

Identifying Obstructive Sleep Apnea in Administrative Data

A Study of Diagnostic Accuracy

Daniel I. McIsaac, M.D., M.P.H., F.R.C.P.C., Andrea Gershon, M.D., M.Sc., F.R.C.P.C.,
Duminda Wijeyesundera, M.D., Ph.D., F.R.C.P.C., Gregory L. Bryson, M.D., M.Sc., F.R.C.P.C.,
Neal Badner, M.D., F.R.C.P.C., Carl van Walraven, M.D., M.Sc., F.R.C.P.C.

ABSTRACT

Background: Health administrative (HA) databases are increasingly used to identify surgical patients with obstructive sleep apnea (OSA) for research purposes, primarily using diagnostic codes. Such means to identify patients with OSA are not validated. The authors determined the accuracy of case-ascertainment algorithms for identifying patients with OSA with the use of HA data.

Methods: Clinical data derived from an academic health sciences network within a universal health insurance plan were used as the reference standard. The authors linked patients to HA data and retrieved all claims in the 2 yr before surgery to determine the presence of any diagnostic codes, diagnostic procedures, or therapeutic interventions consistent with OSA.

Results: The authors identified 4,965 patients (2003 to 2012) who underwent preoperative polysomnogram. Of these, 4,353 patients were linked to HA data; 2,427 of these (56%) had OSA based on diagnosis by a sleep physician or the apnea hypopnea index. A claim for a polysomnogram and receipt of a positive airway pressure device had a sensitivity, specificity, and positive likelihood ratio (+LR) for OSA of 19, 98, and 10.9%, respectively. An *International Classification of Diseases, Tenth Revision*, code for sleep apnea in hospitalization abstracts was 9% sensitive and 98% specific (+LR, 4.5). A physician billing claim for OSA (*International Classification of Diseases, Ninth Revision*, 780.5) was 58% sensitive and 38% specific (+LR, 0.9). A polysomnogram and a positive airway pressure device or any code for OSA was 70% sensitive and 36% specific (+LR, 1.1).

Conclusions: No code or combination of codes provided a +LR high enough to adequately identify patients with OSA. Existing studies using administrative codes to identify OSA should be interpreted with caution. (ANESTHESIOLOGY 2015; 123:253-63)

OBSTRUCTIVE sleep apnea (OSA) has been identified as a significant issue for physicians caring for patients in the perioperative period.¹ Numerous societies provide guidelines to support the care of patients with OSA undergoing surgery,^{2,3} and the relation between OSA and perioperative outcomes has been extensively studied.⁴ Despite the large number of studies that examine OSA in the perioperative period, significant limitations exist in the quality of the current literature and in our overall understanding of the impact of OSA on surgical and anesthetic outcomes.

A 2012 meta-analysis found a significant association between OSA and adverse cardiac and pulmonary outcomes after surgery.⁴ Importantly, this review also highlighted a number of knowledge gaps and methodological limitations in the perioperative OSA literature. First, most studies are single centered, and many lack the sample size to (1) detect

What We Already Know about This Topic

- Considerable research on perioperative risks and outcomes in patients with presumed obstructive sleep apnea (OSA) uses health administrative databases and codes for OSA, but the accuracy of these codes for presence or absence of OSA is unknown

What This Article Tells Us That Is New

- In approximately 5,000 patients who underwent preoperative polysomnography, 56% met criteria for a diagnosis of obstructive sleep apnea (OSA)
- In these patients with known or excluded OSA, none of the health administrative diagnostic codes, diagnostic procedures, or therapeutic interventions by themselves or in combination identified OSA with adequately high sensitivity and specificity
- Existing studies using administrative codes to identify OSA should be interpreted with caution

This article is featured in "This Month in Anesthesiology," page 1A. Corresponding article on page 243.

Submitted for publication November 12, 2014. Accepted for publication March 6, 2015. From the Department of Anesthesiology, University of Ottawa, Department of Anesthesiology, The Ottawa Hospital, and Institute for Clinical Evaluative Sciences at uOttawa, Ottawa, Ontario, Canada (D.I.M.); Sunnybrook Health Sciences Centre, Institute for Clinical Evaluative Sciences, and University of Toronto, Toronto, Ontario, Canada (A.G.); Li Ka Shing Knowledge Institute of St. Michael's Hospital, Department of Anesthesia and Pain Management, Toronto General Hospital, Department of Anesthesia and Institute of Health Policy Management and Evaluation, University of Toronto, and Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada (D.W.); Department of Anesthesiology, University of Ottawa, and Department of Anesthesiology, The Ottawa Hospital, Ottawa, Ontario, Canada (G.L.B.); Department of Anesthesiology, Pharmacology, and Therapeutics, University of British Columbia, and Kelowna General Hospital, Kelowna, British Columbia, Canada (N.B.); and Department of Medicine, University of Ottawa, and Institute for Clinical Evaluative Sciences at uOttawa, Ottawa, Ontario, Canada (C.v.W.).

Copyright © 2015, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. All Rights Reserved. Anesthesiology 2015; 123:253-63

important differences in mortality and major morbidity; and (2) adequately control for important confounders in the OSA–outcome relation. Second, available studies lack consistent measurement of well-defined clinical outcomes. Recent commentary supports the continued existence of significant knowledge gaps regarding perioperative OSA⁵ and ongoing need for studies that provide long-term follow-up and that capture out-of-hospital events.

Population-based studies using health administrative (HA) data have the potential to address specific knowledge gaps related to perioperative OSA outcomes because these data provide a relatively inexpensive and practical means to study disease exposure, processes of care, and health outcomes at a population level. HA data studies are often large and provide the statistical power to study rare outcomes. Accordingly, numerous HA data studies of OSA have been recently published.^{6–9} Unfortunately, available studies of OSA using HA data may suffer from misclassification bias because single diagnostic codes have been used to identify people with^{6–8} or without OSA.⁹ *International Classification of Diseases* (ICD) codes are frequently used, but the diagnostic accuracy of these codes has not been assessed. Because diagnostic codes within HA data demonstrate high variation in their ability to accurately identify true disease status,^{10,11} the misclassification of patients with regard to OSA status could influence study results in important ways.

To address the potentially significant methodological issues present in existing HA data studies of perioperative OSA, we undertook a study with three objectives: (1) to measure the accuracy of methods used in the current literature to identify patients with OSA in HA data; (2) to measure the accuracy of other case-ascertainment algorithms based on a combination of diagnostic codes, diagnostic procedures (polysomnogram), and therapeutic interventions (positive airway pressure [PAP] devices) available in HA data, which we predicted would allow highly accurate identification of patients with OSA within HA databases; and (3) to describe the types of patients identified using these approaches to better understand the possible effects of misclassification on study outcomes described in the current literature. We conducted this study in the Canadian province of Ontario because high-quality, linked, population-based HA data exist, which allow identification of chronic disease prevalence cohorts,^{12–16} provide detailed hospitalization and procedural records, and validate outcome measures and longitudinal follow-up.

