
大型資料庫之介紹 及精算上之應用

政治大學統計系余清祥

1

演講大綱

- 什麼是「大數據（Big Data）」？
- 大數據的特性
- 大數據的限制
- 大數據與保險業
- 大數據的未來發展

2

什麼是大數據？

▣ 大數據（Big Data；或譯為「巨量資料」）在2010年由IBM所提出，大致涵蓋了三個V：

→ 大量化(Volume)：至少TB及PB以上

→ 多樣化(Variety)：視頻、GIS等多樣性

→ 快速化(Velocity)：即時處理

註：根據科技研究公司IDC作出的估測，數據一直都在以每年50%的速度增長。

3

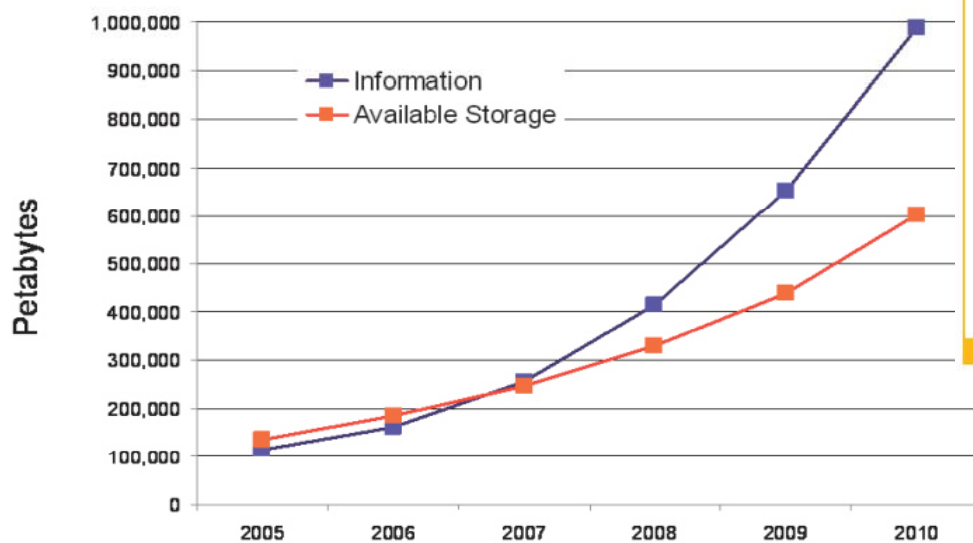
電腦容量單位的演變（資料爆炸！）

Bit (b)	1 or 0
Byte (B)	8 bits
Kilobyte (KB)	1,000 bytes
Megabyte (MB)	1,000 KB
Gigabyte (GB)	1,000 MB
Terabyte (TB)	1,000, GB
Petabyte (PB)	1,000 TB
Exabyte (EB)	1,000 PB
Zettabyte (ZB)	1,000 EB

4

「資料大爆炸」時代的來臨！

Information Versus Available Storage



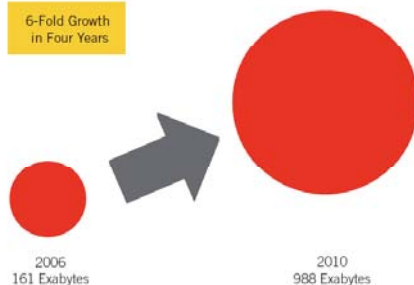
Source: IDC, 2007

出處：The Expanding Digital Universe,
A Forecast of Worldwide Information Growth Through 2010,
March 2007, An IDC White Paper - sponsored by EMC
<http://www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf>

Figure 1

Information Created, Captured and Replicated

6-Fold Growth
in Four Years



Source: IDC, 2007

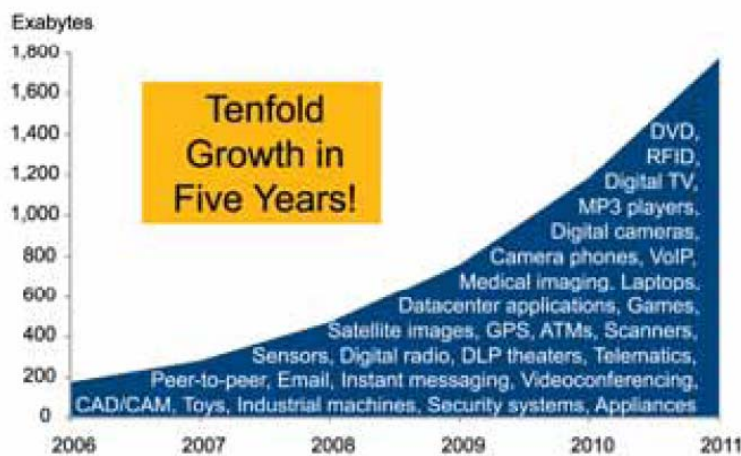
2007年，IDC預估
2010年會成長**六倍**！
(與2006年比較)

2006 161 EB
2010 988 EB (預測)
Note: International Data
Corporation (IDC)

「資料大爆炸」時代始於2007年

Figure 1

Digital Information Created, Captured, Replicated Worldwide



Source: IDC, 2008

2009年，IDC預估2011
年會成長**十倍**！
(與2006年比較)

2006 161 EB
2007 281 EB
2010 988 EB (預測)
2011 1773 EB (預測)

出處：The Diverse and Exploding Digital Universe,
An Updated Forecast of Worldwide Information Growth Through 2011
March 2008, An IDC White Paper - sponsored by EMC
<http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>

巨量資料時代的來臨！

- 資料累積速度遠超過以往。。。。。
 - 人類基因密碼重組歷經十年，才完成三十億個鹼基對的定序，現在只需一天；
 - 美國股市每日成交約七十億股，其中三分之二由數學模型分析後自動交易；
 - 臉書每小時上傳至少一千萬張照片，每天超過三十億次「讚」，可供分析使用者喜好。
-

7

大數據的特性



http://www.abc.net.au/reslib/201204/r926508_9670000.jpg

8

資訊與決策

- ❑ 過去擁有土地、資金等的資本家，透過勞力及資本密集而具有絕對優勢；21世紀有另一個重要轉變，將成為知識經濟的時代，早一步擁有充分資訊的人，會握有絕對優勢。(Yahoo、Google！)
 - 未來優勢為資料、處理即分析資料的能力（軟實力），因為被複製的可能性高，需要法律保障（智慧財產權、版權）。
-

9

資訊的價值

- ❑ 1990年代後期Amazon網站雇用十幾位書評及編輯，提供推薦閱讀的書單，其銷售量卻比不上以讀者回饋產生的建議，最後解散書評團隊。（現在1/3業績來自於電腦推薦。）
 - ❑ 2004美國Walmart開始運用歷史交易記錄，發現颶風來臨前，手電筒、小甜點Pop-Tarts銷售量大增。（註：尿布與啤酒也是另一個知名範例！）
-

10

沃馬特量販店(Wal-Mart)

❑ 沃馬特最先蒐集、分析顧客資料，並以整理所得的資訊，提高銷售業績。

→ 分析發現美國消費者在週末購物時，許多人會同時購買尿布及啤酒。

❑ 問題：為什麼這兩種商品會一起購買？又如何將這份資訊轉變為業績？

註：沃馬特從美國西南部發跡，剛開始只是一家五金行，現在已是全美最大的百貨零售業者。

資訊與知識的價值

❑ 資料挖掘(Data Mining)的範例：\$ \$ \$ \$

→ 協助超級市場促銷及陳設商品

Milk, eggs, sugar,
bread



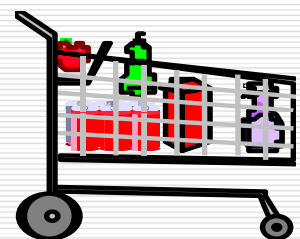
Customer1

Milk, eggs, cereal, bread



Customer2

Eggs, sugar



Customer3

「尿布與啤酒」的延伸價值

- ❑ 尿布與啤酒屬於關連性(Association)的關係，與常見的因果關係(Causality)不同。
 - ❑ 關連性的價值未必低於因果關係，像是尿布與啤酒的關連，可用於：
 - 商品定價與促銷；
 - 商品擺設（商場動線）；
 - 商品倉儲。
-

大數據的幾個基本特性

- ❑ 除了數量龐大外，大數據還有以下幾個特性：
 - 「樣本=母體」
 - 不精確且含有雜訊的資料
 - 相關性而非因果關係



「樣本=母體」

□ 有別於以往資料蒐集限制，大數據強調完整的資料，將不只依賴由樣本推論出整個母體的特性。

→ 更能即時反映資料的特性（傳統普查）；

→ 去除抽樣造成的偏誤；

→ 樣本數不足（局部解析度）。

□ 大數據也需有配套措施，像是更新、儲存、分析資料庫等問題。



<http://www.smartercomputingblog.com/wp-content/uploads/2012/06/where-is-your-data.png>



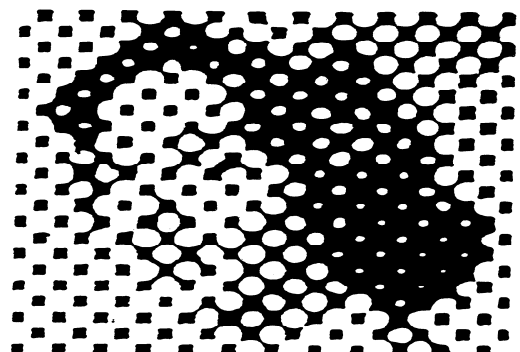
500,000



2,000



1,000



250

樣本對母體之代表性

不精確且含有雜訊的資料

- 資料量愈多、愈有可能不正確，除此之外，大數據背後還隱藏其他類型的「雜亂」：
 - 結合不同源頭、不同類型的資料，產生的相容性問題。(Meta Analysis；整合型分析)
 - 資料格式不一致的調整。
- 資料數量比資料品質重要！
 - CPI官方統計vs.網頁五十萬項商品價格(例：雷曼破產後通貨緊縮)



http://blog.backupify.com/2013/04/01/the-5-but-really-the-6-mistakes-made-in-big-data/00009-08_sept253-jpg/

相關性而非因果關係

- 因果關係未必可由資料分析、數理模型中看出，像是小學生的拼字能力與腳丫大小，背後可能還有其他原因。
- 完整資料、更好的工具，可讓我們能更快、更有效率地找出相關性。（預警！）
 - UPS從2000年起進行「預防性維修」；
 - 紐約市「檢修人員出入孔」的爆炸預測。

大數據的限制

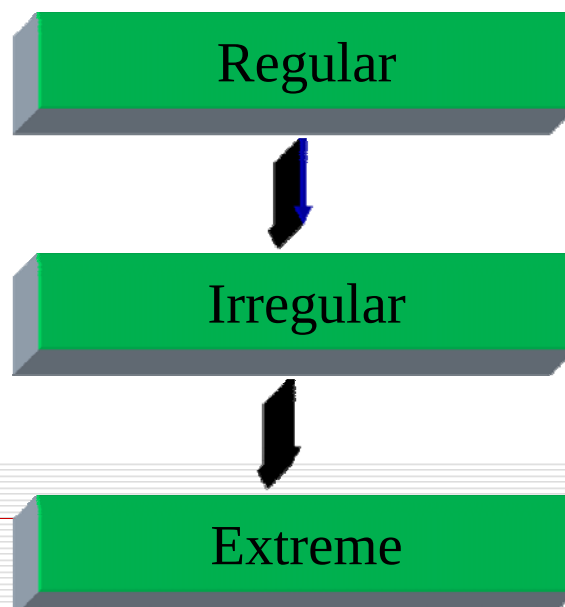


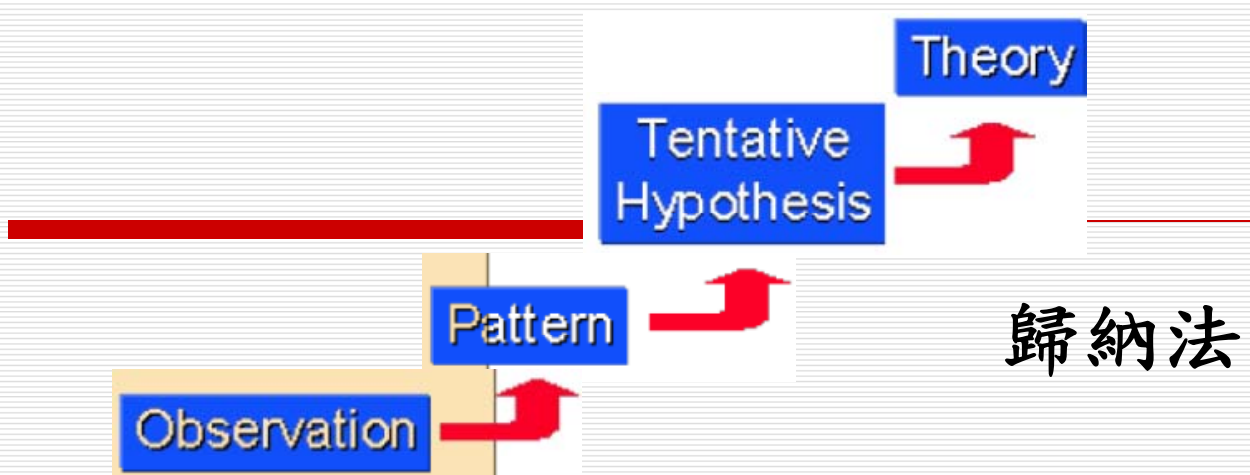
<http://www.tnooz.com/2012/01/04/how-to/big-data-and-the-infinite-possibilities-for-the-travel-industry/>

19

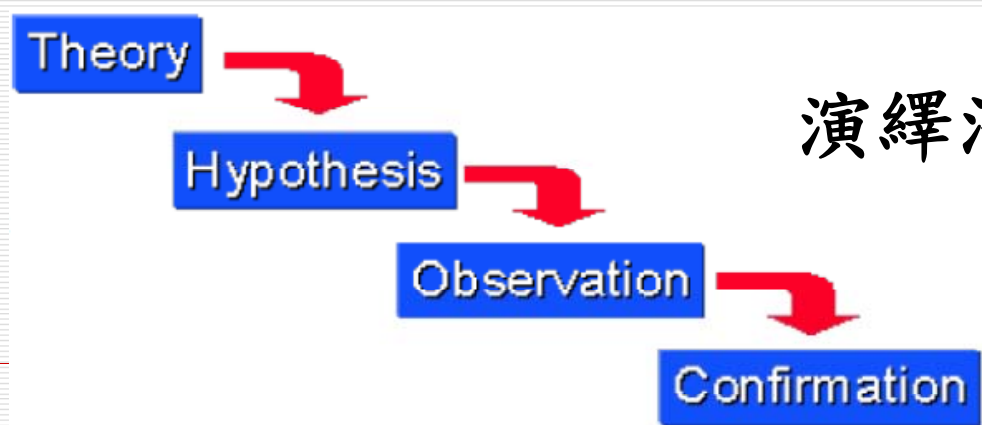
統計與知識

- 統計等計量分析方法屬於歸納法(Induction)，從龐雜的資料找出共同趨勢，並區分資料具有以下哪一種特性：





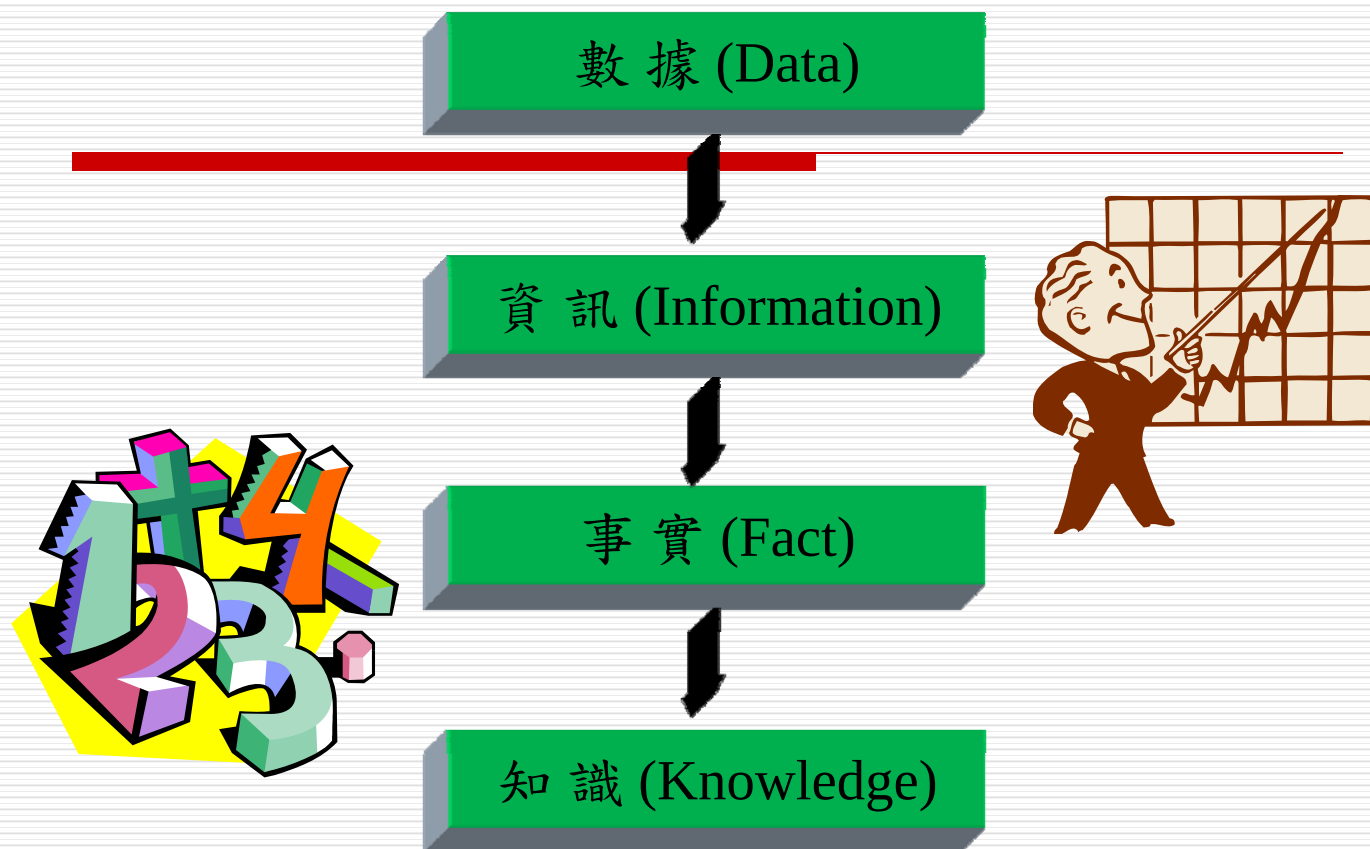
歸納法



演繹法

資訊爆炸 vs. 分析解讀

- 日常生活中到處充斥「資訊」，哪些真正需要的關鍵因素？
- 哪些是必要資訊？
- 資料品質？（「Garbage in, garbage out」）
- 如何根據既有資訊判斷？（哪些決策的風險較高、如何降低風險？）



試誤法(Trial and Error)

- 因為大數據的龐雜，不易規範資料分析的標準操作程序(SOP)，較為可行的相關性分析或許必須透過試誤法。
- 卡古(Kaggle)公司常舉辦資料挖掘比賽，2012年對二手車的相關分析中，發現橘色烤漆的車故障率較低。
- 以2008年紐約人紀錄預測2009年危險人孔蓋名單，高達44%正確，其中最重要的指標是「電纜年份」。

資料蒐集的方式

□ 常見資料分類之一，是依據資料來源：

1. 實驗設計(Experimental Design)

2. 觀察研究(Observational Study)

→ 兩者的差異在於資料蒐集者的參與，蒐集資料並不影響觀察研究，像是研究股市、利率、房地產價格，與實驗設計控制變因獲得觀察值不同。

註：建議資料分析著重於「相關性」！

大數據與保險業



<http://h30507.www3.hp.com/t5/Journey-through-Enterprise-IT/Analyze-This-Big-Data-is-insurance-against-losing-a-competitive/ba-p/143577#.UgZmpLQVEqQ>

大數據與保險業

- ❑ Google在2009年分析搜尋引擎中常見的關鍵字（每天至少三十億筆），與美國疾病管制局(CDC)在2003年至2008年的流感傳播資料比對，比CDC的通報資料（延遲一週）更早發現H1N1的散佈。
 - ❑ 英國學者John Snow在1854年，透過研究倫敦下水道及街道，發現霍亂傳播與受污染的水有關。
- Spatial Statistics（空間統計）



保險業新的發展方向？

- ❑ 大數據可保障競爭優勢(Big Data is insurance against losing a competitive edge)
- 除了利率風險外，保險業未來面臨的風險更為多樣，像是年金險及健康險的延壽風險，或是各類商品皆會遭遇的弱體、超乎預期的投資風險。
- 註：在此僅以保障型商品為目標，考量及討論與訂價相關的風險因素。
-

保險業相關的大型資料庫

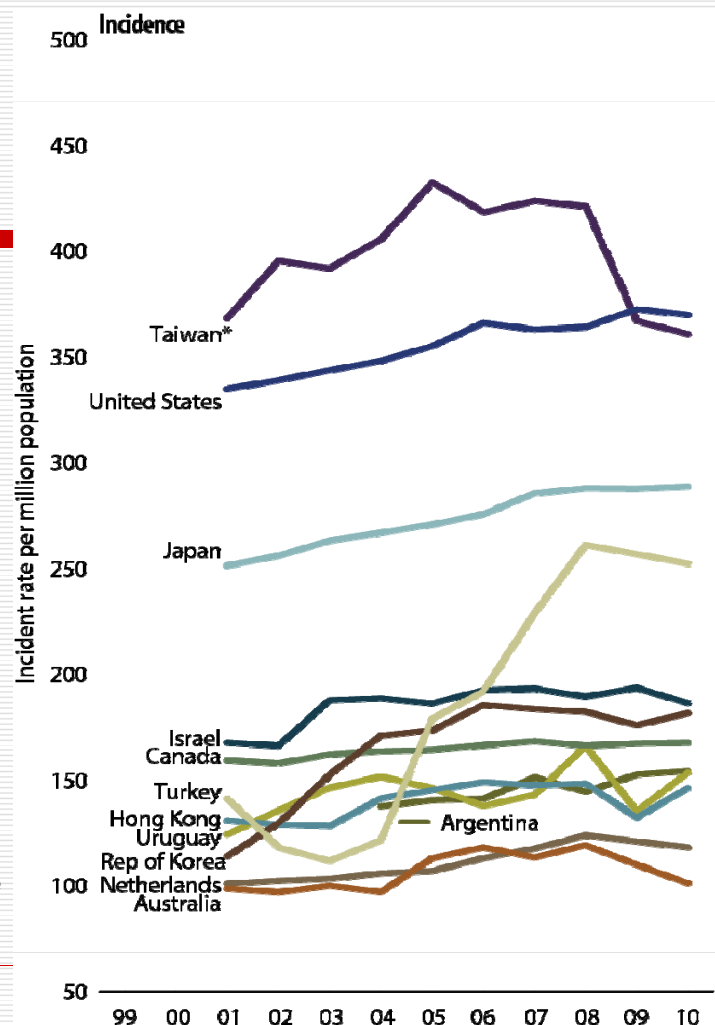
❑ 除了整個壽險業、各保險公司本身、再保公司提供的經驗發生率，國內外也有相關的大型資料庫，可供壽險業者設計商品、評估風險等目的：

→ 壽險：內政部統計處、Human Mortality Database (HMD)；

→ 健康險：健保資料庫（健康加值中心）、美國腎臟學會、美國精算學會(SOA)。

Comparison of unadjusted ESRD incidence worldwide
Figure 12.1 (Volume 2)

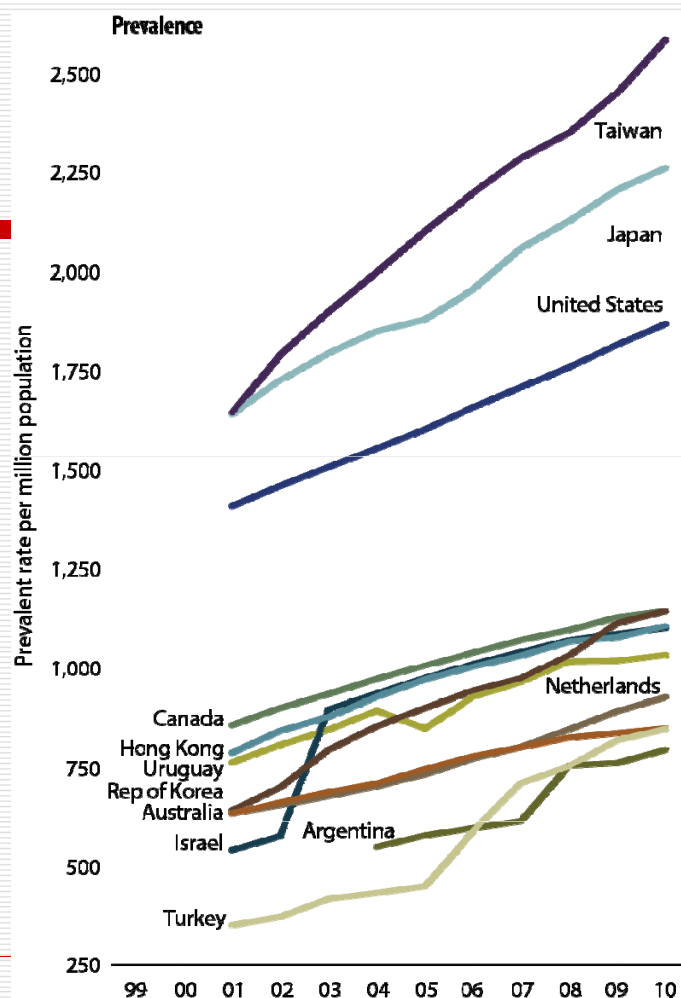
All rates are unadjusted. Data from Argentina (2005–2007), Japan, & Taiwan are dialysis only. *Downturn in incident rates is due to changes in criteria for incidence and in the payment system.



Comparison of unadjusted ESRD prevalence worldwide

Figure 12.1, continued (Volume 2)

All rates are unadjusted. Data from Argentina (2005–2007), Japan, & Taiwan are dialysis only.

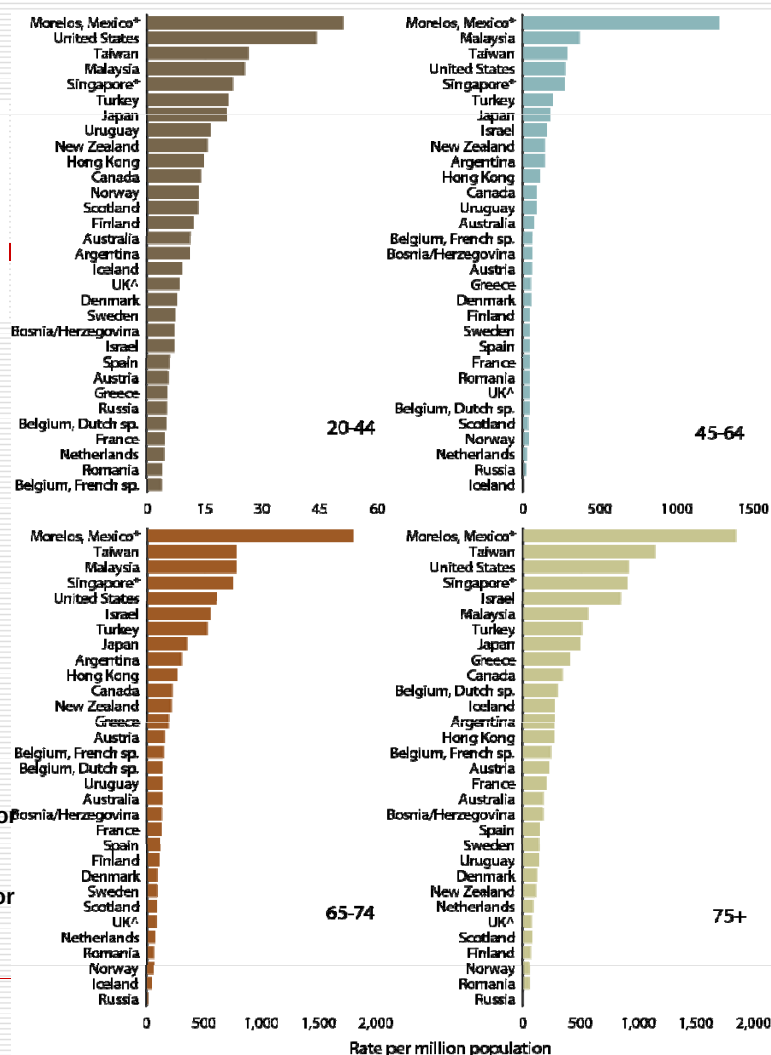


USRDS 2012 ADR

Percentage of incident patients with ESRD due to diabetes, by age, 2010

Figure 12.5 (Volume 2)

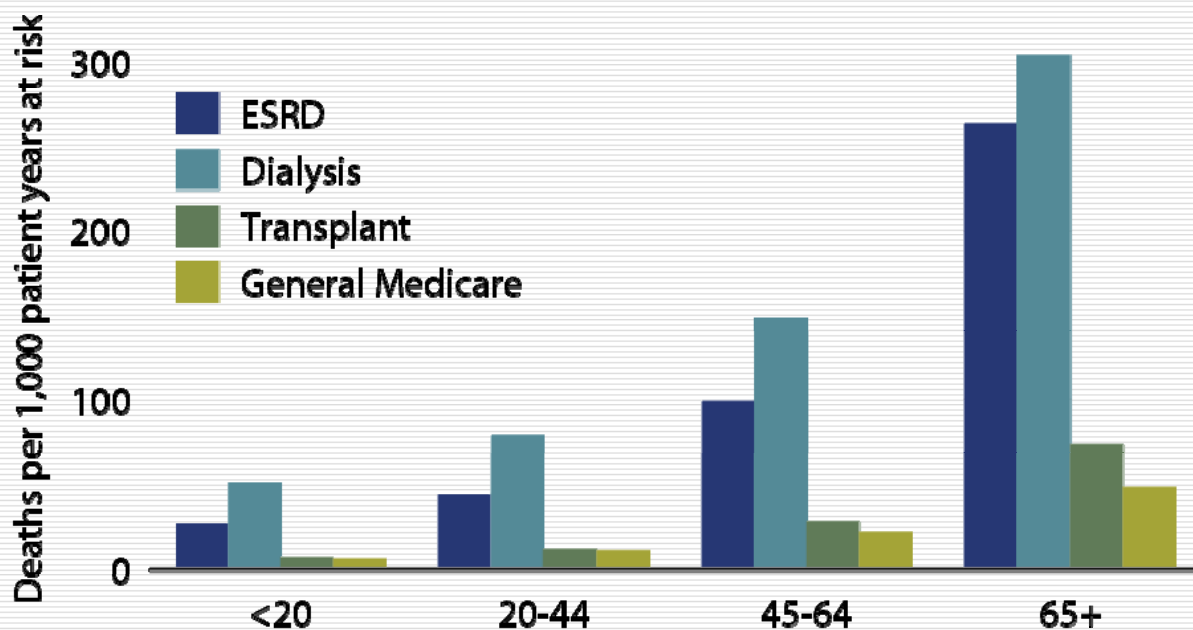
Data presented only for countries from which relevant information was available. All rates unadjusted. ^UK: England, Wales, & Northern Ireland (Scotland data reported separately). Data for Belgium & England/Wales/Northern Ireland do not include patients younger than 18. *Latest data for Singapore & Morelos (Mexico) are for 2009. Data for France include 23 regions in 2010.



