Methodological Quality of Surgical Mortality Studies Using Large Hospital Databases

A Systematic Review

Cécile Payet, MPH,*† Jean-Christophe Lifante, MD, PhD,†‡ Matthew J. Carty, MD,§
Muriel Rabilloud, MD, PhD,¶|| and Antoine Duclos, MD, PhD*†§

Objective: To review the methodology employed in surgical mortality studies to control for potential confounders.

Summary Background Data: Nationwide hospital data are increasingly used to investigate surgical outcomes. However, poor data granularity and coding inaccuracies may lead to flawed findings.

Methods: We conducted a systematic review in accordance with the PRISMA statement in 6 major journals (*NEJM*, *Lancet*, *BMJ*, *JAMA*, *Medical Care*, *Annals of Surgery*) using PubMed from its inception until December 31, 2014. Two reviewers independently reviewed citations. Using a predesigned data collection form, we extracted information about study aim and design, data source, selected population, outcome definition, patient and hospital adjustment, statistics, and sensitivity analyses. The methodological quality of studies was assessed based on 5 criteria and explored over time.

Results: Among 89 included studies from 1987 to 2014, 54 explored surgical mortality determinants, 13 compared surgical procedure effectiveness, 13 evaluated the impact of healthcare policy, and 9 described outcome trends for specific procedures. A total of 89% (n = 79) of studies did not describe population selection criteria at patient and hospital level, 64% (n = 57) did not consider secular trends, 52% (n = 46) neglected hospital clustering or characteristics, 21% (n = 19) did not perform sensitivity analyses, and 4% did not adjust outcomes for patient risk (n = 4). The percentage of studies satisfying at least 3 of these criteria increased significantly from 44% before 1999 to 52% between 2000 and 2009 and 78% after 2010 ($P = 0.008$).

Conclusions: Although methodological quality of studies has improved over time, confounder control could be improved through better study design, homogeneous population selection, the consideration of hospital factors and secular trends influencing surgical mortality, and the systematic performance of sensitivity analyses.

Keywords: hospital databases, methodological quality, surgical mortality, systematic review

(*Ann Surg* 2017;265:1113–1118)

From the *Department of Medical Information Evaluation and Research, Lyon University Hospital, Lyon, France; †Health Services and Performance Research Lab (EA 7425 HESPER), Lyon 1 Claude Bernard University, Lyon, France; ‡Department of General and Endocrine Surgery, Lyon University Hospital, Lyon, France; §Center for Surgery and Public Health, Brigham and Women's Hospital—Harvard Medical School, Boston, MA; ¶Department of Biostatistics, Lyon University Hospital, Lyon, France; and ||Biometry and Evolutionary Biology Laboratory (CNRS, UMR5558), Lyon 1 Claude Bernard University, Villeurbanne, France.

The authors report no conflicts of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.annalsofsurgery.com).

Reprints: Cécile Payet, MPH, Pôle Information Médicale Evaluation Recherche des Hospices Civils de Lyon, 162, avenue Lacassagne, Lyon, F-69424, France. E-mail: cecile.payet02@chu-lyon.fr.

ISSN: 0003-4932/16/26506-1113
DOI: 10.1097/SLA.0000000000002119

Nationwide hospital databases are frequently used to explore time trends in surgical outcomes, identify factors associated with better care, compare procedural effectiveness, or measure the impact of health policies.[1] These large datasets are readily available in most industrialized countries, allowing population-based investigations of routine surgical practice at low cost. However, there are several precautions to adopt when using these data for research. Hospital administrative data is typically generated for billing purposes. As a consequence, studies based on this data may be flawed by the lack of accuracy and high variability in data collection depending on coder motivation.[2–4] Furthermore, insufficient data granularity or unrecorded variables may limit outcome adjustment for specific confounders.[5]

Surgical mortality has been widely utilized to depict quality of care in hospitals because it is a critical outcome for patients and a reliably accurate metric in administrative databases.[6] However, several factors may influence mortality independent of surgical care delivery, which may limit a fair interpretation of its variations. When comparing surgical mortality between predefined groups, systematic differences can exist in the absence of randomization.[5] Risk of death can be heavily influenced by the interplay of characteristics specific to patient, procedure, surgeon, operating room team, and hospital.[7] Furthermore, patients treated and care provided within a particular hospital tend to be more similar than those in another hospital, with the result being that observed outcomes for different individuals cannot be regarded as independent due to clustering effect.[8] Finally, the confluence of many improvements in surgical practice with coding variation across hospitals can also influence surgical mortality over time under a secular trend.[9] Ideally, all these factors should be considered in analyses. However, many confounders are missing or poorly described in administrative databases, which may be partly resolved in research investigations through appropriate study design and control group choice.[1]

We conducted a methodological evaluation of studies investigating surgical mortality through assessment of administrative databases. Specifically, we aimed to describe study design, mortality definitions, population selection strategies, statistical modeling, and sensitivity analyses employed to control for potential confounders.

