Improving Reliability, Transparency, and Reproducibility of database research without transmitting patient-level databases

Sebastian Schneeweiss, MD, ScD
Professor of Medicine and Epidemiology
Division of Pharmacoepidemiology and Pharmacoeconomics,
Dept of Medicine, Brigham & Women’s Hospital/ Harvard Medical School

Potential conflicts of interest
- PI, Harvard-Brigham & Women’s Hospital Drug Safety Research Center (FDA)
- Chair, Methods Core of the FDA Sentinel System
- Member, national PCORI Methods Committee
- Consulting in past year:
  - WHISCOR LLC, Acton Inc. (incl. equity)
  - Investigator-initiated research grants to the Brigham & Women’s Hospital from Novartis, Genentech, Boehringer Ingelheim
  - Grants/contracts from NIH, AHRQ, PCORI, FDA, IMI

Reproducibility in the life sciences

How can we make drug safety database studies more trusted?

1) Reduce bias in Healthcare database analyses:
   - Reduce confounding
   - Reduce time-related biases
   - Reduce measurement-related biases

2) Reduce investigator error
   - Have user interface and state-of-the-art workflows that ensure valid and transparent choices re. design & analysis

3) Make studies reproducible w/o sharing data
   - Have complete reporting enabling 100% reproducibility
   - Share analytic environments not data

Opioid prescribing by multiple providers in Medicare: retrospective observational study of insurance claims

To study adverse outcomes associated with prescribing of opioids by multiple providers, we estimated a beneficiary level logistic regression of the association between multiple provider prescribing and any admission related to opioid use in 2010. Admissions were identified from the linked Medicare provider.

Multiple provider prescribing was positively associated with annual rates of admission to hospital related to opioid use in both unadjusted and adjusted analyses (table 4). Among 314

2) Reduce Investigator Error
Reliable Causal Analyses

Guidance for Industry and FDA Staff
Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data

Intrinsic Study Characteristics
- Internal validity (bias)
- External validity (generalizability, transportability)
- Precision
- Heterogeneity in risk or benefit (personalized evidence)
- Ethical consideration (equipoise)

External Study Characteristics
- Timeliness (rapidly changing technology, policy needs)
- Logistical constraints (study size, complexity, cost)
- Data availability, quality, completeness

From the PCORI Methods Committee report

Basic Study Design for safety studies

Basic Design Consideration

Cohort study (case-control, case-cohort sampling)

Yes

Cohort Definition

No

Exposure Definition

Case validation necessary?

Specificity and sensitivity of measurement

Schneeweiss 2010

Define best practice workflows

Example: A typical Drug Safety Study Workflow

Dynamic study planning

Rapid Cycle Analytics

Data source

Cohort

Measures

Analysis Plan

FINDINGS

Report

Confidential

Define the analytic approach (ITT vs. AT), covariate identification period, follow-up time period, censoring etc....

UIs that guide the investigator along a problem-based workflow

Selecting data source and population in transparent and reproducible ways

Deciding on comparison group

Deciding on risk adjustment

Deciding on follow-up plan
Specify the outcome model for the primary and secondary outcomes, propensity score matching, trimming, stratifying approaches...

Avoiding obviously wrong choices will reduce heterogeneity of results

Limited heterogeneity from valid design and analysis choices

Extreme and unnecessary heterogeneity from invalid choices (eg some of OMOP's choices)*

* Susan Gruber, 2014 OMOP Symposium

3) Make Studies Reproducible
Shared cloud-based analytics

Solution for investigators that cannot store data on their cloud:

Every analysis generates a comprehensive and readable report that allows 100% reproduction of the research ...

Methods
p3-13

Results
p14-20

Appendix
p31-65

... by providing all details regarding coding and methods

Good News

- Reliability of database research can be improved through structured approaches
- Reproducibility can be achieved if
  - We completely and precisely record all choices made during design and analysis
  - We share analytic code (R, SAS, etc.)
  - We share data
- Sharing the analytic environment gets around the inability to freely share most healthcare databases