

以多類別關聯規則探勘方法預測疾病分類

巫俊卿
中興大學
資訊科學與工程學系
wjc@vghtc.gov.tw

廖宜恩
中興大學
資訊科學與工程學系
ieliao@nchu.edu.tw

高國峰
修平技術學院
資訊網路技術系
kfkao@mail.hit.edu.tw

摘要

台灣目前的健保給付制度，針對每次住院診療行為，設定了多種不同的給付標準。其中，病患的疾病分類編碼，是決定給付標準的重要因素。然而，由於部分醫療人員對衛生署所認可的國際疾病分類編碼系統並不熟悉，因此建置一個預測疾病分類編碼的系統來協助醫療人員，便成為一件重要的事情。本研究運用資料探勘技術，將病患疾病分類編碼問題，轉化為一個關聯式規則的問題；我們將每一次診療行為中涉及的藥品、手術、檢驗檢查以及疾病分類編碼，視為一筆交易資料；並限制所挖掘出規則的右手邊項目，只能是疾病分類編碼；同時，我們也提出了一個規則重要度的評分辦法，做為評選關聯規則的依據。該模型設置完成後，我們以中部某醫療中心2010年七萬多筆實際的診療紀錄，做為訓練資料，建構出一套全新的疾病分類系統。在測試當中，本系統只對手術類疾病分類編碼的召回率為84.2%、對整體疾病分類編碼的召回率為64.2%，均較許多現行之系統為佳。本系統的建構，預期將對醫療院所的管理，有相當的助益。

關鍵詞：疾病分類、多類別關聯規則探勘、多類別標籤分類、DRG

1. 前言

台灣目前的健保給付制度，針對每次住院診療行為，設定了多種不同的給付標準。目前施行的是自2010年1月1日起開始實施住院診斷關聯群(DRGs; Diagnosis Related Groups)支付制度。DRGs支付制度在給付醫療費用時，考量住院病患的疾病分類、手術項目、治療處置、年齡、性別、有無合併或併發症及出院狀等因素，認定

在這些因素相似的情況下，治療方式與醫療資源使用之情形也會雷同。該制度據此將住院病患區分為不同的診斷群組，再依每個診斷群組資源耗用度的不同設定相對權值(RW, Relative Weight)作為醫療給付標準，此種給付方式又稱「包裹式給付」。台灣目前實施的DRGs簡稱為TW-DRGs除特殊個案外，將病患分為967種診斷群組。例如說，

TW-DRGs代碼306為攝護腺切除術有合併症或併發症，RW值1.7503、標準住院天數9天；代碼307為攝護腺切除術無合併症或併發症，RW值1.1213、標準住院天數5天。以上例而言，一但合併症或併發症疾病分類編碼錯誤，就會造成給付金額之間高達56%的差異。因此正確的疾病分類，對醫療院所的正常運作，是非常關鍵的。

目前國內使用之疾病分類系統為國際疾病分類第9臨床修訂版(International Classification of Diseases, 9th Revision, Clinical Modification, ICD-9-CM)以下將ICD-9-CM簡稱為ICD9CM。該系統是美國國家衛生統計中心(National Center for Health Statistics)於1977年修訂1975年世界衛生組織發表的ICD-9疾病分類系統，特別強調臨床(Clinical)的適用性，ICD9CM編碼可區分為二類，分別為疾病(diseases)分類代碼以下簡稱**診斷類 ICD9CM**共13,677種，以及處置手術(procedures)分類代碼以下簡稱**手術類 ICD9CM**共3,768種[1]。

在病人住院期間，醫院為了管理目的必須於病患入院起，盡可能即早得知病患所屬TW-DRGs分類。為了要預測DRGs落點，醫師必須於病患住院期間進行

ICD9CM 編碼，但醫師並非專業疾病分類人員，會因為醫療用語與疾病分類用語的不同，造成醫師與專業疾病分類人員所編列之 ICD9CM 編碼相互之間有不一致的情況，造成 TW-DRGs 分類歸屬之不同與管理上的誤差。就疾病分類人員與醫師而言，如何以資訊系統協助疾病分類作業之進行，增加 ICD9CM 編碼的及時性與準確性，在 TW-DRGs 支付制度下是一件重要的議題。

病歷為疾病分類唯一且重要之資料來源，疾病分類人員從事疾病分類工作前，必須閱讀全本病歷，病歷內容中與疾病分類編碼相關部份有 10 項。其資料項及資料型態依發生的順序如下所列：1.急診病歷(文字、ICD9CM 編碼)、2.醫令(醫令代碼)、3.護理記錄(文字)、4.住院摘要(文字)、5.病程記錄(文字)、6.開刀記錄(文字)、7.手術前後之麻醉記錄(文字)、8.病理報告(文字)、9.檢查驗報告(文字、數值)、10.出院摘要(文字)。若想建立自動化疾病分類編碼系統，針對病歷中文字內容部份，須先使用語義處理技術來進行前置處理，其過程較為複雜，本研究捨棄這些模糊的文字說明，只採用藥品、手術代碼...等醫令以及疾病分類編碼等精確的資料做為我們的原始資料，並導入資料探勘的技術，以建構一個準確的疾病分類系統。

大部分的分類問題為單一分類問題，也就是訓練資料只和類別標籤集合 L 中某一個類別標籤值 λ 有關。對實際應用面而言通常訓練資料會與類別標籤集合 L 中一個以上的檢籤值 Y 有關， Y 為 L 的子集合，這類資料稱為多類別標籤資料(multi-label data)[6]。就本研究來說，住院個案為訓練資料而每個住院個案與 1 至 10 個不等的 ICD9CM 編碼有關，可視為多類別標籤資料探勘問題。

多類別標籤分類的運作方式，可分為建立多類別標籤分類器(MLC, multi-label classification)與建立類別標籤排名(LR, label ranking)兩種。多類別標籤分類器在處理查詢案件時，其輸出資訊是將類別標籤以是否與查詢個案相關為標準，進行分割

後的二個集子集合；類別標籤排名在處理查詢案件時，其輸出資料是將與查詢案件相關的標籤以分數高低排序表列；若同時處理 MLC 與 LR 的作業稱為多類別標籤排名(MLR, multi-label ranking)。本研究提出的做法，屬於 MLR。我們以挖掘關聯式規則為基礎，經過規則評分，輾轉化為分類問題。我們的系統會先將每一次診療行為中涉及的藥品、手術以及疾病類別，視為一筆的交易資料；並限制所挖掘出規則的右手邊項目，只能是疾病類別；在挖掘出所有可用規則後，所有規則需由一個重要度評分公式計算重要度，以做為最後分類的依據。

此外，本系統尚具有可以處理低支持度資料的特色。分析 74,356 個案發現有 60% 幾種 ICD9CM 發生頻次在 10 人次以下，也就是說有 60% 幾種疾病或手術處置只有 10 個以下病人罹患或執行。必須設計出一個可用適用於低支持度資料集的類別關聯規則採勘演算法，才能從住院病患之醫令與 ICD9CM 資料中探勘出疾病分類關聯規則集，本研究將此演算法稱為 CARFLS,(Class Association Rules mining For dataset with Low Support count,CARFLS)。

使用疾病分類關聯規則集建立疾病分類編碼推薦與檢核系統，預期將可獲得以下的效益：

- 1.降低醫師與疾病分類人員之間診斷編碼不一致率，提高病患於住院期間 TW-DRGs 落點預測的正確性，以利進行治療規劃與醫療資源使用管理。

- 2.避免應編列之 ICD9CM 編碼未予編列，漏列 ICD9CM 編碼之 Undercoding 發生，造成醫院損失。

- 3.避免不應編列之 ICD9CM 編碼而予編列，多列 ICD9CM 編碼之 Upcoding 發生，造成溢領給付。

本論文於第二節介紹相關研究、第三節說明 CARFLS 演算法、第四節說明疾病分類推薦與檢核系統與實作、第五節為實驗結果與討論、第六節為結論與未來研究

2. 相關研究

自動化或半自動化疾病分類編碼研究可分為二大類，第一類是以電子病歷文字內容為基礎結合自然語言處理技術(NLP, Natural Language Processing)與機器學習技術推論可能的疾病分類編碼，第二類是以藥品或手術醫令為基礎以資料探勘技術推論可能的疾病分類編碼。

