

## 壹、研究背景與目的

臺灣慢性腎臟病的好發率達 11.9%，平均每 10 人中，至少有 1 人罹患，經行政院衛生署（2010）統計顯示各縣市民眾因腎臟相關病症就診人數皆不在少數，國人從 1986 ~ 2007 年間因腎炎、腎徵候群、腎性病變等因素死亡者逐年明顯的增加中。依據中央健保局 2002 年的統計資料顯示，在健保的醫療支出中，洗腎經費占總醫療支出 7.21%，醫療費用支出均呈現成長趨勢（黃妙雲，2006）。因應臺灣腎炎、腎徵候群及腎性病變等病患有日益增加之趨勢，而腎炎、腎徵候群及腎性病變等病症早期症狀較不易發現，即使能夠早期發現病變，通常病患因自覺身體並無嚴重之不適情形，往往不會特別去注意及自我照護，進而導致病變情形快速惡化。本研究擬以應用資料探勘技術（決策樹及類神經網路）分析現有病患個案之資料，找出影響病患各階段變化顯著之要項，提供醫護人員做為診療、照護行為時之相關數據，以利預測病患各階段的演變時程。

## 貳、文獻探討

### 一、慢性腎臟病之定義

腎臟病的種類繁多，較常見的有免疫傷害引起的腎絲球腎炎以及與細菌感染有關的腎盂腎炎等；另外，糖尿病、高血壓及全身性紅斑性狼瘡等病人也常併發腎臟病變。慢性腎功能障礙（統稱慢性腎衰竭）是所有腎臟病的共同結果，此病是一種所謂的「進行性」疾病，也就是說一旦診斷確定以後，只會惡化不可能痊癒。影響腎功能的危險因子中，藥物占有極重要角色，其中較主要的是一些止痛劑及抗生素（陳鴻鈞，2001）。腎臟受損超過 3 個月，導致其結構或功能無法恢復正常，便可稱為慢性腎臟病。慢性腎臟病分為五個階段（如表 1），這個過程可能非常長久；也有可能很快地進入了階段 5——末期腎衰竭，甚至要進行透析或換腎的治療（臺灣腎臟醫學會，2006）。

### 二、資料探勘之定義

所謂的資料探勘是一大量自動化的過程，其運用統計分析從大量資料庫中挖掘出潛在、非顯然的、未知的、潛在的「可能」有用資訊之過程（Frawley, Paitetsky-Shapiro, & Matheus, 1991）。而 Grupe 與 Owrang（1995）則認為資料探勘是指從已經存在的資料庫當中挖掘出專家仍未知的新事實；Fayyad（1996）

表 1

慢性腎臟病的五個階段

期數	病理及併發症	腎絲球濾過率	症狀及因應之道
1	腎功能正常但併有蛋白尿、血尿等腎臟損傷狀況	腎絲球濾過率 GFR 90 ~ 100 ml/min/1.73 m <sup>2</sup>	腎臟功能約正常人 60% 以上，注意是否有糖尿病及高血壓，需要控制血糖、血壓與飲食，每半年作腎功能檢查，一般皆能穩住腎功能。但若有腎絲球腎炎之病患必須接受治療
2	輕度慢性腎衰竭，但併有蛋白尿、血尿等	腎絲球濾過率 GFR 60 ~ 89 ml/min/1.73 m <sup>2</sup>	
3	中度慢性腎衰竭	腎絲球濾過率 GFR 30 ~ 59 ml/min/1.73 m <sup>2</sup>	腎臟功能約正常人 15% ~ 59%，積極配合醫師治療，減緩進入第五期腎臟病變
4	重度慢性腎衰竭	腎絲球濾過率 GFR 15 ~ 29 ml/min/1.73 m <sup>2</sup>	
5	末期腎臟病變	腎絲球濾過率 GFR < 15 ml/min/1.73 m <sup>2</sup>	腎臟功能剩正常人 15% 以下，若逐漸無法排除體內代謝廢物和水分，則必需準備與接受透析治療及腎臟移植

資料來源：引自何謂腎臟，臺灣腎臟醫學會，2006，取自 [http://kidney.tsn.org.tw/1\\_1.html](http://kidney.tsn.org.tw/1_1.html)

則定義知識發掘 (knowledge discovery) 為從大量資料中選取合適的資料，進行資料處理、轉換等工作，再進行資料探勘與結果評估的一系列過程，也就是說，資料探勘只是知識發掘過程當中的一個步驟。定義資料探勘為使用自動或半自動的方法，對大量資料作分析，找出有意義的關係或法則 (Berry & Linoff, 1997)。Akaka (2004) 提出資料探勘是一種應用資料庫的技術，如統計分析與建立模型，用以發現資料中隱藏的模式與隱約的關係，並進行推論以預測未來結果。美國人工智慧協會 (American Association for Artificial Intelligence [AAAI], 2006) 近期指出，資料探勘是一種很強大的人工智慧工具，它可以從資料庫中發現有用的資訊，並可用於改善商業行為與決策模式。

### 三、決策樹

決策樹是一種預測模型；含有根部 (root)、子節點 (child node) 及葉部節點 (leaf node)。「根部」：資料從根部的節點進入決策樹；「子節點」：每一個節點代表「是」或「否」的問題點，答案便是前往下一個問題的前進路徑。決策樹是資料探勘中一個普通的方法。每個決策樹可以靠資料庫的分割進行數據測試。這個過程可以遞歸式的方式對樹進行修剪。當不能再進行分割或一個單獨的類應用於某一支時，遞歸過程就完成了。決策樹是功能強大且相當受