

Administrative and clinical databases: General thoracic surgery perspective on approaches and pitfalls



Biniam Kidane, MD, MSc,^{a,b,c} Elliot Wakeam, MD, MPH,^d Robert A. Meguid, MD, MPH,^{e,f} and David D. Odell, MD, MMSc,^g for Thoracic Surgery Outcomes Research Network (ThORN) Inc

Database research has become so common that it has essentially become the lingua franca of clinical research. Since the vast majority of studies using databases are observational in nature, various statistical methods are employed to control for the inherent biases in such nonrandomized data. However, even the most sophisticated causal inference technique cannot overcome limitations that are inherent to the nature of the data and how they are collected. Most databases are created with a particular purpose, which leads to limitations in the ways they can be used for research. Organizations and professional societies have created databases for national and international benchmarking and research, the archetypical example being the Society of Thoracic Surgeons (STS) cardiac surgery database, which was subsequently expanded to include general thoracic surgery and pediatric cardiac surgery. These databases all have some element of “pay to play”; they depend on subscription fees paid by the institutions that comprise the membership and contribute data. As such, they cannot be considered population-based, as they only represent the cases contributed to them. However, in general, they are clinical, outcomes-focused, high-quality, and reliable and in certain instances can be used for high-quality policy, clinical epidemiology, and quality improvement work.

Thus, we posit that databases are often created to serve 1 of 2 fundamental functions: (1) research to answer questions and (2) benchmarking/quality. In this paper, we discuss various examples of different types of big clinical and administrative databases; more importantly, we discuss how the



Left: Biniam Kidane, MD, MSc (top), Robert A. Meguid, MD, MPH; right: Elliot Wakeam, MD, MPH (top), David D. Odell, MD, MMSc

CENTRAL MESSAGE

Databases are created to serve 1 of 2 fundamental functions: (1) research and (2) benchmarking/quality. Their construction and nature affects the extent to which they can accomplish these functions.

This Invited Expert Opinion provides a perspective on the following paper: *J Am Coll Surg*. 2009 Nov;209(5):551-6. <https://doi.org/10.1016/j.jamcollsurg.2009.08.008>; *J Am Coll Surg*. 2016 Oct;223(4):551-557.e4. <https://doi.org/10.1016/j.jamcollsurg.2016.06.393>.

See Commentaries on pages 1154 and 1155.

From the ^aSection of Thoracic Surgery, Department of Surgery, Rady Faculty of Health Sciences, and ^bDepartment of Community Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada; ^cResearch Institute in Oncology and Hematology, Cancer Care Manitoba, Winnipeg, Manitoba, Canada; ^dSection of Thoracic Surgery, Department of Surgery, University of Michigan, Ann Arbor, Mich; ^eSurgical Outcomes and Applied Research Program, Department of Surgery, and ^fAdult and Child Center for Health Outcomes Research and Delivery Science, University of Colorado School of Medicine, Aurora, Colo; and ^gSurgical Outcomes and Quality Improvement Center, Department of Surgery, Northwestern University Feinberg School of Medicine, Chicago, Ill.

Received for publication Nov 11, 2020; revisions received March 7, 2021; accepted for publication March 13, 2021; available ahead of print March 19, 2021.

Address for reprints: Biniam Kidane, MD, MSc, GH611—Health Sciences Centre, 820 Sherbrook St, Winnipeg, Manitoba, R3A 1R9 Canada (E-mail: bkidane@hsc.mb.ca).

J Thorac Cardiovasc Surg 2021;162:1146-53
0022-5223/\$36.00

Copyright © 2021 by The American Association for Thoracic Surgery
<https://doi.org/10.1016/j.jtcvs.2021.03.057>

construction and nature of said databases affects the extent to which they can accomplish these fundamental functions. This paper is meant as a companion piece to that of Subramanian and colleagues, which focused on introducing the concept of clinical versus administrative data and described some of the commonly used databases. Thus, readers should refer to that paper for descriptions of commonly used databases such as National Inpatient Sample, Surveillance, Epidemiology, and End Results, and the National Cancer Database.¹

AMERICAN COLLEGE OF SURGEONS NATIONAL SURGICAL QUALITY IMPROVEMENT PROGRAM (ACS NSQIP)

