

AJPH LETTERS AND RESPONSES

DEBUNKING MYTHS WHILE UNDERSTANDING LIMITATIONS

Health insurance claims have been used for many years by academia and industry for health-related research. Cozad et al.¹ importantly noted that these data are often underused in public health research. We agree that myths should be debunked, but in our experience conducting comparative effectiveness research with claims data, we caution that limitations must be considered.²

The first myth is that relevant outcomes are not measured. What is relevant depends on the research question being asked. For some questions, claims data will have sufficient diagnosis or procedure codes. However, the outcomes may not be available for all questions (e.g., disease severity or progression). Proxy variables have been used but often are not validated or fail to meet quality standards.³ Therefore, variables in a database must be evaluated to determine whether the work is feasible before conducting the research, because gaps in data could lead to incorrect conclusions.

The second myth is that evidence is unreliable. Claims data can indeed be used reliably for clinical and drug development decision-making. However, both

researchers and practitioners should be aware that merging claims data with electronic medical records or US Census Bureau records is extremely challenging. When only a subset of the population can be linked, both sample size (and, hence, statistical power) and generalizability of the study are reduced. Therefore, we recommend that when key variables for the research question are missing from a data set, alternative data solutions should be fully investigated to better answer the research question before risking sample size and generalizability because of incomplete linkages.

The third myth is that claims data lack insight into practice. We believe that this is a strength of claims data. These data reflect routine clinical practice. However, one must conduct the research in accordance with accepted standards to reduce the risk of incorrect conclusions resulting from flawed study design. If claims data are used for public health investigations, we recommend that the research follow accepted best practice standards to minimize risk of error.⁴

In summary, we support the overall tenet of Cozad et al.¹ to increase the use of administrative claims for research aimed to improve the nation's health. However, we believe that this cannot be stated without acknowledging the limitations of claims data and without stating the need to adhere to accepted standards of research to minimize risk to study quality. **AJPH**

Lisa M. Hess, PhD
Katherine B. Winfree, PhD
Catherine E. Muehlenbein, MPH
Yajun E. Zhu, MS, MBA
Ana B. Oton, MD
Nicole Princic, MS
Himani Aggarwal, PhD

ABOUT THE AUTHORS

Lisa M. Hess, Katherine B. Winfree, Catherine E. Muehlenbein, Yajun E. Zhu, Ana B. Oton, and Himani Aggarwal are with Eli Lilly and Company, Indianapolis, IN. Nicole Princic is with IBM Watson Health, Cambridge, MA.

Correspondence should be sent to Nicole Princic, MS, IBM Watson Health, 75 Binney St, Cambridge, MA 01242 (e-mail: nprincic@us.ibm.com). Reprints can be ordered at <http://www.ajph.org> by clicking the "Reprints" link.

This letter was accepted January 23, 2020.
doi: 10.2105/AJPH.2020.305603

CONTRIBUTORS

All authors contributed equally to this letter.

CONFLICTS OF INTEREST

L. M. Hess, K. B. Winfree, C. E. Muehlenbein, Y. E. Zhu, A. B. Oton, and H. Aggarwal are employees of Eli Lilly and Company and hold stocks in the company. N. Princic is an employee of IBM Watson Health. The authors have no other conflicts of interest to disclose.

REFERENCES

1. Cozad M, Lindley LC, Eaker C, Carlosh KA, Profant TL. Debunking myths about health insurance claims data for public health research and practice. *Am J Public Health*. 2019;109(11):1584–1585.
2. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet*. 2002;359(9302):248–252.
3. Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources. *Value Health*. 2009;12(8):1053–1061.
4. Motheral B, Brooks J, Clark MA, et al. A checklist for retrospective database studies—report of the ISPOR Task Force on Retrospective Databases. *Value Health*. 2003; 6(2):90–97.

EDITOR'S NOTE

Cozad et al. declined to respond.

Letters to the editor referring to a recent *AJPH* article are encouraged up to 3 months after the article's appearance. By submitting a letter to the editor, the author gives permission for its publication in *AJPH*. Letters should not duplicate material being published or submitted elsewhere. The editors reserve the right to edit and abridge letters and to publish responses.

Text is limited to 400 words and 7 references. Submit online at www.editorialmanager.com/ajph. Queries should be addressed to the Editor-in-Chief, Alfredo Morabia, MD, PhD, at editorajph@qc.cuny.edu.