

# 運用決策樹技術探討急診病患醫療費用之消耗

周歆凱<sup>1,2,3</sup> 蘇 喜<sup>1</sup> 黃興進<sup>1,2,\*</sup>

蔡明足<sup>3</sup> 翁林仲<sup>3,4</sup>

**目標：**傳統上，決策樹之分類技術在市場上多運用於顧客資料的區隔分析。然而，本研究也應用該工具來探討急診病患醫療費用之耗用，期望從大量之費用資料庫中，探勘出病患屬性與其醫療費用消耗之潛在關係。**方法：**本研究收集某醫學中心急診室一年之病患就診資料，並利用資料探勘技術中之決策樹工具來觀察各醫療費用群(低費用組、一般費用組、高費用組)間之病患特質(人口學特質、就醫屬性)的分類；藉由分類規則的建立，可預測病人於就診時可能消耗之醫療費用多寡。**結果：**決策樹以多層次之樹枝分佈及顏色區塊等視覺化方式呈現研究結果；其中資訊增益順序為(滯留時間>疾病分類>離院後動向>檢傷分級>科別)，該資訊增益之順序也代表屬性影響醫療費用分佈之程度，意即滯留時間為決定急診病患醫療費用多寡之首要因素。**結論：**本研究建議個案醫院能針對不同類型之病人，給予個人化的照護服務，期望改善病人再回診的情形、降低滯留急診的時間，同時也能降低病人於急診發生之醫療費用。(台灣衛誌 2006；25(6)：430-439)

**關鍵詞：**急診醫學、資料探勘、決策樹

## 前 言

醫院急診部門提供民眾24小時、全年無休且專業的緊急醫療救護服務，使得病人無論在何時、何地遭遇到任何的緊急事故，均可透過醫院的急診室得到最完善、最專業、最迅速的照顧[1]。近年，急診部門病患日益增多及疾病複雜性高的原因一方面是由於社會型態的急速轉變，意外事件遽增使得緊急醫療照護的需求益形迫切，另一方面是由於醫療服務供應能力的速度趕不上醫療消費

擴增的腳步，民眾為求得更快速便捷的醫療服務，便紛紛利用急診室來滿足需求[2]。

急診人次的與日遽增更衍生出急診部門的潛在危機，其中伴隨而來的更包含急診醫療費用的高漲；以Mustard於加拿大之研究[3]，1985-86之急診率為59.9%，1988-89為64.9%，成長率為8.3%；而李玉春[4]針對「全民健保門診之研究」顯示急診醫療費用則由1996年的362.7億、1997年393.9億到1998年的440.3億，成長率高達11.79%。由以上國內外研究顯示，民眾對於急診醫療之利用有逐年增加的現象。但Waller等人[5]指出由於「急診醫療的成本太高」、「急診室太擁擠」及「急診醫療無法提供連續性服務」，所以當急診醫療的服務量越來越大時，對於健康照護品質呈現明顯的負面影響。

因此，為了改善急診室的擁擠情形、增進醫護人員之工作效率、提供全方位之緊急醫療救護服務，唯有充分了解急診病人之疾

<sup>1</sup> 國立台灣大學公共衛生學院醫療機構管理研究所

<sup>2</sup> 國立中正大學資訊管理研究所

<sup>3</sup> 台北市立聯合醫院仁愛院區

<sup>4</sup> 國立台灣大學醫學院

\* 通訊作者：黃興進

聯絡地址：嘉義縣民雄鄉大學路168號

E-mail: mishgh@ccu.edu.tw

投稿日期：95年2月14日

接受日期：95年10月5日

病特質與就醫屬性之關係，藉此醫護人員即可提供專業化、個人化之醫療服務給急診病患，並針對資源耗用較高的特殊個案(如高醫療費用之病人)擬定策略方針予以重點管理。本研究使用目前企業管理廣泛應用之「資料探勘」技術，企圖挖掘出潛在之急診資源高耗用者，以提供醫院管理者策略擬定之參考。本研究之目的如下：(1)分析急診病患特質屬性與其醫療費用之消耗情形，(2)比較資料探勘工具與統計分析工具之結果，以證明探勘工具之效度，(3)實證資料探勘工具於醫院管理領域之實用性。

資料探勘(Data Mining)是近年來隨著人工智慧和資料庫技術的發展而出現的一門新興技術。它可從大量的資料中，萃取出隱含的、過去不為人知、可信而有效的資訊和知識，如趨勢(Trend)、特徵(Pattern)和相關性(Relationship)，並能將資料去蕪存菁、瀝沙揀金地再分析，以獲得更深入成因的瞭解，同時具有預測功能，可藉由既有資料來預測未來[6]。其主要的貢獻在於可從資料庫中獲取有意義的資訊以及對資料歸納出有結構的模式，以作為企業在進行決策時之參考依據[7]。

