

# 真實世界數據的研究設計-登記研究 (Registry-Based Study)的考量重點

衛生福利部食品藥物管理署

中華民國 113 年

## 目錄 (Table of Contents)

一、 前言.....	1
二、 目的及範疇.....	2
三、 登記研究(REGISTRY-BASED STUDY)和病人登記資料庫 (PATIENT REGISTRY)之間的差異 .....	4
四、 登記研究所收集之數據的適用性 (FIT FOR USE).....	6
(一) 登記資料庫數據關聯性 .....	6
(二) 登記資料庫數據可靠性 .....	9
五、 登記研究的設計與執行 .....	13
(一) 規劃登記研究：可行性評估 .....	13
(二) 登記研究計畫書.....	16
(三) 研究群體.....	17
(四) 數據收集.....	19
(五) 數據品質管理.....	19
(六) 數據分析.....	20
(七) 將登記資料庫串聯到另一個登記資料庫或另一個數據系 統時的考量事項 .....	24
六、 登記研究的應用範圍 .....	27
(一) 上市前證據的補充.....	27

(二) 提供上市後的證據.....	28
(三) 評估在懷孕期使用藥品的影響 .....	29
七、 結語.....	29
八、 參考文獻.....	30

## 一、前言

為因應國內外採用真實世界數據/證據作為查驗登記或變更登記支持性資料之趨勢，衛生福利部食品藥物管理署於109年7月22日公告「真實世界證據支持藥品研發之基本考量」。該指引之真實世界數據係指常規性蒐集與病人健康狀態相關或來自於健康照護過程中所得之多種數據，其中病人、疾病或藥品登記資料庫(Registry)都可以是真實世界數據重要來源。而使用包含登記資料庫在內的真實世界數據為資料來源，經適當分析方法產生的真實世界證據，得用於協助說明藥品之使用及其效益風險。

與其他真實世界數據來源相比，登記資料庫有以下優勢：登記資料庫可以收集具結構性且預先確定的數據元素(Data Element)，並且可以提供關於設定的相關病人群體及疾病歷程、併發症和醫療照護的長期數據。此外，登記資料庫可以是系統性地收集醫療申報數據庫、電子病歷或健康紀錄可能缺乏的病人自我報告數據。

然而，針對法規決策目的，登記資料庫也有其限制：現存的登記資料庫可能僅聚焦於特定一種疾病，即使與其他數據來源串聯，仍可能缺少或僅包含有限的共病資料。由於招募或追蹤病人的困難，登記資料庫納入的病人可能無法對感興趣的目標群體具代表性。例如，患有較嚴重疾病的病人可能比較願意參與並且持續接受

追蹤；相較於感興趣的目標群體，登記資料庫納入的病人群體可能具有不同的社經背景、自我照護方式或支持照護程度。這些問題都可能造成登記研究結果的偏差。其他可能限制包含資料異質性的問題（例如不同來源的群體具不同的臨床特徵）和用來處理資料品質方法上的差異等。

本文為 109 年公告指引的延續，進一步說明以登記資料庫進行研究來產生證據時應注意的事項，期使產學研界規劃或進行相關研究時有所參考。

## **二、目的及範疇**

本文旨在提供產學研界設計登記資料庫或使用現有登記資料庫進行前瞻性或回溯性的登記研究，重要方法學與法規考量的建議，以支持藥品研發。

一般而言，登記資料庫的建置是透過有組織的系統，以標準化的格式，藉由觀察性研究方式收集臨床數據，以評估特定疾病、健康狀況和暴露人群的臨床治療結果。這涉及招募預先定義的群體，並對該群體中的每一位病人收集預先定義的健康相關數據（以個別病人為單位作數據收集）。數據可以直接輸入到登記資料庫中，也可能包含來自其他可識別病人特徵資料庫的額外串聯數據。登記資料庫具有支持藥品開發的潛力，也針對介入性臨床試驗或非介入性

(觀察性) 臨床研究的設計提供了資訊並支持其執行。此外，針對罕見疾病(Rare Disease)或新興先進醫療產品(Advanced Therapy Medicinal Product)治療(包括基因治療)，登記資料庫更是評估其有效性和安全性的重要數據來源。

登記資料庫也可用於執行務實性臨床試驗(Pragmatic Trial)的追蹤與資料收集。

研究委託或申請者對於接受特定藥品或治療的病人之登記資料庫不屬於本文探討的範圍。從法規觀點而言，這類型的登記資料庫收集的數據有可能超出病人常規的追蹤範圍，此等資料庫，宜視為介入性臨床試驗或非介入性臨床研究的數據收集。

登記研究的執行應遵守人體研究法等適用的法規要求。

本文旨在說明法規單位接受以登記資料庫做為數據來源，並用以支持法規決策的原則與期待。申請者宜選擇適合其目的之研究方法，在符合本文所提示的注意事項下，使用適當的方法進行研究。

