

「國民健康訪問調查」抽樣設計

洪永泰

「國民健康訪問調查」包括了三個獨立的調查母體，第一個是台灣地區，第二個是離島，第三個是山地；每一個調查母體都有其特殊性質與工作考量，因此而有不同的設計，首先要介紹的是這三個調查母體的定義。

壹、母體定義

(一) 台灣地區：以民國九十年一月十六日台灣地區戶籍資料之普通戶為抽樣母體。

(二) 離島地區：以民國九十年一月十六日澎湖群島、綠島、蘭嶼、金門、連江等外島地區戶籍資料之普通戶為抽樣母體。

(三) 山地：以民國九十年一月十六日所有三十個山地鄉戶籍資料之原住民戶為抽樣母體（全戶皆非原住民者不列入抽樣母體）。

這三個母體的抽樣作業程序彼此之間完全獨立，也就是「各做各的」，沒有互斥或互補的問題。

貳、操作性定義與抽樣底冊 (frame)

三個調查母體對調查對象的操作性定義原則上相同，凡具有中華民國國籍、設有戶籍之國民都是，但不包括軍事單位、醫療院所、學校、職訓中心、宿舍、監獄等機構內之居民。抽樣底冊是台灣地區戶籍資料之普通戶，中選戶內之所有成員均為受訪對象；未設籍於該戶但長期和該戶成員共同居住者，或已設籍在該戶但未和該戶成員共同居住者皆包括在內，也就是「人在籍在」、「人在籍不在」和「籍在人不在」者都是訪問對象，對於最後一類對象則是進行追蹤訪問。

參、抽樣設計

(一) 台灣地區：採分層、多階段、PPS(Probability Proportional to Size, 抽取率與單位大小成比例) 等機率的抽樣設計，首先將台灣地區的 359 個鄉鎮市區依照地理位置和都市生活圈分成七層，在分層之後，每一層第一階段抽出「鄉鎮市區」(第一層除外，詳見稍後的說明)，第二階段自中選的鄉鎮市區抽出「鄰」，第三階段再自中選的鄰抽出「戶」，中選的戶全查。

(二) 離島地區：採分層、兩階段、PPS、等機率的抽樣設計，以台東縣、澎湖縣、金門縣、連江縣等四個縣每一個縣為一層，每一層第一階段抽出「鄰」，其次再自中選的鄰抽出「戶」，中選的戶全查。

(三) 山地：採分層、兩階段、PPS、等機率的抽樣設計，以三十個山地鄉的每一個鄉為一層，每一層第一階段抽出「鄰」，

其次再自中選的鄰抽出「戶」，中選的戶全查。

肆、分層

(一) 台灣地區：將 359 個鄉鎮市區依照地理位置和都市生活圈分成七層，分別是：

第一層(大台北地區)，包括台北市的大安區、松山區、中正區、大同區、中山區、內湖區、南港區、信義區、文山區、北投區、士林區、萬華區和台北縣的板橋市、中和市、新莊市、三重市、永和市、新店市、汐止市、淡水鎮。

第二層(北縣、基隆)，包括台北縣第一層之外的鄉鎮與基隆市。

第三層(桃竹苗)，包括三個縣所有鄉鎮市與新竹市。

第四層(中彰投)，包括三個縣所有鄉鎮市與台中市。

第五層(雲嘉南)，包括三個縣所有鄉鎮市與嘉義市、台南市。

第六層(高屏澎)，包括三個縣所有鄉鎮市與高雄市。

第七層(宜花東)，包括三個縣所有鄉鎮市。

(二) 離島地區：以台東縣、澎湖縣、金門縣、連江縣等四個縣每一個縣為一層。

(三) 山地：以三十個山地鄉的每一個鄉為一層。

伍、總樣本數、各層樣本數、與各層各階段抽出單位數

(一) 台灣地區：總樣本數為 6600 戶，各層依戶數比例分配樣本數，
 規劃如表 1。

表 1 全國健康訪問調查台灣地區之各層各階段抽出單位數

層別	鄉鎮市區數	戶數	佔總戶數 %	分配樣本戶數	第一階段抽出鄉鎮市區數 *	第二階段每鄉鎮市區抽出鄰數	第三階段每鄰抽出戶數	調整後之抽出樣本戶數
(一)大台北都會區	20	1,689,482	25.36%	1674		418	4	1,672
(二)北縣、基隆	28	464,875	6.98%	461	6	19	4	456
(三)桃竹苗	47	879,692	13.20%	871	12	18	4	864
(四)中彰投	68	1,174,003	17.62%	1163	16	18	4	1,152
(五)雲嘉南	78	990,027	14.86%	981	12	20	4	960
(六)高屏澎	77	1,151,522	17.28%	1141	16	18	4	1,152
(七)宜花東	41	312,589	4.69%	310	6	14	4	336
合計	359	6,662,190	100.00%	6600	68	1,648		6,592

