References

1. Kidane B, Wakeam E, Meguid RA, Odell DD. Administrative and clinical databases: general thoracic surgery perspective on approaches and pitfalls. *J Thorac Cardiovasc Surg.* 2021;162:1146-53.
2. Gliklich RE, Dreyer NA. Registries for evaluating patient outcomes: a user's guide. AHRQ publication No. 07-EHC001. Available at: <https://effectivehealthcare.ahrq.gov/products/registries-guide/research>. Accessed April 6, 2021.

See Article page 1146.

Commentary: Beyond the scope of randomized controlled trials: Navigating the sea of big data

Alexandra L. Potter and Chi-Fu Jeffrey Yang, MD

Researchers in North America and Europe are increasingly using clinical and administrative databases to answer a wide range of research questions, from assessing the efficacy, cost, and complications of interventions to studying rare diseases.¹⁻³ In addition, large-scale databases are a key resource for quality-improvement and have been used to inform quality measures for general thoracic and cardiac operations.^{4,5} While randomized controlled clinical trials (RCTs) are the gold standard for assessing the efficacy of interventions, studies using large-scale databases can complement the results of RCTs. Large clinical registries and administrative databases typically allow for analyses of sample sizes orders of magnitude larger than those of RCTs. Notably, the sizes of the available patient populations studied provide unprecedented statistical power to researchers and enable the performance of clinically meaningful subset analyses that would be difficult to do with RCTs. In addition, analyses of large-scale databases can allow for the study of rare diseases, uncommon clinical scenarios, and infrequently performed procedures. Of note, because analyses of large databases have data from real-world clinical practice, their results can be more generalizable than specialized single-institution studies or even an

Alexandra L. Potter and Chi-Fu Jeffrey Yang, MD

CENTRAL MESSAGE

In this editorial, Kidane and colleagues describe the strengths and limitations of commonly used databases and discuss how the construction and nature of each database influences its potential uses.

RCT, whose cohort may be healthier than the typical patient population. However, while large-scale databases provide many opportunities for well-designed and impactful studies, their potential uses must be mediated by a thorough understanding of their strengths and limitations.

In this issue of the *Journal*, the editorial by Kidane and Colleagues describes the strengths and limitations of several commonly used databases, including the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP), Society of Thoracic Surgeons (STS) databases, European Society of Thoracic Surgeons (ESTS) database, and linked databases in Canada and Denmark.⁶ Of note, this editorial is a companion piece to a previous paper by Subramanian and colleagues⁷ that introduced the concept of clinical versus administrative databases and described other well-known databases such as the National Inpatient Sample, Surveillance, Epidemiology, and End Results database, and the National Cancer Database. In the present editorial, Kidane and Colleagues discuss the ways in which the construction and nature of each database influence its potential uses. Importantly,

From the Division of Thoracic Surgery, Department of Surgery, Massachusetts General Hospital, Boston, Mass.

Disclosures: The authors reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

Received for publication March 30, 2021; revisions received March 30, 2021; accepted for publication March 31, 2021; available ahead of print April 20, 2021.

Address for reprints: Chi-Fu Jeffrey Yang, MD, Division of Thoracic Surgery, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114 (E-mail: cjyang@mgm.harvard.edu).

J Thorac Cardiovasc Surg 2021;162:1155-6
0022-5223/\$36.00

Copyright © 2021 by The American Association for Thoracic Surgery
<https://doi.org/10.1016/j.jtcvs.2021.03.111>

this editorial highlights the value of studies using these databases that are methodologically robust and account for inherent limitations within each database. The authors' insightful discussion shows that, if appropriately used, large-scale databases can be a powerful tool to conduct high-quality research.

