

# 資料倉儲建置之研究-以全民健保資料庫為例

劉育津、陳雋仁

世新大學資管系 教授、研究生

ycliu@cc.shu.edu.tw

ji394su33000@hotmail.com

## 摘要

「全民健康保險」自 1995 年三月開辦，其「健康保險資料庫」紀錄納涵了所有保險人每次就醫的紀錄，其資料量之巨大以及成長速度非一般中小型資料庫可比擬，也正因如此，中央健康保險局也應學界之專家學者請託，以全民健保資料庫為基礎，在保障健康保險民眾的隱私之下，建立一個以學術研究為目的的資訊資料庫，提供給非營利機構之學者專家進行相關研究。

本研究的範圍在於自 1999 年以來，健保局進行數位化後的研究為主體，利用「台灣碩博士論文網」的檢索功能，搜集到國內研究者們針對健保議題的研究，我們在確立範圍之後，得到國內研究者對於健保議題的相關文獻，透過「變項分析」的方式，得到數個主要的研究目標主題，我們發現過去探討健保倉儲架構，鮮少對於健保資料主檔或制度面的資料改動進行記錄，因此喪失許多可以比較的數據。

故我們擬進一步對健保倉儲或超市模型提出進一步的建議，以期後續學者或實務專家能夠以更短的時間得出所需要的結果，也提供更精準的決策之用途。

**關鍵字：**健保資料庫、部分負擔、資料倉儲、變項分析

## 壹. 研究計畫之背景與目的

### 1.1 研究背景

由於健保資料庫建置初期，資料倉儲的進階技術尚未引進，因此採用分散式作業架構，其資料分散，健保總局無法保有一套完整的整合資料，並且因系統龐大，各分局執行大量統計耗時過長，影響正常線上作業，甚至導致某些作業無法進行，分局資料異動也非常頻繁，於不同時間點擷取資料，會導致不同結果，為應付統計及分析需求，各分局會保留各項明細，除影響回應速度，尚有檔案備份的負擔。[3]為了解決上述問題，健保局於民國 86 年引進資料倉儲系統，用以支援規畫及分析需求[1]：

1. 將分局線上作業主機與總局之批次作業主機功能分割，使現有設備作最佳利用，增進整體作業效率。
2. 加速資料統計及分析處理，建立分析模式與決策輔助系統。
3. 透過資料清理及整合，提供決策者、分析者，利用工具快速且便捷的取得一致性的資訊。
4. 建立完整的全國醫療相關資訊系統，提供衛生醫療或學術研究單位的所需資訊。

資料倉儲系統是一套改良過後的決策支援系統，可以產生整合資料並產生系統化、結構化、的高階資料，為了達到組織內部的資訊流通、應付管理問題的目的，藉由建置一個集中的資訊倉庫，並且配合資料分析工具與快速的決策支援軟體，使得資料可以幫助企業的決策者，其決策制定並輔助商案的訂定。

健保資料庫所提供的資料對於公共衛生界、保險產業、資訊研究等都有相當大的貢獻；各界研究者可獲取涵蓋全體國人使用醫療產業的情形及概況，進而預測來年的趨勢，訂定適當的對策、重點預防宣導等決策，並且可以監督醫事機構的資源使用情形，進而配置國家整體醫療資源的分配等對策。[2]

## 1.2 研究目的

本研究藉由國內外相關文獻進行探討，分析國內外研究者於研究健康保險系統的概況，經過我們歸納整理過後，期望能夠獲得國內外研究者關心的主題，進行歸納整理，找出國內研究者之研究目標、其分析的變項，並且建置其主題分析模型，為達到健保資料倉儲系統所應有且完善的架構，經由分析當前健保資料倉儲系統之分析模型，提出優缺點，並提出相關建議。

## 貳.文獻探討

本章節將探討國內外資料倉儲的相關研究、使用的概況等議題，並且蒐集國內針對健保所提出的研究，以建立研究架構，可供相關研究者使用，2.1 節探討資料倉儲的相關理論和介紹，2.2 節則回顧國內以全民健保資料庫進行分析的主題。

### 2.1 資料倉儲概況

企業在資訊化後，就會使用到各式資料庫以及各式系統處理日漸增加的資料量；交易處理系統簡稱 TPS (Transaction Processing System) 是企業中最基礎的資訊系統，主要使用於輔助基層人員操作交易紀錄，此日常累積的交易紀錄往往都是提供給其他高階輔助系統的根據，例如銷售／行銷方面如訂單輸入系統。[6]

線上交易處理系統 (On-Line Transaction Processing)，是傳統關聯式資料庫 (Relational Database) 之主要應用，利用即時運算中心的處理能力，將基礎性的日常事務處理、或是用戶的原始資料查詢等資訊，及時回傳至使用者；線上交易處理的基礎建立在快速且可應付大量的請求端，因此也將會具備一個 OLTP 資料庫來應付這類需求，上述這些交易系統所產生的資料不但日漸增加，並且使用者往往會要求能從不同角度去分析資料，以隨時掌握最新資訊。[5]

線上分析處理，旨在促使分析人員、管理人員或執行人員，都能藉由多角度針對資訊進行快速、一致、交互性之存取，在滿足決策支援或在多維度環境下，進行特定查詢與報表需求，進而深入瞭解資料的一項軟體技術；資料採礦是為要發現出有意義的樣型或規則，而必須從大量資料之中以自動或是半自動方式，且不斷循環的決策支援分析過程，它能夠從組合在一起的資料中，發現出隱藏價值的知識，以提供給企業專業人員參考、探索和分析資料，下圖 2-1 為基於資料倉儲決策支援系統示意圖。[35]

