出國報告（出國類別：研習）

參加2025觀察性健康資料科學與資訊、人工智慧和真實世界實證研習出國報告

服務機關：衛生福利部疾病管制署
姓名職稱：陳龍生研究員
派赴國家/地區：美國紐約州紐約市
出國期間：2025.7.12-2025.7.20
報告日期：2025.8.25

摘要

國際觀察性健康數據科學與資訊學聯盟(The Observational Health Data Sciences and Informatics, OHDSI)首度於 2025 年 7 月 14-18 日於美國紐約州紐約市哥倫比亞大學醫學資訊研究所辦理 Summer School 課程，主要的參加者為臨床醫師、從事觀察性研究的研究者、相關領域(醫學資訊、流行病學及公共衛生學)在學學生等約 40 人參加這次暑期課程。

依據本次課程安排，主要分為理論架構說明、研究方法與工具介紹及藉由實際使用 OHDSI 標準化資料庫進行實作練習。本次研習內容中強調如何透過 OHDSI 的標準化資料庫與開放式研究過程，產生真實世界可靠的臨床實證。其中著重介紹當前透過使用健康照護資料進行觀察性研究面臨的研究方法挑戰，並示範在 OHDSI 架構下的解決方案。

透過參與這次暑期課程，更深入了解OHDSI的標準化資料庫與開放式的研究過程，及其處理當前觀察性研究所面臨之方法學挑戰與解決方法。透過課程，並與不同專科醫師或不同專長學員，共同實作一項觀察性研究分析，以更深入了解 OHDSI 的標準化資料庫與開放式的研究工具之應用。本次所習之相關技術與經驗，可以作為本署之後應用健康照護資料，進行觀察性研究時之重要參考。

目錄

壹、目的

國際觀察性健康數據科學與資訊學聯盟(The Observational Health Data Sciences and Informatics, OHDSI)是由美國哥倫比亞大學為首之全球性、非營利跨學科和開放科學的網絡聯盟，其願景為透過觀察性健康資料，利用大數據分析和人工智慧等方法，提升臨床醫學資料價值，實現跨領域的多方研究合作，以提供真實世界實證，解決健康領域問題。

OHDSI 所建立之研究合作平台、健康數據整合標準、分析技術及產製真實世界證據之開放科學架構，可作為我國推動循證決策（evidence-based policy-making）於新興傳染病防疫及政策評估之重要參考。

過去 OHDSI 官方組織每年主要透過辦理論壇方式，在全球各地推廣 OHDSI 資料標準與架構，今年首次於美國紐約州紐約市哥倫比亞大學醫學資訊研究所辦理為期 1 週的暑期課程，針對從事觀察性研究有興趣之臨床醫師、研究者，提供完整理論與實作課程，藉以推廣，讓更多專業人士使用 OHDSI 之標準與架構及相關研究分析理念。

故規劃參加本次暑期課程，研習該組織跨領域合作機制、所發展之健康數據整合與標準及工具，並實際參與從健康資料到真實世界實證產製之案例研析。

貳、過程

一、出國行程摘要

日期：114 年 7 月 12 日(星期六)至 7 月 20 日(星期日)

日期(臺灣時間)	行程
TPE 7 月 12 日(週六)	臺灣桃園國際機場直飛前往美國紐約州紐約市
TPE 7 月 13 日(週日)	抵達美國紐約州紐約市
TPE 7 月 14 日(週一)	參加 OHDSI 研習第 1 日課程
TPE 7 月 15 日(週二)	參加 OHDSI 研習第 2 日課程
TPE 7 月 16 日(週三)	參加 OHDSI 研習第 3 日課程
TPE 7 月 17 日(週四)	參加 OHDSI 研習第 4 日課程
TPE 7 月 18 日(週五)	參加 OHDSI 研習第 5 日課程，研習課程結束後直接前往紐約市甘迺迪國際機場(無住宿)
TPE 7 月 19 日(週六)	紐約市甘迺迪國際機場搭機直飛臺灣桃園國際機場(航程)
TPE 7 月 20 日(週日)	抵達臺灣桃園國際機場

二、研習課程表

本次研習課程設計，包含理論說明、系統環境介紹及實際分析案例練習，課程安排上，依照研究分析實務流程，先介紹研究資料來源、共通性資料標準模式、系統設計概念與操作介面說明、研究分析的三個類別描述性分析、推論性分析及個人化預測模型簡介與案例分享、實際分析案例則是透過分組討論、提出研究假說與實際驗證假說及研究成果展示等。

	Monday	Tuesday	Wednesday	Thursday	Friday
8am-9am	**IT test:** VM and ATLAS				
900am-1200pm	**Lecture:** Introduction to observational network research **Lecture:** Understanding administrative claims (Merative) **Lecture:** Understanding electronic health records (CUIMC) **Exercise:** Mapping a Patient journey	**Lecture:** Designing an observational network study: Characterization **Lecture:** Designing an observational network study: Estimation	**Lecture:** Phenotype development **Demo:** Build cohorts in ATLAS	**Lecture:** Implementing an observational network study using OHDSI framework **Demo:** Create input specifications, execute package	**Lecture:** Interpreting results from an observational network study and building trust in evidence - Openness and verification - Diagnostics - Evidence Synthesis **Demo:** Review Rshiny app
12pm-1pm	Lunch	Lunch	Lunch	Lunch	Lunch
100pm-500pm	**Lecture:** OMOP Common Data Model **Exercise:** Exploring the OHDSI Standardized Vocabularies in ATLAS **Exercise:** Exploring patient profiles in ATLAS **430pm: IT test**: execute R package	**Lecture:** Designing an observational network study: Prediction **Exercise:** Framing your research question using OHDSI standard questions **Exercise:** Evaluating data network fitness-for-use using ATLAS/Data Sources	**Exercise:** Build cohorts in ATLAS **Lecture:** Phenotype evaluation - CohortDiagnostics - PheValuator - KEEPER **Exercise:** Review CohortDiagnostics results	**Exercise:** Create input specifications, execute Strategus	**Exercise:** Review RShiny app and present findings
5pm-6pm	Group photo from DBMI Terrace				School's out for summer!

三、本署與會人員

衛生福利部疾病管制署：陳龍生研究員。

四、課程重點

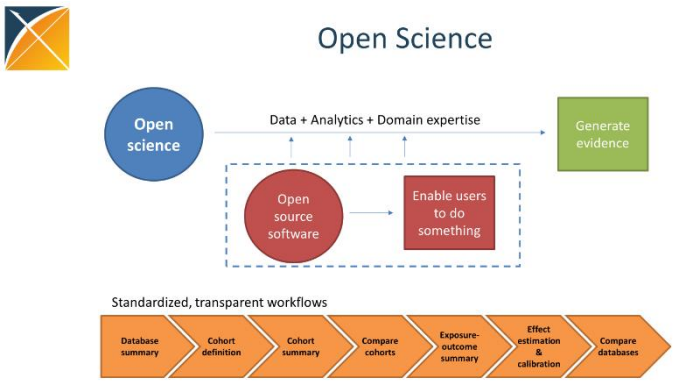
(一)OHDSI 簡介

觀察性健康資料科學與資訊學(Observational Health Data Science and Informatics, OHDSI)計畫是一項多方利益相關者參與的跨學科合作計畫，其目的在透過大規模分析挖掘健康數據的價值。

OHDSI 計畫的使命為透過賦能研究社群，使其能夠協作產生證據，從而促進更明智的健康決策和更優質的醫療服務，以改善健康。多年以來 OHDSI 已建立了一個由研究人員和觀察性健康資料庫組成的國際網絡，並在哥倫比亞大學設有中央協調中心。

目前 OHDSI 共有來自世界 83 個國家的 4,294 位協作者參與，包含醫療資訊學、統計學、流行病學、臨床科學等領域。已發表超過 700 篇研究論文，且對歐洲藥品管理局(EMA)與美國食品藥物管理局(FDA)在應對 COVID-19 疫情時有具體影響力。

OHDSI 計畫採用開放科學架構來生成重要的實證證據，結合資料、分析與領域專家產生真實世界的實證。透過開放原始碼的軟體，使參與的研究者具備相同的研究能力，可以進行相關研究，擴大研究規模，以處理觀察性研究的方法學挑戰。OHDSI 計畫制定一系列與研究有關的工作流程，透過標準化與透明化這些工作流程，達到開放式科學架構。



Data Source: George Hripcsak. 2025. Interpreting results from an observational network study and building trust in evidence. Summer School in OHDSI, AI, and Real-World Evidence.

前述所提到的標準化工作流程中，最為核心的是 OMOP Common Data Model 與 OHDSI standardized vocabularies 兩項。

- OMOP Common Data Model 是觀察性醫學結果夥伴關係通用資料模型的英文簡稱(The Observational Medical Outcomes Partnership Common Data Model)，最早開始於 2008 年，OMOP CDM 成為一開放的資料標準，目的在標準化臨床觀察性醫學資料的結構和內容，以實現高效率的分析，進而產生可靠的證據。

OMOP CDM 目前是由 OHDSI 負責更新與推廣。

- OMOP CDM 的主要目標：
    - 標準化資料：將來自不同來源、格式和術語的醫療資料統一化，解決了資料不一致性的問題。
    - 促進合作研究：由於資料都已經轉換成共同的格式，研究人員可以更容易地分享、整合和分析來自不同機構的資料，從而進行更大規模、更可靠的研究。
    - 支援標準分析工具：這種標準化的結構使得開發和使用一套通用的分析工具成為可能。OHDSI 社群提供了一系列開源工具（如 ATLAS），研究人員可以用它們來進行藥物安全監測、比較療效研究、或建立病患預測模型等。
    - 提高資料互通性：它確保了不同的資料庫可以無縫地一起使用，這對於大型的跨國研究尤其重要。
- OHDSI 標準化詞彙表(OHDSI standardized vocabularies)：OHDSI 的詞彙庫允許對醫療術語進行組織和標準化，以便在 OMOP 通用資料模型的各個臨床領域中使用。簡單來說，這些術語集就像是醫療資料的通用字典或辭典。在醫療領域，有很多不同的編碼系統和術語來描述同一個概念，例如：
    - 疾病： 國際疾病分類 (ICD-10)、SNOMED CT
    - 藥物： RXNORM、ATC
    - 實驗室檢驗： LOINC
    - 醫療處置： CPT4、HCPCS

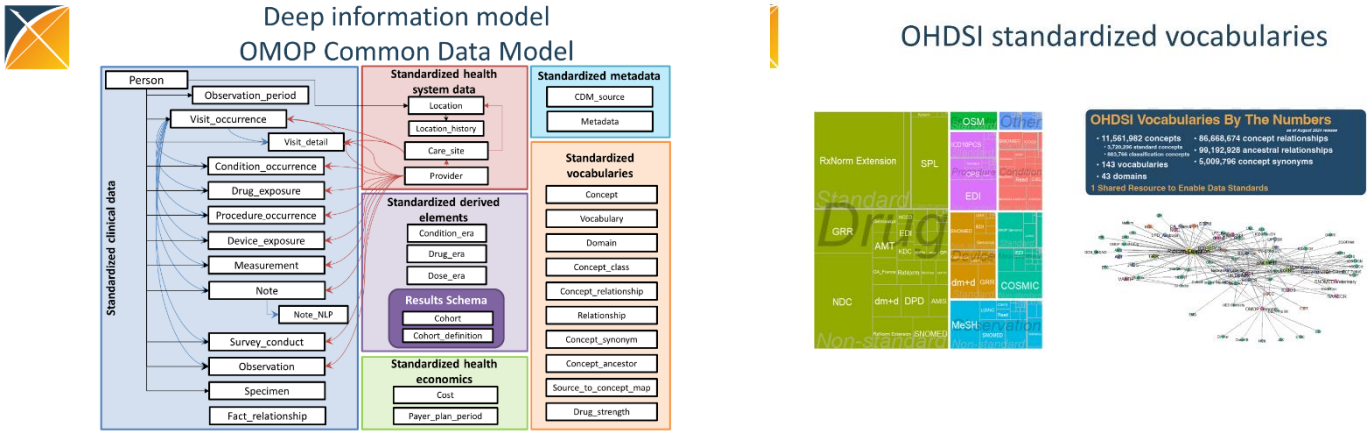
  透過標準化，可以實現跨不同醫療資料庫的協調和互通性，將各種不同的本地編碼系統翻譯成 OMOP CDM 可以理解和用於研究的單一通用語言。
- OHDSI 標準化術語集的核心功能
    - OHDSI 的術語集將這些不同系統的醫學概念，全部映射 (map) 到一個統一的標準概念上。這個過程有兩個關鍵步驟：
    - 收集與整合： OHDSI 持續從世界各地收集各種醫療編碼系統，並將它們整合到一個大型的知識庫中。
    - 映射與標準化： 透過複雜的演算法和人工審核，將不同來源的編碼（例如，ICD-9 的「401.9」和 ICD-10 的「I10」）都映射到一個共同的標準概念，例如 SNOMED CT 中的「高血壓 (Hypertension)」概念。
    - 當你將原始醫療資料轉換成 OMOP CDM 格式時，這些非標準的代碼（Non-standard Concepts）會被轉換為統一的標準概念 (Standard Concepts)。

OHDSI standardized vocabularies 提供了一個標準化架構，可以實現跨不同醫療資料庫的協調和互通性，將各種不同的本地編碼系統翻譯成 OMOP CDM 可以理解和用於研究的單一通用語言。也使得研究人員能夠：

- 定義表徵(phenotypes)：根據一致的標準，準確地識別特定疾病或病症的患者群組，即將臨床狀況轉化為一組標準化代碼。

- 建構特徵變項：創建一致的變數，用於機器學習模型和統計分析。
- 標準化分析：透過確保相同的術語在不同資料來源中具有相同的含義，來促進大規模、分佈式的研究。

同時也消除了研究人員為每個新資料庫編寫自訂代碼的需要，可以加速了科學研究發現的速度。



Data Source: George Hripcsak. (2025). Interpreting results from an observational network study and building trust in evidence. Summer School in OHDSI, AI, and Real-World Evidence.

