗嘗試的方法論在短短數年內躋身於獲得諾貝爾獎的正統科學之列，而這種認可反過來也在吸引更多年輕研究人員進入該領域。

隨著獲得結構的方法從單一實驗擴展到多層次計算，新藥開發早期階段的瓶頸不再是對結構的未知。瓶頸似乎在轉移到從大量預測結構和結合候選物中判斷應先將哪個放在實驗台上的判斷。

輸入單個蛋白質序列並在短時間內獲得預測結構、在該結構上測試數億個化合物、通過物理計算重新驗證有前景的候選物，這些工作現已成為多家製藥公司和初創公司的日常業務。不過，計算機屏幕上完成的結合構象是否在實際人體內保持不變，最終必須在表面電漿共鳴或 X 射線晶體學等傳統實驗台上再次確認。這種預測和實驗之間的往返交流越快，長期以來被認為無法觸及的蛋白質列表就越來越短。

正確確定結構和在該結構上準確測量結合強度是不同的問題。即使 AlphaFold 或 Boltz 系列模型能夠幾乎完全匹配每個原子的位置與實驗結果，指示結合在體內維持的強度和時間的結合親和力預測仍然難以給出答案。這是因為結合親和力是無法從單個靜止照片中捕捉到的值。水分子進出結合位點的程度、鏈折疊時失去的自由度、胺基酸和配體各自所帶的電荷，必須全部加減才能得到一個值。

Boltz-2 在某些驗證中表現與基於物理的計算相當，但在針對形態不斷變化的蛋白質的其他驗證中，即使是具有實際活性的化合物也有時表現與隨機選擇的化合物相似。考慮到即使實驗測量結合強度也存在約 1kcal/mol 的誤差，無論計算多麼精密，從一開始就存在著一道難以逾越的牆。

隨著這些預測工具的廣泛傳播，完全不同的擔憂也隨之增長。因為當可以設計自然界不存在的蛋白質的工具放在任何人都能接觸到的地方時，通過只稍微改變危險蛋白質序列而繞過監視網的途徑也隨之開放。美國政府在 2025 年 5 月通過行政命令指示改進生物學研究的安全和保安，並規定重新制訂政府資助研究中使用的合成核酸的審查標準。五個月後的 10 月，由微軟首席科學官埃里克·霍維茨領導的研究團隊在國際學術期刊《科學》上發表了該審查網中的漏洞。

他們取得包括蓖麻毒素和肉毒神經毒素在內的 72 種危險蛋白質的序列，使用公開的人工智能設計工具進行變換，在計算機中創建了超過 76,000 個設計，並確認其中大量可能通過基因合成公司使用的現有監視軟體。雖然這是計算機內的實驗，實際上並未合成蛋白質，但研究團隊與國際基因合成聯盟一起改進了軟體，之後仍有約 3% 的改變序列未被篩選出去。

計算給出的答案與確認是否可以實際信任該答案的能力之間的差距似乎是該領域反復出現的問題。無論是在測量結合強度還是在識別危險序列中，生成能力超越驗證能力的局面將來可能還會持續一段時間。

2.2

基於深度學習的候選物質生成 (De Novo Design)

2022年，在中國上海的一家製藥公司實驗室裡，一台電腦繪製出了一種可以治療名為特發性肺纖維化疾病的物質。特發性肺纖維化是一種肺部組織在沒有明確原因的情況下逐漸變硬，導致呼吸困難的難治之症。這是一種非常嚴重的疾病，患者在確診後的平均存活時間只有三到四年左右，因此患者們對於新治療藥物的等待非常迫切。設計這種物質的公司是Insilico Medicine，其鎖定的標靶是細胞內負責傳遞發炎與纖維化訊號的酵素TNIK，而這樣誕生的物質被命名為ISM001-055。

從確定標靶的時候開始，到進入以人為對象的臨床一期試驗，所花費的時間是三十個月，還不到三年。由人工智慧包辦從發掘標靶到分子設計，甚至進入臨床試驗的案例，在當時的世界上也是屈指可數。2025年，在國際學術期刊《自然醫學》與美國胸腔醫學會學術大會上，公開了臨床2a期試驗的結果。這項在中國21家醫院召集了71名患者所進行的試驗，達成了安全性的目標，而用來測量肺部能吸入和呼出多少空氣的指標也就是用力肺活量，也隨著給藥劑量的增加而跟著改善。

在2026年7月7日，發表了即將進入臨床三期試驗的消息。人類數十年來透過不斷嘗試與錯誤而四處尋找的新藥候補，現在由電腦從白紙上描繪出來，並在患者的身體上一一確認其結果。這種物質後來甚至獲得了Rentosertib這個國際非專利藥品名稱。國際非專利藥品名稱是世界衛生組織所賦予的官方藥物名稱，這也意味著它已經超越了一家公司的實驗室代碼，開始被視為全世界共同稱呼的正式藥品。

一個新藥候補從發現標靶到通過動物實驗，用傳統的方式需要花費超過4到6年的時間以及龐大的金錢。根據製藥業界經常引用的估算，將一種新藥推出市場所需的費用超過20億美元。在這個過程中，大部分的候補物質都會在中途被淘汰，業界人士將這個階段稱為死亡之谷。業界也會用伊魯姆法則（Eroom's Law）這個詞來指稱這種趨勢。這是將半導體效能會持續提升的摩爾定律反過來拼寫的名字，意思是即使技術進步，開發一種新藥所需的費用和時間反而會越來越多。

即使是進入臨床試驗階段的候補物質，能夠存活到最終批准的比例也不到十分之一，這是業界長久以來通用的經驗法則。當人工智慧最初進入這個產業時，它所擔任的角色是快速翻找像ZINC或Enamine REAL這樣堆滿了數十億個化合物的倉庫。電腦雖然能在眨眼之間掃視多達數兆個化合物並挑選出有希望的候補，但這樣找出來的物質，終究只不過是某人已經製造出來並放進倉庫裡的東西而已。如果倉庫裡沒有正確答案，那麼無論翻找得再快，也找不到正確答案。

深度學習為新藥開發帶來的新事物，不是更快翻找倉庫的技術，而是從根本上全新創造出倉庫中沒有的物品的能力。科學家們估算，理論上具有可以成為藥物性質的分子數量高達 10^{60} 到 10^{100} 個。既然宇宙中的星星數量大約被估算為 10^{24} ，那麼化學所蘊含的可能性大小，就比夜空中的星星還要多到無法估量。這樣的數字是來自於組合的力量。即使只用十種樂高積木，將二十個積木連接在一起的方法數量，就已經增加到我們大腦無法計算的地步。

實際上，根據樂高公司計算的結果，即使只堆疊六個長得完全一樣的積木，可以排出的形狀也超過9億種。然而，構成實際分子的原子種類超過一百種，如果再加上將這些原子以鏈狀、分支狀、環狀連接起來的方式，其組合增加的速度會比樂高還要猛烈得多。人類在實驗室裡花了一個世紀製造並測試過的化合物數量，與這個龐大的空間相比，是小到無法計算的。

在這個廣闊的空間裡，如果電腦要描繪出新的分子，首先必須要有將分子轉換為數字和文字記錄下來的方法。被廣泛使用的方法是SMILES。SMILES是一種將碳、氧、氮等原子以及它們之間的鍵結，用幾個英文字母和符號記錄下來的標記法，會把分子像簡短的句子一樣寫成一整行。就像人類聽寫別人唸出來的內容一樣，電腦會看著SMILES這些文字，學習讀取並寫出分子樣貌的方法。舉例來說，咖啡和茶裡面所含的咖啡因，可以用CN1C=NC2=C1C(=O)N(C(=O)N2C)C這樣一個字串來完整表達。

碳、氮、氧是以什麼順序連接，以及在哪裡形成環狀，都毫無遺漏地包含在這短短的一行之中。問題在於這一行文字非常敏感。就像一個句子中只要有一個拼字錯誤就會導致意思不通一樣，SMILES字串也只要有一個字母出錯，就會製造出在化學上無法存在的假分子。因為生成模型每次大量產出數萬個候補時，都會混雜著這些虛假的東西，使得研究人員長久以來都必須與這種標記法的弱點奮戰。

名為SELFIES的標記法，就是正面針對這個弱點所制定出來的。它在數學上保證，用SELFIES寫下的字串無論如何組合，都必定能轉換為實際可能存在的分子。就像是邊緣形狀預先裁切好而不會互相錯開的拼圖一樣，SELFIES直接消除了被錯誤連接的任何餘地。

隨著這個標記法開始被廣泛使用，訓練人工智慧模型的過程變得顯著穩定，並且可以放心地將每一個字串都轉換成分子來使用。只不過，無論是SMILES還是SELFIES，它們同樣都是單行的字串，因此無法另外告訴我們原子在實際空間中是以什麼形狀排列的。

能夠將分子視為網狀糾纏的結構本身，而不是單行字符串的技術，就是圖神經網絡（GNN）。原子一個一個地置為點，原子之間的化學鍵描繪為連接這些點的線，分子就會呈現出類似地鐵線路圖的網狀形態。圖神經網絡直接閱讀這份線路圖，同時學習每個原子周圍的電子環境和立體形態。被稱為圖卷積神經網絡或圖注意網絡的模型，基於這樣習得的信息，可以相當準確地預先告訴我們某種分子對目標蛋白質的粘著程度有多好，以及在水中的溶解度有多好。

虛擬篩選的速度和準確度大幅提升，這項技術使電腦能夠先行篩選數萬到數十萬個候選分子，而無需人工逐一檢視。

圖卷積神經網絡通過跨多層彙聚鄰近原子的信息，以總結整個分子特徵的方式運作，而圖注意網絡除此之外，還自行給予每個鄰近原子不同的權重，以區分其重要程度。位於目標蛋白質結合口袋中央的原子與位於分子邊緣的原子，對結合力影響不同，注意機制能夠自動區分這種差異。正因為有這樣的預測能力，研究人員無需逐一合成數百萬個候選物質並放入試管進行測試，而是可以先在電腦中篩選大部分候選物，只將真正有潛力的數百個提上實驗台進行測試。

既然已經掌握了讀寫分子的語言，接下來就是用這種語言實際創作新的文章。承擔這項創作工作的三項技術是變分自動編碼器（VAE）、生成對抗網絡（GAN）和擴散模型。變分自動編碼器學習大量化合物，由一個將每個分子的複雜形態壓縮為單一坐標的編碼器，和一個將該坐標重新恢復為完整分子的解碼器組成。在這樣壓縮的坐標空間中，即使只稍微移動坐標，也可以得到與原始分子在骨架上相似，但溶解度或毒性等性質略有改善的分子。

這種方式類似於廚師的筆記，通過逐步改變烹飪材料的配合比例來尋找新菜肴。就像只是減少半勺鹽並稍微增加糖，便能得到一道並非完全不同的菜肴，而是與原菜在味覺框架上相同但大幅改善的菜肴一樣。一個稱為結點樹變分自動編碼器（Junction Tree VAE）的進階模型，不是逐一連接原子，而是以既有化學意義的小片段為單位來構建分子，因此結果更加現實。

生成對抗網絡則以不同的方式競爭。其結構是持續生成虛假分子的生成器與鑑別者相互對抗並共同提升技能，其中鑑別者用來區分該分子是來自真實藥物數據庫還是生成器創造的。生成器為了欺騙鑑別者，會不斷創造越來越難以與真實分子區別的分。Insilico Medicine曾使用GENTRL模型（一種在生成對抗網絡上加入強化學習的模型），在短短21天內就找到了能夠抑制DDR1激酶的候選物質，該激酶是纖維化和神經退行性疾病中的重要靶點。

然而，這項技術也有明顯的弱點。訓練過程在數學上不穩定，模型經常陷入「模式崩潰」現象，即失去多樣性，只反覆產生少數幾種相似形態的分子。由於這種不穩定性，近年來新藥設計領域出現了明確的轉變趨勢，重心從生成對抗網絡轉向擴散模型。

擴散模型是從照片復原過程中獲得靈感的技術。它的訓練過程是：向清晰照片逐步添加噪聲直至完全模糊，然後逐步去除噪聲來復原原始照片。將這一原理直接應用於分子，就能從隨機噪聲開始，在逐步去除噪聲的過程中清晰的分子結構就會浮現。藥物要在人體內真正發揮作用，必須以立體結構而非平面結構與目標蛋白質相結合。最近的三維擴散模型採用經過數學精心調整的結構，使分子即使旋轉或移動也能保持性質不變，從而從一開始就篩除不現實的扭曲分子。

GeoDiff、DiffDock、DiffSBDD等模型能夠讀取目標蛋白質結合口袋的形狀，同時繪製出與其曲率完美匹配的原子排列和三維坐標。DiffLinker模型甚至能創建在三維空間中自然連接分離化合物片段的結構。2026年初發表的GlintDM模型在優化結合力和藥物特性的同時大幅降低了計算成本，同年推出的DiffGui模型不僅停留在電腦模擬，還進行了實際實驗室的濕法實驗，驗證了設計分子確實有效。

直接利用目標蛋白質結合口袋信息來構建分子的方式稱為基於結構的生成，而不使用目標三維結構、僅基於已知活性物質特徵來構建具有相似性質的新分子的方式稱

為基於配體的生成。如果蛋白質結構已通過實驗確定，基於結構的方式會產生更準確的結果，但由於仍有許多結構未知的靶標，這兩種方式在今天仍然並行使用。

無論擴散模型或生成對抗網路製造出多少分子，其中必然會混雜著雖然能緊密結合標的，卻對人體有害的物質。強化學習就是一種能引導方向，讓分子同時滿足多種條件的技術。強化學習是一種學習方式，由電腦裡的代理程式試著稍微改變分子結構，並從給予評分的評估函數中獲得獎賞或懲罰，藉此自己學會累積高分的方法。這時候的評分不會只看藥效好不好。像是能不能在體內良好吸收、分布、分解並排出，以及有沒有毒性，這些綜合指標ADMET（也就是吸收、分布、代謝、排泄、毒性）在水中溶解的程度，或是通過通往大腦的關卡「血腦屏障」的穿透性，以及對心臟造成的負擔等，都會一起反映在評分之中。

AstraZeneca研究團隊所改進的REINVENT 4這個平台，結合了像語言模型一樣處理分子的人工智慧與強化學習，會反覆修飾分子結構，在提升對標的的效果的同時，也降低對其他地方產生作用而引起副作用的風險。2025年以後，研究人員也持續進行著將強化學習與擴散模型完全結合在一起的研究，設計出能同時符合多個標的，甚至具備立體結構的分子。

在2025年與2026年之間，將強化學習與其他技術完全結合在一起的研究大幅增加了。叫做MDRL的方法，就是將擴散模型與強化學習整個綁在一起，設計出能同時符合多種不同標的，甚至具備立體結構的分子。叫做POLO的方法，會像對話一樣，經過好幾次來反映人類的偏好，並逐步修飾候選物質；而叫做MolMem的方法，則是會記住過去嘗試過的結果，幫助我們只要嘗試很少的次數就能接近目標。強化學習本身也正在快速進化，朝著讓電腦先在內部經歷反覆試錯，再把那些記憶用在下一次設計的方向發展。

這樣建造分子的技術，並不僅限於小型的化合物。2025年，一家叫做Chai Discovery的公司所推出的Chai-2模型，從完全空白的狀態設計出了抗體。它不是模仿已知抗體或結合物質，而是單憑作為標的之蛋白質資訊，就描繪出新的抗體與奈米抗體。奈米抗體是只會自然產生在駱駝或羊駝等動物體內，比一般抗體小得多且更容易處理的結合蛋白質。研究團隊挑選了五十二個從未處理過的標的進行測試，結果發現，雖然每個標的只製造了二十個左右的候選物質，卻在其中一半找到了實際能附著的抗體，而且從開始設計到在實驗室確認結果，花費不到兩個星期的時間。

與前一代的電腦設計技術相比，成功率躍升了一百倍以上。過去用來建造新小分子的生成技術，現在已經將觸角延伸到了設計保護身體的複雜蛋白質武器「抗體」上。在這樣完成的抗體候選物質當中，甚至有超過86%獲得了具備足以與實際治療用抗體媲美的穩定性與生產可能性的評價；而在同一年11月的後續研究中，還額外公開了針對更難處理的標的所進行的抗體設計結果。

如果說表達是基礎，那麼生成就是在這個基礎上實際動手去做的作業。以前人類化學家要花上好幾個月用手描繪和修飾分子的工作，現在電腦只要一兩天的時間，就能畫出數萬種替代方案給我們看。當擴散模型和強化學習攜手合作，建造出符合目標的新分子後，剩下的工作就是將這個成果推向世界。

將新建造的分子不斷推向還沒有人去過的化學空間，以及尋找將已經問世的藥物重新用於完全不同的疾病的方法，這兩件事同時發生的階段，就是所謂的擴展。只有到了這個階段，被關在實驗室裡的人工智慧想像力，才終於能轉化為擺在患者面前的藥物。

能讓藥物重新被利用的人工智慧，是從把疾病、標的與藥物之間的關係編織成一張巨大知識網的步驟開始的。圖神經網路和自然語言處理技術會瀏覽高達數百萬篇的醫學論文與基因表現資料，找出人類研究人員沒能注意到的藥物與疾病之間的隱藏連結。像這樣的一張知識網中，有時會包含數千種疾病和數萬種藥物，以及連接它們之間的數百萬個關係。英國的BenevolentAI分析了這張知識網後發現，類風濕性關節炎治療藥物Baricitinib在平息COVID-19引起的過度發炎反應的同時，還能阻斷病毒進入細胞的通道。

這個預測實際延伸到了大規模的臨床試驗，最終甚至取得了作為COVID-19治療藥物的緊急使用授權。為已經確認安全性的藥物賦予新用途的這種方法，比從頭開始製造新分子的方法，能在更短的時間內送到患者面前，因此非常有吸引力。

然而，無論擴散模型和強化學習描繪出多麼像樣的分子，如果無法在實際的實驗室中把它製造出來，那也是毫無用處的。就算在電腦螢幕上原子和結合在邏輯上是吻合的，也不代表就能真的合成出那個分子。這就像在建築設計圖上雖然可以畫出任何形狀的建築物，但要在現實土地上蓋房子的話，就必須重新考慮能不能買到建材，或者能不能立起承受那個重量的柱子，道理是一樣的。在研究人員之間，甚至用

「實驗室裡的大象」來形容這個問題。意思是指大家都心知肚明，卻很難去處理的大問題。用來填補這個落差的技術，就是逆合成計畫。

從目標複雜分子反向逐步分解到市場上便宜的原料物質，尋找合成路徑的工作與探索化學空間並無二致，候選路徑的數量可達數百萬種。AiZynthFinder 和 AskCos 等工具利用深度學習和樹搜索分析這數百萬筆化學反應資料，找出便宜且收率高的路徑。這種樹搜索通過提前探索多條路徑並選擇成功可能性高的方向來運作，這種方式類似於強化學習通過獎勵良好選擇來學習的方式。

提高合成可行性的工作並不像聽起來那麼簡單。若只用已知的反應和材料來構建分子，合成會變得容易，但結果卻會與現有化合物相似，容易失去新穎性。反之，若貪心地深入化學空間陌生的角落，雖然能得到獨創性的分子，但實際能否合成就難以保證。2025 年的一項研究將這個問題框架化為同時追求三個目標的挑戰：多樣性、合成可行性和生物活性。在追求其中一個目標的過程中容易丟失另一個，這是該領域長期以來的困擾。

2026 年 7 月，Insilico Medicine 在首爾舉行的國際機器學習會議（ICML）上發表了名為 ChemCensor 的新評估標準。它不是採用傳統方式機械地計數正確答案，而是試圖更接近實際化學家通過觀察反應中心和官能團來判斷逆合成路徑優劣的方式。最近，PrexSyn 和 SynLlama 這類研究正在探索將合成可行性不作為篩選的最後階段，而是從一開始就嵌入分子設計過程本身。

一些實驗室乾脆將這個過程交給機器人，試驗設計、合成和測試無需人工介入、連續運行的自動化實驗室。由於機器人即使人類不整夜監控實驗也能自動重複預定的操作序列，驗證單個候選物質所需的時間也明顯在縮短。

截至 2026 年，經過人工智慧開發的新藥候選管線全球約有 170 條。然而，並非盡是玫瑰色的前景。2020 年，人工智慧從糖尿病治療候選物質堆中篩選出的物質 halicin 作為具有強抗菌效果的新藥候選，引起了媒體和學術界的極大關注。2025 年發表的後續研究再次確認，halicin 可抑制對多種抗生素有耐藥性的 18 種細菌中的 17 種。

然而，halicin 至今仍未能進入人類臨床試驗。這是因為後來才發現的缺陷：在體內吸收程度低、從體內排出過快，且效果隨感染情況而不穩定。計算機螢幕上看起來

完美的結合力在活的生命體面前行不通的情況確實會發生。

自詡領導人工智慧新藥開發的企業也不例外。美國 Recursion Pharmaceuticals 在 2025 年全年叫停了多個自行發現的藥物候選的開發。針對腦血管畸形海綿狀血管瘤的 REC-994 在中期臨床試驗中顯示的療效未達預期，開發遭中止；針對神經纖維瘤病 2 型的 REC-2282 也走上了同樣的道路。

這家公司的軌跡表明，無論計算能力多麼出眾，其結果都不會直接轉化為好藥。人工智慧在新藥開發中能夠縮短的是尋找和優化候選物質耗費的時間，而非人體本身要求的安全性和療效門檻。

這些挫折反覆出現有共同的原因。深度學習模型以已經成功合成並完成臨床試驗的藥物資料作為學習材料，而這些資料在整個化學空間中分佈在一個非常狹窄且偏向的區域。模型越是跳出這個熟悉的區域去描繪從未見過的陌生分子，其預測與現實偏離的風險也就越大。

研究人員將此稱為域轉移，這也是為什麼將計算機預測的結合力和安全性通過實驗和臨床再次驗證的過程無法跳過的原因。因此，在新藥開發現場，不只依賴計算機選定的單一候選物，還採用將多個性質不同的候選物同時置於實驗臺上以分散風險的策略。

這種競爭對韓國企業也不陌生。韓美藥品使用自主開發的人工智慧平台 HARP 設計了肥胖症治療候選物質 HM17321。它通過不同於既有食慾抑制劑類藥物的機制發揮作用，既能減少脂肪又能保持肌肉，並已獲得美國食品和藥物管理局核准進入臨床 1 期試驗。Syntekabio 通過自主平台已掌握多個重新設計的生物相似藥候選物質，旨在超越廣泛使用的單克隆抗體藥物 Keytruda 的性能，Standam 則與韓美藥品合作開展藥物候選物的發現。

Isomorphic Labs 是 Google DeepMind 的分支，也在 2026 年 2 月推出了以 AlphaFold 3 為基礎的設計引擎 IsoDDE，宣稱具備從零開始繪製抗體的功能，並表示首次人類臨床試驗將在 2026 年內啟動。該公司在 2025 年全年還融資超過 6 億美元。這款引擎在新開發的結合力預測基準 Runs N' Poses 上展現了超過 AlphaFold 3 兩倍的準確性，還在一種叫 cerebellon 的蛋白質上預先僅用序列資訊發現了迄今未知的結合位點，隨後這一預測也被實際後續研究所驗證。

從零開始設計新藥候選的競爭如今已不再只是少數領先企業的故事，而成為全球各地同時展開的趨勢。

製造的手正變得越來越快。如果說 2019 年 GENTRL 在 21 天內找出 DDR1 抑制劑候選藥物讓人感到驚訝，那麼 2026 年的 ISM001-055 則證明了從那指尖產生的物質，已經可以推進到臨床三期。然而，瓶頸並沒有消失，只是轉移到了其他地方而已。如果說以前光是想像出有用的分子本身就很困難，那麼現在，確認想像出來的分子是否能實際製造出來，以及該分子在活體內是否真的安全有效，就成了遺留下來的瓶頸。隨著表現、生成與擴充這三支柱的完備，描繪新藥候選藥物的速度已經超越了人類的雙手，剩下的挑戰，就是將那幅藍圖放在現實的化學與生物學面前，讓它能堅持到最後。

就像 2022 年在上海實驗室開始的一張藍圖，在 2026 年跨進了臨床三期的門檻一樣，在未來的幾年內，我們很有可能迎接從一開始就由電腦描繪出的藥物，被放上藥局貨架的時刻。原本依賴化學家直覺與指尖感覺的新藥設計，其重心正這樣轉移到由數據與演算法共同塑造的工作上，而這個變化，也正與重新編排整個製藥產業研究方式的趨勢相互接軌。

雖然描繪候選物質的手變快了，但是確認那幅藍圖的手卻還沒有那麼快。2026 年出爐的一份產業分析指出，新藥開發的瓶頸已經從尋找候選物質的階段，轉移到了實驗室驗證的階段。美國的 Eli Lilly 從 2020 年開始，就與一家名為 Strateos 的公司合作，將設計、合成、純化與分析結合在一起，共同營運幾乎不需要經過人手就能運轉的機器人實驗室。Hanguk Bogeon Saneop Jinheungwon 也將具備無需人手、由機器人與人工智慧連續進行設計、合成與實驗的自主實驗室，列為未來必須解決的課題。

這個趨勢背後，有一個鮮為人知的隱情。為了讓人工智慧減少反覆試錯，不僅需要讓它學習什麼是行得通的，還必須一起學習什麼是因為什麼原因而行不通的，但失敗的化合物紀錄，卻很少留在論文或公開的資料庫中。企業們會把失敗的候選藥物靜靜地收進抽屜裡，不會向世界告知那次失敗的理由。只收集成功案例來學習的人工智慧，即使失敗近在眼前，也經常無法事先察覺到跡象。這種資料缺乏的問題，即使自主實驗室再怎麼增加，似乎也很難輕易彌補。

生成模型不會只製造能緊緊附著在標靶上的分子。在製造避開已註冊專利申請專利範圍的分子方面，它也變得越來越熟練。最近甚至出現了將這個目標，直接植入強化學習分數計算本身的方法。這種方式是對於與現有專利化合物結構相近的分子給予扣分，而對於找出尚未有人主張權利之空白處的分子給予獎勵。使用這個工具的一方，並不僅限於最初製造出新藥的公司。製造學名藥的公司們，也正在利用相同的技術，設計出避開專利的相似物質，從而提前引發針對原廠公司專利的爭端。美國法院在 2022 年判決，人工智慧不能成為專利發明人。

然而，當描繪分子的一方是人工智慧，而人類只是負責修飾合成路徑時，那個發明人欄位裡到底該填上誰的名字，目前還沒有明確的答案。專利制度是一種約定，以向世界公開新發明為代價，賦予一定期間的獨佔權，但如果機器尋找申請專利範圍漏洞的速度，比人類修改法律的速度還要快上許多，那麼這個約定的份量，將很有可能會變得越來越輕。即使這有著提早帶來便宜學名藥的效果，但保護發明以收回研究費用的專利本意，與這股趨勢相互衝突的問題，在未來似乎也會持續引發喧鬧。

2.3

in silico ADMET 及毒性預測系統

1990年代美國的藥局裡，以Seldane為商品名販售的抗組織胺藥物Terfenadine，被廣泛處方給患有季節性過敏的人們。雖然它在消除打噴嚏和流鼻水的效果上非常出色，但幾位服用了這種藥物的患者，在同時服用特定抗生素、抗真菌藥，甚至葡萄柚汁時，發生了心跳不規律而倒下的事故。其原因在於心臟細胞的鉀離子通道（通常稱為hERG通道）被藥物阻塞，導致心臟跳動的電氣訊號紊亂，從而引發了心律不整。

在開發當時，誰也沒有預料到這種藥物無法在肝臟中被分解，反而會堆積在血液裡威脅心臟。最終，這種藥物在市場上悄悄銷聲匿跡，而製藥業界為了在臨床試驗前，事先找出新藥在人體內會如何運作、在哪裡會產生危險的方法，進行了長期的探索。2020年代後半的人工智慧，正是針對這個問題的解答，它已成長為一種 in silico 預測系統，能在候選物質進入人體之前，事先在電腦中展現出它們的命運。

一種新藥從實驗室的化合物到擺上藥局的陳列架，平均需要花費超過10年的時間與數千億韓元的資金。然而，進入臨床試驗的候選物質中，能夠存活到最終批准的比例還不到十分之一。根據一項分析2010年至2017年臨床試驗失敗案例的研究顯示，失敗原因中有40到50%是因為藥效不如預期，將近30%是因為身體無法承受的毒性，而有10到15%則是因為藥物在體內吸收和消失的方式本身設計錯誤。將毒性與藥物動力學問題加總起來，將近一半的臨床失敗都發生在吸收、分布、代謝、排泄、毒性這五個字，也就是ADMET領域之中。

美國國家衛生院表示，新藥研發費用的22%都花在確認與解決這個ADMET問題上。在製藥業界，新藥開發費用隨著時間推移反而不斷增加的矛盾趨勢，被稱為Eroom's Law，這是將半導體性能持續提升的摩爾定律反過來拼寫而成的。如果候選物質的毒性要等到動物實驗和以人為對象的臨床試驗才顯露出來，那麼這段期間投入的時間和金錢將會白白流失。

長久以來，製藥公司為了事先過濾掉這種風險，一直依賴著將候選物質餵給老鼠和猴子等動物並觀察其反應的方法。問題在於，動物的身體與人類的身體在基因和生理構造上存在著不小的差異。在動物實驗中被確認為安全的藥物，到了以人為對象的臨床試驗時，卻反覆出現引發意想不到副作用的案例，學術界將這種落差稱為轉譯研究的鴻溝，並長期為此感到頭痛。美國食品藥物管理局透過2022年食品藥物管理局現代化法2.0，在法律上廢除了84年來規定新藥批准申請必須包含動物實驗的舊有規定。

2025年4月更進一步提出了發展藍圖，宣布將以抗體治療藥物為起點，在未來的3年到5年內使動物實驗成為例外程序，並以基於人工智慧的毒性預測模型、類器官、器官晶片等新的評估方法來填補其空缺。旨在將這項法律完全納入實際審查規定的食品藥物管理局現代化法3.0法案，也分別於2025年12月和2026年7月通過了參議院與眾議院，截至2026年8月，只剩下參議院的重新表決以及交由總統簽署的程序。

在進入臨床試驗這道昂貴的關卡之前，事先過濾掉有問題的候選物質，也就是所謂的早期過濾（Early Filtering），接手了過去實驗室玻璃瓶裡動物所做的工作，成為了新藥開發的第一道關卡。

ADMET這五個英文字母，本身就是一個關於藥物在人體內經歷的旅行故事。第一個字母A代表吸收，是指藥丸經過胃和腸道，滲入血液的過程。D代表分布，是藥物乘著血液在全身各處遊走，然後停留於肝臟、大腦或肌肉等地方的過程。M代表代謝，意思是肝臟就像工廠的分解生產線一樣，將藥物切碎並轉化為其他物質的過程。

E代表排泄，是使用後剩餘的藥物殘渣經過腎臟，以尿液或糞便的形式排出體外的過程。最後的T代表毒性，指的是在這整趟旅行途中，藥物是否會傷害肝臟、心臟或腎臟等器官。只要在這五個階段中的任何一個環節出了問題，無論效果再好的候選物質，最終都無法誕生為藥物。

僅憑藉一個分子的結構，就能事先猜測其在體內將如何運作的學問，被稱為定量結構-活性關係，簡稱QSAR。這是一門從分子的形狀與性質之間存在一定規則的想法出發的理論。早期的QSAR就像Lipinski的五法則一樣，只停留在考量分子量或脂溶性等幾項數字來衡量其是否具備藥物特性的統計公式水準。然而，今日作為新藥候選而必須檢討的化合物化學空間，已經廣闊到數百萬、數千萬個之多，這個龐大的

空間已經無法僅靠老舊的線性公式來處理。因此，研究人員開始將把分子視為由原子的點與化學鍵的線所組成的網路的圖神經網路（簡稱GNN），以及將分子寫成名為SMILES的字串並解讀其語法的大型語言模型與Transformer架構，結合到QSAR之中。

其結果所產生的ADMET-AI，能在短短幾個小時內瀏覽數百萬個化合物的吸收與毒性；而Deep-PK則利用基於圖的分子特徵，一次產出多達七十三種的藥物動力學、毒性與物理化學指標。這些模型都是以ChEMBL、DrugBank、Tox21等任何人都可以下載的龐大公共資料庫為學習材料而成長起來的。

國際指引 ICH M7 判斷混入藥品中的不純物是否有破壞基因並引發癌症的風險，這是將 QSAR 預測值視為實際法規審查資料的代表性案例。這項指引明文規定，必須同時運行兩種原理不同的預測模型（即將專家的化學知識轉化為規則的模型與統計學習模型），若得出不同結論，則交由人類親自判斷。2025年，在面對陌生的化學結構時，靈敏度比現有機器學習技術高出將近 30% 的 AmesNet 登場，將長久以來成為絆腳石的新型骨架化合物之突變誘發性預測提升到了一個新層次。

這些模型學習的基礎，建構在美國環境保護署營運的 ToxCast 與 Tox21 資料庫之上，2025年8月公開的最新版包含了將近一萬種化學物質的試管內反應資訊，被譽為訓練毒性預測人工智慧的代表性公共資產。製藥公司將這些公開數據與自家化合物資料庫一同用於學習，試圖從合成前的設計階段，就過濾掉疑似具有突變誘發性或致癌性的不純物。

除了預測候選物質的性質之外，從一開始就在電腦中重新繪製具備良好性質的分子，這類嘗試也非常活躍。Retro Drug Design 框架朝著同時滿足藥效、生物利用度、代謝穩定性、低毒性等容易相互衝突的六項 ADMET 條件的方向，自動修飾分子結構，而實際透過這種方式設計的激酶抑制劑候選物，在實驗驗證中也展現了很高的成功率。結合生成式人工智慧 REINVENT 與 ADMET-AI 的 ADMETrix，能在重新繪製分子時立即計算出多項 ADMET 指標，讓設計與評估得以即時交替進行。

DrugEx v2 將在多個目標相互衝突時，不放棄任何一方而尋找平衡點的柏拉圖式強化學習導入藥物設計中，挑選出能緊密結合標靶蛋白質同時避開錯誤標靶的分子。就像廚師設計食譜時會同時考量味道、營養與烹調時間一樣，這些模型等於是同時

