Published in final edited form as:

Anesthesiology. 2015 August ; 123(2): 243–245. doi:10.1097/ALN.0000000000000691.

The importance of validation studies in perioperative database research

Mark D. Neuman, MD, MSc

Department of Anesthesiology and Critical Care, Perelman School of Medicine at the University of Pennsylvania; Department of Internal Medicine, Division of General Internal Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA; Leonard Davis Institute for Health Economics, the University of Pennsylvania, Philadelphia, PA

Over the past decade, administrative data has come to play a major role in research on perioperative care. Beyond providing a basis for studies focusing on the organization and delivery of care for patients undergoing surgery and anesthesia, insurance claims and other administrative records have been used to examine fundamental questions of clinical epidemiology and comparative effectiveness research. Yet despite the growing number of investigators who seek clinical and policy insights from administrative data, efforts to validate such data for identifying specific diagnoses remain relatively uncommon, although important examples do exist within¹ and outside of^{2,3} the perioperative setting. In this issue of *Anesthesiology*, Daniel McIsaac and colleagues⁴ offer quantitative insights that can help readers and investigators make better sense of the opportunities and challenges involved in using administrative data for perioperative research.

McIsaac and colleagues examine the validity of administrative data for obstructive sleep apnea (OSA). OSA is a widespread public health problem that may increase the risk of adverse events after surgery,^{5,6} and research using administrative data has played a major role to date in highlighting it as an important issue for perioperative medicine.^{7–9} Using a large clinical database from one major Canadian teaching hospital, McIsaac and colleagues identified 4,965 surgical patients who underwent a major surgical procedure between 2003 and 2012 and who also had undergone a polysomnogram at that same hospital prior to surgery. The authors linked these clinical data to national and provincial health administrative databases that included claims for inpatient, emergency department, and day-surgery care, as well as for physician services and durable medical equipment. They obtained a successful linkage for 4,353 patients, or 88% of their original cohort, and found 56% of linked patients to have a confirmed diagnosis of OSA based on polysomnogram results or documentation by a sleep medicine physician.

The authors used this linked dataset to study the accuracy of a range of case-ascertainment algorithms for obstructive OSA; these algorithms employed codes from the 9th and 10th

Corresponding Author Address: Mark D. Neuman, M.D., M.Sc., Department of Anesthesiology and Critical Care, Perelman School of Medicine at the University of Pennsylvania, 423 Guardian Drive, 1119A Blockley Hall, Philadelphia, PA 19104, Office: (215) 746-7468, Fax: (215) 349-5078, neumanm@mail.med.upenn.edu.

Conflicts of interest: None

revisions of the International Classification of Diseases (ICD), including ICD codes that have been previously used to identify patients with sleep apnea in administrative data, as well as claims for polysomnography or positive airway pressure devices. Several of the algorithms that the authors assessed demonstrated high specificity (i.e. a high proportion of people who truly lacked OSA who were correctly identified by the algorithm); however, most of the methods assessed, and particularly those that relied exclusively on ICD codes, showed low sensitivity when tested against a standard of a diagnostic polysomnogram or documentation of OSA by a sleep medicine physician (i.e. a low proportion of people who truly had OSA who were correctly identified by the algorithm). Extrapolating their findings to a general surgical population, the authors estimate that only 40% of patients identified as having OSA based on an ICD-9 or ICD-10 diagnosis code in their administrative database could be expected to actually have the condition, while 10% of patients without such a code could still be expected to have OSA.

McIsaac's study has limitations of its own. Their use of data from Canada potentially limits the relevance of their work to studies that use administrative data sources from the US, since coding practices could potentially differ across countries due to variations in the organization and reimbursement of care. Its single-center design may also limit generalizability, since coding practices may vary from institution to institution. And while they present findings for a range of ascertainment algorithms, these algorithms overlap incompletely with methods used in recent large database studies focusing on outcomes for surgical patients with and without OSA, thus limiting direct comparisons.

Yet this work offers important lessons for consumers of perioperative research that employs administrative data, as well as investigators—myself included—who rely on ICD codes to identify patients with specific health conditions. Essentially all empirical studies—regardless of the source of the data used—suffer from some degree of misclassification; in the present context, the low sensitivity of administrative claims for OSA creates the possibility that patients who actually have OSA might be incorrectly classified by common administrative data algorithms as lacking the disease. As McIsaac notes, such misclassification is often cited as being likely to bias study results towards the null hypothesis, or a finding of no effect. For example, in a study aiming to compare outcomes among patients with and without OSA, the failure to identify OSA in a subject who actually has the disease might make the two comparison groups more similar due to the presence in both groups of patients who actually have OSA. Such misclassification would work to diminish the differences in outcomes between groups that would be attributed to the presence or absence of OSA. Study results affected by such bias might be argued to be “conservative,” since the true differences between groups would be expected to be larger than those that such a study would find.