Materials and Methods

This is a validation study of diagnostic test accuracy using clinical data from a multihospital health sciences network as

the reference standard linked to population-based HA data. Clinical data are from The Ottawa Hospital, a 900-bed tertiary care academic health sciences network serving a population of approximately 1.2 million people.* The hospital network consists of three geographically distinct campuses, including two inpatient hospitals and a free-standing ambulatory surgical center. This investigation is reported using the Standards for the Reporting of Diagnostic Accuracy Studies (STARD initiative, appendix).¹⁷ The study was approved by the Ottawa Health Sciences Network Research Ethics Board (OHSN-REB 2008835, 20120813).

Data Sources

Our study used two distinct sets of linked healthcare databases. The first was The Ottawa Hospital Data Warehouse (OHDW), a combination of administrative and clinical data repositories for The Ottawa Hospital beginning in 1996. The OHDW captures reports of all polysomnograms done at the hospital and records all surgeries. The second set of linked databases are housed at The Institute of Clinical Evaluative Sciences, an independent research institute that holds HA data for the province of Ontario, Canada's largest province with a population of more than 12 million people. Ontario provides single-payer universal health insurance to all residents, including coverage for polysomnography and at least partial coverage for PAP devices prescribed for OSA. This study used the following databases: the Discharge Abstract Database (DAD), which provides a detailed information pertaining to hospitalizations, including diagnoses; the National Ambulatory Care Reporting System, which provides a record (including diagnoses) of all emergency room visits; the Same Day Surgery Database, which provides a record of all ambulatory surgeries; the Ontario Health Insurance Plan (OHIP) database, which records fee for service physician claims (including polysomnograms); and the Assistive Devices Program database, which records funding for all durable medical equipment including PAP devices.

Study Cohort

We used the OHDW to retrospectively identify all patients who had undergone a polysomnogram at The Ottawa Hospital between 1996 and 2012. From this cohort, we identified all patients who had surgery at our hospital between January 1, 2003 (to coincide with the introduction of the ICD, Tenth Revision, coding system at our hospital) and December 31, 2012 (the latest date for which all data sources were current). We identified only patients who underwent preoperative polysomnography before surgery to ensure that the OSA status of each patient in our reference population was known. We excluded procedures performed exclusively with monitored anesthesia care, specifically gastrointestinal endoscopy and ophthalmologic surgeries, as we felt that these patients could be significantly different from those having more invasive noncardiac surgery.

* The Ottawa Hospital: About our hospital. Available at: <https://www.ottawahospital.on.ca/wps/portal/Base/TheHospital/About-OurHospital>. Accessed January 31, 2015.

From the polysomnogram most proximal to and preceding surgery, one of two investigators (D.I.M. and G.L.B.) retrospectively abstracted the apnea hypopnea index (AHI) and the diagnosis or exclusion of OSA from the OHDW. Patients were classified with OSA if the sleep medicine physician reporting the study diagnosed the patient with OSA or if the AHI was 5 or greater. OSA severity was that graded by the sleep physician reporting the study; if no severity was reported, criteria from the American Society of Sleep Medicine¹⁸ were applied based on the AHI (AHI 5 to 14 = mild OSA; AHI 15 to 29 = moderate OSA; and AHI >30 = severe OSA). Patients diagnosed with nonobstructive sleep pathology (such as central sleep apnea) were categorized together as having an “other sleep disorder.”

Linking to Provincial Data and Creating the Analytical Dataset

Each patient record was assigned a unique, anonymized identifier and linked to provincial HA databases. We then identified diagnostic codes and claims (for diagnostic procedures and therapeutic interventions) within the provincial datasets, which we hypothesized could indicate a diagnosis of OSA (table 1). For each patient, we identified any claim—within the 2 yr before their surgery—for a diagnostic or therapeutic polysomnogram (from physician billing claims in OHIP); receipt of a PAP device (from the Assistive Devices Program database); and any ICD-9 or ICD-10 code for sleep apnea (from hospitalization data in the DAD or from physician billing claims in OHIP). The ICD-9 codes we used were the same as those used in previous HA data studies of surgical patients with OSA⁶⁻⁹; ICD-10 codes were those that identified obstructive or other sleep apneas. All ICD codes in our analytic dataset were treated separately based on the database of their origin (OHIP *vs.* DAD).

Patient age and sex were determined from DAD or Same Day Surgery Database record for the index surgery. Previously described methods were used to identify the Elixhauser comorbidities based on ICD-9 and ICD-10 codes from the DAD in the 2-yr preceding surgery.¹⁹ A Charlson comorbidity score was calculated for each patient.²⁰

Statistical Analysis

Diagnostic codes, therapeutic intervention claims, and diagnostic procedure claims were coded as binary variables (present or absent) for each patient. Case-ascertainment algorithms were based on the presence of single codes (as in previous studies⁶⁻⁹) or predetermined combinations of codes and service claims, which we hypothesized might improve accuracy. For each case-ascertainment algorithm, each patient was classified as true positive (OSA present, algorithm positive), false negative (OSA present, algorithm negative), true negative (OSA absent, algorithm negative), or false positive (OSA absent, algorithm positive).

Because we anticipated that the prevalence of OSA in our reference population would be higher than the general population (as reference individuals were identified by a physician as being at risk of OSA, hence their referral for a polysomnogram), we measured the association of each case-ascertainment algorithm with OSA by calculating the positive and negative likelihood ratios. Compared with the positive and negative predictive values, likelihood ratios are relatively impervious to disease prevalence and should, therefore, provide a more accurate representation of the ability of each case-ascertainment algorithm to predict the true presence or absence of OSA.²¹ The 95% CIs for these likelihood ratios were calculated according to the method by Simel *et al.*²² Positive and negative predictive values are

Table 1. Diagnostic Codes Used in Case-ascertainment Algorithms

Polysomnography Code	Database	Description
J696	OHIP	Diagnostic polysomnogram
J896	OHIP	Diagnostic polysomnogram
J890	OHIP	Diagnostic polysomnogram
J690	OHIP	Diagnostic polysomnogram
J898	OHIP	Diagnostic polysomnogram (incomplete) <1 h
J899	OHIP	Diagnostic polysomnogram (incomplete) 1–4 h
J990	OHIP	Diagnostic polysomnogram (incomplete) >4 h
J897	OHIP	Diagnostic polysomnogram (repeat)
J697	OHIP	Diagnostic polysomnogram (repeat)
J889	OHIP	Therapeutic polysomnogram
J689	OHIP	Therapeutic polysomnogram
Diagnostic Codes	Database	Description
G4730	DAD	ICD-10 Sleep apnea, obstructed
G4738	DAD	ICD-10 Other sleep apnea
780.5	OHIP and DAD	ICD-9 Unspecified sleep apnea
327.2	OHIP and DAD	ICD-9 Sleep apnea

DAD = Discharge Abstract Database of the Canadian Institute for Health Information; ICD = *International Classification of Diseases*; OHIP = Ontario Health Insurance Plan.

not reported as these values would be significantly biased for external populations by the high prevalence of OSA in our reference population.²³ To describe operating characteristics of various algorithms within this population, we calculated sensitivity and specificity with 95% CIs using the binomial distribution.