透過 OHDSI 所進行的觀察性研究，以產製真實世界實證，主要可以分為三大類型：

臨床特徵描述 – 描述性研究(tally)

- 自然病史：誰患有糖尿病，誰服用 metformin？
- 品質改善：患有糖尿病的患者中，有多少比例會出現併發症？

群體效應評估 – 因果性推論與估計(cause)

- 藥品安全性監測：服用 metformin 會導致乳酸性酸中毒(lactic acidosis)嗎？
- 藥品效果比較：服用 metformin 比 glyburide 更容易導致 lactic acidosis 嗎？

病人癒後預測 - 預測(predict)

- 精準醫療：根據所有關於個別病人已知的資訊，評估如果病人服用 metformin，發生 lactic acidosis 的機率為何？
- 疾病干預：根據所有關於個別病人已知資訊，評估個別病人罹患糖尿病的機率為何？

OHDSI LEGEND (Large-Scale Evidence Generation and Evaluation in a Network of Databases)

LEGEND 是 OHDSI 社群開發的一個自動化分析框架。它不是一個資料庫或資料模型，而是一套系統性、大規模的實證研究方法論。LEGEND 的核心理念是：利用標準化的 OMOP 通用資料模型 (CDM)，對全球多個分散的醫療資料庫，進行大規模、系統性、且自動化的實證生成與評估。傳統的醫學研究通常是針對一個特定問題，手動設計並執行一個研究，例如：「比較 A 藥和 B 藥對某種疾病的療效」。LEGEND 則更進一步，它能自動化地同時比較數千種藥物、數百種疾病，以及各種

不同的結果，並且在多個資料庫中進行重複驗證，以確保研究結果的可靠性。

LEGEND 的主要特點

1. **大規模 (Large-Scale)**：LEGEND 的分析規模非常龐大。它能同時評估數以萬計的藥物治療方案和醫學觀察結果，而這在傳統研究中是幾乎不可能實現的。
2. **自動化 (Automated)**：這套框架將研究設計、資料分析、結果生成等步驟高度自動化。研究人員只需要定義好分析問題，LEGEND 就能自動處理後續的繁瑣工作。
3. **系統性 (Systematic)**：LEGEND 採用一套標準化的研究設計和統計方法，以確保分析的嚴謹性和可重複性。它能夠系統性地生成因果效應估計，並自動進行敏感度分析，以檢測結果的穩健性。
4. **網路化 (Network)**：LEGEND 的設計能夠在一個由多個 **OMOP CDM 資料庫**組成的全球網路中運行。這意味著，同樣的分析可以在美國、歐洲、亞洲等多個資料庫中同時執行，然後將結果進行匯總，而不需要直接交換底層的病患資料，大大保護了資料的隱私。

LEGEND 主要應用於**藥物流行病學**和**比較性療效研究**。透過 LEGEND，研究人員可以：

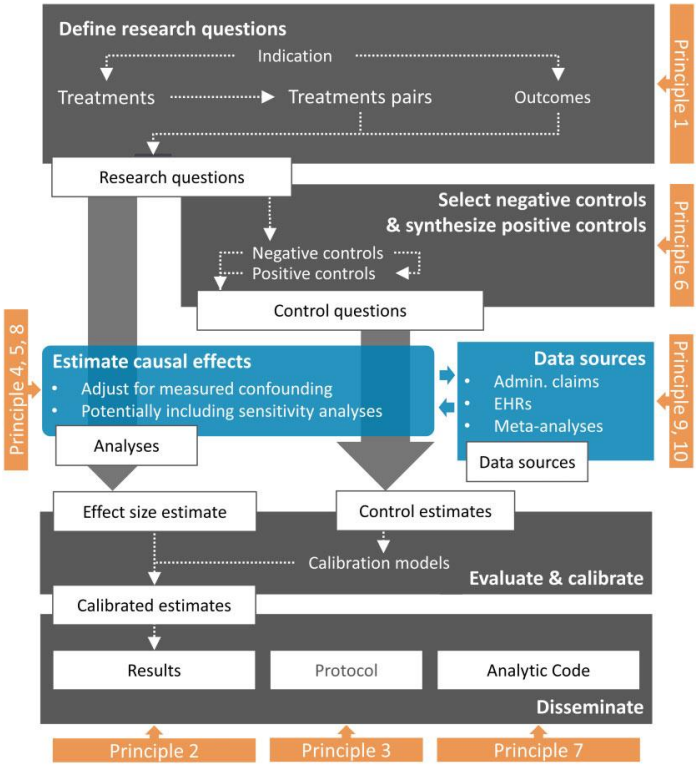
- **比較數種藥物的療效與安全性**：例如，比較所有用於治療高血壓的藥物，哪一種在特定族群中效果最好且副作用最少。
- **檢測未知的藥物副作用**：自動化地分析藥物與各種潛在不良事件之間的關聯性，從而發現傳統研究可能忽略的風險。
- **生成真實世界證據 (Real-World Evidence, RWE)**：為醫療決策者和監管機構提供來自真實世界病患資料的可靠證據。

總結來說，**OHDSI LEGEND** 是一項創新性的科學方法，它結合了 **OMOP CDM** 的標準化優勢，利用自動化和大規模分析的能力，旨在系統性地、大規模地從全球真實世界的醫療資料中，生成可靠且可重複的醫學證據。

OHDSI LEGEND 提出十項原則，用於解決觀察性研究中常見的偏誤(Bias)，包含因未測量的干擾因素而產生的偏差(residual confounding)、P 值操弄(P hacking)，以及出版偏誤(publication bias)。透過開放原始碼軟體和跨資料庫網路的協作，LEGEND 強調透明度、可再現性和結果的一致性，同時確保病患資料的機密性。研究中採用已知答案的對照問題來評估系統性誤差，並對統計結果進行校準，進而提升生成證據的可信度與應用性。

(OHDSI LEGEND)產生可靠證據的 10 項原則：

1. LEGEND 將大規模地產生證據。
2. 證據的傳播不會取決於估計的效應。
3. LEGEND 將使用預先指定的分析設計來產生證據。
4. LEGEND 將透過對所有研究問題一致地應用系統化流程來產生證據。
5. 不會有人為操控結果的情況。
6. LEGEND 將使用最佳實踐來產生證據。
7. LEGEND 將透過使用控制問題來進行實證評估。
8. LEGEND 將使用開源軟體來產生證據，該軟體可供所有人免費使用。
9. LEGEND 不會用於評估新方法。
10. LEGEND 將透過多個資料庫的網絡來產生證據，並將維護資料機密性；網絡中的站點之間不會共享患者層級的資料。



Data Source: Martijn J. Schuemie, Patrick B. Ryan, Nicole Pratt etc.. (2020). Principles of Large-scale Evidence Generation and Evaluation across a Network of Databases (LEGEND). JAMIA, 27(8), 1331–1337.

(二)Summer School 上課方式與授課講師

觀察性健康資料科學與資訊學、人工智慧及真實世界證據 Summer School 將透過

- 大型資料庫實作，進行臨床特徵描述、群體效應評估與病人癒後預測研究
- 依照觀察性研究設計及 LEGEND 10 項原則，為每個應用案例設計觀察性研究，應用 OHDSI 社群的開源工具，並使用真實世界資料來源，進行分析與產製實證。
- 過程中會搭配分析方法的基礎講座，及由講師們帶領的實務操作與小組互動式練習，來完成研究專案。

本次 Summer School 授課講師：

George Hripcsak, MD, MS

Vivian Beaumont Allen Professor, Department of Biomedical Informatics (DBMI) Columbia University

Patrick Ryan, PhD

Adjunct Assistant Professor, Department of Biomedical Informatics (DBMI) Columbia University, Johnson & Johnson

Anna Ostropolets, PhD
Adjunct Assistant Professor, Department of Biomedical Informatics (DBMI) Columbia University, Johnson & Johnson
Karthik Natarajan, PhD Assistant Professor, Department of Biomedical Informatics (DBMI) Columbia University

負責協助實做電腦環境設定之技術人員：
Mark Velez
Clinical Research Programmer Analyst, Department of Biomedical Informatics (DBMI) Columbia University

(三)課程所使用之健康資料庫介紹

本次課程實作部分，會使用美國聯邦醫療保險(Medicare)、美國聯邦醫療補助保險(Medicaid)費用申報資料庫及哥倫比亞大學醫學中心(Columbia University Irving Medical Center, CUIMC)的電子健康紀錄資料庫(Electronic Health Records, EHR)。

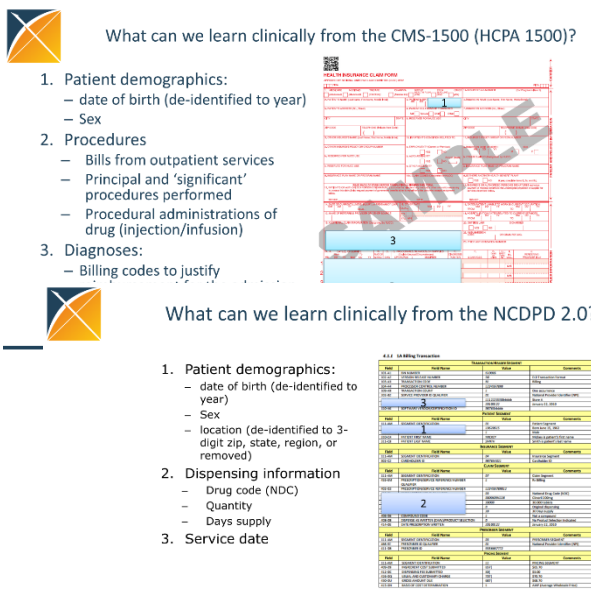
醫療保險費用申報資料庫(administrative claims database)

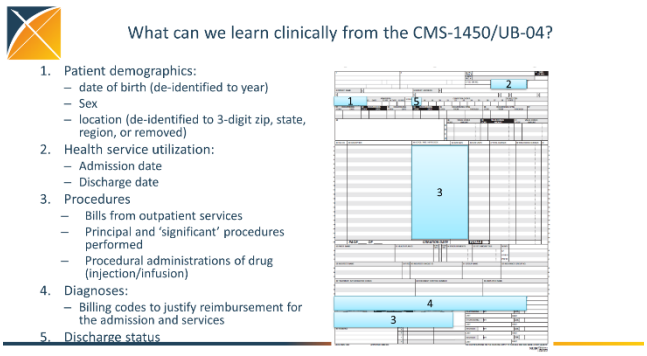
美國醫療保險費用申報資料庫，有幾個共通性的表格格式，包含

- CMS 1450 (UB-04)：醫院、護理機構、住院及其他設施供應商使用的表格。
- CMS 1500 (HCFA-1500)：個體醫生和供應商、護士及專業人士，包括治療師、脊椎治療師和門診診所使用的表格。
- NCPDP D.0：門診藥局處方配藥索賠的電子傳輸標準。

此外，也有統一使用之代碼系統，如：

- 疾病診斷碼：ICD-9/10-CM
- 治療與處置碼：CPT-4、HCPCS、ICD-9/10-PCS
- 藥物：NDC





Data Source: Patrick Ryan. (2025). Understanding administrative claims and the Merative MarketScan databases. Summer School in OHDSI, AI, and Real-World Evidence.

下表簡單說明了在美國不同健康保險制度的特質：

Feature	Medicaid	Medicare	Private Insurance
Population Covered	Low-income, vulnerable groups	Seniors, disabled	Employer-based, individual
Funding Source	Federal + State	Federal	Employers, individuals
Administering Entities	States	Federal (CMS)	Private insurance plans

Data Source: Patrick Ryan. (2025). Understanding administrative claims and the Merative MarketScan databases. Summer School in OHDSI, AI, and Real-World Evidence.

- ●美國聯邦醫療補助保險(Medicaid)
  - ■是一項由聯邦政府與州政府共同資助的醫療計畫，為低收入個體、兒童、孕婦、老年人及身心障礙者提供醫療保險。
  - ■資金來源：聯邦政府會依據各州情況，以不同比例撥款來補助州政府的支出。此計畫由各州獨立管理。
  - ■主要特點：
    - 提供全面性福利，包括住院、門診、藥物等。
    - 各州可自行決定資格規定、承保範圍以及對醫療服務提供者的付款方式。
    - 快速統計：承保人數超過近 8,000 萬名美國民眾。
- ●美國聯邦醫療保險(Medicare)
  - ■ 是一項聯邦醫療保險計畫，主要為 65 歲及以上的個人、患有特定身心障礙的年輕人以下、末期腎臟病 (ESRD) 患者提供健康保險；
  - ■ Medicare 分為四個部分：
    - A 部份（醫院保險）：涵蓋住院、專業護理機構及部分居家照護/安寧療護服務。
      - 費用：若您已繳納 Medicare 稅費至少 10 年，通常是免費的。
    - B 部份（醫療保險）：涵蓋門診服務，如看診、預防保健、醫療用品和診斷檢查。
      - 費用：須支付月費，並有額外的自付額和共同保險（通常是 80/20 的費用分擔）。
    - C 部份（Medicare Advantage，優勢計畫）：經 Medicare 批准的私人保險計畫，用以取代傳統 Medicare（A 和 B 部份），通常包含 D 部份、牙科、視力或聽力服務。（這是標準 Medicare 的替代方案。）
    - D 部份（處方藥物承保）：獨立的處方藥物保險計畫。

- Medicare 不承保的項目：
  - 牙科、視力、助聽器、長期照護（例如：護理之家）。
  - 過高的自付費用，如自付額、共同支付額和共同保險。

本次課程所使用的美國醫療保險資料庫及醫學中心資料庫，主要是透過 Merative MarketScan Research Databases 所取得的研究資料庫。該研究資料庫提供去識別化的、長期追蹤、超過 2.73 億名獨立病患的個體層級健康保險申報和專科治療資料。MarketScan 系列資料庫包含三個核心健康保險申報資料庫，其一為醫院出院資料庫，以及數個連結之健康保險申報資料庫與其他病患和員工個人基本資料庫。
Merative MarketScan Research Databases 資料庫一覽表：

Database	Content	Covered Lives	Tables
Commercial Claims and Encounters (CCAE)	Health care coverage eligibility and service use of individuals in plans or product lines with fee-for-service plans and fully capitated or partially capitated plans	Active employees and dependents, early (non-Medicare) retirees and dependents, COBRA continuees	• Medical/Surgical ○ Inpatient Admissions (I) ○ Facility Header (F) ○ Inpatient Services (S) ○ Outpatient Services(O) • Prescription Drug (D) • Enrollment (A,T)
Medicare Supplemental and Coordination of Benefits (COB) (MDCR)	Health care coverage eligibility and service use of individuals in plans or product lines with fee-for-service plans and fully capitated or partially capitated plans	Medicare-eligible active and retired employees and their Medicare-eligible dependents from employer-sponsored supplemental plans	• Medical/Surgical ○ Inpatient Admissions (I) ○ Facility Header (F) ○ Inpatient Services (S) ○ Outpatient Services (O) • Prescription Drug (D) • Enrollment (A,T)

Data Source: Patrick Ryan. (2025). Understanding administrative claims and the Merative MarketScan databases. Summer School in OHDSI, AI, and Real-World Evidence.

