

第二型糖尿病患者骨密度校正模型： 提升骨質疏鬆與骨折風險評估準確性

組員：方奕程、張日豪、侯言蓁、邱宥然、陳宥蓁、李品瑤

指導教授：朱是鍇

研究結論與未來發展

研究透過 **TAIWAN BIOBANK** 資料分析發現，第二型糖尿病患者的 **DXA** 骨密度測量值偏高，量測方法存在技術性偏誤而不能實際反映骨質健康狀況。為此，研究開發兩種骨密度校正的數學模型，結果顯示，糖尿病患者於脊椎與髕關節三部位的骨密度數值整體呈現下調趨勢，其中脊椎部位被高估約 **0.2** 至 **0.3**，顯示該部位的偏差最為明顯。

接續任務將申請健保資料庫 (**NHIRD**) 進行縱向追蹤研究，進一步確認校正後骨密度與實際骨折風險的關聯性，以驗證校正方法於臨床風險評估的準確性。

研究動機與目的

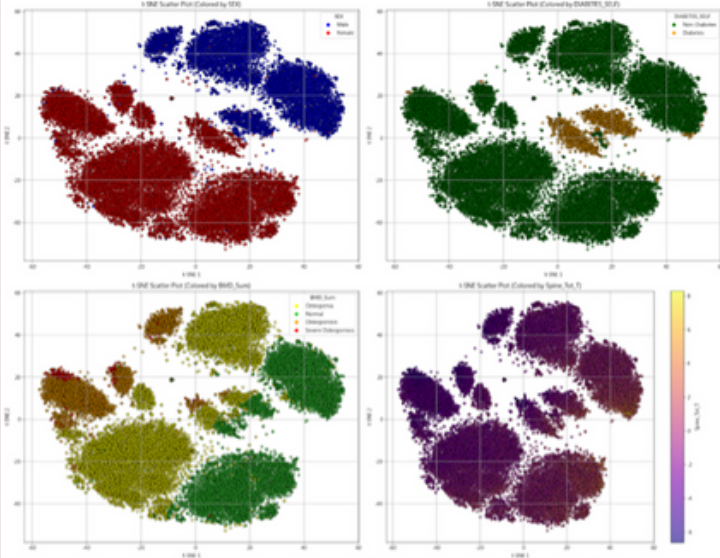
一般情況下，骨密度數值越低代表骨質流失嚴重，骨折風險越高；然而研究指出在第二型糖尿病患者身上，儘管其骨密度數值常與非糖尿病者相當甚至更高，實際骨折風險卻更高，形成所謂「骨質矛盾」的悖論現象。此現象代表現行最主要的骨密度量測技術(**DXA**)與評估指標可能低估糖尿病患者的骨折風險與骨質疏鬆程度。

本研究旨在分析骨密度數值的偏誤來源，進一步建立適用於臺灣第二型糖尿病患者的骨質狀態正確評估方法。

探索性分析

t-SNE(視覺化高維度資料與維度縮減)

研究首先應用t-SNE維度縮減技術揭示高維資料中潛在的非線性群聚結構，資料經過降維後在二維空間的分布呈現數個清晰的群體結構，整體樣本形成類似花瓣狀的分布；進一步根據性別、糖尿病狀態與骨質狀態的變數分別進行上色，初步掌握樣本數據的分布狀況。