METHODS

Data Sources and Search Strategy

The systematic review was conducted in compliance with PRISMA guidelines.[10] A comprehensive search was undertaken using PubMed from its inception through December 31, 2014. The search strategy was devised with the assistance of a research librarian. The search included 3 domains of MeSH terms and keywords combined using "AND," whereas each domain was created using "OR." The first domain contained terms related to surgery and surgical procedures, the second contained terms related to mortality,

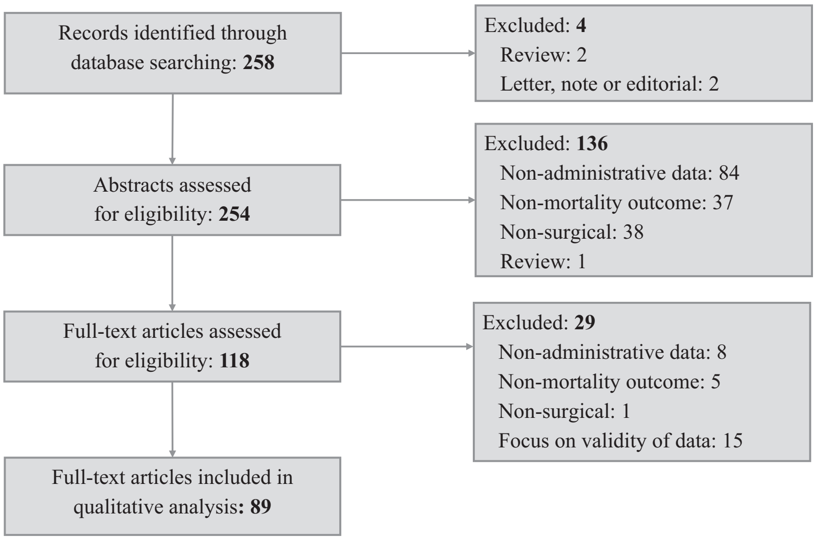

FIGURE 1. Flow chart of study selection.

and the third contained terms related to administrative databases (Appendix S1, http://links.lww.com/SLA/B162).

We focused our search strategy on the 4 generalist journals presenting the highest number of articles published in our theme (*New England Journal of Medicine*, *Lancet*, *Journal of the American Medical Association*, and *British Medical Journal*) within the category ''Medicine, General and Internal'' of the Web of Science Journal Citation Report (JCR). Based on the same rule, we also selected 2 specialist journals, 1 journal of the JCR category ''Surgery'' (*Annals of Surgery*), and 1 journal of the JCR ''Health Care Sciences & Services'' (*Medical Care*).

Study Selection

Two investigators (CP and AD) reviewed the titles and abstracts of all studies resulting from the database search. Nonoriginal articles (eg, letters, reviews, and editorials), studies not evaluating surgical mortality as well as studies using nonadministrative database were excluded. To ensure that all potentially relevant studies were considered for inclusion, every study with uncertain eligibility at title and abstract review was retained for further examination, especially when this initial review did not allow us to categorize paper as an original article based on surgical mortality from administrative data. Full-text documents were obtained for selected articles and 2 investigators (CP and AD) independently examined all studies to decide final inclusion in this review. Any disagreements were resolved by consensus.

Data Extraction

Data were extracted from the selected articles using a predesigned collection form and checked independently by 2 authors (CP and AD). Specific data points collected included: study aim and design, data source, selected population, surgical mortality definition, patient and hospital confounders, and statistical/sensitivity analyses.

Papers were classified into 4 categories depending on their major study aim, as follows: time trend analysis, outcome determinants exploration, comparative effectiveness research, or health policy impact. Time trend studies described patterns of change in surgical mortality over time. Determinant studies identified factors influencing surgical mortality at the patient or hospital level. Comparative effectiveness studies compared the relationship between different surgical procedures and postoperative mortality. Health policy studies evaluated the impact of healthcare programs on surgical mortality.