電子健康紀錄資料庫(Electronic Health Records, EHR)

本次課程使用哥倫比亞大學醫學中心(Columbia University Irving Medical Center, CUIMC)的電子健康紀錄資料庫(Electronic Health Records, EHR)。電子健康紀錄(EHR)，是依據美國在 2009 年透過《健康資訊科技促進經濟與臨床健康法案》（Health Information Technology for Economic and Clinical Health, HITECH）開始強制要求醫療院所使用電子健康紀錄 (EHR)，EHR 中擷取的資料能反映潛在的臨床工作流程，而擷取這些資料的目的也各有不同。主要的目的有帳務：帳務處理、臨床照護：提供臨床醫療服務、法律：法律相關事務、營運：日常營運管理、法規：符合法規要求及研究：學術研究（例如：臨床試驗筆記）。
主要特點：

- 現有的編碼資料主要是因應法規要求。
- 包含細緻到如血氧飽和度的數據與高數量數據（例如：手術室內的生命徵象）。

目前 CUIMC EHR 資料庫包含約 610 萬名病人資料(女性約占 56%、年齡分布 0-19 歲、20-44 歲、45-64 歲、65-74 歲及 75 歲以上者，分別占 9%、26%、27%13%及 25%)。與生理量測有關資料約 17 億筆、診斷疾病與症狀有關資料約 1.4 億筆、藥品相關資料 7 千 1 百萬筆、手術處置相關資料約 4 千 9 百萬筆、與就醫有關資料約 2 千

1 百萬筆。

CUIMC EHR 屬於 CUIMC Clinical Information Ecosystem 中的一部分，整個 CUIMC Clinical Information Ecosystem 包含三個部分：

輔助系統 (Ancillaries)

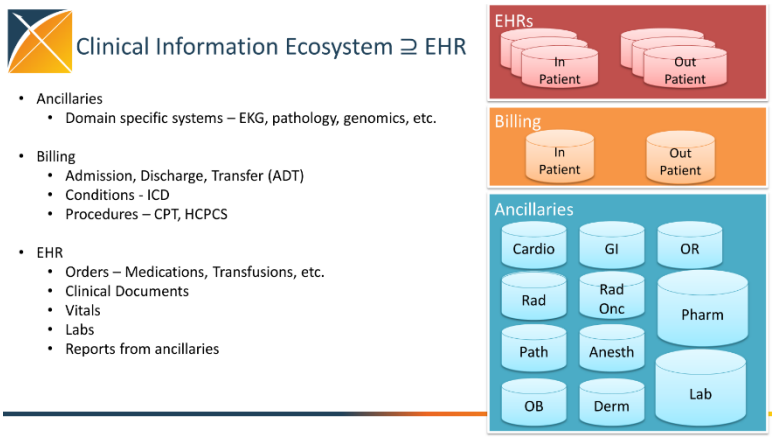
- 特定領域系統，如心電圖 (EKG)、病理學、基因組學等。

計價系統 (Billing)

- 入院、出院、轉院 (ADT)
- 病情診斷 - ICD
- 處置程序 - CPT、HCPCS

電子健康紀錄 (EHR)

- 醫囑 - 藥物、輸血等
- 臨床文件
- 生命徵象
- 實驗室
- 檢驗報告（來自輔助系統）



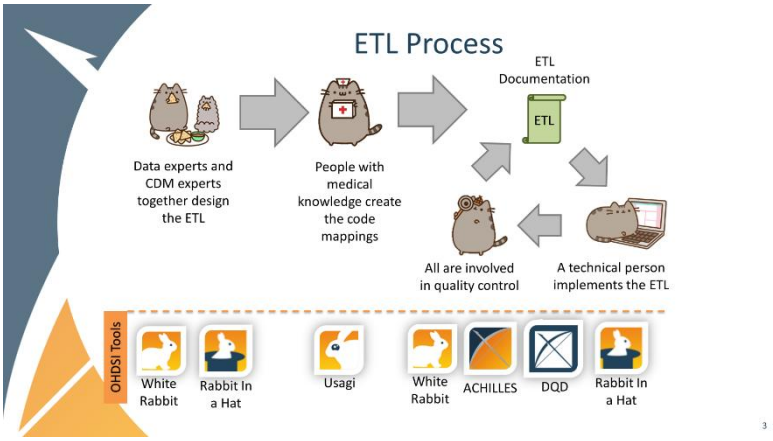
Data Source: Karthik Natarajan. (2025). Understanding Electronic Health Records (CUIMC). Summer School in OHDSI, AI, and Real-World Evidence.

(四)Extract, transform, load (ETL) to OMOP CDM and Evaluating Data Quality

為了將原始資料匯入轉成 OMOP 通用資料模型(CDM)，必須建立一個提取、轉換和載入(ETL)流程。此流程會將資料重構成 CDM 標準結構，並新增與標準化詞彙表的對應。ETL 流程通常以一組自動化腳本（例如 SQL 腳本）的形式來進行。ETL 流程的重要原則是應可重複性，以便在原始資料更新時重新執行。

建立一個 ETL（萃取、轉換、載入）通常是一個浩大的工程。OHSDI 多年來已發展出一套由四個主要步驟組成的最佳執行流程：

1. 資料專家和 CDM（通用資料模型）專家共同設計 ETL。
2. 具有醫學知識的人員建立代碼對應。
3. 技術人員實作 ETL。
4. 所有人都參與品質控制。



Data Source: Karthik Natarajan. (2025). Extract, transform, load (ETL) to OMOP CDM and Evaluating Data Quality. Summer School in OHDSI, AI, and Real-World Evidence.

為了更有效率將原始資料轉換至 OMOP CDM，OHDSI 計畫團隊發展了相應的工具來協助進行 ETL。

「White Rabbit」是一個軟體工具，用於協助將縱向醫療保健資料庫，經由 ETL（萃取、轉換、載入）的過程，轉入至 OMOP 通用資料模型（CDM）。透過 White Rabbit 工具掃描原始資料的資料表、欄位及對應欄位值，並產生一份包含所有必要資訊的報告，以提供設計 ETL 過程。

「Rabbit-In-a-Hat」是一款設計用於讀取和顯示經由「White Rabbit」掃描原始資料後所得到之報告文件的工具。White Rabbit 產生關於來源資料的資訊，而 Rabbit-In-a-Hat 則利用這些資訊，透過圖形使用者介面讓使用者能夠將來源資料連接到 CDM（通用資料模型）內的表格和欄位。Rabbit-In-a-Hat 會為 ETL（萃取、轉換、載入）過程產生掃描報告，但這個階段，尚不會產生用於建立 ETL 的程式碼。此處所產生的掃描報告，內容將包含將原始資料的資料表與 CDM 資料表建立連結，並將原始資料的資料表內容，提供做為 CDM 資料表的資訊；再依據所建立的每一個連結，進一步定義源列到 CDM 列的詳細資訊的連接。

在 Rabbit-In-a-Hat 中開啟 White Rabbit 掃描報告後，就可以開始設計和編寫將來源資料轉換為 OMOP CDM 的邏輯了。

「Usagi」是一款用於輔助手動建立代碼對應過程的工具。它可以根據代碼描述的文本相似性來提供對應建議。如果自動建議不正確，Usagi 允許使用者搜尋適當的目標概念。

ETL 的 Quality Control

針對 ETL 過程，品質控制是反覆進行的。典型的模式是：編寫邏輯 -> 實作邏輯 -> 測試邏輯 -> 修復/編寫邏輯。測試 CDM 的步驟：

- 審查 ETL 設計文件、電腦程式碼和代碼對應：
  任何人都有可能犯錯，因此應至少有一位其他人審查已完成的工作。電腦程式碼中最大的問題往往來自於如何將原始資料中的來源代碼對應到標準概念。請務必仔細檢查所有進行對應的區域，以確保正確的來源詞彙表被轉換成適當的概念性 ID。
- 手動比對來源資料和目標資料中一小部分人員的所有資訊：
  追蹤單一個人的資料，理想情況下是選擇一位擁有大量獨特記錄的人，追蹤單一個人的資料，如果 CDM 中的資料樣貌與根據既定邏輯所預期的不同，則可以榮容易找到問題。
- 比較來源資料和目標資料的總資料筆數：
  根據您選擇如何處理某些問題，資料筆數上可能會存在一些預期的差異。例如，有些合作夥伴選擇刪除性別為 NULL 的人員，因為這些人無論如何都不會被納入分析中。此外，CDM 中的就診記錄可能與原始資料中的就診或接觸記錄的建構方式不同。因此，在比較來源資料和 CDM 資料的總資料筆數時，務必考慮並

預期這些差異。

- ●在 CDM 版本上，試著重現一個已經在來源資料上進行過的研究：
  這是了解來源資料和 CDM 版本之間任何重大差異的好方法。
- ●創建單元測試，用來重現來源資料中應在 ETL 中處理的一種模式：
  例如，若 ETL 規定應刪除沒有性別資訊的患者，則建立一個沒有性別的人的單元測試，並評估 ETL 邏輯如何處理它。
  - ■ 單元測試在評估 ETL 轉換的品質和準確性時非常方便：創建一個較小的資料集，該資料集模仿您正在轉換的來源資料的結構。資料集中的每個人或每條記錄都應測試 ETL 文件中寫明的一個特定邏輯片段。使用這種方法，可以輕鬆追溯問題並識別出失敗的邏輯。小巧的規模也讓電腦程式碼能夠快速執行，從而加快迭代和錯誤識別的速度。

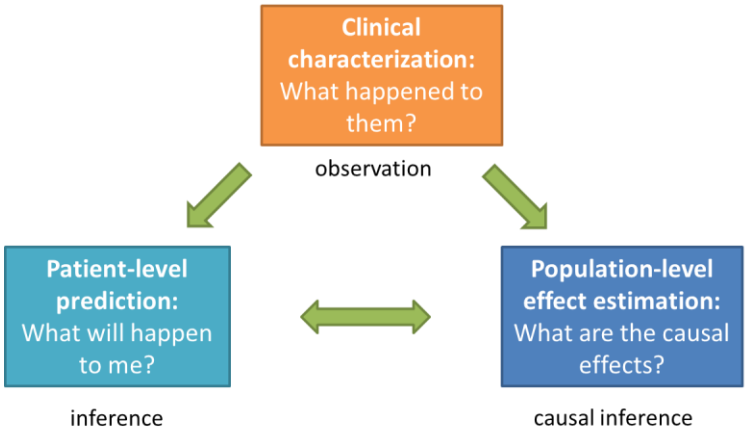
ETL Conventions 與 THEMIS

CDM 的目標是將醫療保健資料標準化，但如果每個團體都以不同的方式處理特定的資料情境，將會使跨網路系統化使用資料變得更加困難。
OHDSI 社群開始記錄慣例，以提高 CDM 之間的一致性。這些經 OHDSI 社群同意並定義的慣例，可作為設計 ETL 時參考。例如，容許病人缺少出生月份或日期，但如果缺少出生年份，則應該將該人員捨棄。在設計 ETL 時，參考這些慣例，將可幫助做出與社群保持一致的特定設計決策。

THEMIS 由社群中的個人組成，他們負責收集慣例、澄清慣例、與社群分享以徵求意見，然後在 CDM Wiki 中記錄最終確定的慣例。

(五)應用觀察性健康資料設計 3 項觀察性研究介紹

觀察性健康資料，透過觀察性研究設計，以產生對於病人治療旅程中重要的實證。針對所希望回答的命題，主要有 3 種研究設計類型：臨床特徵描述研究(Characterization)、群體(因果)效應評估研究(Estimation)及病人為基礎之預測研究(Prediction)。



Data Source: Anna Ostropolets. (2025). Designing an observational network study: Characterization. Summer School in OHDSI, AI, and Real-World Evidence.

根據 3 種研究設計類型，可以進一步區分 8 種次分類研究

Analytic use case	Type	Structure	Example
Clinical characterization	Disease Natural History	Amongst patients who are diagnosed with <insert your favorite disease>, what are the patient's characteristics from their medical history?	Amongst patients with rheumatoid arthritis, what are their demographics (age, gender), prior conditions, medications, and health service utilization behaviors?
	Treatment utilization	Amongst patients who have <insert your favorite disease>, which treatments were patients exposed to amongst <list of treatments for disease> and in which sequence?	Amongst patients with depression, which treatments were patients exposed to SSRI, SNRI, TCA, bupropion, esketamine and in which sequence?
	Outcome incidence	Amongst patients who are new users of <insert your favorite drug>, how many patients experienced <insert your favorite known adverse event from the drug profile> within <time horizon following exposure start>?	Amongst patients who are new users of methylphenidate, how many patients experienced psychosis within 1 year of initiating treatment?
Population-level effect estimation	Safety surveillance	Does exposure to <insert your favorite drug> increase the risk of experiencing <insert an adverse event> within <time horizon following exposure start>?	Does exposure to ACE inhibitor increase the risk of experiencing Angioedema within 1 month after exposure start?
	Comparative effectiveness	Does exposure to <insert your favorite drug> have a different risk of experiencing <insert any outcome (safety or benefit) > within <time horizon following exposure start>, relative to <insert your comparator treatment>?	Does exposure to ACE inhibitor have a different risk of experiencing acute myocardial infarction while on treatment, relative to thiazide diuretic?
Patient level prediction	Disease onset and progression	For a given patient who is diagnosed with <insert your favorite disease>, what is the probability that they will go on to have <another disease or related complication> within <time horizon from diagnosis>?	For a given patient who is newly diagnosed with atrial fibrillation, what is the probability that they will go onto to have ischemic stroke in next 3 years?
	Treatment response	For a given patient who is a new user of <insert your favorite chronically-used drug>, what is the probability that they will <insert desired effect> in <time window>?	For a given patient with T2DM who start on metformin , what is the probability that they will maintain HbA1C<6.5% after 3 years?
	Treatment safety	For a given patient who is a new user of <insert your favorite drug>, what is the probability that they will experience <insert adverse event > within <time horizon following exposure>?	For a given patients who is a new user of warfarin, what is the probability that they will have GI bleed in 1 year?

