

出國報告（出國類別：進修）

神經內科疾病分析進修報告

服務機關：臺中榮民總醫院

姓名職稱：趙怡婷 / 主治醫師

派赴國家/地區：美國 / 國家衛生研究院

出國期間：113 年 9 月 3 日至 114 年 9 月 2 日

報告日期：114 年 10 月 2 日

摘要

為發展神經內科的疾病整理及基因診斷分析系統，本人於 2024 年 9 月 3 日至 2025 年 9 月 2 日至美國國家衛生研究院(National Institute of Health, National Institute of Neurological Disorders and Stroke (NINDS) Stroke Branch)進修。學習基因組學資料分析的技術與分析平台，與人工智慧於醫療方面的應用。例如 NGS, PRS, RNA-seq, ChIP-seq 的原理和運用，以及 MATLAB, BigQuery, Galaxy 的操作方法流程。並且接觸 NIH 在 Bioinformation 的分析平台例如 Biowulf，以及練習運用程式語言，使用 Anaconda, Jupyter Notebook 建立資料分析環境。

關鍵字：資料科學、 基因組學、人工智慧

目 次

一、 目的.....	1
二、 過程.....	<u>2</u>
三、 心得.....	<u>10</u>
四、 建議事項.....	<u>11</u>
五、 附錄.....	<u>13</u>

一、 目的

目前基因診斷與基因治療正迅速蓬勃地發展當中，這些技術已被廣泛應用於遺傳疾病的篩查、攜帶者檢測、疾病易感性評估、個體化醫療方案的制定、癌症基因治療以及基因編輯等多個重要領域。同時，隨著大數據技術的不斷進步，將大量醫療資料進行整合與分析，進一步推動了疾病預測模型和監控系統的建立，提升了臨床診斷與治療的精準度與效率。為了推動神經內科疾病的資料分類與基因分析，本人於 2024 年 9 月 3 日至 2025 年 9 月 2 日期間，赴美國國家衛生研究院（National Institute of Health, NIH）旗下的國家神經疾病與中風研究所（National Institute of Neurological Disorders and Stroke, NINDS）中風部門（Stroke Branch）進行進修學習。此次進修的主要目標包括：掌握神經科疾病相關的基因診斷技術與基因治療方法，並了解機器學習與自然語言處理在醫療領域中的多樣化應用。

美國國家衛生研究院是全球頂尖的醫學研究中心之一，整合了臨床疾病研究、藥物研發、診斷技術以及病患追蹤系統等多項技術，研究範疇涵蓋流行病學與統計學、基因體學、蛋白質體學、代謝體學以及多體學，並且積極引入人工智慧等新興技術進行跨領域創新。院內匯聚了眾多專業研究人員，致力於各個研究層面的突破與發展。我所跟隨的指導教授范揚政老師（Dr. Fann），是人工智慧與機器學習領域的專家，同時也是 NINDS 資訊技術部門的主任，負責構建部分組織架構，並熟悉各大 API 介面與雲端運算的應用，使我得以近距離學習這些前沿技術的實際操作。

在 NIH 進修期間，我並傾注時間學習基因組學資料的分析技術與相關平台的運用。近年來，人工智慧在醫療領域的應用日益廣泛且深入，我也利用在美國國家衛生研究院進修的寶貴機會，了解人工智慧在不同層面的應用方式及操作流程。技術層面上，我學習次世代定序（Next-Generation Sequencing, NGS）、多基因風險評估（Polygenic Risk Score, PRS）、RNA 序列分析（RNA-seq）及染色質免疫沉澱測序（ChIP-seq）等技術的原理與實際應用方法。在平台操作方面，我練習了 MATLAB、Google BigQuery 以及 Galaxy 等分析工具的使用流程。此外，在指導教授的協助下，申請並通過審核取得了 NIH 生物資訊分析平台（如 Biowulf）的帳號，並練習使用程式語言，透過 Anaconda 與 Jupyter Notebook 搭建資料分析環境，進一步加強實務操作能力。

雖然本次進修期間，由於 NIH 經費調整，導致其下的一些教學與研究單位，如科學進階教育基金會（The Foundation for Advanced Education in the Sciences, FAES）所開設的操作班與學分班不再提供學費補助，但我仍然把握機會，參與各類院內（intramural）與院外（extramural）的研討會與學術交流活動，透過這些會議學習新的基因分析技術和研究成果，增添自己的知識與實戰經驗。

