[研究動態報導]

運用網格空間單元分析疾病之時空變化

臺灣大學生物環境系統工程學系　蘇明道
中央研究院台灣史研究所　劉士永
中央研究院人文社會科學研究中心　詹大千

一、計畫簡介

空間資料處理與分析技術的發展，使得具空間特性資料的分析與應用越來越成熟普遍，尤其對在空間上非均勻分布的事件（spatial uneven events；例如疾病、犯罪、污染等）的空間分布特性分析與繪圖更形方便，不僅可以使公共事務的管理與決策者對如污染事件、疫情、公安、公衛事件…等的起源與擴散分布過程更加了解，也可以形成更多的假說作爲進一步調查與研究的基礎。更重要的是，它可以在空間上確定的高風險領域以有效進行針對性的環境危害控制和公共衛生預防。但雖然空間資訊技術(GIS)的發展使得地圖的繪製更加容易，事件資料的空間圖繪只能提供視覺上的檢視，而無法明確的展現空間資料群聚特性與現象，也無法展示資料間的空間相關性(spatial correlations)，空間統計(spatial statistics)則已被廣泛用於事件與空間位置關係的研討(Croner *et al*., 2000; Aziz *et al*., 2012; Shafie 2011)。

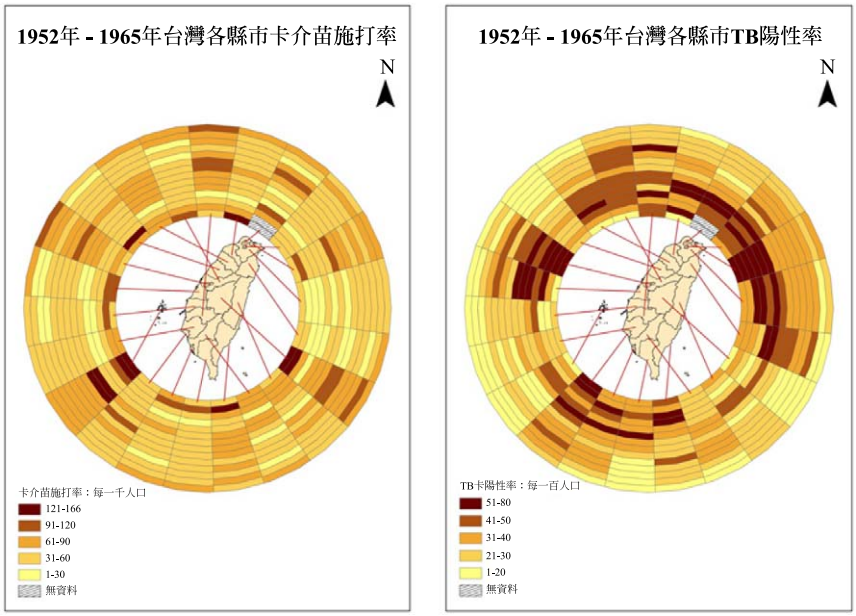
點資料分析(point pattern analysis)是研討與偵測空間群聚現象(clustering patterns)常用的方法，但與社會經濟相關的點資料常會涉及事件的實際位置而引發隱私保護的議題(Maheswan and Haining, 2004)。因此如核密度分析(kernel estimation)與地理統計(geostatistical methods)乃經常被用來進行風險分布面(risk surface estimation)的研討，利用網格(raster or cell-based)資料模型來建立連續分布的面資料，以減緩隱私暴露的疑慮同時又能忠實的將空間分布特性保留下來 (Bithell, 1990, 1999; Diggle, 2000).

然而不均勻活動的空間群集分布圖只是某一時間的一個靜態快照，而忽略的時間動態(temporal kinetics)，因而很難提供對事件未來發展的資訊。因此如 ARIMA 模式(autoregressive integrated moving average model)等時間序列統計方法則被用來預估預期的發病率(Zeger *et al*., 2004; Earnest *et al*., 2005)，時空分析應用於疾病資料的分析也越來越受到重視（Wen *et al*., 2006; Eisen & Lozano-Fuentes 2009; Jeefoo, Tripathi, and Souris 2011; Alzahrani *et al*., 2013)。但是這類方法並不能提供空間風險區域範圍的識別。有鑑於此乃提出「運用網格空間單元分析疾病之時空變化」跨領域整合型計畫，總共包含三個子計畫，分別由臺灣大學與中央研究院的團隊執行。

本計畫的整體目標在於運用網格空間單元爲基礎的計算方法，讓疾病的分布可以與其他不同層級的社會、經濟、人口、環境資料進行串接分析，再針對早期與近代的傳染病、慢性病進行時空分析，並依照不同的疾病特性與分布，找出環境、社會、經濟等因子對於該疾病的影響，對於台灣的學界與衛生行政機關，更可以利用本研究的分析結果，進一步設立疾病預警區，並找出可能造成疾病成因的環境與社會經濟等因子，對於預防醫學將有重大的幫助。