School in OHDSI, AI, and Real-World Evidence.

此外，可以進一步設計研究題目，來了解醫療介入的效果。

Analytic use case	Type	Questions
Clinical characterization	Disease Natural History	• Amongst patients who are diagnosed with Hypertension, what are the patient's characteristics from their medical history? • Amongst patients who are new users of ACE inhibitors with prior diagnosis with Hypertension, what are the patient's characteristics from their medical history?
	Treatment utilization	• Amongst patients who are diagnosed with Hypertension, which treatments were patients exposed to amongst antihypertensive drugs (ACE, ARB, CCB, TZD, BB) and in which sequence?
	Outcome incidence	• Amongst patients who are new users of ACE inhibitors with prior diagnosis with Hypertension, how many patients experienced Acute myocardial infarction within on-treatment period from drug exposure start + 1 day to drug exposure end? • Amongst patients who are new users of ACE inhibitors with prior diagnosis with Hypertension, how many patients experienced Angioedema within on-treatment period from drug exposure start + 1 day to drug exposure end?
Population-level effect estimation	Safety surveillance	• Does exposure to ACE inhibitors increase the risk of experiencing Angioedema within on-treatment period from drug exposure start + 1 day to drug exposure end?
	Comparative effectiveness	• Does exposure to ACE inhibitors have a different risk of experiencing Acute myocardial infarction within on-treatment period from drug exposure start + 1 day to drug exposure end, relative to Angiotensin Receptor Blockers (ARBs)?
Patient level prediction	Disease onset and progression	• For a given patient who is diagnosed with who are diagnosed with Hypertension, what is the probability that they will go on to have Acute myocardial infarction within 1 year following treatment initiation ?
	Treatment response	• For a given patient who is new users of ACE inhibitors with prior diagnosis with Hypertension, what is the probability that they will have Acute myocardial infarction in 1 year following treatment initiation?
	Treatment safety	• For a given patient who is new users of ACE inhibitors with prior diagnosis with Hypertension, what is the probability that they will have Angioedema in 1 year following treatment initiation?

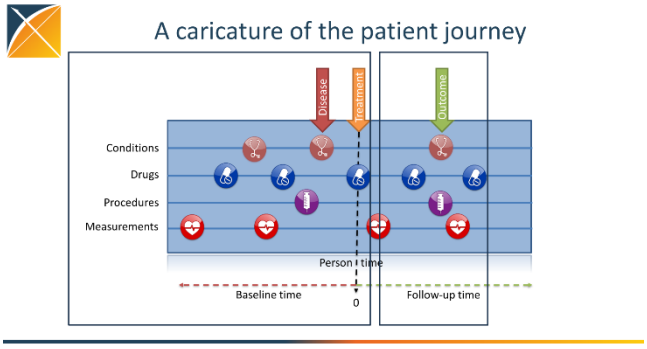
Data Source: Anna Ostropolets. (2025). Designing an observational network study: Characterization. Summer School in OHDSI, AI, and Real-World Evidence.

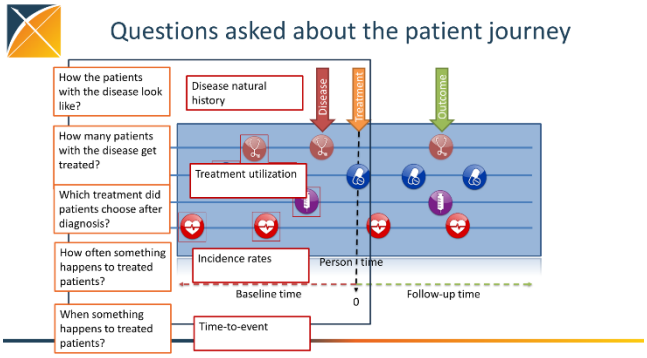
依據前述兩張表格所列出之範例，研究者可以很有結構且明確的提出觀察性研究想要回答的問題為何？並很清楚掌握研究設計中重要關鍵變項。

I. 設計一個基於觀察性研究網路之臨床特徵描述研究(Characterization)

底下兩張圖示是一個虛擬的病人治療旅程，左圖顯示的是該病人就醫的情況與時

間點，列出病人所有就醫過程的 profile；右圖顯示的是關於該病人所有就醫過程，可能需要了解的問題。





可以看出有 4 類型的 5 個研究問題被提出：Disease natural history、Treatment utilization、Incidence rates 及 Time-to-event。底下就根據這 5 個研究問題，需要提問及需要定義的研究目標族群與特徵，進行說明。

●Disease natural history

Type	Structure	Example
Disease Natural History	Amongst patients who are diagnosed with <insert your favorite disease>, what are the patient's characteristics from their medical history?	Amongst patients with myocardial infarction, what are their demographics (age, gender), prior conditions, medications, and health service utilization behaviours?

在 OHDSI 的系統設定中，在 **FeatureExtraction** 中有一組預設的 **Features** 特徵集：

Component	Description
Target population (T):	Who do you want to do the characterization for?
Features:	What are the characteristics you want to look at?
Time windows:	What are the time windows you want to look at?

人口統計資料：性別、年齡組別、種族、族裔、索引年份、索引月份；

過往病症組別 / 藥物組別 / 處置程序 / 醫療器材 / 測量 / 觀察：短期（30 天）和長期（365 天）資料；

共病症風險評分：Charlson Score、DCSI、CHADS2VASC

●Treatment utilization

Type	Structure	Example
Treatment utilization	Amongst patients who have <insert your favorite disease>, which treatments were patients exposed to amongst <list of treatments for disease> and in which sequence?	Amongst patients with myocardial infarction, which treatments were patients exposed to amongst diuretics, beta-blockers, ACE inhibitors, other antihypertensive drugs and in which sequence?

Component	Description
Target population (T):	Who do you want to do the treatment utilization for?
Outcome (O):	What are the treatments you want to look at?
Stratification	Are there subgroups you want to study (age, gender, year)?

●Incidence rates

Type	Structure	Example
Outcome incidence	Amongst patients who are new users of <insert your favorite drug>, how many patients experienced <insert your favorite known adverse event from the drug profile> within <time horizon following exposure start>?	Amongst patients who are new users of ACE inhibitors, how many patients experienced myocardial infarction within 1 year of initiating treatment?

Component	Description
Target population (T):	Who do you want to do the IRs for?
Outcome (O):	What are you estimating the rate of?
Time-at-risk (TAR):	When are you estimating?
Statas	Are there subgroups you want to study (age, gender, year)?

Proportion: (# people with outcome during TAR)/(# people)

Rate: (#outcomes during TAR)/(total person days)

●Time-to-event

Type	Structure	Example
Outcome incidence	Amongst patients who are new users of <insert your favorite drug>, how many patients experienced <insert your favorite known adverse event from the drug profile> within <time horizon following exposure start>?	Amongst patients who are new users of ACE inhibitors, how many patients experienced myocardial infarction within 1 year of initiating treatment? When do users of ACE inhibitors experience myocardial infarction?

Component	Description
Target population (T):	What is the population you want to look at?
Outcome (O):	What are the events you want to look at?

●Risk factors

Type	Structure	Example
Outcome incidence	Amongst patients who are new users of <insert your favorite drug>, how many patients experienced <insert your favorite known adverse event from the drug profile> within <time horizon following exposure start>?	Amongst patients who are new users of ACE inhibitors, how many patients experienced myocardial infarction within 1 year of initiating treatment? When do users of ACE inhibitors experience myocardial infarction? What are the differences between users of ACE inhibitors who go on to have myocardial infarction versus those who do not?

Describe patients with and without the outcome during time-at-risk。

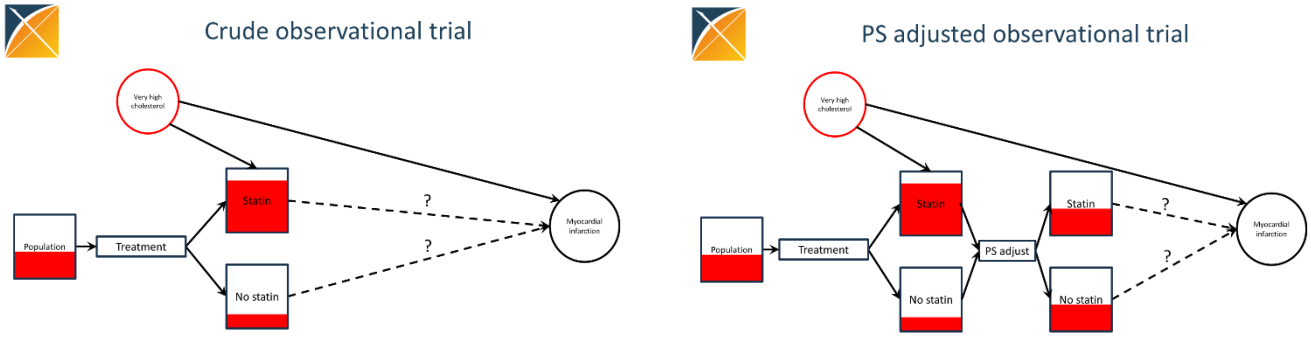
Component	Description
Target population (T):	Patients with an outcome
Comparator population (C)	Patients without an outcome
Time-at-risk (TAR):	When are you looking for the outcome?

II. 設計一個基於觀察性研究網路之比較世代設計群體(因果)效應評估研究 (Estimation)

- Confounders 校正

有別於之前所進行之臨床特徵描述研究(Characterization)，在進行因果效應評估研究時，為避免得到偏誤推論，需要進一步進行 Confounders 校正。因此，這部分研究設計會著重於對於 confounding 的處理。其中在觀察性研究中，最常使用的便是傾向分數(Propensity score)法。Propensity score 是指在給定基線共變數的情況下，病患屬於目標群組相對於對照群組的機率。可將 Propensity score 作為「平衡分數 (balancing score)」，並因其具有使共變數和是否治療變項彼此獨立性質，所以各組的共變數分佈(covariate distribution)應相似。因此，透過平衡傾向分數 (propensity) 來達到平衡共變數 (covariates)；進而透過平衡共變數來得到結果的可比較性，做出因果斷言 (causal assertion)推論，即結果必定是由於治療所致。

底下圖例說明傾向分數應用於藥物效果評估研究之機制：



Data Source: George Hripcsak. (2025). Designing an observational network study: Estimation using Comparative Cohort design. Summer School in OHDSI, AI, and Real-World Evidence.

使用傾向分數法來進行 Confounders 校正，需要注意測量的基線共變數在治療組和未治療組之間的傾向分數分布是否相似？

在藥物流行病學中，使用傾向分數的五個原因：

1.理論上的優勢：適應症混淆(Confounding by indication)是影響研究效度的主要威脅。傾向分數直接針對所研究藥物的使用與非使用適應症，進行有效處理。

2.傾向分數在配對或篩選族群方面的價值：傾向分數可以消除「不可比較」的對照組，而無需假設傾向分數與結果之間存在線性關係。

3.在結果事件較少時改善估計：傾向分數允許研究者僅根據一個純量值進行配對，而非需要為所有共變數配置自由度。這在結果事件個案數較少時特別有用，能提升統計效率。
4.傾向分數與治療的交互作用：透過傾向分數，研究者可以探索不同病患層級對治療反應的異質性。
5.傾向分數校正以修正測量誤差：傾向分數校正可以幫助修正測量誤差，進一步提高研究結果的準確性。

除了前述提到之以傾向分數法校正 confounders 外，另一個進行觀察性因果推論研究必須謹慎考量的是，如何選擇需要納入的 confounders。常見的是透過文獻搜尋方式列出須納入的 confounders，另一種則是 Empirical selection。

Empirical selection of confounders 是透過 Large-scale propensity score (LSPS)方法來達成，LSPS 是一個系統化的方法來進行傾向分數調整；
1. 使用大量的共變數：通常介於 10,000 到 100,000 個之間。
2. 篩選共變數：

雖然使用了大量的共變數，但並非要平衡所有變項。需要特別注意以下幾類變數，不可納入傾向分數法校正：

- 中介變數 (Mediators)：位於治療和結果路徑之間，應被排除。
- 簡單碰撞變數 (Simple colliders)：如果同時是治療和結果的共同結果，應被排除。
- 工具變數 (Instruments)：根據診斷或領域知識判斷，應被排除。
- M 型偏誤 (M-bias)：與潛在病因相關的變數也需注意排除。

3. 建立傾向分數模型：由於變數數量（#variables）多於個案數量（#cases），建議使用 LASSO (正規化迴歸) 來建立傾向分數模型。
4. 根據傾向分數進行配對或分層：完成模型建立後，根據計算出的傾向分數進行配對或分層。
5. 診斷性檢查：最後，進行診斷性檢查，確認所有觀察到的變數都已達到平衡，以確保分析結果可靠性。

LSPS 方法的優勢：

1.傳統方法 vs. 傾向分數：通常在傳統的研究中，我們會試圖挑選出干擾因子(confounders)來進行調整；但在此處，我們的目標是納入所有可用的治療前變數(pre-treatment variables)，這與一些試圖篩選干擾因子的方法（例如 HDPS，高維度傾向分數）不同。

2.變數多於個案：由於變數數量通常多於個案數量，所以使用 LASSO (最小絕對收縮與選擇運算子)這種正規化迴歸方法。不只是能透過挑選出有資訊量的變數來進行維度縮減，更重要的是，它試圖將所有變數的資訊都納入模型中。

3.可調整無法直接測量的 confounders：當我們同時納入許多變數進行調整時，即使某些變數沒有被直接測量，也可能因為其他相關變數的存在而被間接調整到。

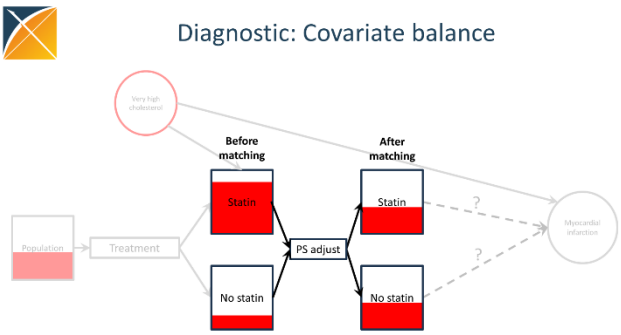
使用傾向分數校正 confounding 的 4 種形式：

- Regression adjustment：將傾向分數作為其中一個共變項，納入迴歸模式調整。
- Matching：將治療組和未治療組的受試者其傾向分數相似者形成匹配的樣本組，並直接比較匹配樣本中治療組和未治療組之間的結果來估計治療效果。
- Stratification：將受試者根據其估計的傾向分數劃分為互斥的子集。在每個傾向分數分層內，治療組和未治療組的受試者將具有大致相似的傾向分數，因此觀察到的基線共變數分布也大致相似，再比較各分層內治療組和未治療組的治療效果。
- Inverse Probability Weighting：使用基於傾向分數的權重來建立一組合成樣本，其中測量的基線共變數的分布與治療分配無關。受試者的權重等於其實際接受治療機率的倒數，進行加權調整。