The ACS NSQIP evolved out of the Veterans Health Administration Surgical Quality Improvement Program initially in cardiac surgery beginning in 1987 and a broader noncardiac surgical patient population beginning in 1991.² Between 2002 and 2004, a grant from the Agency for Healthcare Research and Quality facilitated limited study in 14 academic hospitals.³ In 2005, the ACS adopted NSQIP, making it available nationally and internationally.⁴ The ACS NSQIP was conceived as a quality-improvement program. However, the resulting dataset, the ACS NSQIP Participant Use Data File (PUF), has become one of the gold standard databases for surgical outcomes analysis in addition to the STS Adult Cardiac Surgery Database.⁵ Participating hospitals prospectively contribute data to the ACS NSQIP on a randomly sampled selection of patients undergoing both inpatient and outpatient operations in 9 surgical specialties. Trained, certified surgical clinical reviewers, typically with a background in nursing, collect preoperative (60 variables), intraoperative (12 variables), and postoperative data, including 22 outcomes up to 30 days after the index operation. Data are sampled from more than 3000 different procedures defined by a Current Procedural Terminology (CPT) code list, based on a standardized sampling protocol over 8-day cycles. A limited number of high-frequency procedures, including laparoscopic cholecystectomy and laparoscopic appendectomy, can be collected during each cycle, ensuring representation from other procedures. Three different data-collection protocols exist for participating hospitals—the Essentials program, which abstracts between 13% and 16% of operations performed by surgeons identified as 1 of 9 specialties (general, orthopedic, vascular, thoracic, plastic reconstructive, urologic, gynecologic, otolaryngologic, and neurosurgical. cardiac surgery is not specifically abstracted); the Procedure Targeted program, which allows participating programs to annually select CPT codes of which to collect a 100% sample; and Rural program, for smaller hospitals, collecting 100% of all procedures.⁶⁻⁸

Data Collection and Access

Surgical clinical reviewers are encouraged to contact patients and their families within 90 days of surgery to complete follow-up data. They are also encouraged to contact the local surgeon champion or surgical staff to answer questions about individual cases. Other resources, such as online help and chat groups, exist to answer further questions about nuances of data abstraction. The ACS provides risk-adjusted reports to participating institutions quarterly. Each report contains 12 months' worth of data. Reports are delivered approximately 6 months after completion of

the rolling 12-month cycles. Thus, the report provides data that are between at least 6 and 18 months old. Data are audited, surgical clinical reviewers undergo periodic recertification to ensure adherence to standardized data abstraction protocols, and annual updates to data abstraction protocols are disseminated among participating hospitals periodically and at the annual meeting associated with the ACS NSQIP, the ACS Quality and Safety Conference. Each year's worth of data is released at the end of the following year.

Data Elements

Thirty-day outcomes measured by the ACS NSQIP are mortality, length of stay, discharge destination, unplanned related readmission, acute renal failure requiring dialysis or hemofiltration, bleeding requiring transfusion, cardiac arrest requiring cardiopulmonary resuscitation, deep incisional surgical-site infection (SSI), deep venous thrombosis/thrombophlebitis requiring treatment, organ/space SSI, pneumonia, progressive renal insufficiency, prolonged intubation >48 hours, pulmonary embolism, myocardial infarction, systemic inflammatory response syndrome/sepsis/septic shock, stroke/cerebrovascular accident with neurologic deficit, superficial SSI, unplanned intubation, unplanned reoperation, urinary tract infection, and wound disruption. During earlier years, the ACS NSQIP also collected outcomes on graft/prosthesis/flap failure, peripheral nerve injury, and coma >24 hours.

It is possible for each measured complication to be present before surgery. An indicator variable, "present at time of surgery—PATOS," is included in the PUF. There are variations in data completeness for the different fields, and there exist a small amount of data that appears incorrect. An example of the latter is a male patient coded as being pregnant at time of surgery.⁹

Data Structure

The total database, 2005-2018, includes around 8 million records at the time of writing. Data reported includes CPT codes and *International Classification of Diseases* codes, a series of preoperative comorbidities, and preoperative laboratory values. Every year, the ACS NSQIP modifies the list of included procedures slightly. Thus, there is gradual migration in the >3000 different CPT codes included in the ACS NSQIP. The annual PUF sample size has increased from 152,490 records from 121 hospitals in the 2005-2006 PUF to 1,020,511 records from 722 hospitals in the 2018 PUF.¹⁰ Our analysis of the frequency of distribution of the procedures in the ACS NSQIP PUF combined from 2009 to 2018 shows that this 10-year cohort of 6.9 million records is composed of 48.6% general surgery, 20.3% orthopedic surgery, and less than 10% each of the remaining specialties. The least represented specialty is thoracic surgery, constituting 1.2% of records (n > 81,000).

Definitions of the outcomes measured in the ACS NSQIP PUF are periodically updated. Notable changes include the inclusion of cardiac surgery, podiatry, and interventional radiology in the 2005 PUF, dropped in subsequent years. During 2012, the outcome of unplanned readmission was added. This includes an indicator for whether or not this was related to the index operation in the opinion of the surgical clinical reviewer.