In their editorial, the authors describe several considerations when using large databases to conduct thoracic research. First, the scope of data included in the database can determine the types of analyses that can be performed. For example, while the ACS NSQIP and STS databases provide high-quality data to assess certain postoperative outcomes, they are not designed to evaluate survival because data are only collected up to 30 days after the index operation.⁸ Second, the generalizability of databases to national practice patterns can vary. While some databases, such as the STS Adult Cardiac Surgery database, are considered to be representative of cardiac surgery in the United States, others, such as the STS General Thoracic Surgery Database, ACS NSQIP, and ESTS database, are less representative of national practice patterns.^{9,10} Third, the quality of data in the database relies on the method of data collection and auditing. Of note, one of the strengths of the NSQIP is that data are collected and audited by trained, certified surgical clinical reviewers who often have a background in nursing.

In addition to the aforementioned considerations, researchers should understand the differences in the population-sampling methods for each database. Importantly, the number of hospitals contributing data to the NSQIP, ESTS, and STS can change each year. As a result, studies examining population incidence cannot be performed using these databases. However, as the authors noted, some health systems with universal socialized health care, such as Denmark and Canadian provinces, have developed administrative databases with linkages to other health and nonhealth-related datasets. These databases offer a

considerable strength in that they can be used to provide population-based estimates of effect.

With the growing use of clinical registries and administrative databases in thoracic research, a thoughtful understanding of their strengths and limitations is important. Although subject to inherent limitations, high-quality studies using large databases can improve clinical care, practice management, and public policy. Thus, researchers should carefully consider the functions of each database and should choose a database that is appropriately equipped to answer their research question.

References

1. Lawson EH, Louie R, Zingmond DS, Brook RH, Hall BL, Han L, et al. A comparison of clinical registry versus administrative claims data for reporting of 30-day surgical complications. *Ann Surg*. 2012;256:973-81.
2. Bilimoria KY, Stewart AK, Winchester DP, Ko CY. The National Cancer database: a powerful initiative to improve cancer care in the United States. *Ann Surg Oncol*. 2008;15:683-90.
3. Ruffini E, Detterbeck F, Van Raemdonck D, Rocco G, Thomas P, Weder W, et al. Thymic carcinoma: a cohort study of patients from the European Society of Thoracic Surgeons database. *J Thorac Oncol*. 2014;9:541-8.
4. Seder CW, Magee MJ, Broderick SR, Brown LM, Blasberg JD, Kozower BD, et al. The Society of Thoracic Surgeons General Thoracic Surgery database: 2019 update on outcomes and quality. *Ann Thorac Surg*. 2019;107:1302-6.
5. Badhwar V, Rankin JS, Thourani VH, D'Agostino RS, Habib RH, Shahian DM, et al. The Society of Thoracic Surgeons Adult Cardiac Surgery database: 2018 update on research: outcomes analysis, quality improvement, and patient safety. *Ann Thorac Surg*. 2018;106:8-13.
6. Kidane B, Wakeam E, Meguid RA, Odell DD. Administrative and clinical databases: general thoracic surgery perspective on approaches and pitfalls. *J Thorac Cardiovasc Surg*. 2021;162:1146-53.
7. Subramanian MP, Hu Y, Puri V, Kozower BD. Administrative versus clinical databases. *J Thorac Cardiovasc Surg*. June 24, 2020 [Epub ahead of print].
8. Bhagat R, Bronsert MR, Ward AN, Martin J, Juarez-Colunga E, Glebova NO, et al. National analysis of unplanned readmissions after thoracoscopic versus open lung cancer resection. *Ann Thorac Surg*. 2017;104:1782-90.
9. Allen MS, Blackmon S, Nichols FC, Cassivi SD, Shen KR, Wigle DA. Comparison of two national databases for general thoracic surgery. *Ann Thorac Surg*. 2015;100:1155-61; discussion 1161-2.
10. Tong BC, Kim S, Kosinski A, Onaitis MW, Boffa DJ, Habib RH, et al. Penetration, completeness, and representativeness of the Society of Thoracic Surgeons General Thoracic Surgery database for lobectomy. *Ann Thorac Surg*. 2019;107:897-902.