目前在 OHDSI 系統中，針對 Matching 與 Stratification，有發展對應之應用程式 CohortMethod R package。

共變數平衡診斷(Covariate Balance Diagnostics)

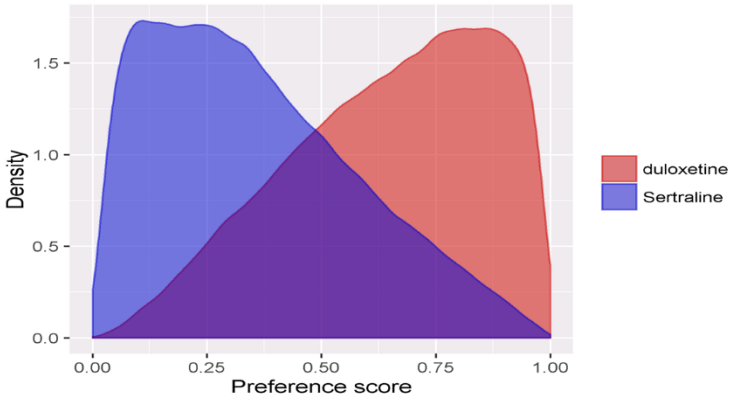
透過**傾向分數校正後，是否達成共變數平衡，可以透過幾個方法來診斷。**



1、Standardized difference of mean：標準化差異(standardized difference)來量化組間差異。標準化差異小於 0.1 通常被認為是共變數均值或流行率之間可忽略不計的差異。

2、Distribution equipoise：測量的基線共變數在治療組和未治療組之間傾向分數分布是否相似。

如果重疊程度太小，代表兩個群組的病患特徵差異太大，這會導致，影響研究結果的外推性、推計推論**不穩定性。**



Data Source: George Hripcsak. (2025). Designing an observational network study: Estimation using Comparative Cohort design. Summer School in OHDSI, AI, and Real-World Evidence.

其他重要不可納入的影響變數

Instruments(工具變項)：在統計學和流行病學中用於處理因果推論中的干擾因子 (confounders) 問題。它是一種特殊的變數，滿足以下三個條件：

1. 與治療變數相關 (Associated with the treatment)：工具變數必須與正在研究的治療(treatment)或暴露(exposure)相關。換句話說，它能預測一個人是否會接受某種治療。
2. 與結果變數無關(Unassociated with the outcome)：工具變數必須與結果(outcome)無關，除非是透過治療變數的影響。
3. 獨立於未測量的干擾因子(Independent of unmeasured confounders)：工具變數必須獨立於所有未測量且同時影響治療與結果的干擾因子。這一項是最難滿足的條件。

為什麼需要工具變數？

在觀察性研究中，很難確保治療組和非治療組在所有方面都完全相似。未測量的干擾因子（例如病患的健康習慣、基因等）可能會同時影響他們是否接受治療以及他們的結果，導致我們錯誤地歸因因果關係。透過工具變數，處理存在未測量的干擾因子的情況下，幫助研究者推斷出更準確的因果關係。然而，要找到一個完美的工具變數是非常困難的。

為什麼工具變數對傾向分數模型有害？

傾向分數模型的目標是平衡治療組與非治療組之間干擾變數 (confounding variables) 的分佈。所謂的干擾變數，是同時與治療和結果都有關的變數。

然而，根據定義，工具變數 (instrument) 並不是干擾變數。當你在傾向分數模型中納入一個工具變數時，你其實是在「控制」治療分配中那些與結果無關的變異。這種做法會扭曲其餘變數與結果之間的關係，可能導致偏差放大 (bias amplification)。

在傾向分數分析中，我們的目的是透過平衡所有潛在干擾變數來確保治療組和非治療組是可比較的。如果將工具變數（它並非干擾變數）納入模型，反而會破壞這種平衡，使得估計結果比完全不納入這個變數時更差。

通常我們可以透過 Distribution equipoise 來檢查是否存在 strong instruments，來避免因工具變項的存在，造成偏差放大 (bias amplification)。

中介變項(Mediators)：中介變數是因果關係中的一個關鍵環節，它解釋了自變數與應變數之間的關係。並透過中介變項說明了為什麼或如何產生了某種效應。

因果路徑：可以將中介變數想像成一連串事件中的「中間人」。自變數（原因）並非直接影響應變數（結果），而是先影響中介變數，再由中介變數去影響應變數。自變數 → 中介變數 → 應變數

中介變數與干擾變數不同，中介變數是因果關係的一部分。它是介於原因和結果之間的一個必要步驟。干擾變數則是一個獨立的變數，它同時影響自變數和應變數，從而造成一種虛假或誤導性的關聯性，非屬於因果路徑中的影響變數。關鍵區別在於因果

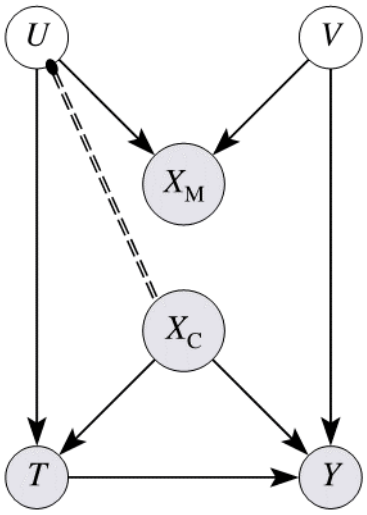
方向：中介變數是由自變數所引起，而干擾變數則是自變數和應變數的共同原因。

簡單對撞變數(Simple Colliders)：對撞變數(Collider)是指一個變數，同時是兩個或多個其他變數的共同結果。想像一下兩條因果路徑像火車軌道一樣，從不同方向駛來，在某一個節點「對撞」在一起，這個對撞點就是對撞變數。
「簡單對撞變數」指的是最基本的對撞變數類型。它通常滿足以下因果路徑：
變數 A → 對撞變數 C ← 變數 B
在這裡，變數 A 和變數 B 都是導致對撞變數 C 的原因。
為什麼要注意碰撞變數？
在統計分析中，控制(controlling for)或調整(adjusting for)一個對撞變數會創造出一種虛假的關聯。這種現象稱為「對撞偏差 (collider bias)」。

- 對撞變數是兩個或多個變數的共同結果。
- 在一般的因果分析中，不應該調整碰撞變數。
- 調整或控制一個對撞變數會產生偏差，創造出原本不存在的虛假關聯。

簡單來說，在進行統計分析時，只要記得一個原則：避免將對撞變數納入模型中，除非你有非常特定的研究目的，否則可能會得出錯誤的結論。

M 型偏差(M-bias)：是一種特殊的 confounders，它因其在因果路徑(causal diagram)上呈現類似英文字母「M」的形狀而得名。
當調整(或控制)一個變數，而這個變數是兩個未測量干擾因子(unmeasured confounders)的共同結果，就會產生 M 型偏差。其因果路徑圖如右：
在這個圖中，C 就是那個中間變數。如果我們將 C 納入統計模型中進行調整，就會產生 M 型偏差。



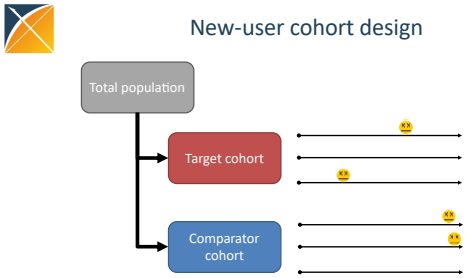
M 型偏差的運作原理

- U 和 V 是無法測量的兩個干擾因子。
- U 影響治療 T 和中間變數 X_C。
- V 影響結果 Y 和中間變數 X_C。
- C 是 U 和 V 的共同結果。

雖然 C 表面上看起來與治療 T 和結果 Y 沒有直接關聯，但當在模型中控制 X_C 時，你會在 U 和 V 之間創造出一條虛假的關聯路徑。這條新的路徑，會進一步將 U 和 V 的干擾效應傳遞給治療 T 和結果 Y，最終導致對治療效應的估計產生偏差。
簡單來說，M 型偏差的產生是因為錯誤地在模型中包含了不應被調整的變數。這個變數(X_C)本身不是干擾因子，但它被兩個未測量的干擾因子(U 和 V)所影響。一旦控制了 X_C，就會「打開」一條新的、原本不存在的偏差路徑。
如何避免 M 型偏差？
要避免 M 型偏差，最直接的方法就是不要在模型中調整這個中間變數 X_C。

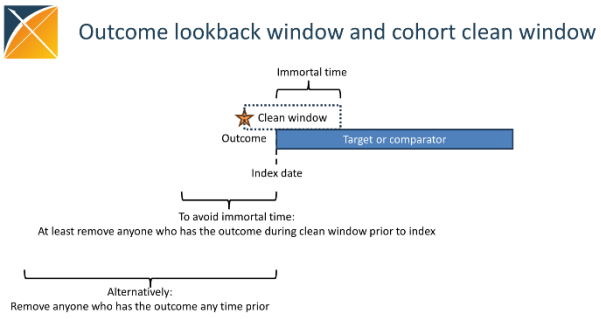
在因果推論中，識別出正確的因果路徑圖至關重要。這包括了解哪些變數是治療和結果之間的真正干擾因子，哪些是中間變數或碰撞變數。只有在正確地識別並處理這些變數後，才能得出更可靠的因果結論。

- **比較性世代設計(Comparative Cohort Design)**
  - 定義兩個需要比較的世代，通常是兩種介入(或治療)的世代，如使用 ACEi vs ARB 的兩群病人。
  - 使用新使用者世代追蹤設計(new-user cohort design)，之後再以傾向分數方法，使兩個世代具有可比較性。
  - 比較兩個世代的健康結果，如 AMI。
  - 回歸分析來產生影響效果估計。



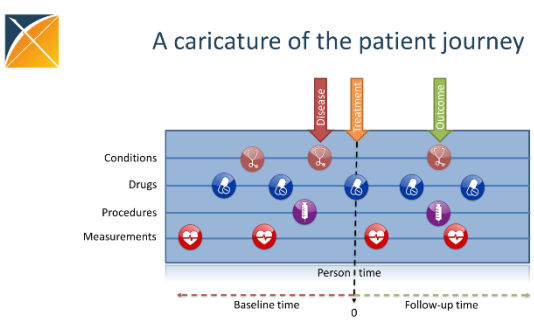
另一個觀察性追蹤世代研究需要注意的問題 Immortal time

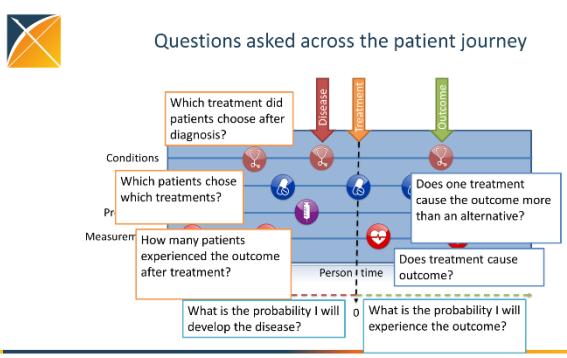
世代追蹤開始時，會有部分個案，出現 clean window，也就是在那段時間中，不會出現任何 outcome，此即所謂的 immortal time bias。研究者錯誤地將一段**病患不可能死亡**的時間納入了治療組的觀察期中。這段「**不可能死亡**」的時間通常是從病患被診斷或被納入研究開始，直到他們真正開始接受某項治療為止。因此，為避免造成這樣偏差，研究者需要調整研究設計，確保治療組和對照組的追蹤期從**相同的時間點(index date)**開始，或者設定排除條件，排除出現 immortal time bias 的個案。



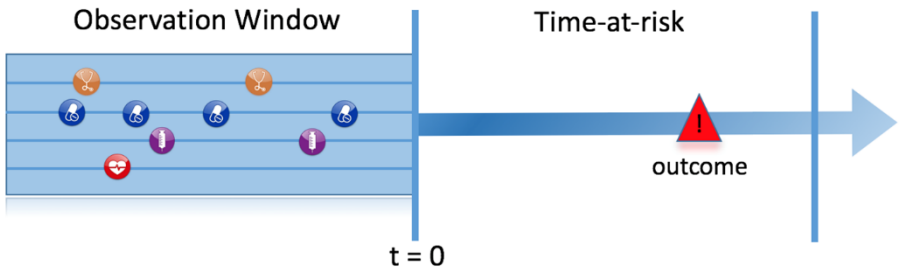
III. 設計一個基於觀察性研究網路之病人為基礎之預測研究(Prediction)

底下兩張圖示是一個虛擬的病人治療旅程，左圖顯示的是該病人就醫的情況與時間點，列出病人所有就醫過程的 profile；右圖顯示的是關於該病人所有就醫過程，對於個別病人可能需要了解的問題。該病人最終發展為疾病之機率？有病者最終發生 outcome 的機率？





個別病人之預測性研究，是指在一個目標人群（T）中，我們希望預測哪些病人在一個特定時間點（t=0）後，會在「風險期間(Time-at-risk)」內發生某種結果（O）。所有預測將僅使用該時間點之前「觀察視窗」內的病人資訊。

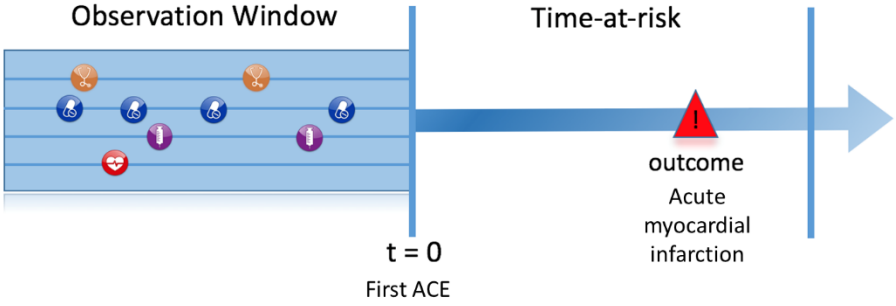


健康照護研究中，關於預測問題的類型

Type	Structure	Example
Disease onset and progression	Amongst patients who are newly diagnosed with <insert your favorite disease>, which patients will go on to have <another disease or related complication> within <time horizon from diagnosis>?	Among newly diagnosed AFib patients, which will go onto to have ischemic stroke in next 3 years?
Treatment choice	Amongst patients with <indicated disease> who are treated with either <treatment 1> or <treatment 2>, which patients were treated with <treatment 1> (on day 0)?	Among Hypertension patients who took either ACE inhibitor or ARB, which patients got ACE? (as defined for propensity score model)
Treatment response	Amongst patients who are new users of <insert your favorite chronically-used drug>, which patients will <insert desired effect> in <time window>?	Among new users of lisinopril, which patients will have acute myocardial infarction in 1 year?
Treatment safety	Amongst patients who are new users of <insert your favorite drug>, which patients will experience <insert your favorite known adverse event from the drug profile> within <time horizon following exposure start>?	Among new users of lisinopril, which patients will have angioedema in 1 year?
Treatment adherence	Amongst patients who are new users of <insert your favorite chronically-used drug>, which patients will achieve <adherence metric threshold> at <time horizon>?	Which patients with T2DM who start on metformin achieve >=80% proportion of days covered at 1 year?

