

研究紀要

# 就醫行為的潛在類型與健康測量： 全民健康保險資料的應用

張苙雲 楊孟麗 趙景雲

張苙雲 中央研究院社會學研究所研究員（[gacloud@gate.sinica.edu.tw](mailto:gacloud@gate.sinica.edu.tw)）；楊孟麗 中央研究院調查研究專題中心助研究員；趙景雲 中央研究院社會學研究所研究助理。這篇論文是第一作者為計畫主持人之國科會計畫（NSC 95-2627-H-001-006-MY2）之部分成果。

本研究資料來源為衛生署中央健康保險局提供、財團法人國家衛生研究院管理之「全民健康保險研究資料庫」，以及行政院衛生署國民健康局及財團法人國家衛生研究院提供之「國民健康訪問調查」原始資料檔。文中任何闡釋或結論為作者之觀點，不代表經費支持機構和資源提供機構之觀點。

收稿日期：2010/7/13，接受刊登：2011/2/16。

## 中文摘要

健康測量是研究健康不平等課題的核心，長久以來，學界對自陳評量的效度，保持某種程度的批判，並用以反省健康不平等的研究發現。然而，因著就醫紀錄的不易取得和複雜性，學界仍以受訪者自陳資料，作為衡量個人健康的主要根據，甚少使用就醫紀錄作為健康測量指標。本文建議運用潛在側圖分析法，分析就醫紀錄，從行為面推測健康情況，以解決就醫紀錄過於複雜的挑戰。本研究使用的資料有二：具全國代表性的「2000年全民健康保險承保抽樣歸人檔」及「2001年國民健康調查」。「承保抽樣歸人檔」近93萬樣本於2001年全年的就醫紀錄之側圖分析，結果顯示四組的分類是最適模型，依其就醫特徵與疾病診斷，其健康等級依序為「零就醫組」、「門診使用組」、「短期住院組」和「重度使用組」。文中除了討論分組結果的信度外，並將此分組與疾病診斷、就醫文獻相比對，以檢驗其效度。最後，受惠於全民健保資料和調查訪問資料的串連，將分組結果與調查所得之自評健康進行比較，彰顯本分析結果在理論和政策應用上的潛力。

**關鍵詞：**健康測量、醫療利用、健康不平等、「承保抽樣歸人檔」、潛在側圖分析

## **Latent Healthcare Utilization Classification and Objective Health Measures**

Ly-yun CHANG

Institute of Sociology, Academia Sinica

Meng-li YANG

Center for Survey Research, Academia Sinica

Jing-yun JHAO

Institute of Sociology, Academia Sinica

### **Abstract**

Health measures sit at the core of health inequality studies. Although objective, medical care utilization is seldom used as a measure of health status because of the unavailability of data and the obscure relationship between medical care utilization and health. A latent profile analysis technique was used to classify medical care utilization, and in turn to investigate the validity of different utilization types as ranked measures of general health. Data were taken from Large Health Insurance Dataset of 2000” (LHID 2000) and the “National Health Interview Survey for 2000 (n=923,892). Four membership groups were identified: “zero utilization,” “outpatient visits only,” “short term hospitalized,” and “heavy users.” Discrepancies between latent types and self-reported health data from a National Health Interview Survey (n=19,732) were noted in terms of age, gender, education, employment status, and income. The validity of classification as a rank order of health status and potential applications are discussed.

**Keywords:** Health measures, health care utilization, health disparities, latent profile analysis

## 一、前言

健康不平等文獻中，就醫行為一直很少用來作為測量個人健康狀況的指標。究其最直接的原因，在於就醫類型與頻率和健康之間的對應關係，並不明確。雖然一般人對門診、住院、急診及手術等就醫行為背後的疾病嚴重程度，不是沒有看法，例如住院相對於看門診，前者的疾病嚴重些。然而，就醫的頻率和類型所形成的各種組合，使得就醫行為和健康之間的對應關係，顯得複雜而難以有效斷定。何況就醫行為的出現不純然因為生病，有可能混含著其他社會面向的因素：國內外就醫行為文獻均指出，除了健康出現狀況以外，時間、醫療資源的方便性、個人性別、年齡、家庭結構、教育程度或收入、社會網絡等，非關個人健康的因素，都影響個人是否就醫及就醫的頻率（Dracup, Moser, Eisenberg, Meischke, Alonzo and Braslow 1995; Van der Meer, Van den Bos and Mackenbach 1996; Dunlop, Coyte and McIsaac 2000; Van der Heyden, Demarest, Tafforeau, Van Oyen 2003; Ahmad, Dag and Kurt 2004; Deri 2005; Curtis and MacMinn 2008；林芸芸 1992；楊文山 1992；劉仲冬 1997）。

然而，雖然時間、金錢、醫療資源等因素可能使需要就醫者沒有就醫，但這並不等於「時間多、錢多、醫療資源多，就會（沒事）也去看病」。求醫畢竟是有成本的行為，必須耗費體力、時間與金錢。雖然回應身體不適的方式，的確可能受個人的條件和外客觀資源的影響，但是求醫行動的出現，本質上是個人對自身健康警訊作出的回應。個人的健康狀況架構著當事人的求醫模式，所以研究者觀察到的就醫次數和類型的組合，可以合理推論是個人意識到身或心出現或大或小的異常狀態時所作的反應。循此思路，有別於就醫行為文獻的取徑，我們不從醫療可近性的角度切入探討，而試圖綜合運用就醫頻率和類型及其組合的資

訊，以指涉健康程度，並探討它和疾病的關係，以及與主觀的健康自評的關聯，使我們有機會一方面能拓展就醫資訊的應用潛力，另一方面檢討或回應健康不平等研究領域中對核心健康的測量。

本研究以國家衛生研究院釋出近百萬人的隨機樣本在2001年的就醫全紀錄為分析資料，根據這近百萬人的就醫類別、頻率和組合，使用潛在側圖分析方法（latent profile analysis，或稱潛在類別群聚分析〔latent class cluster analysis〕，或有限混合模型〔finite mixture model〕）（Vermunt and Magidson 2002），將之分組（membership identification），並解讀各組成員就醫資訊，推論其健康排序，成為行為面的健康測量指標。所得之健康指標將與疾病診斷比較，並檢討內部的異質性，以及其蘊含的理論與政策的意涵。

## 二、健康測量的文獻與挑戰

個人的健康狀況攸關其幸福與生產力，是研究的主要依變項。世界衛生組織則將健康定義為「個人生理、心理和社會面向的安適狀況」，而不只是狹義地根據「沒有生病」來界定是否健康。這定義雖是學界的共識，但是要根據這個定義衡量一個人是否健康，卻不是件容易的工作。因為健康是多面向的「事實狀態」，「多面向」的部分，可以透過不同面向題組設計來達成；更根本的則是「事實狀態」的呈現，必須依賴當事人自己，或透過口頭表達，或透過表現出的行為，才得以推斷（Bergner 1985; Bergner and Rothman 1987）。這些外顯的口述或行為，或透過就醫而由醫療專業人員記載於病歷的資訊，或透過調查研究，而記載於問卷。嚴格來說，無論是病歷資料或調查結果，都是當事人自陳的結果，只是前者經過專業人員的判斷。透過調查方法，直接向

受訪者取得健康（疾病）相關資訊，一般稱為自評健康。自評健康的內容，可以僅止於過去某段時間的大致健康狀況，也就是主觀綜合性健康評量。相對於這種綜合性單一題目的評量，過去30年以來，學者另外發展了不少包括身體功能表現、疼痛情形、心理健康、或日常生活活動等多面向的健康評量，以回應單一面向測量的不足。較為著名的量表包括「日常生活的活動能力」（Activities of Daily Living）（Katz, Ford, Moskowitz, Jackson and Jaffe 1963）、“Sickness Impact Profile”（Bergner, Bobbitt, Carter and Gilson 1981）、“Nottingham Health Profile”（Hunt, McEwen and McKenna 1985）、“SF36”（Ware 1988）、「身體功能測驗」（Physical Performance Test）（Guralnik, Branch, Cummings and Curb 1989），以及“Health Utility Index Mark III”（HUI3）（Furlong, Feeny, Torrance, Goldsmith, DePauw, Zhu, Denton and Boyle 1998）等。綜合而言，涵蓋面向愈廣且題數愈多，愈能用多重指標來逼近健康的「真實」狀況，測量結果的信度與效度就比較理想（McHorney 1997; Garratt, Schmidt, Mackintosh and Fitzpatrick 2002; Lindeboom and Van Doorslaer 2004; Terwee, Bot, De Boer, Van der Windt, Knol, Dekker, Bouter and De Vet 2007；盧瑞芬、曾旭民、蔡益堅 2003）。

疾病的有無一般被視為較客觀<sup>1</sup>的自評健康測量，透過調查訪問，研究者詢問受訪者是否罹患任何疾病或慢性病，或詢問疾病的名稱，或有多少種疾病等。但從疾病的有無界定健康時，由於疾病種類繁多，各類疾病和健康之間的關係不一，甚至即使同一疾病，每個人的嚴重程度也可能差距極大，很難作跨疾病之間健康程度的比較。即或調查問卷能詳細完整紀錄受訪者所罹患的各種病症，研究者仍難以利用疾病的有