In subgroup analyses, the accuracy of each algorithm was evaluated separately for all patients diagnosed with OSA (*i.e.*, mild, moderate, and severe disease) as well as for prespecified subgroups including patients diagnosed with moderate and severe disease only (*i.e.*, those in whom PAP therapy is recommended), any patient with sleep-disordered breathing (*i.e.*, OSA or other sleep disorders), and categorized by inpatient *versus* ambulatory surgery. We performed further *post hoc* analyses restricted to individuals whose polysomnogram had been performed at most 3 yr before surgery to eliminate the bias that might be present due to changes in patient disease status over a long time lag between testing and surgery (a 3-yr time window was chosen because this was the approximate 75th percentile of polysomnogram surgery time lag for our population and because longitudinal studies support stability or increase in AHI for most individuals over 4- to 5-yr follow-up periods^{24–26}). Finally, we limited analyses to patients whose surgeries were in the first half of our study period (2003 to 2007) or the second half (2008 to 2012) to examine the impact of possible changes in coding patterns over time.

To allow a qualitative assessment of the effect of misclassification presented by different methods to identify OSA in HA data, we also documented the demographic details, severity of OSA, and prevalence of chronic medical conditions that may confound the OSA–outcome relation for each case-ascertainment algorithm based on the test result (*i.e.*, categorized for true positives, false positives, true negatives, and false negatives). The chronic medical conditions considered included heart disease (defined as any history of valvular disease,

congestive heart failure, arrhythmia, or history of a myocardial infarction), respiratory disease (defined as any history of chronic obstructive pulmonary disease, asthma, or chronic pulmonary disease), and diabetes (defined as a history of complicated or uncomplicated diabetes). All analyses were conducted using SAS version 9.3 for UNIX (SAS Institute, Inc., USA).

Results

Study Cohort

We identified 4,965 patients who underwent a preoperative polysomnogram (fig. 1). Of these, 4,353 patients (88%) were linked successfully to our HA databases; the other patients were unable to be linked to provincial databases because our hospital also serves patients without OHIP coverage, such as patients, mostly from Quebec, with universal health insurance separate from Ontario's or, rarely, patients with other forms of health insurance. Although we could not identify the specific reason for nonlinkage in all patients, even those in figure 1 without a valid OHIP number were likely from Quebec. OSA was diagnosed in 2,427 patients (56%) with the patients distributed evenly between mild, moderate, and severe OSA. Nonobstructive sleep pathology was diagnosed in 77 patients (2%). Forty-five percent of patients underwent ambulatory surgery with the remainder undergoing inpatient surgery (table 2). Compared with patients without the diagnosis, those with OSA were more likely to be male, diabetic, and hypertensive (table 2), which is in keeping with previous cross-sectional analyses.²⁷ Patients with OSA were less likely to receive ambulatory surgery than those without OSA.

Accuracy of OSA Case-ascertainment Algorithms

None of the diagnostic codes, diagnostic procedures, or therapeutic interventions, by themselves or in combination, identified OSA with adequately high sensitivity and

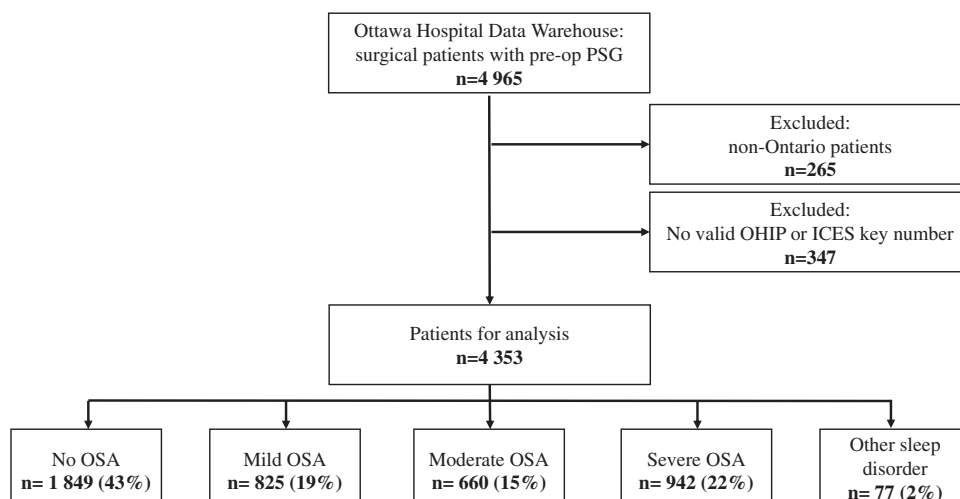


Fig. 1. Study flow. ICES = Institute for Clinical Evaluative Sciences; OHIP = Ontario Health Insurance Plan; OSA = obstructive sleep apnea; PSG = polysomnogram.