在醫療產業上，黃仁貴[8]、黃勝崇[9]等學者將資料探勘的功能分為五類：分類、推估、預測、關連分組以及群集化，各功能均涵蓋了各種的探勘技術，如決策樹(Decision Trees)、記憶基礎推理(Memory-Based Reasoning)、類神經網路(Artificial Neural Networks)、基因演算法(Genetic Algorithms)、鍊結分析(Link Analysis)、購物籃分析(Market Basket Analysis)、叢集分析(Cluster Detection)等技術。其中各功能也應用於醫療資訊之建立及臨床研究之發展，如支援醫療照會決策系統[10]、臨床專家系統之建立[11]、病例與藥品之關聯性[12]、臨床路徑之建立[13]、健保費用之稽核[14]、病患就醫特性之研究[15]、醫療服務量之預測[16]。因此本研究欲利用資料探勘之分類功能來探討醫院急診室病人醫療費用消耗情形。

決策樹，乃是應用資料探勘中的分類

(classification)技術，將資料切割成數個集合，每個集合有簡單的規則。利用CART(classification and regression tree)、CHAID(chi-squared automatic induction)及C4.5等演算技術來建立模式，依照所分析對象的屬性分門別類並加以定義，建立類組。應用研究包含，將銀行借貸客戶之信用狀況區分為高風險、中度風險及低風險[17]，或是將顧客區分為高貢獻度族群、高忠誠度族群等[18]。在醫療研究上，蘇格蘭及英國曾利用決策樹輔助急診醫師對於心肌梗塞病患之臨床診斷[19]，幫助醫師能依據某些表徵快速判斷急診病人可能為心肌梗塞。決策樹的主要優點之一就是模式具有相當的解釋性，因為具有明顯的規則及視覺化的輸出，因此易於評估其結果，以及辨識過程中重要的屬性。

## 材料與方法

本研究主要探討急診室病患之特質與其資源耗用情形，藉由取得急診室病患之基本資料(人口學特性)、就醫情形(就醫屬性、健康狀況)及其醫療過程中之耗用資源(醫療費用、滯留時間)來找出隱含其中的訊息及規則，希望藉由規則及模式之建立來預測病患可能消耗之醫療費用。在研究變項之選擇，不僅參考李芳年[20]在2000年針對某區域教學醫院急診之研究，同時也諮詢急診專家醫師潛在之重要變項。本研究的資料來源是以台北市某醫學中心急診部之就診病患為研究對象，該院每日平均急診人次為250至300人次/日；研究期間為2001年1月至2002年3月，以醫院電腦資訊系統登入急診就診病患的人口學資料、疾病特質與其資源耗用的情形；共收集111,514筆樣本資料，在經過資料除錯後，排除電腦系統的代入錯誤、人為作業的鍵入錯誤、資料記載不完全等三類疏失，取得有效樣本為85,330筆。

本研究在資料分析的過程中共分為兩階段：第一階段為「一般之統計分析」，將資料作前處理以利日後資料探勘之用，使用SPSS 10.0 for Windows統計軟體為研究工

具，將性別、年齡、科別、檢傷分級、滯留時間、回診註記、疾病分類、來院方式、離院後動向等變項作描述性分析、雙變項分析及多變項分析；第二階段則進入「資料探勘分析工作」，藉由PolyAnalyst 4.0資料探勘軟體中之分類(classification)技術—決策樹(Decision Tree)工具；該強大分類功能可使用多層次之樹枝分佈及顏色區塊等「視覺化」(visualization)的方式呈現急診病患醫療費用之分組情形；將來院方式、檢傷分級、科別、性別、離院後動向、滯留急診時間、疾病分類投入決策樹作為分隔屬性，而醫療費用消耗群(高費用組、一般費用組、低費用組)作為目標屬性，進行決策樹分析。藉此探勘出高醫療費用病患之特性，期望兩階段之分析結果能相輔相成以獲得最佳的探勘模式；藉由傳統之統計方法與新興之探勘科技相融合，以提供急診部門在管理決策上所需的資訊及建議。