### 三、 登記研究(REGISTRY-BASED STUDY)和病人登記資料庫(PATIENT REGISTRY)之間的差異

登記研究是基於病人登記資料庫所進行的研究。登記研究和病人登記資料庫之間重要方法學上的差異整理如下表：

主題	登記研究	病人登記資料庫
定義	設定數據收集架構以前瞻性地收集病人數據；或是從既有的登記資料庫來收集病人數據，以進行研究問題或假說的探討	為有組織的數據收集系統，針對由特定疾病或狀態所定義的群體，為了特定的目的而設立。且用來執行基於登記資料庫的研究。
時間軸	時間軸係由特定研究目標相關的研究數據的收集／擷取和分析來決定。	時間軸係由常規數據收集時程和為了登記所執行的任何預期數據分析來決定。
病人招募	由研究目標定義，可能包含登記資料庫群體的子集；若為基於病人登記資料庫之隨機對照試驗 ( Registry-Based Randomized Controlled	在登記目的範圍內有完整的招募；須記錄登記資料庫數據的代表性與外推性。

	Trials) 的情況，將記錄治療組別的分派(即隨機分派)。並分析紀錄研究結果的代表性和外推性。	
數據收集	僅限於研究問題需要的數據，包含潛在干擾因子和修飾因子的數據；可能需要收集其他不在登記資料庫中的數據；如果這類其他數據包含仿單或一般臨床實務之外的受試者監測，則需要參考臨床試驗相關的法規；除了使用次級數據外，研究可能還要收集初級數據。	根據病人登記資料庫的目的來收集廣泛的數據；欲收集的核心數據元素須有共識，且須有相同的定義、編碼系統和數據輸入程序。
分析計畫	詳細的統計考量通常在研究計畫書和登記計畫書之外的單獨文件中定	統計分析計畫包含的分析通常為描述性的，且根據記錄病人的時間間隔而定期



	義；為基於假說的統計分析計畫。	進行，或在登記計畫書中定義時間表。
數據品質管控	以基於風險的方法前瞻性地定義和評估特定研究的數據品質管控。	定期使用於收集程序和數據品質，尤其著重於數據元素的核心集、數據系統以確保數據的完整性(Integrity)。

舉例來說，國內有衛生福利部委託台灣癌症登記中心收集資料所建立的癌症登記資料庫，也有學界利用此癌症登記資料庫進行研究，進而發表論文。國內醫學會也有針對特定疾病或治療方式，進行資料的收集並建立病人登記資料庫。

#### 四、登記研究所收集之數據的適用性 (FIT FOR USE)

在採用登記研究做為法規決策之證據前，研究委託或申請者須考量登記資料庫數據的適用性 (Fit For Use)。透過評估數據的關聯性 (Data Relevance) 和可靠性 (Data Reliability) 來考量數據是否適用。數據關聯性包含關鍵數據元素 (病人特徵、暴露、結果) 的可用性和對研究病人的充分代表性，而可靠性包含數據準確度、完整性、來源的可追溯性。研究委託或申請者亦可參考「真實世界數據—關聯性與可靠性之評估考量」指引。

##### (一) 登記資料庫數據關聯性

在考量是否要使用現有的登記資料庫進行登記研究以做為法規決策之證據前，研究委託或申請者對登記資料庫關聯性的整體評估，須考量登記資料庫是否能充分評估研究的科學目標，進而仔細考量登記資料庫所須擷取的數據元素。

須由登記資料庫擷取特定的數據元素取決於研究委託或申請者意圖使用登記資料庫的目的。例如，與透過登記資料庫來招募介入性研究的受試者相比，將登記資料庫做為單臂試驗之外部對照組數據來源，所需收集的數據元素必須更為全面完整。登記資料庫須保留任何不再收集的數據元素紀錄，或開始要收集的新數據元素資訊。研究委託或申請者也須擬定計畫以減少登記資料庫參與者之缺失數據產生。

對數據關聯性的評估取決於其背景。例如，當考慮將登記資料庫做為法規決策的目的時，研究委託或申請者須考量病人選擇的方法，以及此方法對登記資料庫中病人群體代表性的影響。特別是登記資料庫原先的納入和排除條件，可能會導致在登記研究中的病人群體與藥品開發計畫的目標群體出現差異。此外，留在登記資料庫中的病人可能與退出登記資料庫的病人不同（例如，經歷不同的不良事件）。

登記資料庫對納入的病人所擷取的數據元素通常包含病人特

徵、接受的治療和健康結果。這類數據一般包括唯一的病人識別碼、病人同意參與登記資料庫的日期和當時病人的基線特徵，如人口學因子、共病症、疾病史和其他資訊。研究委託或申請者須根據他們預定的用途來考慮登記資料庫應具備的數據元素。

