

衛生署與內政部死亡資料連結比對分析

朱慧凡¹ 吳肖琪^{2,*}

HUI-FAN CHU¹, SHIAO-CHI WU^{2,*}

¹ 國立陽明大學公共衛生研究所衛生福利與政策管理組

Division of Health Welfare Policy and Management, Institute of Public Health, National Yang-Ming University, Taipei, Taiwan, R.O.C.

² 國立陽明大學衛生福利研究所，台北市立農街二段155號

Institute of Health and Welfare Policy, National Yang-Ming University, No.155, Li-Long St, Sec 2, Shih-Pai, Taipei, Taiwan, R.O.C.

*通訊作者Correspondence author. E-mail: scwu@ym.edu.tw

目標：以內政部與衛生署死亡資料連結進行比對，探討兩單位死亡資料之差異，以為研究者使用該資料庫與相關單位改善資料庫品質之參考。**方法：**以內政部統計與衛生署統計公布之歷年(民國71-91年)死亡人數統計比較人數差異之趨勢，並連結民國88年內政部死亡檔與衛生署死因檔比較實際差異之情形。**結果：**若以歷年內政部戶政司與衛生署統計室公布死亡人數統計做比較，可發現各年度檔案人數之符合率皆高達98%以上。但若以實際資料逐筆連結比對，則88年死因檔符合率由內政部與衛生署統計死亡人數的98.7%降為95.09%，內政部與衛生署檔案中未連結上之人數分別為6,203人與4,747人。進一步分析發現不同年齡別與地區別之連結率有顯著差異。**結論：**本研究發現內政部與衛生署公佈之死亡人數總數差異約為2%，逐筆連結比對後差異約為4%，此差異會隨死亡年齡增加而減少且有地區性差異。建議未來使用者應小心年齡與地區差異。關於造成此比對差異的原因與政策意涵，有待未來研究進一步釐清。(台灣衛誌 2004；23(6)：447-452)

關鍵詞：內政部、衛生署、死亡登記、資料庫

Comparison of the number of deaths released by the registration data of Department of Health and Ministry of Interior: a linkage study

Objectives: To explore the difference between the datasets of the death registration of the Department of Health (DOH) and Ministry of Interior (MOI) for the reference of the researchers in the future. **Methods:** We used the annual number of deaths published by the MOI and DOH during 1982-2002 to analysis the difference; and linking the datasets of the death registration of MOI and DOH, published in 1999, by ID number to verify the actual difference in the number of cases. **Results:** The concordant rate of number of deaths published by DOH and MOI was 98.7%. When the datasets were linked by ID number, the concordant rate was dropped to from 98.7% to 95.9%. There were 6,203 cases from the MOI dataset and 4,747 cases from the DOH dataset not matched by ID number. The rate of mismatch varied by age and area. **Conclusions:** The difference in the number of death published by DOH and MOI was about 2%. When the datasets were linked by each ID number, the difference in the number of death was increased to 4%. The concordant rate between the datasets of DOH and MOI varied by age and area. When using the dataset of death registration from the DOH or MOI for study, researchers should be aware of the variance. The causes of the difference and policy implication of this should be investigated in further study. (*Taiwan J Public Health*. 2004;23(6):447-452)

Key Words: Ministry of Interior, Department of Health, death registration, dataset

前言

隨著資訊科技的進步，資料庫不但可提供管理者監控體系與制定政策的重要資訊，更是研究者重要的研究資源。

死亡資料庫是最常被應用的健康資料庫，資料庫的完整性與正確性對於研究結果有重大的影響。目前國內死亡診斷書分別經由內政部戶政司與衛生署建檔。內政部死亡資料來源為戶政事務所之戶籍登記資料彙整之電子檔，由各級戶政單位逐層通報內政部戶政司，審核無誤後轉送內政部統計處，再轉入內政部公務統計資料庫[1]，檔案擷取的對象則包括檔案年度至次年三月申報死亡資料者，檔案中會包含死亡日期為前兩年死亡者或次年一至三月者。