二、 過程

(一) 進修單位介紹

美國國家衛生研究院（National Institutes of Health，簡稱 NIH）是美國極為重要且具領導地位的醫學研究重鎮，其研究範圍涵蓋疾病的成因、診斷、預防與治療，並探討人類成長與發展過程。此外，NIH 也積極研究環境污染物對生物體的影響、精神健康與成癮問題，以及醫學與健康資訊的收集、傳播與交流等領域。NIH 由 27 個不同的研究單位（Institutes and Centers）組成，總部設立於美國馬里蘭州貝塞斯達（Bethesda）。其研究機構幅員廣大，擁有超過 75 座建築，佔地面積超過 300 英畝。部分研究工作在 campus 內的高端實驗室中進行，但超過 80% 的研究活動是由分布於全美各州乃至全球的科學家共同合作完成。近年來，NIH 尤其致力於推動精準醫學的發展，積極投入人工智慧與大數據的應用，以及資料科學領域的創新研究，期望藉此促進醫療科技的突破與提升公共健康水平。

我所隸屬的單位是神經與中風研究院（National Institute of Neurological Disorders and Stroke，簡稱 NINDS），這是一個致力於神經系統疾病及中風研究的機構。NINDS 內部架構，由多個職能部門間協調運作。首先是神經科學部門（Division of Neuroscience, DON），負責規劃並指導對外合作項目與各項研究計畫，推動基礎神經科學的發展；其次是外部活動部門（Division of Extramural Activities, DEA），主要負責制定與協調院外研究計畫的管理，確保外部資源與研究方向符合 NINDS 的整體策略。第三是臨床研究部門（Division of Clinical Research, DCR），其核心任務是主導臨床試驗的開發與執行，涵蓋早期及進階階段的臨床試驗、比較效益研究，以及涵蓋整個神經疾病發展週期的流行病學研究。最後是轉譯研究部門（Division of Translational Research, DTR），專注於推動生物標記的發現、小分子藥物治療、生物製劑療法以及罕見基因治療等創新領域，致力於將基礎研究成果快速且有效地轉化為臨床治療方案，促進患者的康復與健康改善。

(二) 神經科臨床疾病討論會及研究會議

NIH 在大華府地區有三家主要的合作醫院，這些醫療機構與 NIH 緊密合作，共同推動臨床與基礎研究的整合發展。除了這些核心合作夥伴外，NIH 在全美各地設有許多院內研究部門（Division of Intramural Research）以及院外研究部門（Division of Extramural Research），並與眾多獲得資金支持的臨床機構或研究單位建立合作關係。這樣的合作網絡使得 NIH 的研究成果能夠連結不同地域和領域，並促進醫學科學的跨界交流與應用。

雖然基於我們持有的研究學者簽證規定，不允許直接接觸病人，但 NIH 經常舉辦多種形式的研究會議，這些會議內容常常以臨床案例為切入點，結合最新的學術

研究成果，實現臨床經驗與科研知識的對接。我特別安排時間參與這些會議，從多位領域先進、學術巨擘以及志同道合的同好們的分享中，吸取寶貴的臨床經驗與前沿知識，這對我深入理解疾病的臨床表現與世界最新研究方向頗有助益。

這些研究會議大致可分為兩種主要形式：一是個案討論與 grand rounds，這類會議通常聚焦於具代表性的臨床病例，分別分析病歷、影像及治療方案；二是依疾病類別分門別類舉行的專題會議，如中風會議(stroke conference)、失智症會議(dementia conference)、帕金森氏症會議(Parkinson's disease conference)等，這些會議聚焦於特定疾病的最新研究進展與臨床應用，提供深入的學術交流機會。透過參加這些多元的會議，我得以持續更新自身的專業知識。

1. 中風

(1) 個案討論：約每 1~3 週舉行一次

根據當時待討論的個案數量，NIH 合作醫院通常會每隔 1 至 3 週在晨會時間安排個案討論會議。這些會議由醫師報告近期住院的病人情況，並深入探討相關的診療與處理策略。所挑選報告的病例多為較特殊或具挑戰性的疾病與臨床狀況，特別是那些未直接涵蓋在既有治療指引中的案例。例如，會討論如白血病(leukemia)病人在中風後的處理、女性中風患者的特殊狀況、紅斑性狼瘡患者出現中風的案例、因頸動脈外傷導致血栓形成而引發中風的病人，甚至是針對嚴重程度較低(NIH 中風量表評分較低)的患者直接施行動脈取栓治療等多元且複雜的臨床情境。