---

<sup>1</sup> 之所以用「較客觀」，是因為，只要經過當事人的陳述，就不能全然排除主觀的成分，且民衆自陳罹患病名，可以合理推斷經過醫療專業人員的確認。

無、數量或種類，綜合成為可以跨疾病比較的健康測量。因此，研究者多以單一疾病為研究（如Johnston, Propper and Shields 2009），或以所罹患疾病的總數作為衡量健康程度的指標。

無論單題式的綜合評量、多面向量表形式、單一面向、或罹病數、就醫行為等，都是透過面訪或電訪由受訪者自述。然而，對於調查所得到的健康資料的困境，學界早有認知。首先，調查訪問先天就已排除因健康狀況極度不佳而不能接受訪問的人。其次，受訪者的記憶因素與理解力、及調查結果是否涉及受訪者的利益等等因素，都可能使自評健康產生測量誤差（Krause and Jay 1994; Crossley and Kennedy 2002; Singh-Manoux, Martikainen, Ferrie, Zins, Marmot and Goldberg 2006）。自評的主觀評量和較為具體客觀的健康測量，分別有不同的測量誤差（Bound 1991），且自評健康會因為性別和年齡而產生位移現象（Lindeboom and Van Doorslaer 2004），亦即不同性別或不同年齡層，對於健康的標準不同。第三，調查實務上，總有問卷空間和訪問時間的限制，往往不能全然蒐集到研究者所需的各種資訊。第四，即或調查訪問能涵蓋多重健康資訊，這些資料之間的關係，以及它們和社會人口變項之間的關係，往往有程度和方向不一的問題，使得研究者在綜合研究發現時，遇到很大的挑戰，不易解讀研究結果。以Culter及Lleras-Muney（2006）的研究為例，他們的重點是教育對健康的影響，使用1990、1991和2000年美國National Health Interview Survey中的20項健康資訊（包含綜合主觀評量、功能限制、疾病診斷或死亡等），分析結果顯示，這20項健康測量與教育呈現不同方向或強度的關係，且其間的機制亦有所不同，這使研究者面臨理論收斂上的高度挑戰。

如前所言，許多原因，像是就醫的非疾病因素多、就醫資料不易取得，及就醫紀錄的使用不易等，使得研究者在作健康不平等的跨群體或

跨國比較時，還是以主觀的健康量表或死亡作為依變項，而很少直接利用疾病診斷和就醫的資訊當作依變項。健康不平等領域面對的課題是，儘管個人的自評健康對於極端的結果，如死亡，的確有良好的預測力（Idler and Benyamini 1997; DeSalvo, Bloser, Reynolds, He and Muntner 2006; Huisman, Lenthe and Mackenbach 2007），但學界仍然對主觀自評的信度和效度，保持高度批判態度（Baker, Stabile and Deri 2004; Subramanian and Ertel 2008），因為，研究發現，個人不但對健康的主觀定義與要求不一，也因年齡、教育程度和族群而有所不同（Krause and Jay 1994）。例如，長期追蹤調查中，年紀大的人，在自評健康上顯得穩定，似乎表示健康狀況一如過去（Leinonen, Heikkinen and Jylhä 1998），但實際健康狀況卻已經大不如前。學者認為，這是因為人們在評量自己的健康狀況時，常以想像中（較不健康）的同儕情況作比較，先行調整比較基準（Suls, Marco and Tobin 1991）所致。

事實上，實際罹病狀況和主觀感受之間的落差，不一定就該以實際罹病為基準線，而認為是主觀評量的誤差。我們認為，二者之間出現落差，不必然是對錯與否的問題，反而需要更深一層問：為什麼有些人頻頻看病，卻自覺健康狀況良好？又為什麼有人很少看病，卻說自己健康狀況不好？在公衛政策、社會學、或心理學上，這些現象有什麼意義？這些角度因為資料不易得之故，目前文獻較少討論，卻是值得發展的方向。

國際上在檢討各種測量指標時，無論是透過實驗方法、深度訪談或利用與故事中人物的比較（vignettes），基本上都是在受訪者自陳的架構上進行，很少能與「客觀」的病歷資料比對，只有Baker等人（2004）所作的研究例外。但是，就理解民衆自評健康資訊與實際行為之間的一致／差異這個課題而言，臺灣目前有很大的優勢。拜全民健保

之賜，健保資料庫有完整的民衆就醫全紀錄，且在保護個人隱私的條件下，有系統地串連健保資料和調查訪問資料，開放給學界分析。健保資料的就醫全紀錄，排除了自評健康的主觀與自陳記憶的問題，也將因健康不佳而無法受訪的人納入資料，避免了調查可能有的樣本偏誤問題。

只是，研究者若要充分運用這些資料，面臨的挑戰是：如何化繁為簡，以就醫行為的各類組合和頻率為分析資料，將底層所蘊含的健康狀態，簡要地抽離出來。事實上，長期以來，「求醫行為」或較窄的「醫療利用」都被放在「理性行動理論」（Theory of reasoned action）（Fishbein and Ajzen 1975）的架構下作詮釋（Kleinman, Kunstadter and Gale 1974; Kleinman 1980；吳就君 1981）。細緻的就醫歷程分析指出，就醫時機的選擇有相當「權變」的意味：除了問題的持續且惡化外，真正能發揮示警作用的，主要還是在明顯可見的症狀（張苙雲 1998）。雖然就醫行為的文獻指出時間、醫療資源的方便性等非關健康的因素，會影響就醫的可能性和就醫類型，然而，既然症狀的產生啟動了人們的求醫行為，且就醫本身是有成本的行動，需要人力和時間的投入，那麼就醫頻率和類型的資訊，應可用以推論個人健康狀況。

問題是，如何從複雜的就醫資料，辨明比較各人的健康狀況。就醫行為，包括門診、住院、急診、或是手術等就醫類別，雖不若疾病種類繁多複雜，但可能因疾病的嚴重和緊急程度，甚至個人的就醫習慣，而產生各種組合。例如小恙可能掛門診，門診次數多可能表示情況較為嚴重；大病則需要住院甚或開刀，住院天數或開刀次數也反映疾病的嚴重程度。但有時又似乎難以利用就醫行為論斷疾病的嚴重程度。例如：門診次數多和短期住院，哪一種的健康狀況比較差？

基於此，本文提出以潛在側圖分析方法，希望充分發揮臺灣全民健保擁有就醫全紀錄的優勢，嘗試以之分類健康狀況。本文的分析是在

兩個簡單的預設下鋪陳的：預設一是，健康是個隱性變項（unobserved variable），其分配（distribution of health levels）是未知且無法直接觀察。預設二是，真實健康的尺度是單一的（a unique metric）。這個未能直接觀察的健康變項，反合性地（paradoxically）具有多面向的可能，隨著個人特質、社經條件、居住地點等而有不同的分配。在這兩個預設下，本文將以全民健保資料就醫全紀錄，利用潛在側圖模型，依據樣本的就醫行為頻率和類型的組合，將樣本類型化，並透過與自評健康狀況的比較，探索兩者間的差距及其相關變項，及討論分析結果所顯示對未來研究的重要性。

### 三、資料與分析方法

#### （一）資料

本研究的資料來源有二，其一是中央健康保險局授權國家衛生研究院（以下簡稱國衛院）對外公開釋出的全民健康保險「承保抽樣歸人檔」（以下簡稱「承保歸人檔」）。其二，為示範這種以就醫行為所定義的健康程度，可以用來跟自評健康作比較，並瞭解造成其中差異的因素，我們採用國民健康局（以下簡稱國健局）2001年的「國民健康訪問調查」（以下簡稱為「健康調查」）資料。並且，為配合國健局的資料調查時間，我們以「承保歸人檔」樣本於2001年的就醫資料作分析。

全民健保資料原本是為了醫療費用申報之用而建立。由於我國目前全民健保納保率達到99%以上，使得健保資料成為醫藥衛生相關領域研究中具有代表性的實證資料。因此，國衛院建置「全民健康保險研究資料庫」，自2000年起開始提供學者研究之用。提供的檔案主要分為原始

就醫資料檔與基本資料檔兩大類。

本研究使用的2000年「承保歸人檔」，包含個人醫療診斷、處方紀錄及其相關費用申報。該檔係以2000年納入全民健保的人作為抽樣母群體，隨機選取100萬人，並擷取這些人在1996年至2008年間所有的就醫資料。根據國衛院的分析，此樣本與抽樣母群體，及內政部公布之2000年人口資料，不論在年齡層分布或性別比率上，都具代表性，其詳細抽樣程序以及樣本代表性檢定請參閱國衛院網站（財團法人國家衛生研究院 2010）。