衛生署死亡資料來源為每月由衛生所人員向戶政單位影印死亡證明書，經縣市衛生局蒐集彙整進行基本審查與註記後，再送交衛生署中部辦公室統計科進行死因註碼與人工複審後鍵入死因管理系統中，最後轉入衛生署統計室[2]，衛生署死亡檔擷取對象為在檔案年度中死亡之個案。

基於研究之目的與檔案取得之方便性，衛生署死亡資料庫是衛生領域相關研究者較常使用之資料庫。然過去國內對死亡資料庫的分析主要著重於死因譯碼的信度與效度[4-8]。謝功毅等[3]雖有利用內政部死亡資料庫對衛生署死亡資料進行確認與補正，然該研究乃利用民國69-86年之死亡資料進行分析，無法看出各年度之差異情況。且目前國內在「個人資料保護法」規定下，多數資料檔之身分證字號皆經相關單位加密，以健保資料為例，若要連結死亡資料，則死亡檔之身分證字號也必須加密，如此則無法進行身分證字號檢測以進行補正。另外，內政部自民國87年開始由原本委託財稅資料中心建檔轉為自行建檔，資料庫之建檔方式與品質可能與之前委外時有所差異。故本研究將以內政部與衛生署死亡資料連結進行比對，以為研究者使用該資料庫與相關單位改善資料庫品

投稿日期：92年7月23日

接受日期：93年5月24日

質之參考。

材料與方法

本研究資料來源為民國71-91年內政統計年報與衛生統計公布之各年度死亡人數[9-11]、88年內政部死亡資料檔案(僅取得身分證字號變項)與86-89年衛生署死因檔，死亡資料檔案之個案身分證字號皆經健保局加密。

資料之處理與分析分為兩部分，第一部分是內政統計年報與衛生統計公布之各年度死亡人數進行比較，以兩資料差異人數與符合率進行年度趨勢比較，符合率之定義為衛生統計死亡人數佔內政統計死亡人數之百分比。第二部分則是以內政部與衛生署死亡資料，進行連結比對，各檔案分別依加密身分證字號歸戶後，將88年內政部死亡資料依加密身分證字號分別與89年、87年、86年衛生署死亡資料連結以排除非88年死亡個案，而未連結上之內政部88年死亡個案再與88年衛生署死因檔連結，以比對內政部與衛生署死亡個案實際之差異狀況。進一步再分析88年衛生署死因檔中未能與內政部死亡檔連結者之基本特質，並利用卡方檢定比較各特質未連結比率之差異。

結果

表一為歷年內政部戶政司與衛生署統計室公布死亡人數統計之比較。結果發現除民國71年人數差異較小(973人)，82年人數差異較大(2020人)，其他各年度人數差異接維持穩定狀態(約一千多人)，各年度符合率皆高達98%以上。

表二為88年內政部死亡資料連結各年度衛生署死因資料結果。88年內政部死亡資料(包括88年1月-89年3月間申報死亡者，含死亡日期為88年以前2年或89年1-3月者)先依同一加密身分證字號歸戶，排除373筆身分證字號重複個案後共有159,144人，再與衛生署死因檔資料連結將89年1-3月與87、86年死亡但延遲至88年申報之個案排除，其他視為88年內政部死亡檔中實際為88年死亡者(共

表一 民國71 91年內政部與衛生署統計之死亡人數比較

年度	內政部 ^a	衛生署 ^b	差異人數 ^c	符合率(%) ^d
71	87,425	86,452	973	98.9
72	91,303	89,831	1,472	98.4
73	89,909	88,772	1,137	98.7
74	92,247	91,121	1,126	98.8
75	95,064	93,865	1,199	98.7
76	96,650	95,354	1,296	98.7
77	102,312	101,055	1,257	98.8
78	103,653	102,242	1,411	98.6
79	105,658	104,089	1,569	98.5
80	105,979	104,461	1,518	98.6
81	110,413	108,810	1,603	98.5
82	111,289	109,269	2,020	98.2
83	113,896	112,238	1,658	98.5
84	119,693	117,954	1,739	98.5
85	121,933	120,605	1,328	98.9
86	121,014	119,385	1,629	98.7
87	123,489	121,946	1,543	98.8
88	126,654	124,991	1,663	98.7
89	126,016	124,481	1,535	98.8
90	127,892	126,667	1,225	99.0
91	128,357	126,936	1,421	98.9

