

## 生物統計研究專題

李揚<sup>1,2,3</sup> 趙青<sup>4</sup> 馬雙鴿<sup>5,6</sup>

1. 中國人民大學應用統計科學研究中心
2. 中國人民大學統計學院
3. 中國人民大學統計諮詢研究中心
4. Merck Research Laboratories
5. 美國耶魯大學生物統計系
6. 太原理工大學數學學院統計系

\* 本文（“生物統計的研究進展與挑戰”）原刊載於《統計研究》33 卷 6 期，本專題略作修改，已經作者同意刊登於生統 eNews

- 李揚，男，33 歲，2010 年畢業於中國人民大學統計學院，獲經濟學博士學位，現為中國人民大學統計學院副教授，國際統計學會推選會員，中國人民大學統計諮詢研究中心主任，北京生物醫學統計與資料管理研究會副秘書長。研究方向為生物統計、決策與預測。
- 趙青，女，26 歲，2015 年畢業於美國耶魯大學生物統計系，獲生物統計系博士學位，現擔任 Merck Research Laboratories (MRL) 生物統計師。研究方向為藥物研發試驗設計、統計分析和決策分析，以及癌症生物標記研究和疾病預測。
- 馬雙鴿，男，38 歲，2004 年畢業於美國威斯康辛大學統計學系，獲統計學博士學位，現為美國耶魯大學生物統計系終身教授。研究方向為生物統計、衛生經濟學與癌症研究等。

#### 【第四部分：存活資料分析研究】

存活分析的研究對象是樣本的存活時間（從規定的觀察起點到某一特定終點事件出現的時間長度）。存活分析起源於第二次世界大戰武器裝備的可靠性研究中，但是在大量醫學研究和臨床試驗資料中得到了廣泛應用與快速發展。隨著研究發展，存活分佈的估計從常見的指數分佈、Weibull 分佈、Gompertz 分佈、浴盆風險（Bathtub-Hazard）分佈、逆高斯（Inverse Gaussian）分佈等有母數模型，逐漸向更加靈活的半母數、無母數模型發展。其中 Kaplan-Meier 無母數估計和 D.R. Cox 提出的 Cox 比例風險（Proportional Hazards）迴歸模型的半母數估計是應用最廣泛的經典方法。經典 Cox 模型假設在任一時間點兩組自變數不同的群體的風險比例是固定的，但實際應用時情況往往十分複雜，比如自變數在試驗中不固定（譬如愛滋病研究中，病人的 CD4 細胞數量隨著每次隨訪而變化）。Fisher & Lin（1999）討論了時依（time dependent）自變數的估計及其在實際應用中的重要問題。之後越來越多的研究者將存在內部相關的存活分析和縱向資料（Longitudinal Data）分析結合起來：Tsiatis（2004）對通過潛在隨機效應將序列資訊和事件發生聯合建模這一類方法進行了總結，Ibrahim（2004）提出了用層級貝氏模型（Hierarchical Bayes model）來聯合建模。其他常用的存活模型還包括加速失效時間模型（Accelerated Failure Time Model; Louis et al., 1981）、加性風險模型（Additive Hazards Model; Lin & Ying, 1994）和比例勝算模型（Proportional Odds Model; Murphy, 1997）等。

存活分析的另一個複雜性來源於對設限資料（censored data）的處理。設限指在追蹤性研究（follow-up study）中由於某種原因未能觀察到事件最終發生時間，即確切的存活時間不可知。常見的產生設限資料的原因有失去追蹤、研究結束終點事件尚未發生、死於其他原因和由於嚴重的藥物反應而終止觀察等。在實際應用中最常見的是右設限，即終點事件發生在最後一個追蹤觀察時刻的右方。這個時間點不能丟棄也不能直接作為存活時間的觀測值，否則估計是有偏的。經典的無母數和半母數模型都能分析帶有設限資料的資料。另一類備受關注的是區間設限，包括 I 型區間設限資料（Current Status data；現時狀態資料）和 II 型區間設限資料。Robertson（1998）提出了對於現狀資料的無母數最大似估計（Maximum Likelihood），Huang（1996）的插入法（plug-in approach）、Ma（2005）的加權