Component	Description	Example
Target population (T):	Who do you want to do the prediction for?	New users of ACE inhibitors
Outcome (O):	What are you predicting?	Acute myocardial infarction
Time-at-risk (TAR):	When are you predicting?	1 day to 365 days after ACE inhibitor initiation

針對 ACE inhibitors 降血壓藥物的新使用者**目標族群（T）**，我們的目標是預測哪些病患在首次使用 ACE 抑制劑（t=0），**會在索引日期後 1 到 365 天的期間內發生**急性心肌梗塞（O）。所有預測變項僅使用該時間點前「觀察視窗」內的病患資訊。



而在使用觀察性健康照護資料進行預測研究中，可以參考底下的架構，進行訓練資料集與驗證資料集的配置。

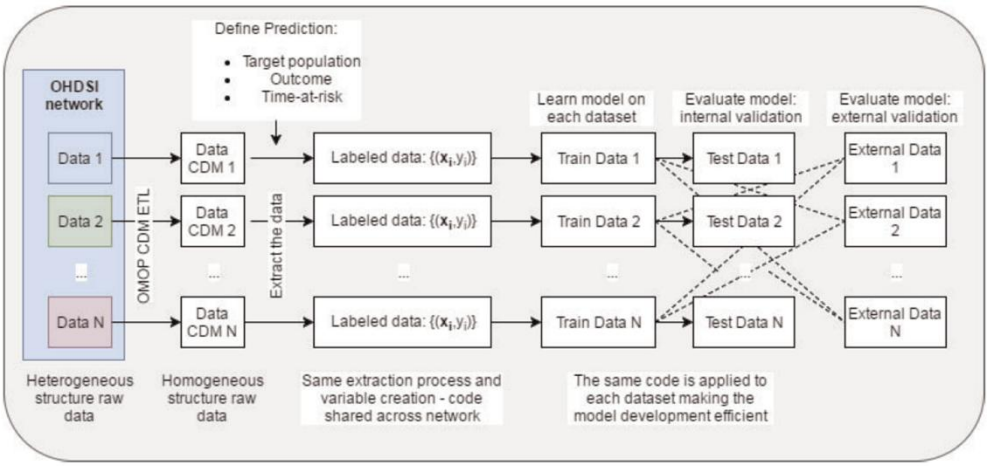
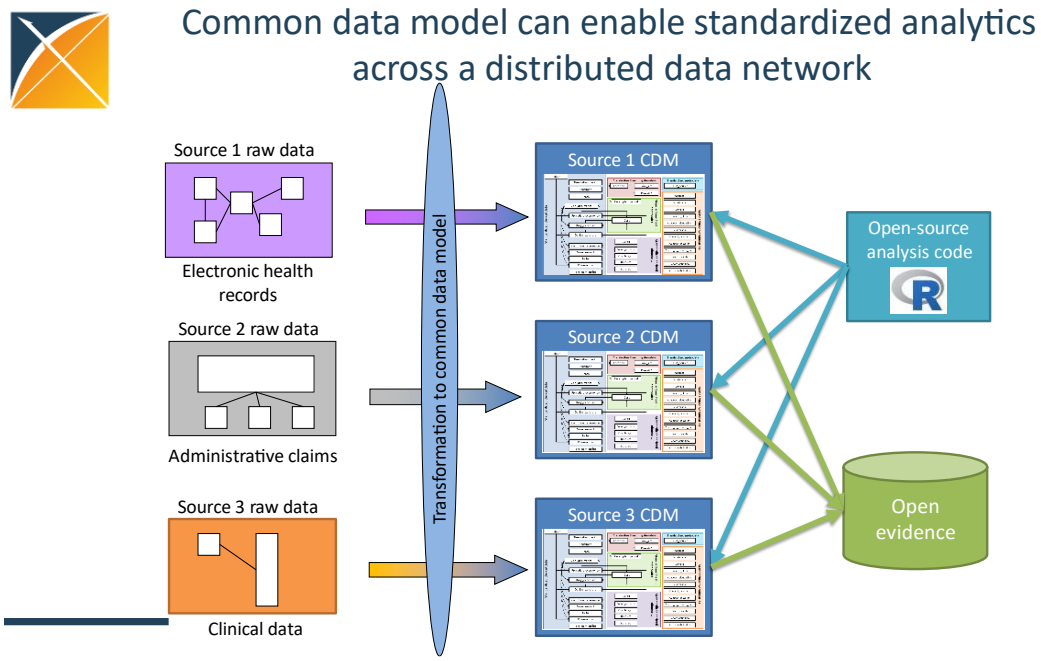


Figure 1. Illustration of how the homogeneous structure of the OMOP common data model enables sharing of model development code.

Data Source: Reps JM et al. (2018). Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *JAMIA*, *25*(8), 969–975.

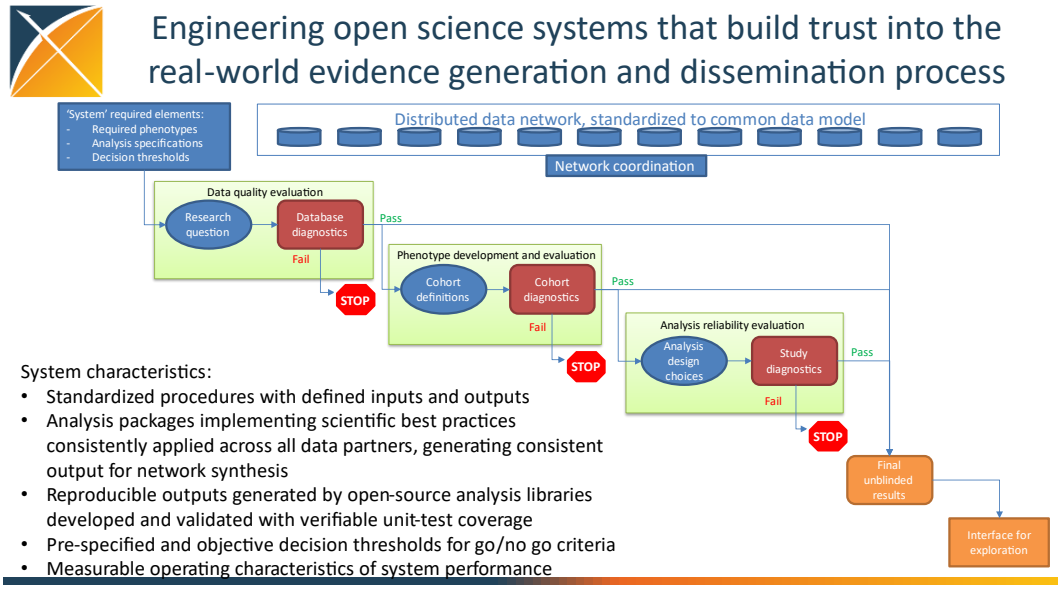
(六)觀察性研究的 phenotype 發展與評估

OHDSI 透過共通性資料模式(CDM)將不同的健康資料進行標準化相關名詞與語彙後，即可進行標準化的分析。如右圖：



而在執行一個觀察性研究中產生實證的過程中，會有幾個重要的步驟，包含使用 CDM 標準化後的資料，清楚定義納入研究的世代、依研究假設選擇標準化分析工具進行分析，然後產生具有影響力的實證。在 OHDSI 開放科學的系統架構下，有幾個重要特質：

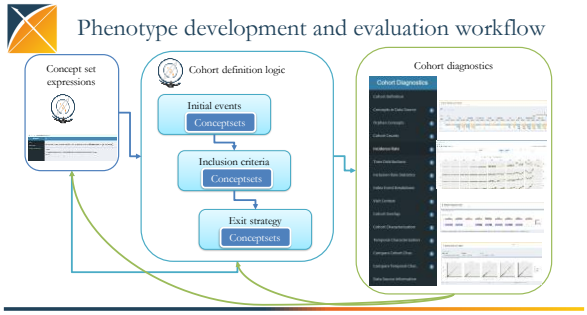
- **標準化流程**，具有明確的輸入和輸出定義。
- **資料分析套件**採用在所有資料合作夥伴間一致應用的科學最佳實踐，為不同資料來源的研究網絡綜合生成一致的輸出結果。
- 透過使用**開源資料分析套件庫**產生可重現的結果，且該套件庫經過開發和驗證，具有可驗證的單元測試覆蓋率。
- 針對取捨標準，設置了**預先指定和客觀的決策閾值**。
- **可測量的系統性能操作特性**。

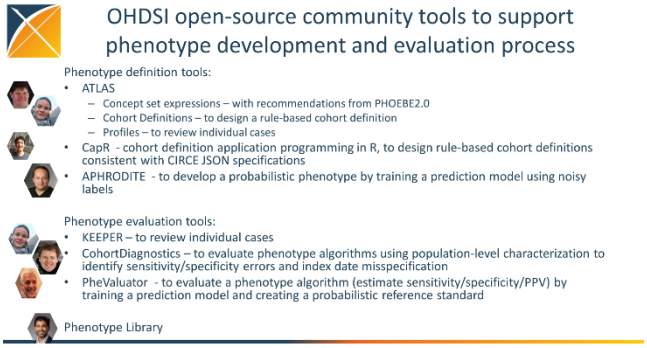


System characteristics:

- Standardized procedures with defined inputs and outputs
- Analysis packages implementing scientific best practices consistently applied across all data partners, generating consistent output for network synthesis
- Reproducible outputs generated by open-source analysis libraries developed and validated with verifiable unit-test coverage
- Pre-specified and objective decision thresholds for go/no go criteria
- Measurable operating characteristics of system performance

因此，OHDSI 將這一系列標準化後的系統，依研究主題與架設，設定規格後所發展形成的研究計畫規格和分析套件，稱之為「phenotype」。並藉由 OHDSI 所發展的輔助工具來指定研究的 phenotype 並進行評估。OHDSI 所發展的輔助工具，包含 ATLAS、CapR、KEEPER、CohortDiagnostics。





在 OHDSI 系統中，一個研究的 phenotype 的評估，其目標是評估透過該 phenotype 所推斷的結果，與病患真實健康狀況的一致程度，並透過敏感度(Sensitivity)、特異度 (Specificity)、陽性預測值 (Positive Predictive Value)及陰性預測值 (Negative Predictive Value)等**指標測量測量誤差；並判斷**研究的 phenotype 是否適用(判斷的原則是依該研究的 phenotype 所設定的資料庫中，研究的測量誤差足夠小，以至於不會對分析結果的解讀產生負面影響)。

完成評估的研究的 phenotype，可以進入下一步直接進行分析，產製分析結果的階段，並進一步提供研究夥伴，進行不同資料來源，但一致性的分析。

(七)使用 OHDSI 架構導入觀察性網絡研究

在一個導入 OHDSI 系統，並完成原始健康資料庫透過 CDM 標準化的組織中，可以很快速應用其他單位所產生的研究 phenotype，即可進行相同的研究分析。同時因為使用相關的研究規格，兩個組織間的研究結果還具有高度可比較性。因此，可以很容易形成所謂的觀察性網絡研究，將不同單位組織的資料庫，各自依據統一的研究 phenotype，進行分析，再將各單位研究結果整合，形成網絡式研究。

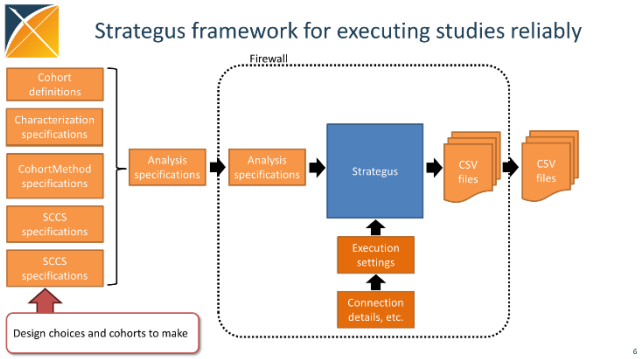
這個部分介紹如何透過 OHDSI 架構，實際設計一項觀察性網絡研究。OHDSI 架構透過標準化程序，來產生可靠的實證(reliable evidence)。OHDSI standardization 標準化程序包含：

- 標準化詞彙(Standardized vocabularies)：將來自世界各地的術語對應到一個通用集合

- Common data model：相同的 schema，包含相同的詞彙
- Cohort generation：在將原始資料對應到概念時做出類似的選擇
- Common study design approach：基於大規模比較的研究設計選擇及需要特別注意的偏差
- Common diagnostic approach：共享測試和閾值
- Library of analytic functions：使用相同的分析函數套件