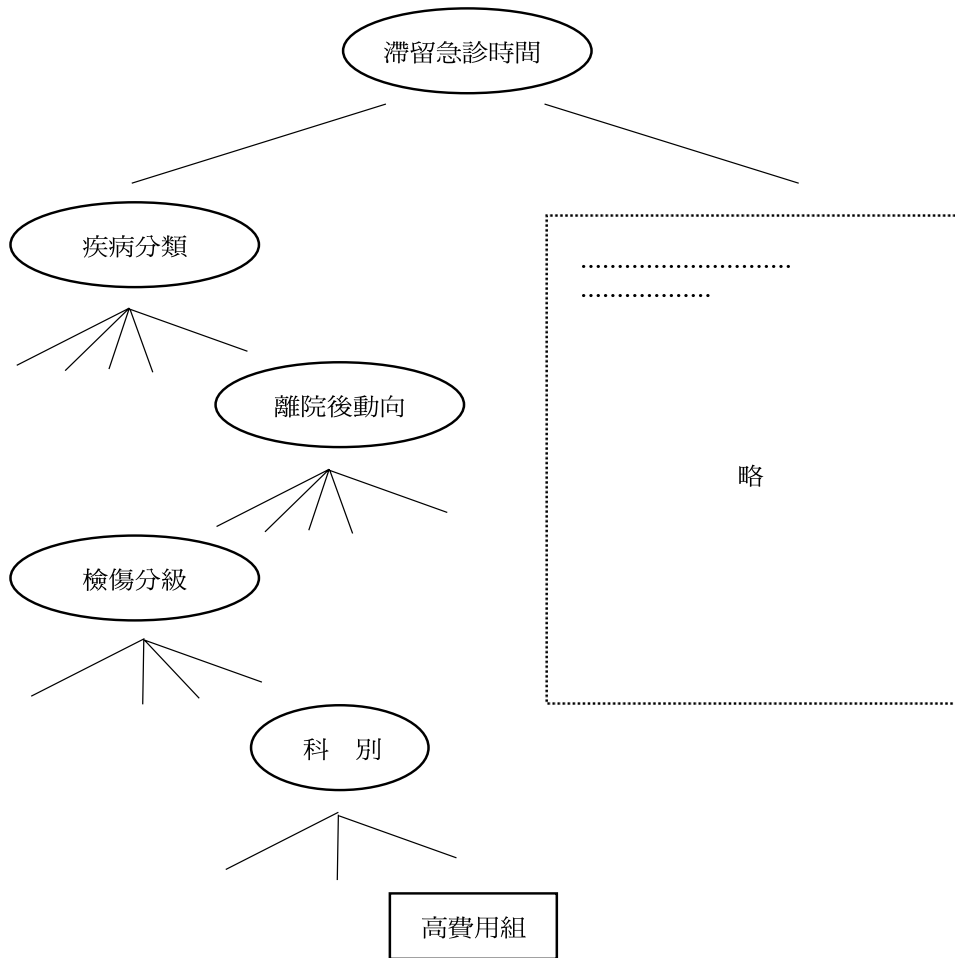
## 結 果

本研究個案醫院之「描述性統計分析」中，男性與女性就診人次所佔之比例各居一半，分別為50.8%及49.2%，其中明顯發現急診之就診病患主要發生在15~24歲青少年(14.9%)、25~34歲壯年(14.8%)及0~4歲嬰幼兒(13.7%)等族群，老年人之急診發生頻率則較少。就診病患中，將近有一半為內科之病人(47.9%)、其次為外科(31.1%)、兒科病人則位居第三(19.0%)，產科(0.5%)及牙科(1.4%)病患則最少。以檢傷分級來看，急診病患中將近一半以上均為檢傷三級的病患(57.2%)，其次為二級(35.7%)，至於最嚴重的一級病患(6.5%)及俗稱假急診的四級病患(0.6%)均屬少數。該個案醫院中再急診率(72小時再回急診)為2.9%，滯留急診超過24小時之病患也佔18.1%。依疾病發生頻率的高低排列順序，以「消化循環系統疾病」為最高(17.9%)；「呼吸系統疾病」居次為16.0%、第三為「症候，徵候及診斷欠明之各種病況」(12.9%)、第四為「損傷及中毒」(9.1%)。

病患自來院進入急診部接受醫療服務，到完成治療並離開急診部之整個就醫流程中，急診病患平均耗用醫療費用合計為2882.86元、標準差7117.70元，其中平均耗用處置費1754.36元、診療費488.23元、藥費599.96元、藥事服務費40.63元；在「急診病患耗用醫療費用之相關分析」中，以病患屬性來看，男性(3264.05元)其平均醫療費用消耗較女性(2492.09元)高，整整多了771.96元；以年齡層看病患醫療費用之消耗，發現年齡與費用消耗成正比；年齡越高之病患，其醫療費用消耗也就越多，75歲以上之老年人的費用消耗最多(5213.64元)，其次為65~74歲(4752.23元)，0~4歲的嬰幼兒醫療費用則最低(1326.00元)。在就醫屬性中，科別以內科病人費用最高(3856.99元)，其次為外科的病人(2216.81)，而產科病人在急診之耗用費用則是最少(830.49元)；從檢傷分級中可看出，病情越嚴重之病人其消耗之醫療費用越多，一級病人花費高達6937.00元、二級病人僅3679.04元、三級與四級的病人則只消耗1913.53與1452.23元；在病人身分註記中發現，72小時再回診病患(3760.48元)其費用消耗較一般病人(2858.02元)多；此外，滯留超過24小時之病人(7773.77元)其費用更是遠高於滯留急診24小時內之病人(1829.83元)。急診病患系統疾病與其醫療費用之相關分析結果所示，病患之系統疾病分類與其醫療費用之消耗有達到統計上顯著差異；以血液和造血器官疾病所消耗之費用最高(17118.00元)是所有急診病患平均耗用(2882.86元)的三倍之多，其次為患有腫瘤之病人(6456.05元)，循環系統疾病則排名第三(4908.90元)。