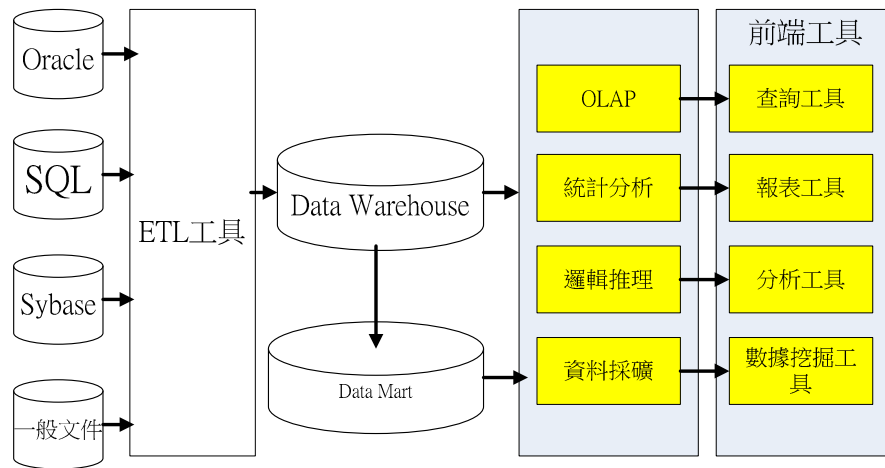


圖 2-1 資料倉儲與決策支援系統結構圖

資料倉儲根據 W. H. Inmon(1994)[37]的定義：「資料倉儲具備 1. 主題導向(Subject Oriented)、2. 不變動性(Nonvolatile)、3. 整合性(Integrated)、4. 時間差異性(Time-Variant)，等特質，主要用來協助企業決策過程。」

1. 主題導向(Subject Oriented)：以協助建置之資料目標為主體，建置出以該資料體為主題的資料倉儲系統，經人員設計後排除不必要的雜訊，建置出具有特定功能的資料倉儲系統。
2. 不變動性(Nonvolatile)：與一般資料庫不同，資料倉儲系統儲存的資料內容屬於長時間儲存之下的歷史性資料，不因目標屬性變動進而牽動整體資料，也不因時間失效更不會變更，使用者除了修正不作資料行更改。
3. 整合性(Integrated)：資料倉儲可自傳統式資料庫整合，傳統式資料庫可能具有分散各地的狀況，使用者所操作之資料格式也可能因人因時而異，建立資料倉儲系統之前必須經過資料前處理、資料萃取(Data Extract)，資料淨化(Data Clean)，資料整合(Data Integrate)，如此才不會有產生同筆資料出現不同描述，或者資訊粒度(Granularity)不一致的問題。
4. 時間差異性(Time-Variant)：傳統式資料庫即使具有紀錄時間的欄位，也不能用以參照時間變化之下，資料內容產生的差異性，而資料倉儲具有累計快照事實表格等時間差異下的變化追蹤技術。

根據資料倉儲的組成元素定義：「資料倉儲的組成元素由萃取、轉換、淨化、載入、彙總、抽象資料、資料倉儲資料庫、查詢工具。」[36]

藉由儲存大量的目標資料，透過展示介面或線上分析處理的工具，提供給管理者作決策參考時使用[35]。而這些資料往往以多維度的(Multi-Dimensional)資料方格(Data Cubes)模式來組織和呈現。

## 2.2 國內全民健保資料研究概況

我們將研究範圍鎖定在國內的健保資料庫，所以我們從國家圖書館的全國碩博士論文網中，以"健保"為其關鍵字進行搜尋，並以有附加電子全文檔的論文為主體，搜尋結果合計 75 篇，依關鍵字分類由主題篇次由多至少排序後得出下表 2-1。[4]

表 2-1 健保關鍵字查詢結果

編號	分類名稱	相關研究篇數	編號	分類名稱	相關研究篇數
1.	全民健保	75	8.	部分負擔	7
2.	全民健康保險	20	9.	資料探勘	6
3.	醫療利用	14	10.	總額預算	6
4.	總額支付制度	13	11.	社會保險	6
5.	醫療費用	11	12.	醫療糾紛	5
6.	健保資料庫	10	13.	總額預算支付制度	5
7.	健保 IC 卡	10	14.	疾病管理	5
…其他篇數過少					

資料來源:本研究整理

然而上述依關鍵字分類時主題會有重疊，或在某些關鍵字主題下的研究過少，故排除重複和過少的主題後，我們利用該系統針對全民健保所提供的主題知識地圖功能後 [4]，得到如下圖 2-2 所示之「知識地圖」，依其「次關鍵字」的「相關度」由近至遠分別為「醫療利用」、「醫療費用」、「總額支付制度」、「部分負擔」和「門診」，故透過該知識地圖我們進一步將健保整體搜尋結果，縮小範圍至上述五種主要類別的研究。

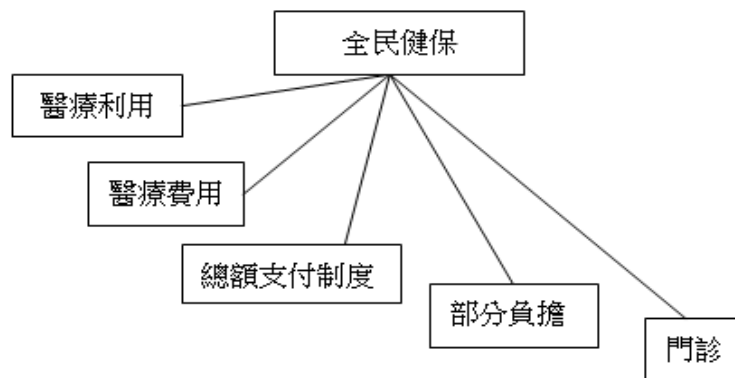


圖 2-2 知識地圖

依上圖 2-2 的知識地圖，我們針對「全民健保」的研究分類之下的五個主要相關研究，進行第二階搜尋，得到:1. 醫療利用 14 篇(表 2-2)、2 醫療費用 11 篇(表 2-3)、3. 總額支付制度 13 篇(表 2-4)4. 部分負擔 7 篇(表 2-5)5. 門診 4 篇(表 2-6)，共計 49 篇次。

我們藉由探討這些論文的目標，希望能夠得到研究者關心的主題，為了了解健保議題的國內研究者的研究內容，如何利用健保資料倉儲的功能進行資料挖掘，我們在第三章將進行變項分析等作業，藉由了解各項研究的變項，建置出各類研究所需之功能導向的分析模型。