註：^a 資料來源：[9]

^b 資料來源：71 90年資料：[10]

91年資料：[11]

^c 差異 = 內政部死亡人數 - 衛生署死亡人數。

^d 符合率 = (衛生署死亡檔人數 / 內政部死亡檔人數) × 100%。

126,447人)，其中可與88年衛生署死因檔連結上者有120,244人(佔95.09%個案)，未連結上者有6,203人(佔4.91%個案)。而衛生署88年死亡檔中未能與內政部88年死亡個案連結上之人數為4,747人，佔3.80%之個案。

表三為88年衛生署死因檔中未能與內政部死亡檔連結者人口與區域特質，整體來說，88年衛生署死因檔中有3.8%未能與內政部死亡檔連結。進一步分析發現年齡層低者未連結率較高，地區別則以中區與東區未連結率較高，而性別部分則差異不大，而年齡與地區別未連結率皆達到統計顯著差異。

討 論

若從每年內政統計與衛生統計公布之死亡數字來看(表一)，內政部與衛生署各年度公布之死亡人數差異量皆不到2%。此差異經與相關單位主管人員討論後發現，可能與兩單位統計擷取延遲申報者之時間點不同，若88年死亡個案延遲至89年2 3月才到戶政申報，衛生署不會納入89年分析，亦不會納入88年死亡檔內，針對此問題，謝功毅等[3]研究是利用內政部死亡資料來填補衛生署未納入之延遲申報個案。

另一方面由於衛生單位向戶政單位取得死亡證明影印本，再進行人工死因譯碼。相

表二 88內政部死亡資料連結各年度衛生署死因資料結果

	88年死亡檔	
	內政部 ^a (百分比%) ^b	衛生署 (百分比%) ^c
死亡檔資料筆數 ^d	159,517	124,991
依同一人歸戶後	159,144	124,991 (100.00)
可與89年衛生署死因檔連結上者	29,658	
可與87年衛生署死因檔連結上者	3,027	
可與86年衛生署死因檔連結上者	12	
內政部88年死亡檔扣除可與	126,447 (100.00)	
89、87或86年衛生署死因檔連結上者		
可與88年衛生署死因檔連結上者	120,244 (95.09)	120,244 (96.20)
未連結上者	6,203 (4.91)	4,747 (3.80)

註：^a 88年內政部死亡檔包括88年1月 89年3月間申報死亡者，含死亡日期為88年以前或89年1 3月者。

^b 百分比之計算是以88年內政部死亡資料中扣除可與89、87、及86年衛生署死亡資料連結上之人數126,447人作為分母。

^c 百分比之計算是以88年衛生署死因檔死亡人數(124,991)為分母。

^d 內政部與衛生署死亡檔案皆依身分證字號歸戶後再做連結。

表三 88年衛生署死亡檔中未能與內政部死亡檔連結上者之基本資料分析

	88年衛生署死亡檔		
	筆數	未與內政部死亡 檔連結上筆數	未連結率 ^b (%)
合計	124,991	4,747	3.80
年齡別			***
< 15	3,397	188	5.53
15 34	7,082	360	5.08
35 64	35,669	1,446	4.05
65歲以上	78,843	2,753	3.49
性別			
男性	77,153	2,905	3.77
女性	47,838	1,842	3.85
地區別 ^a			***
台北	32,658	1,098	3.36
北區	16,161	571	3.53
中區	25,133	1,198	4.77
南區	23,486	862	3.67
高屏	22,252	797	3.58
東區	5,301	221	4.17

註：^a 地區別以健保局各分局地區劃分

^b 未連結率=(未與內政部死亡檔連結上筆數/衛生署死亡檔筆數)*100%

^c 卡方檢定：*p<0.01，**p<0.001，***p<0.0001

Taiwan Public Health Association
台灣公共衛生學會

關單位主管人員表示若死亡證明書上之醫院章若蓋到身份證字號上，則影印本上之身分證字號可能因辨識差而造成譯碼上的錯誤。謝功毅等[3]研究比對衛生署與內政部死亡檔案亦發現衛生署資料庫之身分證字號變項有缺漏、重複率及檢查碼錯誤較內政部高，以資料準確性來說，內政部死亡資料檔除死因資料未鍵入外，應較衛生署為佳。