USRDS 2012 ADR

Adjusted all-cause mortality in the ESRD & general populations, by age, 2010

Figure 5.2 (Volume 2)

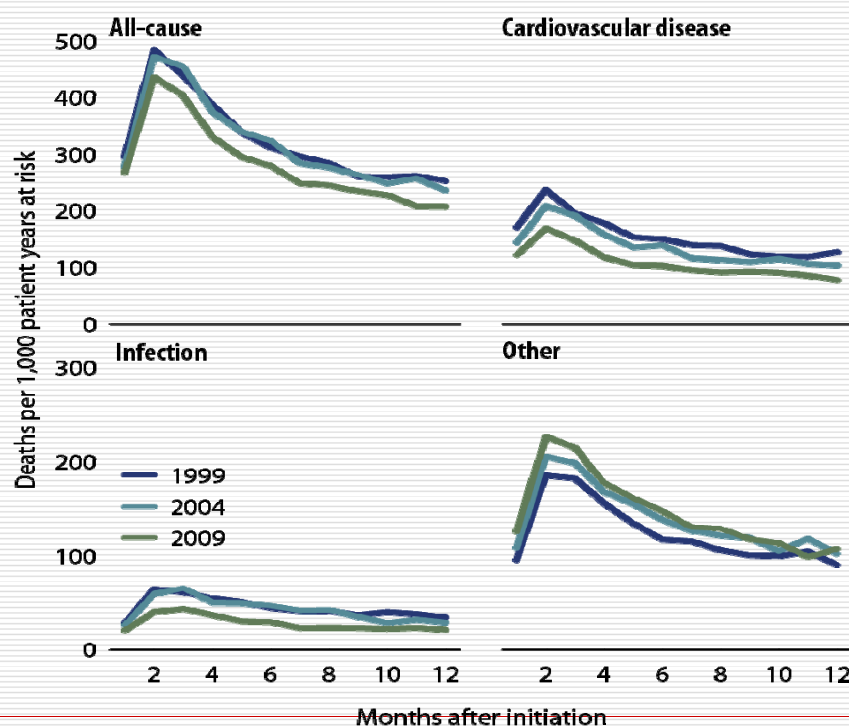


Prevalent ESRD & general Medicare (non-ESRD) patients. Adj: gender/race; ref: Medicare patients, 2010.

USRDS 2012 ADR

Adjusted all-cause & cause specific mortality (from day one) in the first year of hemodialysis

Figure 5.3 (Volume 2)

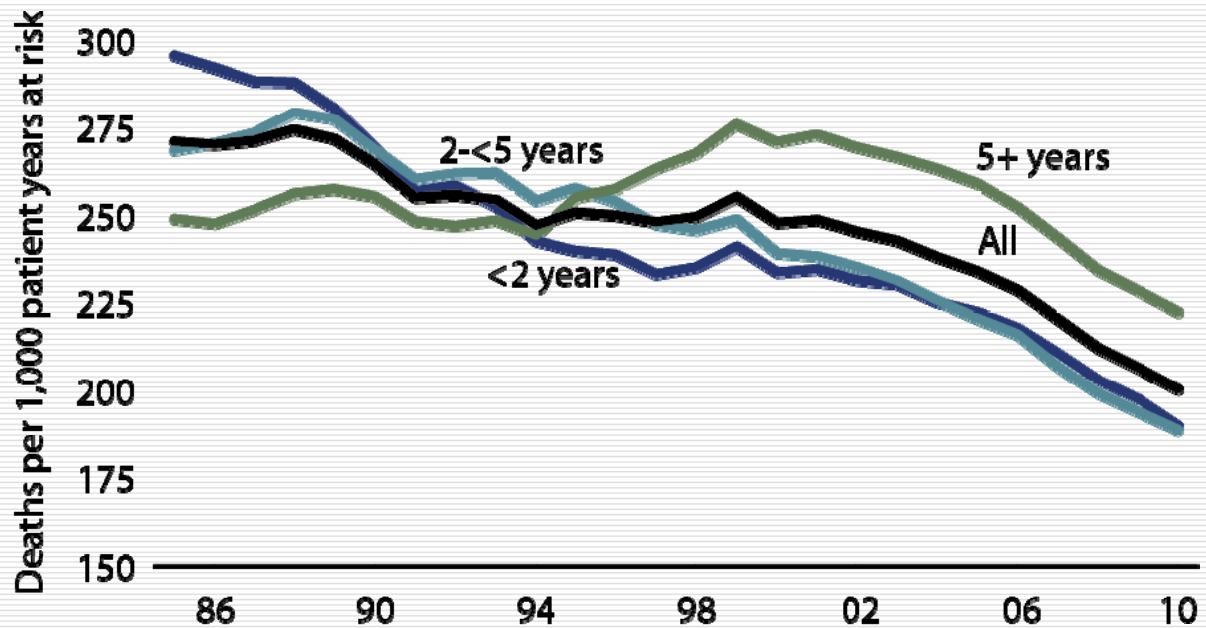


Incident hemodialysis patients defined on the day of dialysis onset, without the 60-day rule. Adj: age/gender/race/Hispanic ethnicity/ primary diagnosis; ref: incident hemodialysis patients, 2005.

USRDS 2012 ADR

Adjusted all-cause mortality in prevalent hemodialysis patients, by vintage

Figure 5.4 (Volume 2)



Incident hemodialysis patients. Adj: age/gender/race/primary diagnosis; ref: incident hemodialysis patients, 2005.

USRDS 2012 ADR

大數據
的未來發展！



<http://cdn.marketingtechblog.com/wp-content/uploads/2013/05/Screen-Shot-2013-05-28-at-11.22.05-AM.png>

未來發展的幾個考量因素

- ❑ 未來保險商品訂價，會與大數據愈來愈有
關連，但仍須注意以下事項：
 - 資料的取得與更新：公司內部、同業合作
（財產權？）、政府與再保資料
 - 資料的維護與使用：安全性（雲端？）
IRB
 - 資料的分析與整合：R&D、產學合作、
Bayesian Credibility
-

人體試驗委員會 (IRB)

- 人體試驗委員會
IRB : Institutional Review Board
 - 研究倫理委員會
REC : Research Ethics Committee
 - 倫理審查委員會 - 人體研究法
-

赫爾辛基宣言 (Declaration of Helsinki)

- 國際醫學協會針對人體試驗，在1964年提出了赫爾辛基宣言。
 - 倫理相關規定類似紐倫堡公約
 - 研究計畫的設計與執行需經獨立的委員會審查
 - 著名雜誌的編輯開始要求研究需依赫爾辛基宣言，結果的發表需有審查委員會的同意函
-

39

紐倫堡公約 (Nuremberg Code)

- 得到參與試驗者的自願同意是絕對必要的
 - 試驗的目的必須能為社會帶來福祉
 - 人體試驗必需有動物實驗的結果為依據
 - 試驗過程應避免不必要的身體或心智痛苦和傷害
 - 任何預知可能造成死亡或傷害的試驗絕不可進行
 - 受試者的風險必須低於試驗可能帶來的益處
 - 必須有合格的人員執行試驗
 - 不可造成不必要的身心傷害
 - 必須有妥善的準備與安全措施
 - 受試者可隨時要求退出試驗
-

40

大數據也有黑暗面

- ❑ 資訊蒐集對隱私造成威脅（類似歐威爾的小說「1984」）
 - ❑ 避免迷信因果關係而做出「預防性的懲罰」（例如：「關鍵報告」）
 - ❑ 類似「訴諸權威」的誤謬，因為迷戀資料分析、濫用、以致於形成「資料獨裁」
- 「考試成績並不等於真正的學識」，適度使用量化後的資訊。
-

41

使用大數據的建議

- 監管大數據的三個策略：
- 個資保護由「個人同意制」，轉成資料使用者的「使用責任制」；
 - 運用大數據進行預測時，需尊重個人的「能動性」；
 - 培養「演算學家」以處理大數據的審計師。
- 註：除了精算師，保險公司也需積極培養大數據的專家（IT、分析師等）
-

42

謝謝您的傾聽！

Q & A

全民健康保險資料庫在癌症資料分析之相關應用

1

林麗芬 教授

逢甲大學統計與精算研究所

「全民健康保險資料庫在醫療保險之應用」研討會

主辦單位:中華民國精算學會
政治大學風險管理與保險學系及統計學系

2013年8月29日

大綱

2

- 全民健康保險資料庫(NHIRD)介紹
- 癌症相關使用資料庫
- 資料庫處理程序
- 癌症相關應用研究
- 結論

全民健康保險資料庫(NHIRD)介紹

3

- 資料時間:1996-2012
- 費用檔：門診處方及治療明細檔 (CD)
住院醫療費用清單明細檔 (DD)
- 醫令檔：門診處方醫令明細檔 (OO)
住院醫療費用醫令清單明細檔 (DO)
- 基本資料檔：醫事人員基本資料檔 (PER)
專科醫師證書主檔 (DOC)
- 醫院相關：醫事機構基本資料檔 (HOSB)
醫事機構副檔資料檔 (HOSX)
醫事機構診療科別明細檔 (DETA)
醫事機構病床主檔 (BED)
門診費用申請總表主檔 (CT)
住院費用申請總表主檔 (DT)
- 病人相關：承保資料檔 (ID)
重大傷病證明明細檔(HV)

癌症相關使用資料庫

4

- ▶ 承保資料檔(ID)
- ▶ 重大傷病證明明細檔(HV)
- ▶ 癌症門診處方及治療明細檔 (CN_CD)
- ▶ 癌症住院醫療費用清單明細檔(CN_DD)

癌症相關使用資料庫

主題	內容	中文檔名	英文檔名	年度	資料筆數	檔案大小	總資料筆數	總檔案大小(MB)
癌症	CN	癌症住院醫療費用清單明細檔	5 CN_DD	1996	216863	101.96	1006316	473.14
				1997	235358	110.66		
				1998	265544	124.85		
				1999	288551	135.67		
癌症	CN	癌症門診處方及治療明細檔	CN_CD	1996	3309411	946.82	100962270	28946.32
				1997	3632661	1039.31		
				1998	4153252	1188.27		
				1999	4688702	1341.47		
				2000	4806448	1375.15		
				2001	5318780	1521.72		
				2002	5706850	1632.74		
				2003	5514536	1577.71		
				2004	6335874	1818.75		
				2005	6588431	1891.15		
				2006	6952396	1995.72		
				2007	7547200	2166.45		
				2008	8076733	2318.48		
				2009	8817105	2531.01		
				2010	9436512	2708.8		
				2011	10077379	2892.77		
重大傷病醫療資源使用	HV	重大傷病醫療使用資源住院醫療費用清單明細檔	HV_DD	1996	280892	132.33	1336730	559.62
				1997	301043	141.54		
				1998	375583	107.46		
				1999	379212	178.29		
重大傷病醫療資源使用	HV	重大傷病醫療使用資源門診處方及治療明細檔	HV_CD	1996	3071993	878.9	112040178	32124.68
				1997	3478998	995.35		
				1998	4064823	1162.96		
				1999	4518433	1292.73		
				2000	4937040	1412.49		
				2001	5514320	1577.66		
				2002	6284145	1797.9		
				2003	7004731	2004.05		
				2004	8221393	2360		
				2005	8870528	2546.33		
				2006	9013807	2587.47		
				2007	8494779	2438.48		
				2008	8981612	2578.21		
				2009	9416362	2703.02		
				2010	9856982	2829.5		
				2011	10310232	2959.63		
承保檔	ID	承保資料檔	ID	2002	71657657	7312.15	310233675	31657.17
				2003	25130381	2564.38		
				2004	25003589	2551.45		
				2005	25942193	2647.21		
				2006	26336288	2687.44		
				2007	26619011	2716.28		
				2008	26970229	2752.12		
				2009	27223012	2777.93		
				2010	27509909	2807.19		
				2011	27841406	2841.02		

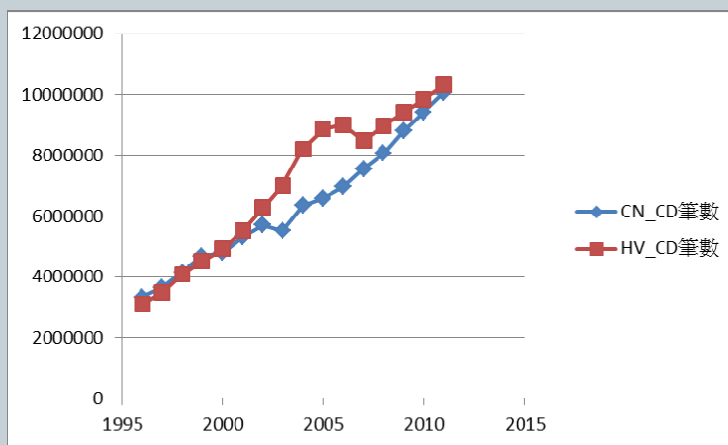
全民健康保險資料庫在癌症資料分析之相關應用 林麗芬

2013/8/29

各年度CN_CD HV_CD門診筆數統計

6

年度	CN_CD筆數	HV_CD筆數
1996	3309411	3071993
1997	3632661	3478998
1998	4153252	4064823
1999	4688702	4518433
2000	4806448	4937040
2001	5318780	5514320
2002	5706850	6284145
2003	5514536	7004731
2004	6335874	8221393
2005	6588431	8870528
2006	6952396	9013807
2007	7547200	8494779
2008	8076733	8981612
2009	8817105	9416362
2010	9436512	9856982
2011	10077379	10310232



NHIRD Studies

7

- 題目方向
 - 研究設計
 - 變數選取
 - 數據分析
-
- Data processing takes 60-75% of the task.

NHIRD研究分析軟體工具

8

- 資料庫管理系統:
ORACLE; My SQL; SQL Server (Microsoft)
- 統計套裝軟體
R; SAS; SPSS; STATA...

SQL 語法

9

- SQL (Structured Query Language)
- 1970 由 IBM 發表，用於大型資料庫中資料的定義/操作/查詢/控制
- 適合資料串連，幾乎遍及所有資料庫系統 (MS SQL Server, Oracle, MySQL...等等)
- SAS 6.0 以後加入專門資料操作模組 (Proc SQL)

資料處理程序

10

- **資料轉換 (READ)**

將資料庫轉換成適合分析的形式，必須搭配資料清潔(Data Cleaning)，將資料有遺失或錯誤的，整理後再匯入。

- **轉換資料庫→資料集 (SELECT FILTER UNION JOIN SORT)**

將不同資料庫檔案進行串連

經驗:極少數 (約1~2%) 就診資料無法正確串連，有問題的樣本並不多，應不會對整體分析造成影響

- **結果分析 (AGGREGATE OUTPUT STATISTICS)**

癌症相關應用研究

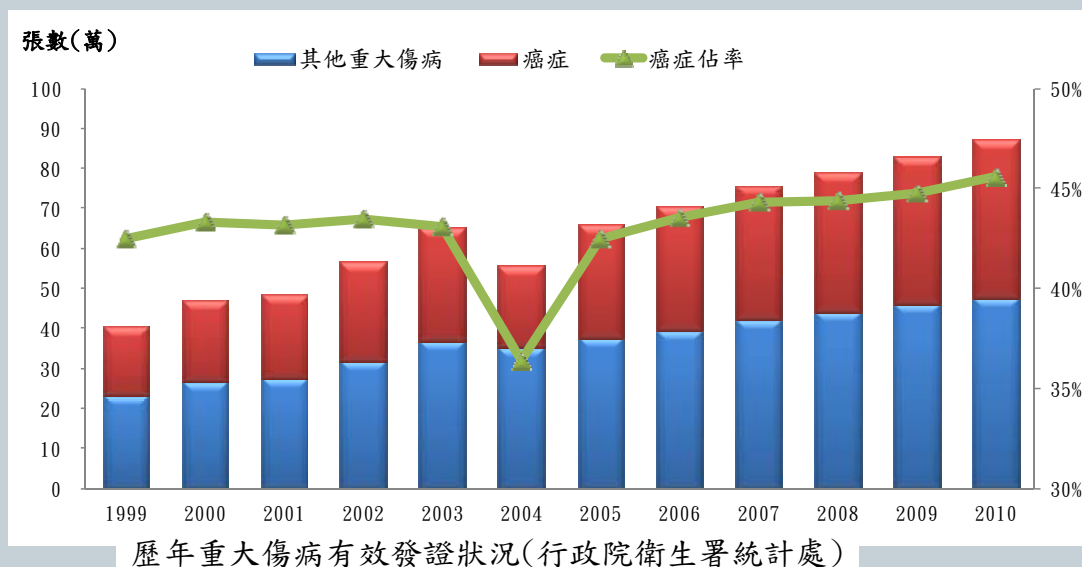
11

- 台灣癌症發生率之研究(林麗芬、游智涵)
- 女性乳癌緩解與復發資料統計分析(林麗芬、李宗道)
- 台灣癌症住院費用與住院天數之長期追蹤資料分析(林麗芬、何韻贊)
- 多重狀態模型在台灣癌症機率之推估(林麗芬、蔣忠儒)
- 台灣女性乳癌之死亡率研究(林麗芬、黃瑞笙)
- 台灣女性乳癌首發之疾病壓縮或疾病擴張研究(林麗芬、鄧柏文)

1.台灣癌症發生率之研究

12

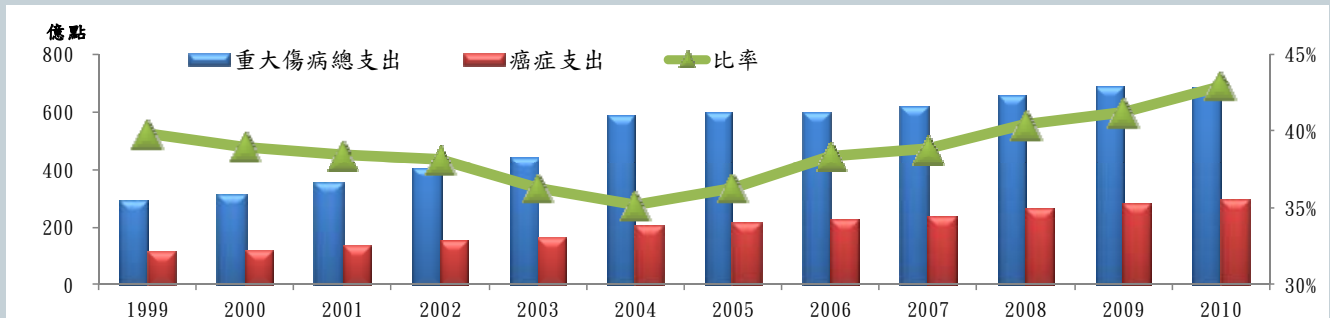
- 癌症蟬連三十年為死亡原因之首
- 重大傷病證明領卡人數持續上升



台灣癌症發生率之研究

13

● 健保對於癌症醫療費用的支出不斷提高



重大傷病住院醫療費用申報狀況



重大傷病門診醫療費用申報狀況

全民健康保險資料庫在癌症資料分析之相關應用 林麗芬

2013/8/29

文獻探討

14

● 喬治華等人(2010):保發中心

研究1996年至2007年各年度首次罹癌率的統計及趨勢。

● 癌症登記資料庫(1997)

衛生署針對50床以上的醫院，所建立的癌症登記系統。

● 陳寬政(2009)

須以考量治癒率的模型才能準確估算癌症盛行率。

● 林麗芬、游智涵(2010)

1997-2008年癌症發生率比癌症登記報告統計，約高7%左右。

資料來源

15

全民健保資料庫

- 重大傷病證明明細檔
- 重大傷病醫療使用資源門診處方及治療明細檔
- 重大傷病醫療使用資源住院醫療費用清單明細檔
- 承保檔

研究對象：
重大傷病類別第一類-
需積極或長期治療的癌症病患

資料庫

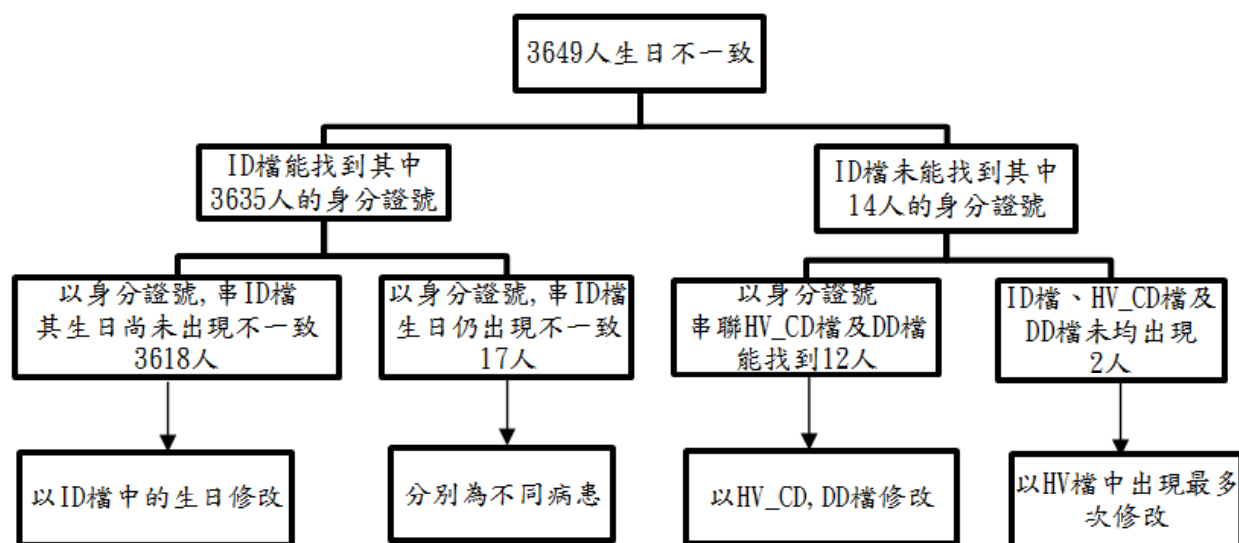
16

- 重大傷病證明明細檔(HV檔)
主要使用的欄位為身分證號、診斷代碼、重大傷病代碼、出生日期、性別、申請日期、死亡註記、死亡日期
- 重大傷病醫療使用資源門診處方及治療明細檔 (HV_CD檔)
主要使用欄位為身分證號、出生日期、性別、就醫日期，並使用病患的身分證字號串連
- 住院醫療費用清單明細檔 (HV_DD檔)
從2000年開始不再提供重大傷病醫療使用資源[住院醫療費用清單明細檔](#)，將使用住院醫療費用清單明細檔，使用欄位為身分證號、出生日期、性別、入院日、出院日、轉歸代碼，用身分證字號串聯此檔案觀察病患的住院期間以及出院的原因。
- 承保檔(ID檔)
檔案內容包含在2010年底前加保之在保及停保記錄，而退保者，僅提供最後一筆異動記錄。使用欄位為身分證號、出生日期、性別、退保日期、退保別，用身分證字號串聯此檔案觀察病患的退保日期及退保原因。

研究對象:94萬名癌症病患

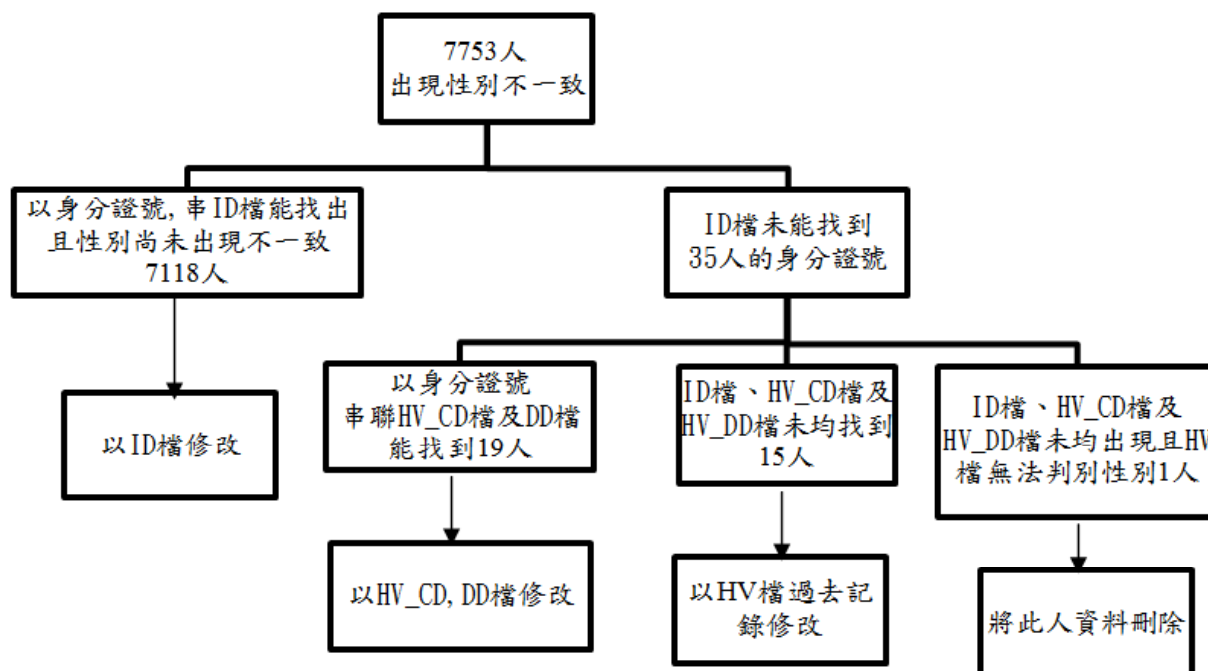
17

生日欄位修改流程



性別欄位修改流程

18



資料處理-死亡日期

19

最後一次出院原因為死亡、病危、自殺的出院日期



最後一次異動記錄為退保的退保日期



重大傷病檔中的死亡日期

資料合併流程

20

癌症病患定義：挑選出在 HV 檔中
重大傷病類別代碼為 01 的病患資料

身分證字號



定義：病患持續一年未就醫為康復

合併成治療期
將記錄日期間格小於一年的日期合併

定義：治療期小於一周為回診

刪除治療期小於一周



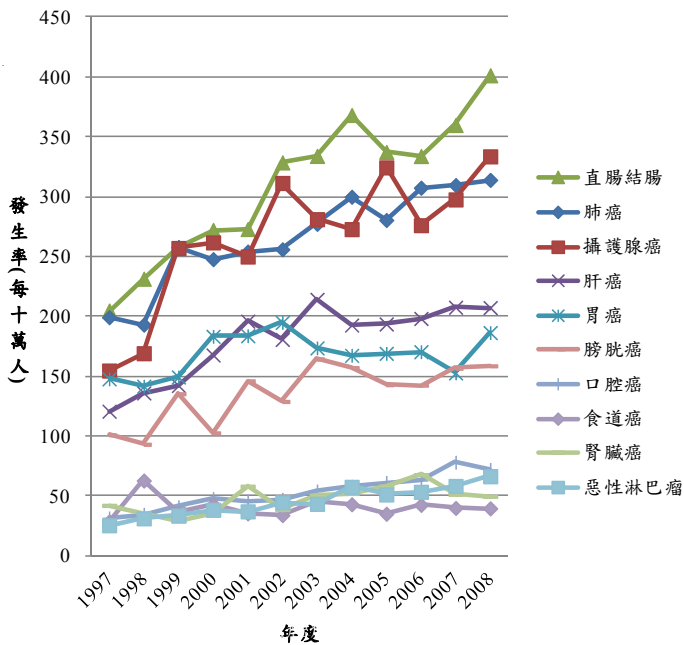
計算癌症發生率

研究結果

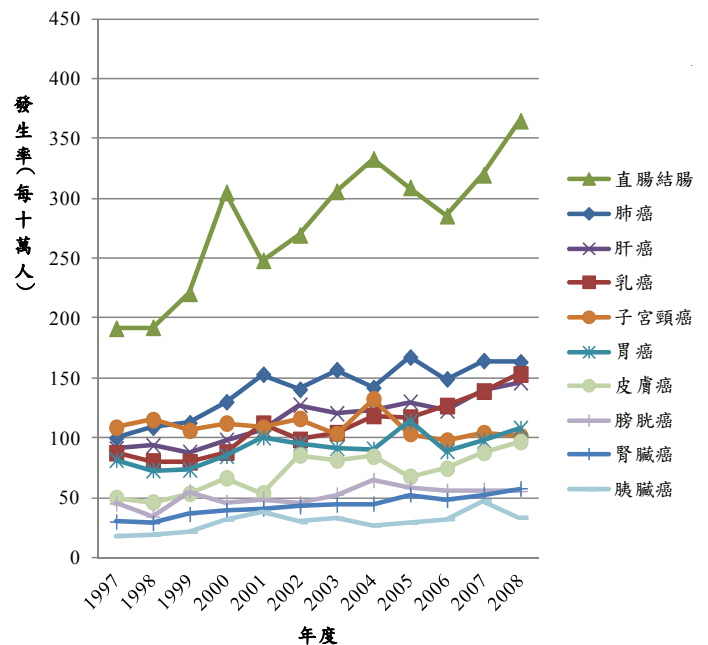
21

● 老年(65歲以上)

男性-老年



女性-老年

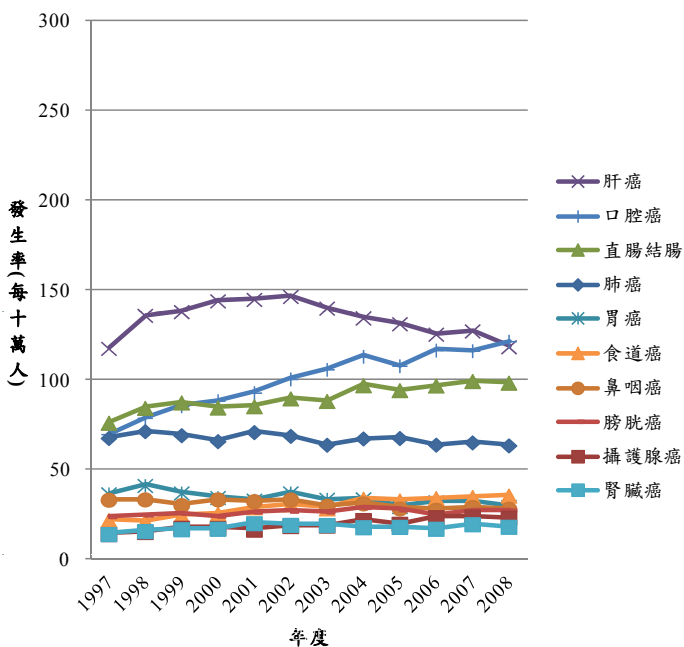


研究結果

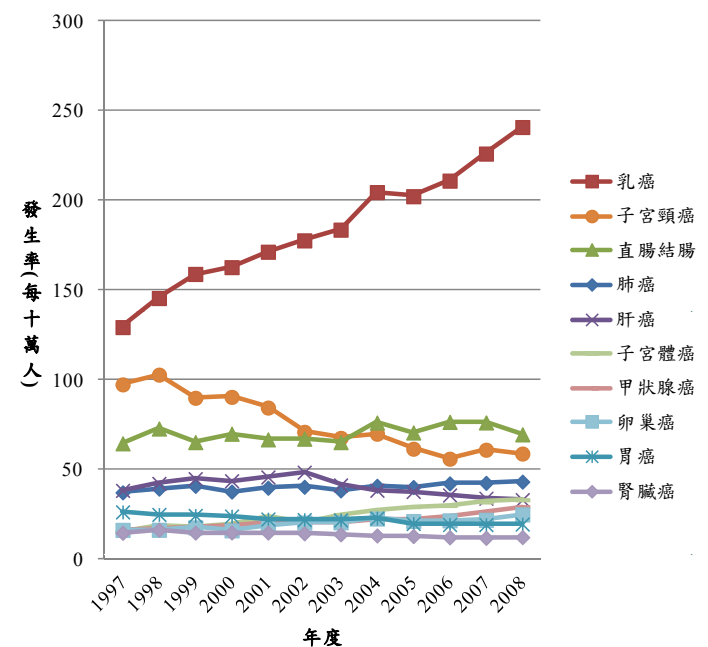
22

● 中年(45-64歲)