Study design was described according to comparator choice to achieve comparable groups, time trend integration, and matching scheme between patients or hospitals. Three different study designs were reported based on comparator group choice. Before-and-after studies compared outcome before and after exposure within a unique group. Studies with contemporaneous comparators compared outcomes after exposure between exposed and nonexposed groups. Difference-in-differences studies compared outcome before and after exposure between exposed and nonexposed groups.

To define surgical mortality, we distinguished intra- from extra-hospital death. We also specified the period for death occurrence from start (at hospital admission or the day of surgical procedure) until the end (at hospital discharge, 30-d, 90-d, 1-yr). Descriptions of statistical analyses included adjustment on the characteristics of patients or providers and on secular trends, model type and methods to take into account the clustering of patients within hospitals. Finally, to assess the methodological quality of each study, we defined a 5-item score. These items included 1) the presentation of selection criteria at the patient and hospital level; 2) the adjustment for patient characteristics in multivariate models; 3) the consideration of hospital characteristics or clustering in these models; 4) the control of secular trends in statistical analyses; and 5) the performance of sensitivity analyses to test the robustness of main findings.

Statistics

Characteristics of the selected studies were presented using absolute and relative frequencies for categorical variables. Continuous variables were described using mean and standard deviation or median with minimum and maximum.

To explore trends in studies' characteristics and methodological quality over time, papers were classified according to their respective aim and the year of publication. We used the Cochran–Armitage test to highlight trends between categories. Data manipulation and analyses were performed using SAS software (version 9.3; SAS Institute Inc, Cary, NC).

RESULTS

Characteristics of Studies

The search strategy retrieved 258 citations (Fig. 1). After screening titles, abstracts, and manuscripts, 89 studies were included

TABLE 1. Study Characteristics

	n = 89	%
Aim		
Time trend	9	10
Determinant exploration	54	61
Comparative effectiveness	13	15
Health policy	13	15
Data and population		
Study period in years, median (min–max)	6 (1–18)	—
No. of patients, median (min–max)	55,830 (502–9,474,879)	—
No. of hospitals, median (min–max)*	531 (39–2691)	—
Outcome†		
In-hospital mortality	44	49
In/out-hospital mortality	44	49
Not specified	6	7
Patient confounders considered	85	96
Demographic	82	96
Comorbidities	74	87
Emergency	38	45
Socioeconomic	32	38
Year of procedure	30	35
Procedure type	18	21
Primary diagnosis	16	19
Admission day	3	4
Procedure complexity	1	1
Hospital confounders considered	31	35
Volume	20	65
Hospital status	16	52
Bed size	13	42
Rural/urban	13	42
Geographic area	12	39

*Forty-nine studies did not report hospital number.
†Studies could be used in-hospital mortality and in/out-hospital mortality.

for analysis (Appendix S2, http://links.lww.com/SLA/B162). Among these studies, 18 were published through 1999, 31 between 2000 and 2009, and 40 from 2010. Overall, 74 studies originated from the United States and used primarily Medicare data or nationwide inpatient samples, 11originated from the United Kingdom and were mainly based on hospital episode statistics, 4 originated from Canada, and 1 from France. Regarding study aim, 61% (54/89) of studies explored surgical mortality determinants while 39% (35/89) focused on time trend analysis, procedural effectiveness, or health policy impact (Table 1, Fig. 2).

Study populations included procedures focused on the following organ systems: digestive (53 studies), cardiothoracic (42), orthopedic (15), respiratory (13), urinary (6), nervous (1), and genital (1). Four studies did not discriminate procedure type and therefore presumably included interventions spanning multiple organ systems.

There was no consensus in the definition of surgical mortality among reviewed papers. Surgical mortality was not defined in 6 studies. Twenty-four studies failed to identify whether surgical mortality was defined from the time of hospital admission or from the time of surgical procedure. Among 44 studies investigating in-hospital mortality, 77% (34/44) defined mortality as deaths occurring before hospital discharge, 25% (11/44) within 30 days, and 1% (1/44) within 14 days. It was not specified whether in-hospital mortality included all postoperative deaths or only focused on those occurring within surgical wards. Among 44 studies investigating in/out-hospital mortality, 45% (20/44) defined mortality as deaths occurring within 30 days, 9% (4/44) within 90 days, 11% (5/44) within 1 year, and 41% (18/44) of studies measured deaths before hospital discharge or within 30 days for patient discharged before that date.