端出了療效與安全性等多道菜餚。

了解每一個分子的性質，與描繪該分子在實際人體內隨時間如何起伏變化，是不同層次的問題。填補這道鴻溝的技術，就是生理學基礎藥物動力學模型，簡稱 PBPK 模型。這可以簡單理解為，將人體視為由心臟、肝臟、腎臟、肌肉等多個城市組成的地圖，並將血流視為連接這些城市的道路，藉此計算藥物這輛汽車隨著時間推移，會在哪個城市停留多久的一種方式。人工智慧能分析分子的 SMILES 結構或指紋資訊，瞬間萃取出分配係數與水溶度、在肝微粒體中的分解速度、與血漿蛋白質結合的比例等核心數值，只要將這些數值輸入 PBPK 模擬引擎，就能描繪出靜脈注射或口服後，血液中濃度隨時間起伏的曲線。

將化學語言模型或圖形卷積神經網路與 PBPK 結合的驗證研究中，最高血中濃度或其到達時間、曲線下面積、半衰期等指標的預測值，與實際值呈現了判定係數 0.61 到 0.93 的相關性，代表誤差大小的絕對平均折疊誤差也停留在 1.84 到 2.8 之間，展現了不亞於現有動物實驗外推法的精確度。在美國，2020年到2024年間獲得批准的 245 件新藥中，有 65 件（比例為 26.5%）將 PBPK 模型作為審查的核心依據提交，其中用於藥物交互作用評估的案例高達 81.9%，在抗癌藥物領域中的使用更是特別活躍，佔了 42%。

這種精確度不僅用於為首次將新藥投予人體的臨床試驗決定安全的起始劑量，更被用作為病患量身打造的精準醫療之基礎，藉以預先估算肝功能衰退的長者，或是體重與器官大小異於成人的兒童等難以直接參與臨床試驗的族群之藥物暴露量。美國食品藥物管理局順應這股趨勢，已另行制定指引，正式承認 PBPK 模型在藥物交互作用評估中的作用。

心臟毒性預測被視為新藥開發領域必須緊急處理的課題。正如 Terfenadine 的案例所示，如果藥物阻斷了心臟細胞的鉀離子通道（即 hERG 通道），心跳的電氣訊號就會變得紊亂，進而引發危及生命的心律不整。幫助腸胃蠕動的 Cisapride 也因為類似原因退出市場，這個案例顯示了與 hERG 相關的心律不整並非僅限於單一抗組織胺藥物類別，而是會跨越不同治療領域反覆出現的風險。Mamoshina 研究團隊建立的基於人工智慧之定量連續評分系統，透過學習藥物銀行與醫學字典資料庫中龐大的化合物資訊，在驗證資料中以曲線下面積 79% 的水準，預先篩選出了潛在的心臟毒性。

2025年，同時結合多項特徵的圖形神經網路模型 hERG-MFFGNN 以能夠說明判斷依據的形式登場，而將值得信賴的心臟毒性數據進行標準化彙整的資料集 DICTrank 也一併發布。同年，結合平衡學習數據與多樣變數的 XGBoost 共識模型獲得了靈敏度 0.83、特異度 0.90 的成績，隨著導入大型語言模型嵌入技術的 hERG-LTN 問世，預測方式正逐年變得更加多樣化。

在一篇彙整並探討 25 件相關研究的論文中，確認了有多個模型在外部驗證中的曲線下面積超過 0.70，或者準確度高於 0.75，獲得了已大步邁向可於實際現場使用水準的評價。

肝臟毒性，特別是藥物引起的肝臟損傷，即DILI，其發病原因並不是單一的，因此預測格外困難。由於每個人肝臟酶的數量和種類不同，即使是相同的藥物，對某些人可能完全無害，對另一些人卻可能會損傷肝臟，反應因人而異。細胞色素類酶基因中常見的多態性，比如CYP2D6或CYP2C19的活性在每個人身上存在先天差異，這意味著同樣劑量的藥物對某些人可能過量，對另一些人可能劑量不足。

最近，研究人員正在嘗試將這些遺傳型信息與化合物結構信息一起進行學習，以便識別出在特定遺傳型群體中才會顯著出現的肝臟毒性風險。最近的人工智慧系統不僅考慮藥物的物理化學特性，還整合了包括基因、蛋白質、代謝產物在內的多組學數據，以及脫靶相互作用網絡，能夠以高達89%的準確度識別出藥物引起的肝臟損傷。

2026年初，一項整合了ChemBERTa、RoBERTa、SMILES2Vec等不同語言模型產出信息的多模態深度學習方法在學術期刊上發表，相比於過去只檢視分子結構的方法，現在已經可以進行更加精細的預測。也發生過象徵性事件。2025年，美國食品藥物管理局接納了Absentia Labs公司提交的人工智慧開發的數字肝臟模型，作為ISTAND計劃（該計劃審查創新性新藥開發方法）的認證對象，這是人工智慧開發的預測工具首次進入該認證程序。

雖然只是通過了三階段認證程序的第一關，但是虛擬肝臟模型開始代替真實人類肝臟填充新藥審查文件的部分內容，這無疑是一個意義重大的事件。

藥物原本不會只溫順地附著在它所針對的單一目標蛋白質上。就像一把鑰匙只能歪斜地塞入多把鎖孔一樣，新藥候選物質也會意外地結合到原本的目標以外的其他受

體或酶上，從而引發意料之外的副作用，這種現象被稱為脫靶效應。Liu研究小組使用能夠同時處理多個預測任務的多任務圖神經網絡，提出了精確捕捉化合物脫靶相互作用的方法，並通過該方法展示了可以在候選物質優化階段提前篩除藥物引起的腎臟損傷等嚴重風險。

深入探究結構與毒性之間關係的ProTox系列網絡服務器也在快速增長。最新版本ProTox 3.0結合分子相似性和機器學習，能夠計算包括半數致死量、致癌性、致突變性和免疫毒性在內的61種毒性指標。在進行成本高昂的動物毒性測試之前，僅用這一個免費網絡工具就能夠將存在嚴重缺陷的化合物從候選物質中篩除。

隨著高齡人口增加和慢性病患者增多，一個人同時服用多種藥物的情況變得越來越普遍，因此預先發現藥物之間在體內發生的相互作用（即DDI）的工作已經成為關係到患者生命安全的問題。就像特非那定與抗生素或葡萄柚汁相遇時那樣，一種藥物可能會阻止或促進另一種藥物代謝的肝酶，導致第二種藥物的濃度比預期高得多或低得多。過去，這類相互作用只能通過逐一搜索現有文獻的方式進行確認，但現在人工智慧已經可以同時讀取大量醫學論文、電子健康記錄和化合物的三維結構，代替人工完成這項工作。

BERT-DDI等大型語言模型能夠在數百萬份臨床記錄和不良反應報告中發現藥物間的風險信號，並在句子層次進行識別。DTiGEMS+的想法是通過圖表繪製藥物與靶蛋白質的關係，在關係網絡中推斷新的結合方式，這種思想進而發展為SumGNN等後續知識圖神經網絡，已經達到了能夠提前識別臨床上尚未報告過的新型藥物組合風險的水平。通過這種方式獲得的預測結果已經連接到臨床決策支持系統，該系統在醫生開處方箋的瞬間就會在屏幕上顯示警告，因此不再只是計算機內部的計算，而是正在逐步成為診療室中實際起作用的安全措施。

藥物上市後，人工智慧也在持續監測電子健康記錄和保險理賠數據等實際臨床數據，並在上市後監測工作中識別在臨床試驗階段未曾發現的藥物組合風險信號，這項工作也在同步進行。

輸入分子的SMILES字符串或三維結構，以及熔點或脂溶性等物理化學特性後，系統會輸出ADMET特性的預測值，同時也會輸出應用範圍（Applicability Domain），用以說明這項預測可信度如何。應用範圍是一個指標，表明輸入的分子與模型學習

過的化合物之間的相似程度如何，經濟合作與發展組織在驗證QSAR模型時也建議必須明確說明這個範圍。就像進入導航地圖上不存在的陌生道路時，導航的指引會突然開始出錯一樣，當遇到與學習數據相差很大的新化學結構時，即使是再精妙的模型，預測也可能會大幅偏離。

目前廣泛使用的混合平台之一 ADMETlab 3.0，從 ChEMBL、PubChem 和 OCHEM 等公開資料庫中蒐集了超過 40 萬筆化合物資訊，並以 SMILES 作為標準輸入值，一次計算出 119 種性質。由於其結構將同時處理多項預測任務的多任務定向訊息傳遞神經網路與傳統分子描述符並行結合，因此獲得了兼顧速度與準確度的評價。然而，2026 年發表的一篇批判性回顧論文指出，在代表性基準資料庫 Therapeutics Data Commons (TDC) 的 ADMET 排行榜上，於 22 項標準評估指標中能成功重現的頂尖模型屈指可數，有些模型甚至存在未公開程式碼，或是訓練資料與測試資料混雜的缺失。這留下了我們在全盤相信華麗的效能數據前，必須再進一步深究的理由。

預測與驗證相互交替的循環也變得越來越緊密。人工智慧挑選出有潛力的候選物質後，機器人實驗室隨即進行合成，並自動完成試管試驗，然後將結果回饋給預測模型。這種所謂的設計、製造、測試與分析的循環結構，正逐漸普及至多個研究機構。有別於過去需要人工逐一訂購化合物並安排實驗日程的方式，這種循環一天可以運行好幾次，持續不斷地完善預測模型本身。

隨著 ADMET 預測變得更加精密，這個循環的速度也隨之加快，而這樣累積起來的新實驗資料，反過來又將下一代的預測模型鍛造得更加完善。

即使是看起來如此快速且精密的人工智慧預測，要進入臨床與法規審查現場，仍有很高的門檻需要跨越。在內部測試中決定係數高達 0.8 左右的優秀模型，一旦遇到現有資料庫中不存在的新骨架化合物，其外部驗證效能暴跌至 0.3 左右的情況並不罕見。目前累積的毒性與藥物動力學資料，大多偏向於容易合成且成功商業化的特定藥物群，或是偏向西方人口群體，這也是一大阻礙。利用這些資料訓練出來的模型，在評估少數族裔特有的代謝特徵，或是細胞與基因療法等新興治療方式時，面臨著給出錯誤答案的風險。如果特定人口群體中顯著的代謝酵素變異未能充分反映在訓練資料中，模型就可能遺漏那些只在該群體中特別嚴重的副作用。

許多交織著數千萬個參數的深度學習模型，往往處於無法用人類聽得懂的語言，來解釋為何將特定化合物指認為肝毒性物質，或是哪種結構特徵會引發心律不整的黑盒子狀態。在攸關病患性命的審查過程中，這種不透明性很難被直接接受。因此，人工智慧給出的預測值，必須經過類器官或器官晶片等實體的生物學系統再次確認，才算是發揮了應有的作用並獲得認可。

美國食品藥物管理局在 2025 年 1 月發布了一份草案指南，探討在藥品與生物製劑的法規判斷中應如何使用人工智慧，並提出了一套包含七個步驟的程序，根據模型使用的情況劃分風險等級以驗證其可靠性。這項指南廣泛涵蓋了人工智慧介入的非臨床、臨床、上市後監視以及製造階段，但也明確界定，不會對不影響安全性或品質的業務效率化進行管制。歐洲藥品管理局於 2024 年 9 月發表了一份關於在藥品整個生命週期中應用人工智慧的反思文件，將人工智慧納入現有的優良臨床試驗規範與品質管理原則中，並強烈建議確保可追溯性與透明度。

大韓民國食品醫藥品安全處也從 2026 年開始，逐步將人工智慧引進許可資料摘要、專業翻譯與審查報告草案撰寫等工作中，並以此為跳板，設立了在 240 天內完成新藥許可審查的目標；同時宣布，未來 5 年內將培育 1,100 名專注於人工智慧新藥與生技領域的法規科學專業人才。各國法規機關意識到，如果任由各國採用不同的人工智慧審查標準，跨國共同進行的臨床試驗資料可能會在不同國家受到不同的評價。基於這種問題意識，各國也開始透過國際醫藥法規協和會 (ICH) 等協商機構，展開了制定共通標準的討論。

從電腦給出的一個預測值，到最終成為實際的處方箋，如今這些規定與程序正與技術一樣，左右著發展的速度。

再怎麼出色的模型，如果缺乏供其訓練的資料，也無法發揮作用。然而，醫院與製藥公司將辛苦收集來的臨床紀錄與化合物資料視為核心營業機密，不輕易對外公開。再加上美國的健康保險可攜性與責任法或歐洲的一般資料保護規則等保護病患資訊的法規層層疊加，導致資料被困在各個組織內部並散落各處，形成了所謂的資料孤島現象。在不打破這堵高牆的情況下，能夠匯集並使用不同機構資料的聯邦學習正備受矚目。

其運作方式是將原始資料原封不動地保留在醫院或公司的安全伺服器內，僅將訓練過該資料的人工智慧模型權重進行加密，傳送至中央伺服器進行整合。在大韓民國，自2024年4月至2028年12月，科學技術情報通信部與保健福祉部共同投入了348億韓元推動K-MELLODDY計畫。結合超過20家國內機構擁有的實驗室資料與臨床資料，透過聯邦學習來打造同時預測ADMET與藥物動力學的模型，這項工作從2025年起已擴大為正式的新型計畫。

歐洲多家製藥公司共同開展的MELLODDY項目首先為其奠定了基礎，兩個項目都證明了可以在不洩露患者信息和侵犯商業機密的情況下共同開發預測模型。國內一家製藥公司在與英西里科醫藥（Insilico Medicine）合作僅四個月後，便發現了同時具備吸收、分佈、代謝、排泄、毒性五項特性的新藥候選物質，這項工作在過去通常需要花費數年。

這種趨勢不再局限於ADMET預測的狹隘領域，而是擴展到融合多組學分析、類器官和器官芯片等物理驗證工具，以及實際臨床現場的電子病歷的更大融合生態系統。當針對單一新藥候選物質建立起電腦預測、實驗室驗證和醫院臨床記錄之間相互交換數據的結構時，ADMET預測的準確性將不再取決於某項技術的進步，而是取決於整個生態系統共同深化的速度。

隨著國內政府資助的K-MELLODDY和民間企業協作案例不斷積累，這種融合正在從少數大型製藥企業的專有工具轉變為多個機構共同使用的公共基礎設施。ADMET預測技術的下一個競爭點不在於哪家公司擁有更精巧的算法，而在於誰能更密集地編織不同類型的數據和驗證手段。

此刻世界各地的服務器上正在對數百萬個候選物質的ADMET特性進行通夜計算，其中那些風險信號明確的化合物在與任何動物或患者接觸之前就悄然被從候選名單中刪除。在特非那定還在藥房貨架上的年代，這樣的景象是難以想像的。當然，我們不能保證當今的人工智能能完美過濾所有風險。面對未曾學習過的新型分子結構時，它仍然會動搖，有時也無法解釋自己的判斷原因，最終還需經過類器官和器官芯片等物理驗證才能被認定為可信的數值。

即使如此，臨床試驗前的危險候選物質篩選網也在逐年變得更加密集，編織這張網的每一根纖維都包含著QSAR和PBPK、圖神經網絡和大型語言模型，以及跨越國界

的聯邦學習。新藥候選物質進入人體前必須先通過電腦模擬這道關卡的程序正在從若干實驗室的試點項目演變為新藥開發的標準階段。

2025年9月，基因編輯療法公司英特利亞治療公司（Intellia Therapeutics）開發的nex-z（nexigran diclumemuran）在一名患者身上導致了4級、即危及生命的肝臟指標升高。該患者在醫院接受了治療。這種藥物針對轉甲狀腺素澱粉樣病變（心臟和神經中澱粉樣蛋白堆積的遺傳性疾病），通過直接進入體內改寫基因而起效，是3期臨床試驗MAGNITUDE和MAGNITUDE-2的研究對象。美國食品藥品監督管理局在10月29日對兩項臨床試驗全部發佈了臨床試驗暫停令，11月又發生了一名80多歲男性患者用藥後死亡的事件。

調查發現這起死亡是由感染引起而非肝臟損傷，但整個後期臨床試驗瞬間停止的事實依然存在。這種療法的特點不是基於化合物的化學結構，而是基於基因編輯工具本身在體內的工作方式。那些學習過SMILES字符串和分子指紋的既有預測模型根本無法識別這一目標。兩項臨床試驗在增加更頻繁的肝臟檢測和預先篩選高風險患者等安全措施後，從2026年開始相繼恢復。

隨著細胞和基因療法等全新治療方式的增加，基於分子結構的預測模型的適用範圍似乎反而在縮小，呈現出一種悖論。AI在預測化合物毒性時，基於數十年的小分子數據。但是，能夠捕捉基因編輯和細胞療法在體內引發的免疫反應和脫靶編輯的數據還沒有在世界任何地方充分積累。在相當一段時期內，無論預測模型多麼精巧，都可能無法替代臨床試驗現場直接監測患者並及時捕捉異常信號的監督系統。

韓國也開始了擴大預測範圍的新舉措。食品藥品安全部在2025年12月3日與首爾國立大學數據科學研究生院簽署協議，決定共同構建一個人工智能平台，將國內醫療機構的臨床數據與多組學信息相結合，繪製出藥物反應在不同患者身上的差異。與在實驗室階段預先篩選新藥候選物質的K-MELLODDY不同，這次協議是試圖將藥物進入患者體內後發生的事情也轉化為數據。

兩個機構不僅進行聯合研究，還決定共同舉辦管理數據品質和培養分析人員的工作坊。隨著這兩項事業的不斷推進，預測模型涵蓋的時間跨度也在不斷拓寬。

第 3 部

次世代模式與智慧藥物傳輸系統

3.1

生物藥物與多重標靶治療藥物設計 (ADC)

2026年6月22日，美國加州聖地亞哥會議中心開幕四天期間，擠滿了來自全球製藥和生物科技企業的工作人員。在世界最大的生物產業博覽會 -- 生物國際大會（通常簡稱BIO USA）上，韓國企業賽爾特里翁搭建了展台，在四天內完成了180場以上的商業會議。光是前往展台參觀的訪客就超過2000人。賽爾特里翁展出的不是過往主力推進的生物相似藥 -- 即已過專利期的生物製藥的仿製藥 -- 而是以人工智能為基礎的新藥開發、抗體-藥物偶聯體，以及同時針對多個靶點的多重抗體技術。

這證明了製藥產業的重心正在從用電腦設計單個小分子化合物的時代，向用人工智能重新完整繪製龐大蛋白質分子的時代轉移。

僅在幾年前，設計抗體或細胞治療藥物等體積較大的醫藥品仍是一項依賴研究人員經驗和反覆實驗的艱苦工作。研究人員需要製造一個候選抗體，將其放入實驗皿中觀察反應，失敗後就得從頭開始，這樣的試錯需要重複數百次、甚至數千次。在這個過程中，往往需要花費十多年的時間和接近兆級的成本。但隨著能在電腦螢幕上精確到原子層級預先設計和測試蛋白質的胺基酸序列和立體結構的技術發展，情況發生了改變。

大型語言模型和蛋白質語言模型，以及能同時讀取多層生物數據的多組學分析，處於這一轉變的中心。曾經需要在實驗皿上花費數年的設計工作，如今在許多情況下可以在伺服器內在幾天內完成。

抗體是我們體內免疫系統產生的Y字形蛋白質。它能夠識別進入體內的細菌或病毒等異物，並精確貼附其上做出標記。就像鑰匙適配鎖一樣，每個抗體都經過設計使其只與特定的靶點結合。我們的體內能夠產生數百萬種不同的抗體，其中特異性識別某一特定靶點的抗體被大量複製後用作藥物，這就是單克隆抗體（mAb）。從類風濕性關節炎治療藥到多種癌症治療藥，目前醫院使用的大多數抗體藥物都是通過這種方式製造的。

進一步發展，不僅一個抗體結合一個靶點，而是設計能同時結合兩個或以上不同靶點的抗體也相繼出現。這類抗體被稱為多重抗體或雙特異性抗體。它的一隻臂膀抓住癌細胞表面的特定蛋白質，另一隻臂膀抓住在體內巡邏的T細胞，就像中介人一樣用兩隻手分別握住兩個人的手使他們互相面對，從而將T細胞拉到癌細胞的身邊。

賽爾特里翁在2026年BIO USA大會上新推出的技術正是一款用於設計這類多重抗體並預先評估開發可行性的人工智能工具。隨著靶點從一個增加到兩個、三個，抗體需要識別的組合情況呈指數級增長，這使得研究人員手工逐一篩選候選方案的傳統方式變得難以應對，而人工智能可以代替人工搜索這片設計空間。

抗體-藥物偶聯體（ADC）是在抗體上掛載強力有毒物質的精準制導藥物。用快遞包裹來類比便於理解。包裹內裝著足以殺死癌細胞的強力化學物質，即有效負載。如果把包裹隨意扔到任何地方，後果不堪設想，因此需要一位知道確切地址的快遞員。扮演快遞員角色的正是抗體。

它識別只存在於癌細胞表面的特定蛋白質，並只將包裹送到該處。連接抗體與有效負載的鏈接器則確保包裹在運送過程中始終密封，只有到達正確目標時才打開。抗體、鏈接器和有效負載這三個組件協調配合，才能最終形成一枚在癌細胞內爆炸的精準炸彈。

既往的抗癌化學治療以截然不同的方式運作。它將強力藥物送入全身血管，無差別地攻擊快速分裂的細胞，在這個過程中不可避免地導致脫髮、腸粘膜潰爛和白血球數量下降等副作用。這就像地毯式轟炸。相比之下，抗體-藥物偶聯體則更像精確坐標制導的導彈。

即使使用相同的有毒物質，由於只作用於癌細胞而不觸及正常細胞，因此能夠以遠更低的劑量產生遠更強的效果。將這種精準性真正付諸實施並非易事，長期以來這對抗體-藥物偶聯體的開發造成了瓶頸。

即使在人工智能設計流程大規模啟動之前，抗體-藥物偶聯體就已經在臨床上證明瞭其療效。諸如用於乳癌和胃癌治療的恩喜圖、用於乳癌和膀胱癌治療的特羅德爾維等藥物，為那些既往癌症治療藥物已無計可施的患者開闢了新的選擇，為整個領域建立了信任。人工智能正在發揮作用，將這一已驗證的模式推進到下一代，使其更快速、更精準。

正是在這裡，人工智能正在改變抗體-藥物偶聯體設計的格局。研究人員開始採用強化學習方法，不是分別調整抗體的結合位點、鏈接器的化學結構和有效負載的毒性強度，而是同時運轉這三個要素。人工智能模型提出一個候選組合後，電腦會虛擬計算該組合在血液中的穩定性如何、到達癌細胞時斷裂的精準度如何、向正常組織滲漏的程度如何。

會根據該結果進行評分，如果分數較低，就會改變結構並重新計算，這個過程會重複數萬次。這樣一來，就形成了一個循環，同時追求提高抗體的標靶結合力以及降低全身毒性這兩個目標。這等於是一個抗體、連接子和負載彌補彼此弱點，共同進化的結構。

抗體、連接子和負載所創造的這個循環，實際上可以稱之為一個多重標靶最佳化迴圈。如果更具體地檢視這個循環結構，給予人工智慧的報酬訊號大致可分為兩種。一種是顯示抗體能多強烈、多持久地附著在癌細胞表面抗原的結合指標，另一種則是顯示負載在正常細胞中有多安靜、僅在癌細胞內有多劇烈運作的選擇性毒性指標。

要找出同時提高這兩個指標的組合，是一個變數太多、無法由人類手動計算的多維最佳化問題。人工智慧會同時測試無數的抗體、連接子和負載組合，找出同時滿足這兩個指標的組合，並且只將結果產生的候選物質交給實際的實驗室進行驗證。化學家和免疫學家各自獨立工作的傳統方式很難產生的組合，在這個過程中被發現了許多次。

這個虛擬計算也同時使用了稱為定量系統藥理學的技術。這是一種透過數學模型，預先描繪藥物進入體內後被吸收、擴散和分解的整個過程的方法。如果預先使用這個模型過濾負載殺死細胞的性能以及連接子的穩定性，就能在進入實際動物實驗和臨床試驗之前，篩選出許多會失敗的組合。試圖收購具備這種設計能力的公司的動向，在 2025 年與 2026 年讓製藥界非常熱絡。Gilead Sciences 在 2026 年 4 月決定以 31 億 5 千萬美元的簽約金，加上最多 18 億 5 千萬美元的各階段里程碑款項，總計 50 億美元的規模，收購德國的抗體藥物複合體專業公司 Tubulis。

Eli Lilly 也收購了抗體藥物複合體新創公司 Crossbridge Bio，加入了相同的潮流。這兩筆交易都不是為了公司擁有的特定抗癌藥物候選物質，而是針對設計連接子、負載和結合化學的人工智慧平台本身所進行的收購。

在這個市場中，中國生技公司的存在感是壓倒性的。甚至有評價指出，在全世界的抗體藥物複合體技術出口合約中，來自中國的候選物質在九十件中就占了九件。在韓國公司中，LigaChem Bio（舊名為 LegoChem Biosciences）正處於領先地位。該公司以獨家開發的连接子和負載核心技術為基礎，向 Amgen 出口了約 1 兆 6 千億韓元、向 Janssen 出口了約 2 兆 2 千億韓元，並向日本 Ono Pharmaceutical 出口了超過 9 千 4 百億韓元規模的技術，僅公開的合約加起來就超過了 9 兆 3 千億韓元。

有統計指出，如果加上未公開的合約，累積的技術出口規模將超過 10 兆韓元。LigaChem Bio 與抗體開發商 NextCure 共同開發中的 B7-H4 標靶抗體藥物複合體 LNCB74，已向美國食品藥物管理局申請在全球第一期臨床試驗中增加高劑量投藥組，正邁向了下一個階段。市場調查機構 EvaluatePharma 預測，抗體藥物複合體市場的整體規模將從 2025 年的 150 億美元左右，成長至 2032 年的 570 億美元。

在設計抗體本身的階段，人工智慧也已經進入實戰。基於深度學習的蛋白質語言模型在學習了龐大的抗體序列資料後，會從頭開始重新創造出能強烈附著於特定標靶的新抗體序列。美國的人工智慧新藥開發公司 Absci 將透過生成式人工智慧平台設計的抗體 ABS-101 作為發炎性腸道疾病治療藥物的候選物質，推進到了第一期臨床試驗。

這種被設計用來附著在名為 TL1A 的標靶蛋白質上的抗體，是由人工智慧完整描繪出來後，首次實際投藥於人體的生物藥品，因此受到了業界的矚目。它被設計成每四個月只需在皮下注射一次即可，並且也將重點放在降低免疫原性的風險上。預計在 2025 年下半年將會公布初期的安全性、藥物動力學和免疫原性資料。

專門從事抗體設計的人工智慧新創公司也湧入了龐大的投資資金。獲得 OpenAI 投資的美國新創公司 Chai Discovery，在 2025 年 12 月吸引了 1 億 3 千萬美元規模的系列 B 投資，企業價值被認可為 13 億美元，據悉之後正進行協商，計畫將企業價值提升至 34 億美元的水準，以額外籌集 4 億美元。該公司推出的 Chai-3 模型與 Pfizer 簽訂了技術合約，並在 2026 年 1 月也宣布了與 Eli Lilly 的合作夥伴關係。

從整體業界來看，從 2025 年 6 月到 2026 年 5 月的一年內，光是利用人工智慧設計包含抗體在內的生物製劑的企業，就透過二十多筆的交易，流入了超過 9 億 8 千萬

美元的投資資金。在 2026 年第一季，Earendil 吸引了 7 億 8 千 7 百萬美元，EmendoBio 吸引了 8 千萬美元等，投資熱潮依然沒有減退。

描繪抗體立體結構的人工智慧模型，也顯著地提升了設計速度。從只要輸入蛋白質的胺基酸序列就能預測立體結構的現有技術更進一步，現在已經能夠進行逆向設計，也就是反過來描繪出完全符合所需標靶形狀的結合部位立體結構，並填入與之相符的胺基酸序列。這並不是從身上已經有的衣服中挑選一件來穿，而是先測量客人的尺寸，然後依此裁剪布料，這與訂製西裝店的方式相似。

這種基於結構的設計技術，不僅用於抗體，也同樣被用於製造奈米抗體、胜肽和 PROTAC 的雙重鉤子結構，正逐漸成為貫穿多種次世代模態的共同基礎技術。

體積比抗體小得多的奈米抗體，也是大幅受惠於人工智慧的領域。這種僅在駱駝或羊駝等動物體內發現的微型抗體，因為體積小，很容易深入堅硬的腫瘤組織深處，也很適合利用微生物進行大量生產。最近有一個研究團隊，捨棄了人類，組成了一個由多個人工智慧代理程式組成的虛擬研究團隊，讓它們在電腦中從頭開始設計能使新型冠狀病毒失去作用的奈米抗體序列。

人工智慧代理程式們經歷了分析標靶蛋白質結構、提出結合候選物，以及互相審查設計的過程後，成功在實際實驗室中合成了這個序列並完成驗證。這被視為一個指標性的案例，顯示只要人類研究人員決定標靶並指出方向，剩下的大部分設計工作，人工智慧團隊都能幾乎自主地完成。

在設計胜肽，也就是由數十個胺基酸短暫連接而成的小分子的領域中，也能看到類似的趨勢。有一款名為PepMLM的人工智慧工具，借用了名為遮蔽語言模型的自然語言處理技術，即使完全沒有標靶蛋白質的立體結構資訊，僅憑胺基酸序列資訊，就能設計出能與標靶特異性結合的線性胜肽。一直以來，胜肽設計是一項必須精確了解蛋白質摺疊結構才能開始的工作，但現在確認了即使沒有結構資訊也能挑選出結合力強的候選物，這使得可設計的標靶範圍本身正在逐漸擴大。

在製作抗體或抗體藥物複合體等體積較大的分子會造成負擔的情況下，胜肽作為一種更輕、製造成本更低的替代方案，正受到廣泛關注。

抗體或CAR-T細胞要能認出標靶，就必須有一段像把手一樣能緊緊貼合的序列區段，這被稱為表位。這就像在寬廣的牆壁上，必須準確找到門把才能開門一樣。問題在於，我們身體的免疫系統，經常會將新進入的治療用抗體或細胞本身誤認為是陌生的入侵者，進而發動攻擊。

這被稱為免疫原性，如果這種副作用很嚴重，藥效就會消失，或者會引發危險的過敏反應。人工智慧會預先掃描抗體序列，達到人類肉眼難以辨識的程度，找出會刺激免疫系統的序列區段，並引導將該區段替換為人體已經熟悉的序列。

實際執行這種免疫原性預測的人工智慧工具，也正快速變得更加精密。有一套名為SITA的預測系統，會同時計算抗體整體序列與個別胺基酸部位的免疫原性分數，能清楚區分源自人類的抗體與非源自人類的抗體。一款名為ImmunoStruct的多模態深度學習模型，學習了多達兩萬六千多個胜肽與主要組織相容性複合體的結合數據，它將序列資訊、結構資訊與生化特性一次結合起來，提升了預測的準確度。

像NetMHCIIpan、Hu-mAb、BioPhi等工具，已實際應用在將抗體修飾成對人類更熟悉的序列的「人類化」工作上。這使得新藥候選物在進入動物實驗階段的很久之前，光靠序列就能過濾掉危險訊號，從而達到同時縮減開發成本與時間的效果。

在設計癌症疫苗時，表位預測也是核心關卡。即使是同一種癌症，每位患者癌細胞表面出現的突變蛋白質，也就是所謂的「新生抗原」，形狀都不盡相同。研究人員正利用結合了蛋白質語言模型與卷積神經網路的深度學習模型，預測每位患者的腫瘤組織中，有哪些新生抗原容易被免疫細胞識別。從黑色素瘤患者的腫瘤檢體中提取出的數千個突變蛋白質中，要挑選出真正會引發免疫反應的候選物時，就會使用這種模型，而且許多大學與企業正根據這些結果，進行為單一患者量身打造專屬癌症疫苗或免疫抗癌藥物的設計研究。

對於在運送抗原到細胞表面的第一型主要組織相容性複合體上所呈現的CD8陽性T細胞標靶表位，其預測準確度在最近幾年內也有了顯著的提升。

像這樣找出標靶表位並預先過濾免疫原性的人工智慧流程，並不侷限於抗體。在設計要新安裝於CAR-T細胞上的外部辨識部位人工受體時，也會運行相同方式的預測模型，預先計算患者免疫系統將該受體本身誤認為陌生抗原而引發排斥反應的風險。即使是像PROTAC這種接近小分子的物質，在體內代謝後產生的新副產物，也不能

排除會引發意外免疫反應的可能性，因此同時運行代謝物預測模型與毒性預測模型，在事前過濾掉這些風險的程序，現在已經成為開發初期階段的基本檢查項目。

從找出標靶的階段到驗證免疫原性的階段，即使是不同的模態，人工智慧所執行的驗證程序骨架，都有著相當高的相似度。

在處理作為藥物標靶的蛋白質方面，採取完全不同思維的技術也正在快速成長。PROTAC，是蛋白質降解靶向嵌合體英文首字母縮寫的名稱。過去大部分的藥物，都僅止於抑制有害蛋白質的活動。也就是只有在藥物分子附著於蛋白質的期間才有效，一旦藥物排出體外，或是蛋白質結構發生些微改變，就會重新開始活動。

PROTAC完全捨棄了這種方式。它不抑制有害蛋白質的活動，而是將該蛋白質整個交給細胞內可稱為資源回收處理班的蛋白酶體，使其被分解。這是一種雙鉤分子，一隻手抓住必須消除的蛋白質，另一隻手則招來會貼上分解標籤的酵素。問題在於，這種雙鉤結構體積大且笨重，難以穿透細胞膜，在體內的吸收率也很低。

最近，以表皮生長因子受體為標靶的EGFR-PROPK等機器學習藥動學預測模型問世，使人們能夠提前計算PROTAC分子在體內的吸收和分佈程度，進而精細調整分子結構。

PROTAC之所以備受關注，是因為它能夠針對既有藥物無法觸及的標靶進行治療。某些蛋白質表面光滑，藥物難以牢固附著，這類蛋白質長期以來被新藥開發者稱為無法觸及的標靶。PROTAC無需牢固抓住蛋白質，只需輕輕掛住即可，剩下的工作交由細胞自身的分解機制進行。因此，許多曾被新藥開發者放棄的標靶正重新被列入候選名單。