Bootstrap 法和 Ma (2006) 的最小平方估計(least-squares estimation)也分別從不同角度提出了針對現時狀態資料的有母數統計推斷。II 型區間設限指終點事件發生在某兩個時間點之間，尤其當相鄰檢測點間隔大而且終點事件不易被觀察到，比如說 HIV 感染或者隱性腫瘤的發生。針對 II 型區間設限資料的方法分別有無母數模型 (Groeneboom, 1992) 和半母數模型 (Sun, 2007)。除了上述傳統存活分析模型，生物醫學實證研究中存在許多需要放寬模型假設的複雜存活分析問題。其帶來的統計挑戰是現代存活分析發展的主旋律：

(1) 無母數存活迴歸模型：在存活分析中常用的有：函數型自變數、部分線性和變係數模型。這些模型更具彈性，能包含維數更高的非線性的變數形式及變數與變數之間的交互作用。

(2) 多結局存活分析模型：這類分析的研究物件是一個病人的多個相關結局事件時間，或者由於時間（或空間）原因導致相關病人群體的事件時間。常用的多元聯合存活分佈估計方法是通過 Archimedean Copulas 建立多元聯合分佈與一元邊際分佈的聯繫 (Genest et.al, 1993)，或者針對每個群體基準風險模型建立的共享脆弱模型 (Shared Frailty Model)。

(3) 競爭風險模型：用來描述患者結局與多種風險因素的關係，通常研究者在假設的總風險下分別對特定的危險因子進行建模。Fine & Gary (1999) 提出了 Sub-distribution Hazard Function 方法，後者的迴歸係數由於針對特定風險因子的而具有更好的解釋性。

(4) 治癒模型：當觀測人群中某個次族群(subgroup)病人的存活時間達到一個很長的平穩時期時，常用的存活假設  $S(+\infty)=0$  不成立。常見的模型有兩種，一種是包含兩個不同存活曲線次族群的混合模型，另一種是有界累計風險模型。有母數模型可以直接採用最大概似估計，右設限和區間設限下的無母數模型詳見 Maller & Zhou (1996)。Logistic 迴歸常被用在半母數模型中對病人是否屬於治癒的次族群組進行建模，比如 Fang (2005) 提出的右設限 Cox 和 Ma (2011) 提出的區間設限加性風險模型。

(5) 貝氏 Cox 風險模型：貝氏 Cox 風險模型是一個熱點，其中的研究問題包括 Power 先驗分佈 (Ibrahim et al., 1999) 和對基準風險的估計和計算。貝氏

模型還被應用於多元存活分析和治癒率模型中。

除上述模型外，隨著分子生物學的發展，包含高維自變數的診斷醫學逐漸成為存活分析的熱門領域。存活時間診斷醫學相較於二元分類診斷試驗，最大的複雜性在於時間和設限。上述問題可以從基於特定時間的 Receive Operating Characteristics (ROC) 和基於時間整合的 Area Under Curve (AUC) 兩個角度展開討論。Heagerty 和 Zheng (2005) 提出了根據調整權重方法實現存活時間的密度函數和存活函數估計，逐步得到真陽性比例 (TPF)、假陽性比例 (FPF)、ROC 和 AUC 等一系列精準測量的估計，針對右設限資料的存活函數估計方法有條件 Kaplan-Meier 估計、條件光滑估計和 Dabrowska 估計。Harrell (1982)、Pencina 和 D'Agostino (2004) 討論了基於時間整合的 AUC 與和諧型統計量之間的聯繫。進一步，基於和諧性統計量的無母數變數估計 (Ma et al., 2010) 也被應用在高維自變數存活分析中。