男性-中年



女性-中年

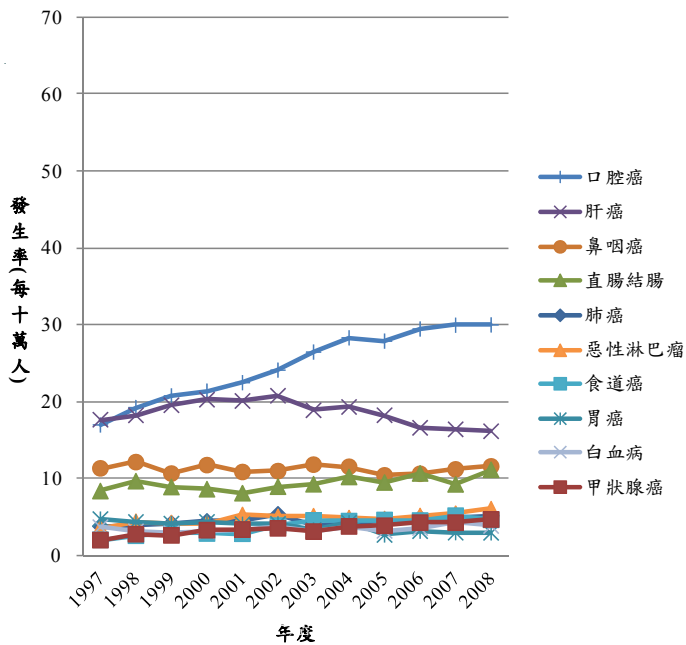


研究結果

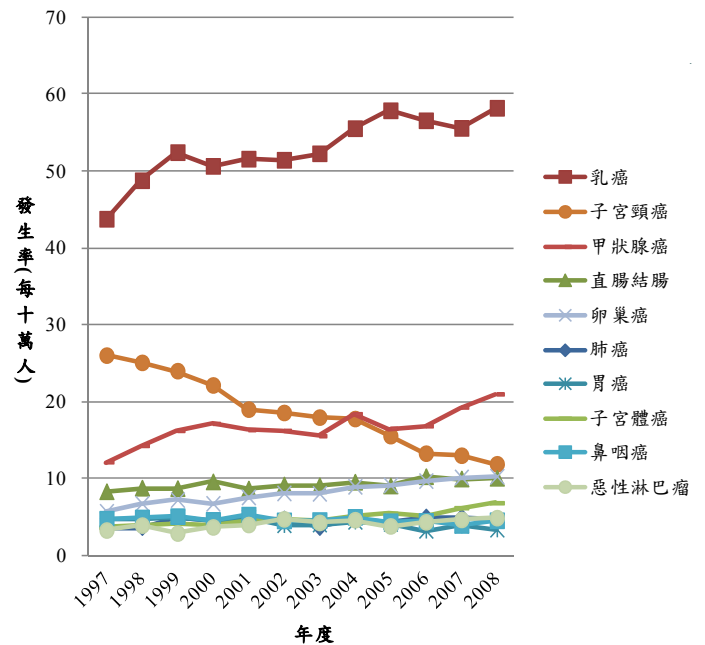
23

● 壯年(25-44歲)

男性-壯年



女性-壯年

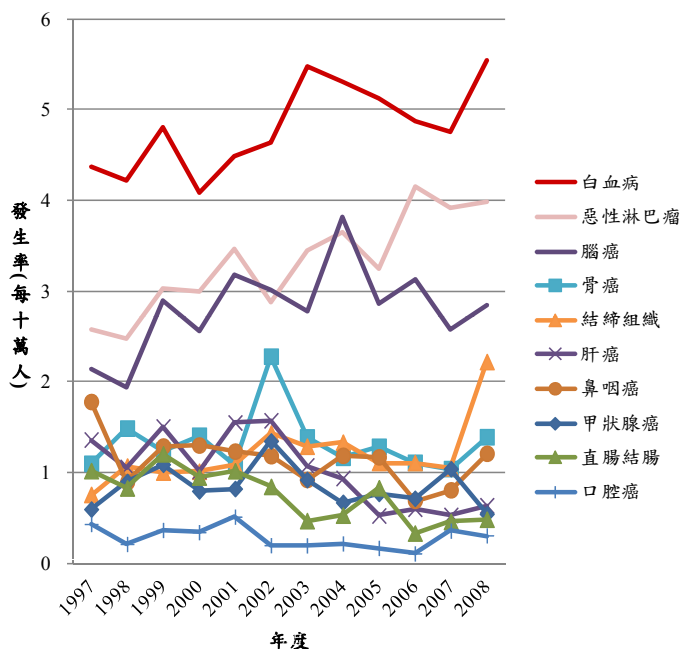


研究結果

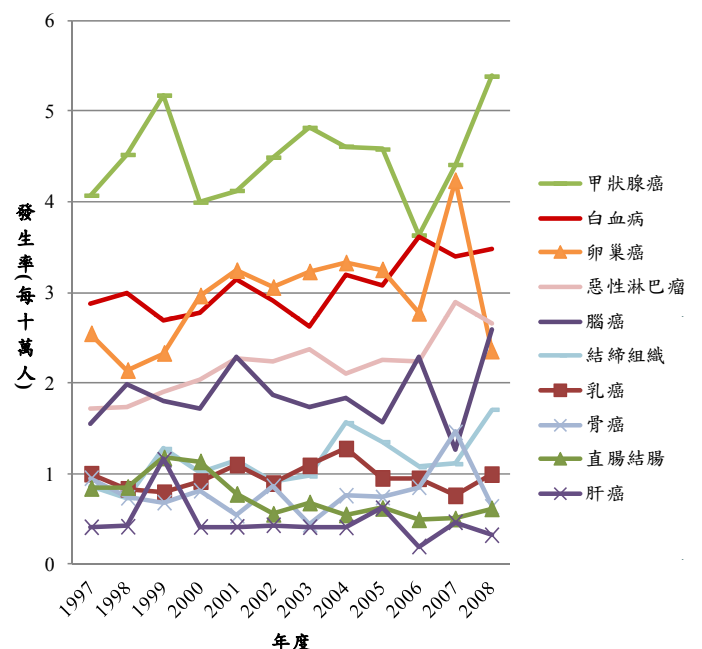
24

● 青少年(15-24歲)

男性-青年



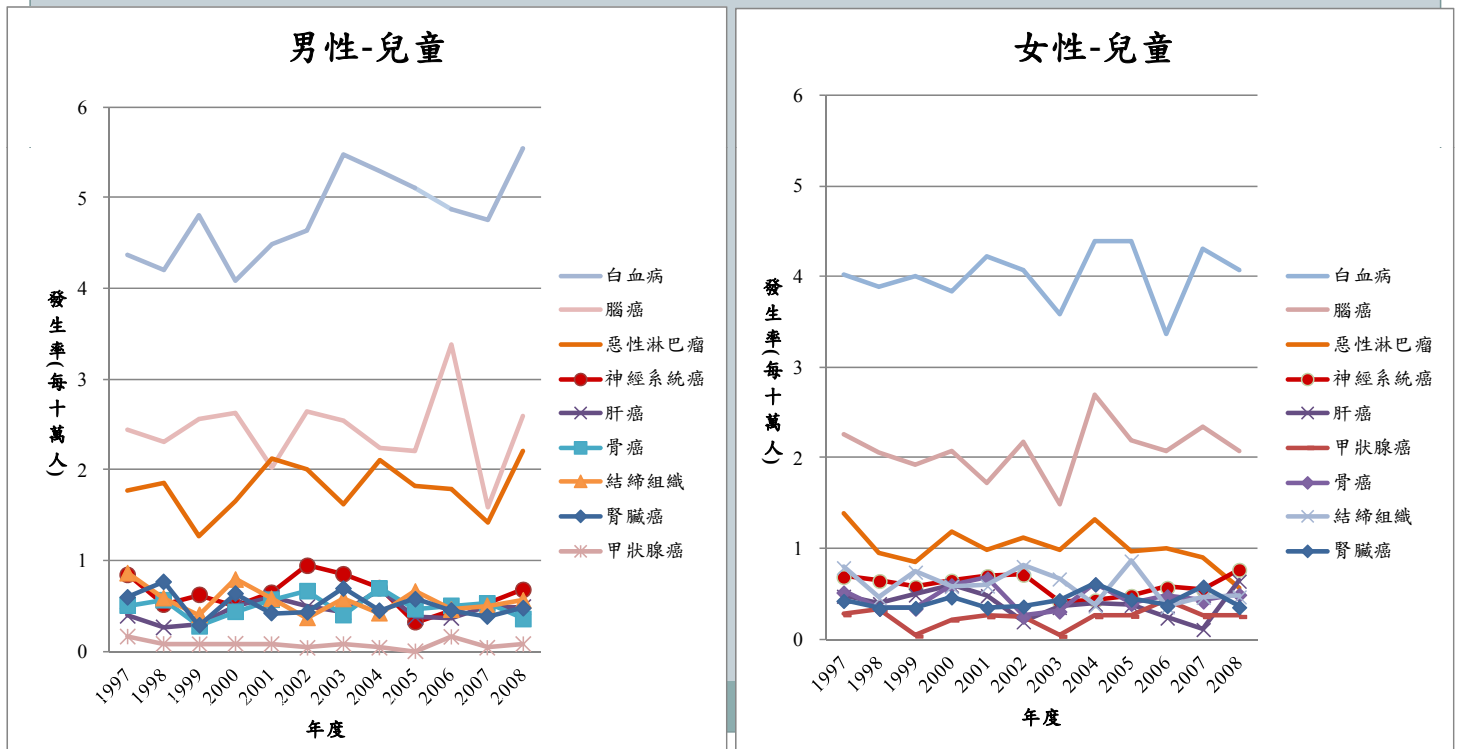
女性-青年



研究結果

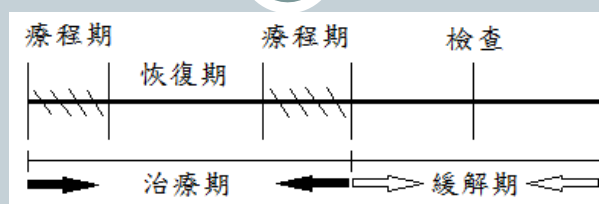
25

● 兒童(15歲以下)



2. 女性乳癌緩解與復發資料統計分析

26



資料來源：

全民健康保險資料庫「癌症門診處方及治療明細檔 (CN_CD)」。

資料範圍：

1996/1/1至2009/12/31。

資料數量：

總資料為10萬3千名乳癌病患、580萬筆病患的門診記錄。

固定八年觀察期間後，擷取資料為4萬1千名病患、220萬筆門診紀錄。

資料與篩選條件

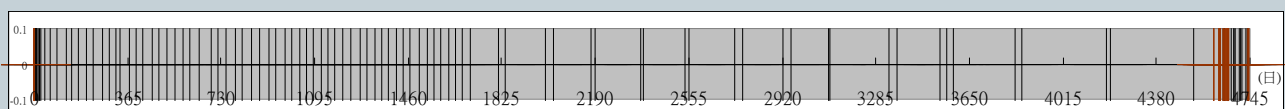
27

- 癌症門診檔擷取條件
 - 特定治療項目代號：12、D1、D2
 - 國際疾病分類號：140 ~ 239；A08 ~ A17
 - 主手術代碼：V57、V58
- 本研究再次篩選條件
 - 國際疾病分類號：174、A113
 - 部分負擔代號：001

條碼圖

28

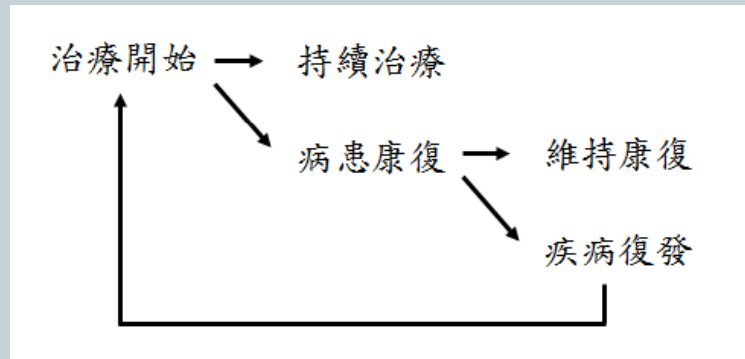
- 「條碼圖」所包含的資訊：
 - 門診時間
 - 住院期間
 - 死亡時點



註：咖啡色的部分為住院時間記錄；黑色的部分為門診治療記錄。
橫軸單位為（日）

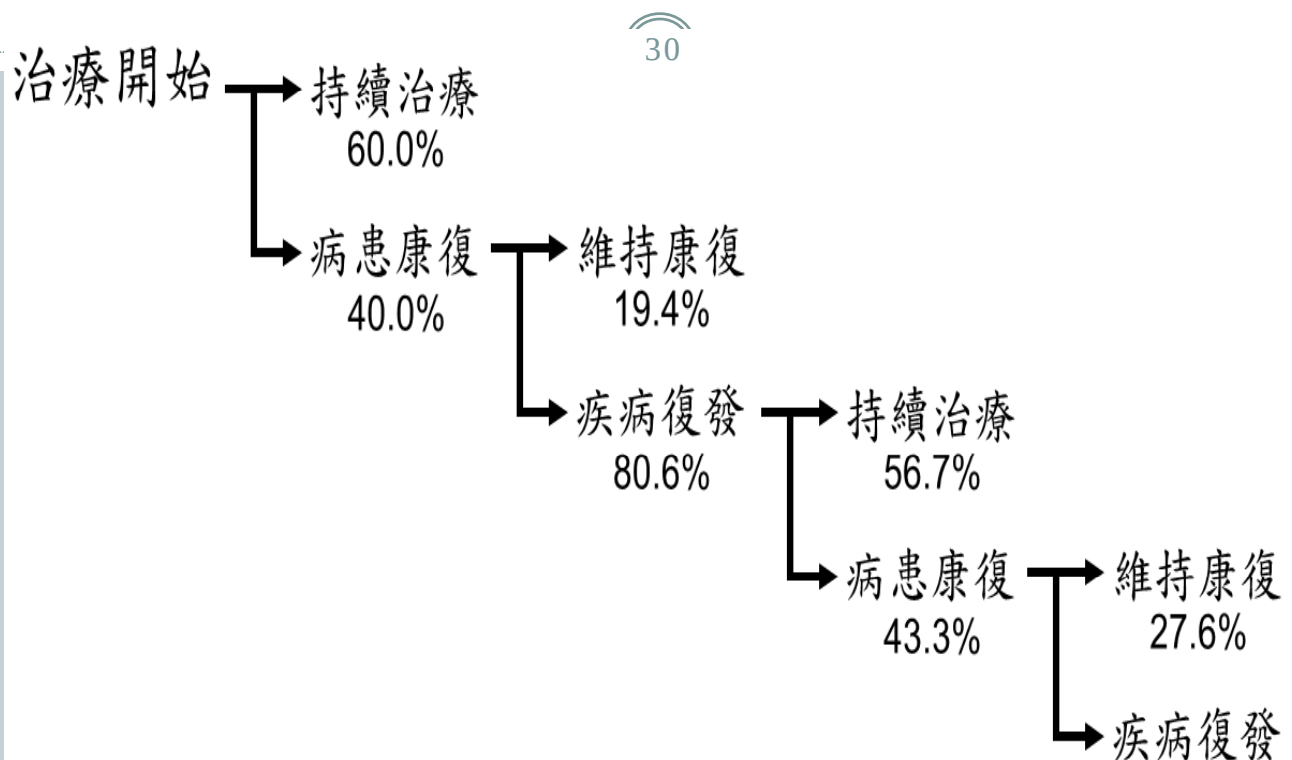
疾病發展狀態

29



疾病發展狀態圖

30



康復率與復發率統計

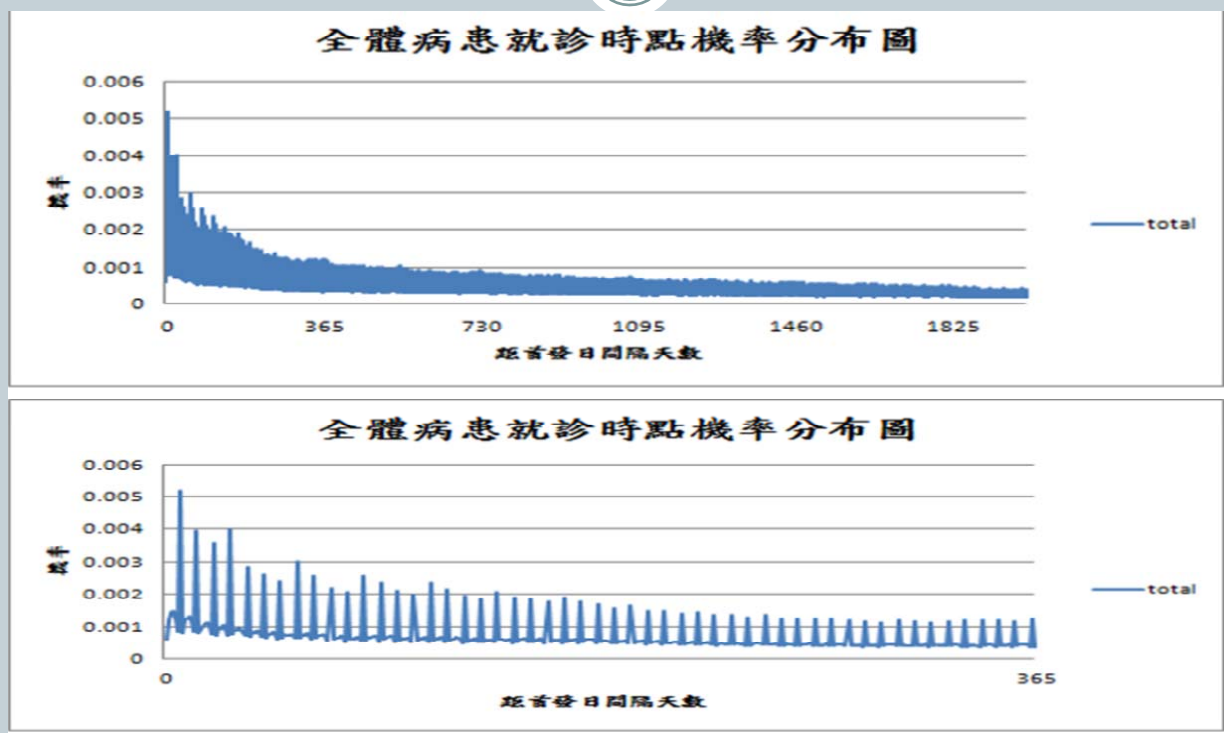
31

● 第一次病況發展循環年齡分層機率分布表

	0~39歲	40~49歲	50~59歲	60~69歲	70歲以上
第一次循環	100% (6879人)	100% (14224人)	100% (9835人)	100% (6514人)	100% (3884人)
→ 持續治療	57.6% (3962人)	55.7% (7921人)	60.3% (5931人)	63.9% (4165人)	70.4% (2733人)
治療滿八年	32.7% (2247人)	35.0% (4974人)	32.6% (3207人)	31.3% (2039人)	24.0% (932人)
治療至死亡	24.9% (1715人)	20.7% (2947人)	27.7% (2724人)	32.6% (2126人)	46.4% (1801人)
→ 病患康復	42.4% (2917人)	44.3% (6303人)	39.7% (3904人)	36.1% (1866人)	29.6% (1151人)
→ 維持康復	8.0% (553人)	8.2% (1172人)	7.4% (731人)	7.4% (483人)	7.1% (277人)
康復滿八年	7.8% (536人)	8.0% (1136人)	7.1% (699人)	6.4% (419人)	5.4% (209人)
康復後死亡	0.2% (17人)	0.3% (36人)	0.3% (32人)	1.0% (64人)	1.8% (68人)
→ 疾病復發	34.4% (2364人)	36.1% (5131人)	32.3% (3173人)	28.6% (1866人)	22.5% (874人)

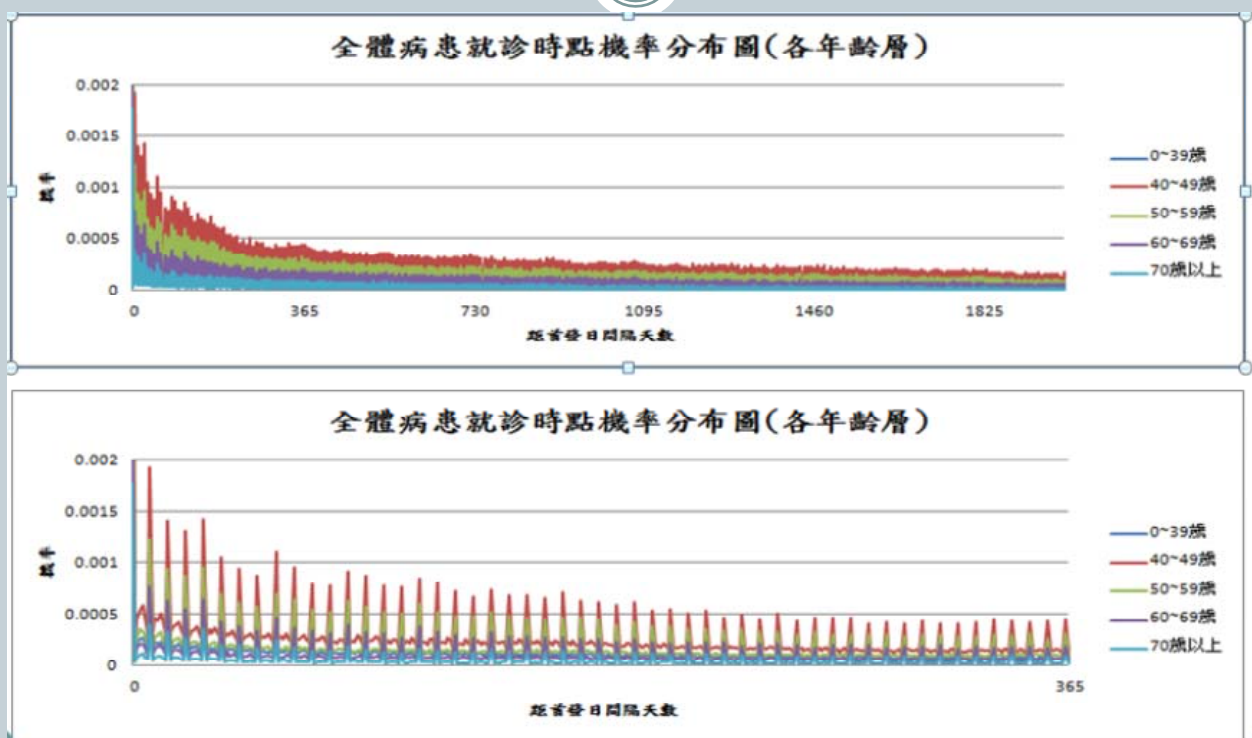
女性乳癌門診就診機率分布

32



女性乳癌各年齡層門診機率分布

33



3.台灣癌症住院費用與住院天數之長期追蹤資料分析

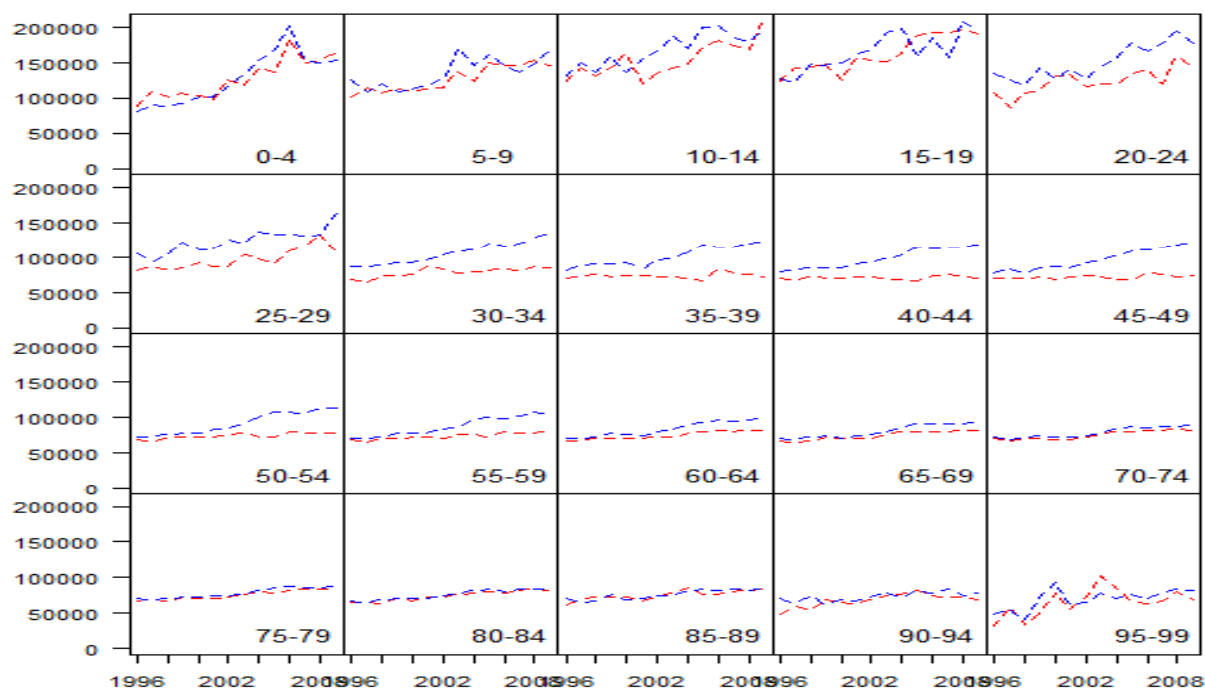
34

- 衛生署2010年統計：
820,505人領取重大傷病卡，佔總投保人口3.56%
醫療費用142.6億，佔總醫療費用27.13%
癌症：醫療費用51.6億，佔重大傷病的36.14%
- 資料庫來源:癌症住院醫療費用清單明細檔(CN_DD)

男性、女性病患各年齡層健保給付的總體平均每次住院費用

藍色虛線: 男性 紅色虛線: 女性

35



健保給付

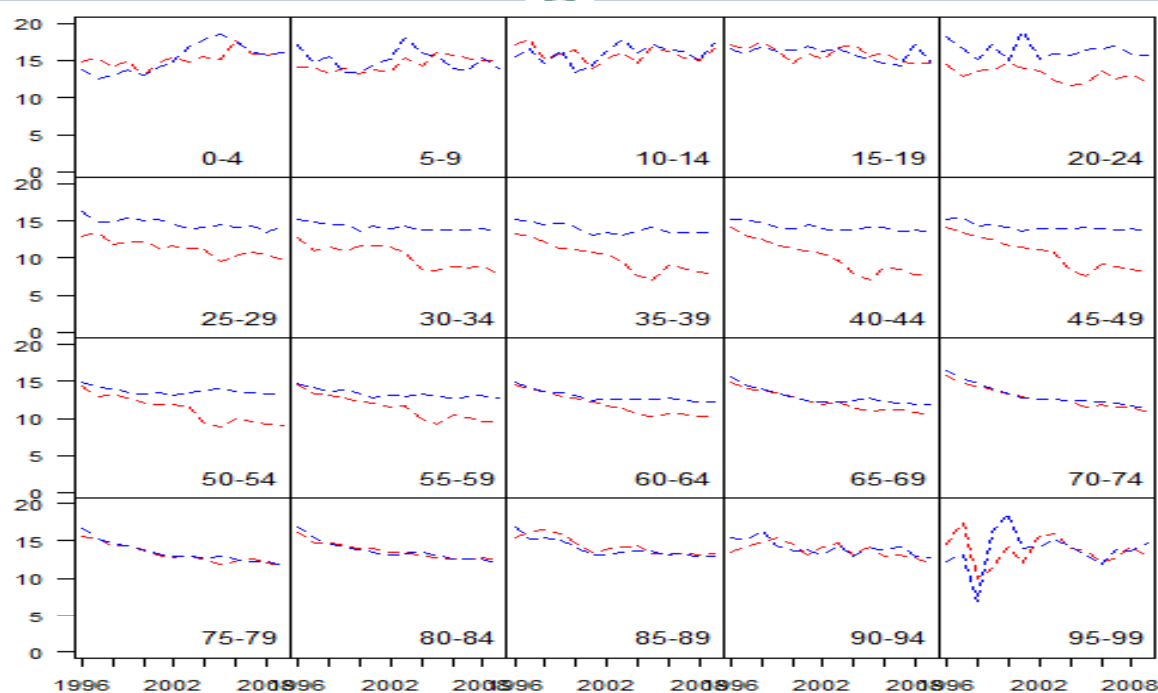
36

- 29歲以下的健保給付平均住院費用偏高，且以10~19歲之間最高
- 0~29歲之平均健保給付高於30~99歲(高出\$21,564)，且隨年齡增加而遞減
- 0到29歲的健保給付平均住院費用成長迅速(男性大約從1996~2009年的\$120,000 → \$170,000，而女性大約從\$100,000 → \$160,000)
- 女性30歲以後就開始平穩(約七萬元),男性則到了60歲以後才開始平穩(約七萬五千元)
- 30至59歲，男女性之間的平均住院費用差別最明顯
- 健保給付平均住院費用:逐年遞增
- 平均住院費用金:男性 > 女性(大約平均高出\$5000)

男、女性病患各年齡層總體平均每次住院天數

藍色虛線: 男性 紅色虛線: 女性

37



住院天數

38

- 住院平均天數:逐年下降
- 住院平均天數:男性 > 女性

十大癌症健保給付和住院天數比較

39

男性癌症病患

1. 基準點癌症攝護腺癌(D185)的健保給付，平均每次住院費用是\$46,027，平均每次住院天數為8.4天
2. 食道癌(D150)最高(健保給付:比攝護腺癌多\$42,852，住院天數:比攝護腺癌多9天)
3. 膀胱癌(D188)與基準點癌症攝護腺癌(D185)沒有顯著的差別
4. 健保給付最低的是肝癌(D155)，攝護腺癌(D185)排倒數第二
5. 多數癌症的每次住院費用逐年增加\$2000~3000

男性病患平均每次住院費用模型結果

40

健保給付的男性病患平均每次住院金額模型結果

代碼	癌症名稱
D140	口腔癌(包含唇、舌、齒齦、口底、口之其他及未明示部位、口咽、下咽、唇口腔及咽之其他及分界不明位置)
D147	鼻咽癌
D150	食道癌
D151	胃癌
D153	結腸癌
D154	直腸癌
D155	肝癌
D162	氣管、支氣管及肺癌
D174	乳癌
D179	子宮癌
D180	子宮頸癌
D183	卵巢癌
D185	攝護腺癌
D188	膀胱癌
D193	甲狀腺癌

參數	變數	估計值	標準誤	t-value	P-value
β_0	intercept	46,027.21	5,907.406	7.791441	0.0000
β_1	t	1,725.04	621.092	2.777436	0.0055
β_2	D140	39,299.71	6,915.650	5.682721	0.0000
β_3	D147	23,148.83	3,479.617	6.652695	0.0000
β_4	D150	42,851.99	7,069.330	6.061676	0.0000
β_5	D151	23,053.61	6,915.650	3.333542	0.0011
β_6	D153	19,087.55	6,915.650	2.760052	0.0066
β_7	D154	21,994.62	6,915.650	3.180413	0.0018
β_8	D155	-4,806.49	6,915.650	-0.695016	0.4882
β_9	D162	5,632.98	6,915.650	0.814526	0.4167
β_{10}	AGE	118.34	81.82	1.446324	0.1503
β_{11}	t * D140	3,098.66	731.444	4.236357	0.0000
β_{12}	t * D150	3,113.49	754.440	4.126890	0.0000
β_{13}	t * D151	2,168.79	731.444	2.965072	0.0031
β_{14}	t * D153	2,788.22	731.444	3.811939	0.0001
β_{15}	t * D154	2,470.23	731.444	3.377188	0.0007
β_{16}	t * D155	1,897.00	731.444	2.593498	0.0096
β_{17}	t * D162	2,602.33	731.444	3.557799	0.0004
β_{18}	t * AGE	-28.44	8.798	-3.232650	0.0012

十大癌症健保給付和住院天數比較

41

女性癌症病患

- 1.基準點癌症乳癌(D174)的健保給付，平均每次住院費用是\$53,203，平均每次住院天數為9.8天
- 2.胃癌(D151)每次住院費用比乳癌多\$22,060；平均每次住院天數比乳癌多5.1天，是女性十大癌症中最高的
- 3.胃癌(D151)直腸癌(D154)和結腸癌(D153)為前三高
- 4.多數癌症的每次住院費用逐年增加\$2200~3300

女性病患平均每次住院費用模型結果

42

健保給付的女性病患平均每次住院金額模型結果

代碼	癌症名稱
D140	口腔癌(包含唇、舌、齒齦、口底、口之其他及未明示部位、口咽、下咽、唇口腔及咽之其他及分界不明位置)
D147	鼻咽癌
D150	食道癌
D151	胃癌
D153	結腸癌
D154	直腸癌
D155	肝癌
D162	氣管、支氣管及肺癌
D174	乳癌
D179	子宮癌
D180	子宮頸癌
D183	卵巢癌
D185	攝護腺癌
D188	膀胱癌
D193	甲狀腺癌