報告者通常為住院醫師，且多數為 PGY 3 以上，負責蒐集並整理最新的治療指引及證據等級較高的研究資料，透過分析比較這些文獻與個案的具體狀況之異同，深入淺出地呈現臨床上會遇到的問題，以及相關研究所提供的可能解決方案。接著，再由主治醫師與其他與會者提出疑問、分享見解，並針對個案處理策略進行評析與交流，促進醫療團隊間的學術互動與臨床判斷的提升。

(2) Grand round：每週一次

Grand Round 通常以期刊報告或最新研究成果的形式呈現，為研究人員和臨床醫師提供持續更新的學術資訊與討論平台。舉例來說，我曾與 NINDS 中風分支(Stroke Branch)的 Dr. Lawrence Latour 合作，Dr. Latour 長期致力於研究中風及外傷後血腦屏障(BBB)破壞的機制。他運用 Dynamic Susceptibility Contrast, DSC 磁共振造影技術，探索不穩定 infarct core 的 biomarker。他的研究發現，血腦屏障破壞越嚴重的患者，隨後大腦白質病變(white matter disease)越顯著，這一指標有潛力成為中風患者預後評估的重要參考依據之一。

在期刊報告部分，像是血栓溶解劑 Tenecteplase (TNK) 的 TRACE III 研究，及其後續的 subgroup analysis 也被納入討論。這些研究結果展示了 TNK 作為除了傳統重組組織型纖維溶酶原激活劑 (rt-PA) 之外，另一個可能有效且安全的治療選項，為臨床中風治療帶來新的希望與選擇。透過每週的 Grand Round，我能及時掌握相關領域的新進展，並與同儕及專家進行學術交流，進一步提升自身的知識與研究視野。

2. 癲癇

幾次的 Grand Round 以癲癇為主題，探討不明原因癲癇背後潛在的成因及相關風險因子。講者特別提到晚發型癲癇 (late-onset epilepsy)，指出糖尿病是其中一項重要的風險因子。此外，失智症與癲癇之間存在雙向風險關係，兩者相互影響。血清中 A β 42/A β 40 比例較高的患者，其癲癇風險亦相對升高，此發現可參考 ARIC 研究 (Atherosclerosis Risk in Communities study) 的結果。

在與死亡率相關的風險因子方面，講者列出了多項包括 SUDEP (癲癇猝死)、中風 (stroke)、失智症 (dementia)、帕金森氏症 (Parkinson's disease，風險略微升高)、以及末期腎衰竭 (ESRD，可能與糖尿病有關)。值得注意的是，癲癇患者因車禍致死的風險並未顯著增加。

NIH 也在進行一項 prospective longitudinal study，追蹤癲癇患者長達 25 年的認知功能變化，並研究低劑量 Keppra 在 subjective cognitive decline 或輕度認知障礙 (mild cognitive impairment) 患者中的治療效果。這項研究希望為保護認知功能提供新的治療策略。

在臨床評估方面，講者強調病患診斷過程中不可或缺的檢查，包括腦電圖 (EEG)、腦部影像學檢查、風險因子篩檢及認知功能測試，必要時亦需安排專科諮詢。治療時則需考慮藥物動力學特性、對器官尤其是肝臟的副作用、對認知功能的影響以及治療成本等因素。

針對最常見的顳葉癲癇 (temporal lobe epilepsy)，除了傳統的藥物治療與前顳葉切除術 (Anterior Temporal Lobectomy, ATL) 之外，講者介紹了幾種新興且有潛力的治療方式，如 Laser Interstitial Thermal Therapy, LITT、Responsive Neurostimulation, RNS。這些治療方法在改善癲癇控制方面展現出不錯成效，並且有機會成為未來的重要輔助治療選項。未來研究可能更重視表觀遺傳學 (epigenetics) 在癲癇發病機制及治療上的應用，期待能開創新的治療思路與策略。

3. 帕金森

α -synuclein 在多種神經退化性疾病中扮演著關鍵的致病角色，特別是在帕金森

氏症 (Parkinson's disease) 及相關疾病的病理機制中， α -synuclein 的異常堆積被視為核心病理變化。來自 Brigham and Women's Hospital 及哈佛醫學院 (Harvard Medical School) 的 Dr. Vikram Khurana，從 synuclein 病理學出發，探討帕金森病臨床表現的多樣性與其背後複雜的致病機制。

Dr. Khurana 強調，帕金森病的臨床狀況並非單一因子所致，而是多基因交互影響以及環境因子共同作用的結果。這些因素交織形成多樣的臨床表型 (phenotypes) 與基因型 (genotypes)，進一步影響疾病的進展與治療反應。基於這種複雜性，傳統的一刀切治療模式已難以全面應對帕金森病的多元特性。