本研究將複迴歸分析中達到統計上顯著差異之變項，如來院方式，檢傷分級，科別，性別，離院後動向，滯留急診時間，疾病分類投入決策樹作為「分隔屬性」，而醫療費用消耗群(高費用組、一般費用組、低費用組)作為「目標屬性」，進行PolyAnalyst Decision Tree分析。其決策樹分類之簡圖如圖一所示。

可從圖一：決策樹分類簡圖中節點



圖一 急診病患醫療費用之決策樹分類簡圖

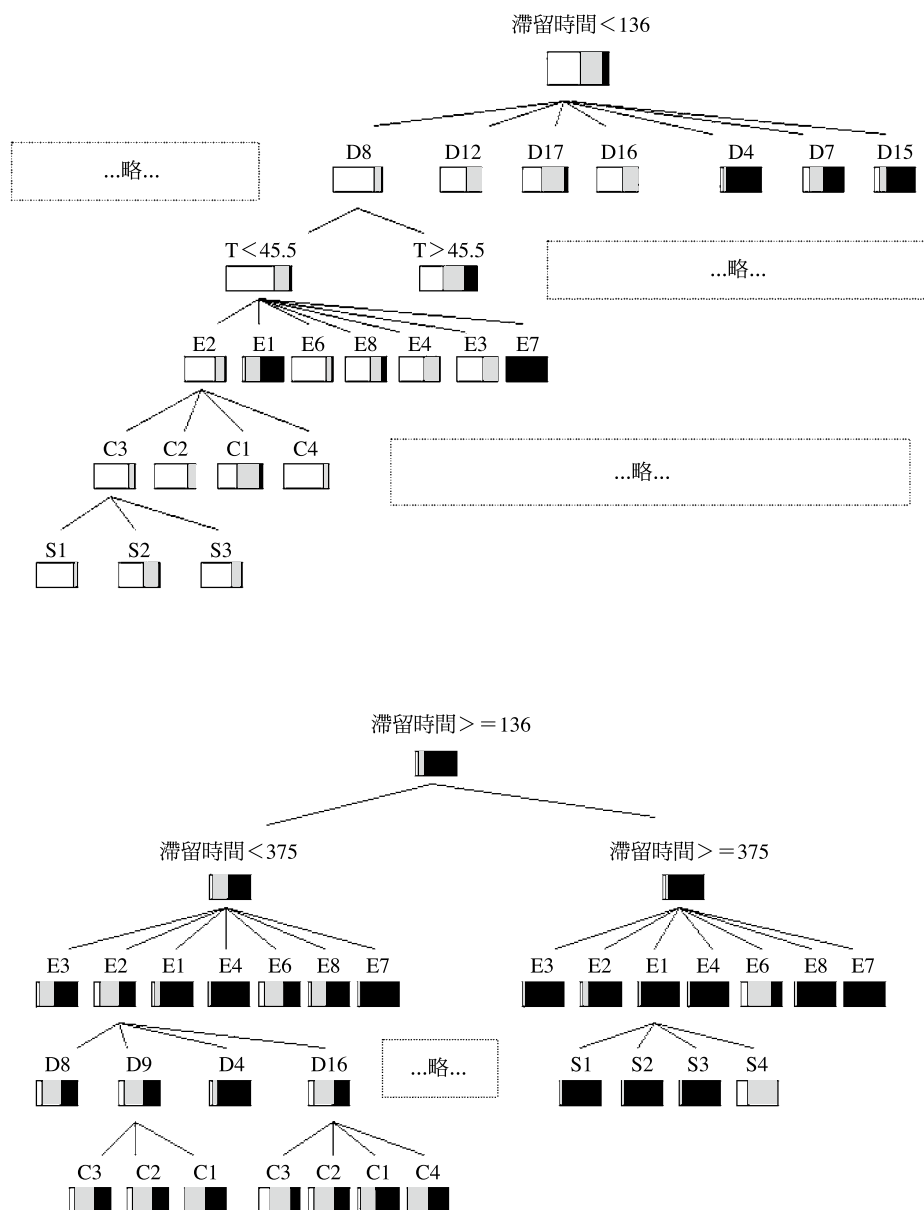
(node)的層次(level)分佈結果得知，該決策樹的資訊增益順序為(滯留時間>疾病分類>離院後動向>檢傷分級>科別)，該資訊增益之順序也代表屬性影響醫療費用分佈之程度，意即滯留時間為決定急診病患醫療費用多寡之首要因素，其次為疾病分類，第三為病患離院後動向，第四為檢傷分級，科別則排名最後。

