

國立臺灣大學醫學院附設醫院研究倫理委員會

使用病歷資料進行研究申請表

1. 若申請病歷資料^註進行研究並申請免除知情同意，請填寫此表。

註：前述所指之病歷資料，包含健檢與影像學檢查等個人健康與病歷資料。

2. 審查方式標準：

*申請病歷個案數 1,000 以上（資料僅在本院之臺大醫療資料分析專區分析之計畫則為 10,000 筆以上）若申請簡易審查，須符合以下全部條件：

a.取得之資料須為完全去識別化資料或資料來源為醫療整合資料中心但不取得病歷號、身份證字號及姓名上述三項可辨識身份的個人資料。

b.申請病歷資料的項目有特定範圍

c.資料僅在院內使用

*以上若有任一不符，則為一般審查。

一、基本資料			
計畫名稱			計畫主持人
所屬機構/科部			
題目	說明		
1.研究計畫類型	☐ 主持人自行發起： ☐ 無院外合作者（研究團隊人員均領有本院員工識別證）。 ☐ 有院外合作者： ☐ 本校系院所之師生； ☐ 校外學術研究機構：_____； ☐ 與廠商合作開發； ☐ 廠商部份贊助； ☐ 其他：_____ ☐ 廠商發起之委託研究計畫 ☐ 其他：_____		
2.資料來源與資料數量	2-1	資料來源/資料庫名稱（可複選）	資料數量（預計人數或筆數）
	☐	醫療整合資料中心	
	☐	影像醫學部	
	☐	電子病歷	
	☐	健保資料庫	
	☐	其他(例如：單位保管的病人清單去串聯其他資料庫、已建立之個別研究用資料庫)，請說明：_____。	
若有其他情形請說明：_____。			
3.樣本數估算 <註：資料數量以最小需求為原則，請說明合理估算方式依據>	請說明：		

4. 資料是否攜出院外 <註：若經審查同意，亦須經本會簽核院方同意後方得攜出，且攜出之資料必須經處理後無可辨識個人資料>	☐ 否，資料僅會在院內使用 ☐ 是，資料須攜出院外，否則計畫無法執行，請說明：	
	(1) 必須攜出之理由： (複選)	☐ AI 模型無法存到本院運算環境，請說明：_____。 ☐ 合作團隊資料無法存到本院環境，請說明：_____。 ☐ 研究使用軟硬體平台無法存到本院運算環境，請說明：_____。 ☐ 因技術專利或授權為合作團擁有，無法使用本院運算環境進行研究，請說明：_____。 合作團隊名稱：_____ 專利名稱：_____ 專利編號：_____ ☐ 其他理由：_____
	(2) 攜出方式：	
5. 資料保存地點及設備擁有者	※此項內容需一併呈現於計畫書中，請說明：記載於計畫書第____頁。	
	5-1	資料保存地點 設備擁有者（電腦含伺服器）
	5-2	資料運算分析地點 設備擁有者（電腦含伺服器）
	5-3	以上設備是否會與外部網路連結或傳輸？ ☐ 否，不會與外部網路連結或傳輸。 ☐ 是，請說明確保資訊安全之措施：_____。
	<註：資料若僅於【臺大醫療資料分析專區】使用，請於計畫書中敘明「向醫療整合資料中心申請之資料僅能於臺大醫療資料分析專區使用」；若非前述狀況或有必要理由，請於計畫書中敘明「遵循資料不得攜出院外」>	
6. 本研究使用之資訊科技(IT)	6-1	資訊科技 ☐ 無使用資訊科技，請說明使用之方式或統計工具：_____。 ☐ 使用資訊科技，如大數據資料探勘、人工智慧學習運算等，請說明提供者：_____。 【例如：XXXX 大學，OOO 教授實驗室；OOOO 公司...等】
	6-2	雲端運算服務 ☐ 無使用雲端運算服務； ☐ 使用雲端運算服務，由誰提供請說明： ☐ 同上 6-1 ☐ 其他：_____。
7. 資料安全及隱私保護措施	【保護措施舉例：資料保存於上鎖之 OOO 研究室，保存資料之電腦與伺服器以密碼保護，資料將於院內保存與運算分析。存取資料之權限管制如 9.1.3.1，可識別身分的個資將以代碼或加密保護。由主持人親自確認所有研究人員(含主持人)完成病歷隱私保護與資訊安全訓練、簽署保密聲明切結書、合作研究合約等。】	