Patient confounders were considered in 85 studies while hospital confounders were considered in 31 studies. Among 74 studies with adjustment for patient comorbidities, 38% (28/74) used

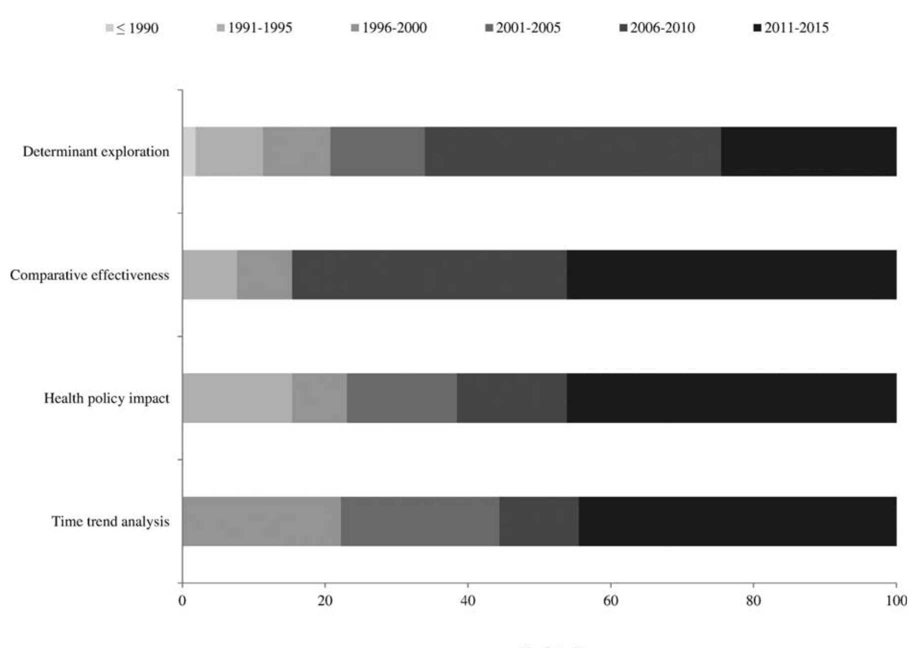

FIGURE 2. Percentage of studies according to study aim and publication year.

TABLE 2. Study Aim and Design

Study Aim	Time Trend Analysis n = 9 (%)	Determinant Exploration n = 54 (%)	Comparative Effectiveness n = 13 (%)	Health Policy Impact n = 13 (%)
Comparator				
Contemporaneous group	0 (0)	54 (100)	13 (100)	8 (61)
Before and after	0 (0)	0 (0)	0 (0)	1 (8)
Difference-in-differences	0 (0)	0 (0)	0 (0)	4 (31)
Time series	9 (100)	16 (30)	3 (23)	4 (31)
Matching	0 (0)	2 (4)	3 (34)	1 (8)

the Charlson index, 31% (23/74) the Elixhauser index, and 31% (23/74) a specific selection of comorbidities. Socioeconomic status was considered in 32 studies using insurance status (53% of studies, 17/32) or other data (56%, 18/32), based on income, education, or employment corresponding to the patient's residence code.

Surgical mortality was modeled using logistic regression (76 studies), cox model (3), and Poisson regression (2). Clustering was considered in 38 studies using Generalized Estimating Equations models (58% of studies, 22/38) or mixed models (34%, 13/38) while 3 studies did not report the model used.

Design and Methodological Quality of Studies

A contemporaneous control group was used as comparator in all the determinant and comparative effectiveness studies (Table 2). Regarding the health policy studies, 8 used a contemporaneous control group, 4 executed a difference-in-differences study design, and 1 employed a before-and-after design. In addition to time trend studies, 23 studies used multiple measurements over time to consider secular trends. Propensity score matching was used in 6 studies. The matching was performed at the patient level in the determinant and comparative effectiveness studies, while hospitals were matched in 1 health policy study.