下列範例為登記資料庫可能包含的數據(非為全部)：

### 1. 人口學和臨床資訊：

(1) 病人人口學因子：包含出生日期、性別、種族、身高、體重、吸菸狀態、酒精使用和消遣性毒品(Recreational Drug)的使用。

(2) 感興趣的主診斷：包含診斷日期、檢測項目和結果、診斷碼，以及基因或其他檢測（若適用）；疾病等級、嚴重度、疾病負擔，以及疾病惡化的重要里程碑。

(3) 病人的共病症：包含這些疾病當前的狀態（例如，併發症、疾病表現）、評估日期和個別共病症的治療。

(4) 認為可以改變疾病嚴重度或惡化特徵的其他相關資訊。

### 2. 感興趣疾病的治療資訊（在適用的情況下）：

(1) 使用藥品的化學名稱和商品名、配方劑量、每項治療的開始和結束日期，以及停藥的理由（若適用）。

(2) 處置類型和日期，或處置過程中併發症的處置。

### 3. 健康相關結果的資訊：

- (1) 特定感興趣的臨床事件（例如心臟病發作、中風、住院、死亡）和發生日期，以及後續是否痊癒出院。
- (2) 其他臨床結果（例如疾病惡化、復發、失能、功能性狀態、生活品質測量）和發生日期。
- (3) 病人治療方式的變更和發生日期。

### 4. 當意圖收集關於懷孕或懷孕結果數據時之懷孕相關資訊：

- (1) 先前的懷孕史。
- (2) 最後一次月經日期（若已知），以及評估妊娠週數的超音波報告。
- (3) 藥品暴露的妊娠時間點。
- (4) 母親的結果，包含子癇前症、子癇症等。
- (5) 懷孕的結果，包含活產、流產、死胎等。
- (6) 胎兒的結果，包含重大先天性畸形、胎兒小於妊娠年齡、早產、低出生體重、任何其他相關的不良胎兒結果等。

## (二) 登記資料庫數據可靠性

在考量使用現有登記資料庫或建立新的登記資料庫時，研究委託或申請者須清楚地瞭解管理登記資料庫的運作、人員的教育和訓練、資源規劃等的流程和程序，以及確保登記資料庫品質的一般實

務操作。此等管理屬性有助於保證登記資料庫可達成其目標，並且應包含但不限於：

1. 對意圖使用登記資料庫數據進行分析者提供一個確立的數據字典(Data Dictionary)和規則，以便登記資料庫數據查詢和編輯檢查的驗證（若適用）。
2. 對登記資料庫定義各項流程和程序，例如：
  - (1) 數據的收集、整理(Curation)、管理和儲存，包含適當的流程，有助於確保登記資料庫中的數據可透過該資料庫的來源數據（若適用）來確認。
  - (2) 病人、研究人員和醫師將如何存取登記資料庫和數據收集系統的數據，並與其互動的計畫。
  - (3) 除登記資料庫建立者以外，其他各方使用登記資料庫數據的條款和條件（例如，研究委託或申請者須滿足的條款和條件，以允許將登記資料庫的數據與另一個來源的數據合併）。
3. 電子紀錄包含存取管控的維護並稽核其路徑，以顯示登記資料庫數據的來源，並支持數據的可追溯性。

在設計登記資料庫，及擬定後續使用登記資料庫數據的計畫書時，研究委託或申請者須確保登記資料庫遵守適用的法規要求，包

含保護病人健康資訊隱私。在建立登記資料庫時，建議諮詢機構審查委員會或獨立倫理委員會，以審視與登記資料庫相關的數據收集和其他程序。

法規單位在評估登記資料庫數據可靠性時，考慮的因素包含數據如何蒐集（Data Accrual）、在數據蒐集和分析期間登記資料庫人員和流程是否能充分保證最少的錯誤和數據有足夠的完整性。研究委託或申請者須關注登記資料庫是否具有適當的隱私和安全管控，以確保數據的機密性和安全性得到保護。

為了支持在登記資料庫內收集到可靠的數據，登記資料庫的數據字典須包含：

1. 數據元素及其定義方式。
2. 數據元素的範圍和容許值。
3. 數據元素參考的來源資料。

建議研究委託或申請者使用常見的數據元素，以促進標準化、一致性、且通用的數據蒐集，如此將有利於登記資料庫數據和其他來源數據的比較或連結。登記資料庫使用的標準化術語，及其相關的數據標準，應與法規單位使用的一致。