另外，本研究從圖二：決策樹分類總圖中，可依各節點(node)中方格的顏色分佈比率來判斷醫療費用的多寡，以下即簡述決策樹中各屬性之分佈特性：

#### 【屬性一】滯留急診時間

決策樹在第一層級(Level 1)中，約略將時間切割為「大於136分鐘」及「小於136分鐘」兩組；當中在「大於136分鐘」的子樹中，我們可發現大部分的節點多呈現灰色(26%)及黑色(67.5%)的分佈，其中若將滯留時間更細分至「大於375分鐘」的節點中，可發現黑色在方格中佔有88%的比率(即滯留超過375分鐘之病人中，有88%落在高醫療費用組中)，因此研究顯示病人滯留急診時間越久，其形成高醫療費用之機會就越高。





圖二 急診病患醫療費用之決策樹分類總圖

## 【屬性二】疾病分類

決策樹在第二層級(level 2)中發現，在停留時間「小於136分鐘」這子集中，大部分之系統疾病均會落在低費用組(白色)或一般費用組(灰色)，僅有些特殊疾病會集中在高費用組(黑色)；例如：血液及造血器官疾

病仍有79.7%屬於高費用組，週產期疾病有63.6%，循環系統疾病有41.7%屬於高費用組。另外，停留時間「大於136分鐘」之子集中，各系統疾病均普遍落在一般費用組及高費用組；但如果停留時間更「大於375分鐘」後，則各疾病均普遍落在高費用組。

### 【屬性三】離院後動向

無論滯留時間「小於136分鐘」或是「大於136分鐘」，病人在任何系統疾病下其離院後動向若為死亡、住院、開刀都會形成較高的醫療費用消耗；另外，如果病人滯留時間「大於375分鐘」，則離院後動向除了逃院所佔比率(34.9%)較低外，其餘如轉住院(97.7%)、轉門診治療(82.4%)、轉院(96.2%)、開刀(98.7%)、死亡(100%)、自動出院(96%)均多數落在高費用組。在此值得討論的是，當滯留時間「大於136分鐘」時，病患「離院後動向」之資訊增益大於「疾病分類」；意即當病患滯留時間超過136分鐘後，病患「離院後動向」比其「疾病分類」更容易影響醫療費用之多寡。

在此小節中，本研究藉由決策樹分類規則的建立，我們可預測病人於急診就診時可能消耗之醫療費用多寡(高費用組或是低費用組)，該預測結果可協助醫院管理者針對高醫療費用者做重點管理，以減少急診資源之耗用。

## 討 論

### 1. 一般統計分析之結果

本研究中急診病患的醫療費用分佈大致上與其滯留時間之分佈相類似；在實務上「醫療費用」與「滯留時間」同屬於急診資源之消耗，而統計上滯留時間與醫療費用亦呈現正相關的現象(即滯留時間越久，所消耗之醫療費用則越多)。就人口學特質與醫療費用之關係，男性較女性高、費用亦隨著年齡層增加而遞增，此外醫學中心所發生之醫療費用平均高於區域醫院700至1000元左右。在「檢傷分級」中，醫療費用也會隨著疾病嚴重度增加而遞增；在「就診科別」中，滯留時間及醫療費用均屬內科病人消耗最多；在「疾病分類」裡，本研究發現高額費用之疾病類型均僅集中於罕見疾病，如血液和造器官官疾病、腫瘤、循環系統疾病、傳染病級寄生蟲病；反之，盛行率較高之消化循環系統疾病和呼吸系統疾病所消耗之費用反而低於樣本平均費用。因此若要監控急

診病患之醫療費用，除了滯留急診時間外，病患之疾病分類也是關鍵因素之一。

### 2. 資料探勘決策樹分析之結果

藉由決策樹之強大分類功能，本研究以多層次之樹枝分佈及顏色區塊等「視覺化」(visualization)的方式呈現急診病患醫療費用之分組情形(低費用組、一般費用組、高費用組)；在龐大之樹狀圖(見圖二)中，臨床醫師能依需求從中挖掘出許多有趣的結果。其中，決策樹之資訊增益順序(滯留時間>疾病分類>離院後動向>檢傷分級>科別)也可提供醫院管理人員預估急診病患醫療費用消耗情形；在滯留時間較短的病患中，需針對特殊疾病患者做適當的處置及治療；若其滯留時間較長，則除了疾病之監控外，急診醫師對該病患之處置(如轉住院、轉開刀、轉至他院)也嚴重影響其醫療費用之多寡。因此若要減少醫療費用之發生，僅能加強急診專科醫師的教育訓練，使其提供高品質、迅速的治療服務；並配合急診作業流程改善、醫護人員專業分工，促使急診工作更有效率，提昇工作人員產能、增加病患之流動率，如此可藉由降低病患滯留時間，有效的減少其於急診發生的醫療費用。