第一類研究大都侷限在以電子病歷中部分內容為資料進行編碼作業、編碼範圍限定在某一類疾病上或只可編列單一診斷編碼。2007年由 Computational Medicine Center 舉辦的競賽其實驗資料為放射檢查報告(Radiology report)包括 978 份訓練資料與 976 份測試資料，分別由三家公司以人工方式編列 ICD-9-CM 編碼，編列的 ICD-9-CM 有 45 種(如 780.6)、94 種不同組合(如 780.6+786.2)，相關的研究論文相當完整[4]，由於編碼種類太少與實際疾病分類作業相距太遠，其實驗結果無法與本研究進行比較。除此之外有二篇與本研究規模較為相近的研究，分別說明如下：

Larkey 等[7]以哈佛大學 Brigham and Women's Hospital 的 11,599 筆出院病歷摘要(Discharge Summary)其中包括 10,902 筆訓練資料、187 筆測試資料、510 筆校準資料，每筆出院病歷摘要包含之字數從 100 至 3000 個字不等，可編列之 ICD-9-CM 個數從 0 到 15 個不等，平均每個案有 4.43 個 ICD-9-CM，去除 ICD-9-CM 出現次數未超過 6 次的疾病編碼後，剩下 157 筆測試資料，ICD-9-CM 種類也由 3,261 種降為 1,068 種，作者結合 k-鄰居分類器(k-nearest-neighbor classifier)、範例查詢(relevance feedback)、簡易貝氏分類器(Bayesian independence classifiers)三種方法來預測這 157 筆病歷的 ICD-9-CM 疾病分類代碼，其評估標準包括【平均精準度】為預測 ICD-9-CM 時預測個數排名在 rank0-10-20-30-40-50-60-70-80- 90-100 這 11 個點精準度的平均值 57%、【主診斷 Top1】為主診斷出現在預測診斷排名第 1 名的比率 46.5%、【主診斷 Top10】為主診斷出現在預測診斷前 10 名的比率 91.1%、

【召回率 Top15】預測 15 個 ICD-9-CM 時預測正確的 ICD-9-CM 數佔全部被預測 ICD-9-CM 之比率 73.2%、【召回率 Top20】為預測 20 個診斷之召回率 77.6%。

Patrick Ruch[9]等，視疾病分類問題為一個對多文件、多類別、多標籤的自動化文件分類作業，並以在 University Hospitals of Geneva 就醫病患中，以採 DRGs 方式給付醫療費的病患為研究對象，探討如何以電腦輔助編碼改善 DRGs 案件的疾病分類編碼作業。作者針對 2004-2006 年間共 78,756 份 DRGs 個案的過去病史(anamnesis)、藥品處方(prescription)、診斷記錄(Diagnostics)等三份文件與平均每個案 5.1 個 French ICD-10 疾病編碼(10,511 種疾病編碼、733,484 個疾病編碼)作為實驗資料集。再由實驗資料集中有過去病史的 18,655 個案(23%)產出{過去病史,ICD-10}組合資料、有藥品處方的 61,372 個案(78%)產出{藥品處方,ICD-10}組合資料及有診斷記錄的 78,756 個案(100%)產生{診斷記錄, ICD 碼}組合資料作為訓練資料，針對三種不同的文件訓練 3 個鄰居法分類器，再以 ICD-10 說明文件為訓練資料出一個 Data-poor 分類器，最後以 800 份抽樣資料為測試資料，以 Fox E.A.所提方法，整合 4 種分類器之預測結果對測試資料進行疾病編碼預測，再以每個案排名前 20 名的疾病編碼為測試個案之預測值，與實際之疾病編碼作比對，計算出召回率(Recall)以 R20 表示，再以排名第 1 名的預測值與所有編碼作比較計算出精準度(Precision)以 P0 表示，其實驗之最佳結果 R20 為 63%、P0 為 75%。作者提出不以傳統的召回率、精準度、F1 值為衡量標準之理由，對疾病分類人員而言，以其專業之判斷，可輕易的從 20 個 ICD-10 疾病編碼中，挑選出正確的疾病編碼，所以系統效能以召回率作為主要考量不需顧及精準度。由於不同的疾病分類人員對相同的個案，在疾病編碼上會有不同的看法，產生的不同編碼結果，作者依 Fabrizio Sebastiani.[5]之理論計算出召回率的理想上限值為 79%，R20 為 63%尚有改善的空間。另 P0 可以用來評估採用全

自動的方式來預測疾病編碼時的效能。整體而言在 P0 為 75%、R20 為 63% 表示預測值中排名第一的編碼，每 4 個就有 3 個是正確、20 個預測值中有 3 個編碼是正確的水準下，對改善 DRGs 疾病編碼作業有相當大助益。

第二類方法中最經典研究成果為中央健康保險局於 2010/08 公告之【全民健康保險手術、處置支付標準與 ICD-9-CM 手術及處置代碼對應檔】[3] 共 6,369 筆資料，每個健保醫令可對應 1-4 個 ICD9CM 的多種組合。例如，有申報健保碼 33033B(內視鏡逆行胰液引流術)可編列 52.13(經內視鏡逆行性胰臟攝影)+52.97(經內視鏡鼻胰引流管植入術)或 52.93(經內視鏡胰管支架導管植入術)如表 1。醫療院所可以利用這些規則，使用申報之醫令碼與 ICD9CM 資料進行手術類 ICD9CM 編碼推薦與檢核作業。

表 1 健保碼與 ICD-9-CM 對應檔

健保碼	手術處置碼 1	手術處置碼 2
33033B	52.13	52.97
33033B	52.93	
33023B	87.53	
33024B	51.10	
33024B	51.11	
33024B	52.13	

陳鎮揮等發表以資料探勘技術建立 ICD-9-CM 編碼決策系統之研究[2]，以 2002-2005 年間某區域醫院 58,507 個住院病患之臨時醫囑、長期醫囑、出院帶藥及疾病分類員所作的 ICD9CM 編碼為訓練資料，依醫囑執行順序以時間序列的概念統計醫囑與疾病分類代碼出現的機率值，再以這些資訊推論出院案件的疾病分類代碼，由 2006 年病患中以抽樣方式取出 395 個個案作為測試資料，實驗結果正確率為 72.8%，其成果為平均每個案所需之推論時間為 2.5 分鐘，所以不適用於線上即時作業，建議以批次作業進行推論。

3. 低支持度資料集類別關聯規則探演算法-CARFLS

本研究之目的是要利用資料探勘方法從住院病患的【醫師醫令】、【健保醫令】、【ICD9CM】歷史資料中探勘出【疾病分類關聯規則集】，再利用這個規則庫建立疾病分類推薦與檢核系統，提供疾病分類人員及醫師使用，如圖 1。

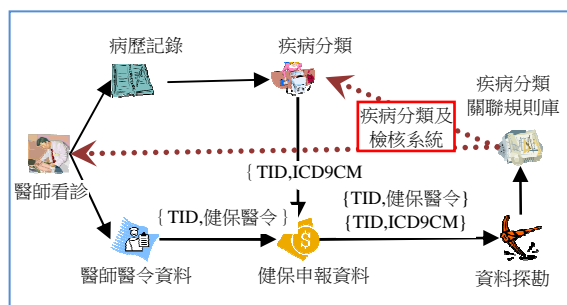


圖 1 疾病分類推薦與檢核系統

表 2 診斷類 ICD-9-CM 頻次分佈

項目	診斷類 ICD-9-CM		
	種類	佔率	累計佔率
1 次	1095 [種]	23.4%	23.4%
2 次	519 [種]	11.1%	34.5%
3-10 次	1379 [種]	29.5%	64.0%
10 以上	1689 [種]	36.0%	100%
合計	4682 [種]	100%	