他提出利用 α -synuclein 作為核心，設計個人化誘導多能幹細胞 (personalized induced pluripotent stem cells, iPSC) 模型，透過這些基於患者自身細胞特性的模型，可以更精確地模擬個別病患的病理過程，進而進行針對性的藥物篩選與治療策略開發，為患者提供更符合個人病理特徵的治療方案，突破傳統研究與治療的限制，為帕金森病的精準醫療開闢新的可能性。

4. 其他周邊神經肌肉疾病

教科書中所描述的典型神經肌肉疾病，例如肌肉失養症 (muscular dystrophy)，在某些情況下，臨床上可根據患者的表現型 (phenotype) 推測其可能對應的基因型 (genotype)，並利用相對簡單且經典的 Sanger sequencing 定序技術進行驗證。然而，實際臨床情況往往較為複雜，表現型與基因型都呈現光譜性變異，兩者之間無法單純以一對一的方式直接對應，這使得診斷過程充滿挑戰。

因此，除了累積臨床知識與經驗外，臨床醫師經常需要輔以更多檢查手段來確診。這些包括組織學檢查如肌肉活檢 (histopathology)、免疫組織化學染色 (immunohistochemistry) 以評估蛋白質表達情況，以及肌肉核磁共振影像 (muscle MRI)，用以觀察肌肉的結構變化和病變分布情形。另一方面，在基因層面，隨著次世代定序技術的普及，臨床常使用整外顯子定序 (whole exome sequencing, WES)、全基因體定序 (whole genome sequencing, WGS) 以及 RNA 定序 (RNA sequencing) 等多種分子診斷方法，以獲得更全面、深入的遺傳資訊，輔助診斷和理解疾病的分子機制，從而提高診斷的準確性及治療的針對性。

5. 期刊討論會：每週

因為 Dr. Fann 的牽線，我參加多場研究會議，並由此結識了 NINDS 專注於腦部影像研究的 PI Dr. Lawrence Latour，以及 Stroke Branch 的主任、專門研究流行病學的 PI Dr. Rebecca Gottesman。在他們的帶領下，我加入了一個由他們兩位麾下 post-doc 和 post-bac 們組成的自主學習 journal club，每週我們會選讀並討論一篇學術期刊文章。

這個討論會的成員們來自不同專業背景，例如 Dr. Gottesman 的 post-doc Egle Marco 擅長生物統計、Dr. Latour 旗下的 fellow Sophia 專精於動態影像層析、Nicole 熟悉研究審核行政流程和規範……。大家彼此分享專長領域的見解，切磋討論各種疾病相關的臨床知識與治療建議。除了內容本身，此 journal club 更加重視研究方法學的嚴謹性與統計工具的適用性，針對不同分析技術的優缺點展開討論。討論範圍還延伸到人工智能領域，如機器學習（machine learning）、計算機視覺（computer vision）等技術。

(三) 基因組學分析與治療的應用

1. RNA-seq and scRNA-seq

我利用幾次週間研究會議的空檔，參加了 National Cancer Institute (NCI) 下屬的計算基因組學與生物資訊部門（The Computational Genomics and Bioinformatics Branch, CGBB）的討論會。該部門主任 Dr. Daoud Meerzaman 詳細介紹了 RNA 測序（RNA-seq）的優勢與多元應用。RNA-seq 能表示細胞或基因在特定條件下的表現情況，並可比較不同基因之間的表現型或表現量，尋找特定基因的表現模式或 allele specific expression，進一步發掘潛在的生物標記、生理病理機制，乃至治療靶點。

Dr. Meerzaman 特別強調，實驗設計是 RNA-seq 成功的關鍵，包括選擇何種樣本（sample）、細胞株（cell line）、疾病類型、樣本數量（sample size）以及 biological replicate 和技術複製 technical replicate 的數量。此外，還需根據研究需求選擇適合的測序平台，目前多以 Illumina 為主流，包括 High Seq、MiSeq、Nova Seq 等，另有 Pac Bio 和 Oxford Nanopore 等平台。這些不同技術各有優缺點，操作和數據解析都需相關的知識與經驗。

講者示範了利用 Galaxy 平台操作 RNA-seq 的完整流程，從載入樣本資料、進行 pre-alignment QC、序列拼接與合併（assembly and merging），到分辨目標基因表現量及數據視覺化等。RNA-seq 的優點包括高通量以及無需預先知道序列，但其結果高度依賴樣本品質，成本較高，且數據分析和解讀需要強大的計算能力和生物資訊專業知識。