本整合型計畫中，三個子計畫的研究簡介分別如下：

第一個子計畫：「台灣疾病轉型之時空演變-健康歷史」的研究範疇將聚焦基本上以 1920 年代日治時期到 1970 年代間，台北地區人口與疾病罹患類型作一長時段的變遷趨勢，由於台北的行政區劃一直有所變動，而且各區的社會經濟活動型態也不盡然一致，因此人口中的疾病罹患結構與類型是否符合這些假設的背景變動，遂成爲一個頗有意思的議題。此等研究一來可以再幫助我們確認疾病類型、人口結構，與社經條件之關



圖一　1952-1965 年台灣各縣市卡介苗施打率與肺結核陽性率（最內圈為 1952 年，最外圈為 1965 年）

係，二來也可探討以疾病類型作爲都市發展下，社會與產業樣態變遷指標的可適性。

第二個子計畫：「傳染病之擴散趨勢與預警區之建立」將以國土資訊系統所建立之社經資料統計區分類系統爲依據，探討該系統空間單元在隱私保護與空間分布特性維持下的效能，作爲未來建立流行病資料監測與發佈應用之依據。計畫中將利用空間分析方法，有效整合病例、病媒、社經資料建立登革熱風險指標，並整合其他兩個子計畫階段性成果，建立登革熱擴散潛勢分析模式。計畫中亦將利用統計區分類系統，在考量隱私保護與保存空間分布型態間取得平衡，建立傳染病資料之發佈方式。

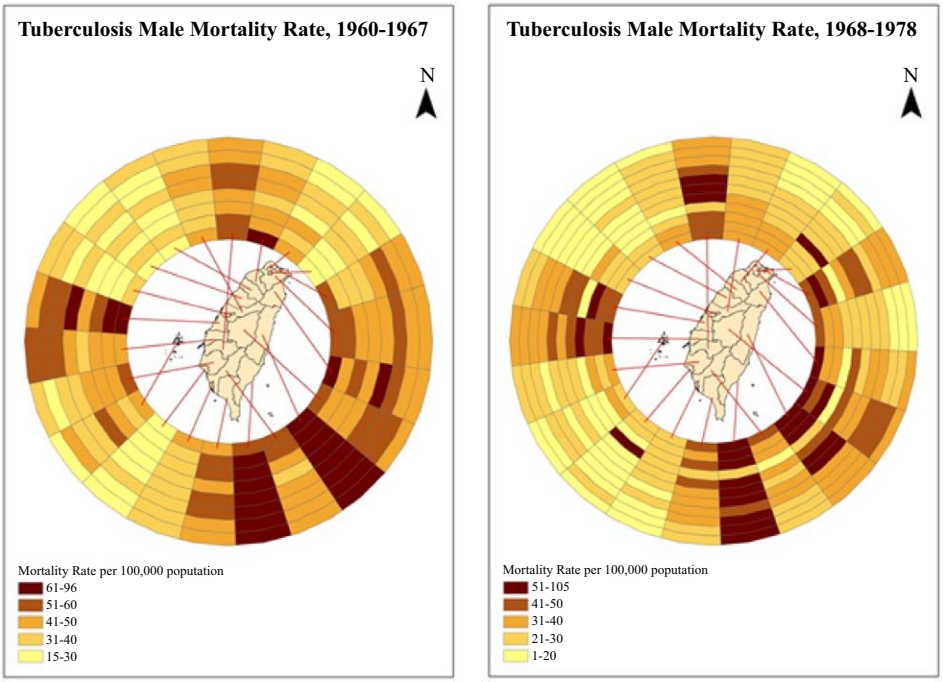
第三個子計畫：「慢性病之時空分析與健康不平等」在健保制度實施後（1995 年），挑選台灣近年來高疾病負擔的幾項慢性病如慢性阻塞性肺病、慢性腎臟病、等分布與所處社會、經濟環境間的關係，透過不同時期、不同空間單元的疾病、人口、社會、經濟等統計資料的整合，運用網格單元與統計區進行模式建構，期望藉由本研究可以清楚地辨識出高風險的網格區域，並尋找可能影響的環境危險因子，可供衛生單位進行後續的介入與追蹤、並提供學界未來進行世代研究的重要參考。

二、研究成果

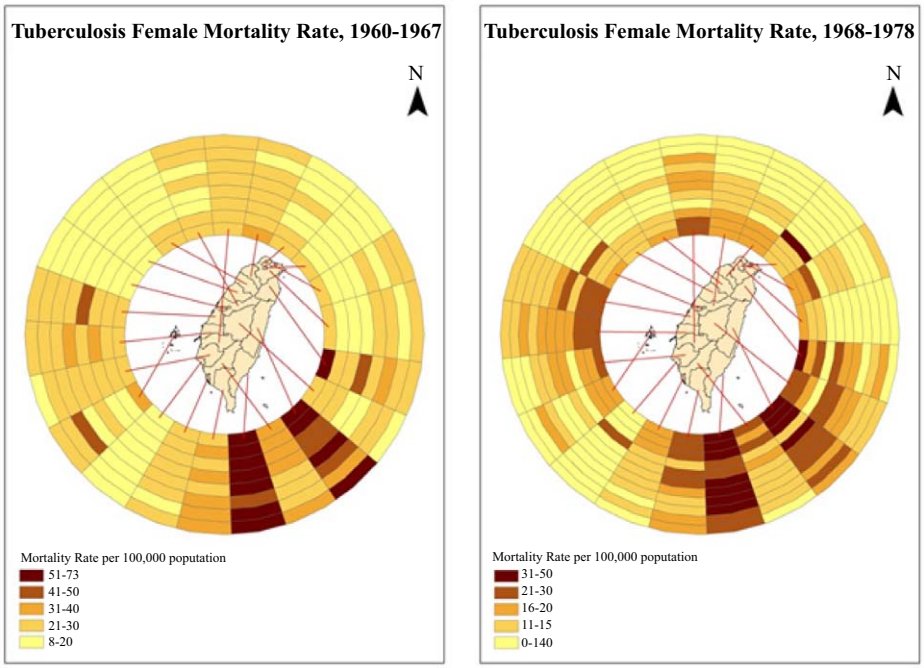
本整合型研究計畫執行期間爲 2 年（2013 年 8 月至 2015 年 7 月），研究成果分別分述如下。