### 參、健保研究之主題變項分析

在此章節我們將第二章提出的相關研究分類主題列表，在本章 3.1 節中進行主題變項分析的歸整，變項分析可歸整成「自變項分析」和「應變項分析」；「應變項」係由

該篇研究探求的目標組成，「自變項」係由研究中所提出會影響應變項的變因組成。由於本研究將重點設立在健保資料倉儲系統上，所以我們在整理時，將這些自變項限定在「健保資料倉儲系統」可取得的資料欄位。而在 3.2 節中，則依前一節的結果提出健保局資料倉儲可再進一步精煉和建置新的架構模型。

### 3.1 變項分析

在此章節首先是依照第二章所提到的系統知識地圖(詳見圖 2-2)，經過整理後依照：分類 1—醫療利用類別(計 13 篇，包含三篇跨主題重覆)、分類 2—醫療費用類別(計 4 篇，包含一篇跨主題重覆)、分類 3—總額支付制度類別(計 6 篇)、分類 4—部分負擔類別(計 2 篇，包含 2 篇跨主題重覆)，以及分類 5—門診類別(計 2 篇)，包含 1 篇跨主題重覆)，總計 27 篇，如表 3-1 所整理。

繼第二章全民健保研究分類，如某研究屬於跨類別，則在其類別分類欄位標上記號；例如「初次診斷憂鬱症患者追蹤三年之醫療利用與費用分析探討」，跨越醫療利用、醫療費用，則在其類別分類第一格與第二格標上記號。

表 3-1 變項分析目標研究清單

類別分類					篇次	研究名稱
1	2	3	4	5		
0					1	全民健保體系下醫療資源分佈對個人醫療使用的影響[8]
0					2	健保部分負擔調整對高血壓病人門診醫療利用之影響[9]
0					3	無職榮民與一般民眾醫療利用之比較分析[10]
0					4	臺灣地區精神疾病醫療利用分析[11]
0	0				5	初次診斷憂鬱症患者追蹤三年之醫療利用與費用分析探討[12]
0					6	台灣全民健保精神疾病盛行率、醫療利用及費用分析研究[13]
0					7	以台灣全民健保承保抽樣歸人檔案分析憂鬱症病患的醫療利用[14]
0					8	婦女醫療利用分析—以剖腹產、子宮切除、子宮頸抹片、乳癌、子宮頸癌為例[15]
0					9	以全民健保 1996-2001 年承保抽樣歸人檔分析下背痛病人之醫療利用[16]
0			0	0	10	全民健保部分負擔調整對民眾門診醫療利用的影響：以門診高利用者為對象[17]
0					11	老年人在全民健保下中西醫就醫行為研究[18]
0			0		12	藥品部分負擔政策對癲癇病人之影響[19]
0					13	全民健保中醫門診利用之研究[20]

類別分類					篇次	研究名稱
1	2	3	4	5		
	0				14	中央健保局年度醫療費用支出之預測模式研究—灰預測理論之應用[21]
	0				15	全民健保醫療費用成長模型建構之研究[22]
	0				16	醫院在參與卓越計畫前後醫療費用變化之探討—以健保局中區分局為例[23]
	0				17	門診藥品部分負擔對醫療院所藥行為及費用影響之研究[24]
		0			18	總額支付制度實施後醫院因應之道與員工認知程度之相關性研究—以某區域教學醫院為例[25]
		0			19	全民健保總額支付制度與醫病關係之影響——某醫學中心之個案研究[26]
		0			20	全民健保之總額支付制度下醫院經營效率之研究—DEA 之應用[27]
		0			21	牙醫門診總額支付制度實施後牙醫健保醫療資源使用狀況分析—以高屏區為例[28]
		0			22	全民健保牙科總額支付制度在新竹市縣城鄉醫療服務差異性的探討[29]
		0			23	牙醫師因應總額支付制度之調適策略探討——以南區健保局之特約診所牙醫師為例[30]
			0		24	全民健保部分負擔制度對幼童、年長者及慢性病患健康之影響[31]
			0		25	全民健保新制部分負擔對不同社會群體影響之探討[32]
				0	26	門診老年失眠症患者之安眠藥處方型態及處方問題[33]
				0	27	全民健保對各型醫院生產力之影響[34]

資料來源:本研究整理

承表 3-1 我們接著進行研究變項分析的歸整，一般來說，變項分析應用的研究是以數據的消長或交叉分析來比照資料前後時程之差異，從而得到該研究所需的佐證資訊。但某些研究，其使用的健保類別概況，並非是以純數據作為主要的描述方法，如此一來，或就不存在明確的變項分析之條件；然而，我們依然可以將其研究目標、使用到的健保資料庫欄位，作為替代應變項、自變項的分析模式，達到為分析模型取得各研究的類別的目的。(如第 1 篇和第 8 篇)

接下來 3.1.1~3.1.5 我們擬以系統知識地圖，分別以不同類別進行相關研究之變項分類。

### 3.1.1 醫療利用類別變項分析

在「醫療利用類別」的部分，我們以第 2 篇「健保部分負擔調整對高血壓病人門診醫療利用之影響」[9]作為範例來說明，該研究的變項分析條件，其應變項為「醫療利用」，包含：「就醫人次」、「醫院層級別就醫次數」、「平均每人每次給藥天數」、「平均每人門診總醫療費用」所組成。

該研究透過上述應變項的數據消長情形，達成「醫療利用」這項研究目標的分析。而在自變項的部分，這些數據需要透過「部分負擔」、「醫院特性」、「病患特質」、「案件類別」等數項健保資料欄位進行差異之比較，上述說明如表 3-2 所示：

表 3-2 醫療利用類別數據化分析範例

研究名稱：健保部分負擔調整對高血壓病人門診醫療利用之影響[9]	
應變項	
就醫人次、層級別就醫次數、平均每人每次給藥天數、平均每人門診總醫療費用	
自變項	
部分負擔	應部分負擔組、低收入戶組、榮民組、原住民組、殘障組
醫院特性	醫院層級別、醫院權屬別、縣市別、區域別
病患特質	性別、年齡、診斷別、投保類別、投保金額、殘障情形
案件類別	01 案件、04 案件、09 專案、D4 案件、E1 案件