Limitations

The ACS NSQIP dataset does have specific limitations in structure, analytic feasibility, and data capture that should be considered when reviewing or planning studies. First, the standard sampling frame collects data from a specific number of cases performed on specific dates for each participating hospital to ensure broad sampling across the institution. While this provides a standardized sampling strategy, it is not truly randomized at the case level and may potentially incorporate some bias. As this sampling strategy was designed to provide accurate assessments of surgical quality at the hospital level, the data are less well suited for the examination of variation at the provider level or within specific patient populations. Indeed, procedures that are less commonly performed (ie, pneumonectomy) may be undersampled and subject to significant bias. In some specific circumstances, this may be offset by participation in additional operation-specific registries developed from the Procedure Targeted ACS NSQIP PUFs, to facilitate disease- and surgeon specialty-specific analyses.⁸ These include 11 different procedures, such as esophagectomy, pancreatectomy, and colectomy, but not lung resections.⁸ Indeed, many hospitals have chosen to participate in these modular programs to address specific quality improvement needs locally. Of note, cardiac surgical procedures are not captured in the ACS NSQIP dataset.

The construction of the dataset does not allow use of some commonly applied statistical techniques. Data are organized by a random patient identifier at the PUF level. ACS policy does not allow identifiable information regarding the contributing institution or provider to be made available to the researcher. Thus, analyses using techniques such as clustering by the institution or provider to account for unmeasured confounding are not possible within the NSQIP PUF data unless specific arrangements are made with the ACS NSQIP governing body whereby they conduct site-identified data analysis and provide extramural researchers with deidentified results.¹¹ However, individual institution's data from before submission to the ACS do include patient identifiers, and the ACS has provided a mechanism for central data analysis in some specific instances. Further, to decrease the possibility of patient identification, the data set does not include information on the date or time of the case beyond the year of the dataset,

making time-sensitive analyses more granular than annual not possible.

Finally, it is important to consider the scope of the patient-level data available when considering projects using the ACS NSQIP PUF. The strength of the data lies within the robust, audited collection of postoperative outcomes data. However, these data are limited to a 30-day period from the index operation, and the dataset does not include longitudinal information regarding patient outcomes. Since count of days from surgery to adverse outcomes is reported, temporal analysis such as timing between surgery and readmission within 30 days of surgery is possible.¹² However, the dataset is ill-suited to examine questions when end points such as survival or recurrence of the disease state are important. The ACS also continues to modify variables year over year, and it is important to understand whether specific outcome variables or important covariates made undergone any of these definitional changes over time, as these may make some questions difficult or even impossible to examine over time. For example, in 2012 the definition of the outcome "transfusion" changed from transfusion greater than 4 units of packed red blood cells to any transfusion of packed red blood cells. Finally, the dataset does not contain some functional information or thoracic-specific outcomes, such as preoperative pulmonary function and postoperative air leaks, which are important to risk stratification and understanding of perioperative complications in thoracic surgery. The ACS NSQIP is piloting a collection of patient-reported outcomes; however, these are currently not collected on a wide scale.

In summary, the ACS NSQIP PUF has become one of the gold standard prospective clinical databases in the United States for analysis of postoperative adverse outcomes. The database has a large sample size and rigorous methodology for data inclusion. Analysis may be performed on subsets of the data based on patient disease or operation. Common postoperative complications are reported, and use of the specialty-specific registries allows for analysis of more granular data. Limitations of the ACS NSQIP PUF include incomplete data capture, as it is a data sample at most hospitals; that the smallest surgeon specialty represented in it is thoracic surgery, and cardiac surgery is not included; broadly applicable outcomes without some thoracic-specific outcomes; outcomes limited to 30-days postoperatively without longer-term data; and limited functional data particular to thoracic surgery.

SOCIETY OF THORACIC SURGEONS (STS)

Commonly regarded as the flagship national clinical repository for cardiothoracic surgery, the STS database is composed of 3 specialty-specific databases: the adult cardiac surgery database (ACSD), the general thoracic surgery database (GTSD), and the congenital heart surgery database.⁵ The original adult cardiac database arose from the

STS in response to concerns about data quality and the methodology of the Health Care Financing Administration's 1986 comprehensive report of identifiable hospital mortality rates. First made available to STS members in 1989, the ACSD now covers more than 95% of adult and pediatric cardiac surgery programs in the United States. The hallmarks of the STS ACSD are detailed clinical data (including detailed comorbidities, concurrent therapies, and physiological parameters), risk-adjusted quality indicators for common cardiac operations, and in-hospital/30-day outcomes such as a suite of complications and other outcomes. Voluntary public reporting for ACSD participants began in 2010.¹³ Given the near-complete penetrance of this database in the United States, it can be considered nationally representative for cardiac surgery in the United States.¹⁴

