carelon

carelonresearch.com

Real-world data Integration for causal inference: benefits, costs, and case studies

Michael Grabner, PhD[1]; Edward Yu, ScD[2]; Ruth Dixon, PhD[1]; Stephan Lanes, PhD[1]; Nathan Hill, DPhil[2]
[1]Carelon Research, Wilmington, DE, USA; [2]Bristol-Myers Squibb, Summit, NJ, USA

MSR96

Background

- Causal inference from Real-World Data (RWD) is growing in importance, driven by the need for rapidly delivered and generalizable evidence to inform regulatory, payer, and patient/provider decision-making.[1,2]
- Integrating multiple sources of RWD for the same patient (e.g., claims and electronic health records) can provide deeper insights into the patient's health journey and facilitate causal research. However, integrating data also presents challenges from technical, ethical, and scientific perspectives.[3,4]
- Recent research is starting to outline the assumptions under which "big data" (consisting of multiple heterogeneous datasets) can be used for causal inference.[5]

Objectives

To summarize considerations for RWD integration aimed at causal inference and to present case studies.

Methods

- A consensus-based decision-making process was used to develop an overview of the benefits and costs of using integrated RWD for causal research.
- We summarized common biases in observational study designs[6] and how integrated RWD could affect these biases and the resulting causal effect estimates.

Confounding bias

Variables that affect both exposure and outcome induce spurious correlation[7]

Measurement bias

Observed values deviate from underlying true values[8]

Selection bias

Study sample selection is related to both exposure and outcome; in other words, the parameter of interest in the target population differs from the parameter in the available analytic sample[9]

Time-related bias

Follow-up time and exposure status are inadequately taken into account (e.g., immortal time bias[10])

- Two case studies were chosen to illuminate the trade-offs associated with using integrated RWD for causal research.

Results

Common challenges with RWD integration

Challenge	Description
Collaboration	Finding common ground across stakeholders
Security, ethics and compliance	Protecting data and patient privacy
Data interoperability	Communicating, exchanging, and using data
Data quality	Ensuring accurate integration and reliability of measures
Data availability	Establishing feasible sample sizes

Possible impacts of integrated RWD on typical biases in observational study designs

Bias	Impact
Confounding	More covariates allow for better adjustment of the causal estimate for confounding bias
Measurement	Information from multiple sources can help estimate and reduce missingness and misclassification of exposures, outcomes, and covariates
Selection	Trade-off: + Effect robustness* can be assessed across multiple databases − Higher potential for introducing selection bias (e.g., when using a subgroup based on available laboratory data)
Time-related	Neutral (design-driven)

*Robustness is defined as sensitivity of inferences to specific biases or changes in assumptions

Case study 1
External Control Arm (ECA)

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Therapeutic area:

Relapsed and refractory multiple myeloma (RRMM)

Challenge:

The pivotal KarMMA-1 trial (NCT03361748) used a single-arm design; due to the number of products previously approved, the European Medicines Agency (EMA) recommended to consider an ECA to demonstrate significant benefit.

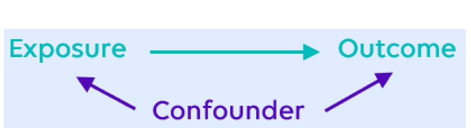
Solution:

A global, non-interventional, real-world study was set up to generate an ECA. Data from clinical sites, registries, and research databases were aggregated into a single data model and further analyzed. The EMA concluded that the efficacy results compared "favorably to those in the matched RW historical cohort as well as those reported in the literature". Limitations of the ECA included a large proportion of missing data and overlap in recruitment for the original study and the ECA at the same study centers.

For more information:

EPAR Assessment Report. EMEA/H/C/004662/0000. CHMP. EMA. 2021. Available at: https://www.ema.europa.eu/en/documents/assessment-report/abecma-epar-public-assessment-report_en.pdf [Accessed 29 Mar 2024]

Case study 2
Directed Acyclic Graphs (DAGs)



Therapeutic area:

First-line treatment for metastatic colorectal cancer (mCRC)

Challenge:

Causal inference from RWD requires many assumptions, and transparency regarding these assumptions is essential for reliable decision-making. This study created a DAG to elucidate causal relationships and applied it to an integrated real-world data source.

