

Review of Evaluating the Impact of Database Heterogeneity on Observational Study Results

Grace Yi Chen

This article systematically evaluated the impact of database selection on observational study conclusions. Many clinical studies only choose one particular database to analyze and report the result. This paper explored 10 observational databases, estimated the relative risks for 53 drug-outcome pairs under both cohort study and self-controlled case series settings. The 10 databases include administrative claims data and electronic health records from Observational Medical Outcomes Partnership project including more than 130 million patients. The authors observed that many drug-outcome pairs under both settings are substantially heterogeneous across databases. The conclusions are not consistent and sometimes even contradicting. These show that clinical studies that use observational databases can be sensitive to the choice of database.

I think the authors did a good job evaluating the impact of database heterogeneity and found that the observational studies are sensitive to the choice of database. With more health claims and EHR data available nowadays, we need to be careful about the analysis method as well as the choice of database. I agree with the author that generalizing one observational study result to other study population is hard. We need to make sure that we always mention this limitation clearly in observational studies. I know that there has been a lot of discussion about this issue and methods such as transfer learning are developed trying to solve it. In addition, meta-analysis using studies from different database could be very useful to assess the consistency of the findings. Forest plots would be a very useful tool showing heterogeneous database-specific estimates. Also, as the authors mentioned, it is still hard to make causal inference from the observational studies even after we adjust for confounders and study heterogeneity using propensity scores and random-effect models. We should be careful when interpreting the causal relationship in these results.

Questions:

1. When doing meta-analysis, we could use linear mixed models to account for the heterogeneity of the studies. However, the authors claim that the causal interpretation of random-effect meta-analytical outputs is problematic. What are the methods that could make causal inference?
2. What are the methods that could solve the generalizability problem in observational studies?