The general thoracic component was not formally implemented until 2002. As of 2013, the GTSD includes more than 800 cardiothoracic and general surgeons. Given the relatively smaller volumes of general thoracic procedures compared with cardiac operations, reporting periods are based on a 3-year window and results are provided to GTSD members semiannually. Like the ACSD, the GTSD includes risk-adjustment models using demographics, comorbidities, and physiologic parameters and focuses on short-term end points of length of stay, mortality, and morbidity. An important feature of STS is that individual institutions are responsible for the data collection and reporting process. The substantial cost associated with self-reporting is one of the principal barriers to complete penetrance among across all surgical providers, especially smaller-volume institutions. The STS databases have traditionally been operative databases with i/short term (ie, in-hospital or 30-day) follow-up. However, they can be linked with other databases such as Medicare, which increases the power of such combined databases to examine longer-term outcomes that can be more useful for policy making and comparative outcome studies. For example, Cowper and colleagues¹⁵ used such a linked database to examine cost outcomes for lung cancer resection in the United States, finding lower costs for minimally invasive approaches and earlier-stage disease. However, it should be noted that these linkage studies require special agreements with the STS to accomplish, and linked datasets are generally not available to users.

True to its benchmarking goal, one important output from the data included in the GTSD is the risk adjustment models and performance metrics that have been developed using the database. With the quarterly reports that are provided to institutions are star ratings, ranging from 1 to 3, for composite outcome measures for lobectomy and esophagectomy. These star ratings are generated at the program level; average performance is assigned 2 stars, and statistically better performance is designated as 3 stars.

The accuracy of the STS-GTSD compares favorably with most other surgical databases. An independent firm audits 5% of randomly selected participant sites for lobectomy cases reported in the GTSD, and across 10 sites and 559 cases, data accuracy approached 95%. While quality of data in GTSD appears reliable, generalizability to national practice patterns is commonly criticized. LaPar and colleagues¹⁶ compared esophagectomy outcomes between the GTSD, National Inpatient Sample, and NSQIP databases. In-hospital mortality rates varied between the 3 databases—3.2% in GTSD compared with 2.6% in NSQIP and 6.1% in the National Inpatient Sample. Perioperative mortality, defined as death before discharge or within 30 days of an operation, is reportedly 2.2% within GTSD, compared to 4.1% within the National Cancer Institute's Surveillance, Epidemiology, and End Results–Medicare linked database.

Certainly, the GTSD represents a selected group of high-performing centers in general thoracic surgery, and ongoing efforts to expand the capture of the database should increase its generalizability in future, particularly with respect to smaller programs and those serving historically underserved populations. Furthermore, the capture and data variables are focused on elective general thoracic surgery, ie, lobectomy and esophagectomy, and may not translate to all areas of general thoracic surgery such as lung transplantation. Despite these concerns, the GTSD provides unique opportunities to execute detailed clinical epidemiology investigations using highly accurate data. For example, Taylor and colleagues¹⁷ identified patients with marginal pulmonary function through the GTSD undergoing resection and assessed rates of complication and survival following lobectomy for cancer. Adjusting for other factors, marginal pulmonary function was not found to be an independent predictor of major complication or early mortality. For carefully selected patients with marginal pulmonary function, the authors concluded that lobectomy remains a viable surgical option to treat NSCLC. Crabtree and colleagues¹⁸ studied the accuracy of clinical staging for T2N0 esophageal cancer and determined the rates of up-staging and down-staging with surgery. The group found that 26% of patients were down-staged, whereas nearly one half were up-staged at the operation. Because T2N0 is a threshold for induction therapy, the group concluded that the accuracy of preoperative clinical staging should be further refined.

EUROPEAN SOCIETY OF THORACIC SURGEONS (ESTS)

The ESTS Database, which is similar in structure to the STS database, was established in 2001 as an attempt to collect data on thoracic surgical procedures for lung cancer from ESTS members. The initial number of participating centers was limited, but the number of included cases has steadily increased over the years. The database is free of

charge for any ESTS member, web-based, prospective, periodically maintained, and updated. The last report in 2017 included more than 110,000 patients from 240 units, 145 of which contributing with more than 100 cases.^{19,20} Like the STS data, it is not strictly speaking population-based, as cases are contributed only from partner institutions. The ESTS database, at this time, is not used in public reporting initiatives in Europe for thoracic surgery. The core dataset (Lung Cancer Registry) was progressively implemented and expanded to satellite databases dealing with other thoracic diseases and procedures. We examine 2 of these satellite registries—the thymic and lung neuroendocrine registries—as examples of clinical database subsets for collecting and analyzing relevant data that can inform clinical practice and are primarily oriented toward clinical epidemiologic research questions rather than policy-based research, which requires a true population-based registry.

Thymic Registry

The Thymic Registry was introduced in 2013 with the intent to provide an up-to-date, online prospective platform for the collection of the claims and electronic medical record data of patients with thymic tumors among the ESTS members.^{19,20} The Thymic Registry is supervised by the ESTS Thymic Working Group and by the ESTS Database Committee, and it is supported and periodically maintained by KData Clinical, the official platform of the ESTS Registry. The ESTS Thymic Working Group was approved and instituted by the ESTS Council in 2010 with the intent of creating an interest group of people in the ESTS dealing with thymic tumors. After the publication of a survey among the ESTS thymic community about the current practice in the management of thymic tumors, the first major initiative was to set up a thymic retrospective database open to all interested ESTS institutions. The group has thus far included more than 2000 patients treated between 1990 and 2010 from 35 institutions. Included variables are shown in Table 1.