Table 2. Patient Characteristics

Characteristics	OSA	No OSA	P Value
	n = 2,427	n = 1,849	
Ambulatory (%)	961 (39.6)	996 (51.7)	<0.0001
Age (mean, SD)	56 (13)	52 (14)	<0.0001
Female (%)	895 (36.9)	1,161 (60.3)	<0.0001
Congestive heart failure (%)	112 (4.6)	47 (2.4)	0.0001
Cardiac arrhythmias (%)	64 (2.6)	37 (1.9)	0.12
Valvular disease (%)	73 (3.0)	38 (2.0)	0.032
Pulmonary circulation disorders (%)	24 (1.0)	12 (0.6)	0.19
Peripheral vascular disorders (%)	70 (2.9)	38 (2.1)	0.055
Hypertension, uncomplicated (%)	327 (13.5)	155 (8.1)	<0.0001
Hypertension, complicated (%)	6 (0.3)	3 (0.2)	0.51
Paralysis (%)	11 (0.5)	7 (0.4)	0.34
Other neurological disorders (%)	36 (1.5)	30 (1.6)	0.84
Chronic pulmonary disease (%)	108 (4.5)	71 (3.7)	0.21
Diabetes, uncomplicated (%)	790 (32.6)	327 (17.0)	<0.0001
Hypothyroidism (%)	29 (1.2)	25 (1.3)	0.76
Renal failure	89 (3.7)	42 (2.2)	0.004
Liver disease	11 (0.5)	10 (0.5)	0.76
Peptic ulcer disease (%)	11 (0.5)	11 (0.6)	0.89
AIDS/HIV (%)	3 (0.1)	0 (0)	0.12
Solid tumor without metastases (%)	194 (8.1)	132 (6.9)	0.12
Metastatic cancer (%)	29 (0.7)	28 (1.5)	0.46
Connective tissue diseases (%)	7 (0.3)	6 (0.3)	0.71
Coagulopathy (%)	30 (1.2)	21 (1.1)	0.66
Obesity (%)	173 (4.0)	64 (3.3)	<0.0001
Weight loss (%)	7 (0.3)	3 (0.2)	0.36
Blood loss anemia (%)	68 (2.8)	42 (2.2)	0.19
Deficiency anemia (%)	8 (0.3)	0 (0)	0.011
Drug abuse (%)	2 (0.1)	8 (0.4)	0.022
Psychoses (%)	2 (0.1)	6 (0.3)	0.08
Depression (%)	39 (1.6)	42 (2.2)	0.16
Charlson score (median, IQR)	0 (0.1)	0 (0.0)	<0.0001
Emergency room visit prior year (%)	1,000 (41.2)	817 (42.4)	0.42

All diagnoses derived from coded hospital abstracts.

AIDS = acquired immune deficiency syndrome; HIV = human immunodeficiency virus; IQR = interquartile range; OSA = obstructive sleep apnea.

specificity (table 3). The combination of a polysomnogram followed by receipt of a PAP device was highly specific for a true diagnosis of OSA (98%) and had the highest positive likelihood ratio (+LR) of all algorithms (10.9). The sensitivities of all diagnostic codes, by themselves or in combination, were very low. The specificity of both ICD-9 and ICD-10 codes for OSA in the DAD exceeded 97%. In contrast, the specificity of OSA diagnostic codes in OHIP was less than 40%. The combination of diagnostic codes for OSA with polysomnogram + PAP maximized sensitivity at 70% but caused the specificity and +LR to both decrease substantially. Negative LR for all algorithms were between 0.8 and 1.1.

Changes in LR, sensitivities, and specificities varied slightly among subgroups (table 4). The polysomnogram + PAP algorithm had a lower +LR in the moderate and severe OSA subgroup, whereas the +LR was minimally different in ambulatory, inpatient, and sleep-disordered breathing subgroups compared with the full population. Sensitivities and specificities for the polysomnogram + PAP algorithm were

also minimally changed in any of the subgroups. Similarly, ICD-9 and ICD-10 codes for sleep apnea applied to the DAD did not change substantially in any subgroup compared with the full population. The ICD-9 code 780.5 also produced similar LR and sensitivities/specificities in all subgroups, as did the algorithm based on polysomnogram + PAP or any ICD code. Limiting analysis to those with a polysomnogram within the 3 yr before surgery did not change LR in any marked way. Likelihood ratios did vary between time periods; however, the changes were not consistent. The polysomnogram + PAP algorithm had higher +LRs after 2007, whereas ICD-10 codes had higher +LR before 2008. ICD-9 codes performed poorly in all time periods.

As compared with the overall cohort of patients with OSA, true positive cases identified using single diagnostic codes or using the combination of a polysomnogram followed by receipt of a PAP device had a higher prevalence of male sex, moderate-to-severe OSA, and important comorbid diseases. For ICD-9 code 780 and the polysomnogram + PAP

Table 3. Diagnostic Test Characteristics

Case Definition	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Likelihood Ratio		Posttest Probability (%)*	
			Positive	Negative	Algorithm+	Algorithm-
Polysomnogram + positive airway pressure device	19 (17–20)	98 (98–99)	10.9 (8.9–13.0)	0.83 (0.81–0.85)	55	8
Sleep apnea, obstructed (ICD-10 4730 [DAD])	7 (6–8)	99 (98–99)	5.5 (3.5–7.5)	0.95 (0.93–0.96)	38	10
Other sleep apnea (ICD-10 G4738 [DAD])	2 (2–3)	99 (99–100)	2.9 (0.6–4.8)	0.98 (0.97–1.1)	25	10
Any ICD-10 code for OSA (DAD)	9 (8–10)	98 (97–99)	4.5 (2.4–6.5)	0.93 (0.92–0.94)	33	9
Unspecified sleep apnea (ICD-9 780.5 [DAD])	3 (2–3)	98 (97–98)	1.1 (0–2.9)	1.00 (0.99–1.0)	11	10
Unspecified sleep apnea (ICD-9 780.5 [OHIP])	58 (56–60)	38 (36–40)	0.9 (0–2.8)	1.1 (1.0–1.2)	9	11
Any unspecified sleep apnea (ICD-9 780.5 [OHIP or DAD])	58 (56–60)	38 (35–40)	0.9 (0–3.0)	1.1 (1.0–1.2)	9	11
Polysomnogram + positive airway pressure device or any ICD code	70 (68–71)	36 (34–38)	1.1 (0–3.1)	0.83 (0.75–0.92)	11	8

Positive and negative predictive values omitted.

* Given an OSA prevalence of 10%.

DAD = Discharge Abstract Database; ICD = *International Classification of Diseases*; OHIP = Ontario Health Insurance Plan; OSA = obstructive sleep apnea.

case-ascertainment algorithms, true positive cases were older than the average age of the full study cohort diagnosed with OSA (table 5).

Discussion

In this validation study of case-ascertainment algorithms, we found that HA data codes for diagnoses, diagnostic procedures, or therapeutic interventions (or combinations thereof) failed to accurately identify surgical patients with OSA. Although the specificities of single diagnostic codes identified in hospital discharge abstracts as well as the combination of a polysomnogram followed by receipt of a PAP device before surgery were high, sensitivities were low, and false-positive results were identified. Such test characteristics mean that the overall accuracy provided by the case-ascertainment algorithms tested in this study are inadequate in identifying people who truly have disease or to rule out disease in people who truly do not have OSA.