2000年「承保歸人檔」的母群體中，不同的兩人，約有千分之一的機會可能共用一個身分辨識代碼（簡稱ID），但兩人的生日與性別不同。我們檢查本文所使用的資料，發現共用ID的比率約千分之二。因此，本研究在串連個人承保資料與相對應的健保就醫紀錄時，是以「承保資料檔」（亦即ID檔）裡的樣本為依據，合併利用ID、生日及性別來進行串檔，避免合併不同人的資料。由於國衛院所釋出之「承保資料檔」2000和2002年是累積性的歷史資料，隨保險對象投保身分別的改變或工作單位的異動而有轉出、轉入的紀錄，意即自健保施行開始的每一筆投保、退保紀錄皆逐年保存於承保檔之中，每一個人可能有多筆紀錄。由於缺少2001年之「承保資料檔」，本研究根據此一特性，自2002年的承保資料檔選取於2001年間曾有承保紀錄之ID作為2001年之分析樣本，也就是排除2001年之前轉出和2002年始轉入者之紀錄，最終得到有效樣本數為923,892人<sup>2</sup>，其中男性占50.7%。受訪者年齡介於0歲至105歲之間，平均約為34歲。

國民健康局2001年的「健康調查」，使用的是12歲以上受訪者的自

---

<sup>2</sup> 我們另外以2000年之「承保資料檔」保留2000年時承保紀錄，得到929,040人。資料使用者亦無法瞭解何以如此。

評健康和他們當年的全民健保就醫資料。2001年「健康調查」抽樣報告指出，「健康調查」所抽出的樣本，在同時考量性別及年齡的情況下，跟當年全國戶籍資料所呈現的分布狀況無顯著差異（NHIS抽樣工作小組 2001: 5）。這項資料有效樣本共19,588人，女性占50.4%，受訪者平均約40歲，年齡介於12歲至98歲之間。

## （二）使用變項

健保資料庫的就醫資料包括門診、門診急診、門診手術、是否急診住院、是否住院開刀，及住院天數等六項。在計算門診次數時，我們作了多次探索性的嘗試，最後將預防保健、預防性疫苗接種、牙科的洗牙與氟化防齲處理、孕婦的產前檢查與正常生產、65歲以上老人流感疫苗接種、及非醫療目的之選擇性手術等六個類別排除於門診的計算之外，但保留有上列情況且同時治療其他疾病的門診資料。然而，資料庫的資訊並無法將牙科的牙齒矯正與皮膚科的美容等項目篩選出來，這是本研究的測量誤差。但這個誤差也許影響不大，因為我們發現，即使不排除健康檢查及洗牙等非疾病的就醫，潛在側圖分析所得的結果與本文所報告的結果，相似度達99.33%以上。

再者，未呈現於文中的初步分析結果顯示，門診急診次數本身在分析上並無區辨力，也就是，分析所得的各組，在該變項的描述統計值都極為類似。從實務上思考，的確，經急診門診求醫者，固然有亟需緊急醫治者，但為省時而掛急診的，也並非沒有。因此，我們認為門診急診不盡然代表疾病的嚴重性，故而將它與一般門診合併計算。最後，門診手術次數也發現無區辨力，故亦併入門診次數計算。住院天數的計算是以個人為單位，同時考量其入院、出院日期，排除重疊時間，以避免重

複計算。

綜合而言，經過多次探索性的嘗試後，我們將門診次數、門診手術次數與急診次數合併為一，再加入住院天數、是否急診住院和住院開刀等三項，總共四項就醫行為作為指標進行分析。

表1列出2001年「承保歸人檔」中，各項就醫指標之描述統計值。這近93萬人在2001年時，門診次數最高達384次，亦即有人平均一天不止看1次門診，最低0次，每人平均看診次數13.65次；急診住院比率是1.56%；45,455人（4.92%）曾動過住院手術；住院天數最高達364天，平均住院天數1.06日（表1）。相對的，2001年「健康調查」受訪者平均每人門診次數約12.41次，最高門診次數達266次（表1）。有1.56%的人曾經由急診而住院，4.67%的人曾動過住院手術，住院天數最高達316天，平均住院天數約0.6天。從表1可看出，接受調查的受訪者，其醫療

**表1 全民健康保險「承保抽樣歸人檔」2001年樣本與2001「國民健康訪問調查」樣本之各變項描述統計**

人口變項	承保歸人檔	健康調查	醫療利用指標變項	承保歸人檔	健康調查
年齡			01年門診次數		
平均值	33.93	39.63	平均值	13.65	12.41
標準差	(20.28)	(17.73)	標準差	(14.03)	(14.87)
全距	0-105	12-98	全距	0-384	0-266
女性	49.30%	50.36%	01年急診住院比例	1.56%	1.56%
受雇		52.37%	01年住院手術比例	4.92%	4.67%
教育年數			01年住院天數		
平均值		9.81	平均值	1.06	0.64
標準差		(4.31)	標準差	(10.55)	(5.38)
全距		0-18	全距	0-364	0-316
個人月收入					
平均值		19690.37			
標準差		(22078.46)			
全距		0-100000			
樣本數	923,892	19,588		923,892	19,588

利用行爲，的確比「承保歸人檔」百萬樣本爲低。

作潛在側圖分析時，門診次數與住院天數，都以連續性變項視之。是否動過住院手術則爲二分變項，“1”代表有手術，“0”代表無，是否急診住院亦爲二分變項，“1”代表曾經由急診而住院，“0”則無。

### （三）分析方法

本文使用潛在側圖分析，將個人的就醫行爲作爲指標，以此分類不同的群體。潛在側圖分析是潛在類別模型（latent class models）的一種，它與潛在類別群聚分析（latent class cluster analysis）的差異是，前者使用的表現變項（manifest variable，也就是指標）可以是連續的，而後者的表現變項都是類別變項。兩者分析得到的結果，則都是不同類別的次群體。在真實的世界裡，許多群體內實際有不同類型的次群體，這些次群體就稱爲潛在側圖或潛在類別。作研究分析時，若忽略這種異質性，得到的結果就可能有誤導之虞。所謂不同類型，是指在某些表現出來的性質（即表現變項）上有差異。潛在側圖模型即在分析這些異質性的來源（表現變項），以找出這些次群體，使得每個次群體具有同質性。因此，它有兩個重要假定：（1）觀察資料混合了來自不同母群的幾個次樣本（即群組），且這些母群的參數（如平均值、標準差等）不同，故樣本在指標的表現也不同（McLachlan and Peel 2000），統計方法可以將這些次樣本分辨出來。（2）研究者知道母群的個數，因此也知道可以分成多少個次樣本。事實上，第二個假定幾乎難以成立，因爲大多數的研究者事前並無法確知母群個數。所以，實際的分析方式是由研究者嘗試指定不同個數的次樣本，再比較不同個數的次樣本所獲得的模型適合度，及參考實務需要，藉以決定分成多少個次樣本最佳。

潛在類別分析已經用於多個實務相關的領域，例如在行銷領域中找出不同經營策略型態的公司類別（DeSarbo, Benedetto, Jedidi and Song 2006）、或找出不同的顧客偏好類型（Desarbo, Jedidi and Sinha 2001）；在農業經濟學領域中用於找出買主對於牛隻特性的喜好類型（Ruto, Garrod and Scarpa 2008）；在土地經濟學領域中用於找出攀岩者的喜好類別（Scarpa and Thiene 2005）；在精神病學領域中用於找出不同類別的飲食異常者（Wade, Crosby and Martin 2006）。

## 四、分析結果

我們以「承保歸人檔」的樣本在2001年全年的門診次數、住院天數、是否曾經急診而住院、及是否曾做過住院手術等四項就醫資料，利用潛在側圖分析法，獲得六種模型，第一個模型只有一組，第二個模型有兩組（兩個次樣本），以此類推。

首先，我們參考各模型所產生之各組的百分比和BIC（Bayes' information criterion）值的變化，以判斷各模型的適合度。文獻指出，潛在側圖分析所產生的若干次樣本之分組結果，BIC或CAIC（consistent Akaike's information criterion）（Bozdogan 1987）是選取模型適合度的重要參考標準（Fraley and Raftery 1998）。BIC與CAIC兩者都是以模型的log-likelihood為基礎，將模型所必須估算的參數及整體樣本數納入考量。事實上，BIC和CAIC兩者的表現通常非常一致（Jedidi, Harsharanjeet and DeSarbo 1997）。雖然所指定的次樣本個數多，可降低模型錯誤，但需要估計的參數卻愈多，模型愈複雜，不夠精簡。BIC與CAIC的意義即在正確率與精簡度之間取得平衡（Jedidi et al. 1997）。

表2列出各模型下各組的人數分布以及BIC值。參考不同模型下各組

所占百分比的變化和BIC降低的速度，我們發現，持續增加組數，最大一組的人數在第三個模型和第四個模型時幾乎維持不變，而是在第三組作了進一步的細分。但是在第五個模型時，各組人數的百分比出現很大的變化，似乎將第一組作了進一步的細分；第六個模型時，各組百分比又呈現相對穩定。也就是說，人數最多與次多的兩組所占的比率，在第五個模型時已相當穩定。各模型BIC降低的幅度愈益縮小，但並無回升（表2）。合併參考這兩個資訊，似乎第五個模型較為適當。

表2 潛在側圖分析之模型適配度：「承保抽樣歸人檔」2001年全年就醫紀錄  
(N=923,892)

