

財團法人演譯基金會

美兆健康資源中心



美兆健康資料庫簡介

MJHRF

技術報告編號：MJHRF-TR-01

2016/08/22

引用本文參考文獻格式:

財團法人演譯基金會(2016)。美兆健康資料庫簡介。財團法人演譯基金會技術報告(編號：MJHRF-TR-01)。

美兆健康資料庫簡介

一、緒論

近年來，非傳染性疾病（NCD）被發現是造成全球疾病負擔的主因。根據 2014 年世界衛生組織對非傳染性疾病的全球狀況報告，比起其他的死亡原因，慢性非傳染性疾病造成了更多的死亡人數，預計到 2030 年將會產生 5,200 萬的死亡案例[1]。造成非傳染性疾病的因素是十分多元的，可分為遺傳因素和環境因素兩方面發展[2、3]，為了更進一步的理解這些疾病的成因和發展機制，已有許多研究在收集具有大規模健康數據的生物樣本，用來推論生理、社會經濟或遺傳因素和慢性疾病之間的潛在關聯[4、5]。

美兆健康資料庫是以健康人口為基礎所建立的長期追蹤資料庫，資料庫內容除了包括社會經濟、行為和健康檢查生化數據的美兆健康數據資料庫，其中有部分參與者更於健檢同時捐贈出他們的血液樣本，從而建立起美兆人體生物資料庫，此兩大資料庫的結合，基本上能滿足大多數大型前瞻性研究材料的要求。以下章節將描述美兆健康數據資料庫以及美兆人體生物資料庫的資料量及其多樣性。

二、美兆健康數據資料庫

美兆健康數據資料庫為一長期追蹤、大規模、並且十分全面的人口健康資料。自 1994 年起，健康數據資料庫收集了於美兆健康管理機構接受健檢服務的參加者，除了健康數據資料，早在 2002 年便有部分參與者同意捐贈血液樣本。美兆健康管理機構在台灣有四個健康檢查中心，分別位於台北（北部地區），桃園（西北地區），台中（中部地區），高雄（南部地區），為台灣和亞洲鄰近地區的民眾提供全面性的健康檢查服務。健康數據資料庫儲存了參與者歷次的身體健康檢查實際數據，資料最早可以追溯到 1994 年的健檢結果，因此健康數據資料庫中有許多參與者已有多次健檢記錄，其行為、生活型態等資訊係透過健康問卷收集，人體與生化檢測數據則透過健康檢查收集，這些數據資料將持續收集下去，沒有終止日期。

美兆健康數據資料庫最早的檢查記錄是在 1994 年，而最早的健康問卷則始於 1996 年。從 1996 年直至 2015 年底，美兆健康數據資料庫已累計超過 150 萬人次之健檢和問卷記錄，共計有 665,156 位參與者。參與者中，年齡 20 歲以上者之人數約占台灣 20 歲以上成人人口的 3%[6]。由於人們可以透過家庭會員或通過雇主擔保的方式參加美兆健康檢查服務，因此美兆健康數據資料庫參與者的來源可依據參與機制區分成家庭或公司群體。

美兆健康數據資料庫反映了中國大陸和台灣的文化和民族多樣性。除了已在島上生活千年的土著，馬來-波里尼西亞民族，自 17 世紀以來，中國不同地區的人民在不同的時間點陸續移民而來，促使台灣形成一個多元文化的社會 [7]。中國人、原住民、荷蘭人、西班牙人和日本人等文化的相互融合，也讓台灣成為在中華社會文化變遷研究的理想實驗室 [8]。為了強調這樣一個獨特的群體特徵，並且利用這些資訊來幫助數據使用者，美兆健康數據資料庫透過自填問卷調查，不僅可得知所有參與者的種族背景，同時也記錄他們的父母和祖父母所屬之族群。

三、健康問卷調查：社會經濟、人口、生活型態與行為

每位參與者在健康檢查之前會完成自填健康問卷。健康問卷資料自 1996 年以來便開始以電子記錄方式儲存著，歷經多次修訂，目前共有 134 道選擇題，分為 7 大類型：基本資料、個人和家族病史、自陳健康狀況、生活習慣、運動習慣、睡眠習慣和飲食習慣（表 1）。

表 1：美兆健康數據資料庫—美兆健康問卷資料內容

題目類型	內 容
基本資料	職業、教育程度、家庭收入、籍貫/種族
個人和家族病史	癌症、慢性疾病
自陳健康狀況	耳鳴、暈眩、排便、腸胃症狀等項
生活習慣	抽煙、飲酒、嚼檳榔等
運動習慣	常做運動、運動次數及時數
睡眠習慣	就寢時間、睡眠情況及時數
飲食習慣	食物、營養素攝取頻率及份量