表 3 手術類 ICD-9-CM 頻次分佈

項目	手術類 ICD-9-CM		
	種類	佔率	累計佔率
1 次	376 [種]	22.5%	22.5%
2 次	182 [種]	10.91%	33.4%
3-10 次	460 [種]	27.5%	60.9%
10 以上	654 [種]	39.1%	100%
合計	1672 [種]	100%	

分析台灣中部某醫學中心 2009 年 1 月至 2010 年 6 月 74,356 住院人次之診斷類 ICD9CM 與手術類 ICD9CM 發生頻次，發現 4,682 種診斷類 ICD9CM 在 74,356 個案中有 1,095 種代碼只有 1 個病人、519 種只有 2 個病人、共有 2,993 種代碼不到 10 個病人，佔全部種類的 64%，如表 2；1,672 種手術類 ICD9CM 在 74,356 個案中有 376 種代碼只有 1 個病人、182 種代碼有 2 個病

人、病人數在 10 人次以內的代碼佔全部種類的 61%，如表 3。

如果要使用類別關聯規則探勘演算法來進行疾病分類關聯規則探勘，不可能使用支持度(Support)作為資料探勘篩選條件之門檻指標，因為若以最小支持為門檻指標當門檻值設為 $0.000134=10/74356$ 時(通常最小支持度不會設定在這麼小的值)，則病患人次小於 10 的 ICD9CM 一定不會在類別關聯規則中出現，也就是使用這些類別關聯規則進行 ICD9CM 推薦時，會有 61% 種手術類及 64% 種診斷類 ICD9CM 不會出現在推薦名單中。因此，必須使用支持度以外指標作為資料探勘篩選條件之門檻指標，減少演算法的探勘範圍。

表 4 類別資料集(ICD-9-CM 資料)

Label	說明	父階層	父階層說明
200.08	多部位淋巴結之網狀細胞肉瘤	200.0	網狀細胞肉瘤
200.11	頭、臉及頸部淋巴結之淋巴肉瘤	200.1	淋巴肉瘤
200.12	胸腔內淋巴結之淋巴肉瘤	200.1	淋巴肉瘤
79.01	肱骨骨折閉鎖性復位未使用內固定器	79	骨折及脫臼復位術
79.03	腕骨及掌骨骨折閉鎖性復位未使用內固定器	79	骨折及脫臼復位術

ICD9CM 編碼採階層式的分類架構。診斷類 ICD9CM 為 3 個整數位及 1-2 小數位最多 5 碼之編碼，前 3 碼從 001 到 999 為類目碼(Category number)、小數部份再依疾病之病因、發病部位或疾病表徵詳細分類，共 995 種類目、13,677 個代碼；手術 ICD9CM 為 2 個整數位及 1-2 個小數位最多 4 碼編碼，前 2 碼 01-99 為類目碼，共 98 種類目、3,768 個代碼。本研究以**類別資料集(Lable)**表示 ICD9CM 其屬性包括 ICD9CM 代碼(Label)、說明、父階層代碼、父階層說明，其中診斷類 ICD9CM 取至小數第 1 位視為父階層、手術 ICD9CM 取前 2 碼視為父階層，如表 4 所示。

醫令是醫師在治療病人過程中，用來調整各各類醫事人員，提供病人所需醫療

服務的溝通管道。一般來說，醫師會依病情判斷病患所罹患疾病來擬定治療計劃，透過醫師處方醫令系統(Computerized Physician Order Entry, CPOE)” 啟動整個治療計劃，達到治療病患的目的。健保醫令是健保局支付醫療費用所製定，這與醫師醫令是以治療病患為目的不同，醫療資訊系統透過健保申報作業將醫師醫令轉換為健保醫令作為申報之用。就醫令與疾病分類關聯度來說，醫師醫令會比健保醫令之關聯度來的高，這點在本究的實驗中也會得到證實。

健保醫令為階層式架構，第一層將醫令區分為藥品處方、病房費、管灌膳食、檢查、放射線診療、治療處置、手術費、復健治療、血液透析、血液血漿、麻醉費、特殊材料、精神科治療，第二層再依類別的不同採不同的分類方式如檢查費可再區分為一般尿液檢查、特殊尿液檢查、糞便檢查、血液學檢查、一般生化學檢查、微量元素測定、血中藥物濃度測定、血中毒物濃度測定、輸血前檢查、免疫學檢查、細菌學與黴菌檢查、病毒學檢查…等；藥品處方依解剖治療化學分類系統(Anatomical Therapeutic Chemical Classification)ATC 分類碼分類；特殊材料類之健保碼長度為 12 碼，第 1 碼為大類碼、第 2 碼為小類碼、第 3-5 為品名碼、第 6-10 碼為規格碼、第 11-12 碼為廠牌代碼。

不論醫師醫令或健保醫令均以**項目資料集(Item)**來表示，屬性包括 Itemcode(醫令代碼)、說明、父階層代碼、父階層說明，其中特殊材料代碼前 10 碼相同時，可視為不同廠牌但作用相同的材料，所以取前 10 碼作為醫令代碼、前 5 碼相同視為相似品項(大類別+小類別+品名碼)以此作為父階層代碼；藥品處方若所屬 ATC 分類相同視為相同藥品，故以 ATC 碼取代原健保碼為醫令代碼、ATC 碼前 5 碼相同視為相似品項以此作為父階層代碼；其他類別依健保局支付標準表之第二階層分為父階層，如表 5。

表 5 項目資料集(醫令資料)

Itemcode	說明	父階層	父階層說明
CGPG1 0648N	冠狀動脈前置導 管-0648N	CGPG1	冠狀動脈前 置導管
CGPG1 0708N	冠狀動脈前置導 管-0708N	CGPG1	冠狀動脈前 置導管
CGPG1 0818N	冠狀動脈前置導 管-0818N	CGPG1	冠狀動脈前 置導管
C04 AD02	ATC-xantinol nicotinate	C04AD	Purine derivatives
C04 AD03	ATC-pentoxifylli ne	C04AD	Purine derivatives
14002B	腺病毒補體結合 抗體	L2-0049	病毒學檢查
14003B	巨大細胞病毒抗 體補體結合法	L2-0049	病毒學檢查
14004B	巨大細胞病毒抗 體酵素免疫法 E I A	L2-0049	病毒學檢查

住院病患 ICD9CM 編碼資料以類別交易資料集表示屬性包括 TID(交易代碼)、Labelcode(ICD9CM)如表 6；醫師醫令資料及健保醫令資料以項目交易資料表示屬性包括 TID(交易代碼)、Itemcode(醫令代碼)，如表 7。

表 6 類別交易資料集(ICD9CM)

TID	Labelcode(ICD9CM)
01852866	581.1
01852866	39.95
01852866	54.93
01852866	38.95
01852865	456.8
01852865	571.2

表 7 項目交易資料集(醫令)

TID	Itemcode(醫令代碼)
01852866	N05AX08
01852866	N04AA02
01852866	05213K
01852865	05213K
01852865	N05AF01

類別關聯規則為 Itemset(醫令組合)與 Label(ICD9CM) 之間的關聯，以 Itemset→Label 表示，例如藥品 A→79.01，屬性值包括：

(1)Confidence 可信度

$$Confidence = \frac{\| Itemset \cap Label \|}{\| Itemset \|}$$

為 Itemem 與 Label 同時出現的病人數佔 Itemset 的病人數，也就是使用某組 Itemset(醫令組)的病人被分類為 Label (ICD9CM)的機率

(2)Recall 召回率

$$Recall = \frac{\| Itemset \cap Label \|}{\| Label \|}$$

為 Itemem 與 Label 同時出現的病人數佔編碼為 Label 的病人數之比率。Recall 函義為使用 Itemset→Label 這個規則可以召回所有標記為 Label 個案之比率。在本研領域另有一個函義為所有標記為某個 ICD9CM 的病人使用到某組醫令組的比率。當 Itemset→Label 之 Recall 值很低表示醫令組被使用到的機率很低，不是病患被標記為 Label 這個類別的主要特徵，這類規則就算 Confidence 值很高，通常 Itemset 與 Label 之間的相關性並不高。