資料來源：本研究整理

我們可以透過此類數據化的研究之應變項與自變項分析，可搭配健保資料庫的資料欄位，建構出類別化的資料倉儲分析模型。

為了進一步說明本節開頭所提到的，某些研究以描述事實的說明方式來達到研究目的，我們以第 1 篇「全民健保體系下醫療資源分佈對個人醫療使用的影響」[8]來進行說明。該研究之目標是探討「地區性醫療資源利用情形」、「年齡、性別醫療資源利用情形」，然而該研究並未從健保資料庫提取所需數據，而是以文獻研究後結果、舉實例佐證等方法，來說明該研究的目標。針對此類型之研究，我們以其研究目標與該研究所關心的健保資料欄位變項等，來做為資料倉儲之設計基礎，舉例說明如表 3-3 所示：

表 3-3 醫療利用類別非數據化分析範例

研究名稱：全民健保體系下醫療資源分佈對個人醫療使用的影響[8]	
研究目標	
地區性醫療資源利用情形、年齡、性別醫療資源利用情形	
健保資料欄位應用變項	
醫院特性	醫院層級別、醫院權屬別、縣市別、區域別、健保分局別
病患特質	性別、年齡

資料來源：本研究整理

由上述之應變項與研究目標進行分析，我們經變項分析之後，可得出以醫療利用為類別的研究，有兩大關係的主題：

- 一、依保險對象醫療利用分析：這些保險對象包含弱勢團體、特殊身分、年齡等差異

性，藉由分析特定族群的醫療利用行為，以此得知某族群是否在醫療利用上具有差異性。(如第 10 篇)

二、依疾病別醫療利用行為:將某病因列為其應變項，探討病患得某病因的病人，在醫療資源的使用上面可能會有何趨勢。(如第 2 篇)

三、依健保收費制度面相關:有關醫療利用行為類型的研究，多半都會針對變更收費制度之後對醫療利用行為造成的影響。(如第 10 篇)

### 3.1.2 醫療費用類別變項分析

在醫療費用類別的部分，我們以第 15 篇「全民健保醫療費用成長模型建構之研究」[22]作為變項分析範例，說明如下表 3-4:

表 3-4 醫療費用類別變項分析範例

---

研究名稱:全民健保醫療費用成長模型建構之研究[22]	
應變項	住院費用、急診費用、門診費用、總體醫療費用
自變項	每月醫療費用
	門診 西醫、中醫、牙醫
	急診
	住院
	醫院特性 醫院層級別、醫院權屬別、醫院型態別
	病患特質 性別、年齡

---

資料來源:本研究整理

由本研究類別我們可以得知，以醫療費用為類別的研究主要探求的三大目標:

一、依保險對象醫療費用分析:藉由分析特定受保險對象族群的醫療費用支出，以此得知某族群佔有全體國人的醫療費用支出的比例，藉此可以分析某族群對於其醫療費用的付出程度。(如第 5 篇)

二、依疾病別醫療費用分析:將某病因列為其應變項，探討患得某病因的病人，在醫療費用上面的付出程度。(如第 17 篇)

三、依醫事機構的醫療費用分析:大多探討某區域或某特定醫事機構的資源使用情形，這類研究藉由醫事機構的看診人次及使用的資源，分析出該機構的資源將該如何配置，以及預測該機構的某些部門未來的資源消耗會有何種趨勢。(如第 16 篇)

四、以健保收費制度面相關:有關醫療費用類型的研究，會針對變更收費制度之後對醫療院所費用收支問題(如第 16 篇)或是對於整體醫療支出造成的影響(如第 14 篇)。

### 3.1.3 總額支付制度類別變項分析

在總額支付制度類別的部分，我們舉第 18 篇「總額支付制度實施後醫院因應之道與員工認知程度之相關性研究-以某區域教學醫院為例」[25]說明其變項分析範例說明如下表 3-5:



表 3-5 總額支付制度類別變項分析範例

研究名稱:總額支付制度實施後醫院因應之道與員工認知程度之相關性研究-  
以某區域教學醫院為例[25]

應變項

看診人次變化、年度收入變化(總醫療費用)

自變項

醫院特性 醫院權屬別、醫院型態別、醫院層級別、門診看診總人  
次、急診總人次、住院總人次

每月醫療費用

門診點數 西醫、中醫、牙醫

住院點數

急診點數

資料來源:本研究整理

由本類別之研究類別我們可以得知，以總額支付制度為類別的研究探求的三大目標:

- 一、依保險對象醫療費用分析: 總額支付制度係政府與醫事機構間談判金錢的制度，與一般民眾較無關聯，所以在自變項使用到一般民眾的項目只在其看診人次，少數有判斷看診民眾年齡的研究，只是探求其年齡與看診次數的關聯性如(第 22 篇)，以便醫事機構針對某年齡層民眾進行資源配置。
- 二、依醫事機構的醫療費用分析: 其應變項多在探討醫療費用支出模式，藉由分析醫事機構每個時程的費用支出，進而預測該機構甚至整體醫療資源費用支出的趨勢。
- 三、以健保收費制度面相關: 有關總額支付制度類型的研究，會針對變更收費制度之後對醫療機構收支費用方面的問題(如第 20 篇)，並且也會針對醫院治療行為模式的改變作探討(如第 19 篇)。

### 3.1.4 部分負擔類別變項分析

在部分負擔類別的部分，我們以第 10 篇「全民健保部分負擔調整對民眾門診醫療利用的影響：以門診高利用者為對象」[17]作為變項分析範例說明如下表 3-6:

表 3-6 部分負擔類別變項分析範例

研究名稱:全民健保部分負擔調整對民眾門診醫療利用的影響：以門診高利用者為對象[17]