隨著生命科學和醫藥領域的突破，臨床存活分析面臨著越來越多的新挑戰。2015 年被譽為東方諾貝爾獎的拉斯克 (Albert Lasker Award) 臨床醫學獎授予了癌症免疫療法的先驅 James P. Allison。在新一代抗癌試劑的臨床試驗中發現，病人的存活曲線相較於傳統的標準療法下存在特有的發展趨勢，亟待生物統計研究者使用更恰當有效的模型進行分析。LinDanyu 教授在 Advanced Medical Statistics (Lu, 2015) 的存活分析章節中也提出了存活研究的進一步發展方向，包括對區間設限、設限依賴、多元存活分析、變換模型 (Transformation Model) 和聯合模型在理論、方法和應用上的進一步拓展。另一方面，隨著高維度資料的普及，高維度自變數存活分析也會持續成為研究的焦點。

## 【第五部分：基因資料分析研究】

繼 2001 年人類“基因組計畫 (The Human Genome Project)”之後，美國政府又提出了“精準醫學計畫 (Precision Medicine Initiative)”，再一次將 DNA 測序和基因組技術推到了萬眾矚目的高度。回顧過去二十年的發展歷程，基因組技術對醫學與製藥領域帶來了重大的影響。人類基因組計畫的完成不僅使我們對曾經無法用傳統家系連鎖分析 (family-based linkage analysis) 檢測出疾病基因的單基因遺傳病研究有了革命性的突破，也給研究者尋找高血壓、糖尿病和癌症等多基因

疾病的遺傳基因帶來了契機。上述疾病相關基因的發現將徹底改變傳統的診斷和治療理念，使未來能夠對合適的病人在合適的時機實現個體化的有針對性治療。在整個醫學與製藥領域對“個性化醫學”這個新興治療理念歡欣鼓舞的時刻，以 DNA 測序為代表的高通量分子生物巨量資料給傳統的生物統計分析理論、方法和應用的研究帶來了前所未有的挑戰。

高通量分子生物巨量資料與傳統資料（比如人口統計資料等）最大的區別在於維度高（數千、數萬乃至數百萬維），結構複雜（多重異質性資料、複雜相關結構等），而樣本數小。因此依賴於  $p < N$  且  $N \rightarrow \infty$  假設的傳統分析方法和理論無法適用。這類問題現在普遍被稱為“ $p \gg N$ ”問題。Donoho(2000)討論了 Bellman 提出的“維度的詛咒”的同時也具體闡述了包括測度的集中性等高維度帶來的益處。針對基因資料的統計方法主要分為資料分析和假設檢定兩個方面。近幾年來，高維度資料分析最熱新方法產生最多的專題就是變數選擇，自 Tibshirani(1996)提出了 Lasso 的思想和演算法之後，Fan (2001)、Zhang (2010)、Zou(2006)相繼提出了 SCAD、MCP、自我調整 Lasso，優化了懲罰最小平方估計(penalized least squares estimation)需要滿足的稀疏性、不偏性和連續性等性質。最常用的演算法包括改進逐漸迴歸演算法的最小角迴歸 (least angle regression, LARS)，用於凹懲罰函數的局部一次近似 (LLA)，還有計算速度最快的座標速降 (CD) 演算法。在實際應用中發現，因為單個生物分子往往對複雜疾病的貢獻很小相關性較弱，引數選擇的稀疏性條件無法完全滿足生物資料的假設，模型對複雜疾病的預測能力有限，所以近年來挖掘高通量分子巨量資料的複雜結構成為了最主要的發展方向。

(1) 高維度引數選擇的統計方法逐步將變數之間的結構融合到模型的構建中。分子與分子往往是通過信號通路來行使功能的。如果能在模型中利用群組效應加大關聯的強度，有效地解決多重共線性，就可以進一步提高模型的可解釋性和預測能力。Zou (2005) 提出了 Elastic Net，這個方法的懲罰函數有效地使兩個強相關的變數趨向於有相同的係數，實現了分組效應。此後，Yuan (2006) 將多元分析的思路引入到高維度資料的結構選擇中，提出了群組稀疏性這個全新的概念。針對 Group Lasso 的塊座標 (Block CD) 下降演算法解決了座標下降演算法對群組正交的限制。由於這些方法能夠很靈活地處理並解釋變數之間的關係，