四、生化資料：健康檢查數據

美兆健康管理機構進行的檢查含蓋 100 多項重要的生理指標，包含人體測量、血液檢查、影像學檢查等，生物化學分析項目包括空腹血糖、血脂肪、腎功能、肝功能、白蛋白、球蛋白、和尿酸分析等，亦有糞便潛血檢查與尿液分析。所有的生物檢體分析都交由各地區的美兆健康檢查中心實驗室進行。除了前述的生物檢體分析外，同時也進行其他項身體檢測，如腹部超音波、X 光、心電圖、肺功能、聽力、視力、眼壓、眼底攝影、婦科檢查和骨質密度（表 2）。

美兆健康管理機構在台灣的四家健康檢查中心皆採用標準化的健檢流程，通過 ISO 9001-2008 品質系統國際驗證的要求。為了確保檢測數據的品質，美兆診所檢驗科還參加了美國病理學會（CAP）的能力測試（Proficiency Test）。另

外由於檢驗設備會隨時間而更換，所以包含設備製造商，型號和使用期限等詳細資訊都被記錄在「MJHRF_TR_06 健康檢查儀器使用技術報告」中[9]。

表 2：美兆健康數據資料庫—美兆生化資料內容

檢驗類型	檢驗項目
一般理學檢查	身高、體重、體脂肪、腰圍、臀圍、血壓、脈搏等
血液檢查	血液常規檢查、血糖、血脂、肝膽功能、腎功能、尿酸、甲狀腺功能、鈣/磷/鐵、B/C 肝炎、發炎因子、腫瘤標記
糞便和尿液分析	尿蛋白、尿糖、糞便潛血等
腹部超音波	肝、腎
X 光	胸部、腹部、上腸胃道
其他	心電圖、肺功能、視力檢查、婦科檢查、骨質密度

五、美兆人體生物資料庫

除了問卷調查和生化資料，後續設立的人體生物資料庫，為探討人體分子生物醫學在縱貫性研究（longitudinal research）上之發展提供一個新的途徑。美兆人體生物資料庫儲存了參與健檢者所捐贈的血液檢體，檢體採集分為兩個階段：2002~2008 年和 2015 年至今。截至 2015 年底，美兆人體生物資料庫已收集了來自 84,542 位捐贈者的 558,104 管血清、221,397 管血漿、以及 261,861 管血沉棕黃層，並長期儲存在-80°C 的冰櫃中。截至 2015 年年底，檢體捐贈者中已有 1,834 人於問卷中陳述在過去曾經罹患過癌症（表 3）。

表 3：檢體捐贈者自陳癌症分布（截至 2015 年年底）

癌症類型	男性	女性	總計
鼻咽癌	48	27	75
肺癌	38	39	77
乳腺癌	1	264	265
胃癌	39	18	57
肝癌	81	24	105
直腸癌	116	63	179
子宮頸癌	0	530	530
前列腺癌	131	0	131
其他癌症	254	271	525
總計	666	1,168	1,834

六、美兆健康資料庫的主要特性

美兆健康數據資料庫的健康數據和生化資料是由健檢參與者所捐贈，每位捐贈者的捐贈總次數、頻率與時間的間隔有所不同。截至 2015 年底捐贈總人數為 665,156 人，其中只進行過一次健檢者計有 381,032 人，占捐贈總人數的 57.28%，其餘參與者皆至少進行 2 次以上的健檢，共有 284,124 位，參與者首次健檢記錄的人口和社會特徵分佈可參見表 4。針對至少進行 2 次以上健檢的參與者，我們以年為單位來計算其初次與最近一次健檢的時間間隔，追蹤時間在 4 年以下的參與者約有 52%，但有 22% 的參與者則追蹤長達 10 年以上的時間。