(3)F1-Measure 簡稱 F1

$$F1 = \frac{2 * Confidence * Recall}{Confidence + Recall}$$

為綜合【若有使用 Itemset 則分類為 Label 這個類別的 Confidence 值】與【若分類為 Label 這個類別則有使用 Itemset 的 Recall 值】之綜合性指標，可取代支持度作為 CARFLS 演算法篩選條件之門檻指標，可篩選掉 Confidence 符合標準、Recall 值太低的規則。

(4)OptimalF1

$$OptiamlF1 = \frac{2 * 100\% * Recall}{100\% + Recall}$$

定理：

對關聯規則 Itemset→Label 而言所有包含 Itemset，型式為 Itemset ∪ X→Label 的關聯規則其 F1 值必小於 OptimalF1，即若 Itemset_x ⊂ Itemset_y 則

$$F1(Itemset_y, L_n) \leq OptimalF1(Itemset_x, L_n)$$

在此原則下若以 F1 作為門檻指標，只要 Itemset→Label 的 OptimalF1 小於 F1 門檻值

則所有的 $Itemset \cup X \rightarrow Label$ 也一定小於 F1 門檻值，一定會被排除在疾病分類關聯規則之外。

證明：

設 $Itemset_x \subset Itemset_y$ 則

1. $Recall(Itemset_y, Label)$

$$= \frac{\| Itemset_y \cup Label \|}{\| Label \|}$$

$$\leq \frac{\| Itemset_x \cup Label \|}{\| Label \|}$$

$$= Recall(Itemset_x, Label)$$

2. $Confidence(Itemset_y, Label) \leq 100\%$

又 F1-Measure 為 Confidence 與 Recall

的正函數，所以 $F1(Itemset_y, L_n)$ 的上限值

為 Confidence 的上限值 100%、Recall

的上限值 $Recall(Itemset_x, Label)$

$F1(Itemset_y, Label)$

$$\leq \frac{2 * 100\% * Recall(Itemset_x, Label)}{100\% + Recall(Itemset_x, Label)}$$

$$= OptimalF1(Itemset_x, Label)$$

故得證。

(5) Score 分數為規則重要性權重值

$$Score = (trunc(Confidence, 2)$$

$$- (trunc(Confidence, 2) \bmod 0.05) * 10000$$

$$+ trunc(F1, 2) * 100$$

例，若 $Confidence=0.5714$ 、 $F1=0.32123$

則為 $Score=5532$

當類別關聯規則中 $Itemset$ 之集合項數為 k 時稱這個規則為 $kItemRule$ ，例如 {藥品 A} \rightarrow 52.13 為 1ItemRule、{藥品 A, 藥品 B} \rightarrow 52.13 為 2ItemRule 依此類推，本研究依類別關聯規則之屬性值 $Itemset$ 之集合項數、Confidence、F1、OptimalF1 等屬性將規則分類為【疾病分類關聯規則】、【高品質疾病分類關聯規則】、【可特殊化類別關聯規則】。

【疾病分類關聯規則】為符合下列 2 個條件中任何 1 個條件的類別關聯規則。

若 $Itemset$ 之集合項數=1 時，即為 1ItemRule 時必需符合：

$$(Confidence \geq ConfidenceThreshold \text{ And } Max_F1 \geq oneItemF1Threshold),$$

其中 Max_F1 為 $F1(Item \text{ 父項目}, Label)$ 、 $F1(Item, Label \text{ 父項目})$ 、 $F1(Item, Label)$ 之最大值

若 $Itemset$ 之集合項數>1 時，時必需符合：

$$(Confidence \geq ConfidenceThreshold \text{ And } F1 \geq F1Threshold)$$

【高品質疾病分類關聯規則】為符合疾病分類關聯規則之條件且

$Confidence \geq QualityConfidence$ 之規則。

【可特殊化類別關聯規則】為符合條件 (OptimalF1 \geq F1Threshold and

$Confidence \geq RefinedConfidence$ and

$Confidence < QualityConfidence$) 之規則。

【疾病分類關聯規則】中有二種型態的規則是多餘的規則 (Redundant Rule)，這二種型態分稱為【特殊化無效多餘規則】與【召回個案相同多餘規則】。【特殊化無效多餘規則】在是指將可特殊化類別關聯規則進行特殊化作業後，特殊化後的分數沒有比原規則大，這個特殊化後的規則就是特殊化無效多餘規則，舉例來 {手術 A} \rightarrow 52.13 的分數為 1535，再加入一個藥品 A 特殊化後之規則為 {手術 A, 藥品 A} \rightarrow 52.13 而其分數亦為 1535，{手術 A, 藥品 A} \rightarrow 52.13 即為特殊化無效多餘規則。【召回個案相同多餘規則】是指當任二個規則 $A \rightarrow L_n$ 及 $B \rightarrow L_n$ 其召回 (Recall) 之個案相同時，二個規則中 Confidence 較小的那一規則就是召回個案相同多餘規則，例如 {手術 A} \rightarrow 52.13 召回之個案為 {X, Y, Z}、{手術 B} \rightarrow 52.13 之召回個案也是 {X, Y, Z}，這二個規則中 Confidence 較小的那個規則就有召回個案相同多餘規則。

低支持度資料集類別關聯規則勘探 (Class Association Rules mining For dataset with Low Support count, CARFLS) 演算法之探勘目的是要由住院病患之醫令與 ICD9CM 資料中探勘出不含【特殊化無效多餘規則】與【召回個案相同多餘規則】的【疾病分類關聯規則集】，CARFLS 演算法有 3 個步驟分別為：

1. 資料預處理，將資料轉換為適合 CARFLS

演算法的資料結構

2. 針對每個 Label 進行 1ItemRule 計算，依屬性值判斷是否將規則加入【疾病分類關聯規則集】與【可特殊化類別關聯規則集】

3. 對【可特殊化類別關聯規則集】內的規則以加入另 1 個 Item 的方式進行特殊化作業，計算特殊化後之規則，依屬性值判斷是否將特殊化規則加入【疾病分類關聯規則集】與【可特殊化類別關聯規則集】

*【疾病分類關聯規則集】是一個稱為 NonRedundant 的特殊物件，透過 Add 功能加入規則，可以保持規則集中不會有【召回個案相同多餘規則】存在其中。

1. 資料預處理:

輸入資料: {Itemcode,說明,Item 父階層碼}、
{Labelcode,說明,Label 父階層碼}、
{Tid,Itemcode}、{Tid,Labelcode}

轉化後資料結構:

- Item 父階層碼 → {Item 代碼} 對應表
- Label 父階層碼 → {Label 代碼} 對應表
- 在讀取交易資料同時，將 Tid、Itemcode、Labelcode 依讀入之先後次序轉為從 1 開始排序的 Tidkey、Itemkey、Labelkey
- Labelkey → Labelcode,Label#,Tidkey[1..Label#], Label 父階層碼,Tot_Recall 之對應表
- Itemkey → (Itemcode,Item#,Tidkey[1..Item#], Item 父階層碼) 之對應表
- BitmapItemTid[1..ItemKinds][1..TidKinds]: 記錄 {Tid,Item} 交易資料，BitmapItemTid[i][t]=true 表示 Tidkey 為 t 之交易有 Itemkey 為 i 之項目

例如:

Item 父階層碼	Label 父階層碼
PI1 {I1,I2,I3}	PL1 {L1,L2}
PI2 {I4,I5,I6}	PL2 {L3,L4}

Tidkey 對應表					
1	T1	3	T3	5	T5
2	T2	4	T4		

Labelkey 對應表					
1	L1,2,[1,3],PL1,0	3	L3,2,[1,5],PL2,0		
2	L2,1,[4],PL1,0	4	L4,3,[1,4,5],PL2,0		