應有適當的政策和程序來支持登記資料庫數據的可靠性，包含預先針對登記資料庫設定數據查詢和編輯檢查的驗證規則，並驗證

收集登記資料庫數據的電子數據系統。且有額外的政策和程序，使得法規單位和有興趣使用登記資料庫數據的人能評估數據品質，包含協助解決諸如編碼錯誤、或解讀一個或多個來源的文件、以及數據輸入、移轉或轉換錯誤等的問題。輸入登記資料庫的數據格式和定義應始終保持一致，隨時間推移任何診斷標準或臨床定義的改變都須說明並記錄在案。

電子資料庫形式的登記資料庫應具備適當的保障措施，包含支持確認數據的數據管理策略。數據管理策略應包含以下流程和程序：

1. 通過記錄日期、時間和最初數據輸入人員，進行預防性和／或矯正性行動來處理數據變更（包含在不刪除錯誤數據的情況下標記錯誤的數據，同時插入修正的數據以供後續使用），以及在不掩蓋先前紀錄的情況下說明任何數據變更的理由等，來實施並維持版本的控制。
2. 確保由另一個數據格式或系統傳輸的數據在轉移過程中不會被更改。
3. 尋求整合登記資料庫中先前使用現在已過時的數據格式或技術（例如作業系統、硬體、軟體）所收集的數據。

4. 說明隨時間臨床資訊的變化（例如疾病診斷標準、癌症分期）。
5. 解釋使用的稽核規則和方法，以及用來減少錯誤的策略。
6. 描述根據稽核識別出的錯誤之類型，以及如何修正數據。

數據的一致性、準確性和完整性須定期進行評估，其頻率取決於登記資料庫收集數據之目的。須進行常規描述性統計分析，來偵測任何缺失數據(Missing Data)、不一致數據、異常值(Outliers)和失去追蹤的程度。

## 五、登記研究的設計與執行

### (一) 規劃登記研究：可行性評估

第一步是先決定所要處理的科學問題以及評估此問題是否適合以登記研究來回答。

規劃一個登記研究可能需要一個到數個登記資料庫，需要由各個登記資料庫及各個單獨中心(不存在中央登記平台的情況下)取得合作協議，最好能找第三方來建立數據庫、數據擷取流程、及數據品質管控。另須了解使用來自不同登記資料庫的數據，是否需要簽署新的受試者同意書。研究委託或申請者應規劃數據處理計畫(Data Management Plan)以說明數據的來源，預期的研究結果，數據如何獲得、互通和重複使用，以及品質管控。

研究委託或申請者在撰寫試驗計畫書(Protocol)前應先進行登記



研究的可行性分析，可行性分析可單獨送件或連同研究計畫書送件。可行性分析應包含以下項目(若適用於該研究)：

1. 敘述所使用的登記資料庫或登記資料庫網絡，相關的疾病流行病學及現行治療。
2. 是否可從登記資料庫中獲得研究所需的核心數據，包含干擾因子(Confounding Factors)和可能影響研究結果的變數，是否可轉用標準醫學用語(例如: Meddra)，記錄頻率，以及必要時是否有收集任何其他數據元素或引進其他數據收集方法的能力。
3. 分析可用的數據元素的品質、完整性、即時性，包含是否有缺失數據及其插補方式，是否出現相同病人資料重複的情況，數據是否經過稽核驗證等。若使用不同的登記資料庫，宜分析登記資料庫之間的差異及其對數據完整性的可能影響，亦須考量登記資料庫間能否互通並整合數據。
4. 登記資料庫中通報不良反應的方式及如何從登記資料庫中收集不良反應，以及在需要的情況下引進額外流程收集或通報不良反應。
5. 納入登記資料庫的醫院數目、參與登記的病人和活躍病人數、每月/每年招募的新病人數、暴露於感興趣藥品的病人數、追蹤持續時間、失去追蹤或缺失數據、是否須取得受試者同意書及

取得受試者同意書的可行性等，將用來評估研究的規模，以及完成病人招募所需的時間。

6. 分析任何可能的資訊偏差(Information Bias) ，包含因醫院和病人納入/排除條件造成的選樣偏差(Selection Bias)、以及登記資料庫內或之間可能的時間偏差(Time Bias)等。評估任何可能產生的潛在干擾因子，特別是無法收集或測量的數據元素。
7. 根據數據特性和研究設計可能產生的分析問題。
8. 是否存在隱私或需要病人知情同意的問題而影響數據的獲取及共享。
9. 綜合評估所使用的登記資料庫對於研究的適用性，將前述任一方面資訊的遺漏都納入考量。