註：內政部八十九年底之戶籍資料，

* 除第一層為兩階段抽樣外（第一階段即直接抽鄰），其餘各層都是三階段抽樣。

(二) 離島地區：總樣本數為 400 戶，各層依戶數比例分配樣本數，
 規劃如表 2。

表 2 全國健康訪問調查離島地區之各層各階段抽出單位數

層別	戶數	佔總 戶數%	分配樣 本戶數	抽出 鄰數	每鄰抽 出戶數	調整後之 分配 樣本戶數
(一)綠島鄉、蘭嶼鄉	1,628	3.34%	14	8	4	32
(二)澎湖縣	27,643	56.69%	227	56	4	224
(三)金門縣	17,985	36.88%	148	36	4	144
(四)連江縣	1,508	3.09%	13	8	4	32
合計	48,764	100.00%	402	108		432

註：內政部八十九年底之戶籍資料

(三) 山地：總樣本數為 600 戶，各層依戶數比例分配樣本數，規劃如表 3。

表 3 全國健康訪問調查山地鄉之各層各階段抽出單位數

層別	戶數	佔總戶數%	分配樣本戶數	抽出鄰數	每鄰抽出戶數	調整後之分配樣本戶數
烏來鄉	1402	2.55%	15	4	4	16
大同鄉	1636	2.97%	18	4	4	16
南澳鄉	1603	2.91%	17	4	4	16
復興鄉	3566	6.48%	39	10	4	40
尖石鄉	2187	3.97%	24	6	4	24
五峰鄉	1494	2.71%	16	4	4	16
泰安鄉	1584	2.88%	17	4	4	16
和平鄉	4465	8.11%	49	12	4	48
信義鄉	4812	8.74%	52	12	4	48
仁愛鄉	4293	7.80%	47	12	4	48
阿里山鄉	1891	3.43%	21	6	4	24
茂林鄉	484	0.88%	5	2	4	8
桃源鄉	1074	1.95%	12	4	4	16
三民鄉	771	1.40%	8	2	4	8
三地門鄉	1820	3.31%	20	6	4	24
霧臺鄉	784	1.42%	9	2	4	8
瑪家鄉	1658	3.01%	18	4	4	16
泰武鄉	909	1.65%	10	2	4	8
來義鄉	1909	3.47%	21	6	4	24
春日鄉	1212	2.20%	13	4	4	16
獅子鄉	1335	2.43%	15	4	4	16
牡丹鄉	1640	2.98%	18	4	4	16
海端鄉	1036	1.88%	11	4	4	16
延平鄉	920	1.67%	10	2	4	8
金峰鄉	858	1.56%	9	2	4	8
達仁鄉	1366	2.48%	15	4	4	16
蘭嶼鄉	915	1.66%	10	2	4	8
秀林鄉	3979	7.23%	43	10	4	40
萬榮鄉	1962	3.56%	21	6	4	24
卓溪鄉	1486	2.70%	16	4	4	16
合計	55051	100.00%	600	152		608

註：內政部八十九年底之戶籍資料

陸、抽樣結果

抽樣工作所使用的母群體資料是衛生署統計室所提供的民國九十年一月十六日全國戶籍資料，總人口數為 22,287,976 人，三個調查母體都依照規劃順利抽出各階段樣本，其中台灣地區各層抽出的鄉鎮市區是(參見圖一)：

- (一) 第一層(大台北都會區)：台北市大安區、松山區、中正區、大同區、中山區、內湖區、南港區、信義區、文山區、北投區、士林區、萬華區，台北縣板橋市、中和市、新莊市、三重市、永和市、新店市、汐止市、淡水鎮；共抽出 418 鄰、1,672 戶、6,024 人。
- (二) 第二層(台北縣、基隆市)：台北縣蘆洲市、樹林市、五股鄉、林口鄉，基隆市安樂區、七堵區；共抽出 114 鄰，456 戶，1,704 人。
- (三) 第三層(桃園、新竹、苗栗地區)：桃園縣桃園市、中壢市、平鎮市、楊梅鎮、大溪鎮，新竹市北區、東區，新竹縣竹北鄉，苗栗縣竹南鎮、後龍鎮；共抽出 216 鄰、864 戶、3,433 人。
- (四) 第四層(台中、彰化、南投地區)：台中市北屯區、北區、南屯區、南區，台中縣神岡鄉、清水鎮、沙鹿鎮、梧棲鎮、豐原市、大里市、太平市，彰化縣伸港鄉、秀水鄉、溪湖鎮、員林鎮、彰化市，南投縣竹山鎮、南投市；共抽出 288 鄰、1,152 戶、5,167 人。