1. 肺結核的時空演變：

原擬聚焦於以 1920 年代日治時期到 1970 年代間，台北地區人口與疾病罹患類型作一長時段的變遷趨勢。此一研究課題，在執行過程中已漸次擴張爲整編日治時期以來台灣地區人、疾病，與地理空間相關之統計資訊，透過這些訊息協助並支援其他兩個子計畫之分析研究外，亦就特定疾病類型分析時空變化下社會與產業樣態變遷指標的可適性。有關肺結核(TB)與卡介苗(BCG)接種之關係，根據 2011 年的數項有關孟加拉的研究顯示，在 1960-1970 年間，該國因衛生部長更迭以致於造成是否強制接 BCG 疫苗的政策有所波動，因此成爲觀察該疫苗造成 TB 僞陽性與爾後 TB 防治問題絕佳之觀察點。據相關研究所示，WHO 主張僅 TB 高風險成員才需注射 BCG 疫苗，而聯合國之 UNICIF 則認爲全面強制接種 BCG 才是正途。據此，台灣的 BCG 疫苗政策，當是在 UNICEF 的引導下進行，而非 WHO 的主張。台灣雖不具備類似孟加拉研究的歷史基礎，但仍舊可以就現有資料，從地理區位、性別、族群的因素，試圖標誌出 TB 感染率的時空變化(如下圖一至圖三所示)，日後或得進而推估台灣在歷經數十年全面 BCG 疫苗強制接種的情況下，



圖二　1960-1967 與 2003-2007 年男性肺結核死亡率

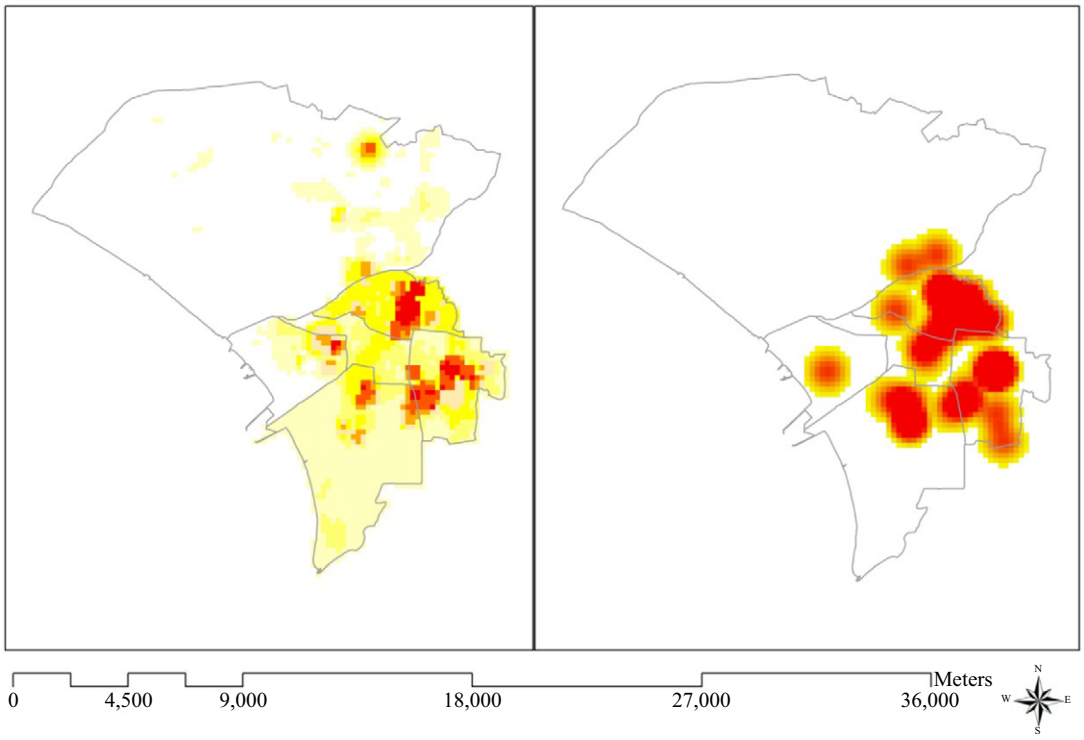


圖三　1960-1967 與 2003-2007 年女性肺結核死亡率

爲何仍有這些差異的出現，是因爲 WHO 之僞陽性假說成立？還是 BCG 效力衰退？抑或有其他非生物性的因素存在？凡此種種，仍然需要更進一步的分析才得進一步探索答案。