The usefulness of a diagnostic test is dependent on the combination of disease prevalence and the accuracy of the test. Likelihood ratios are relatively impervious to disease prevalence and can be used to estimate the usefulness of a diagnostic test at various prevalence levels.²¹ Furthermore, likelihood ratio–based cutoffs have been suggested to guide the assessment of diagnostic tests, such as our case-ascertainment algorithms. Tests with a +LR of 10 or greater and a –LR of 0.1 or less are considered to be very useful; those with a +LR between 5 and 10 and a –LR between 0.1 and 0.2 are considered moderately useful; and those with a +LR between 2 and 5 and a –LR 0.5 are considered somewhat useful. Tests with a +LR less than 2 or a –LR greater than 0.5 are considered useless.²⁸ Existing validated case-ascertainment algorithms used to identify and study chronic diseases in population-based HA data are typically in the

very or moderately useful category (hypertension,²⁹ diabetes,³⁰ previous myocardial infarction,¹⁴ and congestive heart failure¹³). The most accurate case-ascertainment algorithm in our study (the combination of a claim for a preoperative polysomnogram followed by receipt of a PAP device) had a +LR of 10.9 and a –LR of 0.83. For reference, if OSA is conservatively estimated to be present in 10% of surgical patients,³¹ a +LR of 10.9 would increase the pretest probability of OSA from 10 to 55% (*i.e.*, patients meeting these criteria would have only a 55% probability of truly having OSA); a +LR of 40 would be needed to increase the posttest probability to 80%. No subgroup existed in which these case-ascertainment algorithms were any more accurate. Using the single diagnostic code methods used in the literature, we estimate that the highest probability of identifying true OSA would be less than 40%. Furthermore, because –LRs were approximately 1 for all methods tested, the absence of a diagnostic code for OSA would provide almost no indication that the patient did not have OSA. We would therefore expect a prevalence of OSA of approximately 10% in the unexposed group in existing HA data studies of OSA.

Although one would expect misclassification to bias the results of a study toward the null (*i.e.*, no increased impact of OSA on adverse outcomes), the types of patients with OSA identified by diagnostic codes further cloud the study interpretation. Our results demonstrate that the patients who had a diagnostic code for OSA and had a true diagnosis of OSA based on polysomnogram results (*i.e.*, true positives) are those who may be at greater risk of adverse outcomes. These true positive patients had a markedly higher prevalence of diabetes, heart disease, or respiratory disease and were more likely to be male and of increased age. All of these factors are associated with adverse outcomes after surgery.^{32–34} Furthermore, the majority of these patients had moderate or severe

Table 4. Diagnostic Test Characteristics from Subgroup Analysis

Case Definition	Subgroup	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Likelihood Ratio (95% CI)	
				Positive	Negative
Polysomnogram + positive airway pressure device	Moderate-severe	23 (21–25)	96 (95–97)	5.8 (3.8–7.7)	0.80 (0.77–0.83)
	Ambulatory	18 (16–21)	98 (97–99)	10.6 (8.6–12.7)	0.83 (0.80–0.87)
	Inpatient	19 (17–21)	98 (97–99)	11.8 (9.7–13.8)	0.83 (0.80–0.86)
	SDB	18 (17–20)	98 (98–99)	10.6 (8.6–12.6)	0.83 (0.81–0.85)
	Test surgery <3 yr	16 (14–18)	98 (98–99)	8.8 (5.7–13.5)	0.86 (0.84–0.88)
	Surgery 2003–2007	10 (8–12)	99 (98–99)	7.2 (4.0–12.7)	0.91 (0.89–0.91)
Sleep apnea, obstructed (ICD-10 4730 [DAD])	Surgery 2008–2012	24 (22–26)	98 (97–99)	12.7 (8.1–20.0)	0.77 (0.75–0.80)
	Moderate-severe	8 (6–9)	98 (97–98)	3.5 (1.5–5.5)	0.94 (0.93–0.96)
	Ambulatory	4 (2.8–5.4)	99 (99–100)	6.7 (4.6–8.8)	0.97 (0.95–0.98)
	Inpatient	9 (7–9)	98 (98–99)	3.0 (0.95–5.0)	0.98 (0.96–0.99)
	SDB	7 (6–8)	99 (98–99)	5.9 (3.9–7.9)	0.95 (0.93–0.96)
	Test surgery <3 yr	8 (7–9)	99 (98–99)	5.2 (3.2–8.3)	0.93 (0.92–0.95)
Other sleep apnea (ICD-10 G4738 [DAD])	Surgery 2003–2007	11 (9–13)	98 (97–99)	6.8 (4.0–11.7)	0.90 (0.88–0.92)
	Surgery 2008–2012	4 (3–5)	99 (98–100)	4.0 (2.0–8.2)	0.97 (0.96–0.98)
	Moderate-severe	3 (2–4)	99 (99–99)	3.1 (1.1–5.1)	0.98 (0.97–0.99)
	Ambulatory	2 (1–2)	99 (99–100)	3.0 (0.85–5.1)	0.98 (0.97–0.99)
	Inpatient	3 (2–4)	99 (98–100)	3.0 (0.95–5.1)	0.98 (0.97–0.99)
	SDB	3 (2–3)	99 (99–100)	4.3 (2.3–6.4)	0.98 (0.97–0.99)
Any ICD-10 code (DAD)	Test surgery <3 yr	3 (2–4)	99 (99–100)	3.7 (1.8–7.5)	0.98 (0.97–0.99)
	Surgery 2003–2007	4 (3–5)	99 (98–100)	5.1 (2.3–11.3)	0.97 (0.96–0.98)
	Surgery 2008–2012	2 (1–3)	99 (98–100)	2.1 (1.0–4.7)	0.99 (0.98–1.00)
	Moderate-severe	11 (9–12)	97 (96–98)	3.3 (1.3–5.3)	0.91 (0.89–0.93)
	Ambulatory	5 (4–7)	99 (98–100)	4.8 (2.8–6.9)	0.96 (0.94–0.97)
	Inpatient	11 (10–13)	97 (96–98)	3.8 (1.8–5.8)	0.91 (0.89–0.93)
Unspecified sleep apnea (ICD-9 780.5 [DAD]) 0.99 (0.98–1.00) + A32	SDB	9 (8–10)	98 (98–99)	5.0 (3.0–7.0)	0.93 (0.91–0.94)
	Test surgery <3 yr	11 (9–12)	98 (97–99)	4.7 (3.2–6.9)	0.91 (0.90–0.93)
	Surgery 2003–2007	15 (13–17)	98 (96–99)	6.1 (3.9–9.5)	0.87 (0.85–0.89)
	Surgery 2008–2012	5 (4–7)	98 (97–99)	3.1 (1.9–5.3)	0.96 (0.95–0.98)
	Moderate-severe	3 (2–3)	98 (97–98)	1.0 (0–3.0)	1.0 (0.99–1.0)
	Ambulatory	2 (2–4)	98 (96–98)	0.96 (0–3.0)	1.0 (0.99–1.0)
Unspecified sleep apnea (ICD-9 780.5 [OHIP])	Inpatient	3 (2–4)	98 (96–99)	1.1 (0–3.1)	1.0 (0.98–1.0)
	SDB	3 (2–3)	98 (97–98)	1.1 (0–3.1)	1.0 (0.99–1.0)
	Test surgery <3 yr	2 (2–3)	98 (97–98)	0.98 (0.63–1.54)	1.0 (0.99–1.0)
	Surgery 2003–2007	3 (2–5)	97 (96–98)	1.1 (0.7–1.5)	1.0 (0.99–1.0)
	Surgery 2008–2012	2 (1–3)	98 (97–99)	1.1 (0.6–1.9)	1.0 (0.99–1.0)
	Moderate-severe	57 (55–60)	39 (37–41)	0.94 (0–2.9)	1.1 (1.0–1.2)
Any unspecified sleep apnea (ICD-9 780.5 [OHIP or DAD])	Ambulatory	57 (53–60)	39 (37–43)	0.94 (0–2.9)	1.1 (0.99–1.2)
	Inpatient	58 (56–61)	36 (33–39)	0.91 (0–2.9)	1.2 (1.1–1.3)
	SDB	58 (56–60)	38 (36–40)	0.93 (0–2.9)	1.1 (1.0–1.2)
	Test surgery <3 yr	57 (55–60)	39 (36–41)	0.94 (0.88–1.04)	1.1 (1.0–1.2)
	Surgery 2003–2007	55 (51–58)	40 (37–43)	0.91 (0.84–0.99)	1.1 (1.0–1.3)
	Surgery 2008–2012	59 (57–62)	60 (58–63)	1.5 (1.4–1.6)	0.67 (0.62–0.72)
Polysomnogram + positive airway pressure device or any ICD code	Moderate-severe	57 (55–60)	39 (37–41)	0.93 (0–2.9)	1.1 (1.0–1.2)
	Ambulatory	57 (54–60)	39 (36–42)	0.93 (0–1.9)	1.1 (1.0–1.2)
	Inpatient	59 (56–61)	36 (33–39)	0.91 (0–2.9)	1.2 (1.1–1.3)
	SDB	58 (56–60)	38 (35–40)	0.93 (0–2.9)	1.1 (1.0–1.2)
	Test surgery <3 yr	58 (56–60)	38 (36–40)	0.94 (0.88–0.99)	1.1 (1.0–1.2)
	Surgery 2003–2007	55 (52–58)	39 (36–43)	0.91 (0.84–0.98)	1.1 (1.0–1.3)
	Surgery 2008–2012	60 (57–62)	36 (33–39)	0.93 (0.87–0.99)	1.1 (1.0–1.3)
	Moderate-severe	73 (72–75)	36 (34–38)	1.1 (0–3.1)	0.76 (0.66–0.86)
	SDB	70 (68–72)	36 (34–39)	1.1 (0–3.1)	0.83 (0.74–0.95)
	Test surgery <3 yr	70 (67–72)	36 (34–39)	1.1 (1.0–1.2)	0.82 (0.75–0.91)
	Surgery 2003–2007	67 (64–70)	37 (34–41)	1.1 (1.0–1.1)	0.89 (0.78–1.00)
	Surgery 2008–2012	72 (70–74)	37 (32–38)	1.1 (1.0–1.2)	0.81 (0.72–0.91)