人工智慧也被用於找出這些能輕輕掛住的結合位點。它可以掃描蛋白質表面的眾多褶皺，提議在某個位置上掛什麼形狀的鉤子，同時計算該位置應連接什麼樣的連接子及其長度，使兩個蛋白質自然地相互靠近。

得益於這些人工智慧設計工具，PROTAC在實際臨床領域也取得了成效。與輝瑞合作開發乳腺癌治療藥貝法德格塞坦特的Arvinas在2025年美國臨床腫瘤學會公佈了VERITAGE-2三期試驗結果。在既往接受過治療的激素受體陽性、HER2陰性乳腺癌患者中，與既有治療藥物Fulvestrant相比，貝法德格塞坦特明顯延長了無進展生存

期。兩家公司於2025年6月向美國食品藥物管理局提交了新藥申請，並被指定為優先審查。開發KT-621（一種口服STAT6分解劑）的Chimera Therapeutics在異位性皮膚炎臨床Ib期試驗中，在血液和皮膚組織中分別顯示96%和94%的靶蛋白分解率。

Beigene開發的BTK標靶PROTAC BGB-16673在既往治療但復發的慢性淋巴球白血病患者的初期試驗中錄得78%的反應率，隨後在2025年進入三期試驗。Bristol Myers Squibb的雄激素受體標靶PROTAC也在前列腺癌患者中啟動了首項三期試驗。截至2025年中期，進入臨床階段的PROTAC項目在全球共有28項，其中10項已進展至二期或三期。

用經過遺傳改造的細胞本身作為治療劑的細胞療法（CAR-T）是更進一步的方法。正式名稱是嵌合抗原受體T細胞療法。醫療人員從患者血液中提取負責免疫的T細胞，通過遺傳工程修飾這些細胞，使其表達只能識別癌細胞表面蛋白的人工受體。這有點像給武裝警衛發放一份畫著特定人物描述的通緝單，然後將他們重新送回巡邏。

將這樣改造的T細胞重新輸入體內後，它們會精確定位並攻擊血液癌細胞。白血病和淋巴瘤等血液癌已經有多種CAR-T療法獲得批准使用。但對於胃部、肺部和大腸等部位的實體腫瘤情況就不同了。實體腫瘤周圍被厚重的防禦壁所包圍，這道屏障阻止免疫細胞進入和發揮作用，因此CAR-T細胞即使到達也往往無法發揮作用。

人工智慧在設計針對這些實體腫瘤屏障的CAR-T細胞上深度參與。CAR-T細胞進入體內之前，受體本身就會錯誤激活信號導致提前耗盡，這一現象被稱為強直性信號傳導。人工智慧模擬模型可以預先計算哪種受體結構會導致這種過早耗盡，並幫助選擇更能持久的結構。將反映每個患者身體狀況的數字孿生與機器人設備結合起來，以自動化方式製造CAR-T細胞的嘗試也在持續進行。

在將細胞輸入患者體內之前，計算機會先在虛擬中模擬腫瘤細胞和免疫細胞之間的實際戰鬥過程，以及名為細胞因子的信號物質大量釋放導致的細胞因子釋放綜合征的嚴重程度。然後根據這些模擬結果調整細胞設計以降低風險。

在實際臨床領域，針對實體腫瘤的CAR-T細胞療法也開始逐個展現成效。由Allogene Therapeutics開發的現成型CAR-T細胞ALLO-316在腎臟癌患者的初期試驗中表現出持久的反應，這被認為是CAR-T細胞療法在實體腫瘤上也能有效的信號。在STAR-101一期試驗中，針對名為KIR的新標靶的CAR-T細胞在42名患者中即使在

低劑量下也顯示了安全性和療效跡象。鑑於之前沒有任何獲批的細胞療法用於這種癌症類型，這一結果具有重大意義。

中國的UTC Therapeutics利用環形信使核糖核酸開發的CAR-T細胞獲得了中國臨床試驗許可，並在2026年美國癌症研究協會公佈了實體腫瘤一期結果。美國和英國還為患有橫紋肌肉瘤和尤因肉瘤等小兒實體腫瘤的60名患者啟動了同時針對兩種標靶的下一代CAR-T臨床試驗。

儘管取得了這些華麗的設計成果，局限性仍然明確存在。在計算機中完美計算的連接子穩定性或結合親和力並不保證能在實際活細胞和動物模型中完全再現。仍然會出現無法預測的脫靶副作用或免疫反應延遲出現，導致大規模臨床試驗中止的情況，而且用於訓練人工智慧模型的數據本身可能存在偏差，針對特定種族或特定臨床環境進行訓練，使得難以直接應用於其他環境的患者。

無論在設計階段進行了多麼精密的計算，終究無法繞過直接在人體中投予並驗證的程序，這一事實沉重地壓在這整個領域之上。

從抗體、抗體藥物複合體、PROTAC 到 CAR-T 細胞來看，人工智慧所扮演的角色有一個共同的趨勢。找出作為標靶的蛋白質與表位、從頭設計能附著在該標靶上的分子或細胞、預先篩選是否會刺激免疫系統，以及透過模擬預先檢查在體內實際的運作方式，這四項工作在一個循環結構中環環相扣。

與過去精細調整單一種類小分子化合物的新藥開發不同，現在的重心正轉向從多重標靶的觀點出發，一次性設計出混合了抗體、連接子、酬載以及基因改造細胞等多種零件的複合治療藥物。具備能實際承擔這項工作的人工智慧設計能力與數據的企業，很有可能掌握下一代新藥開發的主導權。

在目前進入臨床試驗階段的抗體藥物複合體或 CAR-T 細胞治療候選藥物出現之前，在電腦中已經經歷了數萬個候選組合被創造出來又被丟棄的過程。在注射針頭實際進入患者血管之前，先在電腦模擬環境中驗證抗體的結合力、連接子的穩定性、酬載的毒性，甚至細胞受體的訊號傳遞，這樣的程序已不再是例外，而是成為了標準。與過去只為了最佳化一顆小分子藥丸的新藥開發時代相比，現在正在創造一個完全不同的產業。

抗體藥物複合體一路走來並非一帆風順。在 2000 年被批准作為白血病治療藥物的 Mylotarg，雖然獲得了首款抗體藥物複合體的名號，但由於懸掛藥物的連接子不穩定，導致毒性物質在正常細胞中洩漏，除了引發肝靜脈閉塞症等副作用外，在後續的臨床試驗中也未能再次證明生存益處，最終於 2010 年自願退出市場。即使是目前廣泛使用的 Enhertu，每十名患者中也有一名會出現間質性肺病。毒性物質超越標靶細胞洩漏到周圍組織的旁觀者效應被認為是原因之一。2025 年 12 月，Daiichi Sankyo 與 MSD 共同開發的 Ifinatamab deruxtecan 在小細胞肺癌的第三期臨床試驗中，因為出現多起間質性肺病導致的死亡案例，使得全球的臨床試驗一度全面暫停。

美國食品藥物管理局雖然解除了暫停令，但這兩家公司必須提出安全對策，將高風險患者排除在試驗對象之外，並重新培訓臨床試驗人員。歐洲監管機構則尚未放行。看來無論在電腦螢幕上計算得多麼穩定的連接子，在人體肺組織中會產生什麼樣的反應，人工智慧似乎還無法預先完全篩選出來。

藥效再好，如果價格令人難以負擔，最終還是無法送到患者手中。像 Trodelvy 和 Padcev 這樣的抗體藥物複合體，每次施打都要花費數百萬韓元，在確定納入健康保險給付之前，這些費用都必須由患者自行承擔。韓國健康保險審查評價院藥劑給付評價委員會在審查 Trodelvy 時，儘管其成本效益指標比已經審查完畢的其他藥物還要差，但基於沒有可替代的治療方法，且能顯著延長生存期等理由，最終還是核准了給付。

在 2026 年 5 月 19 日於國會議員會館舉行的女性癌政策討論會上，乳癌病友總聯合會會長 Choi Seung-ran 表示，帶有基因突變的轉移性乳癌患者，即使治療方法就在眼前，也會因為費用和制度的問題而面臨治療的懸崖。即使在設計階段，人工智慧能大幅縮短開發時間並降低成本，但只要這些省下來的成本沒有直接轉化為藥價的降低，患者實際享受到的好處就很可能只會緩慢增加。

韓國政府與企業也正跟上這股潮流。食品藥物安全處在 2025 年 11 月 28 日首次發布了指南，整理出在審查抗體藥物複合體品質時應該注意的事項。內容包含了應如何分開撰寫抗體、連接子及酬載各自的品質資料，以及當製造方法改變時需要重新審查哪些項目。

除了 LigaChem Bio 與 Celltrion 之外，以連接子技術為優勢的 IntoCell 正與 Samsung Bioepis 進行共同開發，同時也與中國 Frontline Biopharma 簽訂了共同開發兩種候選藥物的合約。Pinotbio 則是將五個抗體藥物複合體候選藥物中的三個轉讓給 Celltrion，剩下的兩個則繼續由其自行開發。

3.2

天然物新藥資料庫的AI解讀與混合式發掘

1971年深秋，在北京郊外的一間實驗室裡，屠呦呦正翻閱著大約在西元340年由葛洪所著的醫學書《肘後備急方》。這項實驗並非個人研究，而是隸屬於1967年啟動的中國國家計畫，也就是所謂的523號任務。隨著因瘧疾而倒下的士兵不斷增加，政府召集了全國的研究人員，指示他們尋找抗瘧疾物質。當時，瘧疾是每年奪走數百萬人生命的一種疾病，且對現有治療藥物Chloroquine具有抗藥性的菌株正廣泛傳播。

屠呦呦的團隊雖然已經反覆進行了數百次實驗，但卻沒有取得明顯的成果。古書中一句「不要將青蒿放入熱水中熬煮，而是要像榨汁一樣在低溫下浸泡」的內容吸引了她的目光。她就這樣從古文獻中解讀出，如果加熱，藥效成分就會被破壞的事實。從這個洞察中誕生的物質就是瘧疾治療藥物Artemisinin，而屠呦呦也因此獲得了2015年諾貝爾生理學或醫學獎。

古埃及與希臘人也透過經驗得知，咀嚼柳樹皮可以緩解發燒和疼痛。19世紀末，德國Bayer公司的化學家Felix Hoffmann將樹皮中的成分進行化學修飾，製造出了Aspirin。然而，Aspirin究竟是阻斷了體內的哪種酵素來緩解疼痛與發燒，也就是所謂的作用機轉，直到藥物上市將近70年後的1970年代才在科學上被釐清。1928年，在倫敦的一家醫院實驗室裡，Alexander Fleming休假回來後，看到一個培養皿被青黴菌污染了。

從只有黴菌周圍沒有細菌生長的那個偶然痕跡中，誕生了Penicillin。1960年代，受美國National Cancer Institute委託的植物學家們剝下太平洋紅豆杉的樹皮，找到了抗癌成分，這種物質後來以Paclitaxel之名的抗癌藥物問世。據了解，目前藥局裡販售的小分子藥物中，有很大一部分都根源於這種來自自然的結構。

這份清單並未以這四種結束。從生長在歐洲的植物毛地黃葉片中提取的Digoxin，很久以前就被用來作為調節心臟衰竭患者心跳的藥物。從馬達加斯加的夾竹桃科植物中，提取出了用於治療白血病的Vincristine，而從罌粟果實的乳汁中，則提取出了

至今仍被視為強效止痛藥的Morphine。用來降低膽固醇的Statin類藥物，最初也是源自於黴菌為了保護自己而製造出的物質。

僅憑這幾個例子，就能推估自然界扮演人類藥櫃角色的歲月有多麼深厚。Aspirin、Penicillin、Paclitaxel和Artemisinin，都是人類長久觀察自然，並累積偶然與毅力所獲得的結果。

如今，人工智慧正在電腦螢幕裡重複著這個發現的過程。據推測，自然界中由植物、微生物與海洋生物所製造的化學物質至少有數十萬種，如果人類要將它們一一放上實驗台確認藥效，就算花費好幾個世代也無法完成。在化學資訊學領域中，常常用「理論上可組合的藥物候選分子數量，比宇宙中的星星數量還要多」來比喻，對於人類徒手去尋找來說，那個空間寬廣到難以想像。

雖然人類至今在化學上確認並命名的天然物只有數十萬種，但如果加上連結構都尚未被揭露的化合物，推測其數量將會增加好幾倍。為了應對這個廣大的空間而出現的技術，就是圖神經網路（Graph Neural Network, GNN）。

圖神經網路這項人工智慧技術，將一個分子的原子轉換成點、化學鍵轉換成連接這些點的線所構成的網狀結構，也就是圖（Graph），並將這個網的圖案整個進行學習。就像鑰匙上凹凸不平的溝槽必須與鎖頭內部的結構相吻合才能打開門一樣，分子的立體結構必須與體內蛋白質的特定溝槽相吻合才會產生藥效，而圖神經網路能以原子為單位計算出這種契合度。在最近的研究中報告指出，用這種方式建立的虛擬篩選模型，在挑選出實際具有藥效的化合物的準確度指標上，取得了超過0.90的成績。

據了解，結合了自注意力機制的DTIGN等模型，在預測化合物與標靶蛋白質結合力的任務中，展現了比上一代模型更進一步的效能。不過，這種模型通常只使用已知的化合物與標靶的結合資訊來進行學習，因此也有人指出其侷限性：如果在訓練過程中遇到從未見過的全新類型標靶，預測能力就會大幅下降。人類在實驗室裡花費好幾年來篩選數百種候選物質的工作，這種模型在電腦裡幾天內就能完成。

隨著人工智慧一一解讀植物化學物質，也就是植物為了保護自己而製造出的化學物質結構，過去不為人知的藥理機轉也逐漸浮出檯面。長久以來被用於治療潰瘍性大腸炎的天然化合物Berberine，究竟調節了體內哪條訊息傳遞路徑，正透過網路藥理

學與人工智慧分析才具體描繪出來。Quercetin衍生物Fisetin，也正透過定量結構-活性關係模型，也就是計算分子結構與藥效之間數學關係的模型，來同時評估標靶結合力與毒性。專門被訓練來處理源自自然物分子的骨架大小、分支延伸與立體異構結構的變分自編碼器模型，甚至已經進展到能在維持複雜3D骨架的同時，生成改善了藥物性質的全新天然物相似結構。

全世界實驗室共同累積的質譜資料庫GNPS2中，截至2025年7月為止，已經收集了超過240萬筆光譜，如果將新取得的樣品光譜與這個資料庫進行比對，就能避免把時間浪費在重新發現已知化合物上。人類花費好幾年才從數百種混合物中分離出一種有效成分的工作，人工智慧只要輸入結構式，就能在幾小時內將其縮小到候選範圍。

所有這些預測最終都需要整理良好的數據庫來支撐。名為COCONUT的公開天然物數據庫將散落在全球文獻中的天然物結構信息匯集到一個地方，供所有人下載，天然物圖譜（Natural Products Atlas）則以標準化的格式保存了微生物產生的天然物的來源和生產菌株信息。如果數據保持分散狀態，即使是最出色的人工智能模型也無法找到用於學習的材料，所以不斷更新這類公開數據庫和篩選錯誤雖然被亮麗的新藥發表所掩蓋，但它是不可或缺的基礎工作。

如果將GNPS2等質譜數據庫與這類結構數據庫連接起來，僅通過初次看到的樣本的光譜，就可以提前縮小其中可能含有的化合物範圍。

同樣的解碼工作不僅涉及我們能看到的植物，還延伸到土壤中微生物的基因組。微生物將它要製造的化合物的設計圖儲存在稱為生物合成基因簇的基因組塊中，在整個基因組中尋找這些塊的工作叫做基因組挖掘。2025年發表的一項自動化基因組挖掘研究以鏈黴菌菌株中自我保護的抗性基因為線索，提出了高速尋找生理活性天然物的方法。新開發的基因組挖掘工具NptagM將酶預測準確度比現有廣泛使用的antiSMASH工具提高了三倍，同時將執行時間縮短到八分之一。

像e-DeepBGC和BiGCARP這樣的深度學習模型，只看基因序列就能分類其中隱藏著什麼類型的生物合成途徑、是抗生素還是抗癌物質。已知生活在深海熱液口或極地冰層等人類難以到達的環境中的微生物大多無法在實驗室中培養，但只要獲得基因組序列，人工智能就可以在不進行培養的情況下提前解讀其中沉睡化合物的設計圖。

據了解，即使在已獲得基因組序列的菌株中，到目前為止也只有少數情況揭示了其中生物合成基因簇的實際產品，所以基因組挖掘將來發現的化合物空間仍然非常廣闊。

由於很久以前就發出了沒有新抗生素就連普通細菌感染都治療不了的時代即將到來的警告，能夠解讀這個土壤中的化學工廠就本身而言是一項緊迫的任務。

Halicin和Abaucin都是人工智能首先發現的抗生素候選物質。麻省理工學院詹姆斯·柯林斯實驗室是早期回應這一緊迫需求的地方。2020年，該研究團隊通過機器學習發現，原本作為糖尿病治療藥物候選開發的halicin物質出乎意料地具有殺死多種細菌的抗生作用。2023年，他們發現了對不易受抗生素作用的鮑曼不動桿菌有效的物質abaucin，並揭示了該物質通過抑制細菌的LolE蛋白（負責脂蛋白運輸）而殺死細菌的作用機制。

2025年，該研究團隊放棄了搜索現有化合物列表的方式，轉而使用生成型人工智能繪製從未存在過的分子。他們用計算機創造了超過3,600萬個虛擬化合物並進行抗菌特性篩選，最終發現了針對治療困難的淋球菌和對抗生素耐藥的金黃色葡萄球菌（即MRSA）的候選物質，其結構和作用方式都不同於現有抗生素。以此成果為基礎，他們獲得了美國衛生高等研究計劃局的資助，用於設計15個新型抗生素候選物，並與非營利組織Fairbio繼續推進抗生素人工智能項目。

2025年10月，同一研究團隊用生成型人工智能新發現的抗生素候選物質繪製了其攻擊腸道細菌的具體途徑圖，這被介紹為將通常需要數年才能完成的作用機制闡明縮短到數個月的事例。

在麻省理工學院之外，也在用各種方式進行用人工智能尋找抗生素的研究。2026年2月，一份科學專業期刊介紹了塞薩爾·德拉普埃塔研究團隊，指出該團隊不僅限於化合物列表或單一微生物基因組，而是用人工智能廣泛掃描各種生物的遺傳信息，從而發現了具有抗菌特性的短肽鏈。人體腸道或皮膚中生活的微生物群落本身也是龐大的化學信息庫，人工智能在從這些信息中篩選出僅能殺死病原菌的短氨基酸序列方面越來越熟練。

不將抗生素的原料局限於特定植物或特定微生物，而是將地球上所有生物的遺傳信息視為候選資源庫的這種觀點，正因為有了處理數據的人工智能才得以實現。

在自然的化學多樣性中，有一個厚重的地層是數千年來積累的傳統中醫藥方記錄。這個處方體系不是一種藥材，而是多種藥材按既定比例混合以同時作用於多個靶點，長期以來只被視為經驗醫學，很難獲得科學認可。如果在這裡應用逆向工程技術——通過拆解已完成的結果來追溯設計原理，就可以逐一揭示處方中各種藥材分別作用於哪個靶點，以及它們的力量如何結合產生協同作用。要通過實驗確認一個能延緩衰老的成分，需要逐一將數百個處方和數千個成分塗抹到細胞上進行檢驗，但人工智能可以先在數據層面上篩選這個過程。

人工智慧新藥開發企業 Insilico Medicine 透過學習從人類、小鼠和獼猴等多個物種的組織中提取出的超過 120 萬筆體學數據，即大量測量基因、蛋白質和代謝物的實驗數據，建立了名為 Precious3GPT 的基礎模型。這種為了能廣泛應用於多種任務而預先進行廣泛訓練的模型被稱為基礎模型，Insilico Medicine 將此模型與收集了中醫藥成分和靶點資訊的 BATMAN-TCM2 資料庫進行了結合。該結合模型找出了在人類和其他動物物種中共同出現的十三個與老化相關的基因，並篩選出與這些基因重疊的三十四種中藥成分。

每次細胞分裂時，染色體末端的端粒會逐漸變短，直到某個瞬間細胞本身就會老化。結果，研究在分子層面上確認了這樣一個事實：一種名為 Hua Shan Wu Zi Dan 的古老處方中所含的多種成分，在肝臟、肺部和肌肉組織中會共同作用，以延緩這種端粒的縮短。

除了解讀單一處方之外，將整個中醫學知識轉化為可搜尋地圖的嘗試也在持續進行。資料庫 TCMBank 透過文獻探勘，整理了 9192 種中藥材、6 萬 1966 種成分、1 萬 5179 種靶點蛋白質以及 3 萬 2529 種相關疾病之間的連結關係。不僅如此，還將各成分的三維結構與 DrugBank 和 PubChem 等西方新藥資料庫相互連接，使得中藥成分與已獲批的合成藥物之間的化學相似性也能一目了然地進行比較。得益於這種比較，研究人員無需從頭合成新化合物，只需因為已用於治療其他疾病的藥物在結構上與中藥成分相似，就能將其作為新適應症的候選藥物重新審查，從而同步進行藥物重新定位的工作。

TCM-GPT 是首個僅使用中醫學文獻進行預訓練的 70 億參數級語言模型，在中醫師資格考試題目和診斷任務中，取得了超越現有通用語言模型的成績。北京的人工智慧企業水木分子與清華大學 AIR 研究所攜手，以擁有 1000 億個參數的大型語言模

型為基礎，推出了名為 ChatDD 的對話型新藥開發助手。它分別提供協助發掘靶點和探索候選物質的模組，以及協助設計臨床試驗的模組，並將應用範圍擴展到中醫學現代化研究中。西方醫學將成分、靶點和疾病分別獨立看待，這種單一靶點的途徑無法解決的問題，隨著人工智慧將多種成分如網狀般交織、同時作用於多個靶點的中醫學原理轉化為數據，正在重新迎來突破。

澳門大學的研究團隊針對阿茲海默症這類原因靶點交織在一起的神經退化性疾病，建立了一個整合中醫學文獻和化合物數據的獨立人工智慧平台，正在測試多靶點途徑是否也能對這類難治之症產生療效。

儘管取得了這些成果，要讓人工智慧直接吸收中醫藥數據並不容易。透過把脈、看舌苔、詢問症狀等方式累積的診斷資訊，原本就沒有以數據形式保留下來；而處方記錄在每家醫院和中醫診所的書寫方式也各不相同，單是要整理成一個表格就需要相當大的修整工作。建立資料庫的研究人員必須逐一閱讀論文和古籍，並由人工對成分和靶點資訊進行雙重驗證，原因也正在於此。

如果勉強讓模型學習未經標準化的資料，人工智慧就會有將不存在的相關性呈現得像真的一樣的風險。因此，整理資料的修整工作與華麗的模型開發同樣重要。

Hanguk Gwahak Gisul Yeonguwon 將其擁有的天然物樣本資訊進行數位化，並開發了預測蛋白質靶點的人工智慧平台 NPI-finder。該平台證明了柿葉乙醇萃取物改善乾眼症狀的功效，並成功獲得了 Sikpum Uiyakpum Anjeoncheo 的批准和技術轉移。Son Mi-won 代表於 2020 年成立的 MTera Pharma 建立了一個名為 SyMthomics 的大數據平台，能夠精確揭示由多種成分組成的天然物的作用機制。從這個平台產出的糖尿病性神經病變治療候選物質，在美國二期臨床試驗中獲得了減少 48% 疼痛的結果，目前正在準備三期臨床試驗；而帕金森氏症治療候選物質則於 2022 年完成了美國一期臨床試驗，並引起了 AbbVie 和 MSD 等全球製藥公司的關注。

Yeungnam Daehakgyo 醫學與生命工程學系的 Choi In-ho 教授研究團隊於 2026 年發表，他們結合了人工智慧與電腦模擬分析，找出了治療肌少症的天然物候選物質。他們完全遵循了無需實驗，僅透過電腦分析篩選出候選物質後，再透過細胞和動物實驗進行確認的順序。政府也在將於 2026 年至 2030 年實施的 Je5cha Haniyak

Yukseongbaljeon Jonghapyehoeok 中，將中醫藥人工智慧與數位大轉型定為核心目標，為這一趨勢注入了力量。曾有一段時間被評價為在韓國國內製藥業界中被邊緣化的天然物新藥，由於人工智慧大幅減少了發掘成本和時間，正再次成為備受矚目的資產。

無論人工智慧挑選出多少候選物質，最後的確認仍然必須經過培養皿和動物模型。麻省理工學院的研究團隊在 3600 萬個虛擬化合物中篩選出有望具備抗菌效果的候選物質後，仍然需要實際合成該化合物並將其暴露於細菌中，以觀察細菌是否存活的實驗。Hanguk Gwahak Gisul Yeonguwon 的 NPI-finder 同樣並未僅停留在電腦預測階段，而是在經歷了實際應用柿葉萃取物的試驗階段後，才得以獲得 Sikpum Uiyakpum Anjeoncheo 的批准。

我們將交替重複預測與實驗，並在實驗結果與預測不符時，將誤差重新讓模型學習的這種循環結構，稱為混合發掘。由人工智慧先縮小人類無法負荷的探索範圍，再由人類手動確認這些縮小後的清單，這種分工是目前這個領域的標準做法。只有當我們承認，人工智慧給出的每一個預測並非直接就是標準答案，而是需要透過實驗確認的假說時，這個循環才不會白費力氣，並能順利導向下一個候選物質。

當植物化學物質數據與傳統醫學處方數據，結合深度學習與結合親和力預測演算法（也就是計算藥物分子與標靶蛋白質結合強度的演算法）時，就誕生了自然模仿結構這種全新的設計方式。這不是從白紙開始描繪化合物，而是讓人工智慧學習大自然耗費數億年打磨出的骨架，並透過改變這些骨架來提出新的候選物質。菲利克斯·霍夫曼從柳樹皮中提煉並製造出阿斯匹靈，需要累積數十年的化學知識；但現在，我們可以同時描繪出數千種同系列骨架的變形，並在幾天內從中篩選出極具潛力的候選物質。

據報導，最近公開的 AlphaFold3 系列結構預測模型，在針對評估化合物與蛋白質結合強度的任務進行調整後，能以少得多的運算量，達到過去需要耗費數天進行自由能微擾計算的相似準確度。包含抗生素、抗癌藥物與降血脂藥物等許多已獲批准的治療藥物，追根究底都是直接或間接從自然界骨架衍生而出的物質。考量到這一點，人工智慧學習這些骨架的語法並提出新結構的現有趨勢，似乎是順理成章的下一個階段。

天然物之所以能成為特別的新藥材料，是因為它們的結構已經過數億年演化篩選。植物與微生物為了抵禦掠食者或壓制競爭者，不斷製造出化學物質，在這個過程中，能與活細胞蛋白質完美契合的立體結構自然而然地累積下來。在實驗室裡用電腦從頭設計的化合物，經常會無法穿透細胞膜，或是在體內很快就被分解；相較之下，天然物的骨架已經是通過在生物環境中生存考驗的結構。

目前醫院裡使用的多種藥物系列，如抗生素、抗癌藥物與免疫抑制劑等，都是從這種源自自然的骨架衍生而來，這一點證明了天然物的優勢。在化學資訊學的研究人員中，越來越多的人認同：與隨機生成的虛擬化合物相比，以自然界產生的骨架為出發點時，新藥候選物質存活到臨床階段的機率會更高。

然而，基於同樣的理由，天然物也很難處理。一株植物中混合了數百種成分，很難分辨出是哪一種成分發揮藥效，且有效成分僅佔整體萃取物 100 萬分之 1 的情況也很常見。根據原生地的氣候與土壤條件，成分含量每年都會改變，難以確保重現性；而結構複雜的分子在化學上也不容易大量合成。實際上，在最初生產紫杉醇的時期，數千棵太平洋紫杉的樹皮被剝下，導致該樹種本身面臨資源枯竭的危機。這種為了採集有效成分，反而將作為原料的植物或菌株推向滅絕危機的矛盾，至今仍未完全解決。

長久以來，傳統的探索方式也常被指出一項限制：經常會重複找到相同的化合物，因此無論收集多少樣本，遇見全新結構的機率仍然很低。雖然人工智慧能透過預先過濾已知化合物的質譜指紋，來減少這種重複發現的浪費，但要從實際的植物與微生物中再次取得足夠數量的這些物質，依然需要仰賴人工。

在預測天然物進入體內後如何被吸收、分解以及引發毒性的工作（也就是所謂的 ADMET 預測）中，人工智慧的用途也越來越廣泛。由於多種成分混合的萃取物中，成分之間存在著互相幫助或阻礙吸收的複雜交互作用，因此很難像處理單一化合物那樣來計算毒性。最近，已經陸續發表了一些多成分預測模型，這些模型不僅輸入個別成分的結構，還會將成分的組合比例一併輸入，以計算出引發肝毒性或心臟毒性的機率。

對於試圖向監管機構申請批准原始複合萃取物（而非純化的單一成分）的天然新藥開發公司來說，這種預測能力從設計臨床試驗的階段開始，就成為了降低失敗機率

的實用工具。

將從自然界取得的知識轉換為數位數據的這項工作，也同時引發了一個問題：這些知識原本屬於誰？2010年通過的名古屋議定書，是一項規定若利用某國的生物資源及其相關傳統知識獲取利益，就必須將利益與資源原產國及當地社區分享的國際協定。問題在於，現在人工智慧處理的不是植物標本或種子等實體資源，而是記錄了其基因序列的數位序列資訊。種子需要經過海關，但數據卻不用。在不需要讓任何生物樣本跨越國界的情況下，就能直接從公開的資料庫下載序列資訊來訓練人工智慧，這使得如何向原產國分享利益的爭論一直持續不斷。

源自馬達加斯加夾竹桃科植物的抗癌藥物長春新鹼，經常被用來作為展現這種擔憂的案例。長久以來，人們不斷指出：西方製藥公司利用這種植物開發出創造龐大營收的抗癌藥物，但作為原產地的馬達加斯加，卻幾乎沒有得到相對應的報酬。在談判桌上，擁有豐富基因資源的開發中國家主張，數位序列資訊也應該與實體資源受到相同對待；然而，主要處理這些資訊的已開發國家製藥公司與學術界則反駁，過度嚴格的管制反而會拖慢研究的進度。

2025年2月25日，在義大利羅馬，《生物多樣性公約》《名古屋議定書》的簽約國推出了一項國際基金，名為卡利基金(Cali Fund)，兩天後並通過了關於數位序列資訊的正式決議。利用遺傳資源的數位序列資訊而獲利的企業被建議向該基金貢獻其營收的0.1%或利潤的1%，所籌募的資金用於支援生物多樣性豐富的開發中國家和原住民社群。

雖然目前還很少有企業遵循此建議，法律約束力也很弱，但在只需數位資訊就能設計新藥的時代，這是國際社會試圖以某種方式將利益返回原產地國家的第一步，具有重要意義。該基金用數字重新確認了這個事實：進行韓醫藥方逆向工程和微生物基因組開採的所有工作的根本基礎中，終究是特定的土地和特定的社群長期以來守護的生物資源。

將一個韓醫藥方轉換成圖表並將一個微生物基因組輸入語言模型，現在在技術上已不困難。實際上耗時的是將這樣找到的候選物質轉化為能夠安全用於實際患者的藥物的最後階段，以及從構成該發現根源的知識和生物資源中公正分享利潤的程序。圖神經網絡和大型語言模型最終都運作在人類在數千年間觀察、記錄和失敗而積累

的數據基礎之上。就如同「在低溫下浸提青蒿」這句古老的說法最終導向了諾貝爾獎一樣，現在躺在某個數據庫中的一個處方和一個序列也可能導致下一種新藥。

當人工智能現在讀取數億年進化和數千年經驗留下的記錄時，是否將正當利益返回給最初留下這些記錄的土地和人民，很可能決定此項技術的下一步成敗。一個藥丸的完成必須將螢幕上的圖表、實驗台上的培養皿以及它們根源所在的森林和村莊全部連接起來，這個事實似乎無論人工智能變得多麼精巧都不會改變。

韓國韓醫學研究院在過去10年間通過人工智能相關課題和體質基因組流行病學課題積累了近5萬人的臨床數據。依賴韓醫師的經驗和主觀判斷的診斷方式長期以來一直被指為阻礙韓醫學被認可為科學的因素，但研究院選擇了將遺傳數據、面部照片和穿戴式設備測量值一起蒐集，以數字來支持這些診斷判斷的方向。

除此之外，韓醫界也自行開始了根據國際醫療資訊標準FHIR格式整理臨床記錄的工作。遵循此標準，韓醫診所的臨床記錄也能與醫院電子病歷或健康保險數據等其他醫療資訊相互連接。研究院計畫利用這樣積累的數據來開發韓醫精準醫療技術。

使成分保持一致仍然是通過監管障礙的難題。2015年，審計院指出天然物新藥制度對國內企業給予了過度特權，之後食品藥品監督管理局消除了天然物新藥的單獨分類本身，並將審查標準提升至新藥水平。美國食品藥物管理局也在2016年發布了天然產品藥物開發指南試圖解決此問題，但未經加工、直接被批准為藥物的植物提取物案例至今仍屈指可數，只有2006年的維瑞根和2012年的克羅非勒梅。

這是因為對於成分含量根據採收時間和產地而變化的物質，審查官很難每次都判斷這是同一種藥物。即使人工智能預先計算成分偏差並提供校正值，根據這些值決定是否核准的工作仍然很可能由人類審查官承擔。

一個傳統知識轉化為新藥的案例為我們思考如何分享利益提供了線索。秘魯亞馬遜地區的原住民長期以來一直使用克羅頓·萊瑞樹(*Croton lechleri*)流出的紅色樹液作為治療傷口和腹瀉的藥物。注意到這一知識的製藥公司與19個原住民社群（分布在約30個地點）合作，數十年來以永續方式採集樹液。用這種方式製成的克羅非勒梅在2012年獲得美國食品藥物管理局批准，作為治療艾滋病患者慢性腹瀉的藥物。該公司聲稱，作為回報，一直在採集地區支持教育和清潔飲水項目。

與其說是一份合約，不如說是數十年來積累的信任一直在支撐著這種利益分享。但在當今這個只需序列資訊就能設計新藥的時代，單靠公司的善意來長期保護原產地社群似乎是困難的。