## (二) 登記研究計畫書

研究計畫書須敘述登記資料庫架構及所收納的病人群體如何達到該研究的目的，研究如何進行及研究結果如何驗證。

研究計畫書需先預估為達到目的所需收納的受試者人數。而達到所需受試者人數的可行性分析可從受試者納入排除條件及追蹤時間兩方面來評估。此外也包括估計目標(Estimand)如何定義，如何處理並發事件(Intercurrent Events)及缺失數據，是否需插補，療效大小(Effect Size)及受試者人數(Sample Size)對最後結果的影響，建議參考 ICH E9 (R1)以及財團法人醫藥品查驗中心網站之臨床試驗之估計目標與敏感度分析(臨床試驗統計指導原則之補遺)。若無法達到預期的受試者人數，需考慮是否從其他登記資料庫中收納病人，並綜合考量由多個登記資料庫收納病人的效益與限制。若為描述性研究(Descriptive Studies)，或許不需考量受試者人數對檢定力(Power)的影響，然而研究計畫書仍須論述此一規模的受試者人數能否達到試驗目的及研究結果估計的精確性。

若登記研究使用多個登記資料庫，儘管每個登記資料庫的特性可能有所差異，研究委託或申請者仍應規劃共同的研究設計與研究計畫書。研究計畫書中應敘述每個登記資料庫間的差異，以及相關差異是否影響研究結果，並提出敏感度分析(Sensitivity Analyses)以

從事此項評估。

當有多個登記資料庫可用，惟僅其中幾個用於登記研究時，研究計畫書應說明選擇該等登記資料庫的合理性，例如：受試者納入/排除條件，登記資料庫內數據的獲取及判讀。

若選擇由現有的登記資料庫收集研究特定的初級數據(Primary Data)，研究計畫書須載明由此登記資料庫收納符合資格病人並收集完整數據的程序。

若由現有的登記資料庫收集次級數據(Secondary Data)，研究計畫書應指明已在登記資料庫中收集感興趣的事件，並討論偏差和未測量干擾因子的風險。應使用具有針對性且完整的策略、編碼列表或判斷標準，來精確地定義感興趣的事件。

### **(三) 研究群體**

#### **1. 選擇研究群體**

如何選擇研究群體應視研究目的而定，可能使用登記資料庫內的所有病人或事先定義的次群體病人。例如：當研究目的為某一特定藥品的治療效果，研究群體可能包括：登記資料庫內新診斷且開始接受此藥品治療的病人、先前已診斷此疾病且由其他治療轉換接受此藥品治療的病人、接受此藥品為附加治療(Add-On Therapy)或僅在過去曾接受此藥品治療的病人。在這類情況下，可收集所有使用

此藥品病人的數據，描述整體結果，並識別出各子集間的差異，以便評估所收集數據的同質性(Homogeneity)和對整體的代表性(Representativeness)。

## 2. 知情同意

知情同意為法規要求臨床試驗執行前需符合的道德標準及義務。有關知情同意與同意書的要求，應參考我國的相關法規或公告規定。對於登記研究，除了須有病人同意加入此登記資料庫的同意書，可能也需要病人簽署加入登記研究的同意書。同意書須清楚說明研究目的、預期持續時間、使用病人資料的目的以及如何取得和處理(包括主管機關的監管和查核)。此外，也須說明研究結果可能獲得的資訊和造成的影響。若病人原已同意使用此登記資料庫進行登記研究，進行相關研究時或可免重新簽署同意書。

## 3. 資料保護

進行登記研究須符合我國個人資料保護法對於個人數據取得、處理及互通的要求。個人數據為可直接或間接連結到個人的資訊。當處理數據時，須考量數據是否可能連結到個人。

執行登記研究，處理個人數據的法律標準須先建立。特別是處理較敏感的個人數據應有特殊的考量。登記資料庫處理者(可能是數據庫持有者、研究委託或申請者、試驗主持人)有責任採取技術性或

組織性的手段，確保個人資料能夠以符合法規要求的方式處理。

#### **(四) 數據收集**

登記研究的數據可能來自於既存於登記資料庫的次級數據或基於研究目的而收集的初級數據。

對於既存於登記資料庫的次級數據，應僅擷取用於驗證或推導研究結果的數據，例如：藥品使用時間、治療結果、干擾和修飾因子、和病人群體的特徵。

基於研究目的而進行的初級數據收集，可能會增加登記研究的複雜度。數據收集的方法須清楚記錄於研究計畫書，因為數據收集方法會造成潛在的偏差和干擾，導致缺失數據及影響安全性通報。基於研究目的而進行的初級資料收集也可能影響登記資料庫原本的數據收集，需要稽核及驗證。

#### **(五) 數據品質管理**

登記研究的數據品質管理取決於許多因素，包含計畫如何使用研究的結果，以及使用的為登記資料庫的初級數據或次級數據。儘管登記資料庫的數據品質管理是登記資料庫持有者的責任，研究委託或申請者有責任對登記研究的數據品質進行管理，並且檢視品質管理的結果。數據管理計畫須說明數據如何被驗證及發現相關問題時採取的措施，包含內部或外部稽核。數據管理計畫須預先與登記