2. 建構傳染病風險模式：

傳染病疫情的監測與控制爲維護社會安全重要的一環，然而現今包括病例、病媒、相關社經資料均具有時空分布特性，且資料空間單元不一，需要積極建立相關資料的時空分析方法，並建立疫情風險指標，作爲傳染病監測與控制相關決策之參考。



圖四　登革熱病例密度圖（台南市）：左圖為推估密度，右圖為實際密度

蒐集與疫情相關的人口與社會經濟資料，包括分齡人口、建物、土地利用、基礎設施、公衛醫療、工商活動等。本研究將由國土資訊系統的地理資料倉儲系統整合所需的地理基礎資料如門牌資料、地形圖、使用分區圖、行政區圖、重要地標…等。並利用網格式資料結構建立時間-空間的基礎單元，並以核密度分析方法結合統計分析建立病例、病媒等疫情資訊，提供疫情於時間-空間上的擴散與分布情況，並將人口密度與社會經濟指標整合為易感受體圖資，以供後續分析。

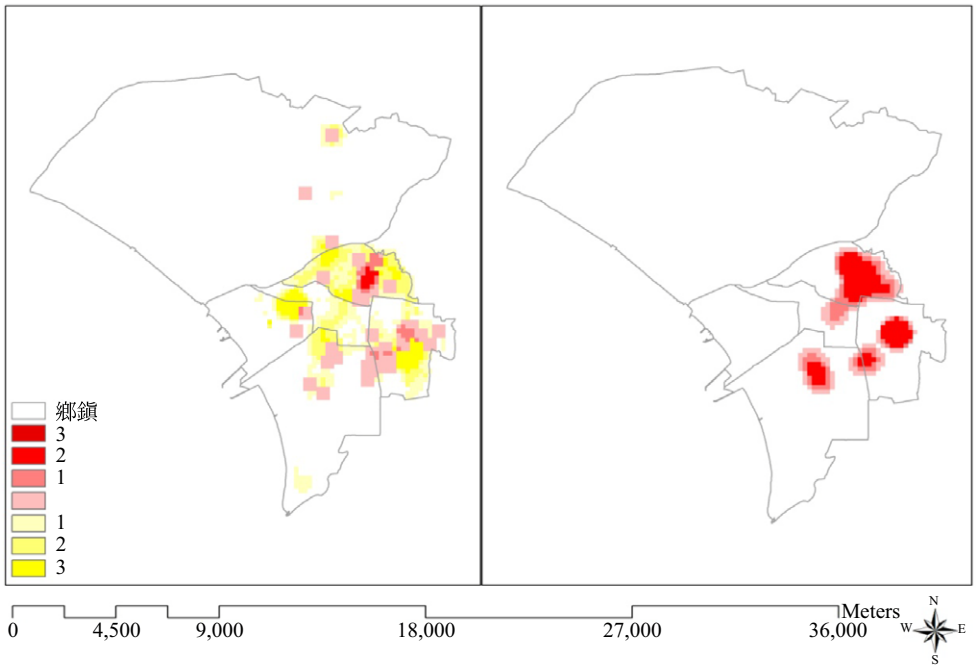
利用上述整合之網格式時間-空間單元，分析前數期期影響因素諸如病媒、病例、易感受體、氣溫等環境因素，對於下一期登革熱病例擴散之影響。建立監測區、警示區之模式，推估下一期登革熱病例密度之可能。以台南市為例：如上圖四所示左邊為推估所得登革熱病例密度，可視為登革熱擴散風險，右邊則為該期實際發生登革熱病例密度，利用推估值提供評估該期登革熱擴散危險區域之參考，建立傳染病擴散趨勢分析模式，以及登革熱風險地圖。

建立區域熱區分析模式

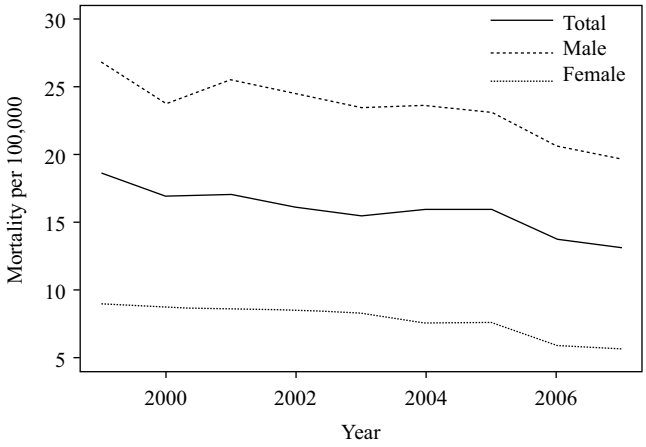
承上述結果建立傳染病流行風險指標，劃分疫區各級風險與熱區圖，如下圖五所示，右邊為該時期實際登革熱熱區圖，共分 3 級，級數越大者顏色越深，代表疫情越嚴峻之地區，左邊為推估疫情熱區，紅色代表警示區，黃色代表疫情未擴散之監測區，利用此分析模式評估各時期登革熱疫情之熱區。