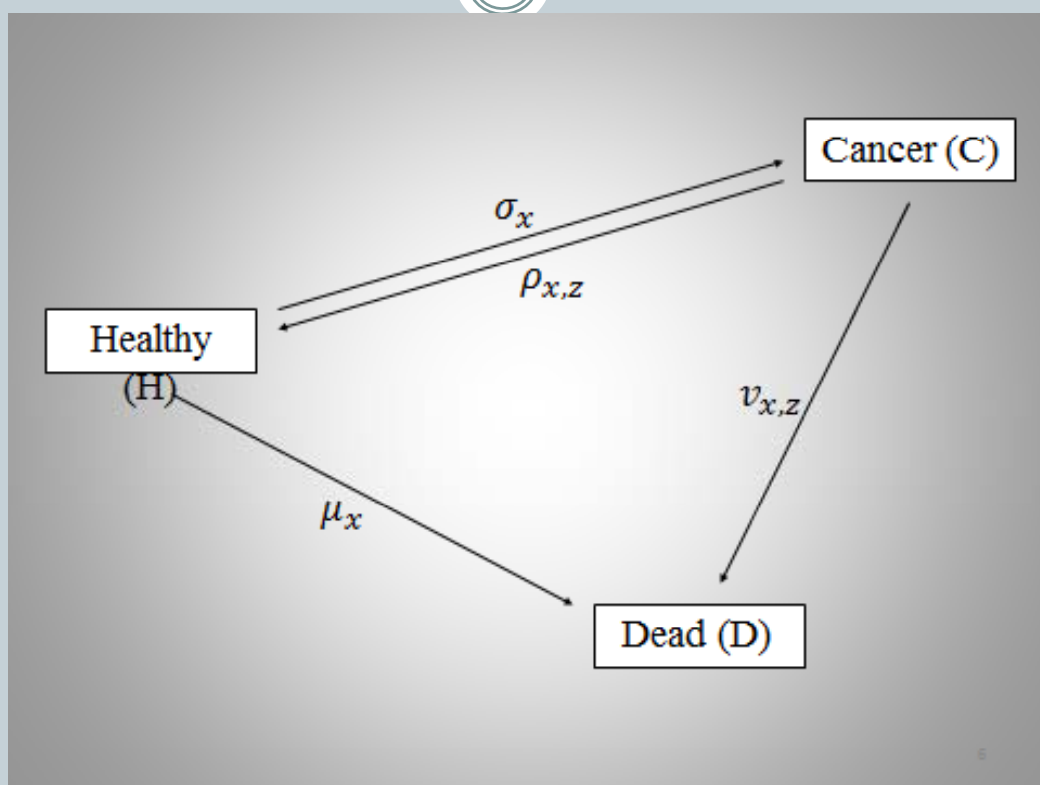
參數	變數	估計值	標準誤	t-value	P-value
β_0	intercept	53202.62	4091.117	13.004422	0.0000
β_1	t	-323.42	431.493	-0.749545	0.4536
β_2	D151	22059.54	5785.714	3.812760	0.0002
β_3	D153	18367.67	5785.714	3.174659	0.0018
β_4	D154	21062.86	5785.714	3.640495	0.0004
β_5	D155	-2791.80	5785.714	-0.482533	0.6302
β_6	D162	619.36	5785.714	0.107050	0.9149
β_7	D179	8896.13	5998.105	1.483157	0.1403
β_8	D180	6301.01	5785.863	1.089035	0.2780
β_9	D183	7867.37	5806.802	1.354853	0.1776
β_{10}	D193	-5734.72	5786.502	-0.991051	0.3234
β_{11}	t * D151	2334.22	610.223	3.825200	0.0001
β_{12}	t * D153	2914.70	610.223	4.776444	0.0000
β_{13}	t * D154	2592.71	610.223	4.248788	0.0000
β_{14}	t * D155	2156.84	610.223	3.534520	0.0004
β_{15}	t * D162	2718.50	610.223	4.454933	0.0000
β_{16}	t * D179	2208.05	641.164	3.443812	0.0006
β_{17}	t * D180	2107.99	610.338	3.453808	0.0006
β_{18}	t * D183	3313.35	612.160	5.411917	0.0000
β_{19}	t * D193	1203.56	610.16	1.972264	0.0487

4.多重狀態模型在台灣癌症機率之推估

43

- Three states in the classification
 - (i) Healthy(H) (ii) Cancer(C) (iii) Dead(D)
 - Notations:
 - (i) σ_x : The transition intensity from the healthy states to the cancer states
 - (ii) μ_x : The transition intensity from the healthy states to the dead states
 - (iii) $\rho_{x,z}$: The transition intensity from cancer states to healthy states
 - (iv) $v_{x,z}$: The transition intensity from cancer states to dead states
- x : age
 z : sickness duration

44



Probabilities

45

We use some notation to denote the probability of the transition between each states as follows:

$$(i) {}_t p_{x,z}^{j,k} = P[Y(x+t) = k | Y(x) = j \text{ and } Z(x) = z]$$

where $j, k = H, C, \text{or } D$ and $t, x, z \geq 0$

$$(ii) {}_t p_x^{\overline{HH}} = P[Y(x+t) = H \text{ and } Z(x+t) \geq t | Y(x) = H]$$

$$(iii) {}_t p_{x,z}^{\overline{CC}} = P[Y(x+t) = C \text{ and } Z(x+t) = z+t | Y(x) = C \text{ and } Z(x) = z]$$

$$(iv) {}_{w,t} p_x^{\overline{HC}} = P[Y(x+t) = C \text{ and } Z(x+t) \leq w | Y(x) = H]$$

Data Collation

46

- The NHIRD: ID, CN_CD and CN_DD
- Selected variables:
id, birthday, gender, date of withdrawal,
out-patient date, the date of admission,
the date of discharge, and diagnosis

Data Processing

47

Correcting the gender error for gender-specific cancers

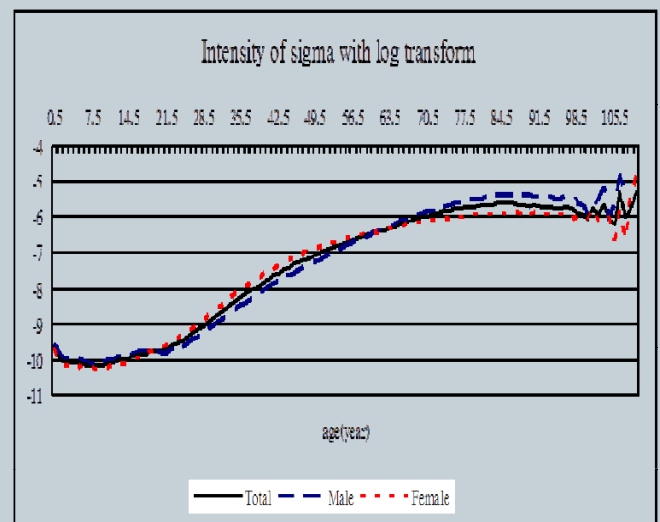
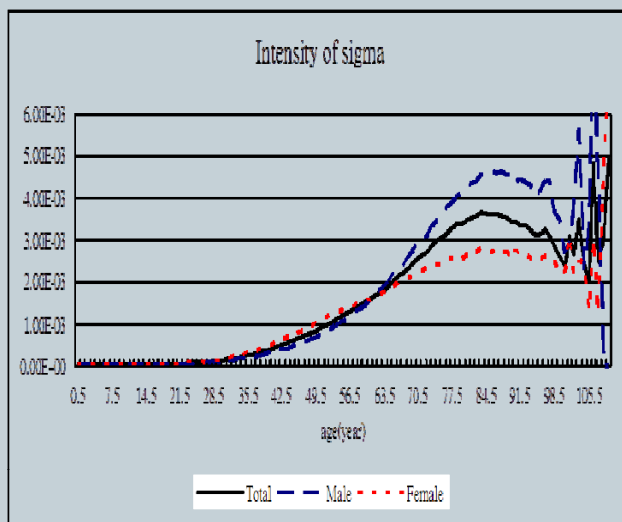
In the cases where multiply data sources showed differing data, we chose the data with the most number of occurrences as the correct data

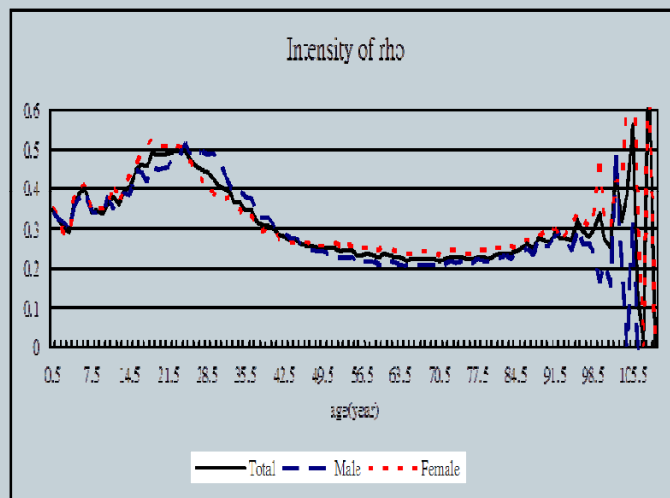
Deleted all the data where the date is still error

Deleted the data in CD and DD where the id is not in ID

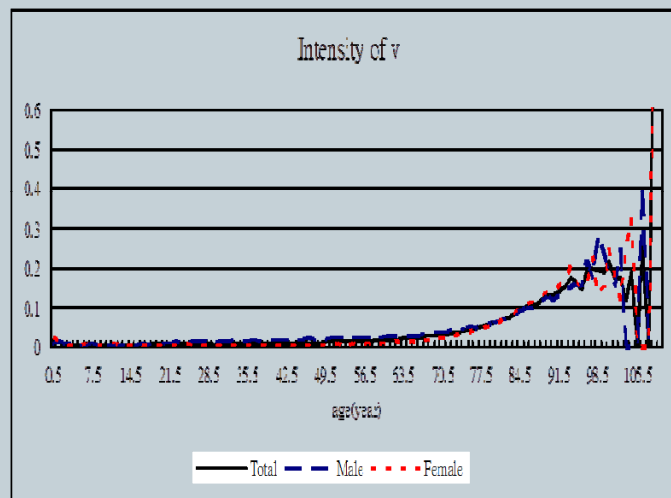
The transition intensity from healthy state to cancer state

48

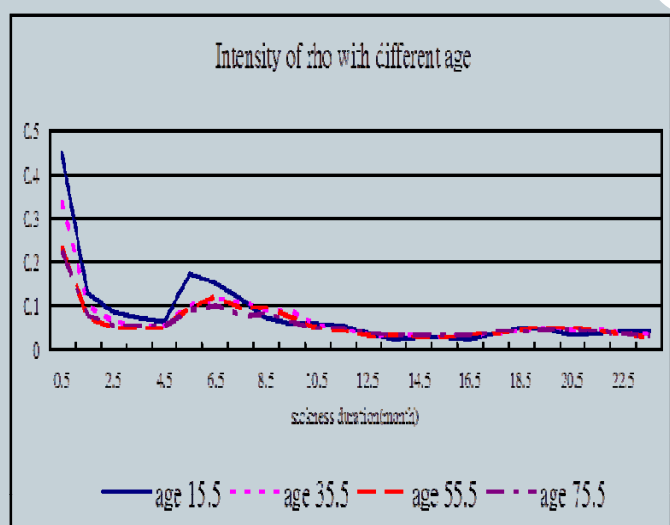




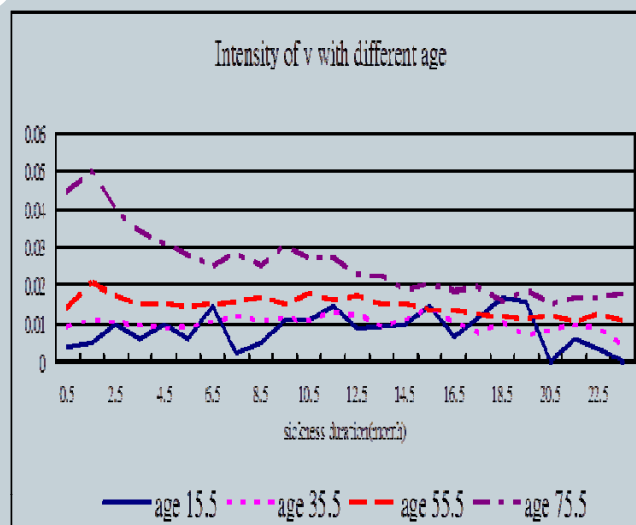
The transition intensity from cancer state to healthy state



The transition intensity from cancer state to dead state



The transition intensity from cancer state to healthy state with different ages



The transition intensity from cancer state to dead state with different ages

NIHIRD使用上的限制

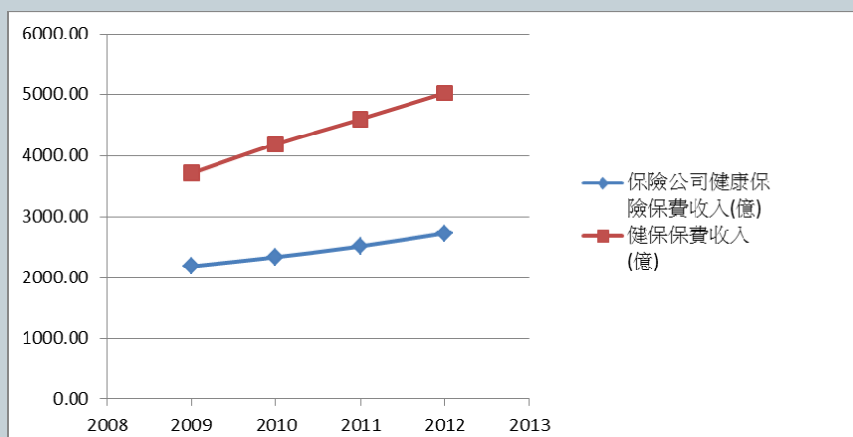
51

- 資料缺失問題:部分病患資料無法取得，例如未公開發行或申請困難資料（如癌症登記檔、出生登記檔、和**死亡登記檔**），抑或是針對某單一議題所作單次問卷
- NIHIRD基本資料檔僅紀錄健保相關門診和住院診療服務總量，並未涵蓋所有細項服務量，也未包括非健保給付服務資料(如手術中使用非健保給付之相關器材、非健保給付的藥品費用)
- 健保資料是以個人為樣本單位，所以不包括家戶組成的相關資料
- 健保資料僅有健保就診診斷資料，不包括投保人自覺健康資訊，亦無家族成員疾病史
- 健保資料對病患特性僅紀錄病患性別及生日，並不包括重要社經（如教育或所得）

建議

52

- 使用健保資料庫的進入障礙相當高，研究者通常需要花費許多時間和心力去試驗和熟悉
- 商業健康保險保費收入與全民健保保費收入比較
- 商業健康保險的後續發展...



**Thanks for
Your Attentions!**



**Can you Imagine
Cancer can be cured
No More Cancer in the World.**

Q & A

從全民健保資料庫分析住院發生率風險因子(草案)

「全民健保資料庫在醫療保險之應用」研討會
中華民國精算學會

真理大學統計與精算系
王信忠

2013/8/29

大綱

2

- 全民健保資料庫資料內容
- 醫療保費計算所需的統計資料
- 住院發生率風險因子
- 討論與建議

全民健保料庫

3

● 全民健保料庫：資料來源

1. 國家衛生研究院：<http://nhird.nhri.org.tw/>

2. 衛生福利部健康資料加值應用協作中心

<http://biostat.cmu.edu.tw/~nhip/indexn.html>

---- 臺灣大學分中心

<http://ntuhdrc.blog.ntu.edu.tw/>

---- 成功大學分中心

<http://healthdata.rsh.ncku.edu.tw/files/11-1294-10917.php>

---- 中國醫藥大學之研究分中心

2013/8/29

全民健保料庫

4

1. 國家衛生研究院資料庫內容：

- 基本資料檔、系統抽樣檔、特定主題分檔、承保抽樣歸人檔、糖尿病人抽樣歸人檔

基本資料檔：

1. 醫事機構病床主檔 (BED)
2. 醫事機構診療科別明細檔 (DETA)
3. 醫事機構基本資料檔 (HOSB)
4. 醫事機構副檔資料檔 (HOSX)
5. 專科醫師證書主檔 (DOC)
6. 醫事人員基本資料檔 (PER)
7. 重大傷病證明明細檔 (HV)
8. 醫事機構服務項目檔 (HOX)
9. 藥品主檔 (DRUG)
10. 承保資料檔 (ID)

原始資料檔：

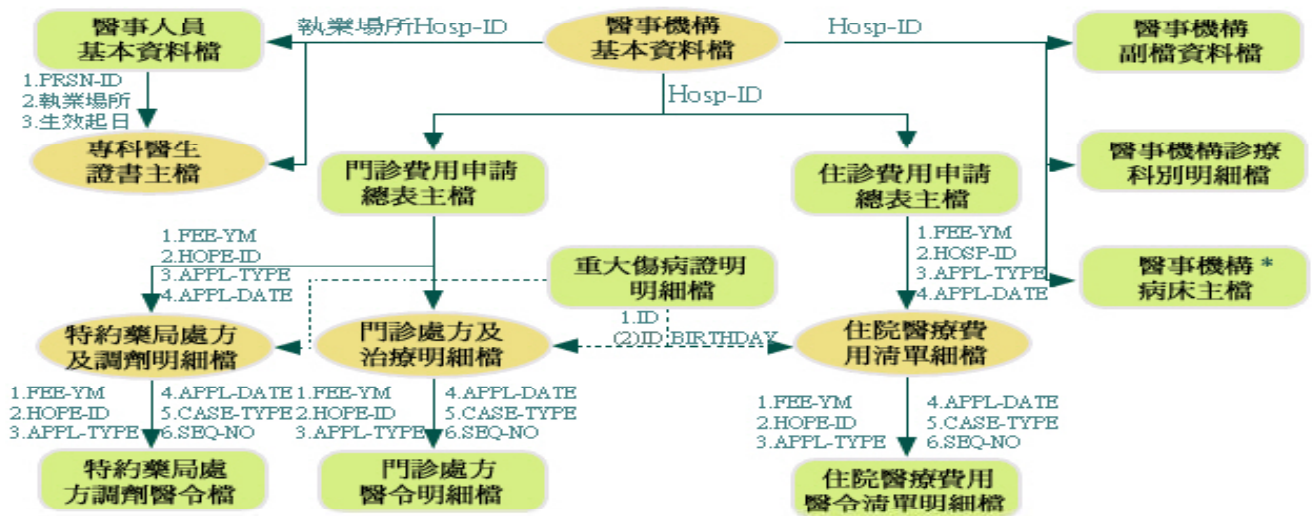
1. 住院費用申請總表主檔 (DT)
2. 門診費用申請總表主檔 (CT)
3. 住院醫療費用清單明細檔 (DD)
4. 住院醫療費用醫令清單明細檔 (DO)
5. 門診處方及治療明細檔 (CD)
6. 門診處方醫令明細檔 (OO)
7. 特約藥局處方及調劑明細檔 (GD)
8. 特約藥局處方醫令檔 (GO)
9. 承保資料檔 (ID)

2013/8/29

全民健保料庫

5

各檔案間串檔變項說明



註:*須注意生效起訖日期

(2)可由ID+BIRTHDAY串檔

→ 各檔案間由所註明變項串檔可獲得對應資訊

→ 各檔案間可由所註明變項串檔,但未必獲得對應資料

2013/8/29

全民健保料庫

6

2. 衛生福利部健康資料加值應用協作中心

目前健康加值中心有二種作業方式:

- (1). A可釋出: 但只有死因檔(不提供身分證字號欄位、出生資料、死亡日期(出生年月、死亡年月有提供),醫療現況及醫院服務量檔----符合收費減免條件者,需檢附公文,並敘明符合減免之事由。
- (2). B協作處理: 需付資料處理費,派人至中心分析費或請中心代分析費。

2013/8/29

全民健保料庫

7

2. 衛生福利部健康資料加值應用協作中心

目前健康加值中心以下為加值中心所有資料檔名稱: 衛生署統計室資料檔清單共37種: 例如下表

資料檔代碼	資料檔名稱	周期	資料期間	資料欄位	資料量	資料提供單位
Health-01	全民健保處方及治療明細檔_門急診	年	1998-2011	包括門診之處方及治療資料, 欄位有費用年月、醫療院所代碼、身分證字號、國際疾病碼、申報費用等	72GB/年	健保局
Health-02	全民健保處方及治療明細檔_住院	年	1998-2011	包括住院之處方及治療資料, 欄位有費用年月、醫療院所代碼、身分證字號、國際疾病碼、申報費用等	1.4GB/年	健保局

2013/8/29

全民健保料庫

8

2. 衛生福利部健康資料加值應用協作中心: 資料檔清單

Release-02	釋出-醫院醫療服務量檔	年	1994-2011	欄位有醫院門診人次、急診人次、手術人次、剖腹產人次、昂貴醫療儀器使用人次等欄位。	1MB/年	統計室
Release-03	釋出-醫療機構現況檔	年	1994-2011	欄位有醫事機構家數、層級別、評鑑等級、醫事人員數、病床數等欄位。	16MB/年	統計室

※另有200萬人之就醫及死因資料檔: 起始年分為兩種, 一89年開始往後10年之資料, 一為94年開始往後5年之資料, 每次申請直接提供健保資料及死因之常用欄位。

2013/8/29

醫療保費計算所需的統計資料

9

- 醫療保費計算所需的統計資料
- 人身保險商品審查應注意事項

第77點

健康保險保單中其送審內容涉及各項給付成本之計算者，應依第一百八十四點規定引用統計資料為基礎評估費率；並就未來趨勢分析加以考量反應。

第79點

長年期健康保險應有給付限額或保費調整機制，並說明其風險控管措施。

2013/8/29

醫療保費計算所需的統計資料

10

- 醫療保費計算所需的統計資料
- 人身保險商品審查應注意事項

第214點

保險商品之送審內容涉及疾病名稱定義時，應由該疾病之相關專科醫師檢視其定義之合理性，但採用中華民國人壽保險商業同業公會研定並報主管機關之重大疾病定義者，不在此限。

2013/8/29

醫療保費計算所需的統計資料

11

- 醫療保費計算所需的統計資料
- 人身保險商品審查應注意事項

第185點

各種計算公式之符號**定義應明確**。

第187點

計算純保費採用安全加成係數（含各項**危險發生率**外加之**標準差**、**惡化率**、**改善率**等）者，應就其合理性詳予說明。

2013/8/29

醫療保費計算所需的統計資料

12

- 醫療保費計算所需的統計資料
- 人身保險商品審查應注意事項

第184點

引用統計資料時，應依下列規定辦理：

（一）所引用之經驗資料，**應採最近三至五年統計資料。但無最近三至五年統計資料者不在此限。**

（二）引用國內外資料者應確實檢附所引用之資料，採用公司本身經驗資料者應檢附統計表報。

（三）依據所引用國內外資料修正或組合訂定者，應敘明計算及其過程；依據公司本身經驗資料修正或組合訂定者亦同。

（四）引用國內外資料訂定預定危險發生率時，應配合保單條款**除外責任及給付條件**，例如，重大疾病有關癌症不包括原位癌症等引用計算發生率資料應不包括原位癌症。**引用國外資料（含再保公司提供）應注意與保單條款給付條件完全一致。**

2013/8/29

醫療保費計算所需的統計資料

13

- 醫療保費計算所需的統計資料

- 外部經驗資料

- 1.台灣壽險業經驗損失率研究報告。
- 2.政府及醫療院所或學會出版刊物。
- 3.再保險公司提供。
- 4.國內或國際學術期刊報告。

2013/8/29

醫療保費計算所需的統計資料

14

- 醫療保費計算所需的統計資料

- 外部經驗資料例如：

- 1.全民健康保險醫療疾病統計月報。
- 2.全民健康保險統計「保險對象人數-按年齡性別分」。
- 3.台灣地區公私立醫院診所診治疾病與傷害調查報告。
- 4.衛生統計(二)醫療機構現況及醫院醫療服務量統計之年度醫療服務量之「歷年醫院住院人次-按病床別分」「歷年醫院平均住院日數-按病床別分」。
- 5.特約醫事服務機構住院醫療費用明細-按性別及年齡別分
- 6.全民健康保險醫療統計年報之「出院患者平均住院日數統計-按疾病別、性別及年齡分」、「住院手術人次統計-按性別及年齡分」、「急診人次統計-按疾病別、性別及年齡分」、「急診件數統計-按疾病別、性別及年齡分」

2013/8/29

醫療保費計算所需的統計資料

15

- 醫療保費計算所需的統計資料

- 外部經驗資料例如：

- 7.全民健康保險統計-特約醫事服務機構住院醫療費用明細(一般案件、特定案件、論病例計酬案件)
- 8.團法人保險事業發展中心出版「癌症相關發生率之研究
- 9.臺北醫務管理學會統計資料「急診轉一般病房率」、「急診轉住加護病房率」
- 10.「台灣壽險業個人壽險經驗死亡率及解約失效率研究報告」
- 11.癌症登記報告
- 12.美國精算協會報告
- 13.美國及香港實際經驗資料
- 14.日本經驗資料

2013/8/29

醫療保費計算所需的統計資料

16

- 醫療保費計算所需的統計資料

- 內部經驗資料

投保人數、保費收入、理賠的人數(件數)，每人平均理賠次數(件數)、每次(件)平均給付金額、相關數據(如各項數據標準差、脫退率、解約率等)

----外部資料通常不能完全滿足醫療保險精算的需要

- 1.保險經營資料的累積和統計分析
- 2.商品種類、定義等之標準化
- 3.理賠資料病名、定義等之標準化(如配合ICD 9,10)

-擁有足夠的內部資料以供保費計算之用。

2013/8/29

醫療保費計算所需的統計資料

17

- 醫療保費計算所需的統計資料
- 健保次級資料使用
 1. 資料係以當年12月的登記資料為切出點
 2. 每月申報資料係原始申報檔而非核付檔
 3. 以保險收支為基礎所建構的健保資料庫中，有關醫療使用者（即病患）的變數收集相當有限
 4. 無法得知病患家戶組成
 5. 僅記錄病患就醫地點，無法得知實際居住地
 6. 在個人特性上僅紀錄病患性別與年齡，其他重要社經特性（如教育、就業、所得等）皆無清楚記載
 7. 病人與醫生及醫療機構層次關係：同一病人可到不同醫院看不同醫生看相同疾病—使用醫療機構服務量資料要注意

醫療保費計算所需的統計資料

18

- 醫療保費計算所需的統計資料
- 健保次級資料使用統計名詞定義(衛生動向):例
 1. 就診人數：係以疾病碼與身分證字號為歸戶鍵值計算而成
 2. 就診件數：係醫療院所向健保局申請之就醫件數。
 3. 手術人次：係包括主、副手術處置碼之手術件數。

資料使用限制：國際疾病碼係採健保欄位中國際疾病分類號為罹病原因，即門診檔中有三組疾病分類號、住院檔有五組診斷代碼，於歸戶時，亦即相同身分證字號但不同疾病碼時，視為不同人數計算，但於計算就醫件數與醫療費用時，則以第一組疾病碼(主診斷碼)為罹病原因。

 4. 門診就診率($=\text{門診人數} \div \text{年中人口數} \times 100\%$)
 5. 每年住院就診率($=\text{住院人數} \div \text{年中人口數} \times 100\%$)

2013/8/29

醫療保費計算所需的統計資料

19

● 醫療保費計算所需的統計資料

● 健保次級資料使用統計名詞定義(醫療服務量):例

- 1.門診人次指到醫療院所診治，經掛號而使用其**病歷表**的次數。
- 2.住院人日：即全年內每日有辦理住院手續之住院病人人數累計。計算方式為**算進不算出**，如為當日住出院者仍以一住院人日計。
- 3.住院人次：全年內前往醫院辦理住院手續之入院人次之累計，但不包括上年留院病人。

2013/8/29

醫療保費計算所需的統計資料

20

● 醫療保費計算所需的統計資料

● 健保次級資料使用統計名詞定義(醫療服務量):例

- 4.平均住院日數=住院總人日數/住院總人次數
- 5.出院人次：全年內每日辦理住院手續之出院人數之累計，包括病癒、轉院及死亡，但不包括**急診觀察病床之出院人次**。
- 6.手術人次：全年內醫院施行手術之次數。
 - (1)分為門診及住院手術人次統計，住院後手術與手術後住院皆列住院手術，其餘為門診手術人次。
 - (2)以手術登記簿之人次為準。
 - (3)含剖腹產及人工流產。

2013/8/29

醫療保費計算所需的統計資料

21

- 醫療保費計算所需的統計資料

- 健保次級資料使用統計名詞定義(醫療給付):例

- 1.門診件數:全年內前往特約醫事機構經掛號後，實際赴門診就醫之件數，**並含急診件數**。如某A同時掛兩個門診
某次看病主診斷碼為糖尿病 次診斷碼為高血壓，另一次看病主診斷碼為急性支氣管炎 次診斷碼為流行性感冒
記錄為人次：4 件數：2 人數：1
- 2.住院件數:全年住院費用之件數，**住院期間醫療院所分數次申報費用**，每一次算一件。
- 3.住院日數:依申報規定以保險對象入院之日起計，包括佔急性病床日數和慢性病床日數，出院當日不計。
- 4.申報點數:費用發生年月之申報點數

2013/8/29

住院發生率風險因子

22

- 醫療險的逆選擇和道德風險較高，計算保費時需將風險特徵不同的人群分組計算

➔減少醫療險的逆選擇

- 統計資料的限制:目前實際只利用年齡、性別兩個主要因素進行分類。

➔當big data時代來臨時

所得,教育,職業,婚姻與住院發生率之相關性

2013/8/29

住院發生率風險因子

23

- 家戶面全民健保利用概況：
- 以居住於臺灣地區內具有中華民國國籍之個人及其所組成之家庭（指營共同經濟生活者所組成之家庭）為調查對象。（不含共同事業戶人口及年內達半年在營義務役軍人）：自民國100年起，採按年辦理及固定樣本15,857 戶。