- 左上：性別
- 右上：糖尿病狀態
- 左下：骨質狀況
- 右下：脊椎骨密度 T 分數

糖尿病患者的特徵有較強的相似性，非糖尿病患者則有更多樣化的特徵模式；骨密度 T 分數於某些群體存在向外擴散梯度形式，推論骨質狀態可能與特定的生理特徵高度相關。

資料來源與處理

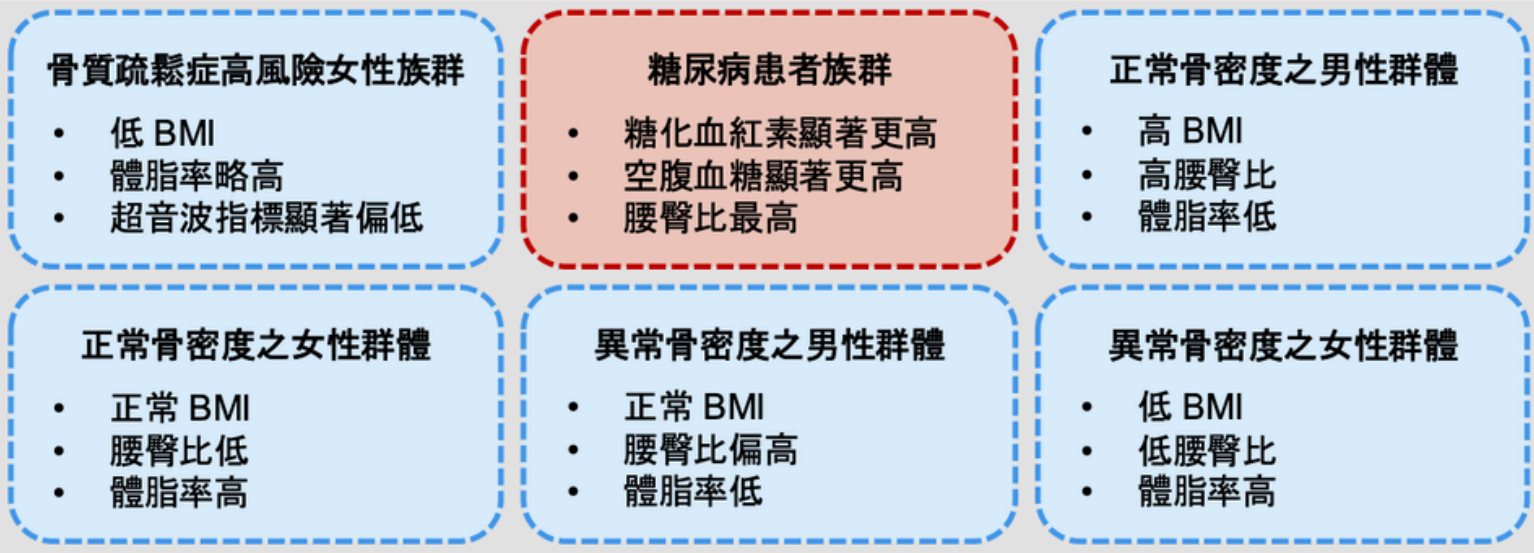
研究資料來自台灣人體生物資料庫(**Taiwan Biobank**)，共納入**24,433**筆個體資料，其中包含**1,837**位第二型糖尿病患者，並擷取**94**項相關變數，如人口學資料、骨質量測數據及糖尿病相關變項等。

資料處理首先針對類別變數重編碼與非結構化欄位的分割與萃取，檢視缺失值與極端值存在；其中缺失值以**KNN**插補法補全，以提升分析的完整性與準確性。

K-means 集群分析(數據差異結構與群集特徵)

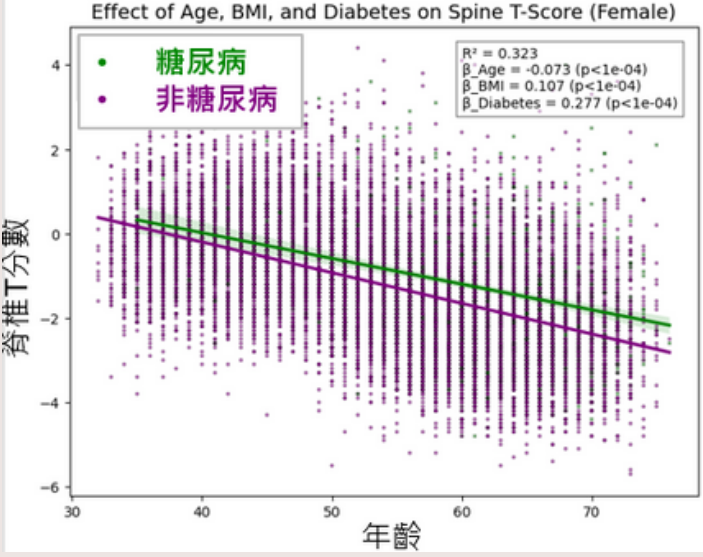
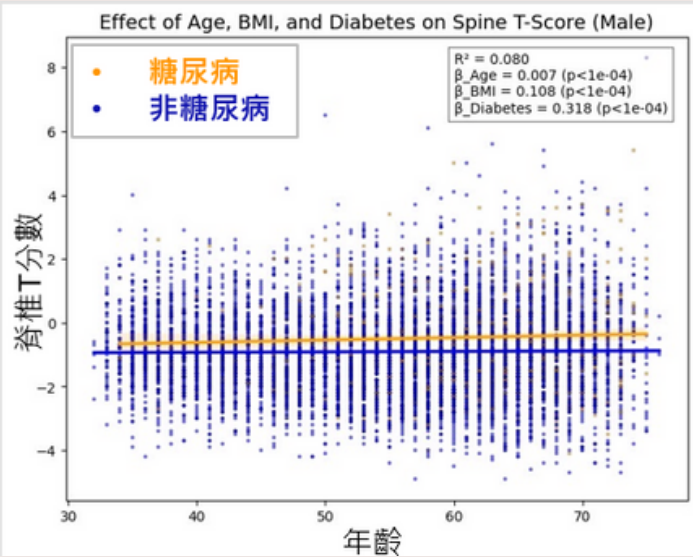
研究透過非階層式**K-means**集群分析法，探索樣本間的潛在結構模式，並系統性闡釋群集差異的關鍵決定變數。

分群結果顯示，糖尿病患者在糖化血紅素、空腹血糖與腰臀比上明顯高於其他群體，並呈現骨密度正常的特殊現象。從分析結果也觀察到肌肉與脂肪的比例及分布模式相較單純 **BMI** 數值具有更佳區別能力