應變項

高診次一般診斷病患、高診次具慢性病連續處方籤、門診看診次數

自變項

醫院特性 醫院權屬別、醫院型態別、醫院層級別

病患特質 診斷別、給藥日份

門診看診次數 西醫、中醫、牙醫門診看診總人次

備註:本篇僅探討門診

資料來源:本研究整理

由應變項分析我們可以得知，以門診為類別的研究會探討的三大目標：

- 一、保險對象部分負擔分析：藉由分析特定族群的部分負擔費用調整後，如何影響就醫行為，以此得知某族群，是否因為部分負擔費用提高後，減少了就醫次數，以此探討是否其就醫行為偏差。(如第 25 篇)
- 二、疾病別部分負擔調整：將某病因列為其應變項，探討患得某病因的病人，在部分負擔費用提高或降低後將會如何影響其就醫行為。(如第 24 篇)
- 三、醫事機構資源分析：部分負擔調整後，醫事機構的營運會有何改變，其門診、急診、住院人數將會有什麼變化，也會是其探討的目標。(如第 10 篇)

### 3.1.5 門診類別變項分析

在門診類別的部分，我們以第 27 篇「全民健保對各型醫院生產力之影響」[34] 作為變項分析範例，說明如下表 3-7：

表 3-7 門診類別變項分析範例

研究名稱：全民健保對各型醫院生產力之影響[34]
研究目標
全民健保制度對各型醫院門診生產力的影響
健保資料欄位應用變項
醫院特性      醫院權屬別、評鑑等級

資料來源：本研究整理

由應變項分析我們可以得知，以門診為類別的研究，探討二大目標：

- 一、保險對象門診利用分析：藉由分析特定族群的在使用門診醫療行為時將會面對的問題，例如健康保險實施後，其特定族群在進行醫療行為之時的費用變化，如何影響在門診方面的就醫行為(如第 27 篇)，以此得知某族群，是否因為費用或其他因素(如第 26 篇)，在進行門診時，所可能遭遇的問題。
- 二、疾病別門診利用分析：將某病因列為其應變項，探討患得某病因的病人，在進行門診醫療利用時，所可能遭遇的問題，如醫院層級可能影響某病因的病人，在門診方面進行醫療行為時所會面臨的選擇(如第 26 篇)。

### 3.2 健保收費制度因素的研究

為了更詳細檢視制度變動相關的研究議題，我們再次探討 3.1 節的各項研究變因，特別是有關收費制度的影響的研究，其判斷的準則是由該研究目標項，使用前後交叉分析的方法，探討「收費制度」或「給付制度」等造成之影響為範圍。

我們從整理過後的整體健保相關研究合計 31 篇次(包含四篇跨主題重覆)中，制度影響各層面的研究佔了 12 篇，列示於表 3-8。代表至少有三成三的研究，致力於「健保相關制度」(外在因素)所造成的影響。但現行的「健保資料倉儲系統」，卻僅能提供使用者們檢視收費制度的變化，或「實施日期」或「中止日期」的前後時程，來判斷其當下造成的影響，無法有系統的對制度變化前後做出有效率的查詢。然而，這類重要的變化查詢正是資料倉儲欲保存或設計的主要目標。因為即使相關機構有外部紀錄，但無法直覺地融入既有的報表查詢上；相對地，一定要倚仗專家採用其它工具進行外在分析和探討。如此一來，當後續的研究者欲追溯當年的變化記錄因果為何，往往在資料的整理

和追溯上曠日費時，並容易受限於過去的研究的成果所影響。有鑑於此，本研究將在第四章中，先針對現行的各項健保倉儲分析模型進行更進一步的細緻化設計，使之包容和記載制度因素所影響之維度主檔資訊，以提供有志於此的研究者或實務分析者，有更精確的資料以進行研究和分析。

表 3-8 有關收費制度相關研究

篇次	收費制度類型	研究名稱
1	保險對象部分負擔收費制度	健保部分負擔調整對高血壓病人門診醫療利用之影響[9]
2	保險對象部分負擔收費制度	全民健保部分負擔調整對民眾門診醫療利用的影響：以門診高利用者為對象[17]
3	藥品部分負擔收費制度	藥品部分負擔政策對癲癇病人之影響[19]
4	健保總額支付制度	醫院在參與卓越計畫前後醫療費用變化之探討—以健保局中區分局為例[23]
5	健保總額支付制度	總額支付制度實施後醫院因應之道與員工認知程度之相關性研究—以某區域教學醫院為例[25]
6	健保總額支付制度	全民健保總額支付制度與醫病關係之影響——某醫學中心之個案研究[26]
7	健保總額支付制度	全民健保之總額支付制度下醫院經營效率之研究—DEA 之應用[27]
8	健保總額支付制度	牙醫門診總額支付制度實施後牙醫健保醫療資源使用狀況分析—以高屏區為例[28]
10	健保總額支付制度	全民健保牙科總額支付制度在新竹市縣城鄉醫療服務差異性的探討[29]
11	保險對象部分負擔收費制度	全民健保部分負擔制度對幼童、年長者及慢性病患健康之影響[31]
12	保險對象部分負擔收費制度	全民健保新制部分負擔對不同社會群體影響之探討[32]
13	保險對象部分負擔收費制度	全民健保對各型醫院生產力之影響[34]

資料來源：本研究整理

透過表 3-8 的健保各項收費制度研究整理，我們可以大致上將其分為三類：1. 「保險對象部分負擔收費制度」、2. 「健保總額支付制度」、3. 「藥品部分負擔收費制度」。我們將在第四章進一步探討有關這些類型的健保收費制度研究。

## 肆、模型建構

在本章將建置 3.2 節所提出的健保收費制度相關因素模型，探討制度因素如何造成影響的主題資料超市之建構。

有關健保資料收費制度方面的研究，可以分為三種類型：1. 「特定身分部分負擔」、2. 「總額支付制度」、3. 「藥品部分負擔」，由於「總額支付制度」為單一的制度變動，例如「醫院在參與卓越計畫前後醫療費用變化之探討——以健保局中區分局為例」[23]，總額支付制度的變動，是健保局與醫事機構之間支付費用計算方式變動，此項收費制度變動不會經常發生，是由健保局在擬定後行諸全國，自民國 87 年以來逐漸針對各醫事機構實施後就不再變動[26]，所以如有變動，也僅需以「SCD2」的方式記錄下來即可。然而有關特定身分部分負擔的收費制度變動和藥品部分負擔收費制度，相對變動比較頻繁，雖亦可使用型 2 的緩時變維度進行紀錄，但鑑於研究者對該類議題的高度重視，因此我們將在 4.1 節建置部分負擔收費制度和藥品部分負擔收費因素模型之資料超市分析模型。