資料庫持有者取得協議始能執行。

應透過可行性分析的指引來選擇適當的方法和特定的措施，以減少得到無效研究結果的風險：

1. 應確認並監測任何數據清理、擷取和轉換的過程，特別是研究使用登記資料庫網絡，且不用中央數據轉換時。透過以風險為基礎(Risk-Based)的方式鑑別重要數據的數據保護及研究結果的可靠性。
2. 應進行數據品質檢查，以警示錯誤、遺漏、超出範圍、以及邏輯不一致的數據，並在必要情況下採取即時的數據驗證和補救措施。
3. 在收集並使用初級數據的研究中，應識別出各種影響品質的因素(例如有限的人力或物力資源，或不充分的訓練)，並加以處理以維持登記研究的完整性。可能的措施包含隨機來源數據的驗證、現場檢視用於數據收集和管理的流程和電腦化系統。

## **(六) 數據分析**

對於感興趣指標的分析方法應預先於研究計畫書及統計分析計畫中設定。研究執行過程中若要改變原先分析方法，應透過計畫書變更(Protocol Amendment)，以修訂研究計畫書及統計分析計畫。所有變更應呈現於試驗報告。

建議參考 ICH E9 (R1) 以及財團法人醫藥品查驗中心網站之臨床試驗之估計目標與敏感度分析(臨床試驗統計指導原則之補遺)，定義估計目標與檢定方法，以及缺失值的處理。可使用敏感度分析探索主要估計目標估計值(Estimate)的穩健性。

數據分析應評估研究群體對於事先定義的目標病人群體的代表性，因為這可能會影響登記研究的外部效度。在收集並使用初級數據的情況，應針對登記研究中符合資格的病人被納入、未被納入和中途退出研究者進行比較，也應比較研究中接受隨機分派的病人和未接受隨機分派的病人。在可能情況下，也應將研究群體和從文獻資料、可用的電子醫療資料庫、其他認為適用於研究但因研究計畫書中所述理由而未用於數據收集的登記資料庫，所識別出的類似群體進行比較，以為補充。

缺失數據可能會導致偏差和干擾，應在研究計畫書和統計分析計畫中詳盡說明其處理方式。並對缺失數據的分布、原因和時間點的假設說明其合理性。

在沒有隨機治療分派的非介入性登記研究中，應考量某些常見的分析問題：



1. 接受不同治療的病人群體其特徵可能會有差異。由於治療的決定可能會受到各種因素的影響，而該等因素(例如，疾病嚴重度或病人的監測)也可能與研究感興趣結果發生的風險相關。雖然有調整干擾因子的方法，但可能無法達到全面性，仍應考慮進行不同的敏感度分析。此外，確定治療改變的原因可能需要在研究過程中收集完整的資訊，假若這樣的觀察須跨越數年，單調整基線共變數可能無法完全解決問題。
2. 登記資料庫和登記研究可能在不同時間點收納並追蹤病人、開始感興趣的治療、以及收集特定事件和其他參數。感興趣事件發生的機率也可能與時間相關。考量這些時間點相當重要，因為它們會影響治療組之間的可比較性。可使用圖示以幫助了解登記研究和登記資料庫的不同時間點。研究治療效果時，當研究開始追蹤時間早於初次治療時間，開始追蹤和開始治療間不可能發生感興趣的結果事件，可能造成不死時間偏差(Immortal Time Bias)。
3. 在與歷史對照組(Historical Control Groups)比較時，可能發生選擇偏差、資訊偏差和時間性偏差。臨床背景(例如：可選擇的治療、診斷方式、依疾病嚴重度選擇治療的醫療實務、病人照護、數據收集的完整性或其他未收集或未知因素)可能隨時間變

化。應識別這些偏差並評估對結果效度的影響。

4. 用以比較的對照組可能選自研究所使用登記資料庫以外的病人群體，例如從其他登記資料庫或藥品尚未上市國家/區域的電子醫療照護紀錄。在此情況下，應確保可能影響研究結果之兩個病人群體間的潛在差異能被充份地評估。由於可能無法識別出所有的潛在差異，資料收集的完整性也可能有差異，這類比較須謹慎判讀。
5. 登記資料庫對於接受感興趣治療的病人與未治療或接受其他治療的病人提供長期比較的機會。開始納入研究追蹤前已接受治療一段時間的病人，有兩項可能的偏差：首先，這些病人是早期治療的「存活者」，如果發生不良反應的風險隨時間變化(例如：治療本身帶有過敏反應的風險或心血管風險)，這可能會引起重大的(選樣)偏差。其次，進入研究時藥品治療相關的共變數(例如疾病嚴重度)，可能會受到先前藥品治療的影響，或病人健康相關的行為也可能有所差異(健康使用者效應)。採用僅納入藥品新使用者的限制(即在研究期間第一次使用感興趣治療於該疾病的病人)之研究設計，可降低這類偏差。然而缺點是樣本數較小，以及具有長期暴露的病人可能減少，進而可能影響結果估計的精確度。