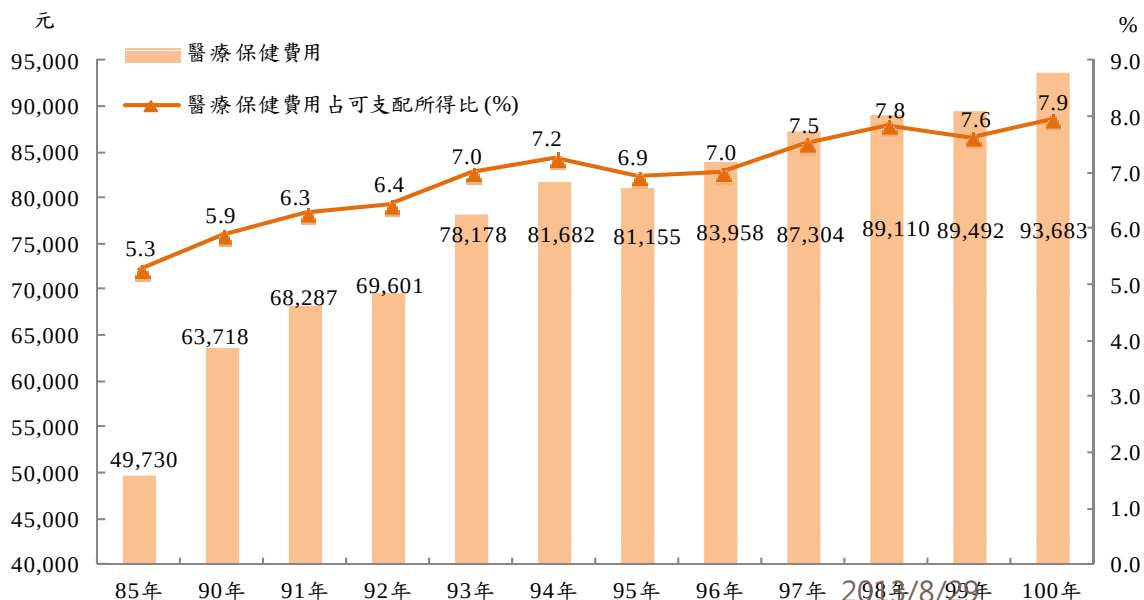
2013/8/29

住院發生率風險因子

24

- 家戶面全民健保利用概況：

平均每戶醫療保健費用占可支配所得比



住院發生率風險因子

25

● 家戶面全民健保利用概況：

平均每人醫療費用

	醫療保健費用			醫療照護自付費用				健保保險受益			
	金額(元)	定基指數	年增率(%)	金額(元)	占率(%)	定基指數	年增率(%)	金額(元)	占率(%)	定基指數	年增率(%)
85年	13,684	100.0	—	3,215	23.5	100.0	—	10,470	76.5	100.0	—
90年	19,198	140.3	1.4	5,252	27.4	163.4	8.9	13,946	72.6	133.2	4.4
91年	20,801	152.0	8.3	5,671	27.3	176.4	8.0	15,130	72.7	144.5	8.5
92年	21,502	157.1	3.4	5,824	27.1	181.2	2.7	15,677	72.9	149.7	3.6
93年	24,486	178.9	13.9	6,478	26.5	201.5	11.2	18,008	73.5	172.0	14.9
94年	25,944	189.6	6.0	7,437	28.7	231.4	14.8	18,507	71.3	176.8	2.8
95年	26,023	190.2	0.3	7,442	28.6	231.5	0.1	18,581	71.4	177.5	0.4
96年	27,222	198.9	4.6	7,969	29.3	247.9	7.1	19,253	70.7	183.9	3.6
97年	28,710	209.8	5.5	8,509	29.6	264.7	6.8	20,201	70.4	192.9	4.9
98年	29,765	217.5	3.7	8,579	28.8	266.9	0.8	21,187	71.2	202.4	4.9
99年	30,436	222.4	2.3	8,730	28.7	271.6	1.8	21,706	71.3	207.3	2.5
100年	32,266	235.8	6.0	9,510	29.5	295.8	8.9	22,756	70.5	217.4	4.8

資料來源：行政院主計總處「家庭收支調查」、中央健康保險局「全民健康保險統計」年報

附註：醫療保健費用=醫療照護自付費用+健保保險受益

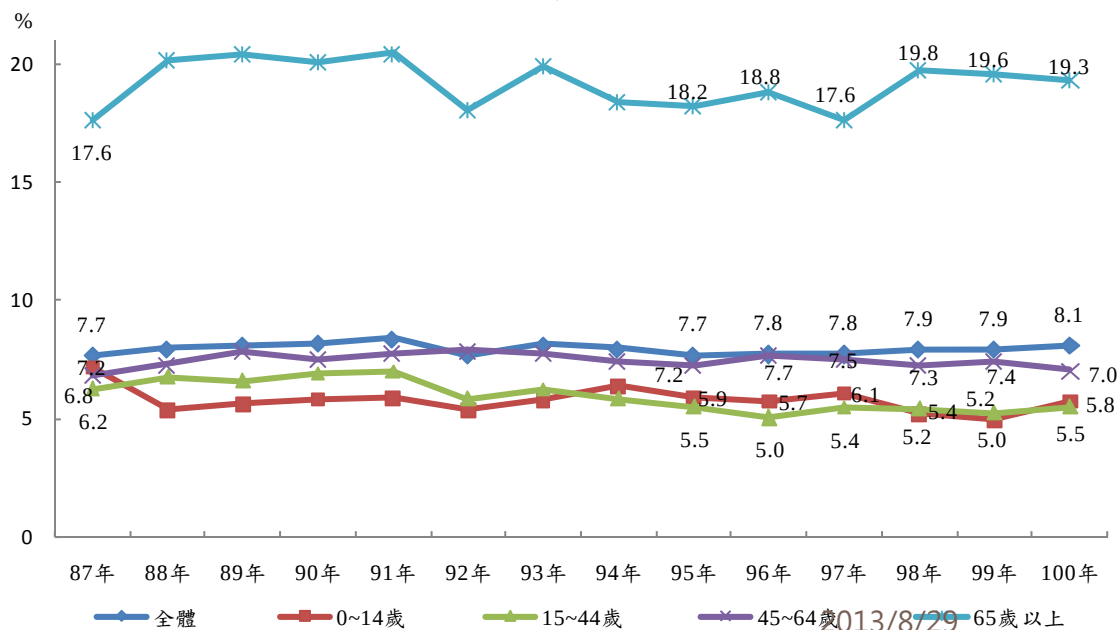
2013/8/29

住院發生率風險因子

26

● 家戶面全民健保利用概況：

住院就診率-按年齡別分



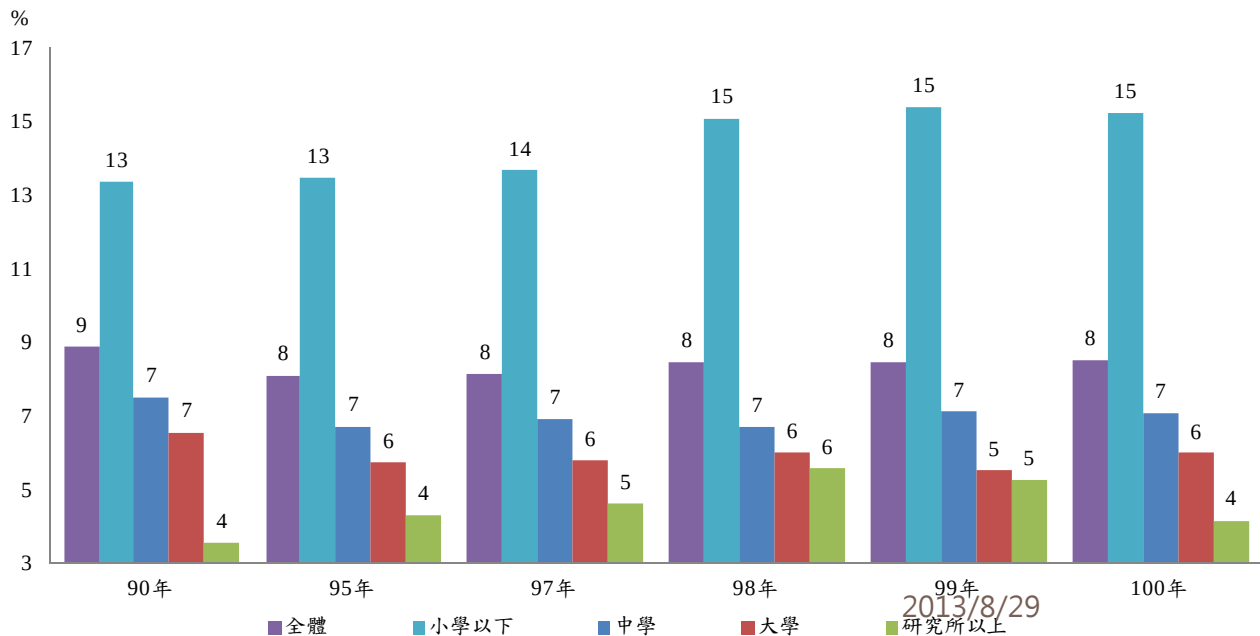
2013/8/29

住院發生率風險因子

27

● 家戶面全民健保利用概況：

15歲以上住院就診率-按教育程度分



住院發生率風險因子

28

● 家戶面全民健保利用概況：

15歲以上住院就診率—按職業別分

單位：%

	全體	無業者	主管人員	專技人員	事務人員	農林漁牧業者	工人	軍人
90年	8.9	12.2	5.4	5.9	6.7	9.5	5.1	1.0
91年	9.0	11.8	5.3	6.5	7.4	10.7	5.4	3.9
92年	8.2	11.1	5.4	5.9	5.7	10.3	4.9	6.3
93年	8.7	11.7	5.0	5.2	5.9	12.1	5.9	3.5
94年	8.3	11.2	5.2	5.1	5.2	10.8	6.1	2.4
95年	8.1	10.9	4.8	5.0	5.7	9.4	5.5	5.4
96年	8.2	11.2	5.4	4.3	5.4	9.6	5.6	2.1
97年	8.1	10.7	6.0	5.3	5.6	9.2	5.6	7.8
98年	8.4	11.3	5.4	5.5	5.9	10.5	4.9	11.2
99年	8.4	11.5	5.4	4.8	5.4	11.2	5.5	1.4
100年	8.5	11.8	4.7	5.2	5.7	8.2	5.4	1.3
歷年平均	8.4	11.4	5.3	5.3	5.9	10.3	5.4	4.5

資料來源：行政院主計總處「家庭收支調查」、行政院衛生署統計室「醫療統計年報」

附註：1.主管人員指民意代表、行政主管、企業主管及經理人員

2.專技人員指專業人員、技術員及助理專業人員

3.事務人員指事務工作人員、事務支援人員、服務及銷售工作人員

4.農林漁牧業者指農事、畜牧、林業、漁業及生產人員

5.工人指技藝有關工作人員、機械設備操作及組裝人員、基層技術工及勞力工

2013/8/29

住院發生率風險因子

29

- 家戶面全民健保利用概況：
- 家庭收支調查**樣本數過低**
- 某些社會邊緣群體如：遊民、受機構安置民眾、無戶籍人口等在這該調查中是無法被抽出的

➔運用分析時應該注意的資料限制,但可**觀察趨勢**

2013/8/29

住院發生率風險因子

30

- 健保實例分析
- 婚姻與住院關係：

1.資料來源使用台灣全民健保資料庫**2005年在保者100萬抽樣歸人檔**與承保檔且**觀察年度為1996-2009年**。

2.資料與研究限制：

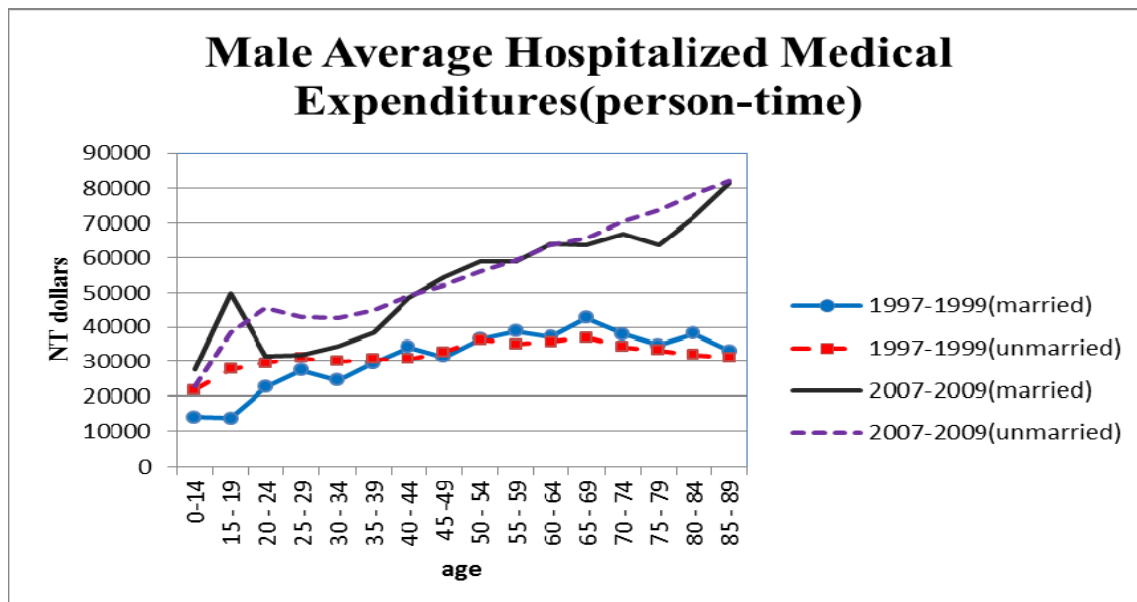
由於100萬抽樣歸人檔在**婚姻狀況別紀錄並不明確**僅能從身分別**INS_ID_TYPE**: 1：雇主無勞保 2：一般 4：榮民/榮民遺眷之家戶代表 5：保險類目為「62」依附6：育嬰 及稱謂代號**INS_RELATION**: 1：配偶 2：父母 3：子女 4：祖父母 5：孫子女 6：外祖父母 7：外孫子女 8：曾祖父母 9：外曾祖父母 0：其他親屬進行判斷,並且搭配承保檔(母體)進行篩選。

2013/8/29

婚姻與住院關係(有無配偶)

31

每次平均住院費用

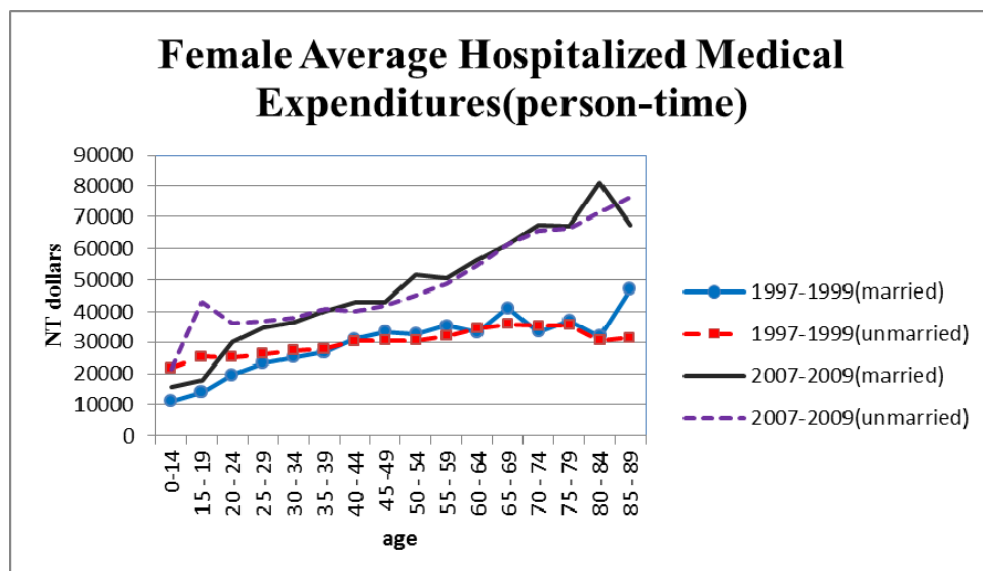


2013/8/29

婚姻與住院關係(有無配偶)

32

每次平均住院費用

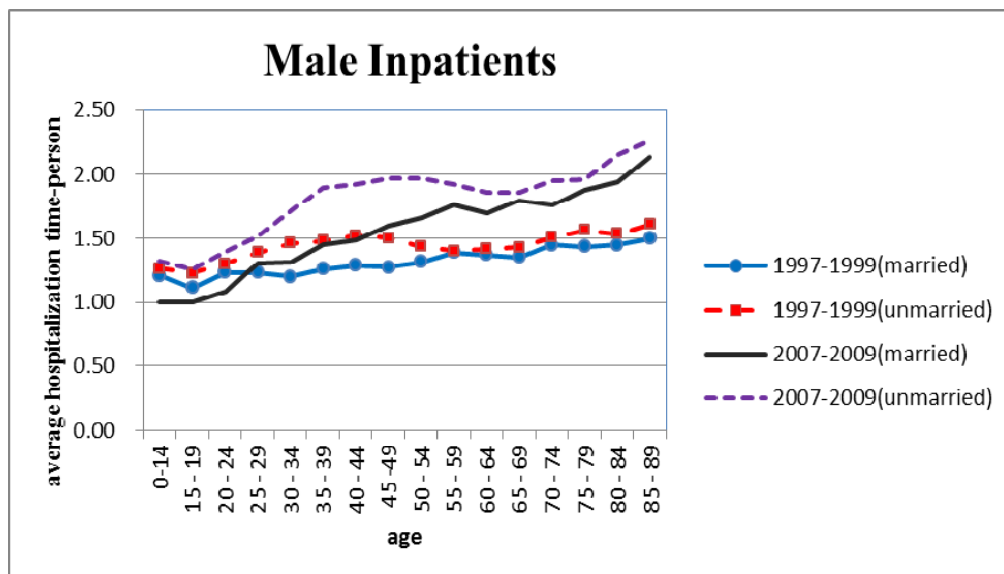


2013/8/29

婚姻與住院關係(有無配偶)

33

每位住院病人住院次數

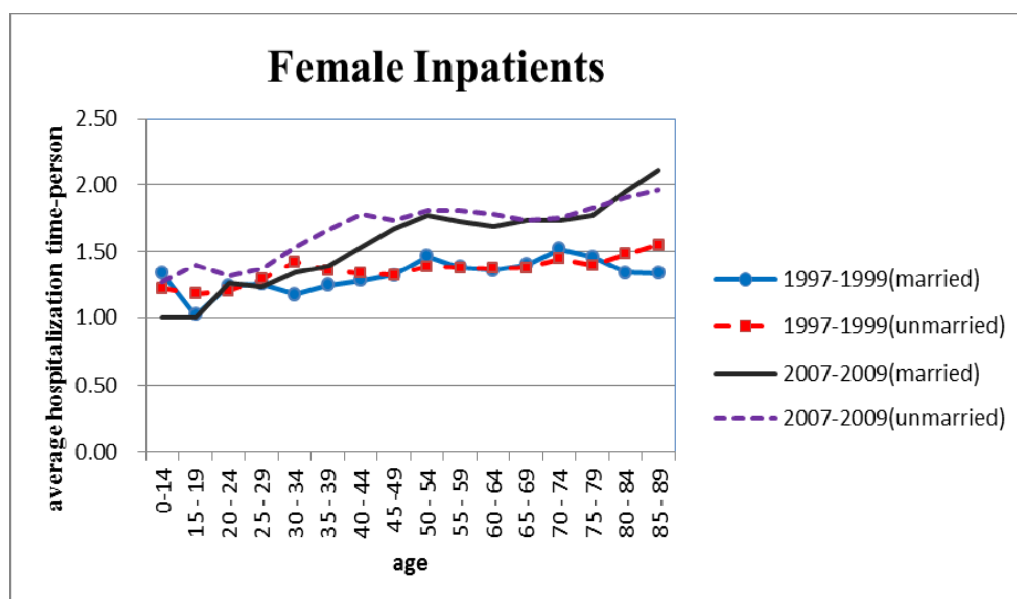


2013/8/29

婚姻與住院關係(有無配偶)

34

每位住院病人住院次數

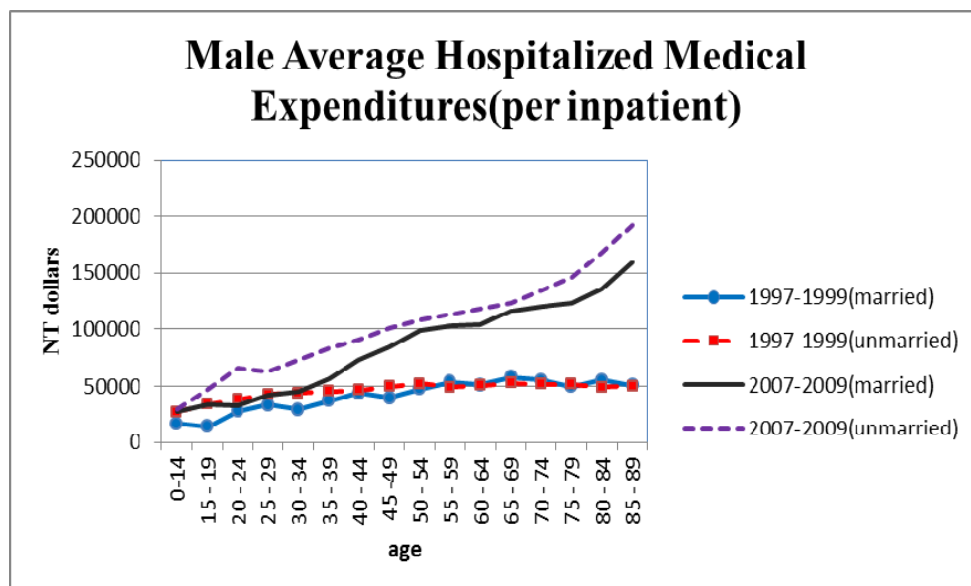


2013/8/29

婚姻與住院關係(有無配偶)

35

每位住院病人平均每年住院費用

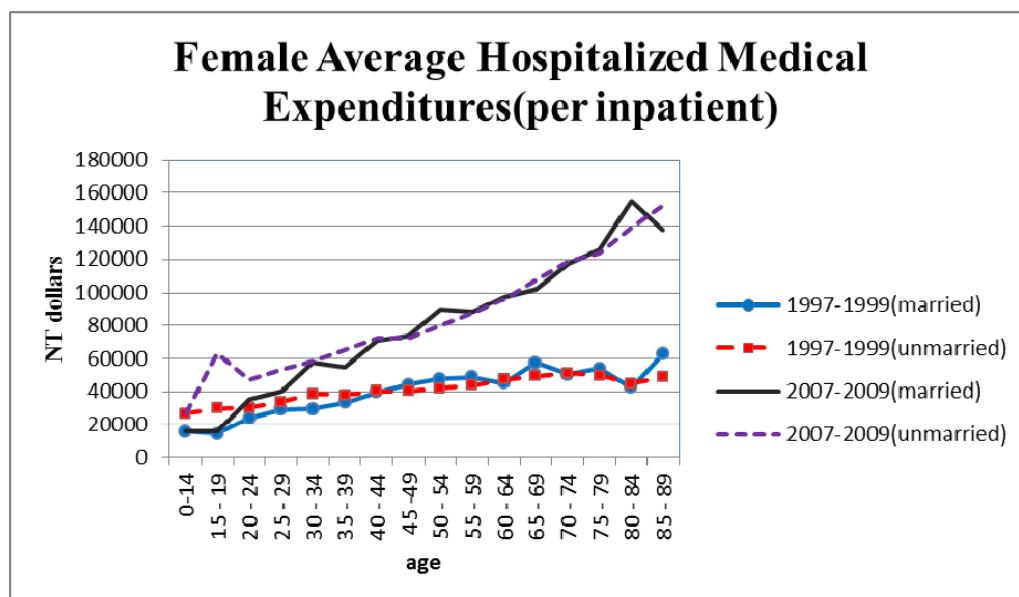


2013/8/29

婚姻與住院關係(有無配偶)

36

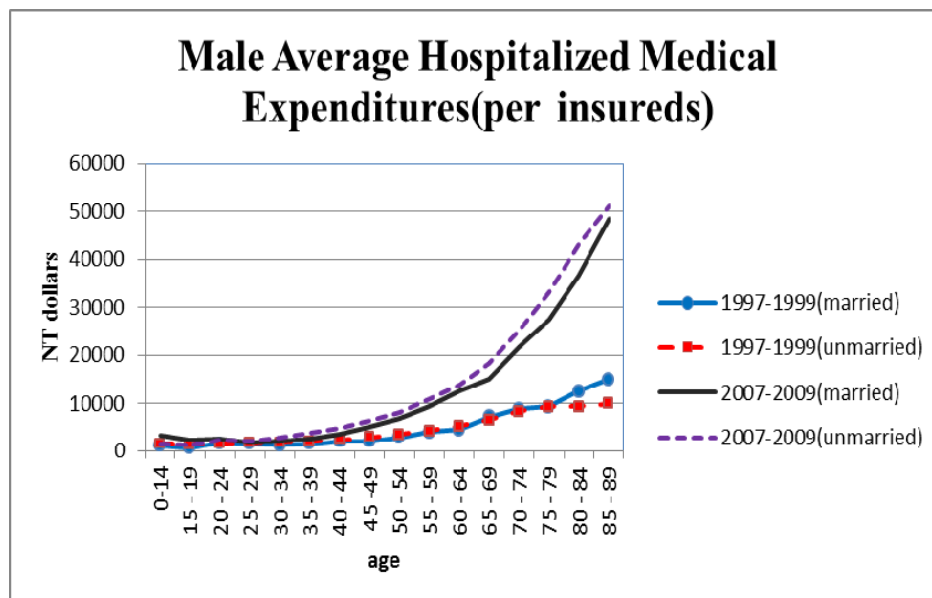
每位住院病人平均每年住院費用



婚姻與住院關係(有無配偶)

37

每人每年平均住院費用(對所有被保險人)

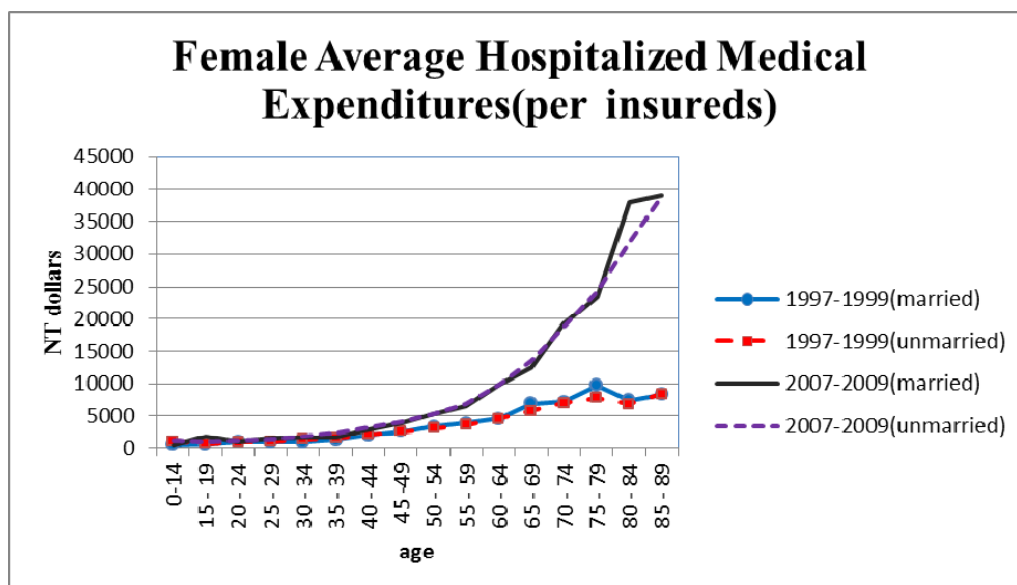


2013/8/29

婚姻與住院關係(有無配偶)

38

每人每年平均住院費用(對所有被保險人)

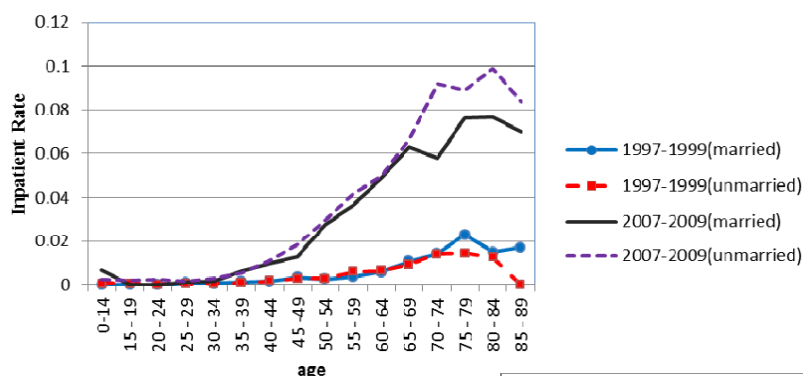


2013/8/29

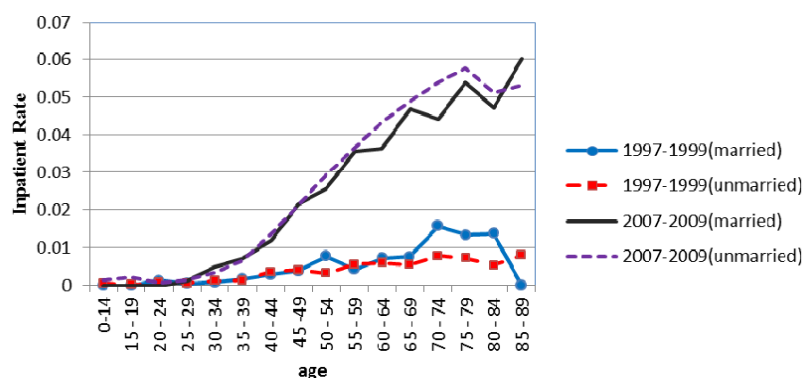
婚姻與住院關係：惡性腫瘤住院發生率 (有無配偶)

39

Male(Malignant neoplasms)



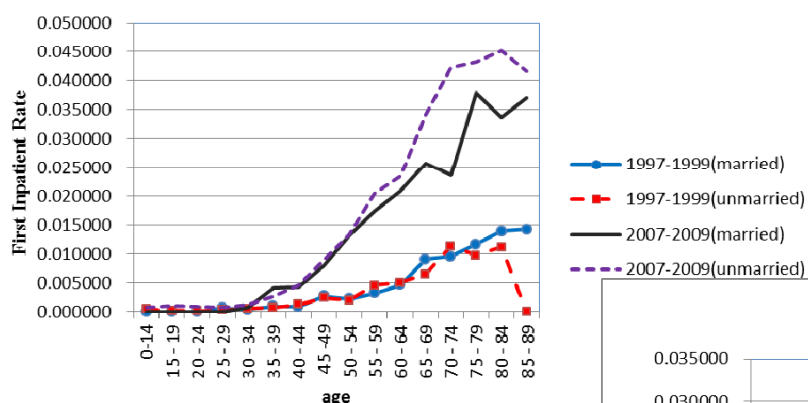
Female(Malignant neoplasms)



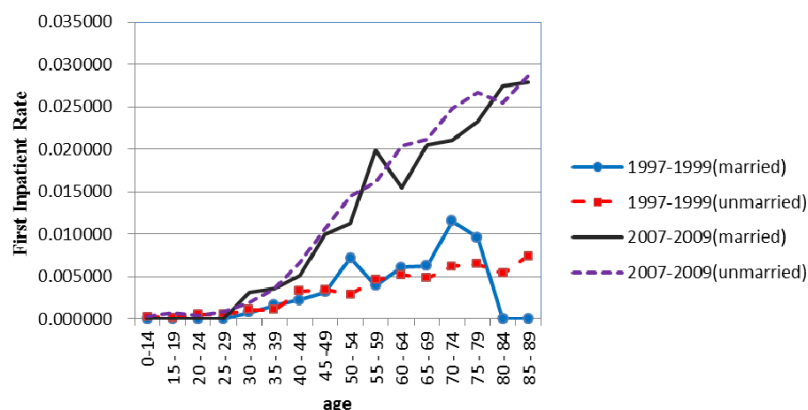
婚姻與住院關係：惡性腫瘤初次住院發生率(有無配偶)

40

Male(malignant tumor)



Female(malignant tumor)



婚姻與住院關係：惡性腫瘤初次住院發生率（有無配偶）

41

惡性腫瘤是台灣十大死因的榜首,也就是國人目前致死率最高的病因.從惡性腫瘤初次住院發生率可以發現男性、女性住院發生率隨著年代有增加趨勢,並且對男性而言早期婚姻關係並無多大影響,但近代以2007-2009年為例,40歲以上無配偶者初次住院發生率較有配偶者高；對女性而言,35歲以上無配偶者初次住院發生率較有配偶者高；男性較女性差異大。以住院發生率也得到相同的結果

2013/8/29

住院發生率風險因子

42

- 健保實例分析
- 職業(投保身分別)與住院關係:

資料來源

本研究使用「全民健康保險資料庫」中的2005年100萬承保抽樣歸人檔之承保資料檔(ID檔)及住院醫療費用清單明細檔(DD檔)。

- ID檔2002年至2010年總筆數共10,795,975。
- DD檔1996年至2010年總筆數共1,574,987。

(住院人次資料，將每個承保對象每次住院醫療紀錄詳細登錄)

2013/8/29

職業(投保身分別)與住院關係: 依單位屬性區分

43

類別	全民健保保險對象分類	
第一類	公務人員、公職人員	
	私校教職員	
	公民營事業、機構等有一定雇主的受雇者	
	雇主、自營業主、專門職業及技術人員自行職業者	
第二類	職業工會會員、外雇船員	
第三類	農、漁民、水利會員	
第四類	義務役軍人、替代役役男	
第五類	合於社會求助法規定的低收入戶成員	
第六類	榮民、榮民遺眷家戶代表	2013/8/29