	※資料安全及隱私個資保護措施需一併呈現於計畫書中，請說明記載於計畫書第____頁。 請說明：	
8.資料保存與結束後處理	8-1 資料保存與期限	☐ 研究結束後____年銷毀資料 ☐ 永久保存，請說明必要理由 _____。
	8-2 最終處理措施 【例如：儲存資料之硬碟將重新格式化銷毀資料等】	
9.資料類型 (可複選)	☐ 9.1.研究者取得的資料含有可識別身分之資料，研究者 <u>不會</u> 將之完全去識別化。 ☐ 9.2.研究者取得的資料含有可識別身分之資料，但研究者 <u>會</u> 將之完全去識別化。 ^註 ☐ 9.3.研究者取得的資料為完全去識別化之資料。 註： 1. 個人資料包括可直接或間接識別身分之個資(identifiers)及其他個資。前者包括最末頁附表一註解之個資項目等。完全去識別化是指資料中不包含可直接或間接識別身分之個資， <u>且</u> 於該研究案之研究者無法從取得之資料識別身分。 2. 如資料來源含「臺大醫院醫療整合資料中心」，請勾選9.1.或9.2.類別。	
	二、個資隱私保護(請依9.資料類型勾選的類別填寫)	
◎勾選「9.1.研究者取得的資料含有可識別身分之資料，研究者 <u>不會</u> 將之完全去識別化」者，請續完成所有9.1之子項：		
※此項內容需一併呈現於計畫書中，請說明：記載於計畫書第____頁。		
9.1	9.1.1 研究需使用可識別身分個資，否則計畫無法執行。	
	請說明必要理由：	
	9.1.2 研究需申請免除知情同意，否則計畫無法執行。	
	請說明必要理由：	
	9.1.3 本研究的風險已降至最小	
	9.1.3.1 有適當的規劃，保護可識別身分的個資，避免不當的使用與揭露？	

將接觸病歷資料之研究人員名單：<註：欄位可請自行增列>

姓名	單位	職稱	病歷隱私個資保護訓練課程的完訓日期	使用資料之權限為何？ 【例如：負責資料統計、負責資訊程式撰寫...】
本院（領有本院員工識別證者）				
本校其他系院所之師生				
其他學術研究機構：機構名稱 OOO				
廠商名稱：OOO				

9.1.3.2 請依研究時程提出第一時間點能銷毀可辨識身分個資(如姓名、病歷號)之時程規劃及預計時間。

<註：除非醫療或研究需要，或依法須保存者可另提說明>

請說明：

_____。

9.1.3.3 請確認本計畫僅收集向研究倫理委員會申請日以前之病人既有病歷資料，不會收集申請日後產生之未來病歷資料

☐是 ☐否(勾選此項請採用設計受試者同意書供審查，並於通過後取得受試者同意書再收集資料)

◎勾選「9.2.研究者取得的資料含有可識別身分之資料，但研究者會將之完全去識別化」者，請續完成所有 9.2 之子項：

9.2 ※此項內容需一併呈現於計畫書中，請說明：記載於計畫書第____頁。

9.2.1 研究者如何進行資料去識別化（包含：去除 free text 或影像資料中的 18 種 identifiers 的時間點、具體流程及資訊技術），請說明：