模型	各組所占人數%						BIC	Entropy R-squared
	第一組	第二組	第三組	第四組	第五組	第六組		
一組	1.0000						14989457.15	1.0000
二組	0.9214	0.0786					2528740.75	0.9934
三組	0.8294	0.0921	0.0784				2047052.39	0.9804
四組	0.8292	0.0926	0.0594	0.0188			1777462.50	0.9739
五組	0.5039	0.3267	0.0929	0.0598	0.0167		1422823.63	0.7567
六組	0.4993	0.3308	0.0931	0.0475	0.0236	0.0057	1353559.91	0.7539

作潛在側圖分析後，最適模型的選擇，除了選取BIC或CAIC值最低的模型外，還要同時以「後驗機率」（posterior probability）所形成的分類交叉表和熵值（entropy R-squared）來評量模型分類結果是否適當。「後驗機率」是指每筆資料被分派到各個次樣本的機率。假定資料檔是由五個次樣本所形成，則每一個觀察值都有五個後驗機率（五者加總等於1）。以最高的後驗機率決定屬於哪個次樣本（組），這種分組標準稱為模型分組。根據分組的結果，可以計算出屬於同一組的人在各次樣本的后驗機率的總和，稱為機率分組。模型分組和機率分組交叉後，就是分類交叉表（classification table）。而熵是以統計值的方式代表分類交叉表所傳遞的訊息，熵值愈大表示錯誤率愈小（Vermunt and

Magidson 2005: 62)。

從表2各模型的統計值來看，第三個模型與第四個模型，熵值都在0.95以上，表示錯誤率相近。但第五個模型及第六個模型時，熵值大幅降低至0.76，錯誤率將近25%。綜合參考四個資訊，我們認為第四個模型應是最佳的選擇，並據以進一步觀察第四個模型的模型分組和機率分組所構成的分類交叉表之表現。

分類交叉表的資訊包括模型分組的各組人數，及各個次樣本的後驗機率總和的分配情形。若某一組在其他次樣本的後驗機率總和愈低，表示被分派到該組的研究對象，被認定為屬於該母群的機會愈高，分類也就愈正確（Pastor, Barron, Miller and Davis 2007）。表3是第四個模型機率分組與模型分組的分類交叉表。模型分組的人數，從第一組到第四組分別為764,850、86,923、55,746和16,373人。第一組的人，在四組的後

**表3 潛在側圖分析模型分組與機率分組交叉表：「承保抽樣歸人檔」  
2001年全年就醫紀錄**

模型分組 \ 機率分組	第一組	第二組	第三組	第四組	機率分組 總和
第一組	764,569.14521 (99.80%) (99.96%)	1,351.1311 (0.18%) (1.55%)	0 (0.00%) (0.00%)	195.2567 (0.03%) (1.19%)	766,115.5329 (100%)
第二組	0 (0.00%) (0.00%)	85,571.85571 (100.00%) ( 98.45%)	0 (0.00%) (0.00%)	0 (0.00%) (0.00%)	85,571.8557 (100%)
第三組	40.5776 (0.07%) (0.01%)	0.0129 (0.00%) (0.00%)	54,174.5627 <sup>1</sup> (98.74%) (97.18%)	649.4763 (1.18%) (3.97%)	54,864.6294 (100%)
第四組	240.2772 (1.39%) (0.03%)	0.0004 (0.00%) (0.00%)	1,571.4373 (9.06%) (2.82%)	15,528.2671 (89.55%) (94.84%)	17,339.9820 (100%)
模型分組人數	764,850	86,923	55,746	16,373	923,892

<sup>1</sup> 有灰色底的組別，是模型分組與機率分組一致性最高之組別。

驗機率總和分別是764,569.1、0、40.6和240.3；也就是說，被歸類（模型分組）到第一組的人，其在第一組的機率總和是764,569.1，因此模型分組屬於第一組的人，其被歸類為第一組的平均機率是99.96%（表3第二欄第四格）；而這一組的機率總和占了所有第一組機率的99.80%（表3第二欄第三格）。以此類推。從表3可看出模型分組與機率分組之間的一致性頗高。實際上分派樣本時，是以模型分組的結果行之。

當然，與其他傳統統計方法一樣，在決定次樣本的個數時，必須同時考量理論和次樣本獨特性質。表4列出第四個模型中，各次樣本的就醫行為的特性，包括該行為的平均值與標準差、及四分位數（包括第25百分位數的人（Q1）和第75百分位數的人（Q3））在該醫療使用變項的數值。潛在側圖分析的結果呈現，係以人數多寡決定組別的呈現次序，組內人數最多的是第一組，次多人數的是第二組，以此類推。但為使後續分析結果的呈現，是依據可能的健康程度排列，我們將表3的第一、二組互換位置，分別命名，列於表4。

表4第一組是完全沒有使用任何醫療資源者，共86,923人，占9.41%，稱為「零就醫組」（即表3的第二組）。其餘三組都使用不同頻率和類別的醫療資源，分別命名為「門診使用組」、「短期住院組」和「重度使用組」，原因說明如下：首先，從門診使用次數的百分位數來看，「門診使用組」和「短期住院組」的門診次數之全距頗接近，分別為1至79和0至76，「重度使用組」的門診次數全距則在0至384之間。雖然各組在這個變項的全距頗寬，但主要的範圍相對頗為集中（觀察Q1和Q3），且各組的平均門診次數之間有顯著的差異：「門診使用組」、「短期住院組」和「重度使用組」，一年的平均門診次數分別是13.82、22.61和47.40。

「門診使用組」共764,850人（82.79%），他們雖然平均使用了

表4 潛在四組之各項指標與人口變項：「承保抽樣歸人檔」2001年全年就醫紀錄

	零就醫組	門診使用組	短期住院組	重度使用組	總和
人數	86,923	764,850	55,746	16,373	923,892
百分比	(9.41%)	(82.79%)	(6.03%)	(1.77%)	(100.00%)
門診次數					
平均值	0.000	13.822	22.614	47.398	13.647
標準差	(0.000)	(11.980)	(14.383)	(36.065)	(14.033)
Q1 <sup>1</sup>	0	5	11	19	4
Q3 <sup>1</sup>	0	19	31	75	19
全距	0-0	1-79	0-76	0-384	0-384
急診住院比率	0.00%	0.00%	24.18%	5.92%	1.56%
動住院手術比率	0.00%	0.00%	63.19%	62.41%	4.92%
住院天數					
平均值	0.000	0.000	4.978	42.646	1.056
標準差	(0.000)	(0.000)	(3.220)	(66.373)	(10.550)
Q1 <sup>1</sup>	0	0	3	13	0
Q3 <sup>1</sup>	0	0	7	42	0
全距	0-0	0-0	1-16	0-364	0-364
罹患慢性疾病率	0.00%	35.02%	65.03%	88.97%	34.49%
罹患7大死因疾病數					
罹患率	0.00%	16.78%	40.35%	77.18%	17.69%
平均值	0.000	0.259	0.778	1.882	0.295
標準差	(0.000)	(0.660)	(1.150)	(1.478)	(0.740)
Q1 <sup>1</sup>	0	0	0	1	0
Q3 <sup>1</sup>	0	0	1	3	0
全距	0-0	0-7	0-7	0-7	0-7
女性 (%)	35.19%	50.98%	50.19%	42.79%	49.30%
年齡					
平均值	34.171	32.938	40.070	58.002	33.929
標準差	(15.931)	(19.989)	(24.051)	(21.600)	(20.278)
Q11	24	17	22	45	18
Q31	43	46	60	74	47
全距	0-105	0-105	1-103	1-104	0-105

<sup>1</sup> Q1是指第25百分位數者的值，Q3是第75百分位數者的值。

13.82次門診，接近全國的平均數，但這一組的人在2001年都不曾住院或手術；相對而言，「短期住院組」和「重度使用組」則都經歷天數不等的住院治療，且有相當比率用過急診住院或動過住院手術。

「短期住院組」的最大特色是：都曾至少住院1天，最高16天，平均住院約5天（4.978天），他們的急診住院比率是四組中最高，占24.18%。因之稱為「短期住院組」（55,746人，占6.03%）。

「重度使用組」不論在平均門診次數或平均住院天數上，都明顯比其他組高（16,373人，占1.77%）。這一組平均使用門診47.4次，平均住院42.65天；急診住院比率較「短期住院組」為低，但手術率則旗鼓相當（63.19%相較於62.41%），顯示這一組的人可能罹患較嚴重的疾病。

綜合而言，「承保歸人檔」近93萬人的樣本，依據他們在2001年所使用的醫療資源多寡，從不曾使用到重度使用，可以作為健康程度排序的推斷基礎：「零就醫組」的健康最好，「重度使用組」最差。「門診使用組」與「短期住院組」之間或有討論的空間，但從門診的平均值與數量分布情形來看，大致上可以說，「短期住院組」所呈現的健康狀態比「門診使用組」稍差。