以 OHDSI 架構進行網絡式研究產製可靠實證的研究歷程：

- 使用標準化的通用資料模型和詞彙表
- 共享研究執行方法
- 共享統計和資料處理的函數庫
- 透過使用者介面以鼓勵良好的設計—包括生成程式碼以 R 套件的形式執行研究
- 依在地特性建立適用之 R 套件版本
- Strategus 架構
  1. 宣告研究設計規格，並包含充足的預設值選項。
  2. 一次確定所有個別研究所需要的軟體版本。
  3. 更快的研究產生速度，更快的當地端研究執行速度。
  4. 目前使用程式碼產生研究設計宣告的規格(declarative specification of study design)，未來朝向使用者介面生成研究設計宣告的規格



OHDSI 的 Strategus 是一個框架 (framework)，採用開源工具與標準化框架，目的在可靠地執行大規模的觀察性研究，其主要目標是將從大型健康資料中生成可靠證據的流程自動化及標準化，特別是針對如**群組特徵描述**、**群體層次效應估計**以及**病患層次預測**等任務。

Strategus **關鍵組成與運作方式**

- **標準化研究設計：該框架要求研究設計以結構化且機器可讀的格式呈現。這份設計包含一個完整研究的所有必要元件，例如研究世代（cohorts，即患者群體）、研究結果（outcomes）及分析設定(analysis settings)的定義。**
- **模組化 R 套件**

此框架利用 OHDSI 社群開發的多個專用 R 語言套件，包括：

CohortMethod：用於估計治療或暴露對特定結果的影響，常使用如傾向分數配對（propensity score matching）或分層（stratification）等方法。

PatientLevelPrediction：用於開發和驗證模型，以預測個別患者的結果的套件。

CohortCharacterization：用於描述和比較患者群組的特徵的套件。

SelfControlledCaseSeries：一種專門用於研究短暫暴露及其影響的特殊方法的套件。

- **自動化執行：一旦研究設計定案，Strategus 便會自動化整個分析流程。它會讀取研究設計，執行適當的 R 套件來進行分析，並儲存結果。這種自動化確保了相同的分析可以在不同的資料庫上執行，無需手動介入，從而提升了一致性與再現性。**
- **可再現性：透過將整個研究定義在一個標準化的 JSON 格式檔案中，並自動化分析過程，Strategus 使研究的再現和與其他研究人員分享變得非常容易。並促進研究的透明度且可信度。**

此外，Strategus 具有下列優勢：

• 加速研究執行：它透過聲明式 (declarative) 的研究設計規範，並提供大量預設值，加速研究的產生和本地執行。

• 軟體版本控制：Strategus 有助於解決軟體版本控制問題，確保所有研究的軟體版本一致。

• 確保研究可靠性：它確保研究分析的一致性、可重現性，並支持開放科學，減少因軟體錯誤或不一致診斷導致的「不良變異」。

在 Strategus 架構下進行網絡研究產製可靠的實證的 8 個步驟及對應的軟體或 R scripts：

1. Generate cohorts e.g. in ATLAS
2. DownloadCohorts.R – pick the cohorts
3. CreateStrategusAnalysisSpecificationTcis.R - …
4. StrategusCodeToRun.R – pick the database
5. CreateResultsDataModel.R
6. UploadResults.R
7. EvidenceSynthesis.R
8. app.R

底下細部說明每個步驟：

Strategus 的研究執行流程涉及多個步驟，從研究世代群體的生成到研究結果發布：

1. 生成研究世代群體 (Generate Cohorts)：
   在 ATLAS 或 CapR 等工具中定義研究所需的研究世代群體 (cohorts)，例如目標群體、比較群體和結果群體。定義群體是一組操作型定義的規格，用於識別符合特定條件並持續一段時間的研究個體的集合。
2. 下載群體定義 (Download Cohorts)：
   使用 DownloadCohorts.R 程式來選擇並下載在第一步中定義的群體。這些群體定義通常儲存在專案的 inst 資料夾中。
3. 建立 Strategus 分析規範 (Create Strategus Analysis Specification)：
   使用 CreateStrategusAnalysisSpecificationTcis.R 程式來定義各項分析的參數。這是一個關鍵步驟，它會生成 Strategus 需要的 JSON 格式規範檔案，用於驅動後續的分析。此步驟會設定：
   - 目標、比較與適應症群體 (Target-Comparator-Indications, TCIs)：定義哪些藥物

或介入是目標組，哪些是比較組，以及它們適用的適應症。

▪ 結果群體與清除窗 (Outcomes and Clean Windows)：定義感興趣的療效或安全性結果，並設定清除窗 (clean window)，以排除在研究開始前發生過的結果事件，避免不朽時間 (Immortal Time) 或重複事件的問題。

▪ 風險時間窗 (Time-at-Risk, TAR)：定義從藥物暴露開始到結果事件發生的時間範圍，例如「治療期間 (On treatment)」、「30 天」或「意圖治療 (Intention to treat)」。

▪ 排除的共變數概念 ID (Excluded Covariate Concept IDs)：在傾向分數模型中排除與治療高度相關或屬於中介變量的共變數。

▪ 負向對照結果概念集 (Negative Control Outcome Concept Set)：用於經驗性地評估系統性誤差。

▪ 研究期間 (Study Period)：可選定限制研究資料的時間範圍。

▪ 傾向分數配對比例 (PS Match Ratio)：設定傾向分數配對的比例，例如 1 對 1 配對或變數比例配對。

▪ 子群組分析 (Subgroup Analysis)：定義需要在哪些特定子群組（如年齡組別）進行分析。

4. 執行 Strategus 分析程式碼 (Strategus Code to Run)：
   使用 StrategusCodeToRun.R 程式來指定要在哪個資料庫上執行分析。Strategus 會將之前定義的分析規範應用於本地的 OMOP CDM 資料庫。
5. 建立結果資料模型 (Create Results Data Model)：
   執行 CreateResultsDataModel.R 程式，準備用於儲存分析結果的資料模型。
6. 上傳結果 (Upload Results)：
   使用 UploadResults.R 程式將本地生成的摘要統計結果上傳到一個中央結果儲存庫。OHDSI 採用分散式資料庫網絡，病患層級資料不會在不同站點間共享，只有摘要統計結果會被彙整。
7. 證據合成 (Evidence Synthesis)：
   執行 EvidenceSynthesis.R 程式來整合來自不同資料庫的結果，進行證據合成，產生綜合性的結果。
8. 運行 Shiny 應用程式 (Run Shiny App)：
   最後可透過 app.R 程式啟動一個互動式的 Shiny 應用程式，供研究人員瀏覽、探索和診斷研究結果，包括功效、流失情況、傾向分數模型、共變數平衡和系統性誤差等資訊。

Strategus 的輸入與輸出

• 輸入：Strategus 的主要輸入是標準化的 OMOP CDM 資料以及詳細的分析規範，這些規範以 CSV 或 JSON 檔案的形式定義了研究的目標、方法和參數。

• 輸出：輸出是摘要統計結果，這些結果會被上傳到結果資料庫，並可透過 Shiny 應用程式進行可視化和診斷，以評估研究的可靠性，例如功效、流失、傾向模型、共變數平衡、系統性誤差和 Kaplan-Meier 曲線等診斷資訊。

透過 Strategus 框架，OHDSI 實現了高度自動化、可重現且透明的真實世界證據

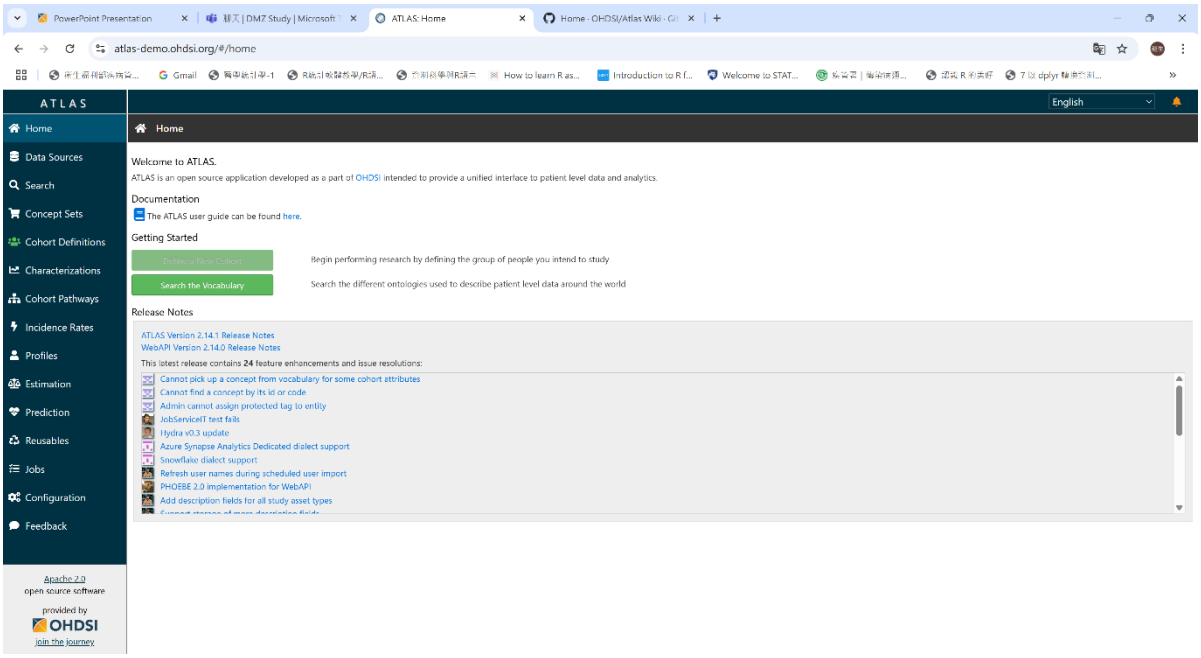
生成流程，大大提升了執行觀察性研究的效率和可信度。

Strategus 架構中的 **ATLAS 工具**

ATLAS 是由 OHDSI 社群開發的免費、公開且以網頁為基礎的開源軟體應用程式，其目的在支援觀察性研究分析的設計與執行，以從病患層級的觀察性資料中產生真實世界的證據。

ATLAS 是一個開源科學分析平台，可安裝於機構內部，以便對一個或多個已標準化為「OMOP 通用資料模型 V5（OMOP Common Data Model V5）」的觀察性資料庫進行分析。該平台還能促進與 OHDSI 社群內採用相同開源科學標準與工具的其他組織交換分析設計。

ATLAS 畫面與網址：https://atlas-demo.ohdsi.org/#/home



Concept sets：能夠建立研究者自己的程式碼清單，以便在整個標準化分析中使用。透過搜尋**標準化詞彙**庫並找出研究者感興趣的一組術語，之後將它們儲存下來，並在所有分析中重複使用。

Cohort Definitions：能夠讓研究者建立一組在特定時間內符合一項或多項條件的世代成員。這些世代成員隨後可作為所有後續分析的基礎輸入。這裡的世代，依研究設計所指定之規格，可以包含 case 組世代、control 組世代、outcome 組世代、indication 組世代等。

Characterizations(臨床特徵描述研究)：使研究者可以透過此功能，針對一個或多個已定義的世代族群，進行描述性分析並總結這些病患群體的特徵。

Cohort Pathways：使研究者可以檢視一個或多個世代群體中發生的臨床事件序列。

Incidence Rates：研究者可以估計目標族群中對於有興趣的結果的發生率

Profiles：研究者探索每一個體病患的長期追蹤觀察性資料，以總結特定個體身上正在

發生的事情。能夠於系統中，以個案事件序列史方式，呈現個案在觀察期間的所有事件，包含門診就醫、服藥、手術、檢驗檢查或住院等事件。
Estimation：研究者可透過此功能進行比較世代設計來進行群體層級的因果效應估計研究，藉此可以針對一系列的結果項目，進行一個或多個目標世代與對照世代之間的比較。
Prediction：研究者可以透過此功能運用機器學習演算法來進行病患層級預測分析，藉此針對任何指定之特定目標暴露，來預測其結果。
Jobs：提供使用者了解目前系統上正在背景執行的程序，並可透過更新按鍵蘭更新研究者的執行程序進度。
Configuration：研究者可以透過此功能了解有哪些已經於系統中配置的資料庫。

- **研究結果之診斷，負向對照(Negative controls)設計**

負向對照 (Negative Controls) 是一種在觀察性研究中，當存在無法測量的干擾因子 (confounders) 時用來檢測和校正偏差(bias)的強大工具。它的概念來自於實驗室科學，建立一個實驗的對照組或對照變數，其對應之結果應該是零、沒有關聯或無效果。如果這個組別的結果偏離了預期的零、沒有關聯或無效果，就代表該研究設計或分析中可能存在偏差。
負向對照主要分為兩種類型：
1. 負向對照暴露 (Negative Control Exposure, NCE)
   - 定義一個與真正感興趣的暴露 (exposure) 具有相似的潛在偏差來源，且已知不會對結果產生因果影響的變數。
   - 如果在研究中發現，這個負向對照暴露與結果之間存在關聯，那就表明可能存在未測量的干擾因子或其他的偏差來源。
   - 例如：假設在進行某種流感疫苗對流感住院率的影響研究時，擔心「尋求健康行為」這個未測量干擾因子對研究結果造成影響。可以使用受傷或創傷住院率作為負向對照結果。因為流感疫苗不應該會影響受傷住院率，如果模型發現流感疫苗與受傷住院率之間有關聯，這就強烈暗示了存在未測量干擾因子，如那些更注重健康的人既會打疫苗，也會更注意安全。
2. 負向對照結果 (Negative Control Outcome, NCO)
   - 定義一個與真正感興趣的結果具有相似的潛在偏差來源，且已知不會被暴露所影響的結果變數。
   - 如果在研究中發現，暴露變數與這個負向對照結果之間存在關聯，那就表示研究設計或分析中的偏差。
   - 舉例：假設在進行某種新藥對心臟病發作的影響研究，擔心有未測量的干擾因子。這時可以使用腳踝骨折作為負向對照結果。因為新藥不應該會影響腳踝骨折的發生率。如果實際模型發現新藥與腳踝骨折之間有關聯，這就證明該模型有問題，可能受到未測量干擾因子的影響。