Subgroup sample sizes: ambulatory: n = 1,962, inpatient: n = 2,391, PSG surgery < 3 yr: n = 3,073, surgery 2003 to 2007: n = 1,870, surgery 2008 to 2013: n = 2,483. Test refers to polysomnogram.

DAD = Discharge Abstract Database; ICD = *International Classification of Diseases*; OHIP = Ontario Health Insurance Plan.

Table 5. Patient Characteristics Based on Test Results of Selected Case-ascertainment Algorithms

	Female	Age (SD)	Moderate or Severe OSA	Heart Disease	Respiratory Disease	Diabetes
Any ICD-10						
True positive	25	51 (13)	52	14.7	34	25
False negative	44	52 (13)	38	8	38	21
True negative	38	52 (13)	N/A	11.5	34	33
False positive	61	48 (13)	N/A	7.5	35	16
ICD-9 780.5 (DAD)						
True positive	37	59 (13)	47	23	62	48
False negative	62	56 (13)	39	15	49	30
True negative	37	52 (13)	N/A	12	47	32
False positive	60	50 (15)	N/A	7	35	17
Polysomnogram + positive airway pressure device						
True positive	35	57 (11)	53	13	33	40
False negative	56	56 (13)	36	9	41	19
True negative	37	52 (13)	N/A	11	34	31
False positive	60	55 (12)	N/A	7	35	17

All values are % unless otherwise stated.

DAD = Discharge Abstract Database; ICD = *International Classification of Diseases*; N/A = not applicable; OSA = obstructive sleep apnea.

OSA based on their AHI, which is also postulated to increase postoperative risk.³⁵ We believe that the results of existing HA data studies examining OSA in the perioperative period should be interpreted with great caution. In summary, we report that diagnostic codes used to determine OSA status appear to identify patients at higher risk of perioperative adverse events, independent of OSA; this bias may overstate the impact of OSA on postoperative outcomes.

As described in the second paragraph of our introduction, the current literature regarding outcomes of surgical patients with OSA contains significant knowledge gaps. HA data studies have the potential to address these gaps, which could explain the recent publication of numerous HA data studies in this area. However, our data raises extensive skepticism about the ability to accurately identify patients with OSA by using the HA data. Until an accurate method to identify OSA in HA data is developed, researchers and knowledge consumers should approach such studies cautiously. Other examples exist in perioperative medicine where diagnostic codes with unknown accuracy for identifying true disease status provided quantitatively different and conflicting results compared with studies using higher-quality prospective data. In the case of statins and postoperative delirium, an initial study, performed using HA data, concluded that the statin use was associated with a 28% *increase* in the odds of postoperative delirium.³⁶ Subsequent prospective study demonstrated that statins were actually protective, with the odds of postoperative delirium being *decreased* by 46% in statin users.³⁷ Validation testing demonstrated that despite high specificity, diagnostic codes for delirium in HA data lack accuracy for identifying delirium, with a +LR 18 and a -LR of 0.82 for a disease that is approximately 10% prevalent.³⁸ Given this example and the new knowledge generated in our study, we suggest that if OSA is felt to be an

important perioperative risk factor, concerted efforts by researchers and health system administrators will be required to allow the accurate identification of patients with OSA in HA data. For researchers, this may involve the development of case-ascertainment algorithms that do not rely exclusively on diagnostic codes, diagnostic procedures, or therapeutic interventions related to OSA. For health system administrators, this may involve improving the coding accuracy of diagnostic codes for OSA in HA databases.