3. 慢性阻塞性肺病 (Chronic Obstructive Pulmonary Disease, COPD)死亡率的時空分佈與危險因子（子計畫 3）：

根據世界衛生組織的統計，COPD 在 2004 年是全世界第四大死因，每年約造成全球 300 萬人死亡，約佔全球死亡數的 5.1%，全球的盛行率為每百萬人 63.6 人，預計在 2020 年將成為全球的第三大死因。而台灣 COPD 相關的死亡率統計，至 1981 年的每十萬人 33.2 降低至 2007 年的每十萬人 3.9。但是由於之前的死亡率統計並未考慮慢性呼吸道阻塞這個診斷(ICD-9：496)，藉由本研究將該診斷納入計算，可以與國際間的 COPD 死亡率進行比較。COPD 的危險因子包含了吸菸、生質燃燒的暴露、空氣污染、年齡、性別、社會經濟地位、肺結核感染與海拔高度。



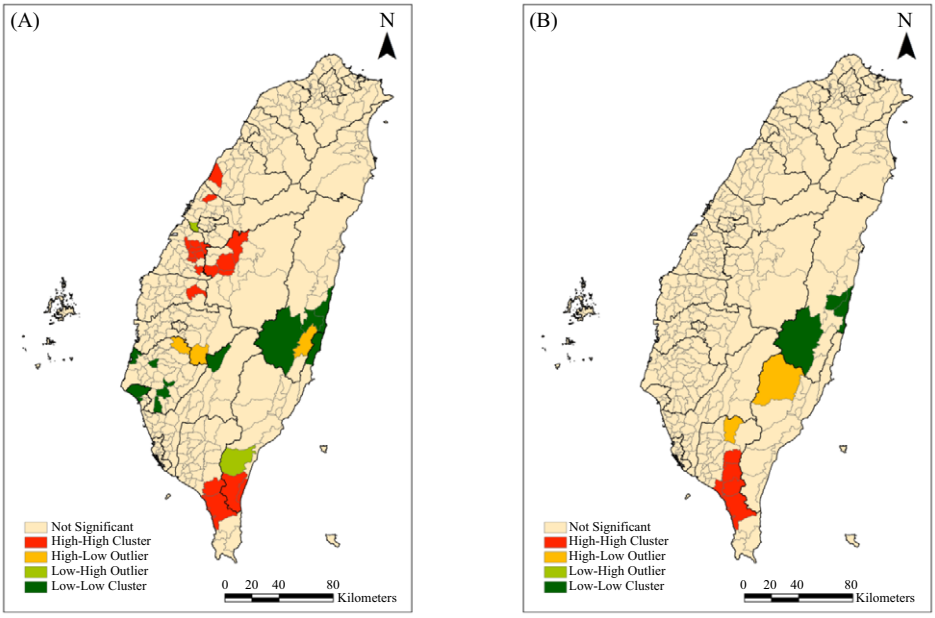
圖五　登革熱風險熱區圖（台南市）：左圖為推估疫情熱區，右圖為實際登革熱熱區圖



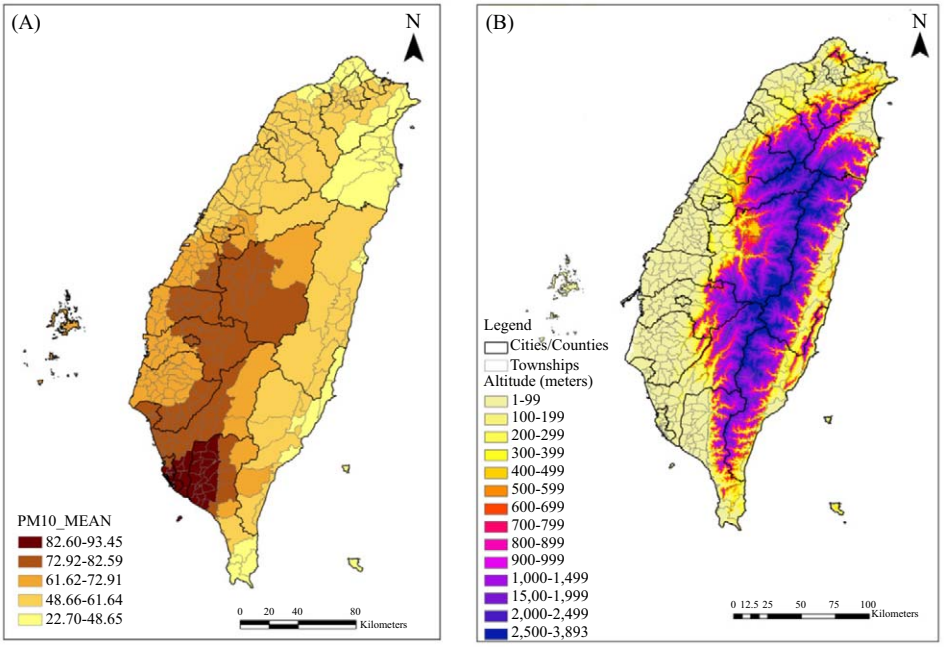
圖六　1999-2007 年慢性阻塞性肺病標準化死亡率趨勢圖

本研究利用 1999 年至 2007 年衛生福利部的死因資料進行年齡標準化死亡率的計算，整體死亡率與男、女個別的死亡率在時間趨勢上皆在下降（圖六），整體而言，男生的 COPD 死亡率是女性的 3 倍。線性迴歸的結果發現，地區原住民的比例、PM_{10}、CO 與 COPD 的死亡率呈現正相關，吸菸率與男性 COPD 死亡率呈現正相關，肺結核死亡率、醫療資源密度、NO_2、海拔高度與男性、女性 COPD 死亡率皆呈現負相關。