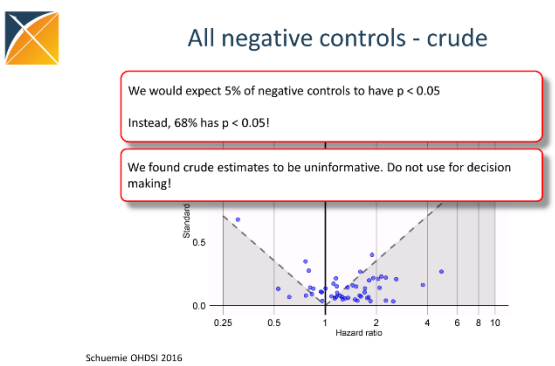
負向對照並不是用來直接修正偏差，而是用來偵測偏差是否存在。它就像一個

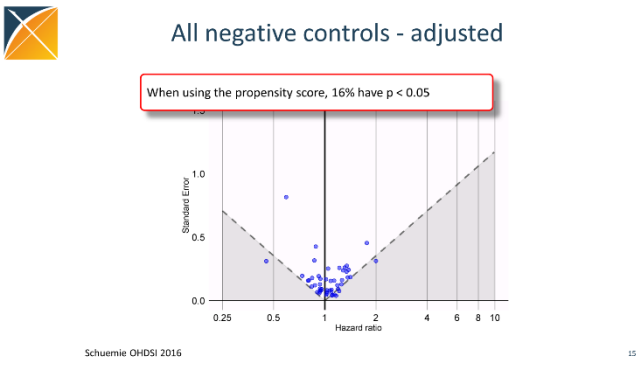
「品質檢查」工具，讓你能夠檢驗研究的假設(不存在未測量的干擾因子)是否成立。如果負向對照的結果如預期般是零關聯，這會增加對該研究結果可信度的信心。反之，如果出現了非零關聯，則需要重新審視研究設計或分析方法，以找出偏差的來源並進行修正。

底下以憂鬱症(Depression)對應可用的 negative control outcomes 為例
研究者列出一系列可作為負向對照結果的疾病：

Acariasis	Ingrowing nail
Amyloidosis	Iridocyclitis
Ankylosing spondylitis	Irritable bowel syndrome
Aseptic necrosis of bone	Lesion of cervix
Astigmatism	Lyme disease
Bell's palsy	Malignant neoplasm of endocrine gland
Benign epithelial neoplasm of skin	Mononeuropathy
Chalazion	Onychomycosis
Chondromalacia	Osteochondropathy
Crohn's disease	Paraplegia
Croup	Polyp of intestine
Diabetic oculopathy	Presbyopia
Endocarditis	Pulmonary tuberculosis
Endometrial hyperplasia	Rectal mass
Enthesopathy	Sarcoidosis
Epicondylitis	Scar
Epstein-Barr virus disease	Seborrheic keratosis
Fracture of upper limb	Septic shock
Gallstone	Sjogren's syndrome
Genital herpes simplex	Tietze's disease
Hemangioma	Tonsillitis
Hodgkin's disease	Toxic goiter
Human papilloma virus infection	Ulcerative colitis
Hypoglycemic coma	Viral conjunctivitis
Hypopituitarism	Viral hepatitis
Impetigo	Visceroptosis

將這些負向控制結果納入回歸模型分析，並將對應的效果估計值呈現於圖形中。底下呈現兩種模型推估所得到的結果。左圖呈現的是模型未經傾向分數校正干擾因子的結果，由圖形中可以看出 crude effect 的估計值，整體呈現正向偏差，且變異大且遠離 null hypothesis 即 HR=1，並且有許多負向控制結果的 effect 其 p-value<0.05；右邊圖形呈現的是模型經傾向分數校正干擾因子的結果，由圖形中可以看出經過校正後的負向控制結果 effect 值，其 p-value 大多>0.05，**偏差很小並且**集中於 HR=1 附近。**這兩張圖呈現透過**負向控制結果，來判斷原始研究設計是否有良好的校正干擾因子。此外，若是有其他未測量的干擾因子存在，也可以透過此結果判斷。





參、心得及應用建議

一、與會心得

本次參加於美國紐約市哥倫比亞大學生物醫學資訊研究所辦理之 OHDSI Summer School 2025 課程研習，此為 OHDSI 社群網絡第一次辦理的暑期研習課程，其目的是為了推廣與介紹如何應用 OHDSI 架構與系統工具，使用健康資料進行觀察性研究。有別於以往透過論壇形式辦理的研討會，Summer School 課程安排上，先有出發前往前收到的課前預習文獻、遠端環境與帳號設定，課程中理論發展與工具介紹及分組實作練習，完整且快速的於研習期間，帶領學員實際完成一項利用健康資料完成觀察性研究得到可靠實證的過程。

透過本次研習系統性的了解當前透過真實世界資料產製具可比較性的實證應用於改善健康照護的相關觀察性研究的方法，及最新的處理研究偏誤的方法；此外，透過 OHDSI 系統，了解如何快速佈署相同研究於其他不同機構或資料來源，並使完成的研究結果具可比較，之後可再透過統合分析的方法，進一步整合多方具可比較性之研究結果，提供可靠的臨床醫學實證，作為政策參採。

有幾點值得提出說明：

I. 高度標準化的資料結構、研究分析定義與流程：
在 OHDSI 架構下，透過 Common Data Model(CDM)，將來不同單位、系統的健康資料(臨床或保險申報資料)進行重新架構並定義詞彙，做為標準化分析的第一步。同時透過 OHDSI 系統將可以研究者設定之研究主題、架構與規格，輸出成的具研究計畫詳細規格和分析步驟之「phenotype」，可先針對該 phenotype 先進行驗證評估，完成後再實際進行資料分析作業。

II. 應用最新研究方法與使用最佳實踐進行研究：
OHDSI 架構的設計，主要目的是為了應用真實世界數據進行觀察性研究，產生穩健可靠的實證。所以在設計時，便將過去進行觀察性研究，可能遇到的研究問題，包含未完整校正干擾因素、選樣偏差、未測量干擾因子偏差、P 值操弄(P hacking)，以及出版偏誤(publication bias)，導入最新研究方法進行診斷與處理，以產生可靠的實證。

III. 透過資訊系統與工具快速產製可比較的實證：
開發資訊系統與工具，簡化研究者了解研究資料、統計分析方法的門檻，讓研究者可以專注於研究題目及要驗證的研究假設，並減少過程中可能的人為操作誤差。經過驗證與實作的 phenotype，亦可以提供其他單位或組織，快速納入進行相同的研究，可以進行多中心、多區域的研究比較。

IV. 更廣泛性評估研究議題：
由於 OHDSI 系統，提供系統性與標準化工具，可以快速導入進行觀察性研究分析，因此，相較過去傳統單一研究針對單一系列研究假設進行驗證的方式，可以同時在一次分析中，將所有相關的研究主題與架設一併納入分析，可以得到更全面性的實證證據，在提供政策參採獲臨床應用上，提供可議導入的空間。

v. 開放科學架構，產生可靠的真實世界實證：
透過資訊工具、標準資料結構與分析架構、開源軟體，將研究成果產製過程透明公開，以增加跨單位可比較性及增進可靠性。

二、應用建議

I. OHDSI 系統，針對觀察性研究產生可靠的實證，已有成熟的技術與推動經驗，建議可以依本署之業務屬性、研究發展主題領域、所擁有的資料結構等面向，整體評估是否適合導入 OHDSI 系統架構。

II. 本次研習除了學習 OHDSI 架構與應用外，課程中講者所介紹協助釐清研究問題之方法，可以納入參考。透過這些結構性問題系統性釐清觀察性研究所需要回答的問題為何？目標族群及主要的結果變數為何？這些概念性問題，於擬定研究計畫階段，能夠協助快速找到方向。此外，在尚未導入 OHDSI 系統架構前，OHDSI 所提出的觀察性研究所需使用之研究方法，如大量干擾因子校正的傾向分數法及負向對照結果法(Negative Control Outcome)，可以優先納入參考，應用於本署相關透過健康資料進行之觀察性研究中，如疫苗效果與安全性研究、治療抗病毒藥物之有效性與安全性等。

附錄

課前預習內容

本次行前課程講師提供了應用 OHDSI LEGEND 架構與資料庫所進行之研究分析專案後，將研究結果發表於專業期刊的兩篇學術文章。摘錄如下：

第一線高血壓用藥類別之有效性與安全性綜合比較：一項系統性、跨國性、大規模的分析

Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis.(Suchard et al., 2019)。

摘要

背景：對於高血壓的最佳單一療法，目前仍存在不確定性。現行臨床用藥指引在沒有共病症的情況下，建議將任何第一線藥物類別作為主要治療，包括：噻嗪類或類噻嗪利尿劑(thiazide or thiazide-like diuretics)、血管張力素轉化酶抑制劑(angiotensin-converting enzyme inhibitors)、血管張力素受體阻斷劑(angiotensin receptor blockers)、二氫吡啶類鈣離子通道阻斷劑(dihydropyridine calcium channel blockers)，以及非二氫吡啶類鈣離子通道阻斷劑(non-dihydropyridine calcium channel blockers)。但現有的隨機對照試驗未能進一步明確哪種藥物更優。

方法：本研究開發了一個全面的真實世界證據（real-world evidence）框架，能夠利用涵蓋數百萬名病患的觀察性數據，對多種藥物進行比較性的有效性和安全性結果評估，同時盡可能地減少固有的偏差。應用此框架，本研究設計採用「新使用者世代研究（new-user cohort design）」進行系統性、大規模的研究，在一個由六個健康保險理賠資料庫和三個電子健康紀錄資料庫組成的全球網絡中，比較所有第一線降血壓藥物類別的三種主要結果（急性心肌梗塞、因心臟衰竭住院、中風）和六種次要有效性結果，以及 46 種安全性結果的相對風險。該框架透過大規模的傾向分數調整、大量的對照結果，以及完全公開所有假設，來解決殘餘混雜因子(residual confounding)、發表偏差（publication bias）和 p 值操弄（p-hacking）等問題。

結果：本研究使用了約 490 萬名病患的數據，生成了 22,000 個經過校準、傾向分數調整後的危險比（hazard ratios），比較了所有藥物類別和結果。大多數的估計顯示，各藥物類別之間的有效性沒有差異；然而，噻嗪類或類噻嗪利尿劑(thiazide or thiazide-like diuretics)在主要有效性方面表現優於血管張力素轉化酶抑制劑(angiotensin-converting enzyme inhibitors)：在初始治療期間，其發生急性心肌梗塞（HR 0.84, 95% CI 0.75–0.95）、因心臟衰竭住院（0.83, 0.74–0.95）和中風（0.83, 0.74–0.95）的風險較低。此外，安全性方面也偏向噻嗪類或類噻嗪利尿劑(thiazide or thiazide-like diuretics)優於血管張力素轉化酶抑制劑(angiotensin-converting enzyme inhibitors)。而非二氫吡啶類鈣離子通道阻斷劑(non-dihydropyridine calcium channel blockers)則明顯劣於其他四種藥物類別。

結論：這個全面的框架為大規模的觀察性醫療照護科學研究提供了一種新方法。本研究的結果支持各藥物類別在開始高血壓單一療法時具有相同有效性，這與現行指南一致；但研究也發現例外，即噻嗪類或類噻嗪利尿劑優(thiazide or thiazide-like diuretics)於血管張力素轉化酶抑制劑(angiotensin-converting enzyme inhibitors)，而非二氫吡啶類鈣離子通道阻斷劑(non-dihydropyridine calcium channel blockers)則較為遜色。

第一線高血壓用藥中血管張力素轉化酶抑制劑與血管張力素受體阻斷劑之有效性與安全性比較：一項跨國性世代研究。

Comparative First-Line Effectiveness and Safety of ACE (Angiotensin-Converting Enzyme) Inhibitors and Angiotensin Receptor Blockers: A Multinational Cohort Study.(Chen et al., 2021)

摘要

血管張力素轉化酶抑制劑（ACE inhibitors）和血管張力素受體阻斷劑（ARBs）皆為臨床用藥指引推薦的高血壓一線治療藥物，然而兩者直接比較的臨床研究卻很少。本研究比較了在真實世界中，ACE inhibitors 與 ARBs 作為高血壓第一線治療藥物的有效性與安全性。

本研究採用回溯性的新使用者世代研究（new-user cohort design）設計，利用大規模的傾向分數調整、經驗校準和完全透明化等技術，來估計危險比，以盡可能減少殘餘混雜因子(residual confounding)和偏差。研究納入了 1996 至 2018 年間，來自美國、德國和韓國等 8 個醫學觀察資料庫中，所有開始接受 ACE inhibitors 或 ARBs 單一療法的高血壓患者。

主要研究結果包括急性心肌梗塞、心臟衰竭、中風和複合式心血管事件及 51 種次要和安全性結果，包括血管性水腫、咳嗽、昏厥和電解質異常等。

在 8 個資料庫中，本研究總共確認了 2,297,881 名開始使用 ACE inhibitors 治療的患者，以及 673,938 名開始使用 ARBs 治療的患者。研究發現，在主要結果上沒有統計學上的顯著差異：急性心肌梗塞（血管張力素轉化酶抑制劑相對於血管張力素受體阻斷劑的危險比為 1.11 [95% 信賴區間：0.95–1.32]）、心臟衰竭（危險比 1.03 [0.87–1.24]）、中風（危險比 1.07 [0.91–1.27]）或複合式心血管事件（危險比 1.06 [0.90–1.25]）。在次要和安全性結果方面，使用 ARBs 的患者在血管性水腫、咳嗽、胰腺炎和腸胃道出血方面的風險顯著較低。

總結來說，在這項大規模的觀察性網絡研究中，ARBs 作為高血壓第一線治療時，在藥物類別層級上與 ACE inhibitors 的有效性沒有統計學上的顯著差異，但安全性方面更佳。這些發現支持在高血壓初始治療時，優先開立 ARBs 而非 ACE inhibitors。

小結：

以上兩篇研究，皆是以 LEGEND-HTN study 研究成果，發表成學術文章，瞭篇文章採用相同之回溯性新使用者比較性世代研究設計，配合大規模分析框架，利用大規

模傾向分數調整、大量對照結果，以及完全公開所檢測的假設，來解決傳統觀察性研究常見之殘餘混雜因子(residual confounding)、發表偏差（publication bias）和 p 值操弄（p-hacking）等問題。同時相關研究結果，並作為修正臨床用藥只因的實證參據。

本次研習課程中，講師們便是透過 LEGEND-HTN study，作為標準範例，講解 OHDSI LEGEND 架構應用於觀察性研究，以產生真實世界之有效性實證。

與課程講師合影

George Hripcsak, MD, MS (左)與 Patrick Ryan, PhD(右)合影