本節將依 Kimball 維度模型的建置步驟[39]分為三個小節：4.1 節將依「業務變動要素分析」，首先確立超市資料主題，4.2 節決定事實表格中的資料粒度(Grain)，粒度為企業流程中最小層級的資料單位，粒度大小會影響資料存放量的多寡，4.3 節將選擇模型中的事實資料，將其納入事實表格。由於我們主要是針對特定身分部分負擔的收費制度變動和藥品部分負擔收費制度的變化，為其分別設計兩個子題。

### 4.1 業務變動要素分析

根據 3.3 節的分析結果，我們可以得知，國內針對收費制度的變化，如何造成影響就醫行為、費用支出的研究類型佔了整體健保研究的一大部分，而影響費用的因素，正是我們重要的分析目標。

#### 子題一：健保部分負擔收費制度

透過填入下表 4-5 我們可以初步了解和健保部分負擔收費制度相關的因素：

表 4-1 健保部分負擔收費制度因素基本項目

健保部分負擔收費制度因素	受保人特殊身分	負擔費率	級距	眷口別	日期	縣市區域
負擔費率變動	○	○	○	○	○	○
補助差額調整	○	○	○	○	○	○
級距調整	○	○	○	○	○	○
眷口別自負率調整	○	○	○	○	○	○

投保單位 負擔率調 整	○	○	○		○	○
所得稅率 級別調整	○	○		○	○	○

資料來源:本研究整理

以第一列來說明表 4-1，當負擔費率變動時，會以「受保人特殊身分」、「負擔費率」、「級距」、「眷口別」、「日期」、「縣市區域」來進行分析。由健保局網站提供的特定身分群組收費標準，我們可以得知，各特定身分群組的「部分負擔」標準不同，而此為最重要的部分負擔收費制度因素，無論是重大傷病或是其他疾病，其特定身分群組間的差異，才是部分負擔費用的計算差別的區分方式。

我們採用列表的方式，點出分析模型的分析目標，藉此可以用於部分負擔收費制度因素可能會影響的目標，以此作為建構部分負擔收費制度因素模型的第一步，所以我們可以初步的描繪出基本的部分負擔收費制度因素模型如下圖 4-1。



圖 4-1 部分負擔收費制度因素初步建構維度模型圖

從上圖我們可以得知子題一是用用來捕捉被保人的收費制度改動所影響的保費金額之事實，我們可以了解影響部分負擔制度因子對被保人投保金額的變化。

### 子題二：藥品部分負擔

藥品部分負擔的部分最原始是健保局依照影響各藥物價格、治療效果和成本的考量以進行必要的調動。一般來說，研究者關心的議題主要是和疾病的相關，所以在業務變動要素分析時，上我們把這個子題訂在被保險人領藥單的明細。

除此之外，健保當有針對被保險人的藥品總額藥價之自付比率進行調整時，這部分則可簡化成一個藥品給付維度再加上緩時變(SCD2)的紀錄方式，連結到現行的主題一到三去做重要量值(KPI)的追蹤。此維度如表 4-2 所示。

表 4-2 藥品自部分負擔費用

藥費	部分負擔費用
100元以下	0元
101~200元	20元
201~300元	40元
301~400元	60元
401~500元	80元
501~600元	100元
601~700元	120元
701~800元	140元
801~900元	160元
901~1000元	180元
1001元以上	200元

## 4.2 部分負擔收費制度因素模型粒度

### 子題一：健保部分負擔收費制度

此部分最主要是記錄各最主要是依照影響各被保人在自身的「特殊身分」、「負擔費率」、「級距」、「眷口別」、「日期」、「縣市區域」等有所變化時即觸動一份快照來記錄事實資料的改動。所以資料精細度即定義為每人、每種特殊身分、每種負擔費率、每種級距、每種眷口別、每所在的縣市區域和每段期間中的保費負擔金額。

除此之外，當研究者若有更進一步的研究興趣—想對每不同保費負擔金額的看診情形有興趣時，可以將子題一的事實資料表在不影響主題一或主題二的資料精細度的情形下，更進一步當成一個維度以連接到主題一或主題二的事實資料表中，進行更多的事實衡量數據分析。

### 子題二：藥品部分負擔

藥品部分負擔的部分因為相比較單純，它是一種維度主檔資料變動的概念，故這部分我們建議以每人、每次看診紀錄、每個病症和每個藥品最為事實表格的精細程度。換言之，以健保局目前主題一、二和三的精細程度而言，它是以單一看診紀錄做為資料粒度，本子題二的粒度要求則是更細節到每次看診的領藥明細之精細程度。

## 4.3 部分負擔收費制度因素模型事實表格

### 子題一：健保部分負擔收費制度

在本節我們將前述的整體事實表格及維度表格中作最後的整合，如圖 4-2 所示。

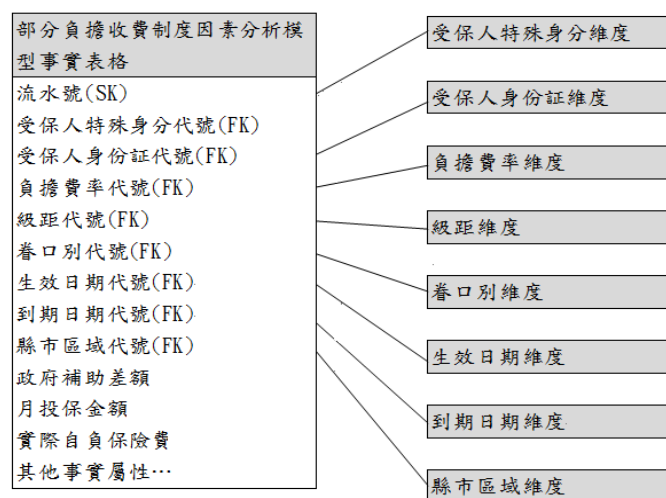


圖 4-2 部分負擔收費制度因素維度模型

「政府補助差額」、「月投保金額」與「實際自負保險費」為事實表格中的數值型欄位，可用來加總或做管理運算之用。此外，我們可利用此事實表格當做主題一到三的維度表格，以利用主題一到三的所有數據衡量對制度面改動的問題做更深入的追蹤。