3.3

奈米醫學及智慧標靶遞送系統 (Nanomedicine)

2020年12月，美國和歐洲的機場陸續迎來了裝載著特殊冷凍容器的貨運飛機，這些容器能夠承受零下70度的低溫。箱子裡裝著Pfizer和BioNTech共同開發的新冠肺炎疫苗BNT162b2，而這種疫苗的每一滴中都漂浮著數十億個看不見的油滴。這些油滴的每一個都是脂質奈米粒子（LNP）。遺傳序列公開後僅25天就完成初期設計的Moderna，以及以創紀錄的8個月速度獲得緊急授權的Pfizer和BioNTech，決定勝負的並非疫苗的核心材料信使核糖核酸（mRNA），而是將這種脆弱的遺傳物質安全運送到體內的油滴 -- 即奈米粒子。

mRNA是一種在體內酶的作用下會迅速分解的脆弱分子，如果沒有這些油滴像傘一樣為其提供保護，疫苗就不會問世。此後，全球數十億人捲起袖子接種了這種油滴，脂質奈米粒子從實驗室中的陌生物質一躍成為全人類親身體驗過的首款商用奈米藥物。這種油滴技術的根源在於加拿大研究人員Pieter Cullis數十年來對脂質奈米粒子的研究，以及他創辦的Acuitas Therapeutics公司，Pfizer和BioNTech所用脂質奈米粒子的許多核心專利均來自這家加拿大公司。

那一天降落跑道上的不僅僅是疫苗盒子，而是奈米醫學跨越實驗室圍牆、注入全人類手臂的時刻。

1納米是1米的十億分之一。如果將一顆玻璃珠算作1納米，那麼1米就相當於整個地球的大小，這是一個極其微小的世界。疫苗中使用的脂質奈米粒子直徑大約在100納米左右，還不到人類頭髮粗度的千分之一。我們身體中的一個紅細胞大約有7000納米，如果將這些奈米粒子排成一列，僅需70多個就能橫跨一個紅細胞。在這樣一個肉眼無法看見、普通顯微鏡也無法單獨觀察的微觀世界裡，奈米醫學的所有故事都由此開始。

但在這樣的微觀世界裡，即使奈米粒子的大小只相差幾納米，它是進入肝臟還是進入肺部、是進入細胞內部還是被過濾出來，都會完全不同。粒子表面附著了什麼分子、表面是帶正電還是帶負電，同樣是改變目的地的決定性因素。就像郵局的包裹

根據大小和外表標記被分配到不同的窗口一樣，奈米粒子也會根據大小和表面性質進入身體內不同的器官。身體內血管和細胞膜上的縫隙大小各不相同，太大的話無法進入，太小的話會被腎臟過濾而快速排出體外。

利用奈米粒子運送藥物的想法本身並不新穎。1995年，美國食品和藥物管理局批准了用油性外殼多層包裹Doxorubicin (Doxil) 的藥物，這是奈米技術首次被用於實際患者的案例。之後雖然出現了將Paclitaxel附著在白蛋白（一種血液蛋白質）上的Abraxane等後續藥物，但進展速度一直緩慢。以2025年為基準，全球獲得臨床應用批准的奈米藥物加起來也只有50到80個左右，而且大多數都沒有超出數十年前建立的化學基本原理。

這是因為從在試管中製造一種新的組合、將其放入細胞和動物體內觀察到獲得結果需要數個月，如果結果不理想就必須從頭開始。這種依靠人類直覺和反覆實驗長達近30年的緩慢進展，在遇到人工智能後才開始加速。

2025年，一個名為NanoSafari的人工智能代理不分晝夜地閱讀了超過2萬篇奈米科學論文。這個代理將論文中分散的奈米粒子合成條件、大小和表面性質數據整合在一起，創建了與大型語言模型相連的知識數據庫。相當於人工智能在一夜之間閱讀了研究人員一生需要閱讀的文獻量，文獻調查從需要幾個月縮短為一天。過去，研究人員需要在實驗台前花費數月甚至數年通過試錯來找到最佳的奈米粒子配方。

現在，一種稱為Nanoarchitectonics的方法正在快速傳播，它像積木一樣逐個堆砌原子和分子來構建具有所需功能的結構，借助人工智能的力量。就像建築師先繪製設計圖然後才能建造建築物一樣，人工智能先繪製分子設計圖，然後按照設計圖組裝奈米粒子的順序正在成為標準。這種設計方法不僅應用於新藥開發，還擴展到化妝品和精細化學品材料的製造，正在拓展為涵蓋整個奈米級設計的方法。

Sungkyunkwan University研究團隊在2025年發表在國際學術期刊《Advanced Materials》上的論文清楚地展示了這一趨勢。從通過生物信息學發現引起疾病的身體內部靶點的階段開始，經過通過機器學習精確設計奈米粒子表面的階段，再到在計算機內（in silico）提前計算這些粒子在體內的分佈位置和藥物釋放時間的階段，形成了一個環形結構。這個環形結構不是單向的，而是循環回到開始階段並反映到下一個設計中的閉環迴圈。

過去需要花費數週才能完成的一次設計、合成與評估週期，現在在電腦中只需一兩天就能重複好幾次。這個循環越緊密，藥物累積在錯誤器官的脫靶副作用就會減少，也更接近針對每位病患身體狀況量身打造的設計。將過去在實驗桌上反覆進行的嘗試與錯誤，轉移到電腦螢幕上的模擬，這似乎是根本上改變奈米醫學研究速度的轉捩點。

仔細觀察這樣設計出的一顆奈米粒子，會發現它就像是精心包裝的禮物盒。最外層的殼上附有只會與特定組織或細胞結合的標靶配體，也就是像蛋白質或醣鏈這樣，能準確附著在體內特定受體上的小標籤。這個標籤能辨識出腫瘤微環境（TME）或血腦屏障（BBB）等棘手的目的地，並為其開門。裡面則是透過機器學習，精細計算出厚度、排列及組成比例的骨架，決定了粒子整體的尺寸、硬度以及分解速度。

而在最內部的中心，裝有實際對身體產生作用的藥物，也就是有效載荷，直到抵達目的地後才會開始發揮作用。其核心在於這三層可以分開設計，只要保留內部材料，僅更換外層標籤，就能創造出針對完全不同器官的全新奈米粒子。這三層中只要有任何一層設計出現偏差，奈米粒子就會在錯誤的器官釋放藥物，或者還沒到達目的地就被排出體外。

名為 Agile 的深度學習平台會自行學習多達 6 萬多個虛擬脂質化合物資料庫，接著透過實際在實驗室中大量合成與測量所得的數據進行微調，藉此預先篩選出可游離脂質將遺傳物質運送至細胞內的效果有多好。2025年8月的國際學術期刊 Nature Nanotechnology 介紹了一款名為 COMET 的 Transformer 結構神經網路模型，該模型透過名為 LANCE 的史上最大規模脂質奈米粒子資料集進行學習，即使是未曾見過的脂質奈米粒子組合，也能預測出其功效。

同年11月，Johns Hopkins 的研究團隊表示，他們高速製作並測試了數百種脂質奈米粒子配方後，透過機器學習找出了能提高基因編輯效率的配方條件。2026年，學界出現了名為 LipoAgent 的系統，由多個大型語言模型代理分工合作，提出新的脂質結構，並互相交叉驗證設計，連同安全性一起進行評估。這些案例充分展現了實驗桌變成數據工廠，而電腦變成設計者的發展趨勢。

在決定要將脂質奈米粒子送往哪個器官時，廣泛使用了名為選擇性器官標靶化（SORT）的技術。就像從一串鑰匙中挑選不同的鑰匙會打開不同的門一樣，根據添加

到脂質奈米粒子表面的脂質分子特性，粒子前往的器官也會有所不同。實驗已經證實，添加帶正電的脂質會讓粒子聚集到肺部，添加帶負電的脂質會聚集到脾臟，而添加可游離的胺基脂質則會聚集到肝臟。其順序是覆蓋在粒子表面的聚乙二醇成分在體內脫落後，血液中的特定蛋白質會新附著在該位置，而這個蛋白質與各器官細胞的受體配對後，就決定了目的地。

隨著能夠選擇並瞄準特定器官，原本應該只在肝臟發揮作用的基因治療藥物蔓延到肺部或心臟所引起的副作用，也呈現減少的趨勢。人工智慧負責從眾多候補中快速篩選出這些微妙的化學組合，取代了由人類手動一一實驗的過程。

不僅是脂質奈米粒子，聚合物與微胞系列的載體也透過人工智慧變得更加精密。結合了在微小流體通道中製造粒子的微流體工程設備與高解析度影像數據進行學習的機器學習模型，能夠自行調整粒子的尺寸、表面電荷甚至是形狀。奈米粒子進入血液的瞬間，血漿中的蛋白質會蜂擁附著在粒子表面，就像穿上一件外套般形成新表面的現象，被稱為蛋白質冠，而這件外套是由什麼編織而成，將決定粒子在血液中停留多久，以及被哪個細胞吸收。

研究人員以收集了 817 種奈米粒子配方與 2497 種蛋白質吸附數據的資料庫為基礎，訓練了隨機森林或 XGBoost 等能夠匯集多個小判斷規則並得出結論的可解釋性演算法，發現粒子尺寸、表面電荷以及在血液中停留的時間，是左右蛋白質冠的核心變數。最近，在將 DNA 鏈摺疊而成的更複雜奈米結構物中預測蛋白質冠的研究也隨之而來，預測對象正超越脂質與聚合物，呈現逐漸擴大的趨勢。

隨著這種預測變得越發精準，就能夠透過不同的表面設計，區分並製造出需要長時間停留在血液中的抗癌藥物載體，以及需要快速被標靶細胞吸收的疫苗載體。這等於是能在進入臨床試驗前，先在電腦中描繪出這件外套的模樣。

當以大腦作為治療對象時，面臨著一道更高的牆。腦血管壁周圍環繞著密集網狀的血腦屏障（BBB），在一百個口服治療候選物質中，通過這道屏障的最多也只有六個，最少只有二、三個。許多針對阿茲海默症或帕金森病開發的新藥候選在動物實驗中顯示療效，卻無法在人體內跨越這道屏障，在臨床試驗末期失敗的案例不勝枚舉。結合了Transformer系列模型MegaMolBART和XGBoost的預測模型，以包含八千多個化合物通過與否信息的資料庫進行訓練，在表示預測與實際相符程度的曲線

下面積指標上獲得了0.88的分數。

2025年8月，美國Lantern Pharma公開了一個名為predictBBB.ai的人工智能模組，宣稱只需化合物的結構信息，就能以高達94%的準確度篩選出是否通過血腦屏障。這些利用圖神經網路讀取分子結構的預測模型，在新藥候選進入動物實驗前，在電腦螢幕上充當篩選工具的角色。據評估，這類預測工具篩選出的失敗案例，使進行動物實驗所需的預算和時間大幅減少。

對於化學上難以直接通過這道屏障的大分子，科學家們採用了完全不同的策略。他們將與腦血管細胞表面大量出現的轉鐵蛋白受體等運輸蛋白結合的抗體或小蛋白附著於奈米粒子表面，就像在木馬中隱藏士兵一樣，這些受體在進行原本的物質運輸工作時，藥物就能悄悄搭乘進入腦內。人工智能被用於找出應該修改抗體的哪個部位以及如何修改，才能與受體緊密結合，同時不引起體內不必要的免疫反應，並從眾多的胺基酸組合中找出最優方案。

這些受體介導運輸技術與預測模型相結合，已形成新的標準程序：先在電腦上篩選出哪種抗體設計能大幅提高通過概率，再通過實驗進行驗證。

面對難以用單純合成奈米粒子穿透的這道屏障，科學家們的關注焦點轉向了活細胞自身分泌的奈米級囊泡——外泌體。由於外泌體是身體本身製造使用的囊泡，免疫系統將其誤認為異物並發起攻擊的風險低於合成奈米粒子，這是臨床上偏好外泌體的原因之一。研究人員將提取自脂肪組織的幹細胞或骨髓間質幹細胞的外泌體，按人工智能計算出的最優比例裝填藥物，然後通過鼻腔噴霧給藥進行臨床試驗。這是因為外泌體從鼻腔黏膜接觸的嗅覺神經束進入，可以不經過全身血液循環而直接到達腦部病變區域。

這相當於一條不經過血液這條主幹道，而是直接通過嗅覺神經這條捷徑進入大腦的路線。研究人員將光驅動蛋白傳遞技術或用於抑制基因表達的CRISPR干擾技術裝載到外泌體中，在患帕金森病的實驗鼠腦神經網路中關閉了導致疾病的蛋白質 α -突觸核蛋白的基因開關，相關實驗也已被報導。由於利用的是身體原本製造使用的運輸工具，外泌體被視為下一代將藥物遞送至大腦的運輸載體。

然而，外泌體從實驗室走向藥局貨架還有相當長的路要走。截至2026年，全球正在開發的外泌體療法候選物約有120種左右，其中略超40%已進入臨床試驗階段，但

與預估市場規模581億美元相比，美國食品藥物管理局批准的外泌體療法至今仍一個都沒有。針對脊髓損傷或視神經損傷的候選物ExoPTEN也在進行前臨床階段，目標是進入人體臨床試驗，隨後的候選物進展也在加快推進。

開發分泌外泌體的心臟來源細胞療法Deramiciel的Capricor Therapeutics，在2026年7月29日召開的食品藥物管理局諮詢委員會上，因未能證明對杜氏肌肉營養不良心肌病的療效而遭遇9比3的負面投票，目前正等待定於8月22日的最終審查結果。諮詢委員會的投票雖然沒有約束力，但這個案例表明，儘管人工智能預測和動物實驗結果再詳盡，面對以人為對象的最後關卡時，仍然會適用謹慎的標準。

業界預測，無論Deramiciel最終結論如何，其結果將長期影響監管當局對外泌體這種治療方式整體的看法。

轉向非腦部的固體腫瘤，又有另一個難題在等待 -- 腫瘤微環境（TME）。腫瘤周圍組織中血管彎曲延伸，氧氣不足，甚至酸鹼度也異於正常組織，單靠一種方法難以將藥物運送到各個角落。研究人員觀察到，奈米粒子在腫瘤組織血管壁間隙（通常稀疏）中能很好地滲透並停留較久的現象，稱為通透性增強與滯留效應。研究人員將從378個腫瘤數據集和200多篇論文中提取的生物學反應信息，與描述奈米粒子化學結構與生物學效果之間關係的結構-活性關係模型，以及用方程式表示藥物在體內如何被吸收和分散過程的生理藥動學模型結合在一起，以24小時後為基準對這一現象進行了計算。

核心是難度而不是複雜度，那才終究是企劃；真正的護城河是代碼而非經驗・信任・數據。

顯示預測值與實際值吻合程度的決定係數高達0.83，準確度非常高。更進一步的生醫材料逆向工程，則將順序完全顛倒過來。只要先輸入想要殺死腫瘤細胞或喚醒免疫細胞的最終目標，演算法就會回溯龐大的資料庫，反向找出符合該目標的聚合物比例、表面配體，甚至是模仿腫瘤細胞膜或紅血球膜的偽裝塗層。只對腫瘤分泌的特定酵素或低酸鹼值產生反應，並且只在該處釋放藥物的聰明載體，就是這樣製造出來的。

只不過每一種腫瘤，甚至同樣的腫瘤在不同患者身上，微環境的細部條件都不盡相同，因此通常很難將在一位患者身上驗證過的設計，直接套用在另一位患者身上。

與將治療藥物送入體內同樣重要的是，提早察覺病情發展到了什麼程度，這項工作也由奈米技術與人工智慧共同負責。像是表面增強拉曼散射或局部表面電漿共振等奈米光學技術，能捕捉血液或腦脊髓液中含量極少的疾病訊號，並以光的形式將其放大。簡單來說，這就像用凹面鏡將手電筒微弱的光線聚集起來，使其變得明亮的原理一樣。人工智慧會即時讀取這些被放大的複雜光譜與螢光訊號，比其他人更早發現阿茲海默症的標記物——類澱粉蛋白 β 與Tau蛋白開始聚集的初期瞬間。

這與傳統診斷方法的不同之處在於，它能在症狀表現出來的幾年前，甚至患者自己都沒有察覺的時候，就找出疾病的苗頭。如果將高解析度奈米影像設備與深度學習影像分析結合使用，就能在體外觀察放入體內的奈米粒子實際上在病變部位聚集了多少，以及藥物正以什麼樣的速度釋放出來。

若再更進一步，下一個目標就是自主驅動型的精準醫療：植入體內的奈米感測器會將患者的生理訊號即時回傳給人工智慧，人工智慧再根據該訊號自行調節藥物釋放的速度。就像測量血糖並自動調整胰島素劑量的儀器已經在陪伴糖尿病患者一樣，有人預測在不久的將來，針對癌症或退化性腦部疾病，體內的奈米感測器與奈米載體也能相互配合，不需要人工介入，就能全天候為患者調整藥量。

只不過，這種自主驅動型系統若要實際進入病房，就必須一併驗證安全裝置，以防止感測器發送錯誤訊號時釋放過多藥物。將診斷、治療以及連結兩者的自主調節功能集於一身的奈米系統，被視為指引奈米醫學下一步發展方向的範例。

能夠用來訓練人工智慧模型的高品質奈米毒性數據嚴重不足，而且各個實驗室的測量方法與標準也不盡相同，因此光是要將數據匯集起來就很不容易。即使是相同的奈米粒子，毒性數值也會因為測量實驗室的不同而有所差異，這類情況屢見不鮮，有時也很難判斷在不同的結果中應該採信哪一方。許多大學與企業各自使用不同程序累積了大量數據，目前就連將這些數據整合為單一標準的工作，都還停留在起步階段。無論預測準確度有多高，深度學習模型都存在著「黑盒子」的問題，無法向臨床醫師或法規主管機關有條理地解釋，為什麼它會判定某種特定的奈米粒子結構不會引發毒性。

如果用微分方程式來解開物質運動的傳統生物物理學模型，無法與直接從數據中尋找答案的人工智慧模型互相配合，那麼在電腦螢幕上看似完美的預測，在實際人體

內就始終存在著引發免疫排斥反應或意外副作用的風險。此外，即使是人工智慧全新提出的奈米粒子結構，在進入人體之前，仍必須承擔長達數年的臨床前試驗與臨床試驗成本，因此，設計階段變快了，並不代表新藥上市的整體時間表也會等比例縮短。

制度的步伐至今依然無法完全跟上技術發展的速度。在 2025 年美國食品藥物管理局批准的 46 款新藥中，有超過一半是採用全球首創機制的治療藥物，這項事實顯示法規當局也正以自己的步調改變審查方式。然而，像奈米藥物這種物質本身由多層結構組成的情況，仍然需要另外制定一套標準。美國食品藥物管理局通過了現代化法案 2.0，修改了過去強制要求進行動物實驗的長久規定，為使用器官晶片或電腦模擬來替代臨床前毒性試驗開啟了一條新路。

然而，包含歐洲藥品管理局在內的許多國家法規機關，對於像外泌體這種源自活細胞的載體，或是能夠穿透血腦屏障的尖端奈米藥物，都各自要求不同的審查標準。因此，在一個國家通過審核的產品，到了另一個國家往往得從頭開始重新準備資料，這類情況十分常見。

人工智慧在大量學習患者基因體資訊的過程中，對於個人資料的同意範圍應該到哪裡，目前界線依然不夠明確。當人工智慧自行設計甚至決定好劑量的奈米粒子，對患者造成嚴重的副作用時，究竟該由軟體開發者、製藥公司，還是開立處方的醫療人員來承擔責任，各自又該分擔多少，目前也沒有標準答案。現有的產品責任法理，習慣將設計缺陷或製造缺陷，縮小範圍歸咎於特定時間點、特定主體的判斷錯誤。但是在人工智慧藉由學習龐大數據，自行改變設計變數的奈米藥物上，想要追溯錯誤究竟是何時、在哪個階段發生的，並不是一件容易的事。

法律界內外越來越多人認為，必須重新制定標準，來釐清開發設計演算法的公司、針對特定患者進行微調的醫院，以及開立最終處方箋的醫療人員之間的責任歸屬。為了讓奈米醫學能夠真正成為臨床現場的標準，我們似乎需要同時進行幾項工作：發展能讓人理解演算法判斷依據的「可解釋人工智慧」技術、經過嚴密的細胞與動物實驗，並且將各國不同的法規整合起來。

這股趨勢也同樣影響著韓國的製藥與生技產業。2026年舉辦的 Nano Korea 是一場匯集奈米生技、3D列印與精密測量儀器的世界級規模活動，其中由人工智慧設計的

標靶傳遞載體與診斷儀器成為了主要話題。在韓國國內，以抗癌新藥和基因治療領域為中心，從奈米粒子設計階段就導入人工智慧工具的企業也正陸續出現。

韓國國內的監管機構也正密切關注美國與歐洲審查標準的變化，並持續完善相關程序。國內的製藥公司、研究機構，以及為他們提供諮詢的法律專家，如果不想在這股洪流中落後，似乎就需要不斷努力追蹤海外的臨床數據與法規動態。

2020年冬天，在零下70度的冷凍箱中首次向世界展現存在感的脂質奈米粒子，如今已經成長為能夠精準找出大腦深處病灶與腫瘤中心的人工智慧設計載體。在這段期間，製造奈米粒子的工作，已經有很大一部分從手持試管的研究人員，轉交給解讀數據的人工智慧運算。儘管如此，這並不代表人類就無事可做了。決定要相信哪些數據、承擔哪些風險，以及優先向患者詢問什麼問題，這些工作依然是由研究人員、醫療人員和監管人員來負責。

剩下的課題，就是搭建一座橋樑，讓這份精密的設計圖能夠順利走出實驗室的螢幕，進入醫院的診療室與藥局的貨架上。這座橋樑何時能夠完成，似乎最終將取決於數據的品質、法規的制定速度，以及在患者身邊堅持確認安全性的人類判斷。似乎只有當這些判斷一點一滴累積起來，奈米醫學才不會只是像 COVID-19 疫苗那樣，僅在危機時刻短暫閃耀的技術，而是能成為平凡診療室裡每天使用的工具。

將目光轉向韓國國內，親自製造並銷售奈米載體的企業，其歷史也不算短。

Samyang Biopharm 於 2007 年在國內獲批的抗癌藥物 Genexol PM，是一種將 Paclitaxel 包覆在聚合物微膠粒中的藥物，被認為是世界上首款獲批的聚合物微膠粒抗癌藥物。這代表國內製藥公司接觸奈米載體已經有將近二十年的時間。最近，ST Pharm 宣布，將以自行開發的脂質奈米粒子劑型技術 STLNP (STLNP) 與信使核糖核酸加帽技術 SmartCap (SmartCap) 為主，在作為寡核苷酸委託開發暨製造服務全球第三大企業的基礎上，將事業領域拓展至涵蓋信使核糖核酸與基因剪刀治療劑的綜合 RNA 委託開發暨製造服務。

在此之前，ST Pharm 也曾與 Ewha Womans University 的研究團隊攜手合作，共同開發過脂質奈米粒子傳遞平台。這展現出原本停留在實驗室設計階段的國內奈米藥物技術，已經成長為有望實現大量生產與出口的事業。

然而，無論人工智慧將奈米粒子設計得再怎麼精密，仍然有一個尚未解決的難題。加拿大 University of Toronto 的 Warren Chan 教授研究團隊在 2016 年，收集了當時過去十年內發表的 117 篇奈米粒子論文，計算了注射到體內的奈米粒子中，實際到達腫瘤的數量。以中位數來看，這個數值僅有 0.7%。這代表注射一千個奈米粒子，只有七個能到達腫瘤，而且研究團隊表示，這個數字在此前十年裡幾乎沒有改變過。

看來，即使是在電腦中描繪得再完美的粒子，人工智慧似乎也無法完全掌握它們順著血管流動時，被肝臟與脾臟過濾、被免疫細胞捕捉的真實體內旅程。這意味著，雖然設計速度變快了，但「傳遞效率」這堵存在已久的高牆，仍然沒有明顯降低，因此，奈米醫學的下一場勝負，很有可能不是取決於繪製新的粒子，而是取決於如何將已經繪製好的粒子安全送達目的地。

大量生產的難關也不容小覷。將實驗室微流體設備每次精細提取幾毫升的奈米粒子，轉移到工廠設備開始以數百公升的規模製造時，粒子大小與藥物包覆率常常會變得不穩定。因為只要管道內流動的速度或溫度有些微改變，脂質奈米粒子的均勻度就會被破壞，業界普遍感嘆很難將實驗室數據原封不動地套用在工廠規格上。人工智慧雖然能在幾天內完成設計，但要將這些設計落實為能讓人安心服用或注射的穩定藥物，依然需要仰賴指尖的經驗與反覆的製程驗證。

第 4 部

臨床前與臨床驗證及精準醫療的革新

4.1

臨床前研究創新：動物實驗替代及生物學建模

一位被診斷出患有瀰漫性胃癌的患者，手臂上掛著點滴管。這景象看起來與一般的臨床試驗第一天沒什麼兩樣。但這個瞬間之所以不尋常，其實另有原因。首次投予這位患者的新藥候選物質，從來沒有在老鼠或猴子的身體上測試過。從設計到安全性驗證的所有過程，全都在電腦裡的人工智慧與培養皿上的類器官中完成。

醫藥專業媒體 BioSpace 報導了這場臨床試驗，稱這是全球首例僅靠人工智慧與類器官就完成臨床前驗證的案例。沒有一隻動物走出實驗室。儘管如此，藥物還是準確地送達了人類的病床邊。

在出現這樣的場景之前，製藥產業長期以來都是以不同的方式運作。在新藥候選物質投予人類之前，研究人員會藉由老鼠、兔子，有時是猴子的身體，來確認藥物的吸收、分解與毒性。數十年來，這種方式是新藥開發的唯一關卡。問題在於，老鼠的身體和人類的身体，並不是用同一套設計圖製造出來的。

就像同一把鑰匙遇到形狀不同的鎖就打不開門一樣，在老鼠肝臟中能正常分解的物質，在人類肝臟中的反應卻截然不同的情況一再發生。由於基因與蛋白質結構在物種間的差異，許多在動物身上確認為安全的藥物，真正在以人類為對象的臨床試驗中，卻引發了副作用或無法產生療效而宣告失敗。

2006年在英國發生的事件，是展現這種差距的罕見案例。一種名為 TGN1412 的抗體治療候選物質，在以猴子為對象的試驗中，即使注入了人類預期投予量500倍的劑量，也沒有出現什麼問題。然而，當六名健康的志願者被注射了遠少於該劑量的藥物後，短短幾小時內，他們全身的免疫系統失控，引發了細胞激素風暴，導致多個器官同時受損。

猴子的免疫細胞和人類的免疫細胞表面上看起來很相似，但在化學性質上卻互不相同，而那微小的差異在實驗台上沒有顯現，直到進入人類的血管中才暴露出來。這起事件發生後，學界與監管機構都痛定思痛地接受了一個事實：靈長類試驗的結果，不能直接套用在人類身上。

動物實驗結果與人類實際反應之間出現的這個落差，業界稱之為轉譯研究的鴻溝。這種鴻溝累積的結果，在數字上也明顯地呈現出來。超過90%的新藥開發專案在臨床階段以失敗告終。這些失敗大多是在候選物質進入人體後，也就是已經花費了大量金錢與時間之後才顯現出來。一種新藥要獲得批准，平均需要花費10年到15年的時間，投入的成本高達26億美元。

製藥業界將這種悖論稱為 Eroom's Law，這是將摩爾定律反過來拼寫的詞彙。它指的是電腦性能逐年提升，但將一種新藥問世所需的成本和時間卻反而不斷增加的現象。這就是為什麼長期依賴動物實驗，一直被指出是這個悖論的根源之一的原因。

減少動物實驗的呼聲，長期以來只被視為生命倫理領域的話題。減少在實驗室裡犧牲的動物數量，這樣的維權主張當然是正當的。然而，在最近幾年之間，這項討論的重心已經明顯地發生了轉移。替代動物實驗的做法，現在不僅被視為倫理問題，同時也是一個提高科學準確度的問題。如果用人類細胞與組織所製成的模型，能比老鼠的身體更準確地預測人類的反應，那麼減少使用動物的選擇，不僅在道德上是正確的，在科學上也是更好的選擇。隨著倫理正當性與科學準確性這兩個理由相互結合，臨床前研究的版圖正在被重新繪製。

這種變化不僅僅是大型製藥公司的專利。要建立並維持動物飼養設施，需要龐大的初期費用和很長的時間，但類器官培養或虛擬細胞模擬，用相對較少的資金就可以開始進行。對於小規模的生技新創公司或大學研究室來說，以前只有大型藥廠才能涉足的臨床前驗證階段，現在也為他們開啟了直接參與的道路。這意味著，決定勝負的關鍵已經不再是資本的規模，而是假設的新穎程度與數據的品質。

這種變化的第一個結，是從法律層面解開的。2022年12月，美國國會通過了《食品藥物管理局現代化法案 2.0》。這項法案廢除了幾十年來申請新藥許可時，必須提交動物實驗資料的義務條款。法律條文中所使用的措辭，也從動物試驗改為非臨床試驗，這短短的一個詞句，成為了在法律上承認人體來源細胞模型、類器官、器官晶片，以及基於人工智慧的電腦模擬為正當替代手段的依據。整個製藥產業為之轟動。動物實驗不再是唯一的標準答案，而成為了眾多選項中的一個。

即使法律開啟了這扇門，但實際將門拓寬的工作，必須由美國食品藥物管理局來承擔。食品藥物管理局在2024年率先表明了將減少動物實驗的方向，並在2025年4月

將其發展為具體的路線圖，提出在未來的3到5年內，要讓動物實驗成為新藥開發中的例外，而非標準。

食品藥物管理局計畫首先從單株抗體，也就是由單一細胞複製而成、具有相同形狀的抗體治療劑，以及其他生物製劑開始減少動物實驗，然後再將範圍擴大到化學合成新藥。在2025年初，他們也一併發布了在藥品與生物製劑審查中，如何處理人工智慧的指南草案。

在發布路線圖一年後，美國食品藥物管理局宣布已實現首年目標。該局發布了在開發抗體治療劑時減少或消除對類似猴子等靈長類動物進行測試的初步指導方針，並更新了用合成物質替代從鬩血液中提取內毒素的指南。僅通過改變這一依賴鬩血液的測試，每年就可以使超過一百萬隻動物無需進入實驗室。美國眾議院不僅如此，還已經通過了進一步的法案，要求加快實施速度。

2025年初發布的人工智能指導方針初稿的核心是根據風險等級對人工智能模型的可信度進行不同審查的理念。即使是相同的人工智能，用於初步篩選新藥候選物的模型和在不進行臨床試驗的情況下對安全性做出最終判斷的模型，也需要以完全不同的重視程度對待。對監管決策的影響越大，就越需要更加仔細地證明該模型是用什麼數據訓練的，以及當預測出錯時是通過什麼程序進行篩選的。這一基於風險的審查框架提前表明，虛擬細胞或系統生物學模型若要在臨床試驗設計中真正介入，需要跨越什麼樣的門檻。

推動這一進展的力量不是來自單一技術，而是來自兩個不同方向共同推動。一方是體外技術，即in vitro技術，用於在實驗室中培養人類細胞和組織。另一方是in silico技術，即在電腦中進行的技術，將人類細胞完整轉移到電腦內。如果類器官和器官晶片從前面在物理上模擬人體，那麼人工智能創造的虛擬細胞和系統生物學模型則從後面以數字方式模擬人體。

這兩種技術像經線和緯線一樣相互交織，彼此填補對方的不足。電腦做出的預測通過活人細胞得到驗證，而培養皿中的實驗結果又被用來精細調整電腦模型。

這一切在2020年代中期集中發生是有原因的。隨著圖形處理器性能每年大幅提升，用於訓練人工智能模型的計算能力增長到了以前難以想象的規模。同時，能在單細胞水平讀取基因活動的實驗設備價格下降，處理速度加快，用於訓練虛擬細胞的生

物學數據終於開始充分累積。

隨著製造類器官和器官晶片的細胞培養和微加工技術的成熟，能夠物理上驗證數字模型預測的實驗工具也隨之具備。如果計算能力、生物學數據以及驗證這些數據的實驗工具這三個條件不是在類似的時期同時具備，即使類器官和器官晶片變得再精巧，也不會有支撐它們的數字對應物出現。

虛擬細胞是將整個細胞完整複製並轉移到電腦中的數字孿生。研究人員從單細胞轉錄體分析、空間轉錄體學和蛋白質體學等多個層面收集數據，並堆疊這些數據，以數學方式重現該細胞的基因表達網絡和代謝途徑。單細胞轉錄體分析能讀出每個細胞使用了多少特定基因，空間轉錄體學將這些活動在組織中的確切位置繪製成圖，而蛋白質體學則測量基因信息最終產生的蛋白質在細胞中的含量。

這與飛機在真正起飛前在電腦中進行數千次模擬飛行類似。通過向這個虛擬細胞投予新藥候選物或進行人為操縱特定基因的*in silico*實驗，研究人員無需實際培養皿或動物，就能提前判斷細胞如何反應，以及是否會錯誤地附著在靶點上並引發副作用。

虛擬細胞要有用處，必須依靠系統生物學建模來支撐，能計算從單個分子運動到整個組織變化的所有過程。這種建模與統計瀏覽數據以尋找模式的機器學習性質不同。它使用人工智能與用微分方程解釋因果關係的生物物理機制模型相結合。細胞內信號傳遞和新陳代謝過程由COPASI或CellNOpt等平台再現，而多個細胞相互作用產生的運動（如腫瘤生長或免疫細胞進入組織）由PhysiCell或CompuCell3D等多細胞模擬器處理。

只有當從分子到細胞、再從細胞到組織這一計算鏈條完整後，虛擬細胞才真正成為可信賴的預測工具。

Recursion Pharmaceuticals可以作為在實際產業中實施這種方法的一個例子。該公司結合實驗室機器人自動化設備和稱為*inVivoPRINT*的機器學習平台，改變了臨床前研究本身的方式。當機器人24小時拍攝細胞和微組織並用傳感器測量數據時，深度學習模型實時讀取這些海量的圖像和數據，提取出用藥細胞顯示的毒性和效力的表型信號。