並且計算效率高，它們被廣泛應用於包括複雜性徵和複雜疾病相關的全基因組和高通量基因表達的資料分析中，找到了一些原來的的方法錯失的極具潛力的致病基因和基因組合(Ma et al., 2007; Cho et al., 2010)。Huang (2009)、Friedman(2010)又提出了同時基於變數和群組的多層次變數選擇，使模型更加靈活，參考綜述Huang (2012)。Li (2008)、Huang (2011)進一步將原始變數之間的關係利用起來，通過資料自身的拉普拉斯矩陣和鄰接矩陣來構建群組形成網路。隨著高通量資料研究資源不斷擴大，為了最大化地整合和利用這些資料提高有效性，多層次變數選擇和圖/網路指向的變數選擇被拓展到多變數迴歸 (Ma et al., 2009; Ma et al., 2011a; Ma et al., 2011b; Shi et al., 2015)，建立多個不同資料平臺複雜結構資訊之間的相互借用，抑或整合類似設計的不同實驗來彙集更多的資訊和結果。

(2) 通過概率圖模型(probabilistic graphical models)來實現對高維度資料複雜結構的挖掘。高斯圖模型(Gaussian Graphical Model)是概率模型無向圖中的一種，它處理變數與變數之間的聯繫更為直接。高斯圖模型關注於對精度矩陣  $\Theta = \Sigma^{-1}$  的估計，求解得到的稀疏精度矩陣反映了節點之間的聯繫。Friedman(2008)通過凸函數的對偶性將其轉化成求解  $W = \Theta^{-1}$  的 Lasso 迴歸問題大大提高了計算能力。Peng (2012)建立了偏相關係數與迴歸係數之間的聯繫，將圖模型轉化成了關於偏相關係數的迴歸模型，也實現了簡化運算的效果。Witten (2009)考慮了  $X^* = (X \ y) \sim N_{p+1}(\mu, \Sigma)$ ，提出了正規化共變異數矩陣(covariance matrix)迴歸問題，證明脊迴歸(ridge regression)、Lasso、還有 Elastic net 都是這類問題的特例。無向圖模型的局限性在於無法描述變數之間的因果關係，而生物分子之間的“方向性”(因果性)是一個很重要的方面，所以有向圖模型中的貝氏網路模型逐漸被應用在基因資料上。由 Wang (2013)提出的 iBag 利用分層貝氏的有向性很自然地處理了 DNA 甲基化資料和基因表達資料之間的調控關係。貝氏網路模型的優點還包括生物資訊可以通過先驗分佈加入到模型中 (Chekouo et al., 2015)，模型可以靈活地對不同來源的資訊的共用和獨立的部分分別進行估計，有效地融合多源資訊 (Ray et al., 2014)。稀疏的先驗分佈也被運用到貝氏網路模型中以達到選擇節點的效果 (Li et al., 2012)。貝氏網路的計算還是一大挑戰，它的計算量隨節點數增多呈指數上升，利用局部化概率估計來提高效率是現在的可行性方案

之一。

除了以上兩大類高維度資料分析方法之外，潛在變數模型在高維度資料中也備受關注。高維度潛在變數模型更多地關注于主成分的稀疏性，凸顯原始變數聚類形成的主要成分，其中稀疏主成分分析 (Sparse PCA)、稀疏奇異值分解 (Sparse SVD)、稀疏典型相關分析 (Sparse CCA) 和稀疏偏最小平方 (Sparse PLS) 已經被廣泛應用於生物基因表達資料和複製數變異資料的監督和無監督聚類分析 (Unsupervised Clustering Analysis)。