Strengths and Limitations

Our study has several strengths. We used definitive standard diagnostic methods to determine the OSA status of each patient in our cohort. We used population-based administrative data, which captured all physician and hospital encounters as well as all polysomnograms and the vast majority of PAPs for all patients. Although many HA data studies of OSA define disease status using administrative data only at the time of surgical hospital admission, we looked back 2 yr before surgery to capture diagnostic codes related to OSA; although we cannot exclude the possibility that a longer look-back window would increase the accuracy, accuracy would have to increase substantially to make codes useful for identifying OSA in HA data. Our validation cohort was large, allowing for subgroup analyses. Our algorithms may be applicable to nonsurgical patients captured in Ontario's, and other similar health systems' HA data, as ICD codes, polysomnography, and PAP devices are not unique to patients with OSA undergoing surgery.

Our study has certain limitations. The sampling frame was a single academic health sciences center and may lack generalizability; however, patient characteristics were consistent with typical patients with OSA from other settings.

Although LRs are relatively immune to disease prevalence, there is little guidance available to predict the full impact of disease prevalence on a binary disease indicator (such as the presence or absence of a diagnostic code). A simulation study of normally distributed disease indicators (such as hemoglobin levels indicating anemia) suggests that +LRs may increase in populations with low disease prevalence, whereas -LRs may tend closer to one.²³ This finding, however, is not directly generalizable to the current study. Our polysomnogram + PAP algorithm may not be applicable in all health systems. Reliance upon PAP device receipt limits the identification of patients with OSA offered other treatment modalities. Our case-ascertainment algorithms contained physician-identified diagnostic codes (which are used for billing submissions and are based on ICD-9); such codes do not require specific diagnostic criteria and could be applied if disease was suspected but not proven (as might be the case in a clinical visit resulting in ordering of a polysomnogram). Finally, 12% of our reference population was not linked to HA data. The majority of unlinked patients were patients with universal health insurance provided by the province of Quebec and who live in the same metropolitan area as our linked patients. It is unlikely that these individuals differed substantively from linked individuals, but such a possibility should be considered when applying our findings.

Conclusion

The use of diagnostic codes, or a combination of diagnostic codes, diagnostic interventions, and therapeutic interventions, did not provide a case-ascertainment algorithm that reliably identified patients with OSA in HA data. Furthermore, the use of diagnostic codes to define OSA may introduce important misclassification bias since patients with OSA identified by these algorithms appear to systematically differ from other patients with OSA with respect to the severity of their OSA and important perioperative prognostic factors. Future research is required to develop methodologies that accurately and consistently identify patients with OSA in HA data. Until such methodologies exist, studies using diagnostic codes to identify OSA exposure should be interpreted with great caution.

Acknowledgments

Dr. McIsaac thanks the salary support from The Ottawa Hospital Anesthesia Alternate Funds Association (Ottawa, Ontario, Canada). Dr. Bryson thanks the support by The Ottawa Hospital Anesthesia Alternate Funds Association (Ottawa, Ontario, Canada). Dr. Wijesundera thanks the support, in part, by a Clinician Scientist Award from the Canadian Institutes of Health Research (Ottawa, Ontario, Canada) and a Merit Award from the Department of Anesthesia at the University of Toronto (Toronto, Ontario, Canada).

This study was funded by the Ontario Thoracics Society (Toronto, Ontario, Canada) and the Canadian Anesthesiology Society (Toronto, Ontario, Canada). This study was also supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an

annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC; Toronto, Ontario, Canada). The opinions, results, and conclusions reported in this article are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended or should be inferred. These data sets are held securely in a linked, deidentified form and analyzed at ICES.

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. McIsaac: Department of Anesthesiology, The Ottawa Hospital, and Institute for Clinical Evaluative Sciences at uOttawa, 1053 Carling Avenue, Room B311, Ottawa, Ontario, K1Y 4E9, Canada. dmcisaac@toh.on.ca. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

References

1. Turner K, VanDenkerkhof E, Lam M, Mackillop W: Perioperative care of patients with obstructive sleep apnea—A survey of Canadian anesthesiologists. *Can J Anaesth* 2006; 53:299–304
2. Joshi GP, Ankichetty SP, Gan TJ, Chung F: Society for Ambulatory Anesthesia consensus statement on preoperative selection of adult patients with obstructive sleep apnea scheduled for ambulatory surgery. *Anesth Analg* 2012; 115:1060–8
3. American Society of Anesthesiologists Task Force on Perioperative Management of Patients with Obstructive Sleep Apnea: Practice guidelines for the perioperative management of patients with obstructive sleep apnea: An updated report by the American Society of Anesthesiologists Task Force on Perioperative Management of patients with obstructive sleep apnea. *ANESTHESIOLOGY* 2014; 120:268–86
4. 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
5. Chung F, Mokhlesi B: Postoperative complications associated with obstructive sleep apnea: Time to wake up! *Anesth Analg* 2014; 118:251–3
6. Memtsoudis SG, Stundner O, Rasul R, Chiu YL, Sun X, Ramchandran SK, Kaw R, Fleischut P, Mazumdar M: The impact of sleep apnea on postoperative utilization of resources and adverse outcomes. *Anesth Analg* 2014; 118:407–18
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
8. 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
9. Mutter TC, Chateau D, Moffatt M, Ramsey C, Roos LL, Kryger M: A matched cohort study of postoperative outcomes in obstructive sleep apnea: Could preoperative diagnosis and treatment prevent complications? *ANESTHESIOLOGY* 2014; 121:707–18
10. Quan H, Li B, Saunders LD, Parsons GA, Nilsson CI, Alibhai A, Ghali WA; IMECCHI Investigators: Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. *Health Serv Res* 2008; 43:1424–41

