

DPCデータ・レセプトデータを用いた臨床研究

康永秀生

東京大学大学院医学系研究科臨床疫学・経済学 教授

近年、様々なリアルワールドデータ (real world data, RWD) を用いた臨床研究が進んでいる。RWDには、(i) 保険データベース (DPCデータ、レセプトデータなど)、(ii) 電子カルテ・データを含むデータベース、(iii) 疾患特異的患者レジストリーなどが含まれる。これらの中でも、臨床研究への利活用が最も進んでいるのが保険データベースである。

Diagnosis Procedure Combination (DPC) データベースには、全国千数百のDPC病院から収集した年間約800万人の退院患者のデータが含まれる。全国レセプトデータ (NDB) は、全国の病院・クリニックのレセプトデータを集積した悉皆データである。本講演では、DPCデータ・レセプトデータを用いた臨床研究の現状、その利点や限界、および今後の展望について論じる。

ランダム化比較試験 (RCT) は最も内的妥当性の高い研究デザインである。しかし一般に、RCTは倫理的・費用的な制約からしばしば実施困難である。またRCTは、リアル・ワールドの臨床とはかけ離れた状況での効果比較となり、外的妥当性に乏しい。

RWD研究は、既に存在する大量の患者データを後方視的に分析するという手法であり、RCTを補完する役割を担っている。データベースによっては桁外れの症例数を確保できる。しかし、後方視的研究であるため、研究に必要なデータ項目が常に得られるとは限らず、リスク調整がいつも十分できるとは限らない。治療選択はランダムではなく、常に交絡 (confounding) の問題がつきまとう。そのため、交絡調整のための高い統計技術が必要となる。

しかし近年は、RWDにおける上記のような困難を克服した日本発の臨床研究論文が、各臨床専門領域のリーディング・ジャーナルに次々に上梓されるようになってきた。本講演ではその具体例を数多く紹介する。さらに今後のデータベース研究の利活用促進について演者の提言を示す。