Neuroendocrine Registry

In 2012, the ESTS launched a new working-group, specifically dedicated to lung neuroendocrine tumor (NETs), with the aim to develop modern knowledge on biology and behavior of these tumors. A dedicated retrospective database was at first developed and sent to all interested ESTS centers.²¹ Thus far, more than 2000 resected NET cases have been collected. The creation of this database, similar to the Thymic Registry, was motivated by the inherent limitations in general databases for the study of specific diseases and treatment.

Research Using the ESTS Registries

The ESTS database has focused heavily in clinical epidemiology. The ESTS lung neuroendocrine database, for

TABLE 1. ESTS thymic minimum dataset

Sex (male/female)
Date of birth
Date of surgery
Paraneoplastic associated syndromes (MG, etc)
Final pathologic diagnosis (thymoma, thymic carcinoma, NETT)
WHO histology
Thymic carcinoma histology
NETT histology
Pathologic Masaoka stage
Final pathologic resection status (R0, R1, R2*)
Pathologic TNM stage (IASLC/ITMIG staging, 8th Ed. TNM)
Outcome at discharge (alive/dead)
Chemotherapy (intent, date initiated)
Radiotherapy (intent, date initiated)
Date of death or last follow-up date
Vital status

These are the minimum standard data fields to be completed for data analysis. *MG*, Myasthenia gravis; *NETT*, neuroendocrine thymic tumor; *WHO*, World Health Organization; *TNM*, tumor–node–metastasis; *IASLC*, International Association for the Study of Lung Cancer; *ITMIG*, International Thymic Malignancy Interest Group. *R0, complete resection; R1, incomplete resection (microscopic); R2, incomplete resection (macroscopic).

example, has been the source of a number of clinically relevant investigations on the treatment and outcomes of lung neuroendocrine malignancy.²² The Thymic Database, likewise, has been used to examine the role of lymph node dissection, margin status, and adjuvant treatment, among other clinical questions related to thymic malignancy.²⁰ The ESTS thymic database played a fundamental role in the development of the 8th edition of the Tumor–Node–Metastasis classification of thymic tumors under the coordination of the International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee.²³

Likewise, the neuroendocrine registry has been used effectively to examine the treatment of lung NETs. Filosso and colleagues²⁴ examined 5-year overall survival rates for atypical carcinoids and large cell NETs and found that they were 77% versus 28% ($P < .001$), respectively. They were able to demonstrate aggressive behaviors and risk factors for recurrence, which can help guide treatment. Other investigations have focused on clinical epidemiology in the treatment and outcomes for lung NETs.

Limitations of the STS and ESTS Registries

Various limitations influence the way in which these databases can be used for research. First, it is important to consider the population sampling methods for any databases so that a researcher can make proper inferences about how the data reflects population incidence. Neither the STS nor the ESTS databases are “population-based” in that they

only enroll patients who have been input into the database rather than some representative sample. Therefore, they are less useful for policy work seeking to examine population incidence. Second, the variables that are included limit how researchers can control for various biases in their analyses. For example, functional status is not well recorded in either the ETST or the STS database. Furthermore, of critical importance are socioeconomic and race variables, which are only crudely captured in the data. Lung function is a critical variable for comparative outcome studies involving thoracic surgery, and although the STS database routinely records forced expiratory volume in 1 second and diffusion capacity for carbon monoxide, other measures for borderline patients such as walk tests, V/Q results or cardiopulmonary exercise testing is not recorded. While in the majority of cases these tests are not done, they may significantly alter operative management when required. These limitations are crucial to consider for researchers attempting to adjust for comorbidity and to study surgical decision-making. Operative procedures in the STS database are also perhaps not ideally coded, and they rely on billing codes to define procedures. Although not ideal, at the level of a large database these groupings are a practical way to collect data and provide enough granularity for risk adjustment at the hospital level, although certainly some complexity of procedural detail is lost in these coding schemes. Lastly, outcome definitions vary between databases. The STS database is most commonly used for research on surgical complications but the complication definitions were not standard between the STS and the ESTS databases. To this end, recent efforts have been made to standardize definitions for comparative research. Recent work has focused on comparing US and European postlobectomy outcomes, demonstrating the utility of these efforts to standardize these databases.²⁵

LINKED DATABASES FROM SOCIALIZED MEDICINE SETTINGS

Some health systems with universal socialized health care have developed robust health administrative databases with established linkages across a spectrum of other health and nonhealth-related datasets. These deidentified databases tend to be truly population-based because they are able to reliably capture essentially every person in that population. As a result of this feature and also others related to reliable linkages, these databases address some of the limitations noted in previous sections of this paper. A major strength of all databases on this type, regardless of the comprehensiveness of linkages, is that they can provide truly population-based estimates of effect.