基因資料的統計方法另一個重要的方面是假設檢定，其中最典型的例子就是全基因組關聯分析 (GWAS)，它是在人類全基因組範圍內找出存在的序列變異，即單核苷酸多型性 (single-nucleotide polymorphism, SNP)，從中篩選出與疾病關聯顯著的遺傳標記。據美國國家人類基因組研究所對 2005-2013 發表的全基因組關聯分析的目錄的統計，發表的文章數呈指數上升，已有超過 2000 個基因位元點被報導與一個或多個複雜性徵在統計上顯著穩健相關。因為每個研究一次同時進行假設檢定的指標數量在數十萬以上，這類統計分析最大的難點是控制因多重檢定而帶來的型 I 錯誤率擴大並且同時提高分析的有效性，最常見的方法主要分為控制總型 I 錯誤率 (FWER) 和控制錯誤發現率 (FDR)。對於像全基因關聯分析這樣假設檢定數目很大的時候，控制總型 I 錯誤率相對保守，控制錯誤發現率更為廣泛地被應用，其中最廣為人知的是 Benjamini & Hochberg (1995) 提出的 BH 逐步上升程式。自從 B&H 提出採用錯誤發現率以來，多重假設檢定的發展都集中在控制錯誤發現率上，其中幾個重要的發展方向包括離散多重假設檢定 (無母數檢定如 Fisher's 精確性檢定; Gilbert, 2005)，自我調整程式 (通過原假設數量的估計來提高有效性; Storey, 2002) 和 FDR 的變形 (局部 FDR, Efron, 2005; 條件 FDR, Tsai, 2003)。還有一個應用廣泛的衍生是是否可以假設這些多重檢定是相互獨立的。這個假設在基因組關聯分析上就有可能不成立，比如說連鎖不平衡 (Linkage Disequilibrium) 現象，也就是指一個群體內不同基因座的等位基因之間的關聯不總是隨機獨立的。Benjamini (2001) 和 Wu (2008) 證明了 BH 程式在檢定正相依時也能控制錯誤發現率。Sun (2009) 進一步提出了在檢定數目很大的情況下利用正相依性來提高多重檢定的有效性，他利用隱馬可夫模型 (Hidden Markov Model, HMM) 將多重檢定與統計決策問題聯繫起來，提出

了優化局部 FDR 的由資料導向的局部顯著指數，Liu (2012) 由此啟發用高維度資料常用的圖模型來進一步探索正相依結構。

基因資料分析是近幾年生物統計最活躍的領域，充滿了機遇，也還有許多尚未解決的難題。高維度變數選擇在統計推斷上有了新的理論發展，但是由於實際應用的統計方法日趨複雜，亟待更系統更嚴格的理論證明。再者，當高維度變數之間存在強相關性時，模型的係數估計和預測依舊面臨著巨大的挑戰。新形式的生物資料的產生將繼續推動借用和整合多種複雜資料的統計分析方法的創新，與此同時也對計算的效率提出了新的要求。最後，作為一個交叉學科，提高模型的可解釋性，符合對生物複雜性的理解，才能使資料分析的結果更具有影響力。對於高維度資料的假設檢定，新一代測序技術會繼續增加一次同時進行假設檢定的指標數量，控制因多重檢定而帶來的假陽性並同時提高分析的有效性依舊是核心所在。其他一些挑戰包括針對在給定人群中中等位元基因發生頻率非常小的時候的罕見變異基因的關聯分析，考慮基因與環境的交互作用或者基因與基因的交互作用的超高維度的變數選擇等等。