數週後，我又參加了來自波士頓大學（Boston University）的 Brian Cleary 教授主講的 single cell seminar。講座內容聚焦於單細胞轉錄組學（single cell transcriptomics），利用 CRISPR 在細胞中 knock-in 或 knock-out 不同基因，每個細胞因而有了這些基因被改變的 perturbation barcode，接著用 scRNA-seq 看基因表現變化；其優點是可以同時分析多個基因被改變後的 transcriptom 變化。

此外，講者還介紹了 Perturb-FISH 技術，這是一種結合 CRISPR 與空間 RNA 分

析（如 fluorescent in situ hybridization, FISH）的方法，不僅能了解基因改變的種類和表現變化，還能定位這些變化在組織中的空間位置。這項技術在臨床上有潛力可應用於免疫反應、癌症轉移研究及探索新的藥物治療靶點。而諸如此類的 computational genomics 和 computational drug discovery 是當前研究的重要焦點。

2. 使用病毒載體(adeno-associated virus [AAV] vector)作基因治療。

AAV (adeno associated virus) 技術已經應用於脊髓性肌肉萎縮症 (SMA) 的基因治療，例如著名的 Zolgensma 療法。講者進一步介紹了利用 in vivo AAV 作為載體，將設計好的基因片段直接送入腦部，以治療在腦部發生的疾病影響。例如針對 GM1 gangliosidosis 這種罕見遺傳疾病，患者因為溶酶體酶 β -galactosidase 缺乏，導致神經元細胞膜上 ganglioside 過度堆積，且在腦脊髓液中濃度升高，最終引發腦部退化、萎縮甚至死亡。利用 AAV 載體送入正常基因，期望能糾正此酶的缺陷，達到治療效果。

目前在動物模型中，已經獲得了這種治療方式的 proof of concept，但在臨床應用之前，必須嚴格評估其安全性。AAV 作為病毒載體注入體內，其用量通常遠高於自然病毒感染的水平，這可能誘發免疫反應，包括炎症反應、細胞激素釋放、補體系統活化，甚至出現血栓性微血管病變 (thrombotic microangiopathy) 等不良副作用。此外，體內產生的某些抗體可能活化 adaptive immune，進一步增加治療風險，這些都為療效和安全性帶來挑戰與變數。

目前針對 GM1 gangliosidosis 的 AAV9-GLB1 基因治療研究已進入第一期臨床試驗 (phase 1)，重點在於安全性評估。初步資料顯示，大多數副作用為輕微且可控，並未出現嚴重不良事件。同時，治療成效也相當令人鼓舞，患者在擴散張量影像 (DTI) 中白質纖維束的數量和體積均較對照組顯著提升，顯示治療可能改善了腦部結構和功能。這項研究不僅為罕見神經退化疾病的治療帶來新希望，也推動基因治療技術在神經科學領域的應用邁出重要一步。

(四) 人工智慧

參加 Anaconda 在 NIH 辦的說明會，這個軟體包含上百個 Python 套件，且具跨平台套件整合功能。並且與先進們切磋學習，練習操作和資料分析應用，用 Jupyter Notebook 建立資料分析環境。

在這一年期間，我參與 NIH 內部的 coding club 與新成立的 AI club，學習使用 RStudio 操作機器學習技術，包括 Decision trees、Survival trees、Random forest 以及資料視覺化。同時也學習使用 Google Colab、Pandas、PyTorch 等工具進行資料分析與模型開發。

隨著大型語言模型（LLM）在醫學領域的快速發展，越來越多應用被開發出來，像是輔助診斷、改善病歷書寫與醫病溝通等。NIH 的國家醫學圖書館（NLM）和國家癌症研究院（NCI）共同合作，開發了一套名為 TrialGPT 的 AI 演算法，該系統能快速且精準地根據患者的臨床資料，從 ClinicalTrials.gov 中篩選並配對合適的臨床試驗，還能生成解釋患者符合或不符合條件的原因，最終提供臨床醫師一份推薦清單，大幅提升試驗配對效率與參與率。

此外，我與同期訪問學者合作，利用 Unsloth 工具微調現有的大型語言模型（GPT 和 Llama），開發基於病歷紀錄預測患者近期中風風險的模型。我們設計生成回答需符合固定邏輯流程（chain of thought），並與實際病人近次是否中風的臨床結果及專家判斷做比較，評估模型的預測能力，並針對模型的邏輯性、思路連貫性及幻覺（hallucination）現象進行優缺點分析。

