

評論：台灣癌症登記與全民健康保險連結資料庫之可行性

本期綜論詳細介紹負責發行美國SEER-Medicare資料庫之相關機構、該資料庫之結構、以及相關研究資源與應用上之限制，並提出對台灣未來發展癌症研究資料庫之建議。台灣目前雖尚無類似SEER-Medicare資料庫之作法，由相關主管單位連結台灣癌症登記與全民健康保險資料庫後成為單一研究資料庫，但根據Medline與Web of Science等文獻索引資料庫之資訊，自2000年至2012年期間，以台灣癌症登記資料為研究材料所發表的研究論文，累計已經超過100篇以上。上述這些研究包含單獨使用癌症資料之研究[1-4]、串連癌症資料與健保資料之研究[5-9]、以及癌症資料串連其他資料(例如臨床檢驗檢查資料)之研究[10-12]，其數量雖不及應用SEER-Medicare資料庫之研究，但足見應用台灣癌症登記資料於研究上之潛力。台灣癌症登記資料庫經多年來的努力，已使得資料庫的品質與完整性達到國際標準，且達到族群為基礎的癌症登記系統(population-based cancer registry)，使用者宜定期查看資料庫的相關說明[13]。當使用次級資料進行研究時，需要先瞭解原資料收集的目的，例如健保資料庫主要是申報費用為主，癌症登記資料庫是癌症發生率的趨勢為主。必須考量原始資料間特性與目的的差異，進行必要的調整或妥協，才能在使用次級資料時增加其價值，並據以提出新的研究假說。

將多種研究資料庫串連成為整合性資料倉儲(integrated data warehouse)是未來醫療資訊系統在癌症研究的發展趨勢[14]，經由必要的資訊萃取(data extraction)、轉換(transformation)、整併(integration)、與除錯(cleansing)，提供符合諸如研究者、臨床

醫事人員、行政管理者等各種利害關係人(stakeholders)各自不同需求的資料。提出SEER-Medicare資料庫的發展與維護為例，這樣的工作仰賴不同機構的協同合作，落實「資料治理」(Data governance)的精神，以確保不同來源資料在彙整後的正確性、完整性與一致性。

資料庫正確性的驗證為建立整合性資料倉儲時重要考量因素之一。台灣癌症登記資料庫採取多種措施以確保收錄資料的正確性，醫院根據病歷專人登錄的資料，先經過癌症系統的資料勘誤檢核，申報至癌症登記工作小組後，會再由研究人員再行勘誤檢核程序，需要時請醫院癌症同仁再修正申報。此外，衛生署國民健康局委託國家衛生研究院執行的「癌症診療品質認證」，也會進行癌症申報資料的隨機病歷再閱(medical records re-abstraction)，經由比對病歷資料與申報資料的一致性，確保癌症登記資料庫的申報品質[7,8]。然整合多重資料來源成為具有加值效果之大型研究資料庫，其資料正確性更要透過多元化的資料驗證措施[15-19]。

個人資料安全也是在探討串連不同資料庫時需要重視的議題。當病人期待要更先進，更好的藥物及醫療照護的同時，也要求他們的資料要被適當的保護。資料維護單位與病人就如同委託與受託的關係，需要以信任做核心，規劃對資料保密、資料安全、互相的盡責(accountability)、資料的擁有權等配套措施。以美國為例，政府針對HIPAA (Health Insurance Portability and Accountability Act, HIPAA)保護個人隱私法案的實施之時，IOM (Institute of Medicine of the National Academies, IOM)強烈建議國會應授權給衛生福利部，有別於一般的HIPAA的規定，針對所有健康有關的研究提出新的一套保護隱私的規定[20]。例如美國國家癌症研究院(National Cancer Institute, NCI)針對21個癌症中心建立多個資料庫的平台multiple data warehouse: the National

國立台灣大學公共衛生學院流行病學與預防醫學研究所

賴美淑

E-mail: mslai@ntu.edu.tw

聯絡地址：台北市中正區徐州路17號

Comprehensive Cancer Network outcomes database (NCCN)，也是一個整合性資料倉儲成功的例子[14]，它建立大多數的癌症病人長期追蹤資料。該資料庫與SEER最大的不同在於其收錄超過300個變項，包括轉移的位置、生物標記(Bio-makers)、治療的持續性、中斷的原因等。但是除了要將病人資料串到檢體收集的情形之外，這個資料庫運作被允許排除在HIPPA限制之外。