將這樣獲得的化合物信號（如分子指紋）與已知毒性物質數據庫進行比較，可以在研究初期階段提前發現在過去會在開發後期才顯現的致命性器官損傷的風險。由於能及早篩選出危險的候選物質，Recursion Pharmaceuticals在項目啟動後僅18個月就將新藥候選物推進至人類臨床試驗階段。

Recursion Pharmaceuticals並非唯一走這條路的公司。所謂的自主實驗室正在多家製藥公司和研究機構逐漸增加，這種實驗室由機器人無需人工幹預即可設計實驗、執行實驗並自行決定下一個實驗。儘管人類研究人員每天最多也只能處理數十個培養皿，但由機器人臂和人工智能相互協作運作的自主實驗室能在相同時間內測試數千種組合，並立即將結果應用到下一個實驗設計中。

得益於這間日夜不停運作的實驗室，篩選新藥候選物質的速度，已經變得比過去人們手工逐一測試的時代快上許多，簡直無法相提並論。

在 2025 年到 2026 年之間，虛擬細胞技術一口氣向前邁進了好幾步。美國的 Arc Institute 在 2025 年中期發表了第一代人工智慧虛擬細胞模型 State3。這個模型是透過學習高達 1 億 7000 萬個細胞資料，以及從超過 70 種細胞株中取得的 1 億個以上的單細胞干擾反應資料所建立而成的。Arc Institute 並未就此止步，他們還舉辦了一場名為虛擬細胞挑戰賽的公開競賽，讓各種人工智慧模型互相競爭，看誰的預測能最貼近真實的生物學實驗結果。

來自 14 個國家的 1200 多支隊伍蜂擁而至，提交了超過 300 件最終成果。學界開始將這場競賽稱為針對虛擬細胞的圖靈測試。就像原本的圖靈測試一樣，當你無法分辨是和電腦還是和人對話時，就代表電腦通過了測試；這意味著這是一場比較虛擬細胞所做出的預測是否精準到無法與真實活細胞實驗結果區分的競賽。

大型科技公司的行動也在差不多同一個時期發生。Chan Zuckerberg Initiative 首次公開了直接從生物學資料中學習而來的細胞與細胞狀態的普遍數位表現形式，也就是所謂的人工智慧虛擬細胞。緊接著，他們與 Nvidia 攜手推出了虛擬細胞平台，為大規模處理建立與培養虛擬細胞模型所需資料，奠定了良好的基礎。

該基金會還同時推出了一個名為 rBio 的推論模型，這是一款當研究人員用普通句子提出生物學問題時，能根據虛擬細胞模擬結果找出答案的工具。Google DeepMind 也表示正在進行自己的虛擬細胞專案。虛擬細胞現在已不再是單一公司的賭博，而

是多家大型科技公司同時投入的競爭舞台。

無論電腦裡的虛擬細胞變得多麼精密，都無法原封不動地取代活人體內 3D 的生理現象。因此，我們需要一個實體夥伴來實際驗證這些預測。這個角色就由類器官來擔任。因為它不是用老鼠細胞，而是用人類細胞製成的，所以這就如同不經過翻譯，由母語人士親自開口說話一樣。類器官是將從患者身上取得的成體幹細胞，或者是將已經完全長成的皮膚細胞等重新恢復到幼年時期萬能狀態的誘導性多功能幹細胞，在培養皿上進行立體培養，從而打造出具備與真實器官相似的細胞組成和結構，甚至具有一定限度生命功能的縮小版器官。

類器官能將平坦的 2D 培養皿或動物身體無法模仿的人類特有的細胞間訊號傳遞與組織微環境，如實地展現出來。作為一種能事先在實驗室裡確認新藥候選物質對人體實際效果的工具，類器官的用途正迅速地擴展當中。

在面對阿茲海默症和帕金森氏症等神經逐漸崩壞的疾病時，類器官展現出了特別大的用處。包含老鼠在內的齧齒類動物大腦，無法完全模仿人類大腦特有的皺褶皮質結構、複雜的神經膠質細胞網絡，以及蛋白質糾結在一起的病理現象。因此，長期以來齧齒類動物的大腦一直被冠上中樞神經系統新藥開發墳墓這樣不名譽的稱號。用人類細胞製成的迷你大腦與中腦類器官則有所不同。它們能夠以 3D 的方式，將帕金森氏症的核心病理，也就是製造多巴胺的神經細胞退化，以及名為 Alpha-synuclein 的蛋白質不正常聚集的現象，極為相似地重現出真實患者大腦的狀況。

最近，甚至還出現了將透過光來調節細胞活動的光遺傳學技術與中腦類器官結合，藉以精確控制 Alpha-synuclein 聚集過程的實驗模型。藉由這個模型，研究人員能夠在人類生物學的脈絡下，直接確認保護神經的新藥物候選物質的效果。

在開發罕見疾病治療藥物時，這些技術的分量就顯得更加吃重。由於患者人數稀少，連動物模型都無法妥善建立的疾病不在少數；在這種情況下，直接從一名患者的細胞中培養出來的類器官，事實上就成為了唯一一種臨床前驗證手段。相較於耗費數年時間來建立並驗證新的動物模型，我們能夠從患者的一個組織樣本開始，在幾週之內培養出類器官並測試藥物反應。這幾週的時間差，對罕見疾病患者的家屬來說是具有決定性意義的。

類器官也有其不足之處。它缺乏血管網絡，缺少與免疫系統的相互作用，也無法呈現體內組織實際承受的物理力量。為了填補這個空缺，結合了微流體工程的器官晶片，也就是微生理系統（MPS），站上了臨床前研究的最前線。在這個領域裡，除了 Emulate 之外，還有像是 Mimetas、TissUse 以及 Hesperos 等公司，也各自以不同的器官為目標投入其中。器官晶片是在拇指大小的透明聚合物晶片上，刻劃出比頭髮還細的流體通道，並在其中將源自人類的組織細胞與血管內皮細胞一起培養出來。

這一小塊晶片原封不動地搬演了人體內發生的力學環境，包含了血流流動時所產生的壓力、腸道蠕動的動作，甚至是肺部膨脹與收縮的節奏。器官晶片所重現的不僅是細胞的形狀，還包含該細胞實際在體內經歷的物理條件，也就是完整的壓力、流動與動作。

每一塊器官晶片並不只侷限於模仿肝臟、心臟、腎臟或肺臟等單一特定器官。如果將好幾個器官晶片用細管連接起來，就能夠在培養皿外連續觀察藥物被身體吸收、代謝直到排出體外的整個過程，也就是完整的藥代動力學。舉例來說，如果將腸道晶片與肝臟晶片並排連接，就能夠按照人體內真實的順序，重現口服藥物在腸道被吸收後，立刻進入肝臟進行分解的過程。

器官與器官之間的這種連接順序是單個器官芯片無法獲得的信息。將腦血管內皮細胞、星形膠質細胞和血管周圍細胞共同培養的血液-腦屏障芯片，能比動物實驗更準確地發現人工智能設計的腦部疾病治療藥物實際穿透環繞大腦的屏障的程度。藥物是否會對肝臟造成損傷或引起心臟跳動異常，由人類細胞製造的器官芯片上放置的多個微細電極實時捕捉。

這些電極能夠以秒為單位讀取細胞發出的微細電信號變化。對於心臟細胞，能立即捕捉到跳動節律混亂的瞬間；對於肝細胞，能立即捕捉到細胞膜受損時滲漏出的信號。當電腦中虛擬細胞做出的預測不確定時，正是這個物理芯片上的數據消除了那種不確定性。計算機模型先指出方向，體外試驗驗證那個方向是否正確，這個循環就這樣完成了。

肝臟之所以特別成為新藥開發的瓶頸，是有原因的。藥物大多經過肝臟分解代謝，在這個過程中經常會產生預想不到的副產物，導致肝細胞受損。藥物引起的肝臟損

傷一直被視為上市後藥物從市場撤出的常見原因，而且在動物實驗中表現完好的物質在人類肝臟中出現問題的情況也並不罕見。

這項技術越過實驗室圍籬在監管現場獲得認可是最近的事。美國食品和藥物管理局在2025年中期將審查創新科學技術方法的通道ISTAND制度正式確立。通過這個通道，埃繆萊特公司的肝芯片正在踏入作為預測藥物引起肝臟損傷工具而獲得認可的最後階段。它正在接受驗證，確認僅通過毒性物質暴露組織的形態變化以及白蛋白、ALT蛋白質數值的波動就能預測實際人類肝臟將發生的損傷。獲得這項資格後，肝芯片將不再是特定一家公司的產品，而是在規定用途內所有新藥開發項目都能使用的經認可工具。

這裡規定的用途，即使用背景，是規範監管語言，明確這個芯片能用來回答哪些具體問題。即使因預測肝臟毒性而獲得認可，也不會自動獲得預測心臟毒性或腎臟毒性的認可。器官芯片產生的數據從個別實驗的參考資料升級為監管審查的正式證據，這個首例就在眼前。

器官體也正在從實驗室轉移向患者身邊。僅在2023年至2025年間，使用患者來源器官體進行的臨床試驗就有139項註冊，這股趨勢在2026年仍在延續。針對無法切除的肝細胞癌患者、提前用器官體測試藥物反應來確定治療方向的二期臨床試驗在2026年初啟動。即使被診斷為同樣的肝細胞癌，每位患者的腫瘤遺傳特性也不相同。這項臨床試驗的方式是先從患者個人腫瘤中提取器官體，在培養皿中觀察它對多種藥物的反應，然後為這個人開具對其器官體有效的藥物。

這與先使用藥物、無效則更換的傳統試錯式處方完全不同順序。器官體現在不再僅停留在實驗室選擇新藥候選物質的工具，而是已經進入決定對特定患者使用已經批准的藥物中哪一種的地步。

韓國的舉動也不例外。韓國食品藥品安全部及其下轄的韓國動物替代試驗法驗證中心正在推動將器官體、器官芯片、人體組織模型等非動物試驗法提升至國際標準，並訂下2028年前納入經濟合作與發展組織國際認可試驗法的目標。國內的器官體科學和新智生物等創業公司也在將器官體基礎新藥評估平台商業化，與這股潮流同步發展。

2026年韓國食藥監規範科學學會春季學術大會特設了處理器官體基礎動物替代試驗法的專題論壇，國內研究者與監管部門正在共同完善這些標準。

儘管取得所有這些進展，仍有許多未解之結。由數千萬個變量構成的人工智能模型在判定某化合物危險的同時，往往無法用人能理解的語言解釋為什麼得出這種判斷。就像答案正確但不寫解題過程的學生一樣。由於監管部門批准新藥時要求的不僅是結果，而是達成那結果的邏輯，所以無論預測多麼準確，不能解釋根據的預測也很難通過審查門檻。

在相似於用於訓練的數據的化合物中表現準確度高的模型，一旦遇到結構完全陌生的新物質或來自不同人群的生物學數據，預測力就會突然崩潰，這種情況也在反覆發生。最近多項研究還指出，由於訓練過程中答案相應的信息誤混入造成數據洩露，會產生看似遠比實際情況準確得多的假陽性結果。

技術之外還存在其他問題。要交換和重新驗證不同研究機構或公司製造的虛擬細胞模型，需要有共同語言，但SBML或CellML之類的標準模型技術語言未能廣泛普及，導致數據和模型仍然分散各處。要恰當地訓練虛擬細胞，需要數十萬人的基因組信息和就醫記錄，這與歐洲的個人數據保護法或美國的健康保險信息保護法直接衝突。

不將原始數據輸出到機構外而只交換加密結果的聯邦學習等替代方案正在討論中，但支撐這些方案的制度和法律條款在許多國家仍未建立。

目前所累積的類器官細胞株與生物學數據，大多集中在特定的地區與人口群體，因此若將這些數據訓練出來的模型，直接應用於遺傳背景不同的患者身上時，其準確度究竟有多高，目前還很難有把握地給出答案。當人工智慧與微生理系統共同做出的判斷，對實際患者造成了意想不到的傷害時，究竟該由製作軟體的公司負責、提供數據的醫院負責，還是由做出最終決定的製藥公司負責，目前也還沒有制定出劃分責任的法律標準。

儘管如此，那位胃癌患者手臂上插著的點滴管，卻已經開啟了另一段故事。沒有經過老鼠或猴子實驗的藥物，已經直接進入了人體的靜脈。如果是在過去，單是動物實驗階段可能就要花費好幾年的時間。這款藥物在未來的幾年內究竟能拯救多少患者，目前還沒有人可以確信。臨床試驗現在才剛邁出第一步。不過，用電腦伺服器

與培養皿來取代實驗室裡的老鼠籠，成為新藥開發第一道關卡的這股趨勢，已經發展到了難以回頭的地步。

在韓國，為了順應這股趨勢，也正持續嘗試制定新的法律。先前雖然 Nam In-sun 議員與 Han Jeong-ae 議員各自提出了 Dongmuldaechesiheombeop 制定草案，但由於 Nongnimchuksansikpumbu、Gihueneojihwangyeongbu、Sikpumuiyakpumanjeoncheo 以及 Nongchonjinheungcheong 之間的立場差異，兩項法案最終都未能跨過國會的門檻而宣告廢止。2025年 12月，Song Ok-ju 議員再次代表提案了這項法案。由於這次是相關部會事先協調好內容後才提出的草案，因此有評價認為，這次的情況與 Nam In-sun、Han Jeong-ae 議員當時有所不同。

法案內容包含政府部會將共同組成協議體並制定基本計畫，並設立按領域驗證 Dongmuldaechesiheombeop 的中心，每 5年 進行一次實地調查。隨著法案的提出，也有一些數字隨之公布。2024年 一整年，韓國國內使用的實驗動物達到了 459 萬 隻，其中超過一半的實驗被分類為痛苦程度最高的等級，這些實驗在不使用麻醉劑或止痛藥的情況下，讓動物暴露於有毒物質中或是引發腫瘤。

類器官所留下的缺口不僅僅是血管與免疫系統。中腦類器官雖然能像人腦一樣展現出製造多巴胺的神經細胞的每一個活動，但目前還無法重現那些神經細胞延伸出長軸突至其他大腦區域以傳遞訊號的迴路。人類的大腦是透過大腦皮質、海馬迴以及基底核互相傳遞訊號，並作為一個單一迴路來運作的，但是在培養皿上單獨生長的一個類器官中，就缺乏了這種長距離的連結。

最近，出現了一些將多個類器官在物理上連接起來，製造成稱為組裝體的結構體，並利用電極來測量在其中穿梭的神經訊號的研究。這個方向是正確的，但目前仍處於起步階段，像是儲存記憶或是引發情緒這種需要整個大腦一起運作才會出現的現象，從一開始就無法在培養皿上觀察到。要完整重現人類大腦這種程度的複雜迴路，仍然是一個遙遠的目標。

減少動物實驗的要求與確實保障人類安全的要求，在未來似乎依然很難輕易達成和解。TGN1412 在猴子體內看起來很安全，卻在人體內引發免疫系統失控的事件，也同時向我們展示了，單靠類器官或器官晶片，仍然有無法完全防範的風險存在。整個免疫系統互相傳遞訊號所產生的反應，或是多個器官同時經歷的慢性毒性，目前

仍然是現有非動物試驗方法難以趕上的領域。

在累積足夠的證據能有把握地過濾掉這些風險之前，比起透過法律一次性全面禁止動物實驗，從風險較低的領域開始逐步接手，很有可能是更為現實的一條路。

4.2

數位雙生 (Digital Twins) 與次世代臨床試驗的現代化

美國食品藥物管理局會議室的螢幕上，一顆人類的心臟正在紅光中跳動。這不是一顆真實的心臟，而是由法國軟體公司達梭系統（Dassault Systèmes）製作並提供的數位複製品，也就是所謂的「活力心臟」（Living Heart）。這個始於2014年的專案已經完成了世界上首個精確複製人類心臟的三維數位心臟，隨後與美國食品藥物管理局簽訂了為期5年的共同研究協議，並被採用為驗證心血管醫療設備的正式虛擬測試環境。大約同一時期，在另一間擺著文件的會議室裡，人工智能新創公司Unlearn.ai提出的建議書正被審查官握在手中。

這份建議書的內容是：用電腦計算產生的虛擬患者資料，來填補不服用新藥的患者群，也就是對照組。在藥物開發商針對臨床試驗設計與規管當局進行的正式C型會議上，這個建議被接受了。這是規管當局首次認可：在臨床試驗中，一半的患者可以不是真人，而是由電腦生成的患者。

傳統臨床試驗至今仍是新藥批准的最後關卡，但通過這道關卡的路途變得越來越艱難。全球80%以上的臨床試驗因為無法及時招募患者而延期，整個製藥業界每年投入的研究開發費超過2,000億美元。儘管投入如此之大，最終獲得批准的新藥還不到12%。那些沒有接受新藥、只服用安慰劑也就是假藥的對照組患者，喪失了最佳的治療機會。隨機對照臨床試驗需要跟蹤數千名患者數年之久，才能得到統計上有意義的結果，這種框架對患者和預算都造成了巨大的負擔。

在減輕這種負擔的努力中，利用人工智能和海量真實世界數據來重新設計臨床試驗的數位孿生和臨床試驗模擬技術正在成為完成精準醫療的驅動力。減少註冊和追蹤一名患者所需的時間和成本的工作，如今已經不是幾家公司的選擇，而接近於整個新藥開發行業的生存條件。歸根結底，這一變化的目的不止於讓一種新藥更快地進入市場，而是為了讓這種新藥以最需要的方式到達真正需要它的患者。

新手飛行員在真正升上天空之前，首先要坐進飛行模擬器。飛行模擬器螢幕上的天空是虛假的，但引擎的聲音、儀表板的震動、強風推動的感覺幾乎與真實飛行一模

一樣，所以飛行員可以在搭載真正乘客之前盡情地犯錯練習。數位孿生就是為每個人創建這樣一個模擬器的技術。將患者的遺傳信息、醫院就診記錄、手腕上智慧手錶測量的心率和睡眠、影像檢查結果、血液中的蛋白質水平等全部輸入電腦，就會產生一個幾乎與該人完全相同方式運動的數位複製品。

醫生和研究人員不再是先將未經驗證的新藥給真實患者服用，而是先給這個數位複製品用藥測試。在提高劑量時會有什麼反應，副作用是從肝臟開始還是從心臟開始，這些都要先在電腦內確認，然後才進行實際處方。就像在模擬器內犯幾十次錯誤都不會傷害真實乘客一樣，經過在數位複製品內測試數百次之後，新藥才會來到實際患者身邊。

類似的嘗試以前也有過。但過去的模型只限於計算某種藥物在整個人口群體中平均如何被吸收和代謝，與個別患者的實時數據沒有連接。它只不過停留在驗證靜態假說的計算器水平上。隨著最近深度學習的引入，這個計算器已經進入了完全不同的層次。現在被稱為「矽基孿生」（in silico twin）的技術，因為它在電腦內運行，可以24小時實時接收來自可穿戴設備、電子健康記錄、高解析度影像、以及同時檢視基因和蛋白質的多組學數據，自我學習患者體內疾病如何進展，並持續更新。

單一傳感器每天產生的數據就達數萬條，因此過去只有到醫院才測量一次的指標，現在幾乎可以每時每刻查看。這意味著昨天的數位孿生和今天的數位孿生可能會給出不同的預測，也意味著它更緊密地跟蹤了實際患者的身體狀況。

在精準腫瘤學這樣的領域，每種腫瘤的性質都完全不同，只有篩選出可能對特定靶向治療藥物有反應的患者，才能決定臨床試驗的成敗。在沒有這種分層的時代，藥物試驗會將反應者和無反應者混在一起測試，結果導致真正有效的新藥往往因為被平均值淹沒而被判為無效。人工智能分析累積在數位孿生中的基因和病理數據，找到具有生物標誌物的患者亞群，也就是擁有能顯示體內變化的特定物質的患者，這些患者極可能對該藥有反應。

正是因為能夠精確到單個基因變異、單個蛋白質水平，即使一種藥物只對十個人中的一個有效，也能準確地找到那一個人，並將臨床試驗引向成功。其結果是，臨床試驗的設計者可以大幅減少所需的患者人數，同時反而能讓證明藥效所需的統計信心變得更加堅實。這裡產生了一個悖論：招募更少的患者，卻獲得更可信的結果。

在挑選患者這項工作的另一面，建立對照組的工作正如同鏡子般與之相對。在現實世界中，人工智慧會篩選出對新藥有反應的患者，但在數位世界裡，同樣的人工智慧會創造出與該患者在統計學上互為雙胞胎的虛擬對照組。其材料是過去許多醫院累積的臨床試驗紀錄、全球疾病登錄檔，以及電子健康紀錄等世界上已存在的真實世界數據（RWD）與真實世界證據（RWE）。研究團隊會以這些數據為基礎，在電腦中重現完全不使用新藥時的病情發展，並將該結果與實際接受新藥的患者病程並列比較。

由人工智慧篩選實際患者數據以建立治療組的前期工作，與由人工智慧篩選真實世界數據以建立虛擬對照組的後期工作，是在兩側重複相同原理的結構。挑選患者的眼睛與創造對照組的眼睛，最終都是同一個人工智慧，這項事實也意味著，每次設計臨床試驗時，都必須同時計算現實世界與數位世界的時代已經來臨。

隨著數位雙生進入臨床試驗，製藥公司已經能在招募實際患者的很久以前，預先估算出一項臨床試驗的成功機率。在電腦中先將開發中新藥的作用機制與目標患者群的數位雙生數據進行比對，並在臨床試驗開始前就篩除掉可能會失敗的設計，這道程序近似於一種虛擬預演。一旦一項 3 期臨床試驗以失敗告終，數年的時間與龐大的費用就會直接付諸流水，因此這項預演所篩除的每一個失敗，從整個業界來看，都會帶來節省大量資源的效果。

雖然在電腦中的成功無法直接保證在實際患者身上的成功，但光是從一開始就篩除掉可能性較低的設計，就足以幫助提升新藥開發的成功率。由於人工智慧從設計臨床試驗的階段就開始參與，因此還有一項優勢，就是能預先修正那些在過去必須等到患者招募結束後才會暴露出來的設計漏洞。

Unlearn.AI 開發的預後共變異數調整（PROCOVA, Prognostic Covariate Adjustment）技術，是將數位雙生所預測的個別患者預後分數放入統計模型中，藉此縮減對照組規模的方法。與過去未反映患者個人預後，僅以平均值進行比較的統計方式不同，這項技術會先將每位患者不同的恢復速度與危險因子納入計算，接著再篩選出新藥純粹的效果。歐洲藥品管理局（EMA）提出了可於 2 型與 3 型臨床試驗中使用此技術的資格認可意見，這是歐洲監管當局首次以這種方式認可基於人工智慧的方法論。

美國食品藥物管理局也在多次的 C 型會議中，對這項技術提出了正面的意見。根據 Unlearn.AI 公布的數據，這能將治療效果估算值的變異數減少 10% 到 20% 左右，這會帶來將對照組所需患者人數減少 20% 到 30%，甚至更多的結果。將實際臨床試驗數據與數位雙生創造的虛擬數據結合使用的混合型臨床試驗，現在已經不再是例外，而是逐漸確立為標準。

除了 Dassault Systèmes 與 Unlearn.AI 之外，Nova In Silico、Deep Intelligent Pharma 等公司也紛紛投入這個市場，使競爭變得更加廣泛。

業界開始將這種將一半的實際臨床試驗與數位雙生創造的另一半合而為一的設計，稱為雙生臨床試驗（TwinRCT）。顧名思義，這是一種一半從人身上，另一半從那個人的數位雙生身上尋求答案的方式。Unlearn.AI 將這個名稱作為自家技術的正式名稱提交給監管當局，而這種命名方式也是一種試圖將尚屬陌生的概念，轉換為臨床現場語言的嘗試。

雙生臨床試驗現在正逐漸跨越少數罕見疾病的研究，蔓延至常見的慢性疾病領域。在設計一項臨床試驗時，該如何分配人與數位雙生的比重，似乎將成為未來臨床試驗設計者們將面臨的新問題。

在罕見疾病領域，這項技術的用處更加顯著。如果是一種就算統整全國，患者人數也僅有數百人的疾病，那麼一開始就沒有餘裕將其平分為安慰劑組與治療組，而這種分類本身在統計學上也是強人所難的要求。在這種疾病中，由數位雙生創造的合成對照組，實際上被視為唯一現實的替代方案。在父母難以爽快同意將子女分配至安慰劑組的兒童罕見疾病領域中，更是如此。

許多國家的罕見疾病新藥制度會考量到這種情況，經常彈性地適用對照組的條件。對全體臨床試驗參加者僅投予新藥的單臂臨床試驗，事後再附加上由數位雙生創造的對照組，這種方式也呈現逐漸增加的趨勢。業界專業雜誌也指出，罕見疾病新藥開發正透過數位雙生與虛擬臨床試驗迎來新的轉機。

對這些患者額外設置安慰劑組，在倫理上往往很難被正當化。讓病情快速進展的患者只服用假藥並等待數年，如果新藥有生效的可能性，這將是更令人難以接受的選擇。數位雙生創造的合成對照組為這些患者開啟了獲得新藥而非安慰劑的機會，同時也能原封不動地確保臨床試驗所要求的比較資料。這種問題意識本身並不新奇。

甚至在人工智慧出現以前，僅憑過去紀錄所建立的歷史對照組就經常捲入可靠性的爭議中，而數位雙生則近似於一種試圖以更精密的統計來彌補那場古老爭論的嘗試。

然而，合成對照組並不能隨便直接使用。根據過去數據製造的虛擬對照組已經有可能與現在的診療環境不同，如果標準治療在幾年間有所改變，過去的數據就難以作為今天的比較基準。因此，美國食品藥物管理局在2023年發佈的指南中闡明了逐案審查外部對照組的原則，並將其使用限制在沒有適當替代治療的重症疾病且具有客觀可測量評估指標的情況下。

適應性臨床試驗是在臨床試驗進行過程中，根據患者反應數據動態調整給藥劑量或目標患者群的方法。強化學習演算法讀取中期分析結果，為某些患者增加劑量，對於無反應或副作用風險大的患者則早期排除出臨床試驗。CURATE.AI針對固體腫瘤患者的預試驗是實際演示這種方式的案例。在用於胃腸癌患者的標準化療方案XELOX或XELIRI中，其中所含口服化療藥物卡培他濱（Capecitabine）的劑量根據患者個人反應數據持續調整。

將一名患者作為一項獨立臨床試驗來處理的這種方式稱為N-of-1，它與傳統處方不同之處在於，不是將患者納入標準用量表，而是實時尋找適合患者的用量。這項預試驗的結果在2025年發表於國際學術期刊npj Precision Oncology，確認了為每位患者量身調整用量是實際可行的。

美國食品藥物管理局在2025年1月發佈的指南草案採取了基於風險的方法，既不無條件信任也不無條件懷疑個別人工智能模型，而是根據該模型的具體使用方式和用途來調整審查強度。如果數位雙生僅用作確定初始劑量的參考資料，相對較輕的驗證就足夠了，但如果用作完全替代安慰劑的合成對照組，則必須通過更嚴格的驗證。新藥開發公司在提交模型之前，必須將該模型用何種數據、何種結構建立，以及未用於訓練的新患者身上是否表現相同等內容整理成文件，並與監管機構協商。

雖然程序繁瑣，但反過來看，這也開啟了人工智能產生的證據能夠與人類產生的證據在同一水平上受審查的道路。這種方法與認證汽車或飛機零件時為每個零件分配不同安全等級的方式差異不大，因此對於監管機構來說並非完全陌生的框架。

這種趨勢並非僅限於少數公司的實驗，從各地正在進行的其他臨床試驗也可以看出這一點。美國克里夫蘭診所正在進行一項隨機對照臨床試驗，以驗證針對第2型糖尿

病患者的全身數位雙生基礎精準治療的效果，結果預計在2026年內推出。在2025年4月，也新啟動了一項驗證數位雙生基礎臨床研究系統本身有效性的個別研究。市場研究機構預測，數位雙生相關全球市場規模將從2024年的177億美元增長到2032年的2,593億美元，每年以接近40%的速度增長。

包括法國製藥公司賽諾菲（Sanofi）在內的大型製藥公司也開始通過自己的媒體公開介紹利用人工智能生成的數位雙生重新設計臨床試驗的工作。隨著協助患者直接追蹤自身治療進展的平台（例如Outcomes4Me）的出現，數位雙生已超越研究人員的工具，成為患者自己使用的工具。

如果合成對照組減少了對照組所需的人數，可穿戴設備則進一步減少了剩餘參與者前往醫院的次數。穿著智能手錶和智能貼片等可穿戴設備，即使患者不前往醫院，心率、睡眠和活動量等數位生物標誌也會24小時累積。這種不受醫院物理位置限制的去中心化臨床試驗（DCT）估計到2026年全球市場規模將超過100億美元。

美國政府生物醫學高等研究開發局（BARDA）與大型連鎖藥局沃爾格林（Walgreens）合作，將社區藥房作為臨床試驗據點，這也是降低醫院門檻的同一趨勢的一部分。美國食品藥物管理局也在2024年4月開設了臨床試驗創新中心，並在同年9月確定了去中心化臨床試驗指南，為這一趨勢增添了監管支持。

這些源源不斷的數據由大型語言模型讀取。人工智能實時讀取非結構化的臨床記錄和患者留下的文本，以超過90%的敏感度檢測不良藥物反應或藥物間相互作用。有研究團隊從報告中自動提取流感疫苗的不良反應，也有研究團隊使用應用Tree of Thoughts推理的大型語言模型從臨床記錄中自動識別心血管不良反應。

通過這種方式，難以前往醫院的農村患者和少數族裔患者也可以參與臨床試驗，因此去中心化臨床試驗不僅解決效率問題，也解決了臨床數據的公平性問題。這種自動檢測相比患者直接輸入症狀的方法的優點在於，它不會遺漏即使患者本人認為不重要而未報告的初期信號。

2024年下半年，Dassault Systemes與FDA合作，出版了全球第一本以ASME的驗證標準V&V 40為基準的in silico臨床試驗指南。Dassault Systemes表示，這項由醫師、工程師與政府相關人員共同合作展開的計畫，已於2025年2月進入了下一代Living Heart模型的Beta測試階段，該模型可以針對每一位患者或特定患者群體進行微調。

繼Living Heart之後，Living Brain也已推出，現在相同的模式正在擴展到包含肝臟和肺臟在內的幾乎所有體內器官。

Living Heart計畫也登上了2025年美國財經雜誌Fast Company所選出的改變世界創意名單中。國際學術期刊Nature Medicine在2026年刊登了一篇論文，指出數位雙生與in silico臨床試驗已進入了新藥開發的轉捩點，而以數位雙生來輔助臨床試驗的研究，也於同一時期發表在學術期刊npj Systems Biology上。

在華麗的成果背後，仍殘留著沉重的絆腳石。首先遇到的問題是，我們無法看透人工智慧做出判斷的內部過程，這也就是所謂的黑盒子問題。由數千萬個參數交織而成的深度學習模型，即使得出特定治療藥物有效的結論，或是警告將會發生肝毒性，如果無法向醫師與監管機構明確解釋該判斷是經過何種生物學路徑得出，就很難在新藥許可審查中被採納為核心依據。

因為關係到患者生命權益的決定，我們不能任由電腦在無人能解釋其判斷原因的狀態下做出。在醫師之間，與其直接遵從電腦給出的結果，他們依然強烈傾向於必須在理解原因後，才願意在處方箋上簽名的態度。因此，最近有一種將最終預測值與幾個對該預測產生重大影響的因素並列呈現的「可解釋的人工智慧」技術，正成為一種現實的替代方案。

訓練數據的侷限性也是一個阻礙。利用偏向特定人種或特定疾病模式的數據來進行訓練的模型，一旦遇到陌生的新藥候選物或其他人群時，其預測準確度就會大幅下降，學界中也有報告指出準確度甚至降至0.3的案例。在訓練過程中，原本應該稍後才出現的臨床結果資訊不小心混入了輸入數據中，導致模型的性能指標被誇大得比實際更高，這種數據外洩問題也屢次受到指責。這類錯誤在臨床試驗階段往往不易被發現，但到了實際開立處方的現場，就會有風險出現將不存在的危險錯誤報告為存在的偽陽性判定。

基於這樣的理由，最近監管機構開始要求在新藥提交資料中，必須包含在完全未用於模型訓練的新患者群體中另外確認性能的結果，也就是所謂的樣本外驗證。因此，學界與製藥界都要求將涵蓋從軟體單元測試到與生物學反應一致性驗證在內的嚴格驗證程序，作為新藥開發的必經關卡。

為了打造精密的數位雙生，必須大量收集如基因體資訊與電子病歷等極度敏感的個人資料。在美國的HIPAA或歐洲的GDPR規範下，醫院與製藥公司將各自擁有的資料轉移至外部伺服器並集中在一起的做法，實際上是受到阻礙的。作為替代方案出現的，是讓患者的原始資料保留在醫院內部伺服器中，僅交換加密後的模型權重的聯邦學習與差分隱私技術。在跨越國界的跨國臨床試驗中，在一個國家合法的資料處理方式，到了另一個國家卻直接變成違法的情況也很常見，這使得負責設計國際共同研究的法務人員的負擔正日益加重。

要將不同醫院的資料格式統一起來，並維持如鐵壁般的網路安全水準，所花費的成本依然不容小覷。因此，大型醫院現在也開始嘗試一種方式，那就是不直接交換彼此的原始資料，而是透過可信賴的第三方機構，僅將完成匿名化驗證的資料用於研究中。

製藥界內部正深刻思考著一個問題，那就是在人工智慧演算法中，該如何以文件記錄要求資料完整性的ALCOA++原則，亦即追蹤是誰在何時建立該資料的歸屬性、準確性、完整性與一致性等標準。除此之外，基於人工智慧設計的虛擬患者與合成對照組而獲准上市的新藥，如果在上市後引發了預期之外的副作用，其責任究竟該歸咎於製作軟體的企業、提供資料的醫院、開立處方的醫師，還是蓋章批准的監管機構，目前仍沒有任何國家制定出明確釐清此問題的法律標準。

目前，新藥開發公司只是透過在合約中預先劃分並註明責任歸屬的方式來填補這個法律漏洞，在法院判例或另行立法確立明確標準之前，這種不確定性很有可能會持續一段時間。雖然多國政府正在檢討相關立法，但要找到已經超越法案討論階段並進入實際實施的國家，目前依然很困難。