Figure 3 shows that population selection criteria at both the patient and hospital levels were not described in 79 studies, without improvement over time. Adjustment for patient characteristics was common, but 46 studies failed to consider the influence of hospital characteristics on mortality or patient clustering. However, the percentage of studies that took into account hospital factors increased from 28% (5/18) until 1999 to 42% (13/31) between 2000 and 2009, and 63% (25/40) after 2010 (Fig. 3, $P = 0.010$). Secular trends during the study period were not considered in 57 studies and there was no improvement over time. Finally, 70 studies conducted sensitivity analyses, increasing from 56% (10/18) until 1999 to 77% (24/31) between 2000 and 2009, and 90% (36/40) after 2010 (Fig. 3, $P = 0.003$). These sensitivity analyses were focused on the testing of secondary outcomes (40 studies), other statistical approaches (29), other definitions of surgical mortality (18), population selection (14), and study design (1).

Regarding the 5-item score, 1 study did not meet any item, 10 met 1 item, 23 met 2 items, 37 met 3 items, 16 met 4 items, and 2 all 5 items. The percentage of studies with at least 3 items increased significantly from 44% (8/18) until 1999 to 52% (16/31) between 2000 and 2009 and 78% (31/40) after 2010 ($P = 0.008$).

DISCUSSION

Nationwide hospital data are increasingly used to investigate surgical mortality, with a special focus on analyzing time trends, comparing procedural effectiveness, or measuring the impact of health policy initiatives. We do not necessarily have tangible explanations to explain the trend toward relative increase of those topics in

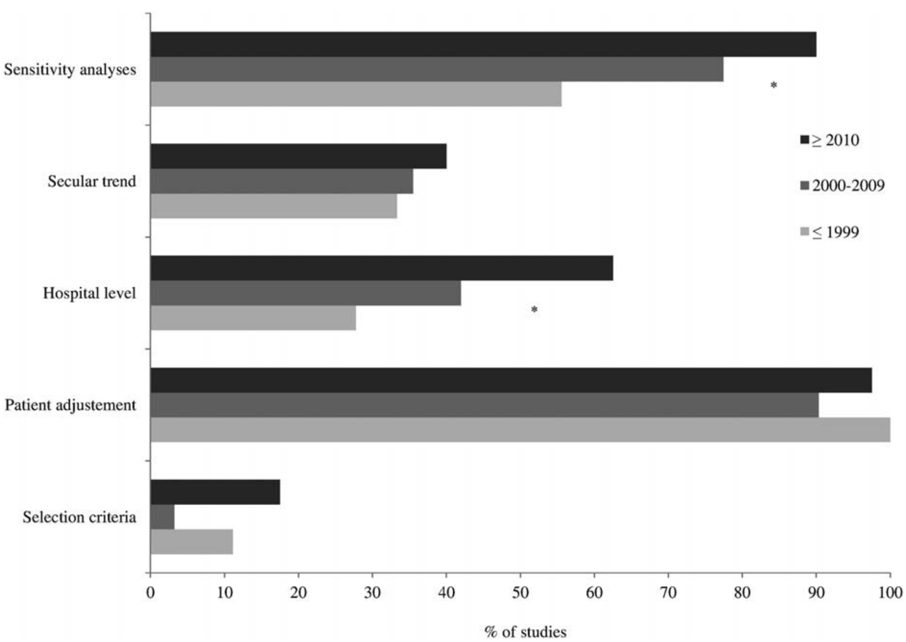

FIGURE 3. Percentage of studies according to methodological quality items and publication year. Selection criteria were at both patient and hospital levels. Hospital level corresponded to hospital adjustment or patient clustering within hospitals. *P value <0.05, Cochran–Armitage test trend.

relation to traditional research on mortality determinants. However, recent emphasis in the fields of health services research with growing funding requires the utilization of large administrative data reflecting routine practice in large populations.[11] The accumulation of nationwide data over decades now allows refinements in designing longitudinal studies with contemporary control groups for investigating real changes in surgical outcomes caused by new policies or procedures. Although the methodology of studies has improved over time, this systematic review of publications from flagship journals highlights the fact that basic quality standards have not been systematically met. Nearly all publications still fail to report both patient and hospital selection criteria to homogenize study populations. A majority of studies do not take into account the influence of hospital effect, secular trend, or coding variation that may alter data interpretation in statistical analyses. Finally, sensitivity analyses are not always performed to test the robustness of study findings. However, several improvements in methodology are possible to manage potential confounders, through appropriate study design, data selection, and analysis.