Itemkey 對應表					
1	I1,2,[1,3],PI1	4	I4,2,[3,5],PI2		
2	I2,3,[2,3,4],PI1	5	I5,4,[1,2,4,5],PI2		
3	I3,3,[2,4],PI1	6	I6,3,[1,2,3],PI2		

BitmapItemTid						
Item Tid \	1	2	3	4	5	6
1	1				1	1
2		1	1		1	1
3	1	1		1		1
4		1	1		1	
5				1	1	

CARFLS

--採一次計算一個 Labelkey 策略減少記憶體空間

- 執行資料預處理
- 設疾病分類關聯規則集為 NonRedundant 物件
- 從 Labelkey 對應表中依序取出一個 Labelkey
- 將 Labelkey.Tidkey[1..label#]轉化為 BitmapLabelTid[1..Tidkinds]之結構
例 Tidkey [1,2,5]轉為[1,1,0,0,1,0,0,...]
- 執行 **Generate_OneItemRule(Labelkey)**
- For k=2 to Max_kRule(建議值為 4)
- 執行 **Generate_kItemRule(labelkey)**
- 輸出疾病分類關聯規則集

Generate_OneItemRule(Labelkey):

- 以 labelkey 產生一個 Itemkeyset={ } 的規則 R0
- for each Itemkey in Itemkey 對應表
- 執行 $R1 = \text{Cal_k+1Rule}(R0, \text{Itemkey})$ 計算出 Confidence、Recall、F1、OptimalF1、Max_F1
- if (R1 符合【疾病分類關聯規則】條件 $\text{Max_F1} > \text{oneItemF1Threshold}$ (建議值 0.01) and $\text{confidence} > \text{ConfidenceThreshold}$ (建議值 0.1)) then 執行疾病分類關聯規則集.Add(R1)
- if (R1 不是召回個案相同多餘規則) then 將 Labelkey.Tot_Recall 加上 Recall 值
- if (R1 符合【可特殊化類別關聯規則】條件 $\text{Optimal_F1} > \text{F1Threshold}$ (建議值 0.1) and $\text{Confidence} > \text{RefinedConfidence}$ (建議值 0.03) and $\text{Confidence} < \text{QualityConfidence}$ (建議值 0.2)) then 將 R1 加入【可特殊化類別關聯規則集】、
將 Itemkey 加入 oneItemkeyset

Generate_kItemRule(Labelkey)

--以【可特殊化類別關聯規則集】與 oneItemkeyset 計算 k+1ItemRule 類別關聯規則

- for each R0 in 【可特殊化類別關聯規則集】
- for each Itemkey in oneItemkeyset
- if (Labelkey.Tot_Recall > Max_Recall (建議值 200%)) then return
- $R1 = \text{Cal_k+1Rule}(R0, \text{Itemkey})$
- if (R1 符合【疾病分類關聯規則】條件 $\text{F1} > \text{F1Threshold}$ and $\text{Confidence} > \text{ConfidenceThreshold}$ and R1 不是【特殊化無效多餘規則】即 $R1.\text{Score} > R0.\text{Score}$) then

- 執行疾病分類關聯規則集.Add(R1)功能
06. if (R1 不是召回個案相同多餘規則) then
將 Labelkey.Tot_Recall 加上 Recall 值
 07. if (R1 符合【可特殊化類別關聯規則】條件
Optimal_F1>F1Threshold and
Confidence>RefinedConfidence and
Confidence<QualityConfidence) then
將 R1 加入【可特殊化類別關聯規則集】

Cal_k+1Rule(Rule R0,Itemkey)

- 01.R1=R0
- 02.add itemkey to R1.Itemkeyset
- 03.Cal_Rule_Cnt(R1)
04. R1.Confidence= R1.ItemLabel# / R1.Item#
05. R1.Recall= R1.ItemLabel# / R1.Label#
06. R1.F1 = F1(R1.Confidence, R1.Recall)
07. R1.OptimalF1 = F1(1,Recall)
08. R1.Score=(trunc(R1.Confidence,2)
-trunc(R1.Confidence,2)mod 0.5)*10000
+trunc(R1.F1,2)*100
- 09.return R1

Cal_Rule_Cnt(Rule R1)

--先從 Itemkeyset 中個案數 Item#最小的 Itemkey，再以此 Itemkey 相關個案之 Tidkey 利用 BitmapItemTid[][]與 BitmapLabelTid[]以 boolean 運算計算 r.Item#,r.ItemLabel#其最多需要 Itemkey.Item#*(Itemkeyset.size()+1)次運算，與交集運算相比可節省可觀的計算時間

01. k = R1.Itemkeyset.size()
02. Itemkey[0..k-1]=r.Itemset.toArray
- 03 從 Itemkey[0..k-1]中找出 Item#最小的 Itemkey0
04. for each Tidkey in Itemkey0.Tidkey[1..Item#]
06. cnt=true,j=0
07. while (cnt=true and j<k)
08. cnt = BitmapItemTid[Itemkey[j]][Tidkey]
09. if (cnt=true)
10. R1.Item#++
11. if (BitmapLabelTid[Tid])
12. R1.ItemLabel#++
13. add Tidkey to R1.ItemLabelTidkey
14. R1.ItemLabelTidkeySum += Tidkey

例如:設 R1 之 Itemkeyset=[1,4,7,9],Labelkey=1

- 1.Itemkeyset 中 Item#最小的為 Itemkey=7
- 2.Itemkey(7)之 Tidkey 為[1,4,10]
- 3.check Tidkey=1,Itemkey=1,4,7,9 之
BitmapItemTid[1,4,7,9][1]之值均為 1 時
cnt=true
If (cnt=true) then R1.Item#++
If (cnt=true and BitmapLabelTid[1]=1) 時
R1.ItemLabel#++
- 4.依序計算 Tidkey=4,Tidkey=10

BitmapItemTid											
Item Tid	1	2	3	4	5	6	7	8	9	10	L1=1
1	1			1			1		1		1
10				1			1	1		1	

NonRedundant 物件(HashMap NRRuleset)

Function Boolean Add(Rule r)

--以則規召回之個案數與個案 Tidkey 之總合為 key 建立結構如下之 HashMap:

HashMap NRRuleset	
Key	Data
"3"+","+"101"	Ruleset1={R1A,R2B,.. }
"4"+","+"123"	Ruleset2={R2A,R2B,.. }
....

Ruleset1 之 Rules 召回個案數=3,Tidkey 總合=101

Ruleset2 之 Rules 召回個案數=4,Tidkey 總合=123

--【召回個案相同多餘規則】之必要條件為召回個案相同，只有 Key 值相同的規則才會有多餘規則發生，有新加規則要加入時，可以先比對 Key 值，再一一比對相同 Key 值的規則，檢視是否有召回個案完全相同完全的規則，若有再比較彼此 Confidence 的大小決定保留那一個規則

- 01.key=r.ItemLabel#+','+r.ItemLabelTidSum
- 02.Get. m#ruleset from NRRuleset with key
03. if (m#ruleset is empty)
04. add r to m#ruleset
05. NRRuleset.Put(key,m#ruleset)
06. return(false)
- 07.for each rule in m#ruleset as r1
08. if (r.ItemLabelTidset = r1.ItemLabelTidSet)
09. if (r1.Item#<r.Item#)
10. remove r1 from m#ruleset
11. add r to m#ruleset
12. retrun(true)
- 13 add r to m#ruleset
- 14 return(false)

4. 疾病分類編碼推薦與檢核系統

就疾病分類編碼推薦系統來說，當醫師或疾病分類人員要求系統對某個住院病人進行 ICD9CM 編碼推薦時，系統利用目前醫師已開立的醫令從【疾病分類關聯規則集】中找出所有符合條件的規則，每個規則都會測預一個 ICD9CM 及分數，每種 ICD9CM 只保留一個分數最高的規則，依分數高低排序，將這些規則推薦給使用者，推薦內容包括{醫令、醫令說明}、ICD9CM、ICD9CM 說明、醫令個案數、