(五) 第五層(雲林、嘉義、台南地區)：雲林縣莿桐鄉、斗南鎮，嘉義市東區，嘉義縣梅山鄉、布袋鎮，台南市安南區、西區、東區，台南縣下營鄉、東山鄉、善化鎮、新營市；共抽出 240 鄰、960 戶、4,345 人。

(六) 第六層(高雄、屏東、澎湖地區)：高雄市三民區、左營區、小港區、前鎮區、苓雅區、新興區、楠梓區，高雄縣大寮鄉、林園鄉、湖內鄉、鳳山市，屏東縣內埔鄉、霧台鄉、屏東市，澎湖縣西嶼鄉；共抽出 288 鄰、1,152 戶、4,580 人。

(七) 第七層(宜蘭、花蓮、台東地區)：宜蘭縣冬山鄉、頭城鎮、宜蘭市，花蓮縣吉安鄉、富里鄉，台東縣卑南鄉；共抽出 84 鄰、336 戶、1,432 人。

總計台灣地區共抽出 88 個鄉鎮市區，1,648 鄰，6,592 戶，26,685 人。離島地區除了金門縣烏坵鄉因戶數過少並未抽中之外，其餘所有鄉鎮市全部中選，共抽出 108 鄰、432 戶、1,954 人。山地鄉則是以鄉為層，自然全部中選，共抽出 152 鄰、608 戶、2,797 人。

柒、抽出樣本的代表性檢定

抽出樣本的代表性檢定依照都市化程度、性別、年齡、和性別交叉年齡幾個項目進行，檢定結果如表 4 至表 6。所有的卡方檢定結果都顯示不論在都市化程度、性別、或是年齡層上，抽樣結果和全國資料都無顯著差異，具有代表性。

表 4 台灣地區第二至第七層抽出樣本鄉鎮市區都市化程度與母體的比較

都市化程度	戶數	母體百分比	樣本鄉鎮市區	樣本百分比
省轄市	1355483	0.2726	19	0.2794
1	101836	0.0205	2	0.0294
2	361718	0.0727	5	0.0735
3	650404	0.1308	8	0.1176
4	1322880	0.2660	19	0.2794
5	992512	0.1996	13	0.1912
6	187875	0.0378	2	0.0294
總計	4972708	100.00	68	100.00

$\chi^2 = 0.5623 < \chi^2(0.05, 6) = 12.59$ ，樣本與母體一致。

表 5 樣本與母群體之同質性檢定：性別

性別	期望值		觀察值	
	人數	母體百分比	人數	樣本百分比
男性	13,646	51.14%	13,647	51.14%
女性	13,039	48.86%	13,038	48.86%

$\chi^2 = 0.00009 < \chi^2(0.05, 1) = 3.84$ ，樣本與母體一致。

表 6 樣本與母群體之同質性檢定：年齡

年齡	期望值		觀察值	
	人數	母體百分比	人數	樣本百分比
0-9	3,716	13.92%	3,723	13.95%
10-19	4,155	15.57%	4,130	15.48%
20-29	4,527	16.97%	4,539	17.01%
30-39	4,579	17.16%	4,513	16.91%
40-49	4,177	15.65%	4,189	15.70%
50-59	2,300	8.62%	2,339	8.77%
60-69	1,712	6.42%	1,765	6.61%
70	1,518	5.69%	1,487	5.57%

$\chi^2 = 4.06 < \chi^2(0.05, 7) = 14.07$ ，樣本與母體一致。

捌、抽樣誤差的估計

由於抽樣設計的規模達到一般所稱「複雜抽樣」的境界，抽樣誤差的估計擬運用隨機分組（random groups）的方式進行，各層第一階段抽出單位數都是偶數即是為此功能預作準備，等資料蒐集工作完成後即可展開。

工作團隊

「國民健康訪問調查」抽樣工作由國家衛生研究院（一）生物統計與生物資訊研究組和（二）醫療保健政策研究組聯合組成抽樣工作小組，成員包括石曜堂主任（醫保組）、熊昭主任（生統組）、劉仁沛研究員（生統組，抽樣及資料中心計畫主持人）、生統組陳淑媛助研究員、翁文舜、陳怡如、謝銘訓、卓恩仔；醫保組張新儀助研究員、蘇文川、樂冠華、和小組顧問台大政治系教授洪永泰。在抽樣執行期間，感謝國衛院生統組劉季鑫先生協助整理全國戶籍資料，以及醫保組張華志先生協助撰寫 excel 巨集抽樣程式；另外感謝醫保組的鍾文慎小姐、黃韻錚小姐、徐郁清小姐、與生統組的薛慧敏博士、于慧芝小姐、李朝棟先生、莊美幸小姐、賴淑娟小姐、林美穗小姐等同仁，以及慈濟大學實習生陳姿臻小姐、蔡瑜珍小姐等人於現場協助抽樣；本文由洪永泰改寫自抽樣工作報告。

圖一、台灣地區分層及樣本鄉鎮市區