由於解釋變數於空間上的影響不一，因此採用地理加權迴歸模型(GWR)進行進階分析，不僅大幅提升模型的解釋力，男性至 49.4%提升到 72.2%，女性至 60.9%提升至 77.9%，此外更可以透過區域型空間自相關指標-Local Moran's I 來偵測殘差的聚集情況，可以發現在屏東縣有幾個鄉鎮有顯著的聚集（圖七），也就是有目前危險因子無法解釋的狀況，在本研究後來的社區研究中有進一步探討地區的危險因子。在本研究中由於空氣污染濃度（圖八(A)）與地形高層資料（圖八(B)），皆爲網格式的資料型態，並需再將網格資料運用區域統計方法計算平均濃度與平均高層，以進行線性迴歸與 GWR 的分析。



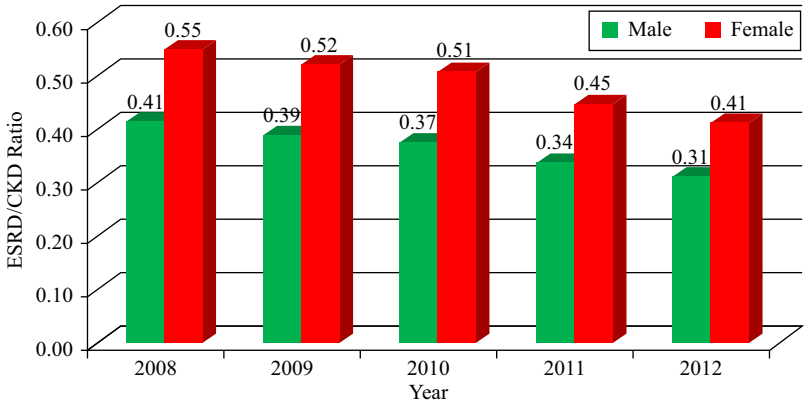
圖七　地理加權迴歸模型殘差值之空間自相關檢定(A)男生(B)女生



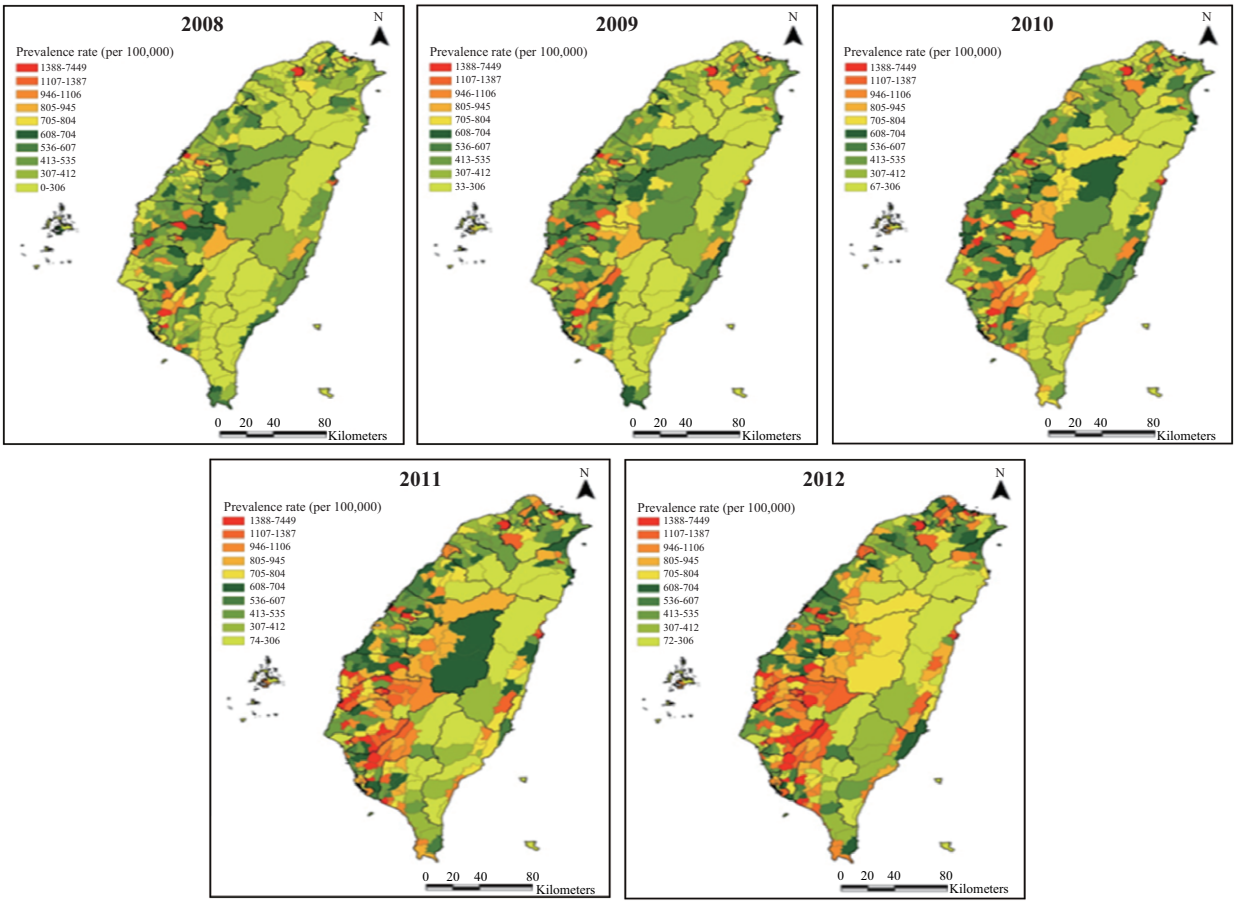
圖八　網格資料之運用(A) 1994-1999 PM_{10} 平均濃度(B)高層資料