Denmark

The quintessential example of such databases linked databases exist in Nordic countries such as Denmark. Each person living in Denmark is assigned a 10-digit unique

identifier by the Civil Registration System (either at birth or on immigration).²⁶ This serves as an effective method for linking data about an individual across many medical and nonmedical databases. It also provides for a reliable method of assuring long-term follow-up until death (or emigration out of Denmark). There is rich variety of databases and they have been grouped into administrative and health databases. Administrative databases include the well-known ones like Vital Statistics (birth, death, etc), hospital admissions or emergency visits but they also include less commonly captured data such as laboratory results and national prescription registry data.²⁶ There is also a wealth of individual-level information about other important determinants of health that are not usually reliably captured in most databases. For example, the Integrated Database for Labour Market Research contains detailed data about education, income, employment, and housing.²⁶ It bears noting that this captures actual/specific data about an individual's income rather than reporting the median income for a particular census tract (which is a proxy used for income in many other databases).²⁶⁻²⁸ Health databases are purely medical registries created for specific purposes and may include biobanks for certain diseases. The Danish Cancer Registry is a well-known disease-specific registry. There can also be procedure-based registries (eg, In Vitro Fertilization registry). A major source of rich data are investigator-initiated cohorts that can then be linked and thereby create potential for nested cohort studies with essentially perpetual long-term follow-up. These cohorts can be based on follow-up of a specific group (eg, the Diet, Cancer and Health Study, which has been following >50,000 patients free of cancer since 1993) or ad hoc population-based surveys on lifestyle factors.²⁹ In essence, this can allow traditionally cross-sectional studies (ie, survey) to take on longitudinal aspects when linked with other databases. An enviable aspect of the health registries in Denmark is that, in addition to robust linked biobanking data, there is also capture of all pathology specimens since inception of the National Pathology Registry in 1997.^{30,31}

Limitations

A major limitation is the lack of data granularity in some elements, which has been a well-described limitation of population-based research. In general, data validity and accuracy are considered quite robust for data elements like Vital Statistics, socioeconomic factors, and medications. Accurate capture of these latter 2 elements is unusual in many databases and would be considered a major strength of the Danish health data. Like many datasets based on discharge data abstraction, however, the validity and accuracy of hospital records and cause of death are more variable and less reliable. To address such limitations, the government undertakes regular validation studies.³²

Although among the most robust of datasets, there are limitations in the available data. Home care and rehabilitation services are not captured reliably in any dataset; thus, this presents a major limitation in assessing the impact of disease and treatments on new functional measures of outcome, such as “days alive and out of hospital” or “institution-free days.”³³ Another major limitation is that primary care data are not captured, and thus this represents a blind spot in this otherwise-robust linked database system.

Canadian Provinces (Ontario and Manitoba)

Similar linked databases exist in Canada; however, they apply only at the provincial level as a consequence of the nature of provincially based health care delivery in Canada. In Canada, each province is the single-payer providing publicly funded universal health insurance. The most notable are the data repositories held and operated under the auspices of the Institute for Clinical Evaluative Sciences in Ontario and the Manitoba Centre for Health Policy in Manitoba.^{27,28} Similar to Danish registries, they contain individual, record-level data for all people who have a provincial health insurance number; this represents virtually everyone in the province, as that is the mechanism to access publicly-funded universal health care.^{27,28} There is a notable exception that will be described in the limitations section to follow. Linkage is facilitated through the use of either a 9-digit (Manitoba) or 10-digit (Ontario) unique identifier for every individual.^{27,28} There is also a rich linkage to a broad range of administrative and health registries in these databases, much like that described for Denmark. A major and unique strength of Manitoba Centre for Health Policy data is that it includes linkages to the criminal justice system and thus can provide even more robust and complete assessment of the nonmedical determinants of health.²⁶ In contrast to Danish registries, these Canadian datasets include information on home care and rehabilitation services.^{27,28}

Limitations

A major limitation is the lack of data granularity in some elements and similar issues arise as described for the Danish data. A specific example related to the Canadian data is that, although data linkages exist with the provincial cancer registries, data granularity related to cancer staging is suboptimal. Recent investigator-initiated efforts have attempted to address this for gastroesophageal cancers, among others.³⁴ Another major limitation is related to the complete capture of people in the population. Although health care is generally funded provincially, aspects of health care for Indigenous peoples are funded federally and thus there is loss of data capture in those aspects of care that are funded through the provincial health insurance number.