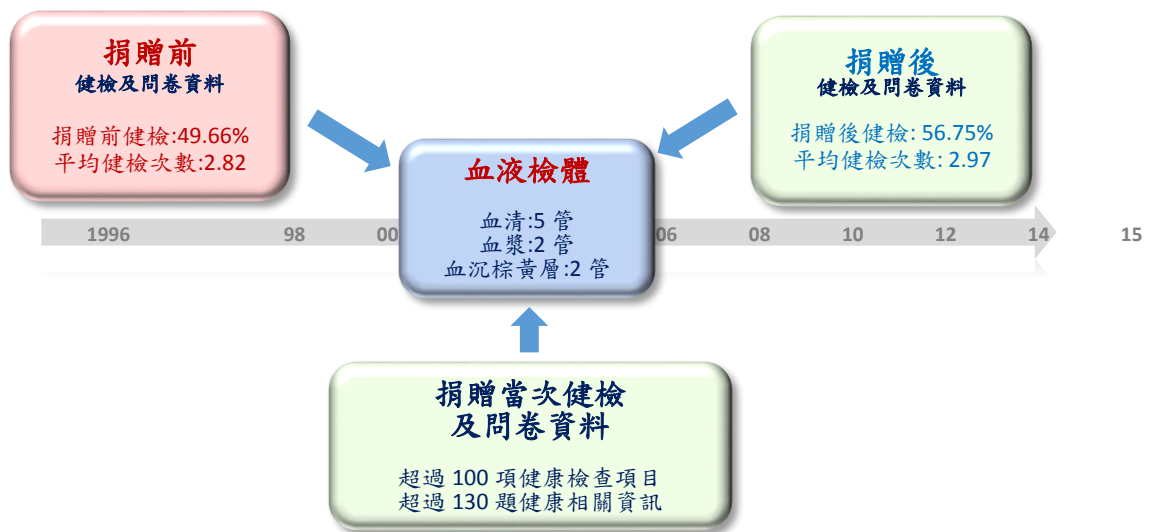
表 4：參與者首次健檢時之人口變項分佈

	男性 (n=318,356)		女性 (n=346,800)		總數 (n=665,156)	
健檢總次數(1-31)	2.3	(2.5)	2.2	(2.3)	2.3	(2.4)
年齡(1-98)	39.3	(14.4)	39.5	(14.3)	39.4	(14.4)
0-19	15,123	4.8%	12,966	3.7%	28,089	4.2%
20-64	281,430	88.4%	314,134	90.6%	595,564	89.5%
65 以上	21,803	6.8%	19,700	5.7%	41,503	6.2%
總數	318,356	100.0%	346,800	100.0%	665,156	100.0%
教育程度						
國中以下	45,932	18.4%	74,049	27.5%	119,981	23.1%
高中	52,810	21.2%	62,215	23.1%	115,025	22.2%
大專以上	150,763	60.4%	133,181	49.4%	283,944	54.7%
總數	249,505	100.0%	269,445	100.0%	518,950	100.0%
婚姻狀況						
無	90,376	37.0%	109,248	41.5%	199,624	39.3%
有	154,084	63.0%	153,964	58.5%	308,048	60.7%
總數	244,460	100.0%	263,212	100.0%	507,672	100.0%
就業情況						
無	28,203	11.6%	93,542	35.9%	121,745	24.1%
有	215,215	88.4%	167,314	64.1%	382,529	75.9%
總數	243,418	100.0%	260,856	100.0%	504,274	100.0%

括號中之數值為標準差。

在血液檢體採集方面，由於美兆健康資料庫是縱貫性質的長期追蹤資料，因此在人體生物資料庫中的參與者所捐贈之血液檢體，便可以同時串接到捐贈者在檢體捐贈當時的健檢資料，以及前後多次健檢的數據。如圖 1 所示，2002 年到 2008 年間捐贈血液檢體的參與者中，有 50% 的人可以往前串連到以往的健檢資料，這群人未捐贈檢體前之平均健康檢查次數為 2.8 次，同時有 57% 的參與者在捐贈血液樣本後直至 2015 年仍持續進行健檢，捐贈後之平均健檢次數約為 3 次。

圖 1：檢體捐贈者在捐贈前、後參與健康檢查的比例



七、個資保護及品質管控

為了保護個人隱私，美兆健康資料庫之資料均加以去識別化，並透過唯一的代碼取代個人識別資訊，個人身分資訊則儲存在專用伺服器中，經授權之人員方可進行使用。關於美兆健康資料庫的去識別化與資訊安全的詳細說明，可以參考「MJHRF_TR_03 資訊安全技術報告」[10]。

為確保資料庫數據品質，所有資料均經過清理，包括針對填寫者的答案進行邏輯檢查、不合理值處理，另外，由於問卷版本的變革，部分問卷題目及/或其回答選項會有些許更動，我們會依據題目變動情形拆解成多題，以便資料使用者進行分析。關於數據清理之詳細介紹，請見「MJHRF_TR_04MJHD 資料整理技術報告[11]」。