職業(投保身分別)與住院關係:研究限制

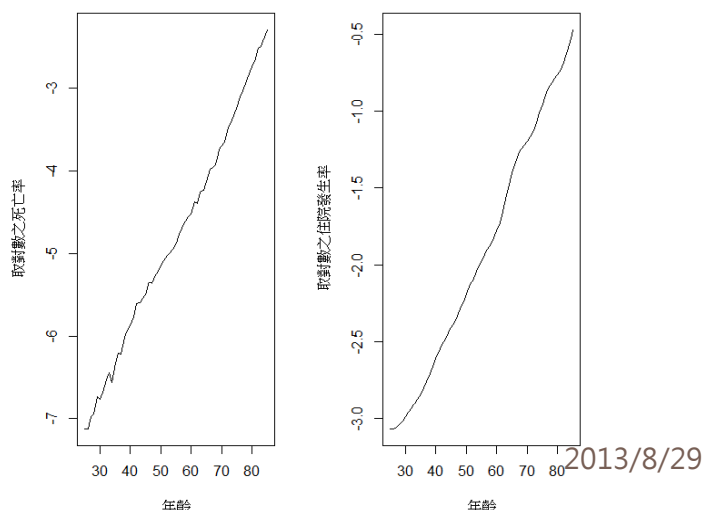
44

- **健保資料庫無職業類別註記**，僅能依投保單位屬性類別區分。
- 2006年以後無新出生人口，依投保對象分類，單齡在低年齡層有遺漏值，無法對低年齡層進行分析；因此本研究採用五齡組資料統計分析。

職業(投保身分別)與住院關係:研究方法 – Lee-Carter模型

45

- 由於Lee-Carter模型是結合年度變化與年齡型態來解釋觀測到的死亡率預測模型，我們所使用之住院發生率年齡趨勢也同時具有與死亡率年齡型態相同趨勢，因此將其應用在疾病發生率及醫療費用的研究上。



職業(投保身分別)與住院關係:研究方法–Lee-Carter模型

46

- 本文使用Lee-Carter模型用於配適與預測住院發生率，基本模型架構如下：

$$\ln(m_{x,t}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t} \cdot \varepsilon_{x,t} \sim N(0, \sigma_\varepsilon^2)$$

α_x ：x歲住院發生率取對數後的平均值；

β_x ：取對數後住院發生率在x歲影響之變化速度；

κ_t ：取對數的住院發生率中年度與年度之間變化之係數；

$\varepsilon_{x,t}$ ：x歲在第t年之誤差項。

職業(投保身分別)與住院關係:平均絕對誤差

47

平均絕對誤差(Mean absolute percentage error, MAPE)

- 主要用於判斷模型配適與預測模型優劣之指標，其值為相對數值，不受實際值或配適值與預測值之單位的影響
- 計算公式如下：

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{A_t - F_t}{A_t} \right| * 100\%$$

A_t ：實際值、 F_t ：配適值或預測值、 N ：樣本總數

MAPE值之評估準則：

MAPE(%)	MAPE評估標準
<10	高準確的預測
10-20	優良的預測
20-50	合理的預測
>50	不準確的預測

2013/8/29

職業(投保身分別)與住院關係:Whittaker修勻法

48

Whittaker修勻法同時考慮適度性與平滑性，基本型態如下：

$$M = F + hS = \sum_{x=1}^n w_x (v_x - u_x)^2 + h \sum_{x=1}^{n-z} (\Delta^z v_x)^2$$

F ：適度性；

S ：平滑性；

u_x ： x 歲的住院發生率；

v_x ：調整後 x 歲的住院發生率；

w_x ： x 歲的權數(本研究假設為1)；

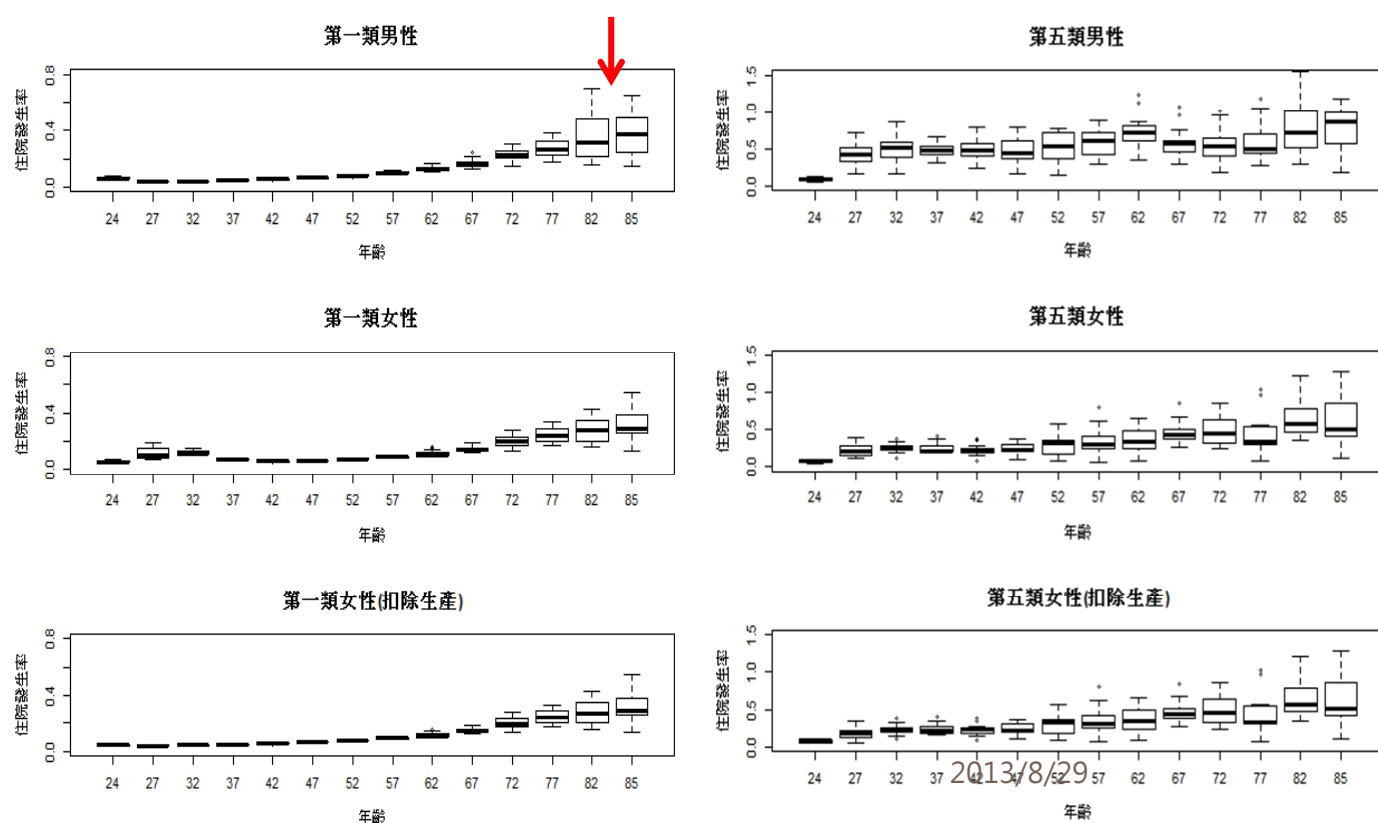
h 及 z 為選定的參數， Δ^z 為第 z 次之差分(difference)。

- 本文採用 z 值為2，在實務應用上重視其平滑性，因此 h 值假設為100。

2013/8/29

職業(投保身分別)與住院關係:住院發生率

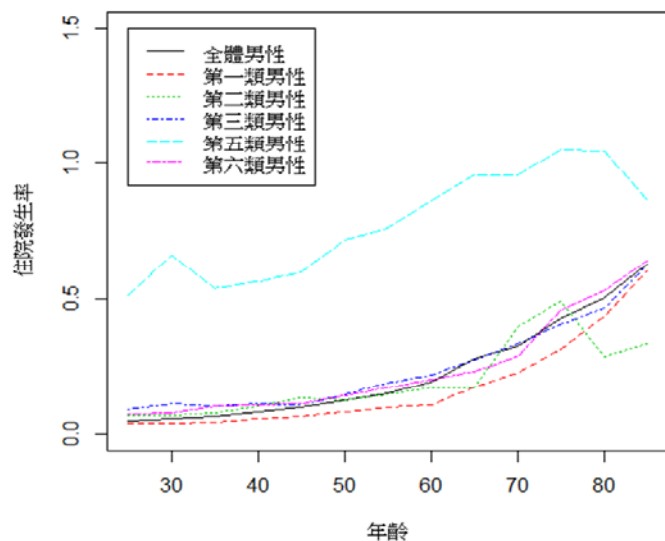
50



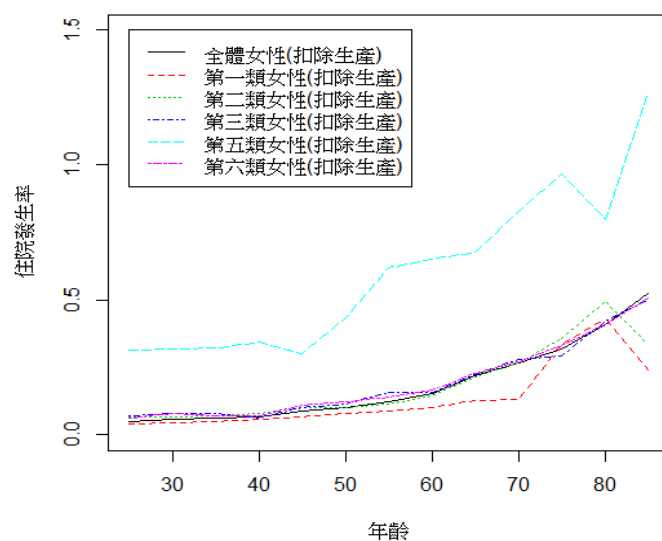
職業(投保身分別)與住院關係: 2010年健保資料庫全體與各投保對象發生率趨勢圖

51

男性



女性(扣除生產)



2013/8/29

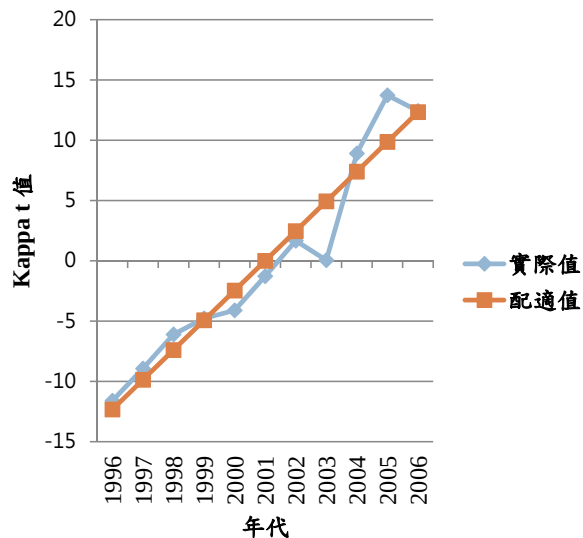
52

2013/8/29

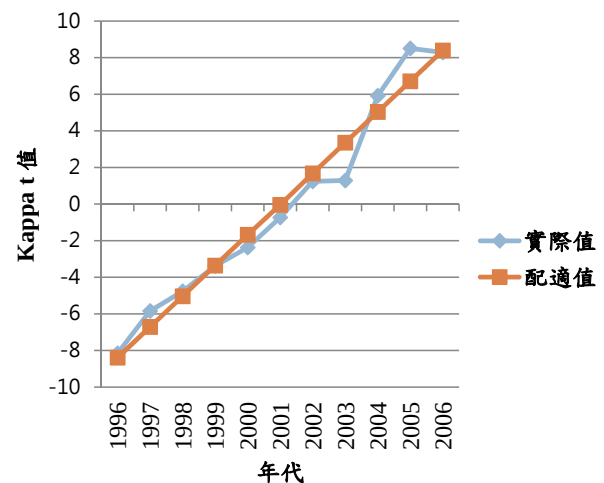
職業(投保身分別)與住院關係:住院發生率之 K_t 實際值與配適值趨勢圖

53

男性



女性(扣除生產)



2013/8/29

職業(投保身分別)與住院關係:住院發生率整體配適與實際MAPE值及K-S檢定p-value

54

➤ Training Period :

1996年~2006年

➤ Testing Period :

2007年~2010年

性別	男性	女性 (扣除生產)
MAPE值	2.68%	2.91%

年代 (p值)	男性	女性 (扣除生產)
1996	0.3879	0.9327
1997	0.8211	0.8211
1998	0.8211	0.8211
1999	0.9878	0.8211
2000	0.9999	0.9327
2001	0.9878	0.9999
2002	0.9878	0.8211
2003	0.9999	0.9878
2004	0.9996	0.9878
2005	0.9999	0.9999
2006	0.9999	0.9996

2013/8/29

職業(投保身分別)與住院關係:住院發生率預測之MAPE值(交叉驗證)

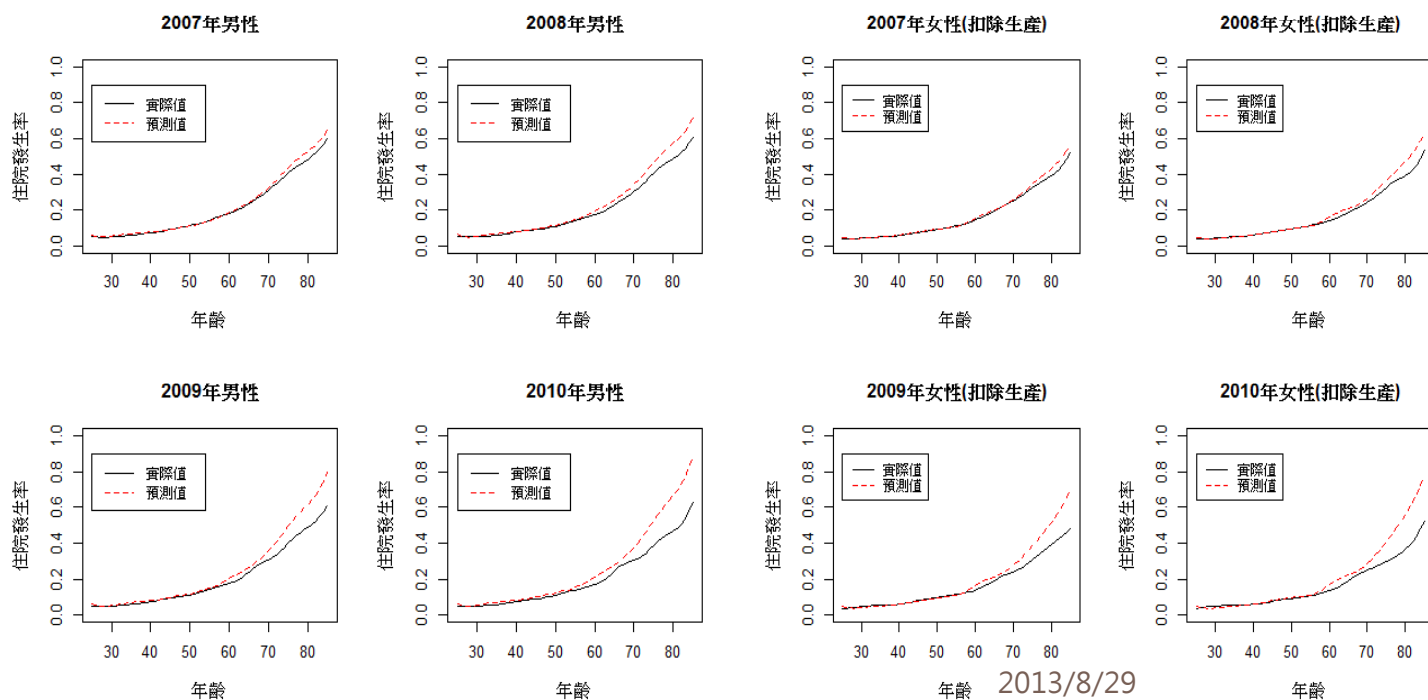
55

年齡層、年度	性別	
	男性	女性(扣除生產)
年齡層		
中齡層(25~64)	9.67%	6.54%
高齡層(65~85以上)	18.87%	19.75%
年度		
2007	5.77%	4.06%
2008	10.41%	7.75%
2009	13.95%	13.20%
2010	21.22%	19.34%
整體	12.84%	11.09%

2013/8/29

職業(投保身分別)與住院關係:住院發生率預測及實際值比較

56



2013/8/29

職業(投保身分別)與住院關係:靜態住院發生率、動態住院發生率

57

- 靜態period住院發生率：

本文選用插捕後的後五年(2006年至2010年)住院發生率平均值做為靜態發生率。

- 動態cohort住院發生率：

考慮年齡因時間推移而不同，本文使用1996年至2010年住院發生率做為歷史資料，以Lee-Carter模型配適後進行預測。

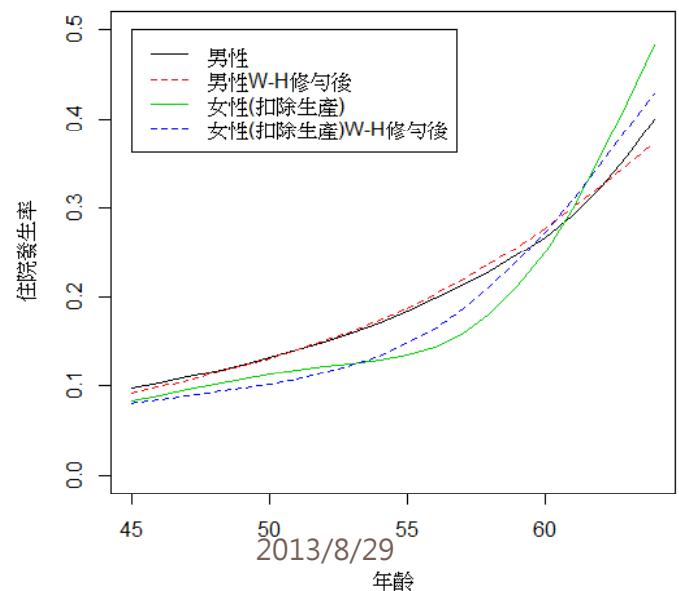
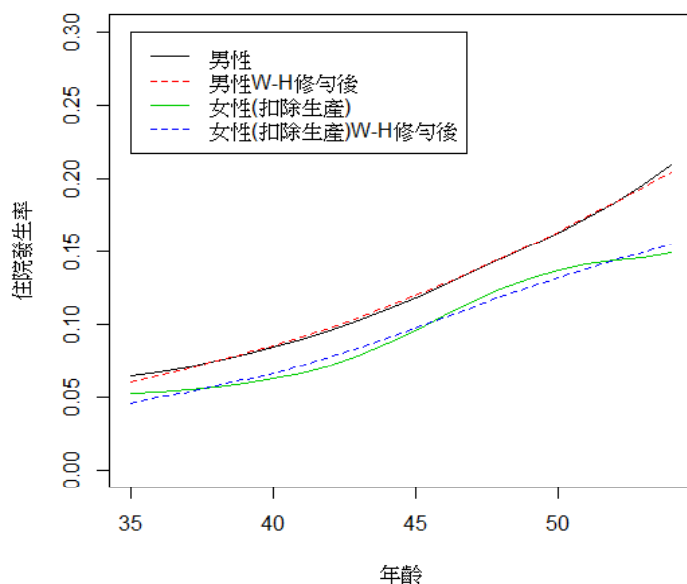
2013/8/29

職業(投保身分別)與住院關係:修勻前後預測20年之動態住院發生率

58

35歲

45歲



- 保險種類與純保險費計算
- 純保險費比較

2013/8/29

32/42

職業(投保身分別)與住院關係:家庭收支調查結果與意外險職業分類費率比

60

家庭收支調查結果分類	比值
全體	1
無業者	1.36
主管、專技、事務人員	0.65
農林漁牧業者	1.23
工人	0.64
軍人	0.54

VS.

意外險職業分類	費率比
第一類(內勤行政)	1
第二類(外勤業務)	1.25
第三類(技術指導)	1.5
第四類(現場操作)	2.25
第五類(高危險)	3.5
第六類(高危險)	4.5

2013/8/29

職業(投保身分別)與住院關係:比值乘數

61

男性	case1(μ)	σ	case2($\mu + \sigma$)	case3($\mu + 2 \times \sigma$)	天數比值
第一類	0.72	0.11	0.84	0.95	0.74
第二類	1.13	0.20	1.34	1.54	0.89
第三類	1.24	0.29	1.54	1.83	0.90
第五類	5.73	3.30	9.03	12.33	1.15
第六類	1.18	0.20	1.39	1.59	0.88

女性 (扣除生產)	case1(μ)	σ	case2($\mu + \sigma$)	case3($\mu + 2 \times \sigma$)	天數比值
第一類	0.77	0.06	0.83	0.89	0.81
第二類	1.14	0.18	1.32	1.51	0.92
第三類	1.15	0.21	1.36	1.58	0.86
第五類	3.70	1.36	5.07	6.43	1.17
第六類	1.17	0.11	1.28	1.40	0.89

2013/8/29

34/42

職業(投保身分別)與住院關係:每千元住院醫療日額保險成本

62

- 計算純保險費前，須先計算X歲時住院日額保險金理賠成本($HCOI_x$)，公式如下：

$$HCOI_x = (1 + 0.3) \times q_x^h \times HD_x$$

q_x^h 為X歲時住院發生率，在計算時需加計30%安全係數，
 HD_x 為X歲時平均住院天數。

- 各年齡每千元住院醫療日額醫療保險成本，公式定義如下：

$$NPR_x = \frac{\sum_{i=0}^{n-1} HCOI_{x+i} \times D_{x+i-1}}{N_x - N_{x+t}} \times R^n$$

其中， D_{x+i-1} 、 N_x 為基數表中計算之基數， R^n 各類的比值乘數。

職業(投保身分別)與住院關係: 20年期住院醫療健康險(預定利率2%)

63

● case1 : 以平均數比值乘數

男性							
年齡	20年期純保費 (i=2%)	全體	第一類	第二類	第三類	第五類	第六類
25	靜態住院發生率	1,104	591	1,114	1,232	7,273	1,156
	動態住院發生率	1,225	657	1,236	1,367	8,073	1,283
35	靜態住院發生率	1,485	796	1,499	1,657	9,786	1,555
	動態住院發生率	1,831	981	1,847	2,043	12,060	1,916
45	靜態住院發生率	2,019	1,082	2,037	2,252	13,299	2,113
	動態住院發生率	2,579	1,382	2,603	2,878	16,994	2,700
女性(扣除生產)							
25	靜態住院發生率	781	489	823	768	3,386	814
	動態住院發生率	757	474	797	745	3,281	789
35	靜態住院發生率	1,122	703	1,182	1,104	4,863	1,169
	動態住院發生率	1,312	822	1,382	1,290	5,686	1,367
45	靜態住院發生率	1,583	992	1,667	1,557	6,861	1,650
	動態住院發生率	2,259	1,415	2,380	2,222	9,792	2,354

男性純保費折扣
46.42%(0.5358折)

女性純保費折扣
37.36%(0.6264折)

2013/8/29

職業(投保身分別)與住院關係: 20年期住院醫療健康險(預定利率2%)

64

● case2 : 以平均數加計標準差

男性							
年齡	20年期純保費 (i=2%)	全體	第一類	第二類	第三類	第五類	第六類
25	靜態住院發生率	1,104	684	1,313	1,521	11,454	1,353
	動態住院發生率	1,225	759	1,457	1,689	12,714	1,501
35	靜態住院發生率	1,485	920	1,767	2,047	15,413	1,820
	動態住院發生率	1,831	1,134	2,177	2,523	18,995	2,243
45	靜態住院發生率	2,019	1,250	2,401	2,782	20,945	2,473
	動態住院發生率	2,579	1,598	3,068	3,555	26,765	3,161
女性(扣除生產)							
25	靜態住院發生率	781	527	954	912	4,632	891
	動態住院發生率	757	511	925	884	4,489	863
35	靜態住院發生率	1,122	757	1,370	1,309	6,653	1,280
	動態住院發生率	1,312	885	1,602	1,531	7,779	1,496
45	靜態住院發生率	1,583	1,068	1,934	1,848	9,387	1,806
	動態住院發生率	2,259	1,524	2,759	2,637	13,396	2,579

男性純保費折扣
38.06%(0.6194折)

女性純保費折扣
32.59%(0.6744折)

2013/8/29

職業(投保身分別)與住院關係: 20年期住院醫療健康險(預定利率2%)

65

● case3：以平均數加計2倍標準差

男性								男性純保費折扣 29.71%(0.7029折)
年齡	20年期純保費 (i=2%)	全體	第一類	第二類	第三類	第五類	第六類	
25	靜態住院發生率	1,104	776	1,512	1,811	15,636	1,550	
	動態住院發生率	1,225	861	1,678	2,010	17,355	1,720	
35	靜態住院發生率	1,485	1,044	2,034	2,437	21,040	2,085	
	動態住院發生率	1,831	1,287	2,507	3,003	25,929	2,570	
45	靜態住院發生率	2,019	1,419	2,765	3,312	28,592	2,834	
	動態住院發生率	2,579	1,813	3,533	4,232	36,535	3,621	
女性(扣除生產)								女性純保費折扣 27.76%(0.7224折)
25	靜態住院發生率	781	564	1,085	1,055	5,878	968	
	動態住院發生率	757	547	1,052	1,022	5,697	938	
35	靜態住院發生率	1,122	811	1,559	1,515	8,443	1,390	
	動態住院發生率	1,312	948	1,823	1,772	9,872	1,626	
45	靜態住院發生率	1,583	1,144	2,200	2,138	11,913	1,962	
	動態住院發生率	2,259	1,632	3,139	3,051	17,000	2,799	

討論與建議

66

1. 住院風險因子除年齡與性別外，可以考量分別與婚姻、職業關係

➔ 在高保費時代商品設計可以考量例如

(1) 與體檢事業合作結合健保身分第一類保單

(2) 婚姻保單

(3) 特定疾病保單

2. 家戶調查樣本數太少可以擴大樣本或普查並與健保資料庫串檔分析

➔ 醫療發生率相關因素研究

2013/8/29

討論與建議

67

3. 保險公司善用實際經驗資料

➔ 理賠成本控制(資料標準化與相關性研究)

4. 在巨量資料下:

(1)核保成本: 以信用紀錄與消費行為取代體檢(英傑華保險公司、保德信與AIG研究)

(2)降低保費: 駕駛人依特定時間地點購買保險(英國)

(3)生病預警等

※支持與贊助健保與跨領域調查資料研究

68

Thank you for your
attention.



全民健保資料庫 — 重大傷病研究

政治大學統計系余清祥



演講大綱

- 社會保險 vs. 商業保險
- 重大傷病介紹
- 臺灣重大傷病現況
- 臺灣健保資料庫介紹
- 臺灣重大傷病實證研究



目標與方向

- 配合政府產業政策，以本土資料為研究之對象與依據。
- 以社會保險等大型資料庫為研究資料來源。
- 建立產學合作橋樑，重新定位商業保險的角色。
- 將本土資料研究之成果舉辦產學研討會，提供主管機關及保險產業參考。
- 提供主管機關政策訂定與監理之參考。

3



補償社會保險之不足

協助開發以社會保險補償性質為目的的商業保險：

- 1.調整社會保險的承保範圍與對象
- 2.採用與社會保險相同之傷病定義
- 3.以調整後的社會保險發生率為基礎
- 4.配合社會保險採取相同的理賠標準
- 5.設計商業保險作為補償社會保險之不足。

4



社會保險 vs. 商業保險



社會保險與商業保險

以保險之基本原理而言：

1. 社會保險與商業保險均為保險，其主要的共同點在於
①均採大數法則，②危險共同分擔，③避免道德危險。
2. 相異之處主要在於①目的、②公平性、③保險費計收基礎、④對被保險人之選擇性、⑤被保險人投保之自由性與⑥政府責任等項。
3. 社會保險的承保範圍 $\neq$ 商業保險的承保範圍



目的與公平性

一、目的

社會保險的目的在於社會安全，而商業保險則為追求利潤，其本質的相異非常清楚。但社會保險之收支即使產生「節餘」，仍屬全體被保險人所共有與共享，若有逆差，也由全體被保險人或社會全體負擔。

二、公平性(個人與社會的公平)

商業保險講求個人的公平，危險性高者，保費高；理賠水準高者，保險費也高。社會保險講求的是社會的公平，同樣所得者，負擔相同的保費；富有者相較於貧者，負擔較高的保費，但在患病就醫時享受的醫療照護相同，即個人付費的金額與提供的醫療服務無關。

7



保費基礎與被保險人之選擇性

三、保費計收的基礎

就保費計收的基礎而言，社會保險通常如前述，同樣所得者，負擔相同的保險費，採用以群體的危險依投保薪資定率分擔；商業保險則依個人的危險性及理賠水準訂定，危險性高者，保險費高。

四、對被保險人之選擇性

商業保險對高危險性者，可以加以拒保或收取極高的保險費。而社會保險其目的為增進社會安全，故只要合乎保險身份資格者，一律納入保險，例如公務員、勞工、農民保險以及目前實施的全民健康保險。

8



投保之自由性與政府責任

五、被保險人投保之自由性

商業保險可拒絕高危險者投保，但如同民眾購買一般商品般，也有權可以自由決定購買與否及選擇購買那一家或那一種的保險。至於社會保險，其目的在於社會保障，屬於強制投保，故保險人不得拒絕任何身份合乎者投保。但同樣的，投保人也沒有選擇的權利。

六、政府之責任

社會保險之目的為社會安全，政府通常高度介入，甚至直接辦理保險。但當保險財務發生逆差，而又不能調整費率時，亦常以國家財務撥補虧損。對於商業性保險，政府主要職責則是訂定規章與監督，確保被保險人能獲得公平的理賠。

以上資料摘錄自楊志良教授之著作「甚麼是健康保險」

9



以商業保險補償社會保險之不足

§開發社會保險補償性質的商業保險：

1. 使用調整後社會保險的承保範圍與對象。
2. 使用社會保險之傷病定義。
3. 採取調整後的社會保險發生率為基礎。
4. 採取與社會保險相同的理賠標準。
5. 以商業保險補償社會保險之不足。舉例：健保自負額、病房費差額、自費用藥、自費醫療器材、居家療養與照護及收入減損等等。

注意：我們無法以商業保險來完整取代社會保險。



重大傷病介紹

全民健康保險與重大傷病

本段資料部份摘錄自：

1. 全民健保重大傷病政策評估 【經社法制論叢/第33 期】
2. 調整全民健保重大傷病給付制度之研究 作者：熊秉元【經社法制論叢/第25 期】

11



重大傷病制度的原由(1)

- 全民健康保險屬於我國社會保險的一環。
- 全民健康保險是社會性健康保險，以全民健康的保障及健康照護為手段，其目的在於避免我國人民因病而貧或因貧而不能就醫的風險，促進社會安全。
- 全民健康保險提供保險對象罹病之財務風險保護，但是為了避免保險帶來道德危害（moral hazard）而影響資源使用效率，各國健康保險制度大多同時具備部分負擔之設計，我國亦復如此。
- 在制度上有針對「特定身分」與「特定醫療」免除部分負擔的雙軌制設計。

12



重大傷病制度的原由(2)

部份負擔之免除：（法源：健保法第48條）

1. 特定身分方面：免除部分負擔者包括低收入戶、榮民及三歲以下兒童。此部分負擔之免除主要是由政府編列的預算補助，與健保本身的財務並無直接關係。
2. 特定醫療項目方面：免除範圍則包括「重大傷病」、分娩、預防保健、山地離島與職業病就醫，其中除了職業病就醫是代辦勞保業務外，其餘均為健保直接提供之優惠免除。此部分負擔之免除則是由健保局直接編列預算補助。

13



重大傷病制度的原由(3)

- 現行健保制度針對住院部分負擔亦訂有單次住院與全年住院上限之規定，但此一措施涵蓋全體保險對象，並無特定適用對象或條件。
- 在所有免部分負擔之相關規定中，又以「重大傷病」最為特殊，不僅影響醫療費用最鉅，同時也是超越公勞保時代健康保險之創舉。