CONCLUSIONS

In summary, we discussed various examples of different types of big clinical and administrative databases. Important details about data provenance were described; more importantly, we discussed how the construction and nature of said databases affects the extent to which they introduce bias and prevent possibility of statistical adjustment, even with the most sophisticated causal inference techniques at our disposal.

Conflict of Interest Statement

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.

William G. Henderson, PhD, MPH, and Karl E. Hammermeister, MD, provided historical insight into Veterans Health Administration Surgical Quality Improvement Program and American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP). Michael R. Bronsert, PhD, MS, provided calculations presented on ACS NSQIP.

References

1. Subramanian MP, Hu Y, Puri V, Kozower BD. Administrative versus clinical databases. *J Thorac Cardiovasc Surg.* June 24, 2020 [Epub ahead of print].
2. Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, et al. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. *Ann Surg.* 1998;228:491-507.
3. Khuri SF, Henderson WG, Daley J, Jonasson O, Jones RS, Campbell DA Jr, et al. Successful implementation of the Department of Veterans Affairs' National Surgical Quality Improvement Program in the private sector: the Patient Safety in Surgery study. *Ann Surg.* 2008;248:329-36.
4. Birkmeyer JD, Shahian DM, Dimick JB, Finlayson SR, Flum DR, Ko CY, et al. Blueprint for a new American College of Surgeons: National Surgical Quality Improvement Program. *J Am Coll Surg.* 2008;207:777-82.
5. Shahian DM, Fernandez FG, Badhwar V. The Society of Thoracic Surgeons National database at 30: honoring our heritage, celebrating the present, evolving for the future. *Ann Thorac Surg.* 2019;107:1259-66.
6. Henderson WG, Daley J. Design and statistical methodology of the National Surgical Quality Improvement Program: why is it what it is? *Am J Surg.* 2009;198(5 Suppl):S19-27.
7. Hall BL, Richards K, Ingraham A, Ko CY. New approaches to the National Surgical Quality Improvement Program: the American College of Surgeons experience. *Am J Surg.* 2009;198(5 Suppl):S56-62.
8. American College of Surgeons. ACS National Surgical Quality Improvement Program. Available at: <https://www.facs.org/quality-programs/acs-nsqip>. Accessed October 2, 2020.
9. Moore HB, Juarez-Colunga E, Bronsert M, Hammermeister KE, Henderson WG, Moore EE, et al. Are pregnant patients undergoing general surgery subject to increased morbidity and mortality? *JAMA Surg.* 2015;150:637-43.
10. American College of Surgeons. ACS NSQIP participant use data file. Available at: <https://www.facs.org/quality-programs/acs-nsqip/participant-use>. Accessed November 1, 2020.
11. Aasen DM, Wiedel C, Maroni P, Cohen ME, Meng X, Meguid RA. Association of hospital altitude and postoperative infectious complications after major operations. *High Alt Med Biol.* 2019;20:421-6.
12. Bhagat R, Bronsert MR, Ward AN, Martin J, Juarez-Colunga E, Glebova NO, et al. National Analysis of Unplanned Readmissions after thoracoscopic vs. open lung cancer resections. *Ann Thorac Surg.* 2017;104:1782-90.

