

論文要旨

大規模電子カルテデータを用いた High-Dimensional Propensity Score の性能評価

生物統計情報学コース

49-186611

細井 宏輝

【背景】 Real World Data に対する、未測定交絡に対処する解析方法の一つとして、high-dimensional propensity score (hdPS) を用いた解析が提案されている。hdPS は、各個人の診断や処方などといった膨大な履歴データの中から、未測定交絡の代理変数として機能しうる変数を抽出することで未測定交絡に対処できると考えられている。しかし、実際に hdPS がどの程度未測定交絡に対処できているかについて、事例に基づいて検討した事例は少ない。例えば、重要な交絡因子であると考えられることの多い血液検査値は、レセプトデータなどでは得られないことが多い。

【目的】 血液検査値が取得できる大規模電子カルテデータを用いて、SGLT2 (sodium dependent glucose co-transporter 2) 阻害薬が尿路感染症のリスクを増加するかどうかを検討する中で、血液検査値を用いずに算出した hdPS によって実際に検査値が比較群間で均一化されているか、および尿路感染症のリスクが検査値を用いて算出した hdPS の結果とどのように異なるかを評価することで、hdPS によって未測定交絡の調整が行われているかを評価する。

【方法】 比較対照として、DPP-4 (dipeptidyl peptidase-4) 阻害薬を設定した。2014 年から 2018 年までの期間中に一度以上 SGLT2 阻害薬または DPP-4 阻害薬を処方された者を研究対象者とし、期間中で先に処方されたのがどちらの種類の薬剤であるかによって群の振り分けを行った。そして、検査値を用いずに、疾患の発症や薬剤の処方に関する情報から算出された hdPS を用いて SGLT2 阻害薬群と DPP-4 阻害薬群のマッチングを行った。また、一部の検査値を算出に用いた hdPS も

他に 2 種類算出して、同様にマッチングを行った。マッチング後の両群の背景因子を比較するために、それぞれの検査値について標準化平均値差を算出し、群間のバランスの指標とした。また、尿路感染症の発症をイベントとし、所属する群を説明変数とする Cox 回帰分析を行い、SGLT2 阻害薬群の DPP-4 阻害薬群に対するハザード比を推定した。

【結果】 研究対象者は 18,216 人（SGLT2 阻害薬群 1,076 人、DPP-4 阻害薬群 17,140 人）であった。検査値を用いずに算出した hdPS を用いたマッチングによって、ほとんどの検査値について標準化平均値差が減少した。その場合のハザード比は 0.97[95%信頼区間:0.77 - 1.22]で、検査値を用いて算出した hdPS を用いた場合の結果とほぼ変わらなかった。

【結論】 SGLT2 阻害薬と尿路感染症の発症との関連についての検討を行う際に、hdPS を用いたマッチングが未測定交絡である検査値についてある程度の調整を行うことができる可能性があることを示した。