建立癌症研究整合性資料倉儲的長期目標，就是期望促進癌症照護的進步，包括高科技藥物、團隊照護、可近性、民眾付擔的起(affordable)的高品質的癌症照護體系等。但是這樣的工作需要凝聚主管機關、民眾、學界、癌症臨床專業人員與癌登人員的共識，投入相當的人力、物力始能達成，以目前人權團體對於健保資料用於研究用途上仍有異議，健保資料、癌登資料與死因登記分屬不同權責單位，且癌症登記工作經費逐年遭主管機關大幅刪減的現況下，除了排除上述困難，冀望台灣衛生主管單位更積極籌畫『建立癌症研究整合性資料倉儲』的長期目標，先以委員會勾勒出願景與行動藍圖，編列預算，結合現有的癌症登記資料庫、健保資料庫和死亡資料庫之外，還可進一步結合篩檢資料庫、疫苗注射資料、生物檢體組織庫等更多的國家資源，及醫院癌症中心資料庫等之外，同時提供資料倉儲之教育、訓練、推廣與結果分享，期使台灣癌症有創新的研究發展，有助提升國人的健康。

參考文獻

1. Chen PT, Kuan FC, Huang CE, et al. Incidence and patterns of second primary malignancies following oral cavity cancers in a prevalent area of betel-nut chewing: a population-based cohort of 26,166 patients in Taiwan. *Jpn J Clin Oncol* 2011;**41**:1336-43.
2. Chiang CJ, Chen YC, Chen CJ, et al. Cancer trends in Taiwan. *Jpn J Clin Oncol* 2010;**40**:897-904.
3. Lin CH, Chen YC, Chiang CJ, et al. The emerging epidemic of estrogen-related cancers in young women in a developing Asian country. *Int J Cancer* 2012;**130**:2629-37.
4. Wu SJ, Huang SY, Lin CT, Lin YJ, Chang CJ, Tien HF. The incidence of chronic lymphocytic leukemia in Taiwan, 1986-2005: a distinct increasing trend with birth-cohort effect. *Blood* 2010;**116**:4430-5.
5. Chang CH, Lin JW, Wu LC, Lai MS, Chuang LM, Chan KA. Association of thiazolidinediones with liver cancer and colorectal cancer in type 2 diabetes mellitus. *Hepatology* 2012;**55**:1462-72.
6. Chen WW, Shao YY, Shau WY, et al. The impact of diabetes mellitus on prognosis of early breast cancer in Asia. *Oncologist* 2012;**17**:485-91.
7. Kuo RN, Chung KP, Lai MS. Effect of the pay-for-performance program for breast cancer care in Taiwan. *Am J Manag Care* 2011;**17**(5 Spec No):e203-11.
8. Kuo RN, Chung KP, Lai MS. Re-examining the significance of surgical volume to breast cancer survival and recurrence versus process quality of care in Taiwan. *Health Serv Res* 2012: doi: 10.1111/j.1475-6773.2012.01430.x. [Epub ahead of print]
9. Tang CH, Pwu RF, Tsai IC, et al. Costs of cervical cancer and precancerous lesions treatment in a publicly financed health care system. *Arch Gynecol Obstet* 2010;**281**:683-95.
10. Chen CJ, Yang HI, Su J, et al. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. *JAMA* 2006;**295**:65-73.
11. Huang YT, Jen CL, Yang HI, et al. Lifetime risk and sex difference of hepatocellular carcinoma among patients with chronic hepatitis B and C. *J Clin Oncol* 2011;**29**:3643-50.
12. Yang HI, Yeh SH, Chen PJ, et al. Associations between hepatitis B virus genotype and mutants and the risk of hepatocellular carcinoma. *J Natl Cancer Inst* 2008;**100**:1134-43.
13. Bureau Health Promotion, Department of Health, Executive Yuan, R.O.C. (Taiwan). Taiwan cancer registry annual. Available at: <http://www.bhp.doh.gov.tw/BHPnet/Portal/StatisticsShow.aspx?No=200911300001>. Accessed August 13, 2012.
14. Nass SJ, Wizemann T. Informatics Needs and Challenges in Cancer Research: Workshop Summary. Washington, DC: Institute of Medicine, 2012.
15. Peabody JW, Luck J, Jain S, Bertenthal D, Glassman P. Assessing the accuracy of administrative data in health information systems. *Med Care* 2004;**42**:1066-72.

16. Earle CC, Nattinger AB, Potosky AL, et al. Identifying cancer relapse using SEER-Medicare Data. *Med Care* 2002;**40**(8 Suppl):IV-75-81.
17. Penberthy L, McClish D, Pugh A, Smith W, Manning C, Retchin S. Using hospital discharge files to enhance cancer surveillance. *Am J Epidemiol* 2003;**158**:27-34.
18. Lamont EB, Herndon JE 2nd, Weeks JC, et al. Criterion validity of Medicare chemotherapy claims in Cancer and Leukemia Group B breast and lung cancer trial participants. *J Natl Cancer Inst* 2005;**97**:1080-3.
19. Leung KM, Hasan AG, Rees KS, Parker RG, Legorreta AP. Patients with newly diagnosed carcinoma of the breast: validation of a claim-based identification algorithm. *J Clin Epidemiol* 1999;**52**:57-64.
20. Nass SJ, Levit LA, Gostin LO. Beyond the HIPAA Privacy Rule : Enhancing Privacy, Improving Health Through Research. Washington, DC: National Academies Press, 2009.