美國國會在2022年與2023年期間通過了FDA Modernization Act 2.0，廢除了數十年來作為新藥許可前提條件的動物實驗義務條款，並正式認可基於人工智慧的in silico模擬資料，作為證明安全性與有效性的替代手段。2025年1月，FDA發布了一份草案，規範在對藥品與生物製劑進行監管判斷時使用人工智慧所必須遵守的指導原則，要求必須預先定義模型的用途，並驗證資料管理、模型結構，甚至在未用於訓練的資料上的性能。

同年12月，監管機構最終確定了基於真實世界證據的醫療器械審查指南，表示將接受匿名化的大規模醫療數據庫作為證據資料，即使沒有能夠識別個別患者的原始數據。歐洲藥品管理局在2024年9月採納了關於整個藥品生命週期內人工智慧應用的反思文件，建議製藥業界透明地解釋模型的結構，以及由人類對結果始終進行責任和監督。2025年7月，將有關外部對照組的單獨指南草案提交公開徵求意見。英國藥品健康管理產品監管局(MHRA)也在2025年5月提出了基於真實世界數據的外部對照組指南草案。

韓國食品藥品安全處也在2025年12月制定了一份指南，其中包含數位設備收集的臨床試驗資料的認可標準。2026年1月，美國食品藥物管理局與歐洲藥品管理局共同發表了十項在新藥開發中應用人工智慧的通用原則，開始協調大西洋兩岸的標準。這在很大程度上相當於在製藥業已經熟悉的良好臨床試驗管理標準 -- 即所謂的GCP -- 傳統上再添加一個人工智慧版本。

許多人認為韓國還有很長的路要走。醫療人工智慧專家指出，即使到了2026年，國內醫療現場使用數位雙胞胎的案例仍然不足，必須先建立支撐信任的監管體系。在美國和歐洲，監管機構和企業經過5年、10年的數據積累，才最終提出了資格認可和指南草案。這種差距與其說是預算或意願的問題，不如說更像是需要時間積累信任的問題。但我們不能讓那段時間白白流逝。

那些數據充分、驗證程序完善的醫院和企業若能先獲得監管機構的信任，其經驗成為後續案例的審查標準，就很有可能縮小這種差距。由數位雙胞胎和合成對照組替代臨床試驗一半的混合模型現已進入監管機構的正式文件中。但要將其作為所有新藥開發的標準做法，還需要時間逐一填補可解釋性、數據公平性和責任歸屬這些要素。

領導春川新創公司Offremed的執行長Hong Suji將外國數位雙胞胎公司的虛擬患者模型大多由白人數據構建這一點視為公司技術的差異化特色。Offremed開發的以人工智慧為基礎的人類數位雙胞胎OPTIVIS針對包括韓國人在內的亞洲患者群體的數據進行了單獨訓練，在全球罕見地創造出亞洲虛擬患者。Offremed在2026年7月於英國倫敦舉辦的阿茲海默症協會國際學術會議(AAIC)上首次公開了該臨床試驗數位雙胞胎平台。

在開發一種新藥平均需要20億美元和10至15年的當今，Offremed表示使用OPTIVIS可以將時間和成本平均減少約30%。Offremed與春川企業創新園區簽訂了業務協議，計劃在那裡新建研究據點。關於以西方人數據為主的累積數據可能不太適合亞洲患者的擔憂，隨著數位雙胞胎在全球擴展，這看起來會成為包括韓國在內的許多國家都會面臨的同樣問題。

實際臨床試驗結果和用真實世界數據模擬的結果並不總是相同的。哈佛醫學院附屬Brigham and Women's Hospital的Jessica Franklin及其同事嘗試僅用真實世界數據重新複現已完成的十項隨機對照臨床試驗。得出與監管判斷相同結論的情況在十項中只有六項，而風險比估計值落在原始臨床試驗信賴區間內的情況有八項。

這項結果於2021年發表在美國學術期刊Circulation上。臨床試驗設計與原始研究的再現接近程度不同，兩個結果之間的準確度有很大差異。在廣泛用數位雙胞胎和合成對照組替代隨機對照臨床試驗之前，逐一辨別在哪些條件下結果會產生偏差似乎更為安全。

最需要數位雙胞胎的患者，其數據反而最稀缺。罕見疾病患者和少數族裔患者過去在臨床試驗中參與不足，因此在用於訓練數位雙胞胎的數據中代表性也不足。學界正在嘗試用統計方法補正患者數量不足，或用人工智慧生成新資料來填補空缺，但無法完全彌補原始資料本身不足的局限。那些在過去臨床試驗中被邊緣化的患者，在新技術中再次被邊緣化的風險，除非積累新數據，否則不會自行消失。

4.3

精準腫瘤學 (Precision Oncology) 與病患客製化治療

由以色列 Sheba 醫療中心與美國 MD Anderson 癌症中心等機構所組成的國際 WIN 聯盟，其所進行的 WINTHER 臨床試驗是一項大型跨國研究。該研究招募了超過三百名的晚期癌症患者，同時採集他們的腫瘤與正常組織，並同時讀取其基因序列與基因表現資訊。在對多發性骨髓瘤與乳癌患者投予名為酪氨酸激酶抑制劑的標靶抗癌藥物時，研究團隊利用電腦，分析了從患者癌細胞中採集到的每一個單細胞的基因活動。患者被分為兩種方式來分配治療藥物：一組透過讀取二百三十六個基因的 DNA 序列分析來分配；另一組則透過比較腫瘤與正常組織基因表現的 RNA 分析來進行分配。

過去只鎖定單一特定基因來挑選藥物的舊方法，早已遇到了瓶頸。將每位患者不同的基因資訊與病理發現結合在一起，找出專屬於該患者治療方法的個人化醫學，正逐漸成為癌症治療的重心。在 WINTHER 試驗中，由能夠同時計算基因、蛋白質與代謝資訊以找出最佳治療藥物的多體學配對演算法，負責將患者與藥物進行配對。腫瘤並不是靜止不動的。追蹤在人體內無時無刻都在改變樣貌的腫瘤，正是這項臨床試驗所提出的問題。

長久以來，癌症治療藥物的開發一直建立在一種信念之上。這種信念認為，如果是因為某個特定基因發生故障而引發癌症，那麼只要製造出能精準阻斷那個故障部位的藥物，疾病就能痊癒。在幾款以特定基因突變為標靶的藥物實際產生戲劇性效果後，這個信念曾一度獲得廣泛的支持。然而，隨著時間推移，即使是帶有相同基因突變的患者，對藥物的反應也出現了分歧；而最初很有效的藥物，在幾個月或幾年後失去療效的案例也逐漸累積。

問題並不在於藥物，而在於看待癌症的框架本身。人們逐漸體認到，腫瘤不單單是一張靜態的照片，而更像是持續播放的影片，必須同時解讀涵蓋患者基因資訊、生活習慣與免疫狀態的更宏觀全貌；隨著這種觀念的普及，個人化醫學便轉移到了癌症治療策略的重心。

即使同樣是乳癌，之所以會因人而異，只要聯想一下同班同學就很容易理解了。就像即使坐在同一個年級、同一個班級裡，每個人的身高、個性和擅長的科目都不盡相同一樣，在「乳癌」這個單一名籤之下，也混合著帶有不同基因突變的患者。腫瘤也不例外。雖然表面上看來只是一個腫塊，但仔細觀察其內部，會發現裡面混雜著各自帶有稍微不同基因突變的癌細胞群，這被稱為腫瘤異質性。就像在同一個教室裡，坐在窗邊的學生和坐在後面的學生會養成不同的習慣一樣，在腫瘤內部，也會根據位置和周遭環境的不同，導致每個細胞的生長速度與對藥物的反應方式有所差異。

如果忽略這種差異，只將整個腫瘤視為單一腫塊，就無法解釋為何最初看似有效的藥物，會隨著時間推移而失去療效。因此，即使是兩位同樣被診斷出罹患乳癌的患者，對其中一人很有效的治療藥物，對另一人卻可能完全無效，這種情況在實際的診間裡經常發生。

癌細胞躲避藥物的過程，可說是自然選擇的縮影。投予抗癌藥物後，大部分的癌細胞會死亡，但少數偶然帶有能抵抗該藥物之基因突變的細胞，卻能不死並存活下來。存活下來的細胞會再次分裂、增生，不知不覺間便讓整個腫瘤充滿了這種突變。抗藥性就是這樣誕生的。這也就是起初有效的藥物，後來變得不再有效的瞬間。一旦出現某個特定的致癌突變，其影響會導致細胞內的訊息傳遞路徑被重組、讀取基因的轉錄環境發生改變，甚至連產生能量的代謝流動也會被重新編排。

這種連鎖反應，是無法僅靠單一藥物阻斷單一分子就能平息的。最近，研究團隊運用了針對林奇症候群患者的單細胞空間免疫分析技術，從細胞層級深入觀察腫瘤內每一個細胞的位置與基因活動。若利用深度學習模型來分析基質細胞、免疫細胞與癌細胞之間的物理距離，以及它們之間相互傳遞的非線性交互作用，就能預先描繪出腫瘤將朝哪個方向演化的軌跡。

在研究人員之間，將人工智慧視為一種能推論腫瘤演化本身的整合引擎，這種觀點也越來越受到重視。這種方法是將在不同時間點取得的資料排列成線，重建癌細胞群的系譜，並將預測未來抗藥性產生方向的任務交給人工智慧。

所有這些預測，都仰賴作為原始材料之資料的解析度。過去，人們會將一塊腫瘤組織絞碎，將其中無數細胞的基因訊號籠統地讀取為一個平均值；但隨著單細胞定序

與空間轉錄體技術的普及，現在已經能夠同時記錄下每一個細胞的基因活動，以及該細胞在組織內確切的位置。這就好比過去只看班級平均分數，現在卻能確認每位學生的姓名與成績一樣。只有累積這樣細緻的資料，人工智慧才不會錯漏，進而捕捉到腫瘤內細胞群的微小差異。

如此掌握到的腫瘤複雜性，將直接連結到對每位患者藥物反應的預測。這裡所謂的多體學，指的是同時觀察細胞內的多個層次，包含了基因序列的基因體、顯示基因被讀取之活躍程度的轉錄體，以及實際製造出來的蛋白質。以深度學習為基礎的模型 CDRscan，是以將在多種癌細胞株中確認的抗癌藥物反應資料與該細胞株的基因體突變資訊進行配對學習的神經網路為基礎，將在細胞實驗階段所累積的知識，轉化為對患者個人的預測。

只從癌症基因體信號就能以高精度提前計算出特定抗癌藥物對該患者的實際效果有多大的方法。名為PERCEPTION的精准腫瘤學管道同時學習了大規模細胞株藥物篩選中獲得的大量表達數據和單細胞表達數據，成功地在針對多發性骨髓瘤、乳腺癌和肺癌患者的實際臨床試驗中預測了對靶向治療藥物的反應以及是否會產生耐藥性。這兩個模型都是在患者組織進入實驗室之前，先在電腦裡預先評估治療結果的做法。

最近，先在多個細胞株中進行學習，然後將該知識轉移應用於實際患者數據的遷移學習方式也被用於揭示藥物耐藥性的分子原因。整合多組學數據來提前篩選出將對免疫檢查點抑制劑有良好反應的患者群的工作也基於同樣的原理進行，這大大減少了過去投予不適合的藥物、浪費數月才確認沒有效果的試錯。

克服藥物耐藥性的努力已經進展到尋找耐藥機制本身。人工智能分析了實際患者的多組學臨床數據後，發現CDK4/6抑制劑的耐藥性分為依賴於雌激素受體的途徑和獨立的途徑兩種。CDK4/6抑制劑是激素受體陽性乳腺癌患者廣泛使用的靶向治療藥物系列，因此揭示這種耐藥途徑的成就對實際臨床應用的影響不容忽視。

VISTA是一種使免疫細胞無法將癌細胞識別為敵人的蛋白質，而STING則是一種感知身體內入侵的外來物質並啟動免疫反應的警報信號途徑。基於這一發現，研究人員設計了阻斷VISTA或刺激STING的聯合治療，填補了單一藥物無法阻斷的耐藥性漏洞。這種同時針對多個靶點的聯合治療可以用烹飪來理解。就像廚師不只是用一種原料調味，而是計算多種原料的配比後一起加入一樣，人工智能也計算出了藥物組合和

劑量比例。

預測化學化合物之間協同效應的深度學習技術也用於同樣的目的，因為癌細胞會在一條途徑被阻斷後轉向另一條途徑存活，所以從一開始就需要制定策略，同時針對多個靶點，從源頭上堵住那些迂迴的道路。

也有研究展示了當靶點不是一個而是多個時，治療效果是如何變化的。由美國加州大學聖地牙哥分校研究團隊主導的I-PREDICT臨床試驗以藥物同時針對患者腫瘤中發現的多少比例的基因突變作為標準，結果表明一致度越高，無進展生存期和總體生存期就越長。僅針對一個突變的做法會導致腫瘤立即轉向其他途徑，但同時抑制多個靶點就能減少逃脫的漏洞。

當病理學基礎模型從組織影像中提取的特徵與基因體分析結果被臨床決策支持系統一起匯總審視時，醫療人員設計多靶點組合以及驗證其證據的速度和精確度明顯提高了。研究人員基於這一結果提出，應該從第一次進行癌症基因體檢測開始，就將考慮多個靶點的設計納入標準診療程序。

曾經坐在顯微鏡前逐張讀取組織切片的病理醫生的工作方式也在改變。基礎模型這個詞可以比作建築物的地基施工。這是一種先用數百萬張組織病理圖像和多種形式的臨床數據打好寬泛的基礎知識，然後在其上逐一加載並訓練胃癌診斷或乳腺癌預後預測等具體任務的方法。這樣創建的大型病理學基礎模型已經開始在癌症診斷和預後預測中得到廣泛應用。

這些模型將細胞的微妙形態和細胞排列的空間配置轉換為病理微觀組學的數值數據，將其與基因體信息或診療記錄結合，在電腦中提前模擬進展期胃癌患者對手術前抗癌化療或免疫治療的反應。芬蘭的Aiforia是用於診斷胃癌、乳腺癌和前列腺癌的人工智能軟體，連續獲得了歐洲體外診斷醫療器械認證（CE-IVD）。分離自荷蘭拉德布德大學醫院的Aiosyn也因其用於計算乳腺癌細胞有絲分裂的軟體獲得了同樣的認證，它將讀取一張切片的時間減少了最多百分之六十。

在國內，由路尼特開發的病理分析平台路尼特 Scope與全球排名前二十的制藥公司中的十五家合作，參與臨床試驗和生物標誌物開發。二〇二五年，它與美國診斷公司LabCorp建立了數字病理領域的合作關係，並在國際學會上發表了聯合研究成果。同年在歐洲腫瘤學會上，利用路尼特 Scope預測免疫抗癌藥治療反應的三項研究同

時公開，提供了病理影像分析結果與患者實際生存期相關聯的證據。

在神經腫瘤學領域，為了以標準化方式評估腦腫瘤患者的治療反應，通過國際研究人員協議體制定了AI-RANO指南，為基於人工智能的影像判讀在臨床試驗中獲得信任提供了證據。

美國精準醫療公司Tempus是展示這一發展規模的好例子。這家在納斯達克上市的公司累積了超過五百拍字節包含分子信息、影像和臨床記錄的數據，其中包括超過四千五百萬患者的診療記錄、超過一百五十萬人的基因序列信息，以及超過四十萬份同時擁有基因體、轉錄組、影像和臨床信息的癌症患者記錄。Tempus推出的Nextela平台不是回顧過去數據的分析，而是標榜在診療進行的那一刻為醫生提供必要信息的實時臨床智能。

2026年在美國臨床腫瘤學會上發表的一項研究指出，在6家美國醫療機構、662名患者中導入該平台後，早期非小細胞肺癌患者的生物標記檢測比例有了顯著的提升。同年7月，宣布收購專門進行微量殘存疾病檢測的Personalis，開始將基於組織的精準腫瘤學與基於血液的殘存疾病追蹤整合在同一家公司內。

Tempus也啟動了一項名為PRECISION Challenge的國家級計畫，將其建構的基礎模型部分研究結果開放給外部研究人員，正朝著與學界共享過去由單一公司獨佔的資料與模型的方向邁進。

龐大的多體學資料與預測演算法所產生的結果，最終都會在主治醫師開立處方的瞬間匯聚在一起。結合Memorial Sloan Kettering癌症中心（MSKCC）的知識與自然語言處理技術，並學習美國國家綜合癌症網絡（NCCN）指南的IBM Watson for Oncology，是一套代表性的認知運算系統；它能從電子病歷中擷取患者資訊，然後以強烈推薦、可考慮及不建議三個等級，向醫師提供治療方案。名為KM2PS的框架則會在開立抗癌藥物之前，分別檢查肝功能與腎功能，以及藥物與基因、藥物與藥物、藥物與保健食品、藥物與生活習慣之間可能發生的交互作用，藉此減少副作用。

像OncoSOLVER這樣篩選臨床試驗合適性的人工智慧，已經發展成為一種虛擬平台，不僅能即時為患者尋找適合的臨床試驗，還能支援由各科醫師組成的分子腫瘤委員會進行討論。美國食品藥物管理局（FDA）正透過名為xDECIDE的決策支援系統，由人類與人工智慧共同審查案例，以此來運作虛擬腫瘤委員會。這就如同在過去

由多位醫師參與的腫瘤委員會中，又多了一位名為人工智慧的與會者。

除此之外，FDA也正在開發一種工具，能從電子病歷、癌症知識庫與臨床試驗註冊資訊中，自動將透過次世代定序（NGS）發現的突變與針對該突變的臨床試驗藥物連結起來。2026年，還出現了試圖將多個人工智慧模型整合在一個指揮系統下，以支援整個腫瘤學決策的協調框架研究，正邁向協調不同模型對同一位患者做出判斷的階段。

在一個地區癌症診療網絡中也有結果顯示，人工智慧建議的治療決策與大型癌症中心腫瘤委員會建議的一致率，從導入前的60%提升到了導入後的94%。

決定治療強度的任務，也不再依賴單一的標準劑量表。名為CURATE.AI的平台，會根據時間重新計算每一位實體腫瘤患者所需的藥物劑量，並在實際臨床上證明了觀察患者反應同時微調劑量的適應性投藥方式的妥當性。這個想法是，即使昨天和今天的體重相同，如果昨天服藥後的效果與副作用有所不同，那麼今天服用的劑量也應該改變。這種方式在療效與毒性之間差距很小、些微誤差就會導致嚴重後果的小兒癌症治療，或是移植後的免疫抑制劑調節中，都有很大的用途。

在奈米醫學領域，結合了涵蓋多層次生物學機制的模型與人工智慧的虛擬腫瘤模型，能以數值計算奈米粒子穿透血管壁並累積在腫瘤組織中的通透性與滯留效應（EPR）；這能在將藥物裝載於奈米粒子送出時，預先模擬其效果在不同患者身上會有何差異。這種計算越是精確，就越能累積證據，以便在面對相同藥物時，為每個人制定不同的奈米粒子設計與給藥途徑。

在透過手術清除肉眼可見的腫瘤後，尋找可能仍殘留在體內某處的癌細胞，長期以來一直是精準腫瘤學的課題。癌細胞在死亡時，會將自身的DNA片段釋放到血液中，這被稱為循環腫瘤DNA（ctDNA）。一滴血就是線索。即使在影像檢查還看不出任何異常的時候，只要追蹤血液中的ctDNA濃度，就能評估復發的可能性；這種手術後殘留在體內的癌症痕跡，就稱為微量殘存疾病（MRD）。

進入2026年，學界報告了一套名為MRDsteer的人工智慧系統；它在ctDNA分析過程中會即時監控突變判讀的可靠性，一旦可靠性降至特定標準以下，就會啟動自主迴圈，只挑選高風險的基因組區段重新進行判讀。驗證資料證實，在捕捉極低濃度的突變方面，這種方式比現有方法更穩定。以乳癌患者為對象、反覆追蹤ctDNA的

GEMINI Breast臨床試驗，也已從2025年10月開始招募患者；而涵蓋胃腸道癌、肺癌與乳癌的多項研究中也反覆證實，手術或治療後被判定為MRD陽性的患者，復發風險越高，存活期也越短。

醫療團隊可以根據這些資訊做出決定：對復發風險高的患者延長抗癌治療時間，對風險低的患者則減少不必要的治療。

癌細胞在生長過程中，會在表面暴露出原本體內不存在的變異蛋白質片段，這被稱為新抗原。由於每位患者的腫瘤突變各不相同，新抗原的形狀也因人而異，就像每個人的臉一樣，成為該患者腫瘤獨有的標記。人工智慧會分析基因序列資料，從這些標記中挑選出體內免疫細胞容易辨識的對象，並以此為基礎，設計出專屬患者個人的疫苗。BioNTech與Genentech共同開發的autogene cevumeran，在對接受胰臟癌手術的患者投與這種個人化疫苗後結果顯示，對疫苗產生反應並製造出T細胞的患者，無復發期大幅延長，而未產生反應的患者則在中位數13.4個月後復發。

三年後的追蹤調查中，確認應答患者的T細胞保持著平均七點七年的長壽命，目前正在擴大至二百六十名規模的二期臨床試驗，驗證化學抗癌藥與免疫治療藥物一起投用的方式。Moderna和Merck在黑色素瘤患者中開發的mRNA-4157（V940），與免疫檢查點抑制劑pembrolizumab一起投用時，將復發或死亡風險降低約四十九%，發布了這樣的臨床結果後進入三期試驗，預計最早在2026年末至2027年間決定首次監管批准。

確認腫瘤基因序列後，實際製造這種個人化定製疫苗通常需要二至四週。一份個人化定製疫苗的價格估計在十萬美元至三十萬美元之間，若投入到復發風險高的手術後輔助治療階段，成本可以用治療效果來證明其正當性，這樣的分析也出現了。使用免疫檢查點抑制劑與個人化定製疫苗的臨床試驗，以黑色素瘤為中心，已超過一百五十項。BioNTech正在開發的另一種個人化定製疫苗BNT122，在包括大腸癌在內的多種實體腫瘤中，與免疫檢查點抑制劑一起進入三期試驗階段，個人化定製疫苗的競爭不局限於一家公司或一種癌症。

當基因體和轉錄體信息、病理影像和放射線影像、診療記錄和血液中的ctDNA、生活習慣資料等不同形式的資料聚集在一起時，人工智能將其綜合，將患者分為細分類型，其結果進而導向個人化醫學和針對每個患者的用藥途徑決定。為了確定一名

患者的治療方向，病理科、腫瘤科和基因體分析團隊聚集在一起的情況現在也不陌生了。在國內，這種趨勢也不再只停留在實驗室內。國立癌症中心參與國際合作臨床研究，積累精準腫瘤學資料，在包括首爾國立大學醫院在內的主要大學醫院中，進行了將嵌合抗原受體T細胞（CAR-T）治療與液體生檢結合的臨床試驗。

標靶治療、mRNA基礎個人化疫苗、細胞治療藥物、液體生檢等四種趨勢在2026年現在開始在一家醫院內同時進行，治療的重心正在從以一種疾病名稱來捆綁患者的方式轉移到以一個一個患者為基準的方式。患者的稱呼方式也逐漸改變。除了肺癌患者或乳腺癌患者這樣的病名，出現了像EGFR陽性肺癌那樣具有特定突變的患者群，進而將其稱為該患者個人的腫瘤指紋的語言也進入了診療室。

然而，無論屏幕上的公式多麼精密，臨床現場的驗證標準要冷酷得多。WINTHER臨床試驗是展示這一現實的案例。成功地為全體患者的三十五%根據多組學資料配對定製藥物，但達到配對治療的無進展生存期應該比之前治療長一點五倍以上這一預先指定目標的患者只有二十二點四%。但是，只看全身狀況良好且配對分數高的患者，總體生存期中位數為二十五點八個月，與全身狀況不佳的患者的四點五個月相比顯示明顯差異。

差距不會自行消失。在用於學習的資料中表現出色的深層神經網絡模型，在遇到陌生醫院的新患者資料時，性能崩潰是常見的，這被稱為虛假泛化問題或領域轉移。在訓練模型過程中，應該之後才知道的治療結果信息無意中混入輸入值導致性能虛報的數據泄漏，以及只學習表面顯現的相關性而非真正因果關係的捷徑學習，也被指出為可能導致實際醫療事故的盲點。

因此，最近研究者開始要求將一家醫院開發的模型用不同醫院、不同種族、不同設備環境的資料重新驗證的外部驗證程序作為論文審查的必要條件。

應對黑箱問題的努力也在增加。分析病理影像的模型在做出最終判斷時，用熱力圖標示組織影像中的哪個部位作為根據，讓病理專家直接用眼睛確認該部位，而處理基因體資料的模型用數值反向顯示眾多變數中哪個基因和哪種相互作用對預測結果產生了重大影響。這樣的解釋技術雖然不是完美的因果證明，但在給予醫生不直接採納人工智能建議、自行判斷其根據的餘地這一點上，發揮了獲得監管部門信任的橋樑作用。

制度與技術之間的差距也可能引發法律糾紛。關於人工智能設計的虛擬患者群資料是否可以代替盲法隨機臨床試驗的對照組，美國食品藥品監督管理局和歐洲藥品管理局根據風險程度預先制定變更計劃並透明地公開的方向，陸續發佈了指南。收集和學習單細胞多組學資料和龐大診療記錄的過程與歐洲通用數據保護條例（GDPR）或美國《健康保險可攜性和問責法》（HIPAA）在多個地方發生衝突，這個問題還沒有找到明顯的解決方案。若資料集中在特定種族或發達國家患者的基因體上，該算法就會冒著向邊緣化群體建議錯誤診斷或不適當藥物的風險，可能反而擴大醫療差異。

人工智慧為什麼推薦該抗癌藥物，為什麼警告特定副作用，無法向醫生提供分子層級證據來解釋的黑盒子問題，仍然是延遲監管批准的絆腳石。當人工智慧臨床決策支援系統發生錯誤導致患者受傷或失去生命時，其責任究竟在於開發演算法的公司、提供資料的醫院，還是開立處方的醫生，目前還沒有任何國家制定出完善的法律條文來釐清這一點。保險給付體系也無法跟上這項變化。因為以過去為單一標準治療定價的方式，很難給出答案來決定如何分攤並認可每位患者不同的合併療法與個人客製化疫苗的費用。

最終，2025年與2026年的精準腫瘤學不能僅停留在提高運算速度或預測準確度的競爭上，還必須同時具備資料的公平性、判斷依據的透明度，以及與監管機構同步的法律安全網，才能邁向向下一個階段。即使在這些技術進步之中，最終做出決定的主體依然是人類醫生，這個事實並沒有改變。

在韓國，精準腫瘤學也不再是個陌生的名詞。由 Korea University 領導的 K-MASTER 事業團，自2017年6月起便與全國55家醫院合作，營運著將癌症患者的基因套組檢測結果與正在進行的臨床試驗連結起來的 Match Master 系統。主治醫生透過這個系統，可以確認是否有適合患者的臨床試驗，以及可以考慮使用哪些標靶治療藥物。

基於次世代定序（NGS）的基因套組檢測，也在2017年3月首次將10種癌症納入健康保險適用對象，並在2019年5月將範圍擴大到所有癌種。不過，檢測費用的一半仍然需要由患者自行負擔，且獲得保險給付認可的檢測，也被限制在初次診斷時，以及癌症復發或轉移到其他部位時，僅限這兩次。

即使找出了基因突變，也並不代表能立刻進行治療。就像帶有 KRAS 突變的肺癌一樣，有些癌種雖然找出了致病基因，但阻斷該基因的藥物尚未問世。即使有藥物，門檻依然存在。研究結果顯示，如果將用於尿路上皮癌的 Padcev 和 Keytruda 聯合投藥，每三名患者中就有一名的癌症會完全消失，但這種合併療法在韓國國內尚未獲得健康保險給付，因此患者必須自行承擔全部的藥費。

新上市的抗癌新藥在一年內引進韓國的比例，100件中僅有5件，甚至達不到經濟合作暨發展組織（OECD）平均的100件中18件的水準。用於肺癌或乳癌的標靶治療藥物，一個月份的費用超過數百萬韓元的情況也並不罕見。雖然找出標靶的技術變快了，但取得適合該標靶的藥物並負擔其費用，依然是另一個層次的挑戰。

這種差距也與醫院的規模息息相關。在2023年新確診的癌症患者中，居住在首都圈的人不到一半，但實際上在首都圈醫院接受治療的癌症患者，卻接近10人中的8人。居住在地方卻前往首爾所謂的 Big 5 大型醫院就診的癌症患者人數，2024年與2022年相比也增加了11.8%。最新的基因體檢測與基於人工智慧的分析，通常會先配備在首爾的幾家大型醫院，而這些設備與人力要平均普及到地區醫院，則需要一段時間。

精準腫瘤學帶來的益處，很有可能先落到那些有能力前往這些醫院就診的患者身上。不亞於提高技術的準確度，擴大該技術能夠觸及的範圍，似乎也是未來留下來的課題。

第 5 部

法規科學、數據倫理與未來融合生態系

5.1

AI新藥開發的監管框架與品質管理治理

2026年1月14日，位於美國馬里蘭州銀泉的食品藥物管理局總部與歐洲藥品管理局辦公室，幾乎同時公開了一份簡短的文件。名稱是「藥物開發之優良人工智慧實踐指導原則」，其中包含的內容僅有十項原則。以人為本的設計、基於風險的途徑、資料治理、生命週期管理等詞彙，密集地分布在這份僅有幾頁的文件中。隔著太平洋的這兩個巨大監管機構攜手並發出同一聲音的場景，是他們雖然稍晚，但正式承認新藥開發的重心正從試管與動物籠轉移至電腦伺服器的瞬間。

在此八個月前的2025年5月，食品藥物管理局宣布導入自行開發的生成式人工智慧Elsa，並在同年6月底全面部署完畢，供所有審查員使用。隨著Elsa僅用六分鐘就完成過去需要花費三天時間的文件比對工作之內部報告曝光，新藥審查速度本身正在發生改變的傳聞隨即被確認為事實。實際運用Elsa的方式有其專屬的安全機制。食品藥物管理局將Elsa封閉在政府專用的雲端環境內，在設計上防止Elsa將製藥公司提交的敏感審查資料吸收學習。

這等於是允許審查員存取內部文件，但阻止了這些內容流入Elsa的學習資料而影響其他公司審查的情況。這意味著監管當局自己也對他們所使用的人工智慧採取了相同的標準。

監管機構將這種變化自行稱之為監管科學。這意味著重新設計判斷新藥是否安全有效的科學依據本身的工作。過去的監管科學著重於以統計學驗證臨床試驗產生的數據。現在的監管科學則必須一併涵蓋驗證產生該數據的人工智慧模型有多可靠的工作。對於審查員而言，僅具備化學與醫學知識已嫌不足，現在還新需要具備能夠一起解讀模型是以什麼資料訓練、在什麼假設之上運作的眼光。

貫穿這所有動向的圖景是，監管正從追趕人工智慧的姿態，轉變為與人工智慧並肩同行的姿態。就在十年前，監管當局還是在新技術率先進入市場後，才遲遲制定規則。現在，食品藥物管理局與歐洲藥品管理局已經轉向，針對指南草案預先與產業界交換意見，並從開發初期階段就開啟了諮詢窗口。這等於是監管機構的角色，正

從要求給出標準答案的考官，轉變為共同解決問題的夥伴。

美國國會在2022年12月通過了食品藥物管理局現代化法2.0，取消了數十年來一直被定為新藥許可申請前提條件的動物實驗義務條款。取而代之的是，在電腦中預先計算化合物活動的電腦模擬、由人體細胞製成的微小器官模型類器官，以及將人體器官功能轉移到半導體晶片上的器官晶片實驗結果，成為了安全性和有效性的正式證據。這等於在法律上開闢了一條道路，即使不經過老鼠和猴子，也能透過電腦預先描繪出新藥候選物質在人體內會如何反應。

這項條款與其說是立即禁止動物實驗，不如說是承認了替代方案。製藥公司依然可以選擇動物實驗，但僅憑電腦模擬資料就能站上審查台的大門首次被開啟，這項事實所具有的份量截然不同。能夠多快消化這項應對策略，預計將決定各家製藥公司未來幾年的臨床開發速度。

食品藥物管理局早在2021年提出處理被分類為軟體型醫療器材（SaMD）的人工智慧與機器學習軟體的執行計畫時，就已準備了兩項機制。其一是稱為優良機器學習實踐（GMLP）的十項原則。這是一種食譜，規定了如何收集和分享資料、如何設計和測試模型、人與人工智慧如何分工，以及在上市後如何繼續監測性能。在十項原則中，必須將訓練用的資料和測試用的資料分開的原則、資料必須均勻代表實際用藥患者群體的原則，以及必須觀察人與人工智慧共同工作之性能而非僅看人工智慧單獨表現的原則，是最常被引用的。

國際醫療器材法規機構論壇（IMDRF）於2025年1月將此原則確立為國際標準文件並發布，美國、加拿大和英國的監管措辭實際上已匯聚為一。另一個則是預先指定變更管制計畫（PCCP）。由於人工智慧是在上市後也會遇到新資料並持續學習改變的軟體，如果每次改變都要從頭重新取得許可，整個產業將會停擺。食品藥物管理局為製造商開啟了一條途徑，允許預先訂定未來可能發生的變更範圍與驗證方法以一次性獲得批准，並在2024年底至2025年期間，陸續確定了針對內含人工智慧之醫療器材軟體的最終指南。

這很像在購買手機時，預先在條款中同意未來發布的軟體更新將自動安裝到什麼範圍的方式。

2025年1月6日，食品藥物管理局發布了題為「關於使用人工智慧支援藥物與生物製劑監管決策的注意事項」的指南草案。該指南的骨幹是由七個階段組成的基於風險的可靠性評估體系。製藥公司必須首先以文字說明要向人工智慧模型提出什麼問題，並界定該模型在新藥開發的哪個階段扮演什麼角色，同時必須將模型判斷錯誤時患者將承受的風險大小進行分級。