值得一提的是，多位人工智慧領域專家認為 LLM 是未來通用人工智慧（AGI，Artificial General Intelligence）最有可能的發展方向。NIH 開發的 GeneAgent 不只是簡單生成內容，而是會主動將生成的資訊與專家審核的權威資料庫進行比對，產出一份詳細的驗證報告，說明資訊的正確性。GeneAgent 主要協助科學家解讀龐大基因資料，找出相關的生物路徑（biological pathways）或功能模組（functional modules），幫助理解疾病與基因網絡間的關聯。

因為現有 LLM 多基於大量網路文字訓練，擅長「生成語句」但不一定能「辨識真假」，導致 AI 幻覺問題屢見不鮮，即使模型自我循環驗證，也可能強化錯誤資訊，這在基因分析這類需高準確度的應用中不容忽視。GeneAgent 透過「自我驗證模組」（self-verification agent module），將 LLM 生成的基因功能描述一一對照外部權威資料庫，確認哪些主張正確、部分正確或錯誤，形成一套自動生成並自我檢查的架構。這套系統在精準醫療、基因體學及藥物開發領域具有極大應用潛力。

（五）資料庫與平台操作技術

1. NIH 的高效能運算 Biowulf

Biowulf 是 NIH 高效能運算團隊（High Performing Computation, HPC）所管理的超大型 Linux cluster，專門為 NIH 內部的研究社群提供強大且高效的計算資源與技術支援。該系統擁有超過 95,000 個處理核心以及超過 40PB 的儲存容量，能夠同時處理大量並行計算任務，例如基因序列比對和分析，或是需要大規模記憶體資源的分子動力學模擬等運算工作。

Biowulf 系統內建並預先安裝多種科學研究常用軟體和資料庫，支援 R、Python 等主流程式語言，方便研究人員快速上手與應用。在架構和功能上，Biowulf 與台灣的國家高速網路與計算中心（National Center for High-performance

Computing, NCHC) 具有異曲同工之妙，兩者皆結合超級運算、高速資料傳輸及雲端服務，並逐步導入 AI 雲計算平台（如 TWCC）來推動科研與產業創新。

在我於 NIH 進修期間，透過指導教授的協助，申請了 Biowulf 的使用權限，並向其他研究員學習其操作流程與技巧，實際運用於基因資料的分析及計算任務。

2. MIMIC III 與 MIMIC IV

在我們利用大型語言模型進行中風風險預測的研究中，主要採用了 MIMIC III (Medical Information Mart for Intensive Care III) 這個開源資料庫。透過運用大數據作為預訓練模型的基礎，最後比較微調後模型與現有大模型語言模型的預測效能。

MIMIC III 是由麻省理工學院 (MIT) 與貝斯以色列女執事醫療中心 (Beth Israel Deaconess Medical Center, BIDMC) 共同建立的一個大型公開資料庫，內含去識別化的重症病患臨床資料，詳細記錄病患基本資料、生命徵象、檢查檢驗報告、用藥紀錄等資訊。此資料庫為研究者提供了豐富且真實的臨床資料，廣泛應用於機器學習與人工智慧在醫療領域的開發與驗證。目前 MIMIC III 已有更新版本 MIMIC IV，包含更近年的臨床數據。

3. All of Us 資料庫

在 NIH 進修期間，我申請並通過審核，取得加入 All of Us Research Program 醫療資料庫的使用權。該計畫是由 NIH 主導的大型國家級研究計畫，旨在透過收集多元族群的健康資料，包括電子病歷 (EHR)、血液檢測數據、長期追蹤結果及穿戴式裝置蒐集的生理數據，協助研究者深入探討遺傳因素、環境、生活方式與健康結果之間的關聯，進而推動個人化精準醫療的發展。

All of Us 在使用權限管理方面實施嚴格把關，從申請階段到每次登入均受到嚴密監控，確保資料安全與使用合規。雖然資料庫持續擴充中，但在離開 NIH 後，使用者的存取權限會自動取消。

該計畫特別強調族群多樣性與包容性，致力納入來自不同族裔、生活背景與地區的參與者，使研究結果更具代表性與普適性。藉由整合基因組學、醫療紀錄、生活型態及環境因素等多維度資料，研究者能更精準地預測適合特定個體的治療或預防方案，推動精準醫療的實現。此外，利用 All of Us 資料，還有望發掘新的疾病風險因子、創新藥物靶點，以及開發針對不同族群的差異化療法。

4. Brain MRI datasets with MATLAB

學習使用 MATLAB 進行腦部 MRI 影像的分析。MATLAB 是由美國 MathWorks 公司開發的高階程式語言與互動式運算環境，廣泛應用於資料分析、視覺化及