No amount of sophisticated statistical analyses can compensate for poor study design. On the one hand, using before-and-after study design to measure the association between the intervention effect and outcome within a unique group is valid only if there are no underlying time-dependent trends in outcomes unrelated to the effect of intervention.[12] On the other hand, using a study with contemporaneous control comparators is valid only if baseline outcomes are the same between exposed and nonexposed groups before the start of intervention. Opting for a difference-in-differences study design can address these issues by comparing outcomes before and after exposure to the intervention between exposed and nonexposed groups.[13] In this way, using a control group that is experiencing the same secular trends but is not exposed to the intervention allows the investigators to subtract out the background changes in outcomes if 2 assumptions are met: 1) that underlying trends in outcomes before the intervention should be the same between groups although these groups may have different outcomes at baseline; and 2) that phenomena occurring concurrent to the intervention will affect both groups equally.[14] These assumptions require multiple measurements for each group to verify time trends and initial comparability between groups and assure their preservation during the entire study period.[15] If these assumptions are not met, multiple measurement design could be used to take into account different trends.

One way to obtain group comparability is to choose well-matched samples of the original exposed and control groups, thereby reducing bias risk due to potential confounders.[16] When the study purpose is to assess the impact of a specific healthcare policy, matching hospitals can be worthwhile for controlling hospital characteristics that may drive surgical outcomes irrespective of policy implementation.[17] If a study aims at measuring the association between patient determinants and mortality or the effectiveness between surgical procedures, patient matching is of great interest;[18] specifically, patient matching can be stratified within hospitals[18] to know whether patients exposed to a new procedure have better outcomes than those treated with the usual procedure by the same surgical team, with the advantage of controlling for potential confounders related to variations across hospitals in perioperative care and data coding. When a large number of variables influence group choice, it can be appropriate to use propensity score analyses to balance groups according to distributions of measured covariates.[16,19] These variables may collectively be proxies for unobserved factors; therefore, the high-dimensional propensity score method has been suggested to identify and prioritize covariables within the investigational database.[5,19,20] Instrumental variables can also be used to control measured and unmeasured confounders. This method employs a variable called an instrument that is predictive of treatment but has no effect on the outcome, except through its influence on treatment assignment.[21]

Given the complexity of the underlying medical and socioeconomic processes that determine surgical outcomes as well as the capacity of nationwide data for exploring healthcare, a substantial uncertainty persists about how to select study populations, to design the study and to specify statistical models. Therefore, systematic sensitivity analyses can be conducted to ensure the consistency of results. Traditionally, sensitivity analyses require changing the measurement, but they can also be performed at each step of the study for testing different control groups, population of patients, surgical procedures, and statistical models.

For example, a feature of large hospital databases is their hierarchical structure, since distinct clusters of patients are treated within different hospitals. Patients admitted to a given hospital may share common characteristics and have correlated outcomes, thus violating the independence assumptions of traditional regression models. The main consequence of not taking account of this correlation is a misinterpretation of statistical tests[22] and inadequate precision of the estimates with too narrow confidence intervals.[23,24] To avoid such a pitfall, random effects models and generalized estimating equations models[23] are commonly used and could be routinely incorporated into statistical analyses or at least into sensitivity analyses. Furthermore, logistic regression is commonly used to model in-hospital mortality, but this is unsuitable when follow-up varies from 1 patient to another or the hospital discharge represents a competing risk of death. A solution to study in-hospital mortality is to employ competing risk models in survival data analyses, integrating time to death and length of hospital stay.[25,26] Finally, when surgical mortality is investigated over several years, it can be interesting to incorporate a time variable in the model because unmeasured confounders can be associated with background changes like coding variations.[12,27]

The main limitation of this review was the noninclusion of all articles published in our theme for feasibility reason. Although we were not representative of studies published in all biomedical literature, we assumed that the editorial policies of selected journals were based on high-quality standards in selecting submitted works. We therefore believe that the methodological quality of papers retained in this review is above average relative to that of the entire range of studies in our theme. A further limitation was the marked heterogeneity in the aims and populations of selected studies. A broad search strategy was used intentionally to capture a full range of studies across all surgical specialties.