11. Quach S, Blais C, Quan H: Administrative data have high variation in validity for recording heart failure. *Can J Cardiol* 2010; 26:306–12
12. Quan H, Khan N, Hemmelgarn BR, Tu K, Chen G, Campbell N, Hill MD, Ghali WA, McAlister FA: Hypertension Outcome and Surveillance Team of the Canadian Hypertension Education Programs: Validation of a case definition to define hypertension using administrative data. *Hypertension* 2009; 54:1423–8
13. Schultz SE, Rothwell DM, Chen Z, Tu K: Identifying cases of congestive heart failure from administrative data: A validation study using primary care patient records. *Chronic Dis Inj Can* 2013; 33:160–6
14. Tu K, Mitiku T, Guo H, Lee DS, Tu JV: Myocardial infarction and the validation of physician billing and hospitalization data using electronic medical records. *Chronic Dis Can* 2010; 30:141–6
15. Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T: Identifying individuals with physician diagnosed COPD in health administrative databases. *COPD* 2009; 6:388–94
16. Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T: Identifying patients with physician-diagnosed asthma in health administrative databases. *Can Respir J* 2009; 16:183–8
17. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC: Standards for Reporting of Diagnostic Accuracy: Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *BMJ* 2003; 326:41–4
18. Epstein LJ, Kristo D, Strollo PJ Jr, Friedman N, Malhotra A, Patil SP, Ramar K, Rogers R, Schwab RJ, Weaver EM, Weinstein MD: Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine: Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med* 2009; 5:263–76
19. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE, Ghali WA: Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005; 43:1130–9
20. Charlson M, Szatrowski TP, Peterson J, Gold J: Validation of a combined comorbidity index. *J Clin Epidemiol* 1994; 47:1245–51
21. Van Walraven C, Austin P: Administrative database research has unique characteristics that can risk biased results. *J Clin Epidemiol* 2012; 65:126–31
22. Simel DL, Samsa GP, Matchar DB: Likelihood ratios with confidence: Sample size estimation for diagnostic test studies. *J Clin Epidemiol* 1991; 44:763–70
23. Brenner H, Gefeller O: Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat Med* 1997; 16:981–91
24. Redline S, Schluchter MD, Larkin EK, Tishler PV: Predictors of longitudinal change in sleep-disordered breathing in a nonclinic population. *Sleep* 2003; 26:703–9
25. Peppard PE, Young T, Palta M, Dempsey J, Skatrud J: Longitudinal study of moderate weight change and sleep-disordered breathing. *JAMA* 2000; 284:3015–21
26. Newman AB, Foster G, Givelber R, Nieto FJ, Redline S, Young T: Progression and regression of sleep-disordered breathing with changes in weight: The Sleep Heart Health Study. *Arch Intern Med* 2005; 165:2408–13
27. Punjabi NM: The epidemiology of adult obstructive sleep apnea. *Proc Am Thorac Soc* 2008; 5:136–43
28. Jaeschke R, Guyatt GH, Sackett DL: Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 1994; 271:703–7
29. Tu K, Campbell NR, Chen ZL, Cauch-Dudek KJ, McAlister FA: Accuracy of administrative databases in identifying patients with hypertension. *Open Med* 2007; 1:e18–26
30. Hux JE, Ivis F, Flintoft V, Bica A: Diabetes in Ontario: Determination of prevalence and incidence using a validated administrative data algorithm. *Diabetes Care* 2002; 25:512–6
31. Singh M, Liao P, Kobah S, Wijesundera DN, Shapiro C, Chung F: Proportion of surgical patients with undiagnosed obstructive sleep apnoea. *Br J Anaesth* 2013; 110:629–36
32. Lee TH, Marcantonio ER, Mangione CM, Thomas EJ, Polanczyk CA, Cook EF, Sugarbaker DJ, Donaldson MC, Poss R, Ho KK, Ludwig LE, Pedan A, Goldman L: Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. *Circulation* 1999; 100:1043–9
33. Grewal K, Wijesundera DN, Carroll J, Tait G, Beattie WS: Gender differences in mortality following non-cardiovascular surgery: An observational study. *Can J Anaesth* 2012; 59:255–62
34. Oresanya LB, Lyons WL, Finlayson E: Preoperative assessment of the older patient: A narrative review. *JAMA* 2014; 311:2110–20
35. Chung SA, Yuan H, Chung F: A systemic review of obstructive sleep apnea and its implications for anesthesiologists. *Anesth Analg* 2008; 107:1543–63
36. Redelmeier DA, Thiruchelvam D, Daneman N: Delirium after elective surgery among elderly patients taking statins. *CMAJ* 2008; 179:645–52
37. Katznelson R, Djaiani GN, Borger MA, Friedman Z, Abbey SE, Fedorko L, Karski J, Mitsakakis N, Carroll J, Beattie WS: Preoperative use of statins is associated with reduced early delirium rates after cardiac surgery. *ANESTHESIOLOGY* 2009; 110:67–73
38. Katznelson R, Djaiani G, Tait G, Wasowicz M, Sutherland AM, Styra R, Lee C, Beattie WS: Hospital administrative database underestimates delirium rate after cardiac surgery. *Can J Anaesth* 2010; 57:898–902

Appendix. Standards for the Reporting of Diagnostic Accuracy Studies Checklist for Reporting of Studies of Diagnostic Accuracy (Version January 2003)

Section and Topic	Item No.		On Page No.
Title/abstract/keywords	1	Identify the article as a study of diagnostic accuracy (recommend MeSH "sensitivity and specificity").	P 3 L70
Introduction	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	P4L 122
Methods			
Participants	3	The study population: The inclusion and exclusion criteria and setting and locations where data were collected.	P 6L 180
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	P 6L 180
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in items 3 and 4? If not, specify how participants were further selected.	P 7 L186
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	P 7 L 187
Test methods	7	The reference standard and its rationale.	P 7 L 190
	8	Technical specifications of materials and methods involved including how and when measurements were taken and/or cite references for index tests and reference standard.	P 7 L 191
	9	Definition of and rationale for the units, cutoffs, and/or categories of the results of the index tests and the reference standard.	P 7 L 191
	10	The number, training, and expertise of the persons executing and reading the index tests and the reference standard.	P 7 L 190
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	N/A
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g., 95% CIs).	P 7 L 211
	13	Methods for calculating test reproducibility, if done.	N/A
Results			
Participants	14	When the study was performed, including beginning and end dates of recruitment.	P 6L 181
	15	Clinical and demographic characteristics of the study population (at least information on age, sex, and spectrum of presenting symptoms).	Table 2
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either tests (a flow diagram is strongly recommended).	Figure 1
Test results	17	Time interval between the index tests and the reference standard and any treatment administered in between.	N/A
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	Table 2
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	Tables 3 and 4
	20	Any adverse events from performing the index tests or the reference standard.	N/A
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g., 95% CIs).	Tables 3 and 4
	22	How indeterminate results, missing data, and outliers of the index tests were handled.	P 7 L 191
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers, or centers, if done.	Table 4
	24	Estimates of test reproducibility, if done.	N/A
Discussion	25	Discuss the clinical applicability of the study findings.	P 11 L 311

MeSH = Medical Subject Heading; N/A = not applicable.