### 3. 一般統計分析與資料探勘之比較

本研究利用資料探勘之決策樹工具來探討各醫療費用群(低費用組、一般費用組、高費用組)間之病患特質(人口學特質、就醫屬性)的分類，由於決策樹之強大分類功能，能夠以多層次之樹枝分佈及顏色區塊來呈現本研究之結果；然而，藉由區塊中各節點(node)的顏色分佈，可看出病患屬性組合的費用分類；在本研究之決策樹分佈中，有許多結果與一般統計分析(如：雙變項分析、複迴歸分析)結果一致，例如：從複迴歸模型中各變項之迴歸係數中，可發現影響醫療費用分佈之程度：滯留時間>疾病分類>檢傷分級≒離院動向>科別；該迴歸結果也與決策樹中資訊增益之順序(滯留時間>疾病分類>離院動向>檢傷分級>科別)不謀而合；從上述複迴歸及決策樹結果得知，本研究也驗證資料探勘工具中決策樹的研究

效度，確實與一般統計所得結果一致。此外，本研究更發現決策樹之另一特殊功能，即是異常資料之稽核；在樹的分佈中，有許多的結果是異於一般統計分析，而此異常結果或許能協助管理者發現醫院急診部門之潛在問題，更期待能從此異常結果中尋求最佳解決方案。以下則概述本研究發現之異常結果，並探討可能原因：

- (1)「滯留時間小於136分鐘、血液及造血器官疾病」該組合之病人無論離院後動向或檢傷分級為何，均有79.7%的病人會落在高醫療費用組。在一般之統計分析中顯示，滯留時間為影響醫療費用分佈之最重要關鍵；病患滯留越久，則所發生之費用越高。在該組合中，雖然其滯留時間很短(小於樣本平均滯留時間431分鐘)，但仍有79.7%的病人屬於高費用組。因此本研究推斷「血液及造血器官疾病」可視為費用監控的重點疾病之一。
- (2)「滯留時間77~136分鐘、呼吸系統疾病、離院後轉門診治療」該組合中，檢傷四級病人所發生之費用比其他級別(一級、二級、三級)為高。在一般統計分析之結果顯示，檢傷分級越高則費用消耗越少；因為疾病嚴重程度較輕，則所需要之處置費及藥費也就越少；然而本研究在決策樹中卻發現到該組特別之結果，其最可能解釋的原因為：醫療人員在檢傷分級時的疏忽，使得檢傷分類結果無真實反應病患疾病嚴重程度。在臨床上，有時當病人進入至急診室時，其病況似乎不嚴重，因此在檢傷分類站被判定為第四級病人，但事後經醫師實際診察、初步檢驗後，卻發現其疾病並非單純，經常得安排進一步的檢查以找出病因，因此就產生較高的醫療費用，這也就是造成此異常結果的可能原因。本研究也檢測決策樹中該屬性組合之樣本資料，發現該組合也僅有一筆樣本而已；因此，在本研究個案醫院中，推論可能存在「檢傷分類疏失」之情形，使

得檢傷分類結果沒有真實反應病患疾病嚴重程度。該推論亦經由急診專科醫師所認同。在此，本研究也證實了決策樹對於異常資料的稽核功能，類似此種罕見的結果(十萬分之一的樣本)或許是一般統計分析所忽略的；僅能靠決策樹中視覺化的呈現結果，使得管理者或研究人員能依據實務經驗立即判斷出異常資料之所在，並且從中挖掘出問題。