4. 慢性腎臟病(Chronic Kidney Disease, CKD)死亡率與盛行率的時空分佈與危險因子（子計畫 3）：

根據美國 2013 年腎臟資料年報的統計，台灣在 2011 年末期腎臟病(ESRD)的發生率全世界第三高，而盛行率則是世界第一高，整體而言，慢性腎臟病在國人的盛行率約爲 9%，是國人十大死因的第十位（2013 年），末期腎臟病人洗腎的醫療支出約佔整體健保支出的 5.89%。本研究使用健保資料進行盛行率與醫療費用的趨勢推估，慢性腎臟病患的盛行率與醫療支出在研究的 2008-2012 年間是逐年遞增，但末期腎臟病患占全體慢性腎臟病患的比例，似乎有隨著時間逐年下降（圖九），有可能是健康保險署自 2006 年開



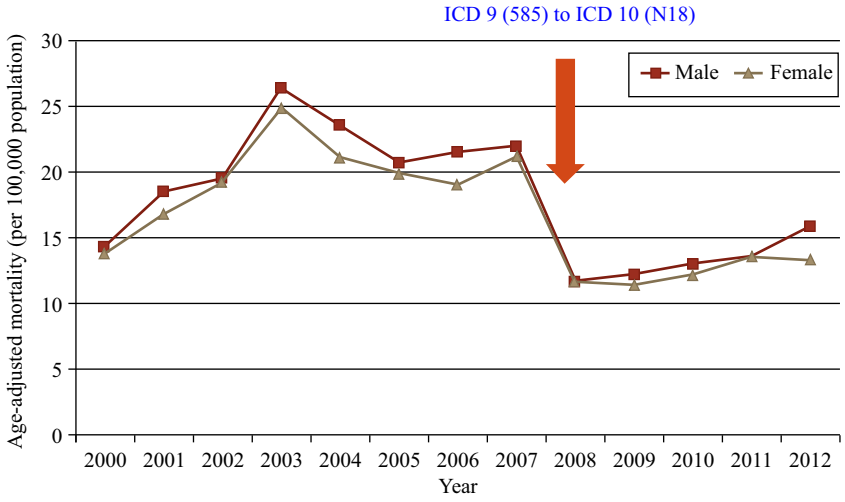
圖九　2008-2012 年末期腎臟病病人佔全部慢性腎臟病病人的比例



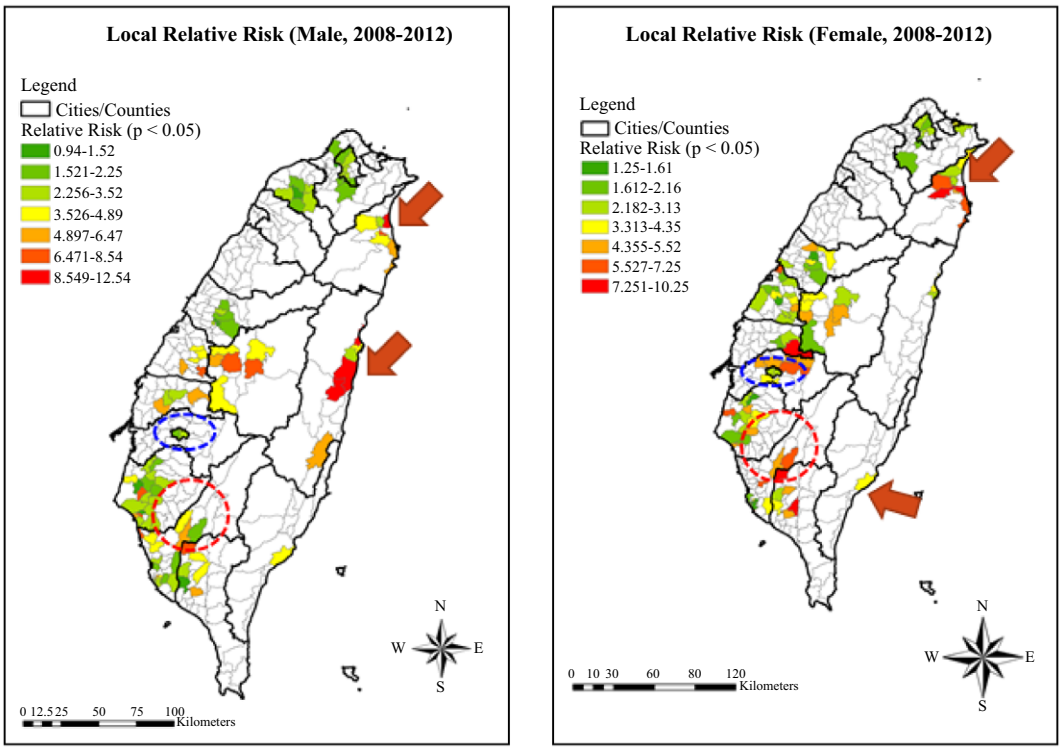
圖十　2008-2012 年慢性腎臟病鄉鎮盛行率估計

始「末期腎臟病前期(Pre-ESRD)之病人照護與衛教計畫」及「初期慢性腎臟病醫療給付改善方案」發揮了早期介入與積極治療的效果，降低慢性腎臟病病人進入末期腎臟病的機會。

依據全民健康保險研究資料庫的慢性腎臟病患資料進行鄉鎮盛行率的推估，由於目前健保資料的空間資訊只有病患的投保鄉鎮與就醫醫療院所的鄉鎮，這對於推估疾病盛行率是一大挑戰，本研究參考以往的研究成果進行推估，慢性腎臟病的盛行率地圖如圖十所列，在西半部以嘉義縣、嘉義市、台南市、高雄市交界的山區鄉鎮爲盛行率較高的地區。而死亡率的趨勢則由於疾



圖十一　2000-2012 年慢性腎臟病年齡標準化死亡率趨勢圖



圖十二　慢性腎臟病盛行率與死亡率的空間差異，底色為死亡率的相對危險性，西部的兩個圈圈代高盛行率地區，東半部的箭頭代表高死亡率的地區

病代碼的改變，在 2008 年從原本的 ICD-9 異動成爲新的 ICD-10，整體的死亡率因此大幅下降（圖十一），因此本研究仍然只分析後面 2008-2012 年的死亡率時空分佈，以避免因爲疾病代碼改變所造成推論的偏差。結果顯示，死亡率相對危險性較高的地區反而是在東部與盛行率較高的西南部地區不太一樣（圖十二），這有可能反應了潛在醫療資源不足，而造成的健康不平等情況發生。