ICD9CM 個案數、醫令+ICD9CM 個案數、可信度、召回率、關聯度、分數等資料，目的是藉由這些資訊告知使用者為什麼會有這樣的推測，讓決策者可輕易決定要選擇那些 ICD9CM，也就說疾病分類編碼推薦系統不是以黑箱作業模式預測 ICD9CM 編碼，這也是本研究與其他研究不同之處。

綜合上述分析，本研究將疾病分類推薦系統歸類為多類別標籤排名 (MLR, multi-label ranking) 問題，先將所有 ICD9CM 區分為相關與不相關二組，再由相關組中以分數排序推薦給使用者。使用系統時先輸入出院日期找出該天出院之病患名單，再名單中選擇要進行編碼的病人，如圖 2，再依設定的診斷類 ICD9CM 最多推薦數(建議值 13)、手術類 ICD9CM 最多推薦數(建議值 7)進行推薦。本實作系統為了檢視測預的正確性，特別挑選疾病分類人員已編碼個案與預測進行比對，在推薦之 ICD9CM 與疾病分類人員相同的 ICD9CM 前以*標記，如圖 3. 整體效能評估於下一節中說明。

病號	病號姓名	醫師姓名	CaseNo	預診日期	出院日期	主診數	次要診數1	次要診數2	次要診數3	次要診數4	Q1	Q2	Q3	Q4	Q5
382	2010-09-16	2010-10-05	574.39	1001.1	402.50	1001.1	77.22	51.38							
214	2010-09-21	2010-10-05	820.21				79.35								
777	2010-09-26	2010-10-05	996.3	414.01	998.91	427.31	250.00	35.22	35.12	36.14	36.15	39.61	37.61	37.22	
899	2010-09-27	2010-10-05	998.59	511.9	912.8	401.9	041.7	34.64							
305	2010-09-28	2010-10-05	733.82	265.4	996.67	250.00	401.9	79.35	76.63						
300	2010-09-28	2010-10-05	532.80	531.90	250.00	102.62	743.81	45.25							
333	2010-09-29	2010-10-05	812.00	813.41	571.2		79.31	79.32							
372	2010-09-29	2010-10-05	738.4	734.02			81.88	81.06	80.31						
100	2010-09-29	2010-10-05	801.02	520.00	710.02	710.02	710.11	79.39							
375	2010-09-30	2010-10-05	715.36	401.9			81.54								

圖 2 病患名單

診斷	診斷	醫令	醫令說明	置信度	說明
1	*414.01	自體性冠狀動脈粥狀硬化	80006009	69/	95/ 可信度: 93%, 使用率: 3%
2	*98.1	接受化學治療	40150	638/	935/ 可信度: 95%, 使用率: 3%
3	*434.91	伴有腦硬變之腦動脈瘤	20006050	624/	1527/ 可信度: 40%, 使用率: 64%
4	*714.0	關節滑膜炎	80300	133/	292/ 可信度: 45%, 使用率: 26%
5	*401.9	肺動脈高血壓	80020070	854/	2093/ 可信度: 40%, 使用率: 7%
6	*272.4	肺動脈高血壓	84880	409/	1083/ 可信度: 37%, 使用率: 25%
7	*424.1	主動脈瓣病變	72020331	54/	161/ 可信度: 33%, 使用率: 20%
8	*518.81	急性呼吸窘迫	45960	299/	922/ 可信度: 32%, 使用率: 18%
9	*428.0	充血性心臟衰竭	6C920	261/	970/ 可信度: 26%, 使用率: 18%
10	*424.0	充血性心臟衰竭	72020333	91/	379/ 可信度: 24%, 使用率: 20%
11	*36.14	四級或多級主動脈冠狀動脈	80006009	86/	95/ 可信度: 90%, 使用率: 97%
12	*37.22	左心導管檢查	80020070	1913/	2093/ 可信度: 91%, 使用率: 34%
13	*88.56	使用二導管之冠狀動脈造影	80020070	1865/	2093/ 可信度: 89%, 使用率: 31%
14	*39.61	開心手術之輔助體外循環	72020333	320/	379/ 可信度: 84%, 使用率: 82%
15	*99.25	應用化學治療物質之注射	40150	782/	935/ 可信度: 83%, 使用率: 4%
16	*36.15	單側內乳動脈至冠狀動脈	80006009	59/	95/ 可信度: 62%, 使用率: 31%
17	*30.22	排列剖切術	72020329	42/	95/ 可信度: 42%, 使用率: 3%

圖 3 疾病分類預測結果

Undercoding 檢核是要找出疾病分類人

員應編碼未編碼之 ICD9CM，當疾病分類人員要求系統對某個住院病人執行 Undercoding 檢核時，使用者必須先設定可信度標準值，系統依據醫師開立之醫令從【疾病分類關聯規則集】中找出所有符合條件而且可信度大於設定值的關聯規則進行 ICD9CM 預測，將所有預測之 ICD9CM 與疾病分類人員編列之 ICD9CM 做比對，凡未出現在疾病分類人員之編碼清單中的 ICD9CM，就有可能是應編碼未編碼的 ICD9CM，此時 Undercoding 檢核系統可將推論出此 ICD9CM 的規則，提供給疾病分類人員做為是否增加 ICD9CM 的參考，如圖 4。可信度標準值的設定會影響發生誤判的程度，建議設在 30%。

疾病分類編碼	Undercoding 醫令	Upcoding 醫令	關立之醫令	規則集	參數設定				
*414.01	*401.9	*200.0		*36.14	*36.15	*39.61	*37.61	*37.22	最小可信度: 3
前414.01 自體性冠狀動脈粥狀硬化									
診斷	診斷	醫令	醫令說明	置信度	說明				
1	*414.01	自體性冠狀動脈粥狀硬化	80020070	69/	95/ 可信度: 93%, 使用率: 3%				
2	*98.1	接受化學治療	40150	638/	935/ 可信度: 95%, 使用率: 3%				
3	*434.91	伴有腦硬變之腦動脈瘤	20006050	624/	1527/ 可信度: 40%, 使用率: 64%				
4	*714.0	關節滑膜炎	80300	133/	292/ 可信度: 45%, 使用率: 26%				
5	*401.9	肺動脈高血壓	80020070	854/	2093/ 可信度: 40%, 使用率: 7%				
6	*272.4	肺動脈高血壓	84880	409/	1083/ 可信度: 37%, 使用率: 25%				
7	*424.1	主動脈瓣病變	72020331	54/	161/ 可信度: 33%, 使用率: 20%				
8	*518.81	急性呼吸窘迫	45960	299/	922/ 可信度: 32%, 使用率: 18%				
9	*428.0	充血性心臟衰竭	6C920	261/	970/ 可信度: 26%, 使用率: 18%				
10	*424.0	充血性心臟衰竭	72020333	91/	379/ 可信度: 24%, 使用率: 20%				
11	*36.14	四級或多級主動脈冠狀動脈	80006009	86/	95/ 可信度: 90%, 使用率: 97%				
12	*37.22	左心導管檢查	80020070	1913/	2093/ 可信度: 91%, 使用率: 34%				
13	*88.56	使用二導管之冠狀動脈造影	80020070	1865/	2093/ 可信度: 89%, 使用率: 31%				
14	*39.61	開心手術之輔助體外循環	72020333	320/	379/ 可信度: 84%, 使用率: 82%				
15	*99.25	應用化學治療物質之注射	40150	782/	935/ 可信度: 83%, 使用率: 4%				
16	*36.15	單側內乳動脈至冠狀動脈	80006009	59/	95/ 可信度: 62%, 使用率: 31%				
17	*30.22	排列剖切術	72020329	42/	95/ 可信度: 42%, 使用率: 3%				