### 子題二：藥品部分負擔

我們將前述的整體事實表格及維度表格中作最後的整合，如圖 4-3 所示。



圖 4-3 藥品部分負擔收費制度因素維度模型

我們將屬於「退化維度」(Degenerate-Dimension)的事實屬性編入事實表格，「藥品部分負擔收費制度因素維度模型」的退化維度包含：「藥台號碼」、「領藥號」、「藥袋數」，「退化維度」不作為獨立維度由事實表向外延伸，根據維度建置原則，因此將其放置於「藥品部分負擔收費制度因素分析模型事實表格」之中。

而「處方天數」、「發藥量」、「請領金額」、「藥品用法」與「藥品自負金額」為事實表格中的數值型欄位，可用來加總或做管理運算之用。

有關記載藥品總額金額負擔制度的改變之資料模型，由於我們的設計可以令其整合至現行之健保資料倉儲系統架構，故此不再贅述。

## 伍、結論與未來研究

「臺灣中央健保局」的資訊量成長十分驚人，以現行納保率達 99% 以上的程度，其資訊成長量可想而知，透過資料倉儲的技術，將這些大量資料轉化為可操作的資訊，並且縮短處理的時間，一向是人們努力的方向，然而因為「健保資料倉儲」能夠提供研究者們處理過後的結構化資訊，使得研究者不必花費大量時間去處理專業領域以外的額外工作，並使得健保方面的決策者、執行者、研究者們，能夠對於民眾們的健康情形有所掌握，在醫療資源方面更能有根據的進行配置，不至於浪費納保民眾的稅金。本章共為二節，5.1 節歸納本研究所獲得的結論；5.2 節則提出未來研究方向與相關建議。

## 5.1 結論

本研究主要是從回顧國內碩博論文對健保局所提供資料中，歸整研究者所關心的議題，並同時回顧健保相關的資料庫倉儲概念後，再對兩者所關注的議題進行比對，以提出可進一步建構的資料超市主題及其模型。主要的研究成果為：

1. 目前的倉儲架構可再將主檔隨時間變化的訊息予以計入，以提供更細緻化的查詢。
2. 有關健保制度面的資料改動部分，這部分有許多研究學者加以關注，但現行倉儲架構卻較為少見，故我們提出「主題四—制度變動之分析模型」，該模型亦可在不破壞現行健保倉儲的「粒度」條件之下，和暨有模型相結合，以期提供更多管理數據做為實務和學術領域的分析與應用。

## 5.2 未來研究

本研究目前關注的議題，是在健保相關研究之中，提出部分負擔的因素在健保議題之中佔了一個重要的部分，但是我們受限於進行國內研究者的研究論文的「變項分析」，因為要齊全的搜集外國的健康保險相關研究論文，不只是蒐集難度的完整性問題，甚至也有不以英文撰寫的解讀上的困難，因為並不是各國都有像臺灣的健康保險制度一樣，擁有大批的官方開放文獻可供研究，也並不是都有齊全的研究資料可供參考，因此在此同時，可能已有許多新議題正在進行，也因此，我們無法去相互比較「臺灣中央健保局」的資料倉儲系統架構與國外的健康保險各項數位化機制的優缺點如何，另外本研究提出的「制度變動之分析模型」，雖然以不破壞粒度性的方式與現行架構作結合，但現行運作中的健保資料倉儲系統所有權仍然屬於「臺灣中央健保局」，在實作的部分，目前的階段尚且未能驗證果效為何，所以在未來研究的方向，可往進行本研究的資料倉儲架構，套入未來受「部分負擔因素」影響的受保險人資料後，所產生的「累計事實」，以此作為根據，進一步探討受保險民眾們，對於醫療資源的付出意願，或者是醫事機構的醫療資源配置是否具有問題等等…更深入的問題進行，或者將本研究作為增加多維度分析構面的範例，更廣泛的深入探討其他領域同樣重要的資料倉儲系統架構。

## 參考文獻

### 中文部分

- [1] 行政院衛生署中央健康保險局；全民健康保險簡介；  
[http://www.nhi.gov.tw/webdata/webdata.asp?menu=17&menu\\_id=659&webdata\\_id=2891&WD\\_ID=897](http://www.nhi.gov.tw/webdata/webdata.asp?menu=17&menu_id=659&webdata_id=2891&WD_ID=897)；2010/12/26。
- [2] 財團法人國家衛生研究院；全民健康保險研究資料庫；  
<http://w3.nhri.org.tw/nhird/>；2010。
- [3] 朱自強；設計製作以網際網路架構為基礎的健保資料倉儲系統-以醫療費用分析為例；國立陽明大學公共衛生研究所碩士論文；2000。
- [4] 台灣博碩士論文知識加值系統；  
<http://ndltd.ncl.edu.tw/cgi-bin/gs32/gsweb.cgi/login?o=dwebmge&cache=1307878267708>，2011。
- [5] 明雲青；線上交易處理(OLTP)；



[http://www.digitimes.com.tw/tw/dt/n/shwnws.asp?Cn1ID=10&id=0000125638\\_6PY53PRA435DQ27EB3KFW](http://www.digitimes.com.tw/tw/dt/n/shwnws.asp?Cn1ID=10&id=0000125638_6PY53PRA435DQ27EB3KFW)； 2009/03/11。