如果判定風險很高，就會收集相應的充分證據來制定並執行驗證計畫，然後透過最後的程序，確認結果是否符合最初設定的目標。如果要向小學生說明，情況會是這樣：不能因為班上同學在考試答案卷上寫出了正確答案，老師就立刻給他一百分。老師還會一併檢視這個答案是經過什麼樣的解題過程得出來的、是不是剛好猜對的，以及在下一次的考試中，是不是也能用同樣的方式得出正確答案。

開發新藥的人工智慧也沒有什麼不同。當人工智慧預測某種特定化合物對肝臟有害時，僅憑預測正確的結果是不夠的。人類必須能夠解讀為什麼會做出這種判斷的過程，並且能夠確認相同的邏輯是否也能一致地應用在其他化合物上。這與法庭上的邏輯並無二致，即使證人的陳述是事實，如果不去追究該陳述出現的經過與可信度，就不會被採納為證據。

監管機構害怕的情況是，人工智慧碰巧給出了正確答案，但實際上卻沒有人能解釋其原因。無法解釋的判斷就無法重複，而當我們將無法重複的判斷原封不動地套用在下一個患者身上時，會產生什麼樣的結果，沒有人能保證。

可靠性評估體系所要求的檔案中，有幾個共通的項目。包括：說明用於訓練與驗證模型的資料從何而來的來源紀錄；以數值表示模型的預測在何種範圍內值得信賴、超出何種範圍就會產生動搖的不確定性管理計畫；以及在模型上市後，定期檢查性能是否下降的生命週期監控計畫。這項機制是為了過濾掉將僅適用於少數幾種化合物的模型，誇大成似乎適用於所有化合物並提出申請的案例。

根據風險程度而要求不同水準的態度，在美國、歐洲與韓國的指南中都很常見。在數千個化合物候補中篩選出潛力候補的初期探索用人工智慧，由於最終判斷會由人類再次審核，因此即使是相對輕鬆的驗證也能通過。相反地，直接建議患者投藥劑量，或將臨床試驗參加者分類為高風險群的人工智慧，因為其判斷會直接對人體造成影響，所以會被要求提供充分得多的證據以及頻繁得多的重新驗證。

也就是說，即使同樣掛著人工智慧的名字，實際背負的監管重量，會根據其判斷所及的範圍而有很大的不同。

同樣的指南也一併探討了模型主檔（MMF）這個制度的引進。像局部作用的藥品一樣，在多次新藥申請中反覆使用類似的人工智慧模型時，如果要求每份申請書都從頭開始重新提交模型驗證資料，對製藥公司和審查員雙方來說，都會累積重複製作和審閱相同檔案的負擔。模型主檔是一種機制，將經過一次確實驗證的人工智慧模型的履歷單獨註冊起來，讓後續的其他申請案件可以將該履歷拿來作為參考檔案使用。

這就像是只要發放過一次大學畢業證書，每次向各家公司應徵時就不必重新考試，而是提交畢業證書影本的方式。產業界與監管當局雖然還處於討論詳細運作方式的階段，但都將其視為減少模型審查重疊之低效率的核心機制。

將目光轉向新藥開發的初期階段，會發現人工智慧正在改變臨床試驗設計的本身。以前是由研究人員憑藉經驗與直覺，來決定要招募幾名患者、投藥量要怎麼分配；現在則是由計算模型透過學習過去的臨床資料，事先提出效率極高的試驗設計方案。這種方法被稱為模型資訊基礎新藥開發，由於無須盲目增加患者人數，就能獲得統計上值得信賴的結論，因為這項優點，它在罕見疾病等患者招募本身就有困難的領域中被廣泛運用。

這也同時減少了尋找候選物質最佳劑量所花費的時間與成本。不過，監管當局不會直接接受這個模型所提出的設計方案。模型假設的統計前提與實際患者群體的特徵有多吻合，以及結果的可信度是否因為樣本減少而下降，都會在前面的可靠性評估體系程序中再次進行探討。

2024年制定，並於2025年8月起陸續生效的歐洲聯盟人工智慧法（EU AI Act），將預測臨床試驗參與者之風險，或控制藥品製造流程品質的人工智慧系統分類為高風險人工智慧。一旦被分類為高風險，在將產品推向歐洲市場之前，就必須接受額外的符合性評估，上市後也必須繼續承擔包含偏見、透明度與網路安全在內的龐大合規義務。符合性評估往往不能僅靠公司自己進行的內部評估來了結。

風險度高的人工智慧醫療產品，必須經過歐洲聯盟指定的獨立審查機構，也就是所謂的 Notified Body 的額外審查，如果沒有該機構發放的證書，產品就不能在歐洲

市場推出。然而，在2026年5月，歐洲聯盟透過所謂 Digital Omnibus 的政治協議修改了這個時程表，將屬於醫療器材的高風險人工智慧系統的適用時間，從原定的2027年8月延後了一年至2028年8月。這項措施是在保持法律要求目標不變的情況下，為企業爭取了準備的時間。

在此期間，包含人工智慧的醫療器材會在現有的醫療器材法規與體外診斷醫療器材法規的框架內先獲得認證，而人工智慧法的高風險義務則在寬限期結束後再重疊適用，目前是以這樣的結構在運作。這不只是一部法律，而是兩部法律對同一產品進行循序漸進、有時甚至是同時規範的雙重結構，因此製藥公司的法務團隊與許可團隊便面臨了必須將兩個時程表並排，藉以擬定應對策略的處境。

歐洲藥品管理局在2023年7月先發佈了反思文件初稿來徵求意見，隨後在2024年9月最終通過了一份涵蓋藥物生命週期全過程中人工智慧應用的反思文件。該文件具體警告了大型語言模型在撰寫規管文件或摘要臨床數據時可能出現的幻覺現象 - 即表面上看似可信但實際並不存在的事實被編造出來。

文件指出，文句越是看似流暢，就越容易在未經驗證的情況下直接相信。因此，該文件明確規定，無論人工智慧做出什麼判斷，最終責任必須由人類承擔。這意味著，即使人工智慧提供了再精巧的初稿，如果沒有人類審查並簽署，這份文件也無法成為規管文件。

無論規管框架多麼嚴密，如果支撐它的數據本身存在明顯偏差，就沒有太大用處。訓練人工智慧模型用的基因組信息和臨床記錄往往集中在西方幾個國家，甚至特定種族和特定收入階層。如果將在這類數據基礎上訓練的模型直接應用於不同種族或具有不同遺傳背景的患者，就會增加誤漏不期望的副作用或錯誤推薦劑量的風險。

為了在不違反歐洲《通用數據保護條例》或美國《健康保險隱私及責任法》等保護患者信息的法律前提下，同時使用多個國家的數據，聯合學習技術受到關注。這種技術將原始數據保留在醫院服務器上，只交換學習結果。然而，由於各國的計算系統差異很大，計算成本也相當高昂，這項技術完全成為規管文件的標準化做法還需要更多時間。

驗證傳統軟體的方式與驗證人工智慧的方式根本不同。計算器程序會對相同的數字輸入產生相同的輸出，所以一旦確認無誤，就不需要重複測試。然而，人工智慧是

一種遇到新數據時會自行調整權重的軟體，因此前一天正確的判斷不能保證今天仍然正確。

因此，監管當局不是對人工智慧進行一次性測試就結束，而是將其視為活的監督對象，在產品留在市場期間進行反覆確認。這類似於購買汽車後不是一次性檢查，而是需要每年進行定期檢查的結構。

當由數千萬個參數組成的人工智慧模型將某種化合物指認為有毒物質時，如果人類無法解釋該判斷背後的因果關係，那麼無論預測多麼準確，也無法通過規管批准的門檻。解釋為什麼僅有準確度是不夠的，這是一項技術性任務，屬於驗證與確認（V&V）領域。業界經常同時使用所謂的SHAP（Shapley加法解釋）技術和LIME（局部可解釋模型無關解釋）技術。

這兩種技術都針對人工智慧做出的單一判斷，反向計算每個輸入值對該判斷的影響程度，並以人類可讀的數字和圖形呈現。雖然無法看透黑盒內部，但可以確定黑盒對哪些輸入賦予了多大的權重。此外，還會採用美國機械工程師協會為醫療領域制定的V&V 40原則，以及長期以來在醫療信息學領域使用的評估報告指南STARE-HI檢查清單。

核心是對模型預測與實際人體內發生的現象進行交叉驗證，並詳細記錄驗證過程本身，使其他研究人員可以完全複製。這就像評分試卷時，不僅公開答案鑰匙，還公開評分標準，使任何人都能按相同標準重新評分一樣。

隨著人工智慧和數位雙生技術進入藥物製造過程，品質設計原則（QbD）和過程分析技術（PAT）開始與人工智慧交融。QbD是指從設計階段開始就將品質內置的原則，PAT是實時監測過程的技術。傳感器和人工智慧實時捕捉溫度微小波動或雜質混入的時刻，並自動調整控制參數。此時，監管當局首先確認的是該數據未被篡改的證據。國際製藥工程協會在2022年修訂的GAMP 5第2版在軟體驗證實務中密集納入了ALCOA++原則，該原則指的是數據的準確性、一致性和可追溯性。

其中包含了對以下要求的詳細規定：必須能夠確認誰在什麼時候創建了數據，數據必須是可讀的，必須在實際事件發生的那一刻立即記錄，必須按原樣保存，必須完整且準確地保留。必須留下審計記錄，確認訓練數據和驗證數據在物理上分離開來，模型沒有發生過度擬合（即像是提前背誦了試題後的作答那樣的過度吻合）。這

類似於廚師單純聲稱按食譜烹飪是不夠的，必須保留購買材料的收據和烹飪過程的照片才能獲得認可。

留下詳細審計痕跡的習慣是在工廠現場而非紙上實踐優秀機器學習實務的方法，同時也是連接到製藥品質管理部門實質競爭力的途徑。

當人工智慧自行搜索知識圖譜並繪製出前所未有的分子結構時，應將該成果的專利權授予誰尚未在國際上得到解決。專利制度原本是以人類創意行為為前提而設計的，因此沒有任何國家的專利局能夠輕易地用相同的標準來評判由算法獨自發現、幾乎沒有人為介入的發明。關於是否應該認可發明人欄無法填寫人名的發明，如果認可，是應該將權利授予開發程序的公司還是提供數據的機構，各國的判斷各不相同。

當事故真正發生時，由誰來負責的問題依然存在。假設監管機構根據人工智慧分析的合成對照組數據或安全性預測結果批准了新藥，但在上市後發生了意想不到的患者死亡事故。在這種情況下，究竟是開發人工智慧的軟體公司負責、提供數據的醫院負責、開出最終處方的醫生負責，還是審查文件並蓋章批准的監管機構負責？目前任何國家都還沒有明確的劃分標準。

就像汽車自動駕駛事故中，駕駛人、製造商和軟體公司之間已經展開了責任歸屬的爭論一樣，新藥開發的人工智慧不久後似乎也會在法庭上面臨相同的問題。

這個領域的工作人員組成也在發生變化。以前是化學家、醫生和統計學家來填寫新藥批准文件，現在則是數據科學家、軟體驗證專家和監管律師圍坐在同一個會議室裡。因為對人工智慧模型的風險進行分級並釐清責任歸屬，最終必須以法律語言來整理，才能完成提交給監管機構的文件。現在已經到了一個如果懂技術的人和懂法律的人不互相學習對方的語言，就連一份文件也無法好好完成的時代。

新藥上市後，人工智慧的角色就轉移到了上市後安全性監視，也就是藥物警戒領域。以前，只有當醫生和患者自發性通報副作用時才能發現問題，因此罕見但危險的藥物交互作用往往很晚才會被發現。現在，自然語言處理技術會即時瀏覽電子病歷、社群媒體貼文和龐大的臨床文獻，及早捕捉人類肉眼容易漏掉的微小信號。在醫院看診和日常生活中就這樣累積起來的這類資料，被稱為真實世界數據。

食品藥物管理局建立的資訊視覺化平台 InfoViP 和 Sentinel 系統，是將這些真實世界數據轉換為圖表，讓審查員可以一目了然的基礎設施。Sentinel 系統在 2024 年 1 月改版為加入了自然語言處理功能的 3.0 版本，至今已執行了超過 250 件的安全性分析。國際醫學科學組織理事會 (CIOMS) 於 2025 年 12 月 4 日最終確定並發表了探討人工智慧與上市後監視的第 14 工作小組報告。這份報告以基於風險的方法、人為監督、模型的穩健性、透明度、數據隱私、公平性和治理這七項原則為骨幹，具體整理了在人工智慧精確挑出罕見副作用、自動過濾重複報告以及自動遮蔽患者個人資訊的實務中，其可信度能達到什麼程度。

上市後觀察到的安全性資訊再次回到初期開發階段，反映在人工智慧訓練數據中的循環結構，正以這樣的方式逐漸完成。

2025 年 1 月 24 日，隨著韓國的數位醫療產品法正式實施，首次制定了法律依據，將包含人工智慧、機器人和虛擬實境技術的尖端數位醫療產品歸入一個獨立類別，從批准到上市後管理都在同一個框架內處理。同一天，食品藥物安全處在全球率先制定並發布了生成式人工智慧醫療器材許可與審查指南。這是組成由學術界、醫療現場和產業界專家組成的協商機構，提前探討生成式人工智慧醫療器材從開發到批准以及上市後管理可能存在哪些風險因素後，所產生的結果。

這股趨勢在 2026 年也延續了下來。食品藥物安全處於 2026 年 6 月 30 日推出了一份新的民眾指南，其中包含了基於大型語言模型 (LLM) 的數位醫療器材許可與審查標準，同時也完善了反映應用人工智慧技術的數位醫療器材特性的變更管理計畫書許可與審查指南。這項措施為產品開闢了道路，只要事先獲得對批准後預期變更事項的核准，就無需每次都重新經過許可和審查，而是可以直接反映在產品上，其骨架與美國的預定變更控制計畫相似。

食品藥物安全處還宣布，計劃從 2026 年到 2030 年的五年內，投入 347 億韓元，培育涵蓋人工智慧新藥開發、尖端生物和數位醫療器材領域的 1,100 名監管科學專業人才。這筆預算將分配給涵蓋學位課程和非學位課程的十個執行機構來執行。這項投資的基礎似乎是基於一種判斷：如果審查員不了解人工智慧模型的運作原理，那麼無論制定多麼精密的指南，在實際審查台上都會失去作用。在美國和歐洲透過法律和指南建立標準的同時，韓國可以說正在把預算投入到培養實際處理這些標準的人才上。

這種趨勢也給韓國製藥公司和生物新創公司帶來了實務上的課題。如果要向美國食品藥物管理局申請新藥，必須遵循美國的可靠性評估體系；如果要進軍歐洲，必須同時滿足人工智慧法的高風險分類和反思文件的原則；而在國內，則必須遵守數位醫療產品法和食品藥物安全處的指南。雖然這三個地區的規定所要求的文件格式和術語略有不同，但其底層的核心問題最終是相通的。如果國內企業從開發初期就按照國際標準累積驗證資料，以後同時向多個國家提交申請時，就能大幅減少重複製作相同資料的麻煩。

國際醫藥法規協和會 (ICH) 於 2026年 1月 29日 最終採納了探討基於模型的新藥開發一般原則的新指南 M15。這項工作創造了一種共通語言，讓各國監管機構在包含人工智慧在內的計算模型用於臨床試驗設計和劑量決定時可以參考。雖然目前還沒有單獨探討人工智慧的獨立國際醫藥法規協和會指南，但如果與 2025年 7月 23日 生效的臨床試驗設計原則 ICH E6(R3) 放在一起看，就可以知道各項指南正朝著同一個方向相互接軌。

美國的現代化法與可靠性評估體系、歐洲的人工智能法與反思文件、國際機構的優良機器學習實踐與上市後監測原則，以及韓國食品藥品安全處的法律與人力培養計劃，儘管各自名稱不同，但所要求的核心是統一的。這就是：人工智能給出的答案是否正確已不再是重點，重要的是人類必須能夠驗證並對達到該答案的過程負責。從法律人的觀點看，這並非陌生的要求。法庭長期以來一直要求檢驗證據的可信性，並將判斷依據記錄在案，如今這套做法正在轉移到實驗室、工廠和醫院的數據流程之中。

但是，法庭的做法與此有所不同。法庭是就已經發生的事件事後進行判決，而新藥監管屬於預防領域，必須在事故發生前建立嚴密的驗證程序。人工智能給出的結果在多大程度上能被人類理解和承擔責任，這個問題看起來將在新藥開發的每個階段被重複提出。

責任歸屬的標準至今尚無，為了填補這一空白，歐洲率先出現了相關動作。歐盟在 2024年11月修訂並公佈了《產品責任指令》，規定軟體和人工智能系統應與有形商品相同視為製造品，並承擔嚴格責任。成員國應在2026年12月9日前將該指令轉化為國內法，受害患者不再需要單獨證明開發商的過失即可請求損害賠償。但是，該時期技術水準無法發現缺陷的開發風險抗辯也同樣適用於軟體，因此人工智能開發

商也獲得了免除責任的新途徑。

韓國的《產品責任法》至今尚未明確將軟體納入製造品範圍，因此即使由人工智能設計的新藥發生事故，針對軟體本身的責任追究在法庭上仍有很大爭議空間。如果國內立法跟不上歐洲的步伐，同樣遭遇事故的韓國患者將面臨更重的舉證責任。

有關監管審查官是否具備驗證人工智能能力的問題，人事統計給出了更為冷酷的答案。美國衛生與公眾服務部在2025年4月1日一次性裁員食品藥物管理局3,500名職員，隨後有報導稱留下的審查官不得不承擔起同事原來的工作。這發生在該部門發佈基於風險的可靠性評估體系還不到三個月的時候。食品藥物管理局似乎是為了彌補這個缺口，在2025年12月1日部署了全員可用的代理式人工智能，該工具可以分擔上市前審查、審查結果驗證以及上市後監測的工作。

以人工智能補足審查人力不足的方式，似乎與「人工智能的判斷必須由人驗證」的原則相悖。韓國食品藥品安全處的處境也不寬裕。截至2025年11月，食品藥品安全處醫療產品審查人員為369人，相比美國食品藥物管理局的9,000人、歐洲藥品管理局的5,500人、日本醫藥品醫療器械綜合機構的1,300人，人員數量相差甚遠。食品藥品安全處正與行政安全部協商追加聘用295名審查官的方案，但其招聘信息網站自2021年起即處於停滯狀態，準備工作進展緩慢。

即便增加人力培養預算，如果坐在審查席上的人數本身就這樣稀少，各國詳細審查每一個人工智能模型的能力就會有很大差異，而且看起來會出現這樣的情況：有些公司會採取先向監管把關較鬆的國家提交申請的策略。

5.2

資料倫理、演算法公平性與聯邦學習生態系統

2024年7月，在首爾的韓國製藥生物協會辦公室，來自十多家製藥公司和大學研究團隊的負責人聚集在一起。該項目名稱為K-MELLODDY，是一項5年期的國家課題，將持續到2028年12月。保健福祉部和科學技術信息通信部投入348億韓元，計畫加速推進以聯邦學習為基礎的新藥開發。會議室的氣氛很微妙。桌子上放著各機構的名牌，但筆記本電腦螢幕上卻沒有任何需要相互展示的共享畫面。

處於競爭關係中的製藥公司研究人員聚在一起，但沒有人想把自己公司的化合物數據交給競爭對手。儘管如此，他們仍然決定共同訓練人工智能模型。並且完全沒有交換任何數據。要想理解這個看似矛盾的承諾如何真正發揮作用，我們首先需要審視新藥開發人工智能目前面臨的根本困境。

要實現精準醫療，需要收集數百萬人的基因組和臨床記錄。單靠一家醫院或一家製藥公司擁有的患者數量，無法解讀罕見疾病的模式。但是，一旦將患者的遺傳信息和臨床記錄集中在一處，該數據庫就會成為黑客和數據洩露的目標。美國通過HIPAA（健康保險攜帶和責任法案）進行限制，歐盟則通過GDPR（一般數據保護規定）嚴格禁止患者信息的跨境流動和對外披露。

從製藥公司的角度來看，製造良好人工智能模型所需的數據在醫院的圍牆外，但法律卻禁止跨越這道圍牆。如果無法打破這種僵局，人工智能新藥開發將只是一個少數大型製藥公司利用自身擁有的數據進行的封閉遊戲。

突破這種僵局的技術是聯邦學習（Federated Learning）。用小學生都能理解的方式來解釋是這樣的。十名同班同學在一起準備考試，他們約定絕對不向他人展示自己的筆記本。相反，他們每天晚上檢查自己解答的問題和正確答案，然後只互相分享哪種解題方法能更好地提高分數。十天後，十人中沒有一個人看到過別人筆記本的內容，但每個人的成績都比獨自學習時有所提高。

聯邦學習也是如此。醫院和製藥公司絕不會將包含患者數據的筆記本移出自己的服務器。相反，人工智能查看了這些筆記本並逐步變得更聰明，然後只將由數字組成

的模型權重加密後發送到中央服務器。中央服務器將來自多個機構的權重合併為一個，然後將變得更聰明的模型返回給各個機構。將多個機構的權重平均併合為一的程序稱為聯邦平均（Federated Averaging）。

這種權重交換方式單憑理論無法保證安全性，因為可能存在從權重本身反向推斷原始數據的攻擊。因此，近期研究人員正在結合兩種技術：差分隱私（Differential Privacy）技術 - - 在將權重發送到中央服務器之前故意混入微小的噪聲，以及安全多方計算技術 - - 防止多個機構只能匹配計算結果而無法看到彼此的輸入值。

2026年發表的PrivMol研究在預測分子特性的圖形人工智能模型中結合了這兩種技術，展示了在不顯著降低預測準確度的同時降低原始數據洩露風險的結果。這就像不只上一把鎖，而是上雙重甚至三重鎖的方式。

首次在新藥開發現場證明聯邦學習確實有效的項目是歐洲的MELLODDY。該項目由歐洲創新醫藥倡議（IMI2）和歐洲製藥工業協會（EFPIA）共同投資1800萬歐元，從2019年領導3年。全球十家製藥公司和七家技術與學術機構參與了該項目。他們共同用於學習的數據包括超過26億個實驗活性值、超過2100萬個化合物和超過4萬個生物學檢測結果。沒有任何公司向競爭對手展示自己的化合物清單。

儘管如此，通過聯邦學習創建的共同模型的預測性能比每家公司僅使用自身數據創建的模型提高了平均5.5%，最多達9.7%。這就像是在不同的冰箱裡各自放入不同的食材，卻通過共同精進一個食譜來提升所有人的烹飪效果。這種改進意味著可以過濾掉的失敗化合物數量增加了，無需在實驗室中直接合成，從而降低了新藥開發初期階段的時間和成本。

相同的理念在韓國以K-MELLODDY的名稱重新誕生。由韓國製藥生物協會於2024年7月啟動的該項目包括三個主要軸：建立聯邦學習平台的任務、收集和提供醫院和製藥公司數據的任務，以及使用這些數據訓練人工智能模型的任務。項目組的預期輸出是名為FAM（聯邦ADMET模型）的預測模型。ADMET是指任何化合物進入人體後的吸收程度、分佈方式、代謝方式以及所引起毒性程度的統稱。

2025年，高麗大學、崇實大學、韓國科學技術院、LG化學等多家機構新承擔並啟動了五項新的子任務。對於國內中等規模製藥公司和生物風險企業而言，由於他們往往缺乏購買數據的資本，這種既能保留各自所有權又能共同學習的方式被認為很有

吸引力。

除了製藥公司之間的合作外，連接醫院與醫院的聯邦學習也在迅速增長。由於罕見疾病在單一國家或單一醫院患者數量很少，人工智能學習模式的樣本本身就不足。例如，在處理腫瘤影像或病理照片的跨國醫院聯盟中，來自不同國家的醫院各自在自己的服務器上保留患者影像，只交換用於區分腫瘤的模型權重。

沒有任何一家醫院會將暴露病患身分的原始影像發送到其他醫院，但模型卻能達到僅靠單一醫院數據無法企及的準確度。這種合作之所以可行，是因為原始數據在物理上並沒有跨越國界。

聯合學習所守護的價值與可解釋人工智慧所守護的價值，就像是放在天平兩側秤盤上的兩個砝碼。一側的秤盤上有著安全性的砝碼。讓原始數據在原地紋風不動，僅讓知識往來交流的聯合學習壓住了這個秤盤。另一側的秤盤上則有著可靠性的砝碼。能夠讓人類看懂並解釋人工智慧為何做出該判斷的XAI（eXplainable AI，可解釋人工智慧）壓住了這個秤盤。

只要任何一側變重，天平就會傾斜。如果數據得到了安全的保護，卻沒有人能解釋模型為何得出那樣的結論，那麼醫師和病患都無法信任並使用該結果。相反地，如果解釋非常完美，但產生該解釋的數據本身卻暴露在洩露的風險中，那麼這個人工智慧從一開始就無法跨過醫院的門檻。

然而，安全地收集了數據，並不保證該數據本身是公平的。散佈在世界各地的臨床試驗與基因體資料庫，大多都集中在WEIRD，也就是西方、受過教育、工業化、富裕且民主國家的總人口上。以2025年年中為基準，參與全基因體關聯分析（GWAS）的人當中，擁有歐洲血統的人口比例高達90.53%。非洲血統人口的比例在過去十年間幾乎沒有增加。

自2009年以來的15年多裡，以歐洲血統為主的究比例僅減少了約5個百分點。如果將僅以北美和歐洲白人病患數據訓練出來的奈米粒子藥物傳遞模型或腫瘤標靶預測模型，直接應用於遺傳和生活習慣背景完全不同的亞洲、非洲或南美洲病患身上，預測性能將會大幅下降，出現意想不到的毒性副作用的風險也會隨之增加。

由於多個用於檢測皮膚癌的人工智慧診斷工具主要是以淺色皮膚病患的照片進行訓練，因此在深色皮膚病患身上的診斷準確度便大幅下降。這不僅僅是人工智慧軟體的問題。在COVID-19大流行期間廣泛使用的脈搏血氧儀，也出現了將深色皮膚病患的血氧飽和度測量得比實際數值還要高的誤差，美國食品藥物管理局曾為此發布了額外的安全性警告。某個診斷細菌性陰道炎的人工智慧工具，在白人女性中準確度很高，但在亞洲女性中準確度卻大幅下降，而對於西班牙裔女性，即使沒有生病也會錯誤判定為有病的偽陽性情況更是特別多。

美國一家大型保險公司所使用的風險預測演算法，系統性地將黑人病患排除在健康管理計畫之外。這是因為該演算法將醫療支出金額作為健康風險的替代指標，而黑人病患即使患有相同的疾病，由於醫療可及性較低，支出金額本身就比較少。演算法即使沒有直接查看人種，也原封不動地複製了基於人種的歧視。在技術上，為了刻意對抗偏誤模式而訓練模型的對抗性學習技術，也正作為減少偏誤的方法被一同使用。

也有研究將數據的公平性以數值呈現出來。使用衡量機器學習模型健康公平性的HEAL框架來拆解某個皮膚疾病診斷模型後發現，在基於性別的次群體中，公平性分數高達92.1%，表現優異；但在患有非癌症皮膚疾病的高齡病患群體中，分數卻降到了0.0%。這意味著在同一個模型中，有些病患能得到精準的診斷，而有些病患則實際上被忽視了。當訓練數據失去平衡時所產生的另一個陷阱，就是捷徑學習。

如果用健康人群樣本佔80%、疾病患者樣本佔20%的數據來訓練模型，模型只需無條件地將所有輸入值判定為健康，就能獲得80%這個看似不錯的準確度。實際上這是一個連一種疾病都無法找出的模型，但光看數字卻會產生表現優異的錯覺。一旦陷入這個陷阱，像罕見疾病患者或有色人種等數據較少的病患，就會完全承受誤診與不當處方的風險。

美國食品藥物管理局（FDA）與歐洲藥品管理局（EMA）在設計臨床試驗的階段，便將要求刻意納入少數人種、不同年齡層與社經弱勢群體的多樣性行動計畫，明定在臨床試驗的草案中。美國食品藥物管理局於2025年8月發布了指導原則，規定了從上市前到上市後，應如何管理包含人工智慧的醫療器材軟體。該指導原則要求對訓練數據的代表性與偏誤進行管理、確保網路安全，以及進行上市後的性能監控，同時也促成了名為「預先決定變更計畫」（PCCP）的制度扎根，讓演算法在預先設

定的範圍內，無需重新取得許可也能進行修改。

2026年也推出了關於臨床決策支援軟體的獨立指導原則。不能僅止於增加數據量，還必須挑選並納入能實際代表人口結構的高品質數據，這項原則現在看來已經超越了幾份文件的建議，逐漸成為新藥許可審查的實務標準。

從根本上修正數據傾斜的嘗試也正在同步進行。一直以來，人類基因體研究都依賴一種方式：將特定人口群體中某一個人的基因體作為標準參考序列，再將其他所有人的基因體與該序列進行比較。由於這個單一參考序列最初並非為了反映人類的多樣性而設計，因此偏離標準的遺傳變異越多的群體，在診斷與治療預測上就越不利。

近年來，將多個人口群體的基因體重疊繪製而成的泛基因體（pangenome）參考地圖，成為了彌補此一侷限的替代方案。只要在泛基因體地圖中充分納入過去研究中被忽視的群體樣本，人工智慧模型從一開始就能將更為均勻的數據作為學習材料。

即使糾正了數據的偏差，深度學習模型中還有另一個需要克服的障礙，那就是黑盒子問題。由數千萬到數十億個參數交織而成的神經網絡，能在瞬間得出某種化合物對肝臟有害，或是具有抗癌標靶潛力的結論。然而，究竟是透過什麼路徑得出這個結論，往往連設計模型的研究人員也無法解釋。這就像是把化合物的資訊放進一個無法打開蓋子的箱子裡，然後只會彈出答案一樣。

沒有人能看到箱子裡進行了什麼樣的計算。在攸關病患性命的藥物核准審查中，這種不透明性是很難被直接接受的。因為如果要讓醫師開立處方、讓監管機構蓋上核准的印章，僅靠預測機率這樣一個數字是不夠的，還必須能夠檢視其背後的判斷過程。

試圖打開這個黑盒子的嘗試，就是可解釋的人工智慧，也就是 XAI。像 SHAP 或 LIME 這樣反向追溯人工智慧結論的技術，會用顏色或數字，具體呈現出化合物的哪種化學結構，或是蛋白質的哪段序列，對結合力或毒性預測做出了多少貢獻。這就像是不只交出一件組裝好的家具，而是連同寫明了在哪裡釘了幾個什麼螺絲的組裝說明書一起交給您一樣。醫師與監管機構人員必須看到這份說明書，才能驗證該結論是否出自真實的生物學依據，而不是偶然的統計相關性。

不過，2026 年一份針對抗生素探索用 XAI 的評估研究指出了一項限制，那就是衡量這些解釋技術效能的標準本身尚未統一。雖然提供解釋的工具越來越多，但用來衡量這些解釋有多可靠的標準，可以說還沒有完全具備。

如果想讓人工智慧預測模型真正在臨床現場扎根，僅靠這種解釋技術是不夠的。就像美國機械工程師學會制定的 V&V 40 原則一樣，驗證計算模型可靠性的程序，以及由人類再次確認最終判斷的監督程序，都必須從設計階段開始就一起規劃進去。近年來，像 CONSORT-AI 與 DECIDE-AI 這樣，要求以文件透明地記錄人工智慧模型效能、訓練環境與臨床應用範圍的報告指引，也正以學界為中心不斷增加。

跳過這些程序就進入市場的人工智慧醫療器材，日後在稽核或訴訟過程中，越來越常因為無法提出佐證資料而陷入困境。不過，這些指引目前還只接近於學術期刊的編輯方針，並不是法律強制的規範。介於建議與義務之間的這道落差，在實際發生醫療事故時，就會直接引發一個問題：到底該向誰追究法律責任？

隨著人工智慧超越了診斷與預測，達到能自行設計新化合物的水準，也開啟了完全不同的法律爭議。生成式人工智慧即使沒有研究人員逐一指示，也能自行學習包含結構與活性之間關係的數據以及蛋白質摺疊資訊，從頭設計出新的分子結構。這就是所謂的 *de novo* 設計。像是 Insilico Medicine 利用自家人工智慧平台進行設計，並將特發性肺纖維化治療候選物質推進到臨床二期的案例；或是 Exscientia 所設計，於 2020 年進入臨床一期的強迫症治療候選物質 DSP-1181，這些由人工智慧從頭到尾參與設計的化合物，已經有許多項進入了臨床階段。

問題在於，為這樣誕生的分子申請專利時，發明人那一欄應該填寫誰的名字。

有一個事件正面碰上了這個問題。美國發明家 Stephen Thaler 宣稱名為 DABUS 的人工智慧自行完成了發明，並將不是人類的 DABUS 列為發明人，向包括美國、歐洲、英國與韓國在內的多個國家申請了專利。2022 年，美國聯邦巡迴上訴法院在 *Thaler v. Vidal* 案件中明確表示，美國專利法所稱的發明人必須是自然人，因而不承認 DABUS 的發明人地位。2025 年 11 月 28 日，美國專利商標局發布了處理人工智慧參與發明之發明人資格的新審查指引，再次確認了一項原則：人工智慧只是像實驗室儀器一樣精密的工具，而發明的核心 -- 構想（*conception*），是只有人類才能進行的行為。

在 2026 年 3 月 4 日，連日本最高裁判所都駁回了 Thaler 的上訴，這使得纏訟 7 年的 DABUS 全球專利攻防戰實質上落下了帷幕。回顧過去，歐洲專利局在 2021 年、英國最高法院在 2023 年 12 月的判決中，就已經得出了相同的結論。到目前為止，包括美國、歐洲、英國、日本在內的幾乎所有國家的專利局與法院，都沒有承認 DABUS 為發明人。

在律師的眼中，這些判例所說明的原則很明確。即使人工智慧實際設計出了候選物質，專利文件的發明人欄位裡，也必須填寫設計並操作該人工智慧，且實際看出化合物價值的研究人員的名字，才能保住專利。問題在於，那些花費數千億韓元建立人工智慧新藥開發平台的企業，該如何將這項原則落實到實務中。因為當人工智慧在一個晚上產出數百萬個候選分子時，研究人員必須在每一份專利申請文件中，逐一證明自己對其中的哪一個分子進行了多少創造性的介入。