(1.1)時間點：	【應於試驗執行前端完成資料收集或資料連結完成，足以進行研究時，即時將可識別個資去識別化】
(1.2)具體流程：	
(1.3)資訊技術：	

9.2.2 刪除 18 種 identifiers 後的資料，研究者是否能藉其知識能力，識別特定個人的身分？

【例如：若資料中含有年齡與職業（台大校長），或含有高齡與多胞胎（新聞事件），則可識別出特定個人的身分。】

☐是，研究者藉其知識能力，無法識別特定個人的身分。

☐研究者藉其知識能力，雖可識別特定個人的身分，但有確實的技術及步驟，將進一步刪除這些資料【例如：使用 OOO 程式，進一步刪除 OOO 資料】，請說明：

_____。

☐不確定，或沒有確實的技術能力及步驟，可刪除這些資料。請改依 9.1.類別填寫。

9.2.3 研究者自評是否可將資料完全去識別化？

☐是，有確實的技術能力及步驟，可將資料完全去識別化。

☐否或不確定。請改依 9.1.類別填寫。

◎勾選「9.3.研究者取得的資料為完全去識別化之資料」者，請續完成所有 9.3 之子項：

9.3	9.3.1 資料提供者確實能完全去識別化？
	☐ 是。資料提供者確實能刪除 18 種 identifiers，包括去除 free text 或影像資料中的 identifiers。 ☐ 若否或不確定。請改依 9.1.或 9.2.類別填寫。
9.3.2	刪除 18 種 identifiers 後的資料，研究者是否能藉其知識能力，識別特定個人的身分？
	【例如：若資料中含有年齡與職業（台大校長），或含有高齡與多胞胎（新聞事件），則可識別出特定個人的身分。】 ☐ 是，研究者藉其知識能力，無法識別特定個人的身分。 ☐ 研究者藉其知識能力，雖可識別特定個人的身分，但有確實的技術及步驟，將進一步刪除這些資料【例如：使用 OOO 程式，進一步刪除 OOO 資料】，請說明： _____。 ☐ 不確定或沒有確實的技術能力及步驟，可刪除這些資料。請改依 9.1.類別填寫。

三、使用大型或小型語言模型生成式 AI 之檢核：

10	10.1 本計畫是否將病歷資料輸入大型或小型語言模型生成式 AI?
	註：生成式 AI 模型是一種電腦程式，旨在創建類似於人類製作(human-made)的新內容。例如 ChatGPT、Bing AI、Bard 和 Midjourney 等。因其大量蒐集、學習與產出之資料，可能涉及智慧財產權、人權或業務機密之侵害，且其生成結果，因受限於所學習資料之品質與數量，而有可能真偽難辨或創造不存在的資訊，須客觀且專業評估其產出資訊與風險。 ☐ 否，請跳答第四項 ☐ 是，生成式AI模型名稱及提供者(個人/機構/廠商)：_____。 答否請跳答四：
	10.2 是否屬封閉式地端部署之生成式 AI 模型？
	註：「封閉式地端」係指機關內不與外部網路連線或資料交換之地端環境，若機關採租用運算資源方式建置生成式 AI 應用，其租用之環境必須符合以下兩項條件，即視為「封閉式地端」： 1. 運算環境座落於臺灣且虛擬主機為租用者及其限定使用者專屬使用。 2. 租用者及其限定使用者採 VPN 或專線方式連入、使用該運算資源。 *機關內係指本醫療體系環境內 ☐ 是 ☐ 否
10.3	是否會與生成式 AI 之提供者(含機構/廠商/個人)簽署資料不會保留於該提供者系統內並確保不會外流之合約？
	☐ 是，請檢附文件，相關約款記載於第____頁；若尚未簽署，預計何時簽署合約或其他說明：____ ☐ 否，請說明如何確保病歷資料不會被保留及外流_____
10.4	資料之種類、型態及風險程度(依衛生福利部制定人體生物資料庫資料國際傳輸之風險程度分類_如附表二)：
	☐ 低風險 ☐ 中風險(部份基因體定序之序列資料(例如：target sequencing, CHIP-seq.)或基因型定型資料、轉譯體定序之序列資料(例如：RNA-Seq,miRNA Seq)、外顯子全基因體定序之序列資料(Exome sequencing) ☐ 高風險(含全基因體定序之序列資料(Whole Genome DNA Sequence)、經全基因體定序所得之全基因體基因型定型資料(Whole Genome Genotypes from Whole Genome Sequencing)、經微晶片陣列所得之全基因體基因型定型資料(Whole Genome SNP Genotyping)、親子鑑定或身份鑑定相關之基因定型圖譜資料) *若屬高風險，原則須以受試者同意書取得同意，並採一般審查