Solution:

Two targeted literature searches identified 94 RCTs and 22 RWD studies, from which 28 variables were extracted. These potential confounders (e.g., tumor characteristics, performance status, health care access) or colliders (e.g., data collection methods) relative to the treatment-outcome relationship were built into the DAG. Using the Healthcare Research Integrated Database (HIRD®), we identified measured and unmeasured confounders and quantified the associations between each potential confounder, the exposure (immuno-oncology therapy vs. chemotherapy), and the outcome (survival).

For more information:

Dixon R, Guzman M, Hopkins K, Lanes S, Grabner M, Hill NR, Dixon M. Treatment and outcomes in metastatic colorectal cancer: A causal study design framework. *Podium presentation at the 2024 US ISPOR Annual Meeting; Wednesday May 8, 8:45-9:45AM, session title: "Novel outcomes research data methods"*

Conclusions

- Using integrated RWD can lower methodological and resource barriers to comparative effectiveness and safety assessments.
- However, integrating data requires trade-offs regarding variable consistency, available sample size, and selection bias.
- Taking these into account will enhance the quality and, thus, the impact of evidence from observational research.

References

1. Purpura CA, Garry EM, Honig N, Case A, Rassen JA. The Role of Real-World Evidence in FDA-Approved New Drug and Biologics License Applications. *Clin Pharmacol Ther*. 2022 Jan;111(1):135-144.
2. Roberts MH, Ferguson GT. Real-World Evidence: Bridging Gaps in Evidence to Guide Payer Decisions. *Pharmacoecon Open*. 2021 Mar;5(1):3-11.
3. Graili P, Guertin JR, Chan KKW, Tadrous M. Integration of real-world evidence from different data sources in health technology assessment. *J Pharm Pharm Sci*. 2023 Jul 17;26:11460.
4. Yu Y, Jiang G, Brandt E, Forsyth T, Dhruva SS, Zhang S, Chen J, Noseworthy PA, Doshi AA, Collison-Farr K, Kim D, Ross JS, Coplan PM, Drozda JP Jr. Integrating real-world data to assess cardiac ablation device outcomes in a multicenter study using the OMOP common data model for regulatory decisions: implementation and evaluation. *JAMIA Open*. 2023 Jan 10;6(1):ooac108.
5. Bareinboim E, Pearl J. Causal inference and the data-fusion problem. *Proc Natl Acad Sci U S A*. 2016 Jul 5;113(27):7345-52.
6. Prada-Ramallal G, Takkouche B, Figueiras A. Bias in pharmacoepidemiologic studies using secondary health care databases: a scoping review. *BMC Med Res Methodol*. 2019 Mar 11;19(1):53
7. Brookhart MA, Stürmer T, Glynn RJ, Rassen J, Schneeweiss S. Confounding control in healthcare database research: challenges and potential approaches. *Med Care*. 2010 Jun;48(6 Suppl):S114-20.
8. Jonsson Funk, M., Landi, S.N. Misclassification in Administrative Claims Data: Quantifying the Impact on Treatment Effect Estimates. Curr Epidemiol Rep 1, 175–185 (2014). https://doi.org/10.1007/s40471-014-0027-z
9. Hernán MA. Invited Commentary: Selection Bias Without Colliders. *Am J Epidemiol*. 2017 Jun 1;185(11):1048-1050.
10. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. 2008 Feb 15;167(4):492-9.

Funding and Disclosures

Carelon Research received funding from Bristol-Myers Squibb for research on which parts of this poster are based.

MG, RD, SL are employees of Carelon Research (a wholly owned subsidiary of Elevance Health), which conducts health outcomes research with both internal and external funding, including a variety of private and public entities. MG is a stockholder of Elevance Health.

EY and NH are employees and stockholders of Bristol-Myers Squibb.

For additional insights on the use of integrated real-world data for causal inference, please review the materials from our workshop presented at the 2024 US ISPOR Annual Meeting, Monday May 6, 5-6 PM. Session code 150.

An introductory step-by-step guide for causal inference using observational data is available at https://www.carelonresearch.com/perspectives/white-paper-designing-rwe-studies-for-causal-inference [Accessed 29 Mar 2024]

Michael Grabner | michael.grabner@carelon.com