6. 若採用上述僅納入藥品新使用者的設計，採用活性對照組可減少不同適應症或不同疾病嚴重度造成的干擾，因不同治療組的病人皆為該適應症須接受治療的病人。然而若研究藥品為新上市藥品，上述設計可能無法找到理想的活性對照組，因新上市藥品處方對象通常為特定預後或符合保險給付條件的病人，因而可能造成偏差。

### **(七) 將登記資料庫串聯到另一個登記資料庫或另一個數據系統時的考量事項**

當登記資料庫未能擷取到所有必要的資訊，以回答感興趣的問題時，研究委託或申請者可能會考慮從其他來源取得補充數據。例如，研究委託或申請者可能會考量將登記資料庫中病人的資料串聯到另一個數據系統或多個系統中的相同病人，例如另一個登記資料庫、電子健康紀錄(Electronic Health Record, EHR)、醫療申報資料庫，或透過數位健康科技(Digital Health Technologies, DHTs) 允許持續或間歇地記錄生理學和/或行為數據的感測器（例如血壓、身體活動 Physical Activity、血糖濃度）、或在通用（General-Purpose）運算平台上運作的軟體應用程式等。

如果登記資料庫與來自另一個數據系統的數據串聯合併，研究委託或申請者應考量額外的數據來源對登記資料庫數據的完整性之

潛在影響。研究委託或申請者應使用策略來修正多餘的數據，解決數據中的任何不一致，並處理其他潛在的問題(例如，在移轉數據時安全地保護病人隱私的能力)。研究委託或申請者應制訂計畫，以解決病人層面串聯的充分性（即確保匹配到同一病人數據）。此外，試圖將病人層級數據串聯到另一個登記或數據系統時，研究委託或申請者還應考量法規的要求(如，特定國家的法律)。

研究委託或申請者也應考量串聯的數據來源是否可以交互運作 (Interoperable)，並以適當的資訊策略 (Informatics Strategies) 支持，以確保數據的完整性。研究委託或申請者應確保 (1) 進行充分的測試來證明數據系統的交互運作性；(2) 以一致且可重複 (Repeatable) 的方式傳輸自動化電子數據元素至登記資料庫；(3) 準確、一致並完整地傳輸數據。應使用預先定義的規則來檢查邏輯一致性和數值範圍，以確認能從串聯數據來源中準確地擷取數據，並且與串聯的數據來源變數的操作定義一致。

用於驗證資料傳輸過程的紀錄，應於研究委託或申請者被查核時提供給法規單位審查。研究委託或申請者也應確保軟體的更新或額外的數據來源不會影響登記數據庫數據的完整性、交互運作性和安全性。例如，應考量串聯數據和登記資料庫數據時間上是否正確對齊。

使用其他數據來源的適當性也取決於研究委託或申請者打算如何使用串聯數據，以及是否有能力對所有病人取得類似的數據。例如，對於每個可能的數據來源，研究委託或申請者應考量：

1. 該串聯數據針對所提出的研究問題是否適用（例如，額外的數據來源提供相關臨床細節和/或長期追蹤資訊）。
2. 數據能否準確地與登記資料庫的病人對應，以及兩個（或多個）資料庫能否準確地執行紀錄的串聯。
3. 感興趣的變數在登記資料庫和額外的數據來源是否具有一致的定義和可靠的確認方法。
4. 收集的數據是否具有足夠的準確度、一致性和完整度，以達到登記資料庫的目標。

在研究委託或申請者決定使用一個或多個額外數據來源來補充登記資料庫後，研究委託或申請者應開發將此等數據串聯至登記數據庫所需的方法和演算法。此外，研究委託或申請者應確定如何評估數據的完整性，包含如何評估任何由串聯所引入的不準確（例如特定數據測量的超額計數）。研究委託或申請者也應使用適當的方法對每個串聯的數據來源進行數據的輸入、編碼、清理和轉換等。

## 六、登記研究的應用範圍

登記研究能否用於法規決策的目的取決於許多因素，包括研究藥品本身的特性、登記資料庫的特性以及登記研究的目的、及設計與分析計畫等。一般而言，登記資料庫較適合用於擷取客觀指標的數據（例如死亡或住院）。雖然登記資料庫也可用於擷取主觀指標數據（例如病人疼痛程度），惟測量方式的標準化是額外的挑戰。此外，旨在收集數據以回答特定研究問題的登記資料庫，可能比原先用於其他目的之現有登記資料庫更具有優勢。開發這類登記資料庫，可針對特定的研究問題收集相關的臨床指標和結果數據，而現有登記資料庫可能須透過與其他數據源串聯的方式才能取得。總結來說，執行登記研究前，建議及早諮詢法規單位的意見。