過去文獻指出[20-22]，檢傷分類的目的是在有限的時間及空間裡，將同時至急診就醫之病患，依其病況的嚴重程度予以先後處理，以確保病患能適時接受最適切的照護。因此，本研究相信「檢傷分類」之結果足以影響整個急診作業之效率；好的檢傷分類工作，可改善急診室醫療處理流程、縮短病患就診時間，增加急診室效率；反之，錯誤的檢傷分類，會加重醫護人員工作負荷、影響醫師的判斷與處置、使得病患滯留過久、增加醫療費用發生、浪費急診資源。因此檢傷分類人員的「繼續教育」與「績效考核」應是目前急診管理需要加強的部分。此外，本研究也推崇急診能夠提供個人化的服務，以強化急診作業效率、避免資源耗用，具體建議如下所述：(1)針對滯留時間較長病患，應加強急診作業流程改善、醫護人員專業分工，促使急診工作更有效率，提昇工作人員產能、增加病患之流動率，如此可藉由降低病患滯留時間，有效的減少其於急診發生的醫療費用。(2)針對滯留時間較短但合併重點疾病之病患(如血液及造血器官疾病、周產期疾病、循環系統疾病)，應加強急診醫師專科訓練，使其提供高品質、迅速的治療服務。(3)針對疾病程度複雜或當時急診內醫療資源不足時(如醫療設備不足、專科人力不足)，應立即協助轉院，避免延誤病情且留觀急診造成不必要的資源耗用。(4)針對疾病輕微(如檢傷四級)或假急診的病人，請急診專科護理師或資深護理人員給予衛教，於檢傷分類時即可輔導離院，避免佔用人力資源、舒緩急診留觀之擁擠現象。



## 致 謝

本研究非常感謝台大醫院急診室石崇良醫師在臨床實務上之經驗分享，支持本研究在方法學上及臨床實務上更有信度及效度。

## 參考文獻

1. 陳音潔：病患多次利用急診醫療之影響因素探討—以中部某醫學中心為例。台中：中國醫藥學院醫管所碩士論文，2002。
2. 姜錦輝：台北市緊急醫療照護系統之評估。台北：台灣大學公共衛生研究所碩士論文，1984。
3. Mustard CA, Kozyskyj AL, Barer ML. Emergency department use as a component of total ambulatory care: a population perspective. *Canadian J Med Association* 1998;**158**:49-55.
4. 李玉春：全民健保支付制度實施現況檢討與改革方向建議。政策月刊 1998；**35**：14-6。
5. Waller AE, Hohenhaus SM, Shah PJ, Stern EA. Development and validation of an emergency department screening and referral protocol for victims of domestic violence. *Ann Emerg Med* 1996;**27**:754-60.
6. 林傑斌、劉明德：資料採掘與OLAP理論與實務。台北：文魁資訊股份公司，2002；9-12。
7. Craven MW, Shavlik JW. Understanding time series networks: a case study in rule extraction. *Int J Neural Syst* 1997;**8**:373-84.
8. 黃仁貴：以叢集分析技術探討病患就診屬性與看診時間之關係。台北：台北醫學大學醫學資訊研究所碩士論文，2002。
9. 黃勝崇：資料探勘應用於醫療院所輔助病患看診指引之研究。嘉義：南華大學資訊管理研究所碩士論文，2001。
10. 陳澤生、林錫璋：運用資料挖掘方法支援醫療照會決策系統。行政院國家科學委員會補助研究計畫成果(計畫編號：NSC88-2213-E006-079)，1999。
11. 鄭振鴻：舌診自動分析系統併入中風證型診斷專家系統之研究。行政院衛生署中醫藥年報 2001；**19**：266-347。
12. 陳世源：資料採礦技術在病例與藥品關聯性之研究。高雄：國立中山大學資訊管理學系研究所碩士論文，1999。
13. 李淑芬：臨床路徑之建立機制—應用資料採礦技術。台中：東海大學工業工程與經營資訊研究所碩士論文，2002。
14. 王俊程：運用資料挖掘技術協助健保財務管理—以承保中斷及繳費不正常資料挖掘為例。台北：行政院衛生署委託研究成果報告，2000。
15. 陳益良：應用資料探勘法探討老人就醫特性—以高雄市三民區為例。高雄：國立高雄醫學大學公共衛生學研究所碩士論文，2002。
16. 紀志賢、曾詠淑、林錫璋、陳澤生、蔡良敏：以資料探勘模式實施緊急救護量預測。慈濟醫誌 1999；**11**：337-42。
17. 劉祥熹、黃日鉦：銀行業借貸風險評估之實證分析—資料探勘與統計模型之比較。華人前瞻研究 2005；**1**：45-71。
18. 王智弘：應用資料探勘技術分析及預測銀行顧客貢獻度之研究。嘉義：國立中正大學資訊工程研究所碩士論文，2001。
19. Tsien CL, Fraser HS, Long WJ, Kennedy RL. Using classification tree and logistic regression methods to diagnose myocardial infarction. *Medinfo* 1998;**9 Pt 1**:493-7.
20. 李芳年：急診病人的資源耗用—以台北市某區域教學醫院為例。台北：國立陽明大學醫務管理研究所碩士論文，2000。
21. Rund DA, Rausch TS. *Triage*. St.Louis: CV Mosby Co., 1981.
22. 繆珣：檢傷分類 *Triage*。榮總護理 1979；**6**：38-41。