## 五、結果評估與應用

利用潛在側圖分析法分析就醫資料所得的結果，在效度上可被接受嗎？能應用在哪些研究？我們首先檢驗各個次樣本之間重疊性的問題，並從疾病診斷資料觀察各組可能的健康等級，後將之與文獻中跟健康相關的社會人口特質作比對，以確認分組結果的有效性。接著，我們深入處理組內異質性的問題，並從中抽離出政策意涵。最後，我們根據就醫行為的分組結果與自評健康交叉比對，探討這項研究結果的應用潛力。

## （一）組間的重疊性是否影響分組之有效性？

「短期住院組」與「重度使用組」在各醫療利用行為上看似有重疊，令人擔心兩者是否有明確差異。回頭檢查原始的就醫資料，我們發現「重度使用組」中，若門診次數為0，其住院天數都比「短期住院組」的最高住院天數高，至少17天。且「重度使用組」沒有住院開刀的人，門診次數都超過80次，因此也跟「門診使用組」不重複。然而，住院天數在16天以內，且門診次數在76次以內者，若是因為急診住院，則被歸為「短期住院組」，否則被分到「重度使用組」；但這類處於兩組邊緣的情況者是極少數。

## （二）與疾病診斷之間的關係

為了確認這四組可視為四個程度的健康狀況，我們使用兩種診斷，分別是「是否罹患慢性病」及「罹患重大疾病的個數」，作為外部效度的參考。慢性病的定義是依據健保局97項慢性病之疾病定義，或 / 和治療過程中領有慢性病處方箋者；所謂重大疾病，則是根據行政院衛生署92年衛生統計的十大死因，排除肺炎、事故傷害、及自殺後的七大疾病，依序為：惡性腫瘤、腦血管疾病、心臟疾病、糖尿病、慢性肝病及肝硬化、腎炎腎徵候群及腎性病變、及高血壓性疾病。表4中，四組罹患慢性病的比率，分別是0%、35%、65%和89%。而患有上述七大疾病的百分比，依次為0%、16.8%、40.4%和77.2%，平均罹病的個數也在四組中依序增加。這些比率或平均值的數值都是由「零就醫組」的最小，到「重度使用組」最高，四組明顯不同，所指涉的健康程度愈益下降，與前述的各組就醫行為所指涉的健康程度一致。

### (三) 醫療使用情形與社會人口變項

文獻指出女性的就醫次數高，但活得長。前述健康分組，是否與文獻的發現一致呢？表4最下面兩欄列出各組的男女比率及年齡分布。表4中，「門診使用組」與「短期住院組」的男女比率相似；「零就醫組」與「重度使用組」則以男性比率較高，尤其前者更甚；似乎符合一般認為「女性比較愛看病或比較會照顧身體」的想像，及「男性常等到嚴重病徵出現才進醫院」的印象。其次，文獻指出健康在出生後隨著年齡增加而改善，中壯年之後健康隨著老化而逐年變差，至於健康狀況往下滑的起始年齡點，則因社會地位不同而不同，但無論如何，健康與年齡之間呈現的負相關，在這裡得到印證：各組的年齡範圍都非常大，但平均年齡的排序則頗為符合一般年齡與健康關係的想像：「零就醫組」和「門診使用組」的平均年齡最低（34.2/32.9）、標準差亦小（15.9/20.0），大致屬於青壯年階段；平均年齡次低的是「短期住院組」（40.1，標準差為24.1），屬於中壯年階段；最後，「重度使用組」的平均年齡則接近老年期（58.0，標準差為21.6）。

### (四) 「門診使用組」的異質性

即便四組的健康程度大致可從就醫模式與疾病罹患情形推斷，但是「門診使用組」占了將近83%的樣本，而同在這一組的人，門診使用次數差異卻非常大（1-79）。把一年內看1次門診和看30次門診，甚或79次門診的人，放在同一組，視為反映同等的健康狀態，似乎不盡合理，確有必要進一步剖析其間的異質性。我們從兩方面探討。

首先從就醫的疾病診斷切入。我們根據就醫紀錄中是否包含罹患

(1) 慢性病（以慢性病處方簽為準）以及(2) 十大死因中之七大疾病之任一種，作為分類的標準，把這一組再分為沒有慢性病也沒有重大疾病、只有慢性病、只有重大疾病、和同時有慢性病和重大疾病等四組，結果列於表5。從表5看出，「門診使用組」的76萬多人中，有62.8%沒有任何慢性病或重大疾病，他們是四組中看病次數最少的，平均一年10.25次，標準差為9.29；慢性病患者較重大疾病患者的平均門診使用率高一些，分別是17.43（標準差為12.10）和14.65（標準差為10.88）；患有兩種疾病的組，平均門診使用率則高達24.01次（標準差為14.58）。各組兩性的比率都非常近似，只有在慢性病一組中，女性比率明顯較高，與Case and Paxson（2005）的研究有類似的發現。再者，四組的平

**表5 門診使用組再依慢性病及七大疾病分組後之門診次數與人口變項分布情形**

	無	慢性病	七大疾病	慢性病及七大疾病	總和
人數	480,118	156,428	16,868	111,436	764,850
百分比	(62.77%)	(20.45%)	(2.21%)	(14.57%)	(100.00%)
門診次數					
平均值	10.252	17.434	14.648	24.010	13.822
標準差	(9.293)	(12.100)	(10.881)	(14.575)	(11.980)
Q1 <sup>1</sup>	4	9	7	13	5
Q2 <sup>1</sup>	7	15	12	21	10
Q3 <sup>1</sup>	14	23	19	31	19
全距	1-79	1-79	1-79	1-79	1-79
女性	49.59%	55.16%	49.51%	51.29%	50.98%
年齡					
平均值	26.817	34.744	44.211	55.072	32.938
標準差	(16.921)	(19.130)	(18.965)	(16.297)	(19.989)
Q1 <sup>1</sup>	13	20	31	44	17
Q2 <sup>1</sup>	25	35	44	56	31
Q3 <sup>1</sup>	38	48	58	68	46
全距	0-100	1-103	1-103	1-105	0-105

<sup>1</sup> Q1是指第25百分位數者的值，Q2是第50百分位數者的值，Q3是第75百分位數者的值。

均年齡也有明顯差距：沒有兩類疾病的組最年輕；患慢性病組次之；平均年齡最大的是患有兩類疾病的一組。然而，儘管平均使用次數有差異，四組的看診次數範圍還是很大（1-79）。

「門診使用組」中高達62.8%的人沒有慢性病也沒有七大疾病，但門診使用次數偏高；這些人在性別或年齡上有任何特性？我們以該組的平均門診使用次數加上兩個標準差（ $10.25+2*9.29=28.79$ ）為切割點，觀察門診使用29次以上者的人口特徵，並以使用次數低於平均值（10次及以下）及兩個標準差以內（11-28次）的兩組相比較。從表6的分類結果，我們發現，女性在較高使用次數的兩組（11-28次，及29次以上）有較高的比率，與就醫行為文獻的發現一致；這三組在年齡分布上亦有

**表6 門診使用組中無慢性病也無七大疾病者三等級的門診次數與人口變項分布情形**

	1-10次	11-28次	29-79次	總和
人數	305,900	149,853	24,365	480,118
百分比	(63.71%)	(31.21%)	(5.07%)	(100.00%)
門診次數				
平均值	4.879	16.726	37.894	10.252
標準差	(2.753)	(4.746)	(9.145)	(9.293)
Q1 <sup>1</sup>	2	13	31	4
Q2 <sup>1</sup>	5	16	35	7
Q3 <sup>1</sup>	7	20	42	14
全距	1-10	11-28	29-79	1-79
女性	45.27%	57.64%	54.36%	49.59%
年齡				
平均值	28.676	24.248	19.277	26.817
標準差	(15.840)	(17.686)	(20.663)	(16.921)
Q1 <sup>1</sup>	16	8	4	13
Q2 <sup>1</sup>	26	22	7	25
Q3 <sup>1</sup>	39	37	34	38
全距	0-100	0-99	1-97	0-100

<sup>1</sup> Q1是指第25百分位數者的值，Q2是第50百分位數者的值，Q3是第75百分位數者的值。

明顯的差距：使用門診29次以上的這百分之五的人，平均年齡最輕，且其第50%的人年齡只有7歲；使用11-28次的人之中，第25%的人年齡只有8歲。也就是說，有相當比率的幼童，平均隔週就看1次門診，但這些是沒有慢性病或七大疾病的兒童或幼童！他們為什麼會大量使用門診？這是值得進一步探索的問題，也值得衛生主管機關注意。

### （五）行為面的健康分組與主觀的自評健康

把以上的四種健康程度跟主觀自評健康相比對，有助於理解民衆就醫行為的主觀意涵，或者反過來說，理解民衆主觀健康評量的行為反應。為說明本文所建構的健康測量的可能應用，我們使用2001年國健局之「健康調查」及其所連結的就醫資料，觀察兩者的關係，作為應用這項結果的例子。「健康調查」的自評健康題項，詢問受訪者在受訪時間點之前四週的健康狀況。根據調查時間的紀錄，訪問的時間是從2001年7月至2002年2月，但約92%的樣本集中在2001年8至10月間完訪（附錄）。