14



健保自行負擔之住院費用上限

行政院衛生署公告

中華民國102 年1 月1 日（補登）

衛署健保字第1012660343 號

主 旨：訂定「一百零二年全民健康保險保險對象應自行負擔之住院費用上」，並自中華民國一百零二年一月一日生效。

依 據：一百零二年一月一日實施之「全民健康保險法」第四十七條第二項。

公告事項：訂定「一百零二年全民健康保險保險對象應自行負擔之住院費用上限」如下：

一、因同一疾病每次住院部分負擔上限：新臺幣三萬一千元。

二、每年住院部分負擔上限：新臺幣五萬二千元。

三、前二項住院部分負擔上限之適用範圍，以保險對象於急性病房住院三十日以下或於慢性病房住院一百八十日以下所應自行負擔之醫療費用為限，不包含全民健康保險法所規定不予給付之項目。

15



健保門診自行負擔之費用

醫院層級別	西醫門診基本部分負擔		急診	牙醫門診	中醫門診
	轉診	逕赴醫院			
醫學中心	二一〇元	三六〇元	四五〇元	五〇元	五〇元
區域醫院	一四〇元	二四〇元	三〇〇元	五〇元	五〇元
地區醫院	五〇元	八〇元	一五〇元	五〇元	五〇元
基層醫療單位	五〇元	五〇元	一五〇元	五〇元	五〇元

16



重大傷病制度的原由(4)

比較我國與國外對於重大傷病定義範圍的差異：

☆國外：

1. 是指需要長時間醫療照護，且醫療費用非常高之中老年慢性病患者。
2. 採取的方案是個案申請，或者訂定一定之醫療費用金額，超過時就給予補助。

☆我國：

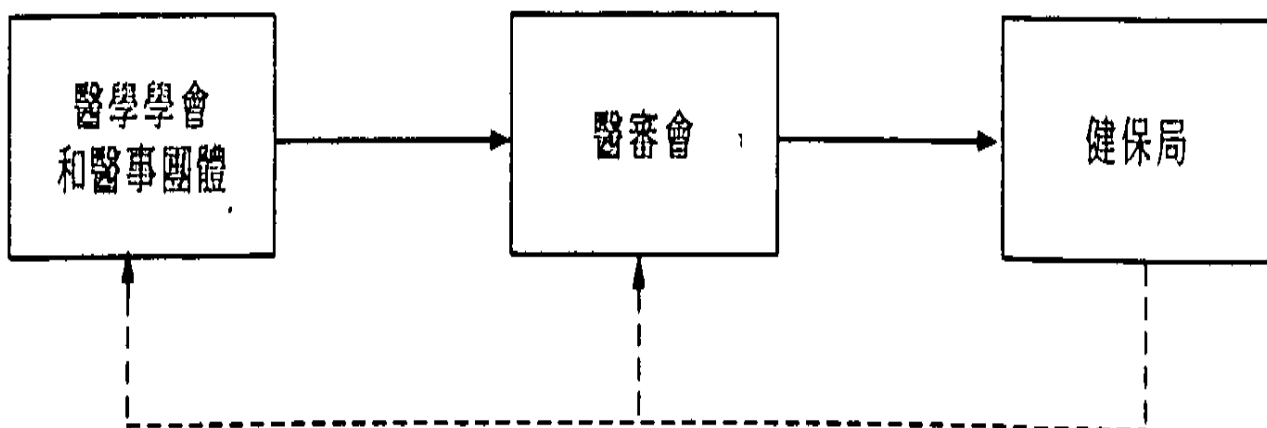
1. 採合議制界定「重大傷病」適用的範圍(如圖五)。
2. 以發給「重大傷病卡」為資格核定之標準，同時採取其他配合之措施(如圖六)。

17



初期重大傷病類別之訂定

圖五 界定重大傷病類別流程

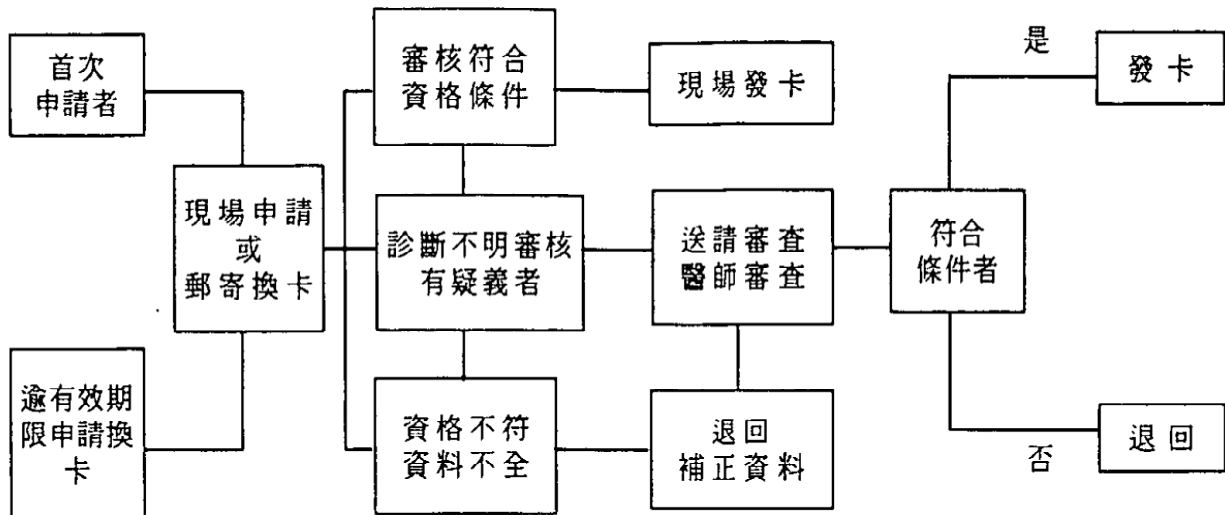


18



初期重大傷病發證流程

圖六 重大傷病證明卡申請作業流程圖



資料來源：「全民健康保險重大傷病制度研討會」會議資料專輯，1998, p. 70。

19



重大傷病制度的原由(5)

- 我國於民國84年3月開辦全民健保。除了將原有公、勞、農保體系下的未保險人口納入健保體制之下，亦同時引進部分負擔制度，以期提高醫療資源之使用效率。
- 為了同時兼顧制度的公平性，全民健康保險法第三十六條(已修訂為四十八條)亦規定：
 1. 保險對象有「重大傷病」情形者，免自行部分負擔。
 2. 至於「重大傷病」之範圍，全民健康保險法並無明文規定，授權由主管機關決定。

20



立法意旨(1)

■ 全民健保採取部分負擔的原因：

1. 部分負擔類似使用者付費（user fee），透過向就醫的消費者直接收費的方式，成為籌措醫療保健財源的管道之一。
2. 部分負擔可提高醫療資源使用效率，避免資源浪費。



立法意旨(2)

■ 重大傷病政策的立法，就是為了減緩部分負擔制度所面臨的「公平」與「效率」之衝突。

1. 以「重大傷病」作為免部分負擔的標準，即是著眼於重大傷病可具體衡量醫療需要的角度，因此可減輕部分負擔制度對醫療服務提供公平性的不利影響。
2. 在健康與所得（或財富）存在有強烈正相關的普遍趨勢下，對重大傷病者免部分負擔之做法，亦有助於紓解部分負擔制度對醫療保健財務籌集公平性的不利影響。
3. 對於已具「重大傷病」身份者，仍有很大部分的醫療支出是不可預測的。這種的不確定性及其所衍生的醫療支出之不確定性，正是健康(社會)保險制度存在的理由。



實施方式(1)

- 全民健康保險主管機關為行政院衛生署，對於「重大傷病」之認定所採取的操作原則，係必須同時符合以下三項條件：一、治療時間長；二、總醫療費用昂貴；三、無濫用之虞(戴桂英，1998)。
- 根據此一操作原則，衛生署在全民健保開辦時，根據開辦前籌備會議中各醫學專科醫學會、團體代表之綜合意見，公告十五類重大傷病，主要項目包括癌症、精神病、全身性自體免疫症候群及慢性腎衰竭等疾病。
- 截至100 年 6 月止，全民健保的「重大傷病範圍」共經歷超過十一次的修訂。

23



實施方式(2)

- 全民健康保險重大傷病範圍之修訂，主要有二個重點：
 1. 增加重大傷病的疾病類別，從開辦初期的十五類傷病，增加到三十項。
 2. 擴大原有重大傷病疾病類別涵蓋的細項。
- 隨著近年來重大傷病類別之增加與涵蓋範圍之擴大、醫療技術之進步與人口老化，符合重大傷病資格的人數呈現逐年增加的趨勢。截至2013年2月，累積領證數為959,229張，歸戶後實際有效領證人數為898,863人。

24



臺灣重大傷病現況介紹



重大傷病制度下之受益者

- 現行重大傷病制度係採疾病別認定，因此一個人可能同時符合二類以上重大傷病，並領有多張證明卡。
- 急性腦血管疾病急性發作一個月內及早產兒所引起之神經、肌肉、骨骼、心臟、肺臟等併發症由醫師逕行認定，則免申請證明。
- 重大傷病資格於全民健康保險制度下的好處：
 - ①可免除就醫時的自負醫療費用。
 - ②可針對特定傷病使用特殊用藥。

法源之依據為：「全民健康保險保險對象免自行負擔費用辦法」



重大傷病人數排序

- 重大傷病就醫人數最多的項目，是衛生署公告的第一大類重大傷病癌症，統計至2011年12月，共有432,344張，約佔全部重大傷病就醫人數的47%，同時佔全國被保險人口約1.8%。
- 次高人數的項目是第六大類重大傷病慢性精神病，有208,432張，約佔所有重大傷病就醫人數的22%。

分析顯示：衛生署公告的重大傷病項目雖多達30大類，但各分類項目人數的分布呈現高度集中的情形。人數最多的前五大類，即佔全部重大傷病就醫人數的九成，而前十大類的佔率則超過九成六。

27



2011年12月底領證數前十大疾病別表

排序	疾病別	領證數	占率	成長率
1	需積極或長期治療之癌症	432,344	47.1%	8.8%
2	慢性精神病	208,432	22.7%	0.4%
3	需終身治療之全身性自體免疫症候群	79,672	8.7%	7.2%
4	慢性腎衰竭(尿毒症)必須定期透析治療者	68,536	7.5%	4.0%
5	心、肺、胃腸、腎臟、神經、骨骼系統等之先天性畸型及染色體異常	35,878	3.9%	1.2%
6	先天性新陳代謝疾病	13,381	1.5%	5.9%
7	小兒麻痺、腦性麻痺、早產兒所引起之神經、肌肉、骨骼、肺臟等之併發症者(其殘障等級在中度以上者)	12,726	1.4%	1.4%
8	因呼吸衰竭需長期使用呼吸器者	11,573	1.3%	-0.7%
9	肝硬化症	10,869	1.2%	-1.2%
10	接受心臟、腎臟及骨髓移植後之追蹤治療	9,349	1.0%	10.4%
	其他	35,960	3.9%	4.6%
	合 計	918,720	100%	5.4%

註：占率為該疾病占合計比率，成長率為較前一年同期成長

資料來源：行政院衛生署網站

28



2012年12月底領證數前十大疾病別表

排序	疾病別	領證數	占率	成長率
1	需積極或長期治療之癌症	463,703	48.2%	7.3%
2	慢性精神病	208,325	21.7%	-0.0%
3	需終身治療之全身性自體免疫症候群	85,473	8.9%	7.3%
4	慢性腎衰竭(尿毒症)必須定期透析治療者	71,468	7.4%	4.3%
5	心、肺、胃腸、腎臟、神經、骨骼系統等之先天性畸型及染色體異常	36,375	3.8%	1.4%
6	先天性新陳代謝疾病	14,054	1.5%	5.0%
7	小兒麻痺、腦性麻痺、早產兒所引起之神經、肌肉、骨骼、肺臟等之併發症者(其殘障等級在中度以上者)	12,857	1.3%	1.0%
8	因呼吸衰竭需長期使用呼吸器者	11,969	1.2%	3.4%
9	肝硬化症	10,673	1.1%	-1.8%
10	接受心臟、腎臟及骨髓移植後之追蹤治療	10,043	1.0%	7.4%
	其他	36,325	3.8%	1.0%
	合 計	961,265	100%	4.6%

註：占率為該疾病占合計比率，成長率為較前一年同期成長

資料來源：行政院衛生署網站

29



總領證數增加分析

- 醫療技術之進步，提早發現某些重大傷病，造成死亡率下降，續發卡人數增加，如癌症。
- 醫療技術之進步，造成某些重大傷病，死亡率下降，累積人數增加，如重大器官移植、長期使用呼吸器、嚴重營養不良者等等。
- 總領證數代表傷病之盛行率，與發生率有某些相關性，但盛行率 $\neq$ 發生率。



各項重大傷病總醫療費用

- 2009年健保重大傷病項目所支付之醫療點數共1378億佔全民健保總醫療點數之26.9%。
- 癌症佔全部重大傷病醫療點數的35%，洗腎佔28%，呼吸器佔11%，慢性精神病佔10%及急性腦血管疾病佔4%。
- 前五大重大傷病總醫療費用的總和，即超過全部重大傷病費用的八成八。前十大的佔率則超過九成五。
- 顯示總醫療費用的集中程度與就醫人數集中程度相同。
- 與人數(卡數)分布情形一樣，總醫療費用的分布情形也是呈現高度集中的現象
- 2011年健保重大傷病醫療費用為1,504億點，約占全國總醫療費用27.2%。

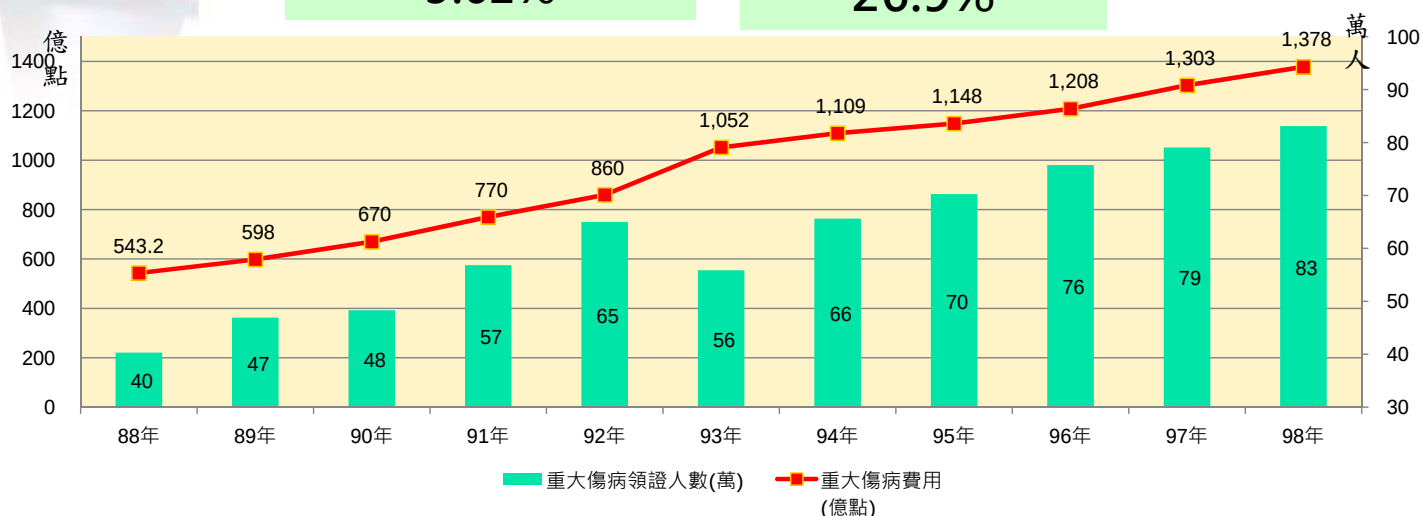
31



健保醫療費用成長2009年-重大傷病

2009年領證數占率
3.62%

醫療費用占率
26.9%



- 健保開辦的時重大傷病人數占總人口的1.5%，現已達3.6%，醫療費用達26.9%。
- 重大傷病者健康的保障，所需醫療花費非一般人所能負擔，例如：癌症病人每人年的醫療費用大約是全人口平均的6.4倍、洗腎為29.4倍、呼吸器依賴者為37.6倍、血友病患者為94.6倍。

資料來源：醫師公會

32



重大傷病平均醫療費用較高者

>觀察2011年的資料：

- 1.每人醫療費用最高的項目是先天性凝血因子異常(血友病)，平均每人一年的花費點數為2,965,147點，九成是發生在門診，是重大傷病中門診費用最昂貴者。
- 2.第二高的是因呼吸衰竭需長期使用呼吸器者，平均每人一年的醫療費用點數高達717,995點，其中絕大部分是花在住院上，是重大傷病中每人住院費用花費最多者。
- 3.每人費用排名第三到第五名的項目，分別是洗腎、腸道切除給予全靜脈營養及燒燙傷症，平均每人費用點數也都在約 40 萬點以上。

33



2011年重大傷病每人平均費用

2011年重大傷病每人平均費用前五大疾病（門住診合計）

排序	病名	就醫人數 (人)	點數 (億)	全年每人平均 醫療費用(點)	使用多少人保 險費支應 (人)
1	血友病	970	28.76	2,965,147	198
2	呼吸器依賴	22,115	158.79	717,995	48
3	洗腎	71,820	4361.67	601,051	40
4	腸道切除給予 全靜脈營養	80	0.43	536,602	36
5	燒傷	463	1.85	399,954	27

註：

平均就醫點數：醫療點數/就醫人數

約需多少人保險費支應(人)：全年每人平均醫療費用/全年平均每人應收保費
(2011年約為14,988元)



重大傷病每人平均費用

2010年健保重大傷病醫療資源利用情形
重大傷病人口占3.4%，使用醫療資源26.99%

類別	醫療費用(點)	平均值倍數
全國每人平均	22,798	1
每一血友病患者	2,729,027	119.7
每一呼吸器患者	714,662	31.3
每一洗腎患者	601,109	26.4
每一癌症患者	136,578	6.0

35



重大傷病每人平均費用

2011年健保重大傷病醫療資源利用情形
重大傷病人口占3.56%，使用醫療資源27.15%

類別	醫療費用(點)	平均值倍數
全國每人平均	23,950	1
每一血友病患者	2,965,147	123.8
每一呼吸器患者	717,995	30.0
每一洗腎患者	601,051	25.1
每一癌症患者	137,071	5.7

36



結論(1)

- 以重大傷病人數的分布來看，發現分配呈現高度集中的情形：人數最多的前十大重大傷病類別，佔所有重大傷病人數的九成六。
- 截至100年12月止重大傷病實際有效領證數計918,720張，較99年增加47,263張，成長5.4%，其中癌症領證數最高432,344張，佔總領證數之47%，次之為慢性精神病208,432張，佔22.7%。歸戶後之領證人數861,826人，約佔保險人口之3.7%，較99年增加41,321人，成長5%。

37



結論(2)

- 100年重大傷病項目中癌症領證數最多，占率為47%，其醫療費用高占36.8%；以100年全年增加醫療費用77億點，其中49.2%（37.9億點）用於癌症治療。
- 觀察全民健保醫療費用逐年成長之原因，重大傷病醫療費用支出之成長為主要之原因之一。
- 全民健保之「重大傷病」≠現行健康險市場之「重大傷病」。

38



臺灣健保資料庫介紹



全民健康保險研究資料庫

National Health Insurance Research Database



NEW

100年全年資料發行

ENGLISH

☐ 非學術界研究類

☐ 學術界研究類

非學術界研究類

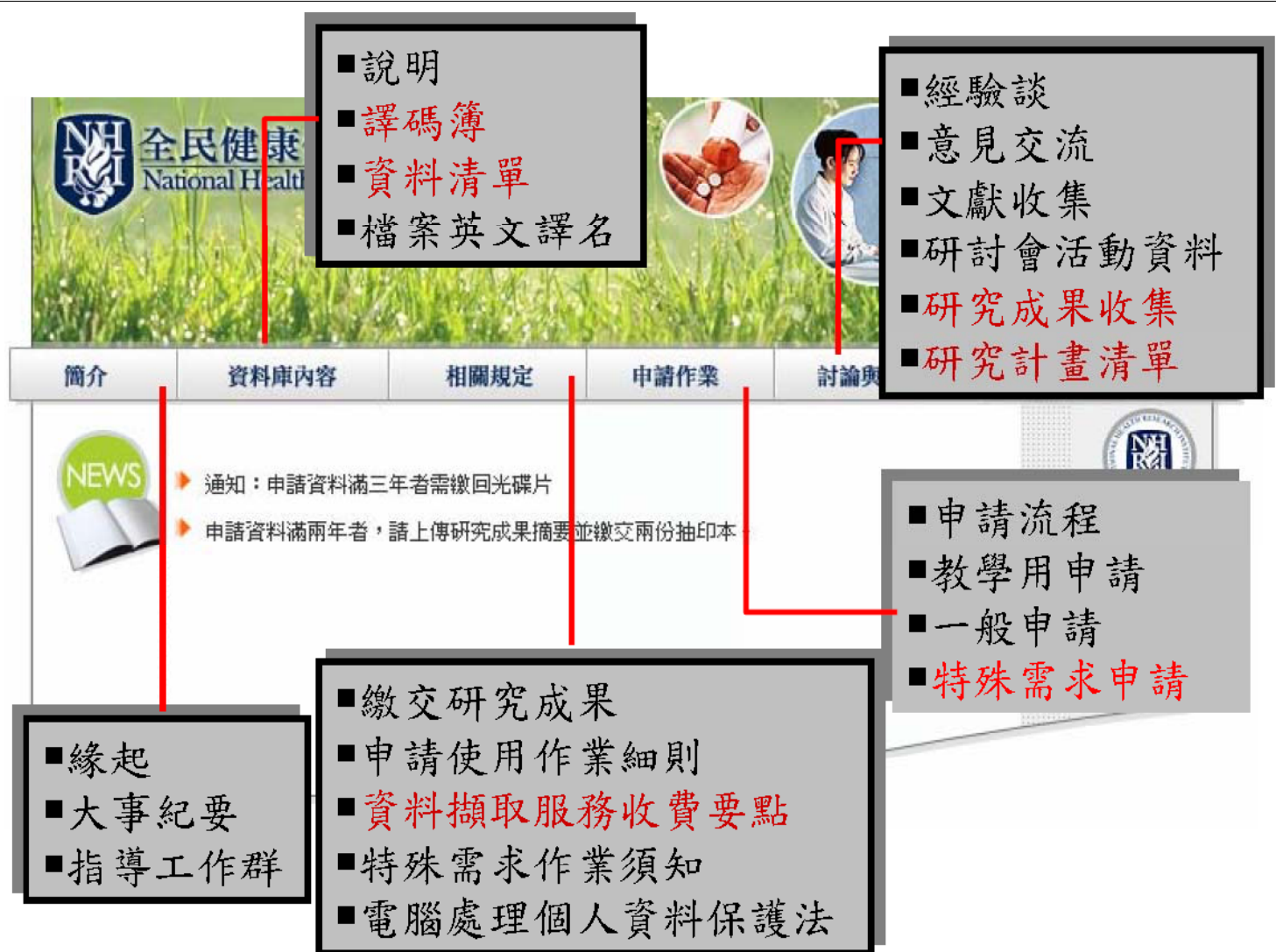
- 1、申請人必須為我國國民或本國登記立案之公司或機構。
- 2、不符合學術研究類資格或具學術研究類資格但未通過學術研究類之審查程序者。

National Health Insurance Research Database

最佳瀏覽效果800x600

Copyright 2003 National Health Research Institutes. All rights reserved.

健保資料庫首頁 <http://nhird.nhri.org.tw/>



全民健康保險研究資料庫 (簡稱「健保資料庫」)

- 由中央健康保險局負責管理
- 民國87年健保局委託國衛院建置健保資料庫
→ 民國89年起開放資料庫給各界單位
- 申請類別：
 - 學術單位：學術機構（團體及個人）都可申請。
 - 非學術單位：國內登記公司或機構，但僅提供部份資料檔案。



關於全民健保資料庫

- 健保局委託國家衛生研究院，維護全民健保資料庫；學術團體可付費申請資料庫，經過加密處理，避免個人資料外洩。
- 申請資料時須附上研究計畫書，從2012年起計畫主持人、協同主持人也得檢附「人體試驗及研究倫理講習」(IRB)的證明。
- 除了HV等特定主題的資料外，近幾年公布的資料有2005年百萬人抽樣檔、2010年百萬人抽樣檔。

43



健保資料庫基本敘述

- 資料時間：自健保開辦(1995)至今都有資料，但考量資料的可用性、完整度(例如：2000年前缺某些資料欄位、資料品質較不佳)，建議使用近年(如：2001年後)的資料。
- 健保資料庫由許多資料(或檔案)組成，其中有完整**母體檔**，也有**抽樣檔**。
- 母體較小(如：重大傷病)可提供完整母體檔，但較大母體(如：全民)則以抽樣檔為原則。

44



較具代表性的抽樣檔

- 抽樣檔(或「抽樣歸人檔」)目前有三個時段：
 - 2000年 (LHID2000, 2002年發行20萬人、2009年發行80萬人)
 - 2005年 (LHID2005, 2007年發行百萬人)
 - 2010年 (LHID2010, 2011年發行百萬人)
- 除重新抽樣外, 更擴大抽樣人數為100萬人, 擷取其所有就醫申報資料製作歸人檔。100萬承保抽樣歸人檔可提升具有統計檢定力(statistical power)之前瞻性研究(Prospective study)及回溯性研究(Retrospective study)之可能性。

45



LHID2000抽樣方法

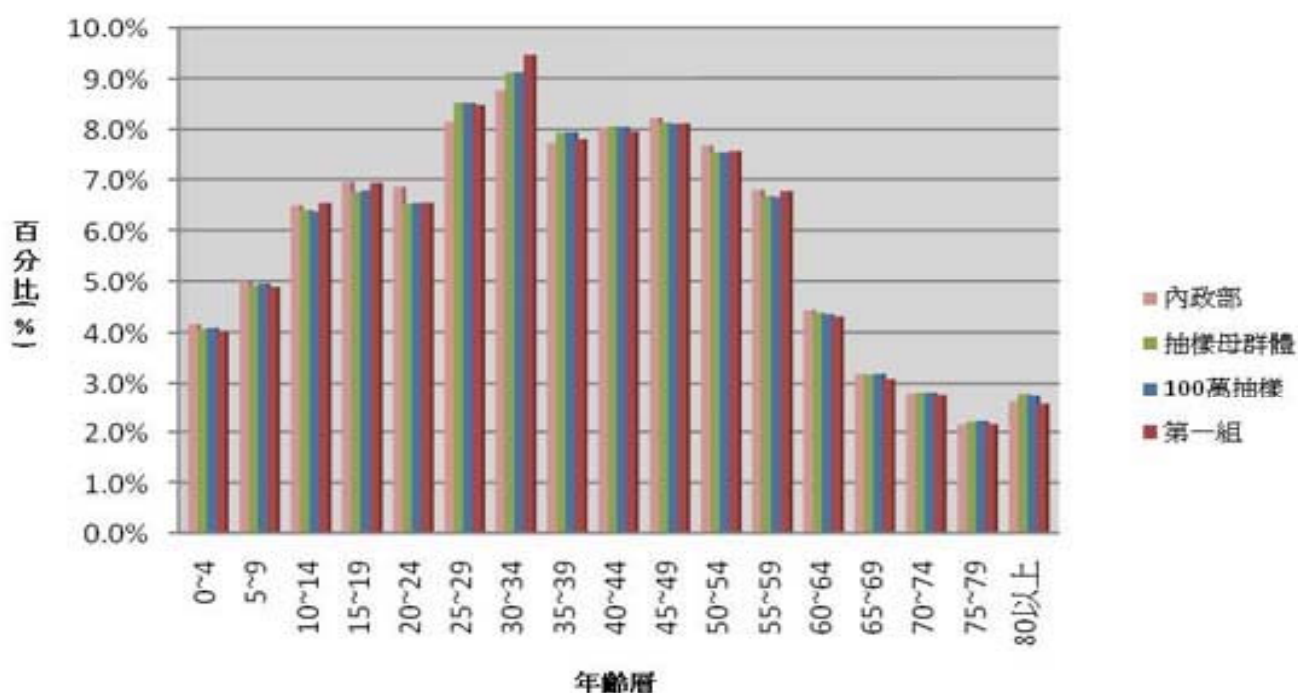
- 隨機抽樣方法為將抽樣母群體23,251,700人賦予流水號, 利用隨機值產生器產生至少100萬個隨機值(實得1,074,263個隨機值), 取與100萬個隨機值相同的流水號, 隨機抽取所需的保險對象樣本, 再剔除身份證字號重複者(共24個), 再補抽至得到100萬人樣本為止。
- 隨機值產生係採Oracle的DBMS_RANDOM套件, 本身有內建隨機值產生器, 可產生八位整數隨機值。1與23,251,700之間產生110萬個隨機值, 剔除重複出現的隨機值(共25,346個)後, 共得到1,074,263筆隨機值。

46



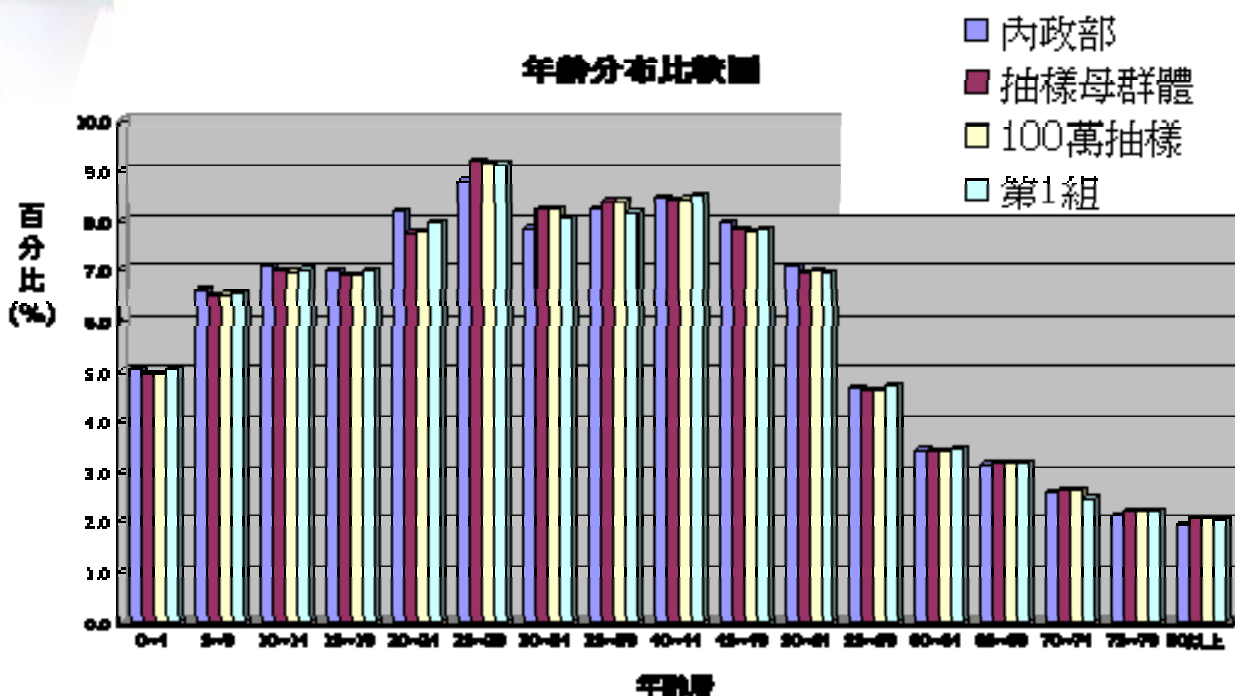
2010年抽樣檔年齡分佈

年齡分布比較圖



2005年抽樣檔年齡分佈

年齡分布比較圖





抽樣檔的基本特性

■抽樣檔是長期追蹤資料，使用上必須注意：

→優點:可以追蹤某人所有就醫行為；

→缺點:存有樣本偏差！

(例如：抽樣年度之前的死亡、移出人口無法被抽到，抽樣年度之後的新生兒、移入人口無法被抽到。)