如果無法證明這一點，好不容易發掘出的化合物就會面臨無法受到專利保護，而直接暴露給競爭對手的風險。相反地，如果過度誇大研究人員的貢獻，就等同於自己製造了日後專利本身被推翻無效的把柄。因為這個緣故，部分企業從一開始就選擇了以營業秘密代替專利來保護化合物的策略，但是營業秘密也有一項限制，那就是只要競爭對手獨自找出了相同的化合物，這層保護傘就會直接消失。在實務上如何拿捏這個平衡，很可能在未來幾年內成為專利訴訟的核心爭議點。

知識產權的衝突並不止於發明者地位的問題。隨著人工智慧利用大型語言模型破譯傳統醫學古籍和原住民社群的藥用植物知識，進而發現新的天然物新藥候選物質的研究日益增加，一個新的問題也隨之浮現：應該如何認可那些幾百年來守護這些知識的社群所擁有的權利？2010年通過的《名古屋議定書》是一項國際條約，規定利用遺傳資源及其相關傳統知識的一方必須事先獲得知識持有者社群的同意，並且公平分享由此產生的利益。

迄今為止，142個國家已經批准了這一條約，130多個國家各自制定了相關的利益分享法律。然而，在人工智慧能夠迅速學習大量傳統知識文獻並提出新化合物的今天，追蹤某項資料來自哪個社群並據此分享利益的程序仍未被妥善確立。南非的聖人族（San）和科伊族（Khoi）與紅茶灌木茶產業達成的利益分享協議經常被引用，作為傳統知識持有者社群實際獲得版稅的案例。

當人工智慧透過傳統知識文獻尋找新化合物候選物質時，也應該進行相同水平的版稅支付和出處標示，這樣的呼聲在學術界日漸高漲。將傳統知識轉變為私有化專利並獲取巨大利潤的動機，與向守護這些知識的社群返還正當代價的要求，正面相撞。

超越專利問題，還有更為尖銳的問題在等待著。當人工智慧的判斷導致患者實際受傷時，責任應該由誰承擔？即使是經過可解釋人工智慧驗證並通過聯邦學習訓練的臨床決策支援系統，也可能因訓練數據中隱藏的細微偏差而對特定患者群體錯誤地推薦不安全的藥物，並判定其為安全的。例如，如果訓練數據未能充分反映僅出現於特定基因型患者群體中的副作用，人工智慧就無法預測該風險，反而可能做出安全的判定。

在這種情況下，責任究竟應該由設計神經網路的資訊科技企業、提供數據的醫院、相信該推薦並開出處方的醫生，還是批准許可的監管部門承擔？一個能夠明確區分這些責任的完整法律體系至今還未在世界任何地方齊備。

歐盟為了填補這一空白，準備了一份獨立的人工智慧責任指令。但在2025年2月宣布將撤回該指令草案後，於同年10月在官報刊登中確定了撤回，並改變方向，決定在修訂後的產品責任指令框架內處理人工智慧相關事故。歐盟《人工智慧法》本身於2024年8月1日生效，自2025年8月起，對開發通用人工智慧模型的機構可以處以罰款。包含醫療設備在內的高風險人工智慧系統的義務因2026年《綜合修訂法案》而被推後，附件3對應的高風險系統自2027年12月2日起適用，附件1系列的產品內置系統自2028年8月2日起適用。

從事高風險人工智慧的企業必須建立風險管理計劃、高質量訓練數據、透明文件、人類監督機制以及部署後的持續監測體系。如果其中任何一項不符，就可能被視為產品責任指令所述的缺陷。這是一個既定義務並因未能履行該義務而在發生事故時難以免責的結構。

歐盟《人工智慧法》對醫療人工智慧進行雙重監管。參與患者診斷或治療決定的人工智慧系統必須獲得醫療設備規定下的認證，同時也必須遵守《人工智慧法》規定的高風險系統義務。在兩項監管中，只要違反其中任何一項，就會被禁止進入市場。這意味著僅憑算法性能是不夠的，必須自始至終記錄下該性能所源自的數據和設計過程的文件，才能進入市場。

參與聯邦學習聯盟的機構之間也日益普遍地在參與合約中預先規定損害賠償和保險負擔。這是在發生事故時於法庭上爭執責任歸屬之前，先通過合約條款分散風險的舉措。

韓國的法律體系也正朝著相似的方向發展。「匿名資訊」是指刪除了姓名、身分證號等能夠直接識別特定人士的項目，或將其替換為其他數值，使得在沒有附加資訊的情況下難以識別身份的個人資訊。韓國的《個人資訊保護法》規定，不同機構持有的匿名資訊不能由相關人員直接進行比對，只有通過稱為「數據結合專門機構」的第三方機構才能進行數據結合。政府依據2022年至2026年的研究開發路線圖，一直在完善多個機構各自持有的匿名資訊相互結合的制度和技術，2026年還將舉行競賽活動來評比利用結合匿名資訊的案例。

保健福祉部在2026年研究開發事業中宣布了將醫療數據和人工智慧技術連結到新藥開發和醫療現場各處的計劃，政府決定到2030年投入4000億韓元用於人工智慧基礎新藥開發和先進腦科學研究。

歐洲方面，與此相應的制度是《歐洲健康數據空間》（EHDS）規定，該規定於2025年3月26日生效。根據該規定，全部規定自2027年3月26日起適用，而用於個別診療的第一次使用和進行研究、政策制定、人工智慧開發等與原始目的不同的第二次使用，則將分別於2029年和2031年逐步適用。患者摘要、電子處方以及大部分研究用數據將於2029年先行開放，而影像、檢查結果、出院記錄、基因組資訊等敏感或複雜的資料則被推後至2031年。

各會員國應設立健康數據存取機構，負責核發數據使用許可並管理相關責任。歐洲資料保護委員會與歐洲資料保護監督機關曾擔憂，此規定可能會削弱《一般資料保護規則》所確立的個人權利，因此最終條文採取了折衷方案，將二次使用的法律基礎重新附加在《一般資料保護規則》之上。保護隱私的規範與收集使用數據的規範，似乎在同一個體系內相互拉扯。

聚集在 K-MELLODDY 會議室的研究人員，之所以能在不向彼此展示筆記的情況下共同提升成績，是因為技術證明了可以同時守護信任與安全這兩種價值。然而，在技術開啟的大門背後，仍有一些尚未解答的問題。例如：該模型是使用哪個族群的數據進行訓練的？人工智慧設計的化合物專利應歸屬於誰的名字？發生事故時應由哪

個機構承擔責任？這些都是需要決定的工作。

從病患與市民的眼中來看，值得記住的是，當這種平衡被打破時，最先暴露於風險之中的，往往是那些數據較少、難以發聲的人們。這項工作並未在實驗室內結束，而是持續延伸至法庭、國會，以及國際組織的談判桌上。在現階段預先將偏見審計條款與發明貢獻度記錄義務納入合約中，似乎能減少日後發生爭議時釐清責任的成本。

如同美國與日本的法院已決定不承認人工智慧為發明人一樣，未來幾年內，處理數據公平性與責任歸屬的規範，也有很大可能會隨著各國的判例與指引而逐步累積。在這些規範變得嚴密之前，聯邦學習與可解釋的人工智慧目前所創造的信任與安全之平衡，似乎將會持續面臨考驗。

保健福祉部與個人情報保護委員會於 2025 年 12 月 31 日，修訂並實施了《保健醫療數據活用指引》。為了解決多間醫院共同進行的研究，或是難以自行設立機關保健醫療情報審議委員會（即 DRB）的中小型機構的困境，新設立了公用 DRB 制度。此外，也制定了將機關生命倫理委員會與 DRB 審議程序合而為一的標準程序。

當使用已故病患的醫療數據進行研究時，也制定了新標準，要求必須刪除會暴露身分的識別碼以及與遺屬相關的代碼後才能使用。對於使用人工智慧尋找新藥候選物之國內新創企業或大學研究團隊而言，從收集多家醫院數據到訓練預測模型所需等待的審議時間，預計將會因本次修訂而縮短。

然而，這所有程序都是以審議委員會的視角在運作，而非提供數據的病患本人。國內的個人資料同意體系，仍停留在醫院最初收集資訊時取得一次性同意的方式，之後經過假名化處理的資訊究竟被用於哪項研究、或是哪家企業的人工智慧學習，目前並沒有提供讓病患事後追溯的管道。此外，法律也已經允許，即使假名資訊被用於統計編製或科學研究等與原目的不同的用途時，亦無須再次取得資訊主體的額外同意。

如同英國國民保健服務，雖然有些國家設有只要資訊主體事後表達拒絕意願便停止使用資訊的 opt-out 機制，但韓國法律尚未發展至該階段。不需要同意，並不代表連監督都不需要，但究竟由誰來負責這項監督工作，目前仍不明確。即使不將數據移出原處的聯邦學習再怎麼安全，如果病患沒有辦法自行確認自己資訊的去向，那

麼這種安全很有可能只是半調子的安全。

即便相信聯邦學習能保護原始數據，這種信念也存在著漏洞。2025 年 3 月，國際學術期刊《Journal of Cheminformatics》刊登了由研究員 Fabian Kruger 帶領的研究團隊，針對預測新藥候選物質性質的神經網絡模型，進行 membership inference attack（成員推論攻擊）的嘗試結果。該方法不需要窺探模型內部的權重，僅透過從外部提問並獲得回答，就能猜出特定化合物是否包含在訓練數據中。

無論在哪種數據集或哪種模型結構中，資訊都會外洩，而且屬於樣本數較少的少數群體化合物，就越容易被識破。在新藥開發中，樣本較少的化合物往往是競爭對手尚未涉足的珍貴候選物，這等於是面臨了最需要保護的資訊反而最先外洩的風險。Kruger 研究團隊將此攻擊手法的程式碼一併公開，讓其他研究人員也能自行檢查自身模型的風險。

5.3

製藥產業的未來：自主驅動實驗室 (Self-driving Labs)

2025年10月，Seoul的一間實驗室悄悄地開門了。在Hanguk Jeyak Baio Hyeophoe旗下AI Sinyak Yeonguwon所打造的這個空間裡，沒有穿著白袍的研究員，而是一支機械手臂站在天平前。這支機械手臂轉開玻璃瓶蓋，精準地取出0.005毫克，相當於一粒鹽巴重量幾百分之一的試劑。混合液體並再次秤重的動作，每隔幾秒鐘就會重複一次。過了晚上十點，研究員們陸續下班，實驗室的燈也隨之熄滅。然而，機器人在那片黑暗中並未停止運作。

人們將這個空間稱為「Dark Lab」。協會相關人士說明，這間實驗室的目標在於，即使在人們熟睡的時候，也能將實驗處理量提升到最大。機器人所處理的每個反應容器，都會自動記錄實驗條件與結果；隔天上班的研究員，便能一邊喝著一杯咖啡，一邊確認昨夜累積的數十件實驗結果，以此展開新的一天。出入口前方的螢幕上，會以即時數字顯示當晚處理的實驗件數與成功率；如果溫度或濃度超出了預先設定的範圍，系統也設計成會自動發出警告音。

在2026年6月首次向媒體公開時，研究員方面也同時表明了要在今年內重新培育150名人力的新計畫，而國內的幾家製藥公司也已經在檢討將這間實驗室的營運方式引進自家的方案。僅僅在幾年前還被視為電影場景般的景象，如今已經成為Seoul一間實驗室裡每晚不斷重複的日常。

2024年，國際學術期刊《Nature》上刊載了一項研究，內容是基於大型語言模型的人工智慧，自行控制光譜儀與液體處理機器人，在不經過人手的情況下，合成了二十九種有機矽化合物。其中有八種是先前世上未曾知曉的化合物。化學家必須在試劑櫃前苦戰好幾天的工作，人工智慧卻能自行尋找反應路徑並協調設備行程，在一個晚上之間就完成了。過去的新藥發掘，是尋找成為疾病原因的標靶，並設計符合該標靶的化合物，進行合成，再透過試管與細胞確認功效，這四個階段必須耗費好幾年來枯燥地反覆進行。

這個通常取設計、合成、測試、分析的第一個字母，被稱為DMTA週期的過程，是一個光靠人類走完一圈就需要花上幾個星期的緩慢循環。跨國製藥公司Janssen將愛滋病治療藥物Prezista的生產製程，更改為基於即時感測器與機器學習的自動調節系統，進而滿足了美國食品藥物管理局與歐洲藥品管理局的嚴格標準，這件事也處於同樣的趨勢之上。從Seoul的一支機械手臂，到地球另一端學術期刊的一頁，實驗這項行為本身的定義正在發生改變。

引領這項變化的力量，是能夠自行思考並行動的Agent AI。這指的是只要給予目標，就能負責制定計畫直到執行的AI。過去的醫療用人工智慧，比較接近於針對一個輸入值給出一個輸出值的被動工具。Agent AI則不同。它會持續接收資料，記住先前的判斷，如果判斷自行制定的計畫有誤，就會修改並寫出下一個計畫。在此加上能夠同時解讀視覺、語言、化學訊號的多模態能力，這種人工智慧已經不再停留在雲端上的軟體，而是提升到了能夠直接指揮光譜儀與機械手臂等實驗室硬體的程度。

名為ChemCrow的初期結構，很好地展現了這個原理。就像人類打電話給各個部門分派業務一樣，這是由大型語言模型透過為每個光譜儀與層析設備所規定的通訊規格，也就是所謂的應用程式介面，來直接下達命令的模式。研究人員只要輸入目標物質與必須遵守的安全條件，人工智慧就會翻找龐大的化學反應資料庫來規畫合成路徑、訂購試劑、協調設備行程，並在執行實驗後分析其結果，自行修改下一個條件。

從閱讀文獻、建立假說，到透過機器人進行合成與分析，這種不經過人手而連成一體的流程，被稱為閉環系統。與只會重複既定順序的現有自動化設備不同，這個系統在情況改變時，會自行重新規畫接下來要執行的程序本身，在這一點上有著根本上的差異。

閉環的第一步，是建立假說的工作。Google DeepMind所推出的「Coscientist」，是基於Gemini大型語言模型，由四個代理程式分擔角色來進行協作的系統。當生成代理程式瀏覽論文與資料並提出多個假說時，反思代理程式就會像同儕審查員一樣找出其中的漏洞。排名代理程式會讓存活下來的假說以淘汰賽的方式進行較量，而演化代理程式則會再次修飾並結合其中優秀的假說。

四個代理程式會互相交換彼此的成果，並好幾次重新檢討同一個假說，這種反覆過程比較接近於模仿人類在學術會議發表前，與同僚們進行好幾個月論戰的結構。英國Imperial College London的Jose Penades教授研究團隊，研究了將近10年關於抗生素抗藥性基因如何在細菌之間傳播的問題，他們試著將這個尚未發表的研討主題直接交給了Coscientist。人工智慧在無法接觸到研究團隊未公開資料的狀態下，僅在48小時內就提出了指向相同結論的假說。

人類研究員必須耗費幾週到幾個月來建立並反駁的假說驗證過程，這四個代理程式能在短得多的時間內自動運轉。問題在於，當假說出現得越快，實際去確認該假說的雙手就會成為新的瓶頸。自主運作實驗室，正是為了填補這個瓶頸而登場的下一個階段。

當假說到達機器人前，接下來就是物理執行。中國的XtalPi在世界各地運作約300台機器人工作站，全天候不停地運行，結合自身開發的約500個人工智能模型和量子力學計算，來合成和評估化合物。該公司正在與禮來進行規模高達2.5億美元的聯合研究，並在國內與JW中外製藥合作，共同完善針對炎症疾病靶蛋白STAT6的先導物質。人工智能演算法繪製的分子結構或蛋白質序列直接傳輸到機器人組裝線和細胞培養系統，轉變為物理物質。

在此過程中，所有產物從電腦中的設計圖到最終生產批次，被賦予連貫的唯一編號，終究被追蹤記錄。這是最起碼的安全網，讓人們之後能夠反向追蹤某個實驗條件如何導致了某個結果。首爾的DarkLab中，機械臂以0.005毫克為單位處理試劑的精準動作，以及地球另一端實驗室的300台機器人軍隊，最終都按照相同的原理運作。人類易於出錯的重複工作，機器人的手臂從不疲倦，能夠將相同條件完全相同地重複上百次、上千次，這一點大大提升了物理執行階段的價值。

實驗結束後，循環不會停止，而是進入下一階段。嵌入生物反應器中的多個傳感器實時讀取細胞的基因表達、電信號和代謝產物的跡象，並將其發送到人工智能控制裝置。強化學習控制器接收到這些信號後，自動調整培養液的濃度和信號物質的量，在品質偏離標準之前提前採取行動。結果分析產生的數據回到最初的假說設計階段，成為下一次實驗條件更精細完善的依據。這樣，假說設計、實驗室自動化、結果分析、數據反饋這四個階段形成無休止的循環。

昨天的失敗實驗成為今天成功實驗的材料，今天的成功又產生了明天的新問題。由於每次實驗運行時系統本身都會逐漸變得更聰明，這個實驗室不再是需要人工操控才能運作的靜態空間，而更接近於一個自我學習、不斷成長的生命體。與以前人類每天工作八小時、處理少數幾個實驗不同的是，現在在相同的時間內，可以同時測試數百個條件，並將結果立即反映在下一個實驗中。這種運作規模的差異正在逐步出現。

自主實驗室的可能性也在大企業的宣傳資料之外得到驗證。多個人工智能代理組成一個團隊，設計了附著於新冠病毒的納米抗體 -- 這是一種根據駱駝或羊駝體內發現的抗體仿製而成的極小結合蛋白 -- 並通過實際實驗驗證了其效果。這項「虛擬實驗室」研究經過同行評審後發表在學術期刊上。美國FutureHouse開發的「Robin」系統將假說建立、實驗設計和結果解釋的過程統一為一，新發現青光眼治療藥物Ripasudil可以作為乾性老年性黃斑變性（即隨著年紀增長而逐漸喪失視力的眼病）的治療候選。

這項研究於2026年5月發表在國際學術期刊《自然》上。兩個案例都經過大學和研究機構的同行研究人員驗證並發表為論文，證實了自主實驗室不是投資融資的宣傳文詞，而是實際運作的科學方法論。不僅獲得龐大資本支持的企業在參與，相對預算有限的大學研究團隊也正在通過租用雲端人工智能和共享實驗室設備的方式加入這股潮流。

生物技術創投公司Flagship Pioneering於2025年3月首次向世界推介名為Lila Sciences的公司，並宣佈該公司獲得2億美元的種子輪融資。同年10月，在英偉達投資部門主導下，該公司又新增融資1.15億美元，累計融資達5.5億美元，企業估值達到13億美元。到2026年6月，該公司估值85億美元的後續融資談判也正在進行中。曾任OpenAI研究部門副總裁的Liam Feder和曾在Google DeepMind擔任研究員的Akin Dous Chubak共同創立的Periodic Labs也於2025年獲得3億美元種子輪融資，正在建設針對包括超導體在內的新材料的自主實驗室。

英國格拉斯哥大學的Lee Cronin教授創立的Chemify公司憑借其「化學電腦」技術 -- 將化學反應轉換為代碼，使機器人能夠準確再現 -- 於2025年獲得了超過5000萬美元的B輪融資。這三家公司並非只瞄準單一新藥。它們的目標是化學、材料和能源領域研究開發本身的自動化。在投資者名單中，還包括英偉達、亞馬遜創始人傑

夫·貝索斯和前谷歌首席執行官埃里克·施密特等人物。僅這三家公司籌集的投資資金加總就達9億美元，而幾年前這個領域還只被視為大學實驗室研究課題的程度，如今能吸引這麼多資金，這一事實本身足以說明該技術所通過的驗證之重。

自主實驗室產生的成果已經在多個國家以具體數字得到驗證。澳洲聯邦科學與工業研究組織（CSIRO）將開發新催化劑所需的時間從5年縮短到了6個月。加拿大的加速促進聯盟通過自主實驗室發現了21個以上的新候選物質。在韓國，韓國科學技術研究院（KIST）自主開發了一個名為「章魚」的操作系統，用以指揮多台機器人設備，並正在進行相關實驗。韓國製藥生物技術協會和延世大學則正通過K-NIBRT項目組單獨構建相關基礎設施。

這兩項計畫目前都還處於只有少數研究人員接觸過的起步階段，但由於這是將政府與大學各自嘗試的實驗自動化整合在一個標準之下的工作，因此扮演著邁向下一階段的橋樑角色。雖然國籍與研究對象各自不同，但共通點在於，原本需要人類投入數年時間的工作，人工智慧與機器人只需幾個月，有時甚至只需幾週就能完成。

Hanguk Gwahak Gisul Jeongbo Tongsinbu 提出了從 2026 年到 2028 年的三年間，投入 495 億韓元建設六座自主實驗室的計畫。一處通用實驗室由 Gathollik Daehakgyo 負責建立「K-Cell」平台，五處特化實驗室則分別由 Daegu Gyeongbuk Gwahak Gisulwon 負責液體活檢、KAIST 負責傳染病、Goryeo Daehakgyo 負責基因載體、POSTECH 負責酵素工程、Ulsan Gwahak Gisulwon 負責基於類器官的藥物評估。領導此計畫的未來戰略技術政策官 Oh Dae-hyun 表示，結合人工智慧與機器人的自主實驗室，將成為從根本上改變生物研發典範的契機。

政府除此之外，為了以國內技術確保物理人工智慧（也就是相當於機器人大腦的技術），也決定在 2 年內追加投入 340 億韓元。首爾的 DarkLab 事先展示的熬夜實驗場景，到了 2028 年，將準備擴展為在全國六處同時展開的畫面。

業界一部分正提出一項構想，要將從閱讀一篇論文到申請專利所需的時間，壓縮到 90 天左右。雖然目前還沒有被公認實際達成此目標的案例，但已經出現了能讓人猜測這個方向並非空喊口號的數字。總部設在香港的 Insilico Medicine 以標靶發掘程式 PandaOmics 與新藥候選物質設計程式 Chemistry42 為首，表示要找出臨床前候選物質（也就是準備好進入動物試驗的物質），平均需要 12 個月到 18 個月的時

間。

這個數值甚至不到傳統新藥發掘所需 2 年半到 4 年時間的一半。該公司進行的一項計畫，在合成不到 80 個候選物質的情況下，只花了 260 萬美元就找出了針對新標靶的新分子，這大約是現有方式所需費用的 10 分之 1 水平。該公司推出的特發性肺纖維化治療候選物質，在人工智慧設計的新藥中，首次於臨床 2 相試驗中確認了改善肺部功能的效果，並在 2025 年將針對疼痛治療的新候選物質 ISM9528 推進到臨床前階段，向年市場規模達 1000 億美元的領域下戰帖。

該公司共同營運的 inClinico 程式，是一種用數據事先評估臨床 2 相試驗能延續至 3 相並取得成功的機率的工具；從尋找候選物質的階段到設計臨床試驗的階段，同一家公司的人工智慧從頭到尾都有參與，在這一點上可以稱之為閉環的縮影。

將量子力學的糾纏與疊加現象與機器學習結合的量子機器學習（QML），能以現有電腦難以應付的精確度，計算出分子之間微小的交互作用。法國的 Qubit Pharmaceuticals 利用 Pasqal 製造的中性原子方式量子電腦 Orion，透過量子演算法計算出了水分子在蛋白質結合口袋中的位置。2026 年 3 月，美國 Cleveland Clinic 與 IBM 攜手宣布，他們成功以穿梭於量子與古典電腦之間的混合方式，計算出了由 303 個原子組成的微型蛋白質 Trp-cage 的電子結構。

就像藥物在體內是否會對膠原蛋白組織引起意外的毒性一樣，在事先解讀多個原子同時糾纏所產生之微妙交互作用的工作中，量子演算法也提出了比現有方式更精確的答案。市場調查機構預測，這個領域的市場規模將從 2025 年的 1 億 2600 萬美元，成長到 2035 年的 6 億 3800 萬美元，年平均成長率為 17.9%。這裡還結合了從投入原料到生產出成品的過程中，不中斷地讓製程流動的連續製造技術。在乾燥與粉末混合、壓製藥錠的打錠過程中發生的微小溫度變化或雜質，會由製程分析技術與數位孿生進行即時感測與模擬，Janssen 在生產線上加上這種體系並滿足監管標準的案例便是代表。

這意味著發現的速度與生產的速度，現在開始以相同的節奏運作了。實驗室裡幾個星期內找到的候選物質，在工廠裡也能在幾個星期內具備大量生產的條件，這使得過去一直存在於發現與生產之間的時間差距本身正在縮小。

產業界將這種整合稱為超越自動化的 Industry 5.0。作為機器取代人類重複勞動的 Industry 4.0 的下一個階段，這意味著人工智慧與機器人成為擴展人類認知與身體極限之伴侶的時期。電腦伺服器裡的設計圖，經過實驗室機器人，無縫銜接至工廠生產線，也是其中的結果之一。隨著人類無法觸及的夜晚與週末也被納入研發之中，用相同的預算與相同的人力所能產出的候選物質數量本身，正在呈數倍增加。然而，即使是伴侶，也需要有必須遵守的規則。

走進這個閉環，我們可以看到一款新藥所走過的道路。在找出與疾病相關蛋白質的標靶發掘階段，人工智慧會瀏覽龐大的基因組與論文數據，篩選出有潛力的標靶。預測出該蛋白質的三維結構後，人工智慧會從頭開始繪製出完全符合該結構的分子，同時預先計算該分子在肝臟中如何分解、如何排出體外，以及是否具有毒性。如果判斷某個候選物質僅靠單一小分子化合物是不夠的，它也可能被重新設計成同時瞄準多個標靶的多標靶抗體或細胞治療藥物的形式。

這些候選物質會被裝載到由脂肪顆粒或高分子製成的奈米粒子上，經過再次修飾，以確保能精準遞送到肝臟或腫瘤等目標組織。用患者細胞在培養皿上培育出的微型器官「類器官」，以及將患者全身原封不動複製到電腦裡的「數位雙生」，成為了在動物實驗之前，預先測試候選物質效果與副作用的舞台。實驗室本身也不例外。在讓機器人使用實際試劑之前，先在電腦裡的虛擬實驗室中運行實驗條件，只篩選出成功機率較高的條件再交給實體機器人的方式，也越來越被廣泛使用。

在這裡存活下來的候選物質，會進入由數位雙生預先設計好的臨床試驗，並與讀取腫瘤基因圖譜、為每一位患者尋找合適治療方法的精準腫瘤學工具相遇。在這些所有階段產生的紀錄，也會同時累積成為提交給監管機構的申請文件草案。從尋找標靶到為患者尋找合適處方，原本在不同部門各自運作的工作，現在都在一個循環中互相交換數據並同時運作。

這個循環看起來並非完美無缺。代理人工智慧因為偏見數據或錯誤，而自行設計出危險物質，或者在連續製造生產線上應用了錯誤的條件，導致大量生產出不良藥品的可能性，是真實存在的威脅。因此，在實驗室現場，越來越多的地方會將人工智慧的活動範圍限制在預先設定好的安全界線內，並額外設置「守衛代理」來進行監視。一旦偵測到試圖越過規定界線的行為，守衛代理就會被設計成能立即停止實驗本身，並將判斷權交給人類，發揮了自主性與安全性互相牽制的雙重裝置作用。

感測器與軟體之間的同步錯誤導致數據出現偏差的技術缺陷也時常發生，而且人工智慧為何會做出那種判斷，人類無法完全解釋的黑盒子問題也依然存在。為了控制這些風險，美國食品藥物管理局與歐洲藥品管理局在 2026 年 1 月，首次發布了在開發新藥使用人工智慧時必須遵守的共同指南。其核心原則是，越是影響品質的判斷，人工智慧的建議就越需要由人類進行最終確認。製藥公司必須依照國際製藥工程協會修訂的電腦系統驗證標準指南，以及在被分類為醫療器材的軟體中應用人工智慧時應遵循的執行計畫，還有必須能夠追蹤數據是誰在何時建立的數據完整性原則，將人工智慧的學習過程與偏見驗證程序記錄成文件。

為了控制人工智慧在自行學習過程中逐漸產生變化的現象，也就是所謂的模型漂移，美國食品藥物管理局也導入了事前變更管理計畫制度，規定在產品上市後，必須預先訂定人工智慧可以安全更新自身的範圍並獲得批准。

進入法律領域後，問題會變得更加尖銳。對於在沒有人類發明者實質介入的情況下，人工智慧閱讀了數千萬篇論文並從白紙上繪製出的分子，是否能賦予專利權，這個問題的答案因國家而異，且目前在世界上尚未有統一的定論。當自主實驗室最佳化出的用藥方法對患者造成了意想不到的副作用時，責任應該歸咎於編寫軟體的技術企業、導入該軟體的製藥公司，還是給予批准的監管機構？目前還沒有任何地方有明確規定這點的法律體系。在美國與歐洲各自以不同速度與方式制定指南的同時，包含韓國在內的許多國家，目前都還沒能為這個問題準備好獨自的答案。

如果各國專利局與法院的判斷出現分歧，對於同一個人工智慧設計出的同一個分子，可能會真實發生在某些國家承認專利，但在某些國家卻遭到拒絕的情況。技術能跨越國界，在一夜之間製造出二十九個新的化合物，但要為該結果負責的法律語言，卻依然以不同的速度在各個國家中緩慢移動。

關於在機器人取代雙手的實驗室裡，人類應該做些什麼，我的判斷是這樣的。研究人員並不會消失。只是他們所扮演的角色正在發生根本性的改變。雖然搬運試劑與測量天平的雙手勞動會交給機器人，但最初要提出什麼問題，以及要相信哪些數據、懷疑哪些數據的決定，依然必須由人類來做。在人工智慧徹夜產出的數百個結果中，分辨出哪些是偶然、哪些是真正發現的眼光；決定系統不該越過的安全界線要畫在哪裡的判斷；在發生事故時，能夠回答該由誰承擔責任的法律與倫理監督角色，這些都是機器人無法取代的。

研究人員似乎正在從實驗台前拿著滴管的技術人員，轉變為設計整個實驗室，並為該實驗室可能犯下的錯誤承擔責任的監督者。對於在大學裡學習化學或生物學的學生，現在也同時被要求具備統計、程式設計與機器人工程的基礎，而營運 Dark Lab 的協會將人才培育放在事業的首要位置，或許也是出於相同的原因。取代透過指尖學習實驗技術的學徒制教育，驗證人工智慧提出的判斷並深入探究其極限的訓練，正逐漸成為下一代研究人員的基本功。

即使在 Dark Lab 熄燈的夜晚，也必須有人向該機器人下達今天要做什麼實驗的命令，而且隔天早上，也必須先決定好要在該結果上簽名的人。

2028年，若投入495億元的韓國自律實驗室六處按計畫啟用，連同首爾的Dark Lab，全國各地實驗室每晚將同時熄燈、啟動機器人，呈現一幅景象。在地球另一端，運轉300台機器人的CrystalFab和獲得85億美元估值的Layla Science，將在同一時刻以相同方式繼續實驗。無論是瞄準超導體的Periodic Labs機器人，還是將化學反應轉化為代碼的ChemiFy的Chemputer，各自針對不同物質，卻都將以相同原理通宵運作。

閱讀文獻、提出假設、合成物質、反饋結果的這一循環，即便在人們睡眠期間也不會停頓。尋找目標的人工智能、繪製分子的人工智能、控制機器臂運動的裝置，如今雖然散佈在不同公司和不同國家，卻像一個循環一樣啣合運轉。但是，當某個清晨時分，自律實驗室整夜優化的某個配置出現了意料之外的數據時，站在它前面的人是誰，還沒有任何國家的法律中明確記載。

機器人不會回答那個問題。那個問題只會靜靜地放在次日早上來上班的研究人員的辦公桌上，旁邊是一整夜積累的數十份數據。準備那個答案的工作，每個實驗室燈光再次打開的早上，都會回到人手中。

美國伯克利國立實驗室創建的自律實驗室A-Lab在2023年在《自然》雜誌發表論文，稱機器人在沒有人工干預的情況下，在十七天內發現了數十種新型無機化合物。幾個月後，倫敦大學化學家Robert Palgrave等外部研究者們重新審視了這些結果。號稱「新發現」的物質中，許多早已在結晶結構數據庫中登錄——也就是收集已知物質原子排列的資料庫。《自然》在2026年1月發表了勘誤聲明。

主導該實驗室的伯克利研究隊伍反駁稱A-Lab確實合成了那些化合物，爭執至今未息。從建立假設、執行實驗到解釋結果，一個人工智能越是獨自承擔的封閉環路，能夠重新審視其判斷的外部眼光反而越少 -- 這起事件似乎說明了這一點。沒有直接接觸試劑經驗、僅通過屏幕上的數字學習化學的下一代研究者，能抓住多少這樣的錯誤，至今還沒人能保證。

防止人工智能設計危險物質的機制曾經經歷過一次測試。2022年，美國製藥公司 Collaborations Pharmaceuticals的研究人員將自家的、受訓以避免毒性的新藥設計人工智能MegaSyn的目標反轉，命令它尋找毒性最大化的分子。僅用一台普通筆記本電腦運行六小時，人工智能就生成了多達四萬多種有毒物質的列表，包括《化學武器公約》禁止的神經毒劑VX。

研究人員並未詳細公開這些結果，而是在《自然機器智能》期刊上僅披露了風險，將具體的化學結構保密。在教來製造新藥的人工智能和教來製造毒藥的人工智能之間，除了人改變一個目標外別無二致 -- 這是這個實驗留下的令人不適的教訓。

逆轉MegaSyn目標所需的只是改動幾行代碼。在機器臂通宵稱量試劑、人工智能自主決定下一步實驗的時代，實驗室裡必須有人注意到那個手勢並將其制止。若某個早上有人把人工智能整夜設計的一種藥物擺在你面前，在追問那種藥是人製造還是機器製造之前，你需要先問一個問題。那就是誰用何種方法確認了那種藥。

AI新藥開發的典範轉移與未來融合生態系

作者 | 金京鎮

發行人 | 金京鎮

出版 | 金京鎮律師出版社

定價: USD 15

© 金京鎮 2026

本書為著作者之智慧財產，禁止未經授權之轉載與複製。

註) 本書部分內容與編輯過程借助了人工智慧工具。

kimkj008@gmail.com

支持這本 AI 書房藏書

若這本書對您有幫助，歡迎透過 PayPal 支持後續的翻譯、
免費公開版製作，以及開放的人工智慧知識計畫。



PayPal

<https://paypal.me/kimkjkj>

請掃描 QR code 或開啟上方連結。