In 2015,the Reporting of studies Conducted using Observational Routinely-collected health Data checklist was created as an extension to the STROBE statement to address reporting items appropriate to observational studies using routinely collected health data.[28] This checklist encourages authors to clearly state codes, algorithms, and linkages between databases used to select study populations and define exposures, outcomes, and confounders. However, it would be worthwhile adapting this checklist for the assessment of large hospital databases that gather different data granularity levels (ie, hospital, patient, stay) across several years. This review provides a thorough evaluation of methodologies employed in studies using these data and highlights several key findings for investigating surgical outcomes. In support to the Reporting of studies Conducted using Observational Routinely-collected health Data statement, confounder control could be improved through appropriate study design, homogeneous population selection, the consideration of hospital factors, and secular trends influencing surgical mortality, and the systematic performance of sensitivity analyses (Appendix S3, http://links.lww.com/SLA/B162).

REFERENCES

1. Dimick J, Greenberg C. *Success in Academic Surgery: Health Services Research.* Berlin, Germany: Springer; 2014.
2. Peabody JW, Luck J, Jain S, et al. Assessing the accuracy of administrative data in health information systems. *Med Care*. 2004;42:1066–1072.
3. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med*. 1997;127(8_Part_2):666–674.
4. Mohammed MA, Deeks JJ, Girling A, et al. Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *BMJ*. 2009;338:b780.
5. Brookhart MA, Stürmer T, Glynn RJ, et al. Confounding control in healthcare database research: challenges and potential approaches. *Med Care*. 2010; 48:S114–S120.
6. Aylin P, Bottle A, Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *BMJ*. 2007;334:1044.
7. Sittig DF, Classen DC. Safe Electronic Health Record use requires a comprehensive monitoring and evaluation framework. *JAMA*. 2010;303: 450–451.
8. Rice N, Leyland A. Multilevel models: applications to health data. *J Health Serv Res Policy*. 1996;1:154–164.
9. Lindenauer PK, Lagu T, Shieh M, et al. Association of diagnostic coding with trends in hospitalizations and mortality of patients with pneumonia, 2003-2009. *JAMA*. 2012;307:1405–1413.
10. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151:264–269.
11. Moses H 3rd, Matheson DH, Cairns-Smith S, et al. The anatomy of medical research: US and international comparisons. *JAMA*. 2015;313: 174–189.
12. Chen Y-F, Hemming K, Stevens AJ, et al. Secular trends and evaluation of complex interventions: the rising tide phenomenon. *BMJ Qual Saf*. 2016;25:303–310.
13. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA*. 2014;312:2401–2402.
14. Ryan AM, Burgess JF, Dimick JB. Why we should not be indifferent to specification choices for difference-in-differences. *Health Serv Res*. 2015; 50:1211–1235.
15. Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-experimental Designs for Generalized Causal Inference.* Boston, MA: Houghton Mifflin; 2002.
16. Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci*. 2010;25:1–21.
17. Reames BN, Scally CP, Thumma JR, et al. Evaluation of the effectiveness of a surgical checklist in medicare patients. *Med care*. 2015;53:87–94.
18. Silber JH, Rosenbaum PR, Kelz RR, et al. Examining causes of racial disparities in general surgical mortality: hospital quality versus patient Risk. *Med Care*. 2015;53:619–629.
19. Seeger JD, Kurth T, Walker AM. Use of propensity score technique to account for exposure-related covariates: an example and lesson. *Med Care*. 2007;45:S143–S148.
20. Schneeweiss S, Rassen JA, Glynn RJ, et al. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology.* 2009;20:512–522.
21. Pezzin LE, Laud P, Yen TWF, et al. Reexamining the relationship of breast cancer hospital and surgical volume to mortality: an instrumental variable analysis. *Med Care*. 2015;53:1033–1039.
22. Austi PC, Tu JV, Alte DA. Comparing hierarchical modeling with traditional logistic regression analysis among patients hospitalized with acute myocardial infarction: should we be analyzing cardiovascular outcomes data differently? *Am Heart J*. 2003;145:27–35.
23. McCulloch CE, Searle SR, Neuhaus JM. *Generalized, Linear and Mixed Models.* New York: Wiley; 2011.
24. Hubbard AE, Ahern J, Fleischer NL, et al. To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. *Epidemiology.* 2010;21:467–474.
25. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol*. 2009;170:244–256.
26. Varadhan R, Weiss CO, Segal JB, et al. Evaluating health outcomes in the presence of competing risks: a review of statistical methods and clinical applications. *Med Care*. 2010;48:S96–S105.
27. Lindenauer PK, Lagu T, Shieh MS, et al. Association of diagnostic coding with trends in hospitalizations and mortality of patients with pneumonia, 2003-2009. *JAMA*. 2012;307:1405–1413.
28. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015;12:e1001885.