圖 4 Undercoding 檢核系統

Upcoding 檢核是找出疾病分類人員不應編列而予編列的 ICD9CM，當疾病分類人員要求系統對某個住院病人做 Upcoding 檢核時，使用者必須先設定 F1-Measure 標準值，系統針對疾病分類人員已編列的 ICD9CM 逐一檢測是否發生 Upcoding，首先從【疾病分類關聯規則集】中找出與該 ICD9CM 相關且 F1 值大於標準值的規則，由這些規則推論出有編這個 ICD9CM 的病人會使用到那些醫令，再與開立的醫令作比對，如果該個案沒有開立相關醫令，則此 ICD9CM 有可能是不應編列而予編列之 ICD9CM，Upcoding 檢核系統可將這個 ICD9CM 應該使用到那些醫令的規則，提供給疾病分類人員做為是否刪除此 ICD9CM 編碼的參考，如圖 5。選擇 F1 而不以可信度作為標準值是因為，此作業之推論模式是以 ICD9CM 推論病人應有之治療醫令，也就是某種 ICD9CM 對醫令的使用比率，使用比率必須達到一定的標準才

可作為檢核的依據，如果以可信度作為標準值容易發生誤判，建議設在 30%。

The screenshot shows a software interface for upcoding checks. It includes several tabs at the top: '疾病分類檢核', 'Upcoding檢核', 'Upcoding檢核', '建立之醫令', '規則集', and '參數設定'. Below the tabs are input fields for '主診斷', '次要診斷1', '次要診斷2', '次要診斷3', '次要診斷4', '手術1', '手術2', '手術3', '手術4', '手術5', and '最小可信度'. A table below lists medical codes with columns for '診斷', '醫令', and '醫令說明'. The table contains two rows of data for diagnosis code 45.25.

圖 5 Upcoding 檢核系統

5. 實驗結果與說明

本研究實驗資料集為台灣中部某醫學中心 2009 年 1 月至 2010 年 6 月之住院病患資料共 74,356 筆，在不涉及病患隱私條件下，每個住院個案僅使用住院案號 TID、醫師開之醫師醫令代碼、健保申報之健保醫令代碼、疾病分類代碼編碼等四個代碼性質的資料，將所有資料轉為醫師醫令資料集 {TID, Ordercode}、健保醫令資料集 {TID, Ordercode}、疾病分類資料集 {TID, ICD9CM} 等三個資料集。

表 8 診斷類 ICD9CM 交易資料集分析

實驗資料集統計分析診斷類 ICD9CM 個案			
項目	總個案	訓練個案 [9/10]	測試個案 [1/10]
個案	74,356	66,921	7,435
ICD9CM 種類數	4,682	4,582	2,290
ICD9CM 編碼數	230,242	207,366	22,8769
每個案 ICD9CM 數	3.096	3.099	3.077
健保醫令種類	5,741	5,669	4,150
健保醫令碼數	4,110,807	3,701,348	409,459
每個案醫令數	55.29	55.31	55.07
醫師醫令種類	6,819	6,724	4,512
醫師醫令碼數	4,432,295	3,990,280	442,015
每個案醫令數	59.61	59.63	59.45

訓練個案與測試個案以 9 比 1 的分配比例進行分配，首先將 74,356 個案區分為有無手術類 ICD9CM 個案，分別將這二類個案依 TID 作排序，有手術類 ICD9CM 之 55,522 個案中排名為 10、20、30...之個案為測試資料其餘為訓練資料，以相同作法

處理沒有手術類 ICD9CM 之 18,834 個案。分析 66,921 個訓練個案有 6,221 種 ICD9CM、294,115 個 ICD9CM 編碼、每個案有 4.39 個 ICD9CM 編碼；5,699 種健保醫令、6,743,302 個健保醫令、每個案有 101 個健保醫令；6,724 種醫師醫令、7,274,822 個醫師醫令、每個案有 109 個醫師醫令，如表 8,表 9。

表 9 手術類 ICD9CM 交易資料集分析

實驗資料集統計分析手術類 ICD9CM 個案			
項目	總個案	訓練個案 [9/10]	測試個案 [1/10]
手術個案	55,522	49,970	5,552
ICD9CM 種類數	1,672	1,639	862
ICD9CM 編碼數	96,370	86,749	9,621
每個案 ICD9CM 數	1.736	1.736	1.733
健保醫令種類	3,753	3,716	2,756
健保醫令碼數	3,379,450	3,041,954	337,496
每個案醫令數	60.87	60.88	60.79
醫師醫令種類	6,684	6,591	4,329
醫師醫令碼數	3,649,182	3,284,542	364,640
每個案醫令數	65.72	65.73	65.68

在 CPU 為 Intel Core 2 Duo T8100 (2.10Ghz)、2G 記憶體、作業系統為 Windows Vista Home Premium 版、Java 版本為 1.6.0_11 版之電腦上，執行以 Java 開發的 CARFLS 演算法，在參數設定為 MAX_kRule=4、oneItemF1Threshold=0.01、F1Threshold=0.1、ConfidenceThreshold=0.10、RefinedConfidence=0.03、QualityConfidence=0.20、Max_Recall=200%

下探勘健保醫令與 ICD9CM 關聯規則需時 17 分鐘、規則數為 50,934 個，探勘醫師醫令與 ICD9CM 關聯規則需時 28 分鐘、規則數為 76,257 個。

由從健保醫令探勘所得之疾病分類關聯規則中，每種 ICD9CM 取一個分數最高的規則，分析這些規則發現與手術類醫令相關的規則佔診斷類 ICD9CM 45%、平均可信度 54%、平均召回率 25%；佔手術類 ICD9CM 63%、平均可信度 66%、平均召回率 44%，如表 10。綜合分析手術、特殊材料、治療處費、檢查、藥品這 5 類醫令與疾病分類之關聯性較高，其中藥品對

ICD9CM 的預測能力排名第 5 名，這與許多研究以藥品來預測疾病分類的想法有著極大的差異。

表 10 健保醫令規則分析表

診斷 ICD9CM 規則分析				
醫令類別	種類數	佔率	平均可信度	平均召回率
手術費	1,279	45%	54%	26%
特殊材料費	468	17%	56%	17%
治療處置費	340	12%	59%	20%
檢查費	282	10%	41%	18%
藥品	275	10%	42%	21%
其他	175	6%	61%	12%
合計	2,819	100%	53%	22%
手術類 ICD9CM 規則分析				
醫令類別	種類數	佔率	平均可信度	平均召回率
手術費	908	63%	66%	44%
特殊材料費	233	16%	77%	17%
治療處置費	147	10%	77%	21%
檢查費	49	3%	54%	26%
藥品	25	2%	74%	17%
其他	71	5%	73%	31%
總計	1,433	100%	69%	35%

將 7,435 個測試個案分別以健保醫令與醫師醫令探勘所得之關聯規則進行測試，在診斷類 ICD9CM 最多預測 13 個、手術類 ICD9CM 最多預測 7 個的條件下，測試結果如表 11、表 12。診斷類 ICD9CM 最佳之召回率為 52.6%，由只使用 1 個醫令的規則到合併使用 2 個醫令的規則時召回率可提升 2% 左右，再合併使用 3 個以上醫令的規則時召回率沒有明顯改善。手術類 ICD9CM 最佳之召回率為 84.8%，由只使用 1 個醫令的規則與合併使用 2 個以上醫令的規則時召回率並沒有明顯的改善。就規則的可理解性而言，在召回率沒有明顯的改善下，本研究建議在建置疾病分類推薦與檢核系統時，與診斷類 ICD9CM 有關的規則只挑 1 個及 2 個醫師醫令的規則，與手術類 ICD9CM 有關的規則只挑 1 個醫師醫令的規則即可。