八、美兆人體生物資料庫的同意和再同意授權

參與者在捐贈健康數據和生物檢體在給美兆健康資料庫之前皆有得到參與者的知情同意。具體來說，所有參與者都會在進行健檢前收到美兆健康數據和人體生物資料庫參與者同意書，參與者可以選擇是否同意將他們的健康資料提供給財團法人演譯基金會作為研究用途，只有經過同意授權的資料才會被納入美兆健康資料庫中。

在 2010 年之前，台灣的法律並沒有針對人體生物資料庫之建立和營運有相關規定。在人體生物資料庫相關法規尚未訂定之時期，美兆人體生物資料庫於 2002 年獲得倫理審查委員會的批准，並在取得參與者簽署同意書之情況下，進行了第一階段的生物檢體採集，而後於 2008 年結束採集，與此同時，政府也開始擬定管理人體生物資料庫的相關法令，直到 2010 年政府頒布人體生物資料庫管理條例[12]，正式開始規範台灣地區人體生物資料庫的建立、管理和應用。

為遵循新的法規，美兆人體生物資料庫設置獨立的倫理委員會（EGC），以監督檢體與相關資料與資訊的用途，同時也依法設計了新的美兆人體生物資料庫參與者同意書。2014 年 9 月獲得衛生福利部的核准後，自同年 10 月起，針對 2002 年至 2008 年間曾捐贈血液檢體的參與者開始進行重簽同意書之工作，此外，在 2015 年 1 月也正式啟動了第二階段的血液檢體採集。截至 2016 年 7 月底，已完成 11,252 份的同意書重簽，占所有須重簽同意書之捐贈者的 15%。

九、美兆健康資料庫的管理

財團法人演譯基金會是台灣非營利性質的健康研究機構，於 2012 年由美兆健康管理機構的母公司—美兆集團捐款成立。華人特有的健康數據對於學術界和產業界的生物醫學研究十分具有價值，演譯基金會以促進生物醫學的尖端創新研究為己任，特別設置美兆健康資源中心，專責儲存和管理美兆健康數據資料庫。此外，演譯基金會擁有完善的保全系統以及先進的檢體儲存設備，並對於敏感性高的健康數據和生物檢體之處理皆擬訂了一套嚴謹的作業流程。所有相關資源的取得和應用都會由董事會督導。關於演譯基金會的完整資訊可參考以下網址：http://www.mjhrf.org/main/page/about/zh_tw/#about01。

十、研究出版概況

自 1990 年起美兆健康管理機構開始與學術和研究機構對各種公共健康研究議題進行合作[13]。2001 年至 2016 年間，利用美兆健康數據資料庫的數據所進行的研究已發表 47 篇學術出版物，其中有些文章更發表於高影響力的期刊。例如 Wen 等人以 428,584 人的世代為基礎，利用年齡、性別、健康史、B 型肝炎

和 C 型肝炎病毒感染、血清麩丙酮酸轉氨基酶（ALT）、血清麩草酸轉氨基酶（AST）、甲型胎兒蛋白（AFP）等變項建立模型，開發肝細胞癌風險預測模型 [14]。另外一篇則是以 416,175 人的樣本研究運動量的多寡與對於減低死亡風險並延長壽命之影響，本篇研究運用美兆健康問卷中的 3 個運動問題來測量運動代謝當量（MET）：最常做的運動類型、每次運動的持續時間、以及過去兩個禮拜的運動頻率 [15]。在另一項研究中則是用統合分析整合 35 個追蹤世代樣本（美兆健康數據資料庫的追蹤世代為其中之一）、170 萬名參與者，在控制年齡、性別、總膽固醇、收縮壓、糖尿病、心血管疾病史、尿蛋白和吸煙狀況的影響之下，腎絲球過濾率（eGFR）與末期腎病（ESRD）、死亡率間的關聯變化 [16]。

對於使用美兆健康數據資料庫數據發表的出版物有興趣者，可在網址：http://www.mjhrf.org/main/page/resource/zh_tw/#resouce08 中找到出版物清單的相關資訊。