登記研究產生的證據對於法規決策可能有用的案例包括：

### (一) 上市前證據的補充

臨床試驗依然是上市前藥品療效與安全科學性評估的核心。在某些情況下，該等評估可由病人登記資料庫的觀察性證據得到支持。此等證據範例包括：臨床實務中該疾病的發生率與疾病結果的決定因素、該疾病的標準治療方式資訊、或目標群體的特徵等。

基於病人登記資料庫的研究也可提供做為無對照組之單一組別臨床試驗的比較組，以獲得試驗結果；也可用於支持基於病人登記

資料庫之隨機對照試驗 (Registry-Based Randomized Controlled Trials)，用於招募受試者（例如，識別出符合納入/排除條件的病人）與收集數據。

## **(二) 提供上市後的證據**

基於病人登記資料庫的研究，可做為上市後執行的隨機對照試驗、非介入性研究、上市後療效研究 (Post-Authorization Efficacy Studies, PAES)和上市後安全性研究 (Post-Authorization Safety Studies, PASS)。先前於隨機分派臨床試驗證實療效的藥品，登記研究有助於評估於臨床實務上使用的劑量調整方案之有效性，以及瞭解藥品在廣泛臨床相關疾病和較具異質性的病人群體中的有效性和安全性。在具有足夠樣本數和適當研究設計情況下，基於病人登記資料庫的上市後安全性研究可提供數據以量化和描述風險、識別出不良反應發生的風險因子、評估藥品長期使用的安全性、或評估藥品使用模式以增加對藥品效益風險概況的瞭解。

新興先進醫療產品很大部分是針對罕見疾病來發展，且經常以無對照組試驗進行療效與安全評估。由於藥品核准上市時提供的療效和安全性資料有限，因此上市後藥品療效和安全性的追蹤是強制性的，且通常會強制執行 PAES 和 PASS 來產生上市後證據。這些研究通常優先在現有病人登記資料庫中進行。

### (三) 評估在懷孕期使用藥品的影響

懷孕登記資料庫納入暴露或未暴露於不同治療的懷孕女性，並進行追蹤，以收集特定藥品的懷孕結果以及新生兒資訊。這類研究的挑戰與懷孕結果資訊的完整性、暴露時間範圍/孕期的確定有關，可能需要與出生缺陷登記、畸胎資訊服務或電子醫療照護紀錄收集的資料庫串聯，如此母親小孩的聯結才有可能。

## 七、結語

建立真實世界證據，須先對欲解決的問題提出研究假說與目標，擬定適當的研究設計與統計分析方法，依照研究設計來蒐集適當的真實世界數據，執行數據分析，方可產生符合研究目標的分析結果。

登記資料庫可作為真實世界證據的數據來源之一，與其他真實世界數據來源相比，登記資料庫有其優勢和限制。研究委託或申請者尋找登記資料庫進行登記研究時，建議參考本文相關考量並及早諮詢法規單位的意見，確保品質良好的數據來源，以獲得可信的分析結果。此外，登記研究的執行應遵守人體研究法等適用的法規要求。



## 八、參考文獻

1. 衛生福利部食品藥物管理署。真實世界證據支持藥品研發之基本考量。中華民國 109 年 7 月。
2. US FDA Real-World Data: Assessing Registries To Support Regulatory Decision-Making For Drug And Biological Products. Guidance For Industry (Draft). November 2021
3. EMA Guideline On Registry-Based Studies. October 2021
4. 衛生福利部食品藥物管理署。真實世界數據—關聯性與可靠性之評估考量。中華民國 110 年 3 月。
5. 衛生福利部食品藥物管理署。採用電子病歷資料進行臨床研究指引。中華民國 109 年 11 月。
6. 衛生福利部食品藥物管理署。使用真實世界數據/真實世界證據作為申請藥品審查技術文件應注意事項。中華民國 110 年 7 月。
7. 衛生福利部食品藥物管理署。真實世界證據的研究設計—務實性臨床試驗的考量重點。中華民國 110 年 1 月。
8. ICH E9 (R1) Addendum On Estimands And Sensitivity Analysis In Clinical Trials To The Guideline On Statistical Principles For Clinical Trials
9. 財團法人醫藥品查驗中心。臨床試驗之估計目標(Estimand)與敏感度分析(Sensitivity Analysis) (臨床試驗統計指導原則之補遺)

(第一版)。中華民國 109 年 11 月。

10. 財團法人醫藥品查驗中心。單臂試驗採用外部對照組之研發策略指導原則(第一版)。中華民國 111 年 10 月。