假設驗證

假設驗證階段旨在確認資料存在糖尿病患之骨質矛盾現象。研究首先以多元線性回歸模型比較**DXA**與超音波兩種方法所量測之骨密度數值，結果顯示糖尿病患的**DXA**骨密度數值不分性別皆顯著較高；超音波骨密度則是只有男性顯著較高；研究進一步以卡方檢定與邏輯斯回歸比較糖尿病相關變數在不同骨質疏鬆評斷指標下呈現的差異，結果顯示只有在**DXA**指標下，糖尿病與骨質疏鬆呈現負向趨勢；另外將性別與年齡進行分層分析，結果同樣顯示只有在**DXA**指標下，糖尿病的骨質疏鬆人數呈現減少趨勢。綜合假設驗證分析結果，判斷糖尿病患者的骨密度數值存在**DXA**量測方法上的技術性偏誤 → 後續研究以建立**數值校正模型**為目標。



校正模型成果

以線性回歸模型校正：

總和 T 分數差距統計	脊椎	左髕關節	右髕關節
平均差 (DXA原始值 - 預測值)	0.235	0.093	0.12
標準差	1.252	0.757	0.744
DXA原始值大於預測值比例	54.22%	54.11%	55.69%

線性回歸校正應用傾向分數匹配(**PSM**)將糖尿病患與非病患族群進行配對，透過捕捉非糖尿病患者族群的資料形式構建回歸模型，將患者的脊椎與左右髕骨部位之骨密度 T 分數進行校正。此方法優勢在於線性回歸模型有較高可解釋性。比較校正前後數值，結果顯示在所有部位中，校正後的 T 分數平均皆低於原始 **DXA** 測量值，顯示 **DXA** 可能存在高估情形。其中以脊椎部位平均差最大，左右髕關節雖差距較小，且校正後數值波動較小、範圍集中。

以多重插補模型校正：

總和 T 分數差距統計	脊椎	左髕關節	右髕關節
平均差 (DXA原始值 - 預測值)	0.291	0.061	0.087
標準差	1.807	1.607	1.036
DXA原始值大於預測值比例	56.23%	50.95%	52.97%

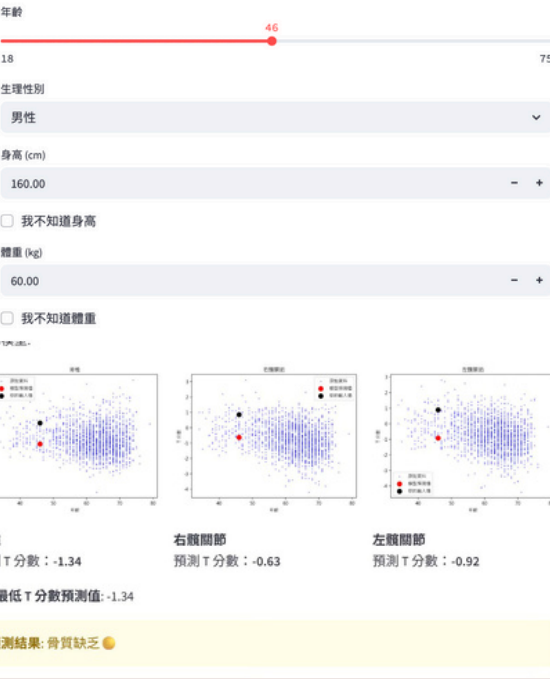
多重插補校正透過將糖尿病患的骨密度 T 分數視為缺失值，應用迭代填補器(**ITERATIVE IMPUTER**)結合貝葉斯嶺回歸(**BAYESIAN RIDGE REGRESSION**)作為核心演算法，在迭代過程中估計變數間的條件概率分布，為糖尿病患生成多組合理的骨密度估計值。此方法優勢在於能處理變數間的高度共線性，同時保留更多原始數據的分布特性，避免單一決定性模型可能帶來的系統性偏誤。比較校正前後數值，結果同樣顯示所有部位校正後的 T 分數平均而言均較 **DXA** 量測結果低，其中以脊椎部位平均差最大，波動程度相較回歸方法的校正數值而言更大；左右髕骨部位校正前後數值平均差極小，顯示糖尿病對髕骨部位的影響相對有限；與線性回歸校正數值比較，左右髕骨的校正前後之數值差距整體而言更小。

網頁工具開發

第二型糖尿病患者 DXA 骨質密度校正

本工具利用臨床資料預測第二型糖尿病患者在脊椎、右髕關節和左髕關節的 T 分數，協助使用者初步評估骨質健康狀態，並提供圖表與結論參考。

- 輸入年齡、性別、身高、體重、體脂率與超音波 T 分數，若不確定，可勾選「我不知道」將使用參考群體平均值。
- 可選擇輸入各部位已知 DXA T 值，與模型預測值比較。



研究針對所建立之數學校正模型開發網頁應用工具，協助第二型糖尿病患者校正其**DXA** 骨密度數值。

使用者輸入基本資料(如年齡、性別、身高、體重、體脂率與超音波數值)後，模型將透過線性回歸與多重插補兩種方式校正其脊椎、左髕與右髕的 **DXA** 測量 T 分數，並根據數值給出骨質狀態的結論。工具會另外透過圖表呈現校正前後的數值差距供使用者進行比較。