13. Shahian DM, Grover FL, Prager RL, Edwards FH, Filardo G, O'Brien SM, et al. The Society of Thoracic Surgeons voluntary public reporting initiative: the first 4 years. *Ann Surg.* 2015;262:526-35; discussion 533-535.
14. Jacobs JP, Shahian DM, He X, O'Brien SM, Badhwar V, Cleveland JC Jr, et al. Penetration, completeness, and representativeness of The Society of Thoracic Surgeons Adult Cardiac Surgery database. *Ann Thorac Surg.* 2016;101:33-41; discussion 41.
15. Cowper PA, Feng L, Kosinski AS, Tong BC, Habib RH, Putnam JB Jr, et al. Initial and longitudinal cost of surgical resection for lung cancer. *Ann Thorac Surg.* October 5, 2020 [Epub ahead of print].
16. LaPar DJ, Bhamidipati CM, Lau CL, Jones DR, Kozower BD. The Society of Thoracic Surgeons General Thoracic Surgery database: establishing generalizability to national lung cancer resection outcomes. *Ann Thorac Surg.* 2012;9:216-21; discussion 221.
17. Taylor MD, LaPar DJ, Isbell JM, Kozower BD, Lau CL, Jones DR, et al. Marginal pulmonary function should not preclude lobectomy in selected patients with non-small cell lung cancer. *J Thorac Cardiovasc Surg.* 2014;147:738-44.
18. Crabtree TD, Kosinski AS, Puri V, Burfeind W, Bharat A, Patterson GA, et al. Evaluation of the reliability of clinical staging of T2N0 esophageal cancer: a review of the Society of Thoracic Surgeons database. *Ann Thorac Surg.* 2013;96:382-90.
19. Ruffini E, Falcoz PE, Guerrero F, Filosso PL, Thomas P, Novoa N, et al. The European Society of Thoracic Surgeons (ESTS) thymic database. *J Thorac Dis.* 2018;10(Suppl 29):S3516-20.
20. Ruffini E, Guerrero F, Brunelli A, Passani S, Pellicano D, Thomas P, et al. Report from the European Society of Thoracic Surgeons prospective thymic database 2017: a powerful resource for a collaborative global effort to manage thymic tumours. *Eur J Cardiothorac Surg.* 2019;55:601-9.
21. Filosso PL, Falcoz PE, Solidoro P, Pellicano D, Passani S, Guerrero F, et al; ESTS Lung Neuroendocrine Working-Group Participating Centers*. The European Society of Thoracic Surgeons (ESTS) lung neuroendocrine tumors (NETs) database. *J Thorac Dis.* 2018;10(Suppl 29):S3528-32.
22. Filosso PL, Ferolla P, Guerrero F, Ruffini E, Travis WD, Rossi G, et al. European Society of Thoracic Surgeons Lung Neuroendocrine Tumors Working-Group Steering Committee. Multidisciplinary management of advanced lung neuroendocrine tumors. *J Thorac Dis.* 2015;7(Suppl 2):S163-71.
23. Deterbeck FC, Stratton K, Giroux D, Asamura H, Crowley J, Falkson C, et al. The IASLC/ITMIG Thymic Epithelial Tumors Staging Project, proposal for an evidence-based stage classification system for the forthcoming (8th) edition of the TNM classification of malignant tumors. *J Thorac Oncol.* 2014;9(9 Suppl 2):S65-72.
24. Filosso PL, Rena O, Guerrero F, Moreno Casado P, Sagan D, Raveglia F, et al; ESTS NETs-WG Steering Committee. Clinical management of atypical carcinoid and large-cell neuroendocrine carcinoma: a multicentre study on behalf of the European Association of Thoracic Surgeons (ESTS) Neuroendocrine Tumours of the Lung Working Group†. *Eur J Cardiothorac Surg.* 2015;48:55-64.
25. Seder CW, Salati M, Kozower BD, Wright CD, Falcoz PE, Brunelli A, et al. Variation in pulmonary resection practices between The Society of Thoracic Surgeons and the European Society of Thoracic Surgeons General Thoracic Surgery databases. *Ann Thorac Surg.* 2016;101:2077-84.
26. Schmidt M, Schmidt SAJ, Adelborg K, Sundbøll J, Laugesen K, Ehrenstein V, et al. The Danish health care system and epidemiological research: from health care contacts to database records. *Clin Epidemiol.* 2019;11:563-91.
27. Kidane B, Jacob B, Gupta V, Peel J, Saskin R, Waddell TK, et al. Medium and long-term emergency department utilization after oesophagectomy: a population-based analysis. *Eur J Cardiothorac Surg.* 2018;54:683-8.
28. Smith M, Roos LL, Burchill C, Turner K, Ostapuk T, Towns D, et al. Health services data: managing the data warehouse: 25 years of experience at the Manitoba Center for Health Policy. In: Sobolev B, Levy A, Goring S, eds. *Data and Measures in Health Services Research. Health Services Research.* Boston, MA: Springer; 2016:1-26.
29. Tjønneland A, Olsen A, Boll K, Stripp C, Christensen J, Engholm G, et al. Study design, exposure variables, and socioeconomic determinants of participation in diet, cancer and health: a population-based prospective cohort study of 57,053 men and women in Denmark. *Scand J Public Health.* 2007;35:432-41.
30. State Serum Institute. The Danish National Biobank. Available at: www.danishnationalbiobank.com. Accessed November 1, 2020.
31. Erichsen R, Lash TL, Hamilton-Dutoit SJ, Bjerregaard B, Vyberg M, Pedersen L. Existing data sources for clinical epidemiology: the Danish National Pathology Registry and Data Bank. *Clin Epidemiol.* 2010;2:51-6.
32. Division of Health Planning at C.F. Møller on behalf of The Danish National Board of Health. Projekt vedrørende datakvalitet i Landspatientregistret [Project concerning data quality in The Danish national patient registry]. Copenhagen: C.F. Møller-Healthcare Service; 2004 [in Danish].
33. Jerath A, Austin PC, Wijeyesundera DN. Days alive and out of hospital: validation of a patient-centered outcome for perioperative medicine. *Anesthesiology.* 2019;131:84-93.
34. Gupta V, Levy J, Allen Ayodabo C, Amirazodi E, Davis L, Li Q, et al. Population Registry of Esophageal and Stomach Tumours in Ontario (PRESTO): protocol for a multicentre clinical and pathological database including 25,000 patients. *BMJ Open.* 2020;10:e032729.

Key Words: database, comparative effectiveness research, statistics