十一、資料申請

美兆健康資料庫的數據資訊和生物檢體可供全球產學界內的研究機構進行生物醫學研究。健康問卷調查與健康檢查數據會依年度更新，目前最新版本已更新至 2015 年，健康數據資料庫的數據變項可參考數據資料譯碼簿，下載網址為：http://www.mjhrf.org/main/page/resource/zh_tw/#resource05。使用健康數據資料庫前，研究人員需經過申請程序，並提交相關申請文件，申請表單可於http://www.mjhrf.org/main/page/release2/zh_tw/#release03 下載。有意申請生物檢體者可至http://www.mjhrf.org/main/page/release1/zh_tw/#release01 網址中得到有關生物檢體申請流程的詳細資訊。申請通過後除了數據資料或檢體的費用外，尚須收取管理和服務費用，最後並提供該申請案件之使用授權碼，於未來發表研究論文時將按照美兆健康資源中心的資料庫引用格式引用本授權碼。任何關於美兆健康數據資料庫數據與人體生物資料庫申請和使用的相關問題，可以來信至 contact_us@mjhrf.org 詢問。

參考文獻

- 1). World Health Organization. Global Status Report on noncommunicable diseases 2014. World Health Organization. 2014; Available: <http://www.who.int/nmh/publications/ncd-status-report-2014/en/>.
- 2). World Health Organization. Genes and Human Disease – Genes and noncommunicable diseases. World Health Organization. Available: <http://www.who.int/genomics/public/geneticdiseases/en/index3.html>.
- 3). Sing CF, Stengård JH, Kardia, SLR. Genes, Environment, and Cardiovascular Disease. *Arterioscler Thromb Vasc Biol.* 2003; 23: 1190-96.
- 4). Chen Z, Chen J, Collins R, Guo Y, Peto R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol.* 2011; 40: 1652-66.
- 5). Jiang C, Thomas GN, Lam TH, Schooling CM, Zhang W, et al. Cohort Profile: The Guangzhou Biobank Cohort Study, a Guangzhou-Hong Kong-Birmingham collaboration. *Int J Epidemiol.* 2006; 35: 844-52.
- 6). Department of Household Registration, M.O.I. Household registration statistics data analysis in December 2015. Ministry of the Interior. Available: http://www.ris.gov.tw/en/web/ris3-english/home/-/asset_publisher/Ebr8/content/household-registration-statistics-data-analysis-in-december-2015?redirect=http%3A%2F%2Fwww.ris.gov.tw%2Fen%2Fweb%2Fris3-english%2Fhome%3Fp_p_id%3D101_INSTANCE_Ebr8%26p_p_lifecycle%3D0%26p_p_state%3Dnormal%26p_p_mode%3Dview%26p_p_col_id%3Dcolumn-2%26p_p_col_pos%3D1%26p_p_col_count%3D2.
- 7). Department of Information Services, Executive Yuan. The Republic of China Yearbook 2015 Taiwan. Executive Yuan. Available: <http://www.ey.gov.tw/en/cp.aspx?n=6CF4694061D32E34>.
- 8). Chen SH. Taiwan as a laboratory for the study of Chinese society and culture. *Bulletin of the Institute of Ethnology (Academia Sinica).* 1966; 14: 1-14.
- 9). Wang, ML. MJ Health Research Foundation Technical Report – Laboratory Equipment (MJHRF-TR-06). MJ Health Research Foundation. 2016; Available: <http://www.mjhrf.org/main/page/resource/en/#resource08>.
- 10). Chuang, YC. MJ Health Research Foundation Technical Report – Information Security (MJHRF-TR-03). MJ Health Research Foundation. 2016; Available: <http://www.mjhrf.org/main/page/resource/en/#resource08>.
- 11). Chuang, YC. MJ Health Research Foundation Technical Report – Data Cleaning (MJHRF-TR-04). MJ Health Research Foundation. 2016; Available: <http://www.mjhrf.org/main/page/resource/en/#resource08>.
- 12). Ministry of Health and Welfare. Human Biobank Management Act. Laws & Regulations Database of the Republic of China. 2012; Available:

<http://law.moj.gov.tw/Eng/LawClass/LawAll.aspx?PCode=L0020164>.

- 13). MJ AMHTS. The Annual Report of MJ Health Screening Center 2012. MJ Health Management Institution; 2013.
- 14). Wen CP, Lin J, Yang YC, Tsai MK, Tsao CK, Etzel C, Huang M, Hsu CY, Ye Y, Lopa M, Hawk E, Wu X. Hepatocellular carcinoma risk prediction model for the general population: the predictive power of transaminases. *J Natl Cancer Inst.* 2012; 104: 1599-1611.
- 15). Wen CP, Wai JP, Tsai MK, Yang YC, Cheng TY, et al. Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. *Lancet.* 2011; 378: 1244-53.
- 16). Coresh J, Turin TC, Matsushita K, Sang Y, Ballew SH, et al. Decline in estimated glomerular filtration rate and subsequent risk of end-stage renal disease and mortality. *JAMA.* 2014; 311: 2518-31.