機器學習等領域。其內建多種 **toolboxes**，專為不同應用領域設計，提供豐富且專業的函式庫與工具，使影像處理與分析更為便利且高效。

5. NCBI 資料庫、INMAP、DeGAUSS

因為自己和其他訪問學者的研究主題需要，我在這段時間也學習利用一些線上開源的大型資料庫，獲取一些去識別化的基因資料，與輔助整合環境因子資料例如區域空氣汙染等，進行資料整合與分析。

三、心得

謝謝院長、副院長及各位院部長官，給予我這次寶貴的機會，能夠前往美國國家衛生研究院（NIH）進行學術進修。特別感謝神經醫學中心的黃金安主任與陳柏霖主任，感謝他們引薦我認識並跟隨 Dr. Fann 學習，並在我於美國期間持續關心我的學習進度與生活狀況，給予許多寶貴的指導與提點。此外，也非常感謝科內所有同仁，在我出國期間協助分攤工作。還有研究部的同仁們，因為我的部分研究主題使用本院自有的資料庫，運算與資料處理也非常仰賴研究部同仁們的大力支持與協助，令我感激不盡。

在 NIH 期間，我參與了許多 **intramural** 與 **extramural** 的研究會議，除了廣泛吸收新知之外，也深刻感受到美國政府對科學研究持續且穩定的資金投入。許多尖端甚至是新興的測試技術、軟硬體設備累積多年，花費甚鉅，但正是因為有這樣充足的經費支持，各個 **intramural** 和 **extramural** 的研究項目才能夠持續推進與發展。此外，業界也積極推廣商業化軟體與平台服務，如各大雲端運算平台 Google 的 GCP、Amazon 的 AWS、Microsoft 的 Azure 及 365 Copilot 等，這些工具與擴充服務大幅提升了研究工作效率。過去我較少接觸或甚至未曾了解的智慧資料整理及檢索功能，也在這一年的進修期間有了實際操作與練習機會。

在此次進修過程中，我與同期的研究人員互相交流學習，彼此分享各自專長。由於我過去覺得統計學是自己的弱項，因此特別選修了幾門由 NIH 圖書館舉辦的生物統計課程，並且參加每週的 **journal club** 自主學習。在討論會中，Dr. Gottesman 實驗室的博士後研究員們會將文章中所使用的統計方法講解得淺顯易懂，甚至比較不同統計方法的優缺點，讓我們這些非統計背景或公共衛生背景的參加者能夠更快速理解文章內容，掌握研究的優勢與限制。此外，我也從多位 PI 及訪問學者身上學習到 **bioinformatics** 資料的前處理、整理、品質控管（QC）及結果解讀與評估，彌補了自己過去在這些環節的知識空缺。

NIH 定期舉辦多場演講與研討會，廣邀國內外該領域的頂尖專家進行講座與授課。這些會議多數皆有錄影，並保存於 NIH 的 **Videocast** 網站及 **YouTube** 平台，方便研究人員會後反覆觀看與複習，或是接觸到先前錯過的內容，對於科研知識的傳播與推廣具有極大的幫助，也讓我在專業知識上有許多收穫。

四、建議事項

(一) 整合全院的資料庫成為中心化的資料庫平台

因政策規範和隱私等因素，NIH clinical center 的資料並不會輕易釋出，訪問學者們幾乎都是使用原本自己的資料庫或線上開源資料庫。此時便深感擁有案量充足品質優良的資料庫的重要性。在過往和現在各位長官的推動下，各資料庫已逐漸發展擴充，但像是做基因數據分析時，數據量不過往往無法得到足以令人信服的結果，且 GWAS-PheWAS study 更是需要結合良好品質的紀錄，與減少不同族群 epidemiology 的歧異，因此持續擴充精進各科的資料庫實屬必要。

據我個人在 NIH 進修及使用 All of Us 等資料庫的經驗，發現他們也是將這些龐大的資訊集中放在一個中心化的資料庫或平台，將所有臨床資料、影像資料、基因資料統一做整合，由研究者向平台管理員提出申請，取得去辨識的資料。

將來若軟硬體功能許可，可考慮比照這幾個超大型資料庫，將這些多項資料整合到統一平台。過程必定耗時費工：許多資料散落在各個部科，而且需要另闢足夠大的儲存空間供各部科上傳資料；上傳的資料需要去識別處理、並轉化成可分析的統一格式；最後還要建置可用易操作的管理與查詢系統模式。