Such an argument assumes that misclassification occurs independently of other important factors (i.e. is “nondifferential”). Yet if other factors that might impact outcomes also influence an individual's likelihood of being classified as having the condition under study, such “differential” misclassification can also work to bias study findings away from the null hypothesis, potentially yielding spuriously large results.¹⁰

McIsaac's work suggests that the misclassification of OSA in administrative data may indeed be related to other important factors; administrative diagnosis codes were more likely to correctly identify OSA in patients who also had other major perioperative risk factors, such as heart disease and diabetes, and who moderate or severe versus mild OSA. Such considerations create the potential that studies relying on administrative data to identify patients with OSA could potentially overstate the overall impact of this condition on postoperative outcomes if those patients classified in the study as having OSA tended to be sicker or to have more severe disease than OSA patients overall.

To their credit, authors of studies that have employed administrative data to examine the impact of OSA on postoperative outcomes have been frank about the limitations of such data sources for health research.^{9,11} Rather than invalidating these past studies, the work of McIsaac and colleagues provides an opportunity to add nuance to their interpretation. In particular, it offers new quantitative data that can help readers and researchers to better gauge the potential extent and nature of misclassification of OSA within administrative data. Future investigators might draw lessons from this work, as well as other investigations^{12,13} that have shown substantial variability in the coding of common diagnoses over time, as to what insights (regarding OSA or other conditions) might or might not be obtainable from administrative data sources. This investigation also offers useful information that can help practitioners and researcher more correctly interpret the results of past studies by newly quantifying and explicating limitations of administrative data sources in this context that have been suspected and debated, yet not fully described. Most importantly, this work offers a strong example of the importance of well-done validation studies in perioperative database research, not only for guiding the types of questions that we choose to investigate through the lens of retrospective data, but also for making better sense of what insights such investigations might be able to convey.

Acknowledgments

STATEMENT OF SUPPORT: Support was provided from institutional/departmental sources and the National Institute on Aging, Bethesda, MD (K08AG043548)

References

1. Lawson EH, Louie R, Zingmond DS, Brook RH, Hall BL, Han L, Rapp M, Ko CY. A comparison of clinical registry versus administrative claims data for reporting of 30-day surgical complications. *Ann Surg.* 2012; 256:973–81. [PubMed: 23095667]
2. Carnahan RM, Moores KG. Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative and claims data: methods and lessons learned. *Pharmacoepidemiol Drug Saf.* 2012; 21(Suppl 1):82–9. [PubMed: 22262596]
3. Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived from ICD-9-CCM administrative data. *Med Care.* 2002; 40:675–85. [PubMed: 12187181]
4. McIsaac D, Gershon A, Wijeyesundera D, Bryson G, Badner N, Van Walraven C. Identifying obstructive sleep apnea in administrative data: a study of diagnostic accuracy. *Anesthesiology.* 2015
5. Kaw R, Chung F, Pasupuleti V, Mehta J, Gay PC, Hernandez AV. Meta-analysis of the association between obstructive sleep apnoea and postoperative outcome. *Br J Anaesth.* 2012; 109:897–906. [PubMed: 22956642]
6. Memtsoudis SG, Besculides MC, Mazumdar M. A rude awakening—the perioperative sleep apnea epidemic. *N Engl J Med.* 2013; 368:2352–3. [PubMed: 23782177]

7. Memtsoudis S, Liu SS, Ma Y, Chiu YL, Walz JM, Gaber-Baylis LK, Mazumdar M. Perioperative pulmonary outcomes in patients with sleep apnea after noncardiac surgery. *Anesth Analg*. 2011; 112:113–21. [PubMed: 21081775]
8. Spurr KF, Graven MA, Gilbert RW. Prevalence of unspecified sleep apnea and the use of continuous positive airway pressure in hospitalized patients, 2004 National Hospital Discharge Survey. *Sleep Breath*. 2008; 12:229–34. [PubMed: 18236092]
9. Mokhlesi B, Hovda MD, Vekhter B, Arora VM, Chung F, Meltzer DO. Sleep-disordered breathing and postoperative outcomes after elective surgery: analysis of the nationwide inpatient sample. *Chest*. 2013; 144:903–14. [PubMed: 23538745]
10. Rothman, KJ.; Greenland, S.; Lash, TL. *Validity in epidemiologic studies*, Modern Epidemiology. 3. Rothman, KJ.; Greenland, S.; Lash, TL., editors. Philadelphia: Lippincott, Williams & Wilkins; 2008.
11. Memtsoudis SG. In reply. *Anesthesiology*. 2014; 120:246–7. [PubMed: 24398748]
12. Lindenauer PK, Lagu T, Shieh MS, Pekow PS, Rothberg MB. Association of diagnostic coding with trends in hospitalizations and mortality of patients with pneumonia, 2003–2009. *JAMA*. 2012; 307:1405–13. [PubMed: 22474204]
13. Rhee C, Murphy MV, Li L, Platt R, Klompas M. Comparison of trends in sepsis incidence and coding using administrative claims versus objective clinical data. *Clin Infect Dis*. 2015; 60:88–95. [PubMed: 25258352]