## 參考文獻

1. Siegel, R. L., Miller, K. D., & Jemal, A. (2016). *Cancer statistics, 2016*. *CA: a cancer journal for clinicians*, 66(1), 7-30.
2. Cohen, M. H., Johnson, J. R., Chen, Y. F., Sridhara, R., & Pazdur, R. (2005). *FDA drug approval summary: erlotinib (Tarceva®) tablets*. *The oncologist*, 10(7), 461-466.
3. Skoulidis, F., Byers, L. A., Diao, L., Papadimitrakopoulou, V. A., Tong, P., Izzo, J., ... & Zhang, J. (2015). *Co-occurring genomic alterations define major subsets of KRAS-mutant lung adenocarcinoma with distinct biology, immune profiles, and therapeutic vulnerabilities*. *Cancer discovery*, 5(8), 860-877.
4. Machin, D., Fayers, P. (2010). *Randomized clinical trials: design, practice and reporting*. John Wiley & Sons, Oxford.
5. Schulz, K.F., Altman, D.G., & Moher, D. (2010). *CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials*. *Annals of internal medicine*, 152(11), 726-732.
6. Gottlieb, A., Menter, A., Mendelsohn, A., Shen, Y.K., Li, S., Guzzo, C., Kavanaugh, A. (2009). *Ustekinumab, a human interleukin 12/23 monoclonal antibody, for psoriatic arthritis: randomised, double-blind, placebo-controlled, crossover trial*. *The Lancet*, 373(9664), 633-640.
7. Puffer, S., Torgerson, D.J., & Watson, J. (2005). *Cluster randomized controlled trials*. *Journal of Evaluation in Clinical Practice*, 11(5), 479-483.
8. Brunoni, A. R., Valiengo, L., Baccaro, A., Zanao, T. A., de Oliveira, J. F., Goulart, A.,

- Fregni, F. (2013). *The sertraline vs electrical current therapy for treating depression clinical study: results from a factorial, randomized, controlled trial. JAMA psychiatry*, 70(4), 383-391.
9. Manaseki-Holland, S., Maroof, Z., Bruce, J., Mughal, M. Z., Masher, M. I., Bhutta, Z. A., Chandramohan, D. (2012). *Effect on the incidence of pneumonia of vitamin D supplementation by quarterly bolus dose to infants in Kabul: a randomised controlled superiority trial. The Lancet*, 379(9824), 1419-1427.
  10. Saver, J. L., Jahan, R., Levy, E. I., Jovin, T. G., Baxter, B., Nogueira, R. G., Zaidat, O. O. (2012). *Solitaire flow restoration device versus the Merci Retriever in patients with acute ischaemic stroke (SWIFT): a randomised, parallel-group, non-inferiority trial. The Lancet*, 380(9849), 1241-1249.
  11. Piaggio, G., Elbourne, D. R., Pocock, S. J., Evans, S. J., Altman, D. G., & CONSORT Group. (2012). *Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. The Journal of the American Medical Association*, 308(24), 2594-2604.
  12. Chow, S. C., & Chang, M. (2008). *Adaptive design methods in clinical trials-a review. Orphanet Journal of Rare Diseases*, 3(11), 169-90.
  13. Rosenberger, W. F., Sverdlov, O., & Hu, F. (2012). *Adaptive randomization for clinical trials. Journal of biopharmaceutical Statistics*, 22(4), 719-736.
  14. Rosenberger, W. F., Lachin, J. L. (2002). *Randomization in Clinical Trials: Theory and Practice*. New York: Wiley.
  15. Ehrhardt, S., Appel, L. J., & Meinert, C. L. (2015). *Trends in National Institutes of Health Funding for Clinical Trials Registered in ClinicalTrials.gov. Journal of the American Medical Association*, 314(23), 2566-2567.
  16. Hu, J., Zhu, H., & Hu, F. (2015). *A unified family of covariate-adjusted response-adaptive designs based on efficiency and ethics. Journal of the American Statistical Association*, 110(509), 357-367.
  17. Huang, T., & Zhu, H. (2016). *Longitudinal Covariate-Adjusted Response-Adaptive Randomization: Impact of Missing Data. Modern Adaptive Randomized Clinical Trials: Statistical and Practical Aspects*, 81, 327. Boca Raton: CRC Press.
  18. Kwong, K. S., Cheung, S. H., & Hayter, A. J. (2013). *Step-up procedures for non-inferiority tests with multiple experimental treatments. Statistical methods in medical research*, 0962280213477767.
  19. Hartley, A. M. (2015). *A Bayesian adaptive blinded sample size adjustment method for risk differences. Pharmaceutical Statistics*, 14(6), 488-514.
  20. Souza, J. P., Gülmezoglu, A. M., Vogel, J., Carroli, G., Lumbiganon, P., Qureshi, Z., ... & Neves, I. (2013). *Moving beyond essential interventions for reduction of maternal mortality (the WHO Multicountry Survey on Maternal and Newborn Health): a cross-sectional study. The Lancet*, 381(9879), 1747-1755.
  21. Dutra, L. M., & Glantz, S. A. (2014). *Electronic cigarettes and conventional cigarette use among US adolescents: a cross-sectional study. JAMA pediatrics*, 168(7), 610-617.
  22. Kotloff, K. L., Nataro, J. P., Blackwelder, W. C., Nasrin, D., Farag, T. H., Panchalingam, S., ... & Faruque, A. S. (2013). *Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. The Lancet*, 382(9888), 209-222.