- [6] 嚴紀中、陳鴻基；管理資訊系統理論科技實務與應用；臺北；松崗電腦圖書資料股份有限公司；1999。
- [7] 梁定澎；資料倉儲與資料專櫃；〈決策支援系統與企業智慧〉；ch. 4；臺北；智勝；2006
- [8] 顧漢凌；全民健保體系下醫療資源分佈對個人醫療使用的影響；東華大學國際經濟研究所碩士論文；2005。
- [9] 陳陸英；健保部分負擔調整對高血壓病人門診醫療利用之影響；東華大學國際企業研究所碩士論文；2008。
- [10] 劉雅婷；無職榮民與一般民眾醫療利用之比較分析；台北護理學院醫護管理研究所碩士論文；2007。
- [11] 陳大申；臺灣地區精神疾病醫療利用分析；長庚大學醫務管理學研究所碩士論文；2002。
- [12] 洪瑞屏；初次診斷憂鬱症患者追蹤三年之醫療利用與費用分析探討；亞洲大學健康管理研究所碩士論文；2005。
- [13] 簡以嘉；台灣全民健保精神疾病盛行率、醫療利用及費用分析研究；陽明大學公共衛生研究所博士論文；2003。
- [14] 蔡述信；以台灣全民健保承保抽樣歸人檔案分析憂鬱症病患的醫療利用；中國醫藥學院醫務管理研究所碩士論文；2002。
- [15] 鍾孝其；婦女醫療利用分析—以剖腹產、子宮切除、子宮頸抹片、乳癌、子宮頸癌為例；中國醫藥學院醫務管理研究所碩士論文；2002。
- [16] 莊媿婷；以全民健保 1996-2001 年承保抽樣歸人檔分析下背痛病人之醫療利用；臺灣大學醫療機構管理研究所碩士論文；2003。
- [17] 涂宜均；全民健保部分負擔調整對民眾門診醫療利用的影響：以門診高利用者為對象；臺灣大學衛生政策與管理研究所碩士論文；2002。
- [18] 鄭瑞英；老年人在全民健保下中西醫就醫行為研究；中國醫藥大學醫務管理研究所碩士論文；2003。
- [19] 陳厚任；藥品部分負擔政策對癲癇病人之影響；臺灣大學醫療機構管理研究所碩士論文；2003。
- [20] 梁淑勤；全民健保中醫門診利用之研究；中國醫藥學院環境醫學研究所碩士論文；2002。
- [21] 劉蓓雯；中央健保局年度醫療費用支出之預測模式研究—灰預測理論之應用；朝陽科技大學保險金融管理研究所碩士論文；2005。
- [22] 余健釗；全民健保醫療費用成長模型建構之研究；台北大學統計學研究所碩士論文；2003。
- [23] 陳炫達；醫院在參與卓越計畫前後醫療費用變化之探討—以健保局中區分局為例；臺中健康暨管理學院健康管理研究所碩士論文；2004。

- [24] 劉琪；門診藥品部分負擔對醫療院所用藥行為及費用影響之研究；臺灣大學會計學研究所碩士論文；1999。
- [25] 楊榮真；總額支付制度實施後醫院因應之道與員工認知程度之相關性研究-以某區域教學醫院為例；中華大學經營管理研究所碩士論文；2006。
- [26] 熊心如；全民健保總額支付制度與醫病關係之影響——某醫學中心之個案研究；東海大學管理碩士學程在職進修專班碩士論文；2005。
- [27] 陳宗義；全民健保之總額支付制度下醫院經營效率之研究—DEA 之應用；義守大學管理科學研究所碩士論文；2003。
- [28] 鍾尚衡；牙醫門診總額支付制度實施後牙醫健保醫療資源使用狀況分析—以高屏區為例；高雄醫學大學口腔衛生科學研究所碩士論文；2002。
- [29] 李俊德；全民健保牙科總額支付制度在新竹市縣城鄉醫療服務差異性的探討；高雄醫學大學口腔衛生科學研究所碩士論文；2002。
- [30] 黃志峰；牙醫師對總額支付制度滿意度之研究；中正大學企業管理研究所碩士論文；2001。
- [31] 黃昱瞳；全民健保部分負擔制度對幼童、年長者及慢性病患健康之影響；陽明大學公共衛生研究所碩士論文；2007。
- [32] 洪明皇；全民健保新制部分負擔對不同社會群體影響之探討；政治大學財政學研究所碩士論文；2000。
- [33] 賴奕菁；門診老年失眠症患者之安眠藥處方型態及處方問題；國立陽明大學衛生福利研究所碩士論文；2003。
- [34] 王慕凡；全民健保對各型醫院生產力之影響；中正大學會計學研究所碩士論文；1999。

#### 外文部分

- [35] Admin；” ETL、Staging Area、Data marts” ；<http://d-h-n.de/blog/?p=556>；2009。
- [36] Kenneth, C. Laudon. ；Management Information Systems；臺北；東華書局；47-49；2002。
- [37] W. H. Inmon；” What is a Data Warehouse” ；New York；John Wiley & Sons Inc. ；1994。
- [38] W. H. Inmon；” Building the Data Warehouse” ；New York；John Wiley & Sons Inc. ；1996。
- [39] Ralph Kimball、Margy Ross；The Data Warehouse Toolkit：the Complete Guide to Dimension Modeling；New York；John Wiley & Sons Inc. ；2002。

# Data Warehousing for National Health Insurance Research

## Database

Yu-Chin Liu, Chun-Jen Chen  
Dep. of Information Management  
Shih Hsin University  
Taipei 116, Taiwan, R.O.C.  
ycliu@cc.shu.edu.tw  
ji394su33000@hotmail.com

### Abstract

Since the Taiwan National Health Insurance Database recorded all medical data from 1999, NHID has already gathered a huge amount of information. Hence, many researchers have done lots work based on NHID. The purpose of our work is to establish a data warehouse system to accommodate most research topics which most researchers are interested. .

We first explore the thesis and dissertations maintained in National Digital Library of Taiwan and several major research themes are recognized. Furthermore, we compare them to existing NHID DW architectures. These literatures show over half of researchers pay lots attentions on events related time dimensions, but however existing NHID DWs don't take care about them. Therefore, our architecture which covers most research interests related to NHID is proposed.

KeyWords: Data Warehouse 、 Premium Contribution 、 Variables Analyzed 、 National Health Insurance Database 、 Data Mart