為比對該月的就醫資料所指涉的健康狀況與自評健康之間的異同，我們的分析步驟如下：首先，以「承保歸人檔」中，12歲及以上的樣本在8到12月間的就醫資料，逐月進行四組的潛在側圖分析，這些按月分析所得到的各月分組方式完全一致<sup>3</sup>。接著我們根據所得四組的就醫行為的特性，分別命名為（1）沒有使用任何醫療資源的「零就醫組」；（2）僅看過1次門診的「使用一次門診組」；（3）門診量介於2至13次

---

<sup>3</sup> 唯一有差異的是「使用多次門診組」，這一組的門診範圍有兩個月的最高門診次數分別為12及14，其他月份都是13次。由於差異甚小，我們直接將這一組的門診次數限制為2至13次。

的「使用多次門診組」；及（4）曾經住院，或未曾住院但門診量至少14次的「重度使用組」，這四組依次代表健康等級排序。第三，將以上四組的分類標準，套用於「健康調查」的受訪者在受訪當月的就醫資料，分類這些受訪者，形成他們在當月的健康狀況等級。我們不直接以「健康調查」受訪者的健保就醫資料進行潛在側圖分析，是考慮到調查資料所包含的健康分布，與母群可能有差異之故。

根據這樣的分析策略所得到的結果顯示，「零就醫組」占全部樣本的57.96%，「使用一次門診組」占17.6%，23.7%的人屬於「使用多次門診組」，「重度使用組」僅0.8%（表7）。這些分布與「承保歸人檔」全年的資料非常不同，但近似「承保歸人檔」逐月的分析結果<sup>4</sup>。

「使用多次門診組」，所占比率頗高，也有門診次數範圍過廣（2-13次）的問題。但由於這不是此處分析的重點，我們就不再分析其異質性，而直接比較這樣的客觀健康分組與自評健康的一致性。

國健局2001年「健康調查」使用的自評健康測量，屬於不平衡設計，答項依序是「極好」、「很好」、「好」、「普通」和「不好」。從「極好」到不好所占比率分別是5.21%，30.5%，25.8%，32.37%，和6.12%。為與我們所建構的客觀健康等級相對應，我們將「極好」和「很好」合併成「很好」，產生四等級的主觀自評健康。

客觀的就醫行為產生的健康分組與主觀自評健康之間的一致性如何？由於每個人的自評健康標準不一，因此在分辨是否與實際狀況一致時，我們採取較為寬鬆的標準。據此，「二者一致」的條件定義如下：

- （1）「零就醫組」若自評健康為「很好」或「好」。
- （2）「使用一次

---

<sup>4</sup> 以「承保抽樣歸人檔」8到12月的資料，逐月進行分析時，「零就醫組」約占總體樣本的50%-55%之間；使用一次門診組約占19%-20%，使用「多次門診組」約占24%-29.5%之間，而重度使用者則在1%上下。

表7 2001年「國民健康訪問調查」當月就醫資料分類結果：依「承保抽樣歸人檔」2001年就醫紀錄檔同月份資料之分析結果

	零就醫組	使用一次門診組	使用多次門診組	重度使用組	總和
人數	11,354	3,443	4,641	150	19,588
百分比	(57.96%)	(17.58%)	(23.69%)	(0.77%)	(100.00%)
門診次數					
平均值	0.000	1.000	3.383	4.673	1.013
標準差	(0.000)	(0.000)	(1.786)	(4.572)	(1.715)
Q1 <sup>1</sup>	0	1	2	2	0
Q3 <sup>1</sup>	0	1	4	6	1
全距	0-0	1-1	2-13	0-32	0-32
有急診住院比率	0.00%	0.00%	0.00%	14.67%	0.11%
有動住院手術比率	0.00%	0.00%	0.00%	56.67%	0.43%
住院天數					
平均值	0.000	0.000	0.000	6.133	0.047
標準差	(0.000)	(0.000)	(0.000)	(6.779)	(0.797)
Q1 <sup>1</sup>	0	0	0	2	0
Q3 <sup>1</sup>	0	0	0	7	0
全距	0-0	0-0	0-0	0-31	0-31
女性 (%)	46.74%	53.33%	57.34%	40.67%	50.36%
年齡					
平均值	36.908	40.242	45.508	49.767	39.630
標準差	(16.561)	(18.029)	(18.637)	(19.225)	(17.735)
範圍	12-92	12-96	12-98	12-84	12-98

<sup>1</sup> Q1是指第25百分位數者的值，Q3是第75百分位數者的值。

門診組」若自評健康「普通」或「好」。(3)「使用多次門診組」，且自評健康「不好」或「普通」。和(4)「重度使用組」且自評健康「不好」(表8)。表8是自評健康的四組與就醫行為四組的交叉表。表8以灰色為底的七格即為主客觀健康程度一致者；無灰色底部分，右上方屬於主觀健康評估差，但從就醫資料而言，是健康狀況良好的一群；左下方則反之。

考量到一般人說話的模糊性及對於健康好壞的標準不一，這裡的就

表8 行為面的健康分組與主觀自評健康狀況：2001年「國民健康訪問調查」樣本

自評健康	行為面健康分組				樣本數
	重度使用組	使用多次門診組	使用一次門診組	零就醫組	
不好	4.09% <sup>1</sup>	44.49% <sup>1</sup>	16.19%	35.23%	100.00%
	32.67%	11.48%	5.63%	3.72%	6.12%
普通	0.90%	29.35% <sup>1</sup>	18.28% <sup>1</sup>	51.47%	100.00%
	38.00%	40.10%	33.66%	28.75%	32.37%
好	0.36%	21.49%	17.75% <sup>1</sup>	60.41% <sup>1</sup>	100.00%
	12.00%	23.40%	26.05%	26.89%	25.80%
很好	0.37%	16.60%	17.06%	65.98% <sup>1</sup>	100.00%
	17.33%	25.02%	34.65%	40.65%	35.71%
樣本數	150	4,641	3,443	11,354	19,588
	0.77%	23.69%	17.58%	57.96%	100.00%
	100.00%	100.00%	100.00%	100.00%	100.00%

$$\chi^2(9) = 902.79^{**}$$

<sup>1</sup> 有灰色底的組，是我們認為主觀自評與客觀分類一致之組別。

醫紀錄與自評健康兩者一致性之定義，根據的是有點常識性的推論，不可否認地，並不絕對嚴謹。舉例而言，一個月內完全沒有就醫紀錄的人，假定此人健康良好的可能性很高，因而也很有可能在自評健康時回答是「很好」或是「好」。相反的，一個月內重度使用醫療資源者，極有可能會認為自己健康狀況「不好」。但較有爭議的情況是，只使用1次門診，或使用多次（2次到13次）門診者，其自評健康需要是「很好」、「好」或「普通」，才算一致？這些並不容易作明確的判斷。為了簡化討論，後續的多項式邏輯迴歸時，就暫時把比較有爭議的不一致組排除於比較組之外，這些包括屬於「零就醫組」但自評「普通」、屬於「使用一次門診組」但自評「很好」、「使用多次門診組」但自評「好」，及「重度使用組」但自評「普通」的人。

我們挑出比較確定屬於「不一致」的四組，以其性別、年齡、

教育程度、受雇情形與收入作為自變項，進行多項式邏輯對數分析（multinomial logistic regression），用以觀察這些變項在一致與不一致組之間所扮演的角色。這四種不一致包括（1）零就醫且自評健康「不好」，（2）使用一次門診且自評健康「不好」，（3）使用多次門診且自評健康「很好」，及（4）重度使用且自評健康「好」或「很好」。其中（1）與（2）是少用醫療資源但自認健康不良者，而（3）與（4）是較常使用醫療資源但自認健康者。我們特意略過上一段所說的爭議較大的四組，是因為作多項式邏輯迴歸的目的，是要將一致與不一致者相對照：將這些屬於灰色地帶的四組排除，讓我們的問題得到較清楚的答案。

表9是以一個多項式邏輯模型，探討自變項與四種不一致者之間的關聯；每一種不一致組的參考組都是之前所定義的那四種一致組。

**表9 行為面的健康分組與主觀自評健康狀況之關係：多項式邏輯對數分析<sup>1</sup>**

	零就醫 且自評不好	使用一次門診 且自評不好	使用多次門診 且自評很好	重度使用 且自評好或很好
女性	0.105 (0.106)	0.044 (0.154)	0.153* (0.063)	-0.686* (0.333)
年齡	0.023*** (0.003)	0.029*** (0.005)	-0.008*** (0.002)	0.015 (0.010)
教育年數	-0.061*** (0.015)	-0.044* (0.021)	0.033** (0.010)	0.114* (0.046)
受雇	-0.239 (0.158)	0.168 (0.225)	-0.276** (0.103)	-0.086 (0.487)
個人月收入	-0.068* (0.034)	-0.147** (0.050)	0.060** (0.022)	-0.063 (0.100)
常數	-3.586*** (0.264)	-4.670*** (0.395)	-2.567*** (0.152)	-6.846*** (0.741)