23. Singh, S., Chang, H. Y., Richards, T. M., Weiner, J. P., Clark, J. M., & Segal, J. B. (2013). Glucagonlike peptide 1–based therapies and risk of hospitalization for acute pancreatitis in type 2 diabetes mellitus: a population-based matched case-control study. *JAMA internal medicine*, 173(7), 534-539.
24. Pearce, M. S., Salotti, J. A., Little, M. P., McHugh, K., Lee, C., Kim, K. P., ...& Parker, L. (2012). Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *The Lancet*, 380(9840), 499-505.
25. Titulaer, M. J., McCracken, L., Gabilondo, I., Armangué, T., Glaser, C., Iizuka, T., ...& Aguilar, E. (2013). Treatment and prognostic factors for long-term outcome in patients with anti-NMDA receptor encephalitis: an observational cohort study. *The Lancet Neurology*, 12(2), 157-165.
26. March, D., & Susser, E. (2006). The eco-in eco-epidemiology. *International Journal of Epidemiology*, 35(6), 1379-1383.
27. Wacholder, S., Chanock, S., Garcia-Closas, M., & Rothman, N. (2004). Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *Journal of the National Cancer Institute*, 96(6), 434-442.
28. Restif, O. (2009). Evolutionary epidemiology 20 years on: challenges and prospects. *Infection, Genetics and Evolution*, 9(1), 108-123.
29. Ostfeld, R. S., Glass, G. E., & Keesing, F. (2005). Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in ecology & evolution*, 20(6), 328-336.
30. Austin, P. C. (2013). The performance of different propensity score methods for estimating marginal hazard ratios. *Statistics in medicine*, 32(16), 2837-2849.
31. Imai, K., & Ratkovic, M. (2014). Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1), 243-263.
32. Maeda, I., Morita, T., Yamaguchi, T., Inoue, S., Ikenaga, M., Matsumoto, Y., ...& Tatara, R. (2016). Effect of continuous deep sedation on survival in patients with advanced cancer (J-Proval): a propensity score-weighted analysis of a prospective cohort study. *The Lancet Oncology*, 17(1), 115-122.
33. Szpiro, A. A., & Paciorek, C. J. (2013). Measurement error in two-stage analyses, with application to air pollution epidemiology. *Environmetrics*, 24(8), 501-517.
34. Alexeeff, S. E., Carroll, R. J., & Coull, B. (2015). Spatial measurement error and correction by spatial SIMEX in linear regression models when using predicted air pollution exposures. *Biostatistics*, kxv048.
35. Szpiro, A. A., Sheppard, L., Adar, S. D., & Kaufman, J. D. (2014). Estimating acute air pollution health effects from cohort study data. *Biometrics*, 70(1), 164-174.