進一步將兩單位死亡檔案依身分證字號連結逐一比對(表二)，可發現88年內政部死亡檔中若排除在其他年度死亡者後，符合88年死亡之個案人數為126,447人，小於內政統計公佈之死亡人數126,654人，但因檔案在歸戶時已排除373筆身分證字號重複個案，故本研究人數較少可能是因身分證字號重複而被排除所造成。而比對結果也發現內政部死亡檔中，88年死亡個案與衛生署88年死亡檔符合個案降為95.09%，內政部與衛生署檔案中各有6,203人與4,747人連結不上。此差異除因衛生署檔案未將失蹤人口與該年度死亡但延遲申報者納入檔案，所以內政部死亡檔之個案人數會較多，但衛生署仍有4,747筆資料未能與內政部資料連結，究竟是否因身分證字號鍵入有誤而造成，或有其他原因，值得相關單位注意。

由於本研究所使用之內政部死亡檔僅有身分證字號變項，無法再進一步加入出生日期作為檔案連結的依據，亦無法用死亡日期變項，未連結上之部分有可能為85年以前死亡個案造成，但由86年衛生署死亡檔可與88年內政部死亡檔連結上者僅12人來看，推估此部分造成之影響應相當低。

研究者在現階段使用衛生署死亡資料時，必須注意年齡與區域別連結率上的差異，並應盡量再與內政部死亡檔案連結以獲得較完整死亡個案。而根本之道是相關單位應加強稽核死亡資料建立與收集之品質，如要求醫院蓋官印位置正確、加強影印本之辨識率、或改用電子死亡證明書，以提高資料之準確性。也建議相關研究應將兩單位死亡檔中不一致個案挑出進一步深入分析，釐清差異之發生原因，並評估所造成之影響與政策意涵。

致 謝

承蒙健保局計畫(DOH89-NH-045)提供加密死因檔資料，以及衛生署統計室之協助，特此致謝。

參考文獻

1. 內政部統計處：戶政類統計資料發布時間表－出生數、出生率、死亡數、死亡率統計資料背景說明。內政統計資訊服務網，2003/7。URL：<http://www.moi.gov.tw/W3/stat/home.asp/>
2. 衛生署統計室：行政院衛生署預告統計資料發布時間表－統計資料背景說明(生命統計主要死因統計)。衛生統計資訊網，2003/7。URL：<http://www.doh.gov.tw/statistic/統計資料背景整理版/生命統計//主要死因統計.doc/>
3. 謝功毅、陳保中、王榮德：公元 1980-1997年本國衛生署死因資料庫準確性之確認與補正。中華衛誌 2002；21：329-38。
4. 呂宗學：為什麼高血壓不是第一大死因？－談原死因選擇規則的相關問題。中華衛誌 2001；20：5-14。
5. Lu TH, Lee MC, Chou MC. Accuracy of cause-of-death coding in Taiwan: types of miscoding and effects on mortality statistics. Int J Epidemiol 2000;29:336-43。
6. 呂宗學、石臺平、賴華山、李麗雪、李孟智、周明智：台灣地區死因統計準確度評估(2)：死亡證明書高頻率開具者特徵分析。中華衛誌 1997；16：273-9。
7. 呂宗學、李孟智、周明智：死因統計品質的國際間比較。中山醫學雜誌 1996；7：57-62。
8. 呂宗學、石臺平、賴華山、李麗雪、李孟智、周明智：台灣地區死因統計準確度評估(1)：可能錯誤來源與評估方法。中華衛誌 1995；14：179-90。
9. 內政統計年報：台閩地區年齡別死亡人數－按發生日期統計。內政統計資訊服務

- 網，2003/7。URL：<http://www.moi.gov.tw/W3/stat/home.asp/>
10. 行政院衛生署：中華民國九十年衛生統計。台北：行政院衛生署，2002。
11. 衛生署統計室：民國91年臺灣地區主要死亡原因。衛生統計資訊網，2002/6。
URL：<http://www.doh.gov.tw/statistic/data/死因摘要/91年/表1.xls>