三、預期效益

本整合型計畫的預期綜合效益如下：

1. 已收集肺結核與戰後 BCG 疫苗接種之相關紀錄，將繼續就兩者之相關性進行分析。具回顧

外文二手資料得知，1950 年代的 BCG 疫苗推廣實有兩派意見，一爲 UNICEF 主張全面強制接種，二爲 WHO 主張僅就高風險家庭進行接種。此兩種意見主要的差異，在於這兩個單位對於 BCG 疫苗可能造成 TB「僞陽性」的風險顧慮不同，以及衍生之人口中控制 TB 的動員方式差異所致。台灣自 1960 年代後開始接受 UNICEF 的援助，全面對學齡兒童進行 BCG 疫苗的接種，並在統計上體現了 TB 感染率的快速下降。然而，自 2000 年起，各種醫院與健保資料陸續顯示，台灣 TB 的感染率似乎有上升的趨勢，而就患者年齡分布來看，確有不少都應該曾經接種過 BCG 疫苗。因此現有之研究重心將與另兩項子計畫延續既往的合作模式，試圖就前 BCG 時代（日治～1960 年代）與 BCG 時代（1960 年代～）做一對照性研究。

2. 推估登革熱疫情風險地圖以提供決策參考，利用空間分析方法，有效整合病例、病媒、社經等資料，探討登革熱於時間-空間上擴散風險，劃分各時期登革熱熱區與風險地圖，期望提供防疫措施有效資訊。建立疾病於時間-空間擴散探討模式，利用網格式時間-空間基礎單元，整合疫情、人口、社會經濟等資料，探討疾病擴散之成因以及其空間分布，此研究模式亦可適用於其他疾病探討其成因，以達到疾病控制之目的。
3. 透過死亡率與盛行率的推估與計算，本研究找出慢性阻塞性肺病與慢性腎臟病的聚集地區，並辨識出環境暴露（如吸菸、空氣污染）、醫療資源差異、原住民比例高的地區等因子對於健康不平等的影響，除了使用空間資訊進行生態分析外，本研究也在慢性阻塞性肺病死亡率高的屏東縣進行社區的流行病學調查，期望透過更細緻的觀察，驗證生態研究的結果，並發現地區性的危險因子，以上的研究成果，除了可提供公共衛生單位介入時瞭解地區的疾病熱區與危險因子外，也可進而在高發生率地區或高死亡率地區進行長期的追蹤研究，對於國人的健康將有助益。