

長庚醫院的疾病診斷 智能編碼系統

◎林口長庚醫療人工智能核心實驗室碩士級研究助理 紀欽益



專長 | 計算神經科學、演算法、
資料前處理

● 國際疾病分類標準

「國際疾病分類標準（ICD）」是世界衛生組織依疾病的解剖部位及病癥等條件，加入編碼分類原則的疾病編碼標準，主要用來蒐集和比較各個國家的疾病資料、提供醫療服務量及治療參考指引。雖然台灣沒有加入世界衛生組織，但也導入此標準編碼，並參考國內本土病情增加部分代碼以提供更完整的醫療照護評值。現行的 ICD-10 診斷編碼有將近 7 萬個代碼，光是常見的痛風就有 100 多個相關代碼，可見此編碼標準的龐雜與精細。

● ICD 的重要性

國際上，ICD 是各國衛生組織數據溝通的語言。在國內，是民眾投保時的重點查核項目、是醫院向衛福部申報健保給付的依據、是醫療品質監控與改善的指標，醫療與公衛相關的決策與統計都依賴著 ICD。

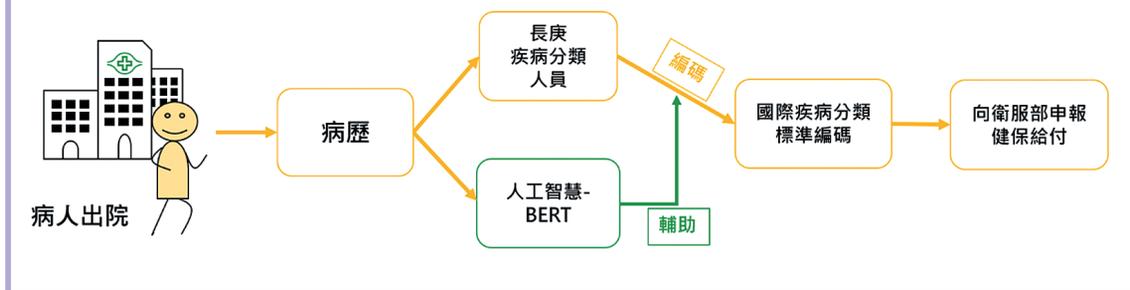
和大眾最有關係的就是保險了。在民眾常申請的健康保險中，保險公司會依據民眾提供的病歷、診斷書上的診斷 / 手術敘述，轉譯為疾病 / 處置分類檢索代碼，評估是否可核保。同樣的在理賠端，保險公司會根據被保險人提供的診斷書其內容來轉譯為疾病分類標準，決定理賠的金額。

● 現行的編碼流程

疾病編碼對醫院來說相當重要，醫院這邊是怎麼做編碼的呢？

首先是門診和急診，醫師在寫診斷時會依據病人病癥從該科的常用編碼中，帶入疾病診斷編碼或自行檢索適合之代碼。而在住院個案上，因病人情況

長庚體系—疾病診斷編碼流程



比較複雜，可能動了手術、出現併發症等多重疾病，以林口院區為例：病人出院時會有的診斷代碼在 5 至 10 個的比率最高。這樣多代碼、跨科別的編碼作業，需要受過訓練（例：住院診斷關聯群 DRG）並取得相關證照（例：疾病分類師、高階疾病分類師）的疾病分類專業人員進行。疾病分類人員逐筆調出出院病歷，以出院病歷摘要、病史、手術紀錄等內容為核心閱讀整本病歷紀錄，再依紀錄完整將疾病診斷與處置編碼轉譯出來。平均每筆資料需耗時 22 分鐘。林口院區平均每月約有一萬一千筆的住院資料，閱讀病歷與轉譯編碼工作量相當繁重，雖然很多資料都已資訊化，但仍有編碼速度與確保編碼正確性的問題待解決。

● 疾病診斷智能編碼系統

為了減輕疾病分類人員的工作負擔，病歷課疾病分類人員、醫療人工智能核心實驗室與長庚醫學科技從 2020 年開始合作開發了疾病分類智能編碼系

統，運用人工智能的自然語言處理技術—BERT，讓人工智能學習疾病分類人員編碼輔助編碼速度與正確性。當病人出院時，完成的病歷內容就會送入人工智能中分析、推論，再將最有可能的疾病分類編碼植入個案編碼輸入畫面。導入人工智能後監控診斷代碼，發現 5 個代碼以下的正確率可達 8 成，初期導入期花費非常多的心力討論與程式修改。目前已全面導入長庚體系疾病分類編碼系統，輔助疾病分類人員較快速的選擇編碼並進而降低編碼時間，也已在初階的疾病分類人員中獲得不錯的迴響。

● 未來規劃

未來除了要致力於正確率的繼續提升，我們也預計將智能編碼系統從出院病歷端逐步推入住院端，減輕醫護人員學習編碼的負擔。在其他應用上，我們也與保險公司合作，從客戶提供的病歷和診斷內容轉換成疾病分類編碼，作為輔助核保和理賠判斷的工具，加速投保與理賠的流程。☞