N=19,588。\* p<.05, \*\* p<.01, \*\*\* p<.001。

<sup>1</sup> 參照組為自評與醫療使用情形一致者，亦即表8中灰色網底的組別。

因此，例如，年齡在「零就醫且自評不好」的次模型中是顯著為正（0.023\*\*\*），其詮釋是：相對於就醫行為與自評健康一致者，年齡愈高，愈容易落在「零就醫且自評健康不好」的一組，分析結果列於表9。觀察表9，有三個主要發現：（1）女性儘管使用多次門診，卻傾向於自評健康很好；相對而言，男性即使重度使用醫療資源，也傾向於自認為健康狀況良好。（2）年紀愈大，教育程度愈低，愈可能屬於「少用醫療資源但自認健康狀況不好」的一組；相對而言，使用多次醫療資源，但卻自認健康狀況良好的人，傾向於年輕或教育程度高的特性。（3）合併受雇與否及個人收入來看，個人收入對於主客觀健康評量的兩種不一致情形，有相反的解釋力量。收入愈低，愈可能屬於「少用醫療資源但自認健康不好」的一組；相對而言，收入愈高，或是未受雇者，愈可能屬於「多使用醫療資源但自認健康狀況佳」的一組。綜合而言，自認身體不好但無就醫者，似乎較多為年紀大、教育程度低、收入少的人，比較像是社會的弱勢群體，既貧且老。相對而言，多次使用醫療資源卻自認健康佳者，年齡較輕，教育程度也較高。當然，貧且老者可能居住在醫療資源較不足的鄉村，而年輕且教育程度較高的人，則很可能居住在醫療資源豐富的都市，因此這些發現也很可能跟醫療可近性的差異有關，而不僅只於個人本身的特徵，醫療可近性的影響，則有待進一步實證研究來評估。

由於此處尚屬初步分析的結果，且這些分析的目的，只是在舉例說明以就醫行為作健康分組後，可能作的進一步分析；在一致與否的標準認定上，仍有商榷之處，我們不能就此斷定，全民健康保險似乎沒有祛除社會弱勢者尋求醫療協助的障礙。但由於這樣的分析，引發了許多值得進一步探索的課題，例如自認健康佳者，為什麼常使用醫療資源？而那些自評健康不好的人，又為什麼沒有就醫？

## 六、結論

本研究依循健康指標量表的傳統，視健康為多重面向的構念，測量上使用多重指標。但不同於健康測量文獻之多由調查得到多面向的測量資訊，本研究採用由「承保歸人檔」產生，包括門診、住院、手術、疾病診斷等就醫資料，運用潛在側圖分析法，分析了2001年近93萬人全年的就醫紀錄檔，辨別出四個群組：（1）「零就醫組」：完全沒有使用任何醫療資源；（2）「門診使用組」：僅使用門診且沒有住院紀錄；（3）「短期住院組」：平均門診量比「門診使用組」高，皆曾短期住院，且有較高比率是經由急診住院；（4）「重度使用組」：長期住院、或高度使用門診的一群。我們觀察到潛在側圖分析所得出的四組，大致以門診的使用、及住院開刀資源的使用，作為兩個切分點，使組內的差異最小。且由慢性病及七大疾病在各組的發生率來看，各組間的健康程度的確有明顯差異。也就是說，依就醫資料，可以推測這四組的健康程度依序由「零就醫組」的最佳、「門診使用組」、「短期住院組」、到「重度使用組」的最差。本文的研究方法，提供學界一個運用現有健保資料，產生測量健康的另類方式。

運用潛在側圖分析法分析就醫紀錄，不僅可以將複雜的就醫紀錄，化繁為簡，將底層的健康程度，抽離顯明，更可透過分析各組內部的異質性，將衛生政策所關切的資訊剖明呈現，而產生政策參考價值。例如由「門診使用組」的異質性，我們發現，高門診次數且無慢性病及重大疾病的人之中，幼兒比率偏高；就醫行為的健康分組與自評健康的分析則顯示，年長、低教育、低收入者自評健康不佳，就醫次數卻相對較低；而年輕、高教育、高收入者，自評健康好，醫療資源使用次數卻相對多。這些發現都有值得追究分析的必要，具有重要政策意涵。不僅是

方法學上的課題，社會階層化理論上亦有值得開拓的可能課題。

本文分析了全年的就醫資料，也分析了8到12月逐月的資料，我們注意到兩者的結果差異頗大，例如在分析全年資料時，僅有將近9.5%的人完全沒有使用醫療資源，而分析逐月資料時，則有超過50%沒有使用任何醫療資源。很明顯的，因為資料時間點的界定，所得到的潛在健康分類的標準和所得之比率亦有差別。畢竟，一年的時間比一個月長了許多，在一個月內完全沒有就醫的人，未必在其他時段不會就醫。這說明了：自評健康的時間參考範圍，究竟應該以年計、月計、或是綜合多少個月來計算，需要研究者視研究主題來判斷。本文在比較行為面的健康排序和自評的健康程度時，也意識到另一個問題：一個人對健康的自覺，大約維持多長的時間？一般自評健康題項多以過去一個月的時間為參考範圍，詢問受訪者自覺的健康程度；這個時間點的決定，是根據國際慣例；但是，多長時間之內的就醫行為，會跟一個人在某一個時間點的自評健康最有關係？這個問題的釐清，可望有助於研究者在設計問卷時，對於自評健康時間參考範圍的設定。

當然，潛在側圖分析之模型分組，不必然要全盤接受，以「門診使用組」為例，門診使用次數的全距頗大，不盡然合理。研究者可以利用其他資訊，將變異量過大且比率過高的組進行細分，或也可能將疾病診斷的資訊，與就醫行為合併一起，進行潛在側圖分析；或者利用常態分布的概念，將門診次數超過平均值兩個標準差的人獨立成一組。無論如何，研究者的判斷，在分組模型的選擇上，是不可少的一環。

健保資料與調查資料的串連，使臺灣的學者有機會深究自評健康資料與就醫行為資料間的異同，並檢視那些以自評健康為依變項的研究結果及意義。藉著如此綜合所產生的健康指標，可用於理解各人因教育、社會經濟地位、性別、族群等變項而產生的健康差距，與自評健康和死

亡為依變項的研究發現對話：如此對於理論的精進與突破，將有相當潛力。具體而言，根據本文建議，由就醫類型和頻率組合而產生客觀的健康指標，有機會減少如 Cutler及Lleras-Muney（2006）一文中，健康測量方式太多的困擾；可以超越Johnston等人（2009）在探討主觀的自評健康和客觀的健康狀況間之異同時，只能使用高血壓作例子的推論限制；更有機會與Baker等人（2004）比較病歷資料與自評健康的研究，進行對話。而之所以有機會進行這類研究，不僅因為調查資料與健保資料的串連，更因為有方法可以綜合呈現複雜的就醫資訊。

### 作者簡介

張苙雲，中央研究院社會學研究所研究員。目前正進行的研究包括教育與健康不平等、生命週期與健康效應、及公平與幸福感等。

楊孟麗，中央研究院調查研究專題中心助研究員。目前正進行的研究包括訪員的人格與訪問特質對於資料品質的影響、及學生的學業成就與心理健康。

趙景雲，東吳大學社會學系碩士。現為中央研究院社會學研究所研究助理。

**附錄 2001年「國民健康訪問調查」完訪年月次數分布**

完訪年月	完訪樣本數	百分比
2001年7月	1	0.00
2001年8月	2,897	13.93
2001年9月	9,130	43.89
2001年10月	7,212	34.67
2001年11月	933	4.49
2001年12月	269	1.29
2002年1月	329	1.58
2002年2月	31	0.15
總樣本數	20,802	100.00

## 參考書目

- NHIS抽樣工作小組，2001，《2001年「國民健康訪問調查」抽樣報告》。NHIS抽樣工作小組。
- 吳就君，1981，〈台灣地區居民社會醫療行為研究〉。《公共衛生》8: 25-49。
- 林芸芸，1992，〈大學生的中、西醫醫療求醫行為〉。《中華民國家庭醫學雜誌》2(4): 176-186。
- 財團法人國家衛生研究院，2010，2000年承保抽樣歸人，LHID2000。  
[http://w3.nhri.org.tw/nhird//date\\_cohort.htm](http://w3.nhri.org.tw/nhird//date_cohort.htm)，取用日期：2010年6月24日。
- 張苙雲，1998，〈「逛醫師」的邏輯：求醫歷程的分析〉。《臺灣社會學刊》21: 59-87。
- 楊文山，1992，〈台灣地區民衆求醫行為之分析〉。《榮總護理》9(2): 121-125。
- 劉仲冬，1997，〈求醫行為研究〉。《醫望雜誌》20: 95-97。
- 盧瑞芬、曾旭民、蔡益堅，2003，〈國人生活品質評量(I)：SF-36 台灣版的發展及心理計量特質分析〉。《台灣衛誌》22(6): 501-511。
- Ahmad, Al-Windi, Elmfeldt Dag and Svardsudd Kurt, 2004, "The Influence of Sociodemographic Characteristics on Health Care Utilization in a Swedish Municipality." *Uppsala Journal of Medical Sciences* 109: 33-42.
- Baker, Michael, Mark Stabile and Catherine Deri, 2004, "What Do Self-Reported, Objective, Measures of Health Measure?" *The Journal of Human Resources* 39(4): 1067-1093.
- Bergner, Marilyn et al., 1981, "The Sickness Impact Profile: Development