(二) 提高出國開會及進修管道多樣性

出國期間結識了幾位來自台灣、在 NIH、John's Hopkins Hospital、跟其他學校與醫院工作或讀書的學者，許多都是拿台灣研究所的獎助學金，或者國科會千里馬計畫補助。對年輕學者而言，的確可大大提高出國參加國際會議或長期進修的意願，開拓自身視野的同時也可為科內注入創新思維。建議各科主治醫師出國進修可增加多樣性，例如對於基因研究有興趣的醫師或研究員，鼓勵參與 American Society of Human Genetics (ASHG) Annual Congress 和 American College of Medical Genetics and Genomics (ACMG) Annual Clinical Genetics Meeting，對於影像尤其磁共振造影研究有興趣者，可參與 ISMRM (International Society for Magnetic Resonance in Medicine) Annual Meetings。

(三) 延攬基因研究人才並整合臨床及數據分析專業

基因診斷與治療的發展一日千里，雖學術原理有機會自行查找，但操作分析時十分仰賴有經驗的 bioinformatics 人才做 coding。我在進修期間也經常請研究部同仁們幫忙。期許能延攬並留住這些優秀的生物資訊學者和研究員們，並且期待未來各部科研究能持續與之合作。建議可以如現行研究部的組織，組成由基因研究人員、臨床醫師及數據分析專家共同參與的多學科團隊，確保研究與臨床需求緊密結合，並且適當補助他們參加專業培訓和研討會的花費，提升其參與意願。

(四) 加強外部專家合作

進修期間除了 Dr. Latour 團隊、Dr. Gottesman 團隊，還結識了幾位來自台灣、跟隨 Dr. Fann 學習的訪問學者。各自的專長有急診醫學、數據分析、高齡醫學、癌症基礎醫學研究、影像分析、自然語言分析等。以及在大華府地區工作或進修的學者，他們進修結束後多數回到台灣在各自的領域耕耘，工作地點遍佈北中南。我也維持和老師及這幾位同好學者的聯繫，持續合作。目前與 NIH 范揚政老師、台北醫學大學簡碩宸博士與高雄榮總徐盈欣醫師合作撰寫文章關於語言模型的應用。影像分析方面，與 NIH Dr. Latour、成功大學電機系詹寶珠教授、邱建裕博士接洽。基因組學分析操作，NIH Dr. Hirotaka Iwaki 可提供協助。期待各界跨領域的交流能碰撞出不一樣的火花。

五、 附錄

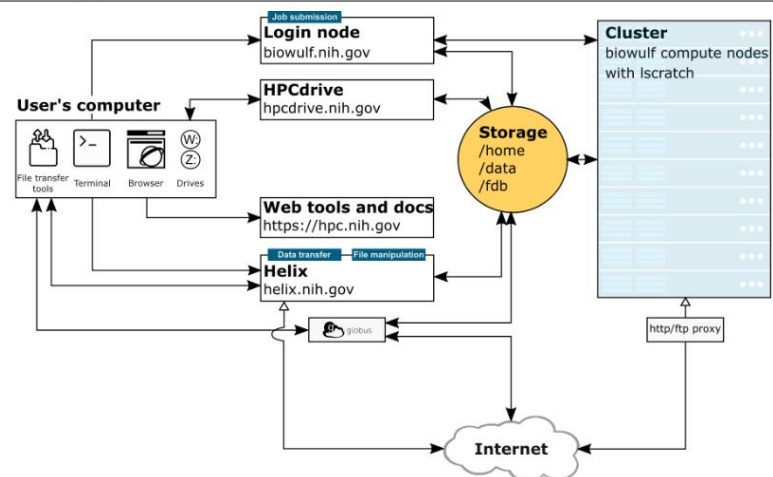


▲Clinical center (building 10)，也是我們最常活動的據點、辦公室所在、多項會議舉辦的地點。



▲NIH Director 和主要首長辦公室所在的 Building 1，Dr. Fann 辦公室也在此

NIH HPC Systems



▲ 高效能運算團隊 High Performing Computation (HPC) group 架構示意圖



▲ 和 NINDS PI Lawrence Latour 合影



▲ 和 CARDS (Center for Alzheimer's and Related Dementias) PI Hirotaka Iwaki 合影，部分研究與之合作



▲和同期訪問學者：高雄榮總高齡醫學科徐盈欣醫師及中國附醫急診科張裕鑫醫師在 Building 10 Clinical Center 合影



▲與來 NIH 參訪的中榮研究部陳一銘主任及研究部同仁們和 Dr. Fann 在 building 1 前合影



▲NIH 工作識別證