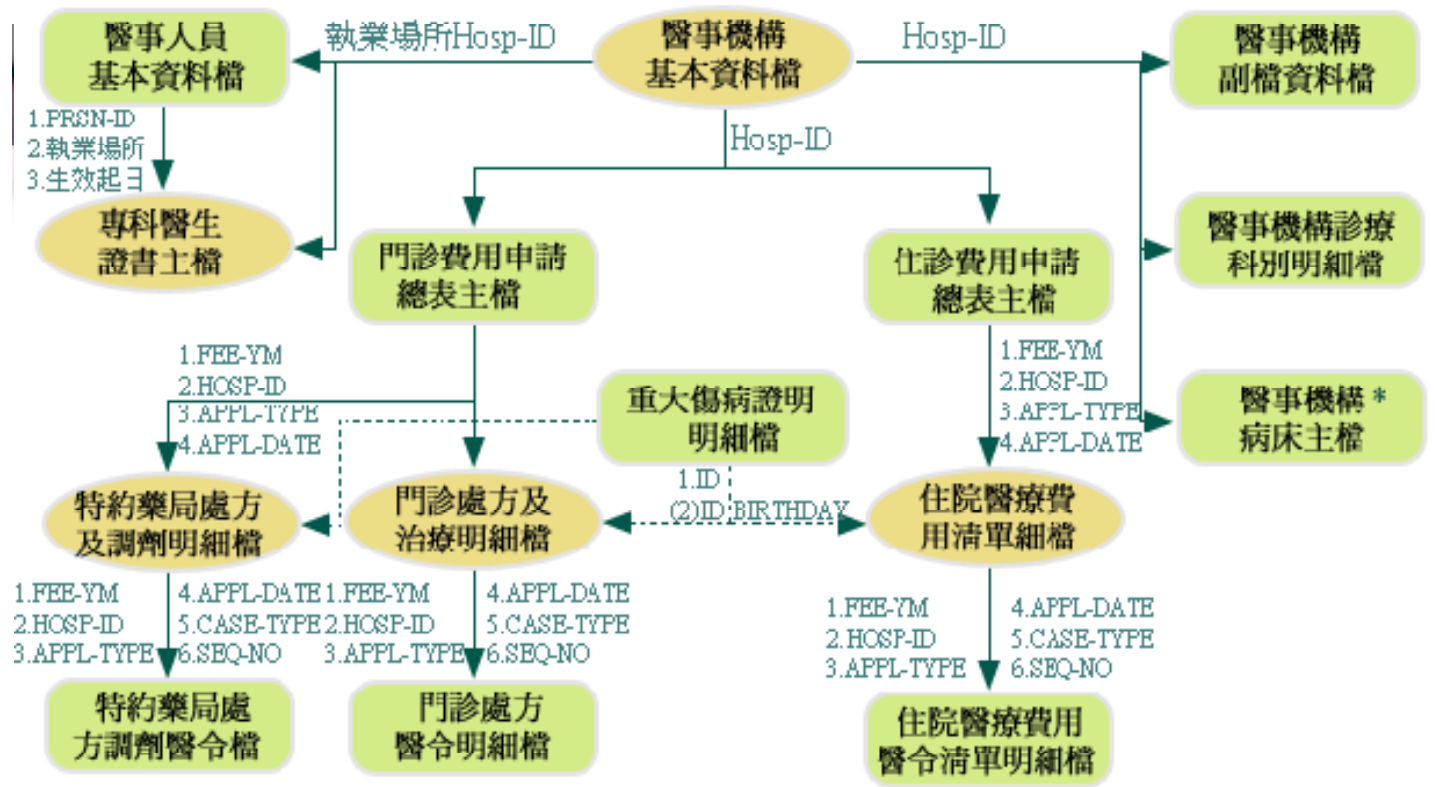
■每年平均資料量約15GB左右，一般統計軟體無法直接分析，需透過資料庫軟體處理，一般使用SQL (Access等軟體也無法處理；健保局另外提供SAS資料採礦軟體)。

49

各年度資料清單可以分為四大類

- 基本資料檔：10類
- 系統抽樣檔：住院醫療費用清單明細檔與門診處方及治療明細檔（含醫令資料）
- 特定主題分檔：14類
- 承保抽樣歸人檔：4組（每組5萬人，共計20萬人）

各檔案間串檔變項說明



註:*須注意生效起訖日期

(2)可由ID+BIRTHDAY串檔

➡ 各檔案間由所註明變項串檔可獲得對應資訊

➡ 各檔案間可由所註明變項串檔,但未必獲得對應資料

系統抽樣檔

- 抽樣之目的以研究者能在PC上使用該資料為原則，儘量完整呈現資料之原始架構，讓使用者對資料進行初步之測試，並未以任何研究目的作為抽樣之考量。對於特定目的之研究，應依研究特性另外選擇適當之抽樣方法，本抽樣方法並不適用於所有的研究目的。



抽樣檔的幾個範例

■ 住院醫療費用清單明細檔(DD)及住院醫療費用醫令清單明細檔(DO)的抽樣方法，大致為母體總數的5%（亦即每20筆抽取1筆），以系統抽樣法按月抽取所需樣本數。

■ 門診處方及治療明細檔(CD)及門診處方醫令明細檔(OO)，抽樣方法與上述類似，依健保局提供之原始檔中西、牙、中醫檔案的分類，以系統抽樣、按月比例抽出，大約母體總數之0.2%（或是每500筆抽取1筆）。

53

※資料清單

92年度資料清單

基本資料檔

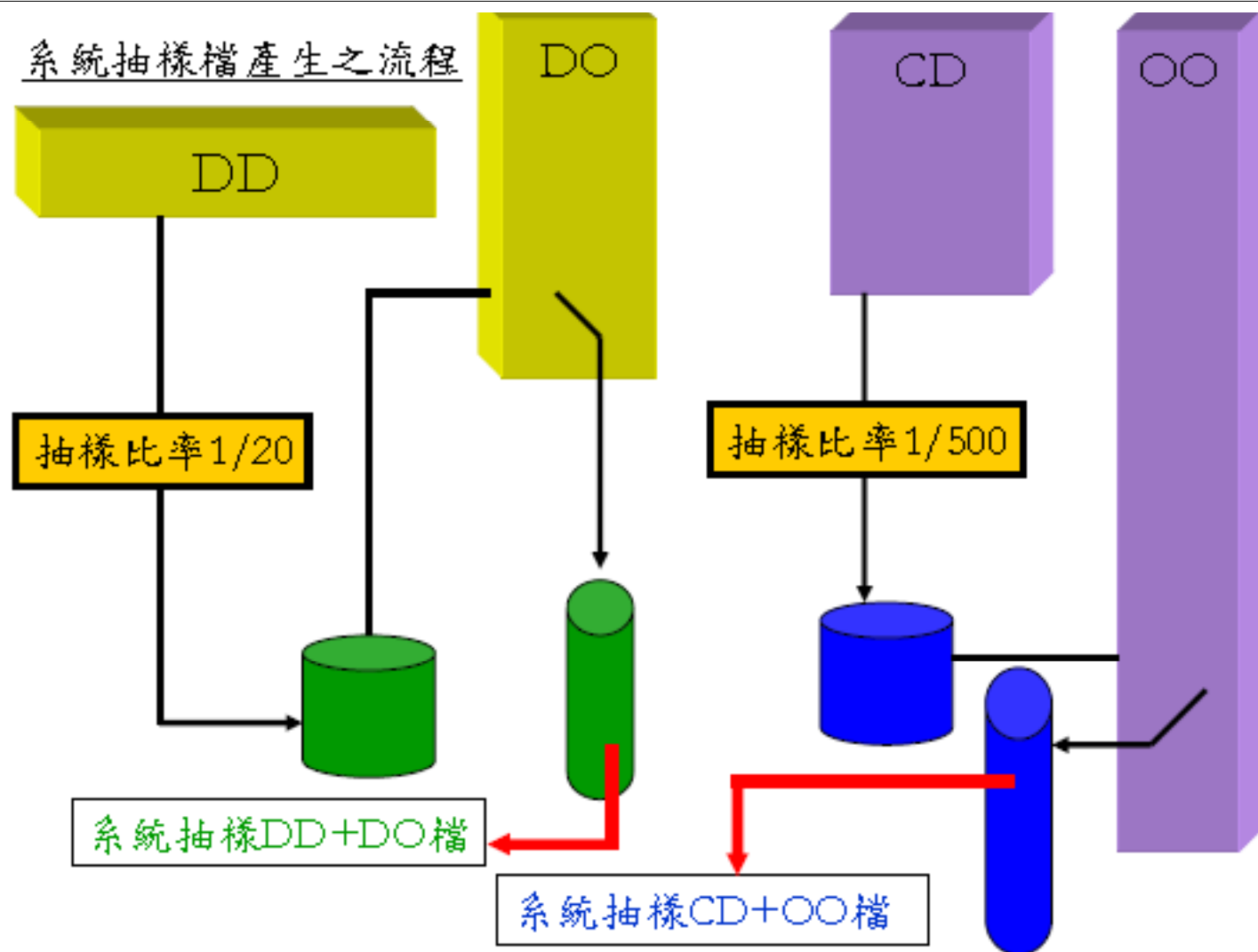
系統抽樣檔

特定主題

■ 系統抽樣檔（一套2片）

光碟編號	中文檔名	英文檔名	資料筆數	檔案大小 (MB)
S9201	門診處方及治療明細檔（西醫）	S_CD20030	516,546	103.45
	門診處方及治療明細檔（牙醫）	S_CD20031	50,638	10.14
	門診處方及治療明細檔（中醫）	S_CD20032	62,516	12.52
	門診處方醫令明細檔（西醫）	S_OO20030	2,361,591	216.21
	門診處方醫令明細檔（牙醫）	S_OO20031	85,837	7.86
	門診處方醫令明細檔（中醫）	S_OO20032	303,134	27.75
S9202	住院醫療費用清單明細檔	S_DD2003	136,699	55.41
	住院醫療費用醫令清單明細檔	S_DO2003	7,731,522	589.87

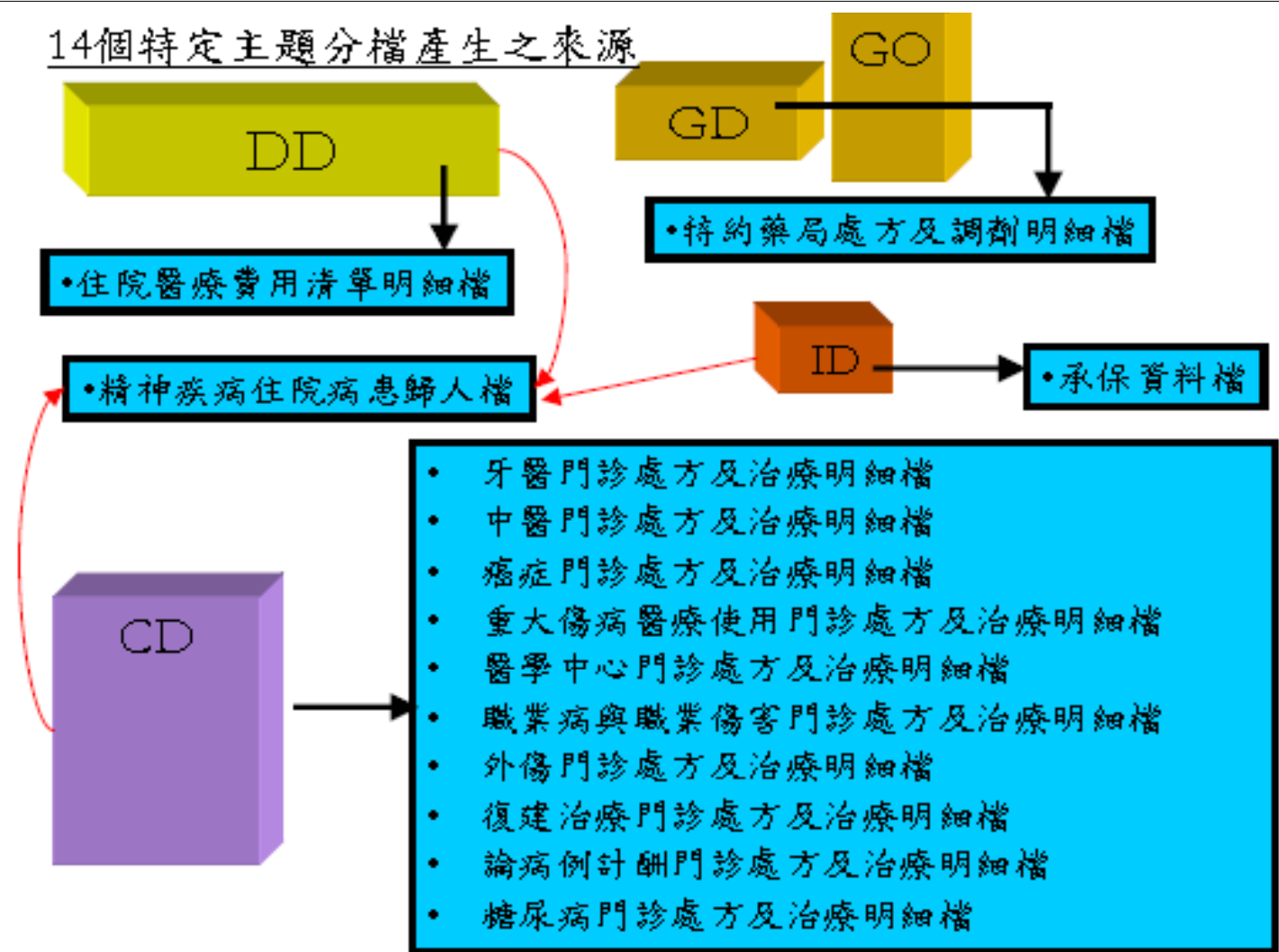
系統抽樣檔產生之流程



特定主題分檔（14類）

- 特約藥局處方及調劑明細檔
- 牙醫門診處方及治療明細檔
- 中醫門診處方及治療明細檔
- 癌症門診處方及治療明細檔
- 重大傷病醫療使用門診處方及治療明細檔
- 醫學中心門診處方及治療明細檔
- 職業病與職業傷害門診處方及治療明細檔
- 外傷門診處方及治療明細檔
- 復建治療門診處方及治療明細檔
- 論病例計酬門診處方及治療明細檔
- 糖尿病門診處方及治療明細檔
- 住院醫療費用清單明細檔
- 承保檔
- 精神疾病住院病患歸人檔

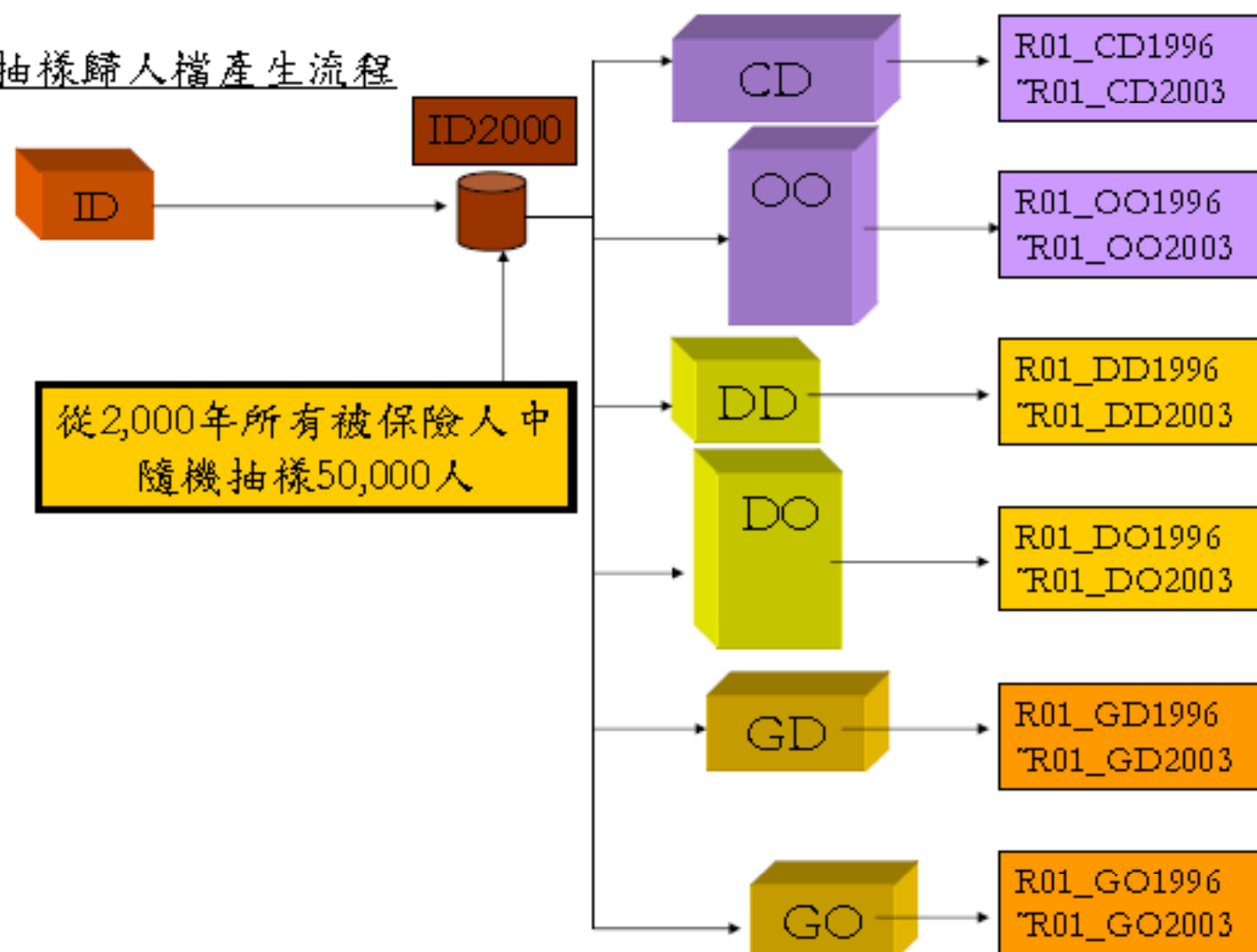
14個特定主題分檔產生之來源



承保抽樣歸人檔

- 以健保承保資料檔為抽樣母群體，依簡單隨機抽樣，決定樣本（人），依此擷取各該人各年所有資料(包括門診、住院及特約藥局)，每組共抽出五萬人，目前共開放四組供研究者申請。
- 抽出樣本之年齡、性別及費用三項變項分佈與母體間並無明顯差異。

抽樣歸人檔產生流程



資料清單

承保抽樣歸人檔

承保抽樣歸人檔組別：

■ 承保抽樣歸人檔 (第一組)

光碟編號	中文檔名	英文檔名	資料筆數	檔案大小 (KB)
R01_1996	承保資料	ID2000	123,093	7.40
	門診處方及治療明細檔	R01_CD1996	560,452	112.24
	住院醫療費用清單明細檔	R01_DD1996	4,861	1.97
	特約藥局處方及調劑明細檔	R01_GD1996	494	0.06
R01_1997	門診處方及治療明細檔	R01_CD1997	608,625	121.89
	門診處方醫令明細檔	R01_OO1997	2,161,335	197.88
	住院醫療費用清單明細檔	R01_DD1997	5,018	2.03
	住院醫療費用醫令清單明細檔	R01_DO1997	264,614	20.19
	特約藥局處方及調劑明細檔	R01_GD1997	4,323	0.54
	特約藥局處方調劑醫令檔	R01_GO1997	16,952	1.23

譯碼簿範例

重大傷病證明明細檔							HV
英文檔案名： 中文檔案名：		HV		長度		245	
序號	英文欄位	中文欄位	資料類型	長度	起始位置	結束位置	資料描述
1	ID	身分證號	C	10	1	10	已轉碼
2	DISE_CODE	診斷代碼	C	5	11	15	請參考ICD-9與A-Code
3	HV_TYPE	重大傷病類別代碼	C	2	16	17	請參考健保局網頁 http://www.nhi.gov.tw/03hospital/hospital_5_01.htm
4	ID_BIRTHDAY	出生日期	C	8	18	25	YYYYMMDD
5	ID_SEX	性別	C	1	26	26	M：男 F：女
6	APPL_DATE	申請日期	C	8	27	34	YYYYMMDD
7	APPL_TYPE	申請類別	C	1	35	35	1：個人申請 2：醫院推薦
8	HOSP_ID	醫事機構代號	C	10	36	45	已轉碼
9	PRSN_ID	醫師代號	C	10	46	55	已轉碼
10	ACPT_DATE	受理日期	C	8	56	63	YYYYMMDD
11	ACPT_NO_YYY	受理編號_年度	C	3	64	66	YYY (民國年)
12	ACPT_NO_B	受理編號_分局	C	1	67	67	0：總局 1：台北分局 2：高雄分局 3：中區分局 4：南區分局 5：基隆分局 6：東區分局
13	ACPT_NO_SEQ	受理編號_流水號	N	7	68	74	
14	DISE_DESC	診斷病名	C	150	75	224	
16	STOP_REASON	退保原因	C	1	225	225	5：停保 B：開辦前退保 D：失聯 E：失聯滿六個月 F：出國 I：監所受刑處分 M：死亡 N：不在保 X：查保 O：棄、轉保人員 U：不具本法第十條資格者
17	STOP_DATE	退保日期	C	8	226	233	YYYYMMDD
18	REPRINT	複印份數次數	N	2	234	235	
19	DUPRINT	複印張數	C	1	236	236	Y：重覆 N：未重覆
20	VALID_E_DATE	有效迄日	C	8	237	244	YYYYMMDD
				244			

一般申請

- 凡學術界研究類之申請案，國衛院於提供該資料檔之同時，得對外公告研究計畫主題及摘要。
- 申請對象為現有之四類檔案內容。
- 申請資料滿兩年者，請上傳研究成果摘要並繳交兩份抽印本。
- 本資料庫申請使用者自取得國衛院所提供資料之日起至屆滿三年之日前，應繳回該資料予國衛院。

擷取資料之收費

學術研究類

• **申請光碟片**：包括基本資料庫、系統抽樣檔及特定主題分檔，每片光碟片以**新台幣500元**計收。

• **特殊需求**：以讀取量為基準，**讀取1GB**的資料收費**新台幣200元**。（請參考表格內容）

- 申請案經審查同意提供研究使用者，在確認申請資料內容後，依規定核定應收取費用，申請人先繳費後取件。
- 確認申請人繳交全部費用後，所有資料一律以掛號方式寄出。
- 本資料庫申請使用者應遵循「電腦處理個人資料保護法」及相關法令之規定，不得侵犯個人隱私或藉以研擬個別單位，亦不得將資料作為營利用途。
- 本費點經院長核定，並送中央健康保險局核備後公告實施，修訂時亦同。

	門診明細檔(CD)	門診醫令檔(DD)	住院明細檔(DD)	住院醫令檔(DD)
85年	10,600元(53GB)		200元(1GB)	
86年	11,400元(57GB)	18,400元(92GB)	200元(1GB)	2,000元(10GB)
87年	12,200元(61GB)	23,000元(115GB)	200元(1GB)	2,000元(10GB)
88年	12,600元(63GB)	28,200元(141GB)	200元(1GB)	2,000元(10GB)
89年	12,400元(62GB)	26,800元(134GB)	200元(1GB)	2,000元(10GB)
90年	12,200元(61GB)	25,200元(126GB)	200元(1GB)	2,200元(11GB)
91年	12,400元(62GB)	25,000元(125GB)	200元(1GB)	2,400元(12GB)
92年	13,200元(66GB)	25,800元(129GB)	200元(1GB)	3,000元(15GB)

3. 收費標準

• **申請光碟片**：每片光碟片以**新台幣2,500元**計收。

• **特殊需求**：以讀取次數及讀取量為基準，**每次每讀取1GB**的資料收費**新台幣800元**。（請參考下表說明）

非學術研究類



臺灣重大傷病實證研究



台灣地區重大傷病醫療費用推估

論文刊登日：2004/10/22（人口學刊）

作者：黃泓智、劉明昌、余清祥

65



重大傷病發生率研究

- 以全民健康保險研究資料庫為基礎，探討台灣地區重大傷病現況。
- 訂定重大傷病之定義及其範圍
 - 研究項目：重大傷病盛行率、發生率、死亡率及門診醫療費用支出
 - 計算重大傷病之盛行率、發生率、死亡率
 - 計算罹患疾病之後的平均餘命及疾病對整體壽命的影響。

66



重大傷病發生率之定義

- 疾病盛行率：單一時點或一段期間內罹患疾病人數佔總人口數的比例。
- 重大傷病發生率（或稱做罹病率）：特定疾病在單一時點或一段時間內，新增病患人數佔暴露該疾病之下人口數的比例，比例愈大則代表罹患該疾病的機率相對較大。
- 通常也稱為初次罹患發生率或首發率。
- 重大傷病盛行率與發生率有高度相關。

67



台灣地區重大傷病發生率概況

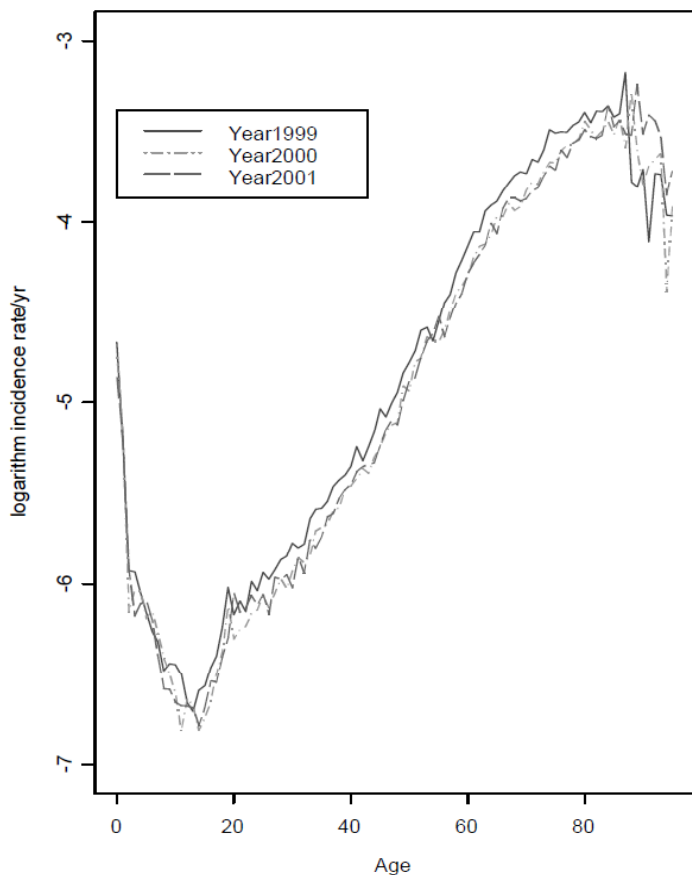
- 盛行於40 歲以上人口，男性重大傷病盛行率普遍高於女性。
- 重大傷病發生率大致隨年齡而上升，符合身體機能隨年齡增加而衰退的認知，也與死亡率的圖形類似。
- 14 歲以下的重大傷病主要為先天性疾病。

68

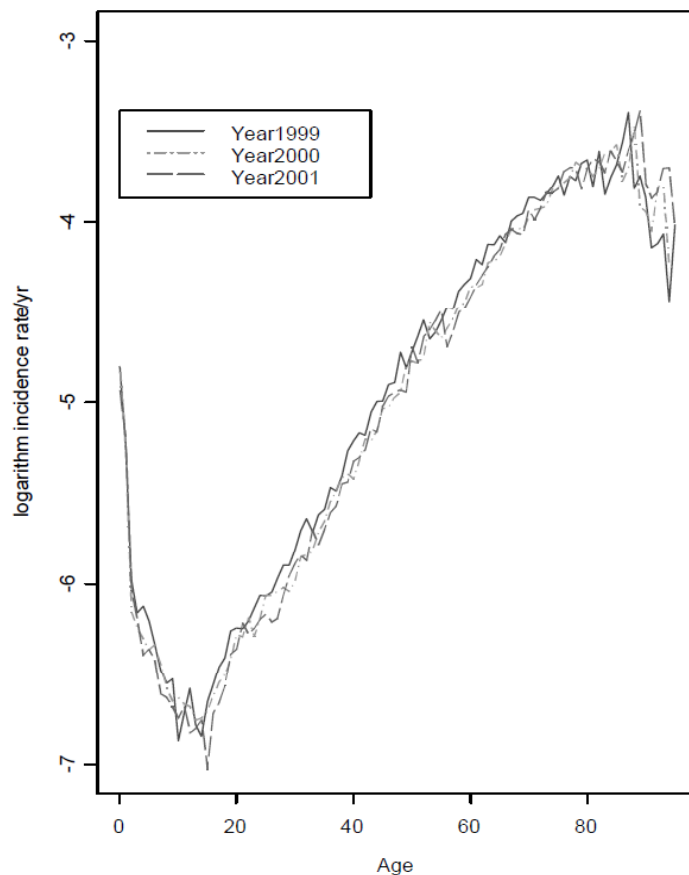


首次罹病率—年齡

1999~2001 log incidence rate (Male)

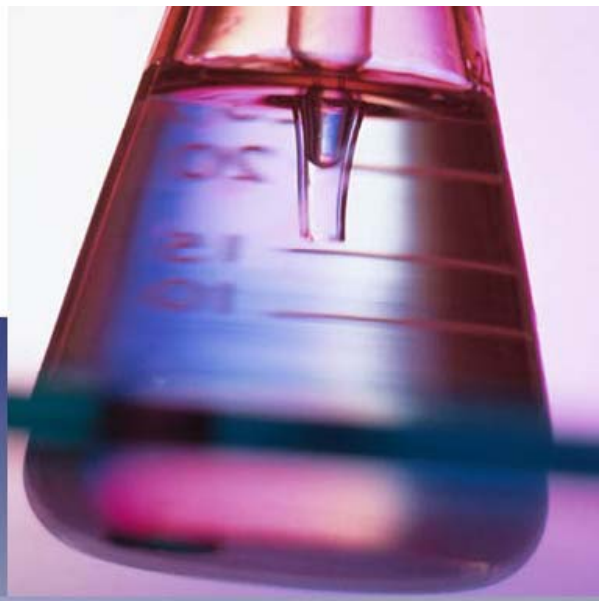


1999~2001 log incidence rate (Female)



實證分析結論

- 本研究先估算台灣居民的重大傷病盛行率、發生率、死亡率以及門診醫療費用現況，再衡量重大傷病對於健康狀況的影響以及推估未來重大傷病門診費用。
- 重大傷病發生率隨年齡提升，盛行率主要集中於40歲以上的人口，男性普遍高於女性。
- 重大傷病門診醫療費用將隨人口老化以及重大傷病死亡率遞減等因素而逐年（直線）上升，建議全民健保費率必須定期檢討、調整。



重大傷病保險 產品開發

71



發生率來源

- 健保局全民健康保險資料庫
- 學術研究機構之研究申請
- 相關論文發表
- 依資料定義發生率
- 由國際再保險公司依經驗調整

72



重大傷病保險之定義

- 「重大傷病」：
本契約所稱「重大傷病」，係指「全民健康保險法」所稱之重大傷病。
- 「重大傷病範圍」：
本契約所稱「重大傷病範圍」，係指中央衛生主管機關公告實施之「全民健康保險保險對象免自行負擔費用辦法」附表『全民健康保險重大傷病範圍』中所載之傷病範圍。其後傷病範圍或項目如有變動，則以中央衛生主管機關最新公告之重大傷病範圍或項目為準。

73



重大傷病保險之定義（續）

- 本契約所稱「首次罹患重大傷病」，係指被保險人自本契約生效日或復效日起，經中央健康保險局(署)特約醫院、特約診所之醫師診斷，確認首次罹患「全民健康保險法」所稱之重大傷病，同時依行政院衛生署所公告實施之「全民健康保險保險對象免自行負擔費用辦法」規定，取得中央健康保險局(署)核定之重大傷病證明者。

74



重大傷病保險產品開發方向

- 重大傷病健康保險
- 重大傷病定期壽險
- 變動給付型重大傷病保險(含死亡給付)
- 還本型重大傷病保險金(含死亡給付)
- 分次給付型重大傷病保險(含死亡給付)
- 特定重大傷病長期照護型健康保險(重症照護保險)

75



謝謝您的傾聽！

Q & A

76