10.5 輸入生成式 AI 之資訊是否已確實刪除 18 種 identifiers，包括去除 free text 或影像資料中的 identifiers。

☐是 ☐否，勾選否，原則不適合執行，請說明必要理由及如何避免疑慮：

10.5.1 刪除 18 種 identifiers 後的資料，使用者是否能藉其知識能力，識別特定個人的身分？

☐否 ☐是，勾選是，原則不適合執行，請說明必要理由及如何避免疑慮：

10.6 使用生成式 AI 是否可能造成智慧財產權或人格權之侵害？

☐否 ☐是，勾選是，原則不適合執行，請說明必要理由及如何避免侵害：_____

10.7 計畫設計是否以未經確認之產出內容直接作成臨床診斷或治療之唯一依據？

☐否 ☐是，勾選是，原則不適合執行，請說明必要理由及如何避免疑慮：_____

四、本院與本院研究人員的權益

11. 是否有院外人員使用本院病歷資料？【例如：本校系院所之師生、校外學研機構或廠商，或使用廠商提供的雲端運算服務。】

☐否

☐是，請續填 11.(1)~(3)。

11 (1)研究成果的歸屬：研發成果商品化時智慧財產權益之歸屬？

請說明：

_____。

(2)研究成果回饋方式為何？

請說明：

_____。

(3)是否已簽署合作研究合約或產學合作研究合約，約定智財權益的分配及成果保密義務，其相關約款為何？

☐是，已簽署合作研究合約或產學合作研究合約，請檢附文件，相關約款記載於第____頁。

☐否，預計何時簽署合約或其他說明：_____

※美國 HIPAA 規定的 18 種 identifiers

附表一

The following identifiers of the individual or of relatives, employers, or household members of the individual, are removed:

	英文名稱	中文名稱
1	Names	姓名
2	All geographic subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code if, according to the current publicly available data from the Bureau of the Census: (1)The geographic unit formed by combining all zip codes	所有小於國家的地理分區，包括街道地址，城市，縣，區，郵遞區號及其等效地理編碼，郵遞區號的前三位數除外，如果根據當局公佈的數據人口普查： (1) 將所有郵遞區號與相同的三個初始數字組合而成的地理單位包含 2 萬多人 以及

	with the same three initial digits contains more than 20,000 people ; and (2)The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000.	(2) 所有包含 20,000 人或更少人的地理單位的郵遞區號的最初三位數字改為 000。
3	All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older	與個人直接相關的日期（年份除外）的所有要素（月，日），包括出生日期，住院日期，出院日期，死亡日期；所有年齡超過 89 歲以及所有指示此年齡的日期（包括年份），除了這些年齡和要素可以合併為 90 歲或以上的單一類別
4	Telephone numbers	電話號碼
5	Fax numbers	傳真號碼
6	Electronic mail addresses	電子郵件地址
7	Social security numbers	身分證字號
8	Medical record numbers	病歷號碼
9	Health plan beneficiary numbers	健保卡號碼
10	Account numbers	帳號（住院帳號、銀行帳號等）
11	Certificate/license numbers	證書/執照號碼
12	Vehicle identifiers and serial numbers, including license plate numbers	車輛識別碼和序列號，包括車牌號碼
13	Device identifiers and serial numbers	醫材識別號和序列號
14	Web Universal Resource Locators (URLs)	網址（URL）
15	Internet Protocol (IP) address numbers	網路（IP）位址
16	Biometric identifiers, including finger and voice prints	生物識別標識，包括指紋和聲紋
17	Full face photographic images and any comparable images	全臉攝影圖像和任何類似的圖像
18	Any other unique identifying number, characteristic, or code, except as permitted by paragraph (c) of this section	除了前述第 3 項所允許的外，任何其他具有唯一識別號，特徵或代碼

附表二

參考衛福部對於資料國際傳輸或生物檢體衍生物輸出之風險程度分類

種類(範例)	型態	風險程度
全基因體定序之序列資料(Whole Genome DNA Sequence)	資料	高
經全基因體定序所得之全基因體基因型定型資料(Whole Genome Genotypes form Whole Genome Sequencing)	資料	高
經微晶片陣列所得之全基因體基因型定型資料(Whole Genome SNP Genotyping)	資料	高
親子鑑定或身份鑑定相關之基因定型圖譜資料	資料	高
部份基因體定序之序列資料(例如：target sequencing, CHIP-seq.)或基因型定型資料	資料	中
轉譯體定序之序列資料(例如：RNA-Seq,miRNA Seq)	資料	中

種類(範例)	型態	風險程度
外顯子全基因體定序之序列資料(Exome sequencing)	資料	中
經基因晶片分析後所產生之基因圖譜資訊(例如：copy number variations)	資料	低
原始參與者的相關臨床或流行病學資訊	資料	低
原始參與者的相關統計資訊	資料	低
蛋白質體實驗資料(proteomics)	資料	低
代謝體實驗資料(metabolomics)	資料	低
基因表現微陣列晶片(gene expression microarray)	資料	低
染色體組型實驗(karyotyping)	資料	低
經統計後的全基因體定序資料	資料	低
經統計後之全基因體定型資料	資料	低
經統計後之轉譯體分析資料	資料	低