# 【附錄說明：決策樹之判讀方式】

決策樹之代碼說明：

## D=疾病分類

1傳染病及寄生蟲病 2腫瘤 3內分泌，營養，新陳代謝和免疫疾病 4血液和造血器官疾病 5精神障礙  
6神經系統和感覺器官疾病 7循環系統疾病 8呼吸系統疾病 9消化循環系統疾病 10泌尿生殖系統疾病  
11懷孕，生產及產後合併症 12皮膚及皮下組織疾病 13肌肉骨骼系統及結締組織疾病 14先天異常  
15週產期疾病 16症狀，徵候及診斷欠明之各種病況 17損傷及中毒 18損傷及中毒之外因補充分類  
19影響健康因素及健康服務之補充分類

## E=離院後動向

1住院 2門診治療 3轉院 4開刀 5轉安養機構 6逃院 7死亡 8自動出院

## S=科別

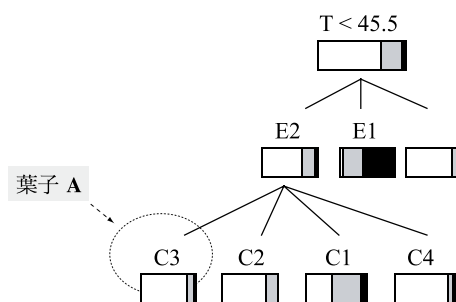
1內科 2外科 3兒科 4產科 5牙科

## C=檢傷分級




1一級 2二級 3三級 4四級

## T=滯留時間

決策樹之分類說明：(以部分枝葉為例)



決策樹中每一個節點(node)均有一小方格呈現各醫療費用組別(低費用組、一般費用組、高費用組)之分布比率，白色代表低費用組、灰色代表一般費用組、黑色代表高費用組。

	→ 白色：	低費用組	占33.3%
	→ 灰色：	一般費用組	占33.3%
	→ 黑色：	高費用組	占33.3%

若小方格中白色所佔面積較大，表示該組別多屬於低費用組；若小方格中黑色所佔面積較大，則表示該組合多屬於高費用組。因此我們可藉由判斷方格中「顏色所佔比例」來決定該組合多屬於何種醫療費用組？以上例舉之決策樹分類結果所示，葉子A該組合『滯留時間小於45.5分鐘、離院後轉門診治療、檢傷三級』之病人其醫療費用消耗大部分屬於低費用組，因此node(節)中多呈現白色區塊(94.8%)。

## Utilization of a decision tree for high expenditure patients in the emergency department

HSIN-KAI CHOU<sup>1,2,3</sup>, SYI SU<sup>1</sup>, HSIN-GINN HWANG<sup>1,2,\*</sup>, MING-TSU TSAI<sup>3</sup>, LIN-CHUNG WOUNG<sup>3,4</sup>

**Objective:** Traditionally, classification via a decision tree has been primarily used to distinguish between types of customers. In the current study, however, a decision tree was used to track a patient's medical expenditures in the emergency department and determine the potential relationship between patient attributes and expenses, as derived from a large database maintained in the emergency department of the hospital. **Method:** Patient records were collected inform the emergency department of a medical center over the course of approximately one year and a decision tree was used to classify patient data based on the magnitude of medical expenses incurred (i.e., lower, average, or higher); in the future, we will be able to predict the potential medical expenditures of emergency department patients according to such a classification. **Result:** The decision tree consisted of multiple levels of branches and color blocks to present the output and the sequence of information gathered (e.g., length of stay > disease classification > mode of departure from the hospital > triage > medical specific) and reflected the degree to which the distribution of medical expenses were influenced. **Conclusion:** This research suggests that the hospital can supply professional and personal services to various patients who have some special needs; at the same time, the hospital also can reduce the number of patients that re-visit the emergency department within 72 hours or remain in the emergency department > 24 hours, thereby decreasing the expenditures within the emergency department. (*Taiwan J Public Health*. 2006;25(6):430-439)

**Key Words:** *Emergency Medicine, Data Mining, Decision Tree*

<sup>1</sup> Graduate Institute of Health Care Organization Administration, College of Public Health, National Taiwan University, Taipei, Taiwan, R.O.C.

<sup>2</sup> Department of Information Management, National Chung Cheng University, No. 168, University Rd, Min-Hsiung, Chia-Yi, Taiwan, R.O.C.

<sup>3</sup> Taipei City Hospital, Ren-Ai Branch, Taipei, Taiwan, R.O.C.

<sup>4</sup> College of Medicine, National Taiwan University, Taipei, Taiwan, R.O.C.

\*Correspondence author. E-mail: mishgh@ccu.edu.tw

Received: Feb 14, 2006 Accepted: Oct 5, 2006