表 11 診斷類 ICD9 測試結果

個案數		7,435				
ICD9CM 數		3.08				
類別	預測	正確	召回率	精準度	F1	
健保醫令	1Item	11.75	1.52	49.4%	12.9%	20.5%
	2Item	12.05	1.59	51.6%	13.2%	21.0%
	3Item	12.05	1.60	51.9%	13.3%	21.2%
	4Item	12.05	1.60	51.9%	13.3%	21.2%
醫師醫令	1Item	11.97	1.54	50.0%	12.9%	20.5%
	2Item	12.26	1.61	52.3%	13.1%	21.0%
	3Item	12.27	1.61	52.3%	13.1%	21.0%
	4Item	12.27	1.62	52.6%	13.2%	21.1%
	前三天	11.86	1.54	50.0%	13.0%	20.6%

表 12 手術類 ICD9 測試結果

個案數		5,552				
ICD9CM 數		1.733				
類別	預測	正確	召回率	精準度	F1	
健保醫令	1Item	5.74	1.42	81.9%	24.7%	38.0%
	2Item	5.78	1.43	82.5%	24.7%	38.0%
	3Item	5.78	1.43	82.5%	24.7%	38.0%
	4Item	5.78	1.44	83.1%	24.9%	38.3%
醫師醫令	1Item	5.99	1.46	84.2%	24.4%	37.8%
	2Item	6.00	1.47	84.8%	24.5%	38.0%
	3Item	6.00	1.47	84.8%	24.5%	38.0%
	4Item	6.00	1.47	84.8%	24.5%	38.0%
	前三天	5.69	1.27	73.3%	22.3%	34.2%

表 13 與健保局規則比較

本研究與健保局提供規則比較		
個案數	5,270	
ICD9CM 數	8,347	
每個案 ICD9CM 數	1.58	
	健保局	本研究
規則數	5,618	6,140
最多推薦數	無上限	7 個
平均推薦數	6.02	5.67
正確數	1.09	1.30
召回率	68.90%	82.27%
正確率	18.13%	22.99%
F1	28.71%	35.94%

使用健保局【全民健康保險手術、處置支付標準與 ICD-9-CM 手術及處置代碼對應檔】之規則，對本研究之測試個案進行預測並本研究方法比較，首先去除 ICD9CM 不在健保局規則內之個案與編碼後，測試個案只剩 5,270 個、ICD9CM 個數

8,347 個，在健保局規則的召回率為 68.90%，本研究的召回率為 82.27%，詳如表 13。由此可知本研究的方法在效能上較健保局提供之規則為佳。

使用[疾病分類關聯規則集]對第一階段實施也就是 2010 年 01 月開始適用 TW-DRGs 給付制度的 111 項 DRGs 個案進行測試，整體召回率為 75.7% 與 Patrick Ruch 所提之召回率理想上限值 79% 相比只有 3.3% 的差距、主診斷正確率高達 83.6%，如表 14 所示。健保局目前聘請疾病編碼專家對 DRGs 個案進行審查，對於不適當的編碼將予以重新編碼後再行核付費用，我們推測健保局若先以這些高品質的規則對第一階段實施的 111 項 DRGs 個案進行 Upcoding 檢測，針對有 Upcoding 之疑的個案進行審查應可提升審查品質。

表 14 以 TW-DRGs 實施階段分析

實施階段	1 (2010)	2 (2011)	3 (2012)	4 (2013)	5 (2014)	非 DRG
個案數	1,131	874	1,035	872	649	2,874
ICD9CM 數	4.01	4.18	3.86	4.43	3.96	4.83
預測數	19.0	16.8	15.5	16.3	15.4	16.8
正確數	3.04	2.61	1.95	2.47	1.97	3.14
主診斷正確率	83.6%	67.3%	44.2%	59.6%	49.0%	75.6%
手術類召回率	88.6%	85.9%	78.5%	85.3%	73.9%	86.1%
診斷類召回率	66.9%	51.3%	41.2%	46.2%	43.4%	56.1%
合計召回率	75.7%	62.4%	50.6%	55.8%	49.7%	65.0%

因 Larkey 等之研究對象 Brigham and Women's Hospital 為波士頓地區高危險妊娠孕婦後送轉診醫院，其病患型態集中在孕婦與新生兒，與本研究實驗資料之背景差異極大，無法相互比較。Patrick Ruch 等之研究對象 University Hospitals of Geneva 為綜合型醫院，實驗資料與本研究資料較為相近，其 R20 也就是推薦 20 個 ICD-10 時的召回率為 63% 與本研究的 64.2% 比相差 1.2%，由此可知本研究所提 CARSLF 演算法是一個可行性很高的方法。

6. 結論與未來研究

本研究採醫令就即時性與方便性而言均比其他病歷內容較佳，對台灣所有醫院而言健保醫令與 ICD9CM 均為電子化資料，在資料取得上不必受限於電子病歷的限制普遍性最佳，與 NLP 方法相較本研究所提之方法，較容易且不受限制。

在成效的比較上與健保局提供之手術類 ICD9CM 規則比較召回率為 82.27% 比 68.90%，本研究之方法明顯高出很多。與國際先進醫療機構之研究成果相比總召回率為 64.2% 比 63%，亦有同等級的表現。

在預測第一階段實施 DRGs 個案時整體召回率 75.7%、手術類 ICD9CM 召回率 88.6%、診斷類 ICD9CM 召回率 66.9%、主診斷正確率 83.6%，健保局若先以這些高品質疾病分類關聯規則對第一階段實施的 111 項 DRGs 個案進行 Upcoding 檢測，針對有 Upcoding 之疑的個案進行審查應可提升審查品質。

在研究方向上可納入檢查驗之報告值作為採勘資料，就領域知識來說，報告值之正常與否與疾病分類有很大的關聯。在作法上可將報告值異常當作一個 Item，加入 CARFLS 演算法進行探勘，例如將 ItemX="血壓檢測值異常"，若有一個病人血壓檢測異常就建立一筆 {TID, ItemX} 之交易資料，我們預測以這樣的方式進行應可以改善整體的召回率及主診斷正確率。

致謝

感謝本研究獲得國科會部分經費補助，計畫編號：NSC 99-2221-E-005-080。

參考文獻

- [1] 總校閱: 范碧玉、主編: 鄭茉莉、張樹棠, *ICD-9-CM 分類規則彙編*, 台灣病歷管理協會, 2006。
- [2] 陳鎮揮、楊鎮嘉、吳效文, 2008, 「以資料探勘技術建立 ICD-9-CM 編碼決策系

統」，光田醫學雜誌，2008，vol.3 no.5，
pp.31-41

- [3]健保局,全民健康保險手術、處置支付標準與 ICD-9-CM 手術及處置代碼對應檔, [http://www.nhi.gov.tw/webdata/AttachFiles/Attach_9718_1_ICD-9-CM\(99.08.16\).xls](http://www.nhi.gov.tw/webdata/AttachFiles/Attach_9718_1_ICD-9-CM(99.08.16).xls)
- [4]Farkas., Szarvas., Automatic construction of rule-based ICD-9-CM coding systems., Proceedings of the Second International Symposium on Languages in Biology and Medicine (LBM), Singapore, 2007.
- [5]Fabrizio Sebastiani: Machine learning in automated text categorization. ACM Comput. Surv.34(1),2002,pp. 1-47
- [6]Grigorios Tsoumakas., Ioannis Katakis., Ioannis Vlahavas., Mining multi-label data. In O. Maimon and L. Rokach, editors, Data Mining and Knowledge Discovery Handbook. Springer, 2nd,2010, pp. 667-685..
- [7]Larkey L., Croft W., Combining classifiers in text categorization. SIGIR '96: Proceedings of the Nineteenth Annual International ACM SIGIR Conference on Research and Development in International Retrieval. NY: ACM Press; 1996. pp. 289-297.
- [8]Liu, B., Hsu, W., Ma, Y. 1998. Integrating Classification and Association Rule Mining. KDD'98, 1998.
- [9]Patrick Ruch., Julien Gobeill., Imad Tbahriti., Antoine Geissbühler., From Episodes of Care to Diagnosis Codes: Automatic Text Categorization for Medico-Economic Encoding, AMIA Annu Symp Proc., 2008, pp. 636-640.