- and Final Revision of a Health Status Measure.” *Medical Care* 19(8): 787-805.
- Bergner, Marilyn, 1985, “Measurement of Health Status.” *Medical Care* 23(5): 696-704.
- Bergner, Marilyn, and Margaret L. Rothman, 1987, “Health Status Measures: An Overview and Guide for Selection.” *Annual Review of Public Health* 8: 191-210.
- Bound, John, 1991, “Self-Reported versus Objective Measures of Health in Retirement Models.” *The Journal of Human Resources* 26(1): 106-138.
- Bozdogan, Hamparsum, 1987, “Model Selection and Akaike’s Information Criteria (AIC): The General Theory and Its Analytical Extensions.” *Psychometrika* 52: 345-370.
- Case, Anne C. and Christina Paxson, 2005, “Sex Differences in Morbidity and Mortality.” *Demography* 42(2): 189-214.
- Crossley, Thomas F. and Steven Kennedy, 2002, “The Reliability of Self-Assessed Health Status.” *Journal of Health Economics* 21(4): 643-658.
- Curtis, Lori J. and William J. MacMinn, 2008, “Health Care Utilization in Canada: Twenty-Five Years of Evidence.” *Canadian Public Policy* 34(1): 65-87.
- Cutler, David M. and Adriana Lleras-Muney, 2006, “Education and Health: Evaluating Theories and Evidence.” *NBER Working Paper*, No. 12352. Cambridge, MA: National Bureau of Economic Research.
- Deri, Catherine, 2005, “Social Networks and Health Service Utilization.” *Journal of Health Economics* 24(6): 1076-1107.
- DeSalvo, Karen B. et al., 2006, “Mortality Prediction with a Single General

Self-Rated Health Question: A Meta-Analysis.” *Journal of General Internal Medicine* 21(3): 267-275.

DeSarbo, Wayne S. et al., 2006, “Identifying Sources of Heterogeneity for Empirically Deriving Strategic Types: A Constrained Finite-Mixture Structural-Equation Methodology.” *Management Science* 52: 909-924.

DeSarbo, Wayne S., Kamel Jedidi and Indrajit Sinha, 2001, “Customer Value Analysis in a Heterogeneous Market.” *Strategic Management Journal* 22: 845-857.

Dracup, Kathleen et al., 1995, “Causes of Delay in Seeking Treatment for Heart Attack Symptoms.” *Social Science & Medicine* 40(3): 379-392.

Dunlop, Sheryl, Peter C. Coyte and Warren McIsaac, 2000, “Socio-Economic Status and the Utilisation of Physicians’ Services: Results from the Canadian National Population Health Survey.” *Social Science & Medicine* 51(1): 123-133.

Fishbein, Martin and Icek Ajzen, 1975, *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley Pub. Co.

Fraley, C. and A. E. Raftery, 1998, “How Many Clusters? Which Clustering Method? Answers via Model-Based Cluster Analysis.” *Technical Report*, No. 329. Seattle: Department of Statistics, University of Washington.

Furlong, William et al., 1998, “Multiplicative Multi-Attribute Utility Function for the Health Utilities Index Mark 3 (HUI3) System: A Technical Report.” *CHEPA Working Paper Series*, No. 1998-11. Hamilton, Canada: Centre for Health Economics and Policy Analysis, McMaster University.

- Garratt, Andrew et al., 2002, "Quality of Life Measurement: Bibliographic Study of Patient Assessed Health Outcome Measures." *British Medical Journal* 324: 1417-1419.
- Guralnik, Jack M. et al., 1989, "Physical Performance Measures in Aging Research." *The Journal of Gerontology* 44(5): M141-M146.
- Huisman, Martijn, Frank van Lenthe and Johan Mackenbach, 2007, "The Predictive Ability of Self-Assessed Health for Mortality in Different Educational Groups." *International Journal of Epidemiology* 36(6): 1207-1213.
- Hunt, S. M., J. McEwen and S. P. McKenna, 1985, "Measuring Health Status: A New Tool for Clinicians and Epidemiologists." *The Journal of the Royal College of General Practitioners* 35: 185-188.
- Idler, Ellen L. and Yael Benyamini, 1997, "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38: 21-37.
- Jedidi, Kamel, Harsharanjeet S. Jagpal and Wayne S. DeSarbo, 1997, "STEMM: A General Finite Mixture Structural Equation Model." *Journal of Classification* 13: 23-50.
- Johnston, David W., Carol Propper and Michael A. Shields, 2009, "Comparing Subjective and Objective Measures of Health: Evidence from Hypertension for the Income/Health Gradient." *Journal of Health Economics* 28: 540-552.
- Katz, Sidney et al., 1963, "Studies of Illness in the Aged. The Index of ADL: A Standardized Measure of Biological and Psychosocial Function." *The Journal of the American Medical Association* 185(12): 914-919.

- Kleinman, Arthur et al., eds., 1974, *Medicine in Chinese Cultures: Comparative Studies of Health Care in Chinese and Other Societies*. Washington, DC: U.S. Department of Health, Education, and Welfare, Public Health Service, National Institutes of Health.
- Kleinman, Arthur, 1980, *Patients and Healers in the Context of Culture: An Exploration of the Borderland between Anthropology, Medicine, and Psychiatry*. Berkeley: University of California Press.
- Krause, Neal M. and Gina M. Jay, 1994, "What Do Global Self-Rated Health Items Measure?" *Medical Care* 32(9): 930-939.
- Leinonen, Raija, Eino Heikkinen and Marja Jylhä, 1998, "Self-Rated Health and Self-Assessed Change in Health in Elderly Men and Women — A Five-Year Longitudinal Study." *Social Science & Medicine* 46(4-5): 591-597.
- Lindeboom, Maarten and Eddy van Doorslaer, 2004, "Cut-Point Shift and Index Shift in Self-Reported Health." *Journal of Health Economics* 23(6): 1083-1099.
- McHorney, Colleen A., 1997, "Generic Health Measurement: Past Accomplishments and a Measurement Paradigm for the 21st Century." *Annals of Internal Medicine* 127(8): 743-750.
- McLachlan, Geoffrey and David Peel, 2000, *Finite Mixture Models*. NY: John Wiley & Sons Inc.
- Pastor, Dena A. et al., 2007, "A Latent Profile Analysis of College Students' Achievement Goal Orientation." *Contemporary Educational Psychology* 32: 8-47.
- Ruto, Eric, Guy Garrod, and Riccardo Scarpa, 2008, "Valuing Animal

- Genetic Resources: A Choice Modeling Application to Indigenous Cattle in Kenya.” *Agricultural Economics* 38: 89-98.
- Scarpa, Riccardo and Mara Thiene, 2005, “Destination Choice Models for Rock Climbing in the Northeastern Alps: A Latent-Class Approach Based on Intensity of Preferences.” *Land Economics* 81: 426-444.
- Singh-Manoux, Archana et al., 2006, “What Does Self Rated Health Measure? Results from the British Whitehall II and French Gazel Cohort Studies.” *Journal of Epidemiology & Community Health* 60(4): 364-372.
- Subramanian, S. V. and Karen Ertel, 2008, “Is the Use of Self-Rated Health Measures to Assess Health Inequalities Misleading?” *International Journal of Epidemiology* 37(6): 1436-1437.
- Suls, Jerry, Christine A. Marco and Sheldon Tobin, 1991, “The Role of Temporal Comparison, Social Comparison, and Direct Appraisal in the Elderly’s Self-Evaluations of Health.” *Journal of Applied Social Psychology* 21(14): 1125-1144.
- Terwee, Caroline B. et al., 2007, “Quality Criteria were Proposed for Measurement Properties of Health Status Questionnaires.” *Journal of Clinical Epidemiology* 60: 34-42.
- Van der Heyden, J. H. A. et al., 2003, “Socio-Economic Differences in the Utilisation of Health Services in Belgium.” *Health Policy* 65: 153-165.
- Van der Meer, Joost B. W., Johannes van den Bos and Johan P. Mackenbach, 1996, “Socioeconomic Differences in the Utilization of Health Services in a Dutch Population: The Contribution of Health Status.” *Health Policy* 37: 1-18.
- Vermunt, Jeroen K. and Jay Magidson, 2002, “Latent Class Cluster Analysis.”

Pp. 89-106 in *Applied Latent Class Analysis*, edited by Jacques A. Hagenaars and Allan L. McCutcheon. Cambridge: Cambridge University Press.

Vermunt, Jeroen K. and Jay Magidson, 2005, *Technical Guide for Latent Gold 4.0: Basic and Advanced*. MA: Statistical Innovations Inc.

Wade, Tracey D., Ross D. Crosby and Nicholas G. Martin, 2006, "Use of Latent Profile Analysis to Identify Eating Disorder Phenotypes in an Adult Australian Twin Cohort." *Archives of General Psychiatry* 63: 1377-1384.

Ware, John E., 1988, *How to Score the Revised MOS Short-Form Health Scale (SF-36)*. Boston, MA: The Health Institute, New England Medical Center Hospitals.

