HHS Public Access
Author manuscript
Pharmacoepidemiol Drug Saf. Author manuscript; available in PMC 2018 August 01.

Published in final edited form as:
Pharmacoepidemiol Drug Saf. 2017 August ; 26(8): 945–953. doi:10.1002/pds.4226.

Evaluation of Algorithms to Identify Delirium in Administrative Claims and Drug Utilization Database

Dae Hyun Kim, MD, MPH, ScD[a,b], **Jung Lee**[b], **Caroline A. Kim, MD, MPH**[b], **Krista F. Huybrechts, MS, PhD**[a], **Brian T. Bateman, MD, MSc**[a,d], **Elisabetta Patorno, MD, DrPH**[a], and **Edward R. Marcantonio, MD, SM**[b,c]

[a]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Boston, MA

[b]Division of Gerontology, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA

[c]Division of General Medicine and Primary Care, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA

[d]Department of Anesthesiology, Perioperative and Pain Medicine, Brigham and Women's Hospital, Boston, MA

Abstract

Corresponding Author: Dae Hyun Kim, MD, MPH, ScD, 1620 Tremont Street, Suite 3030, Boston, MA 02120, Tel: 617-278-0930; Fax: 617-232-8602; dkim12@partners.org.

Conflict of Interest: D.H. Kim is consultant to Alosa Health, a nonprofit educational organization with no relationship to any drug or device manufacturers; K.F. Huybrechts receives grants from Pfizer, Lilly, and GlaxoSmithKline for projects unrelated to the present study; B.T. Bateman receives grants from Pfizer, Lilly, GlaxoSmithKline, and Baxalta for projects unrelated to the present study, and is a consultant to Optum, Inc.; E. Patorno receives grants from GlaxoSmithKline and Boehringer Ingelheim for projects unrelated to the present study; other authors declare no conflict of interest.

Funding Sources: This study was funded by a KL2 Medical Research Investigator Training award (KL2 TR001100-01) to D.H. Kim and an additional support (UL1 TR001102) from Harvard Catalyst / Harvard Clinical and Translational Science Center; the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health; and by the Training Program in Cardiovascular Research award (T32-HL007374) to C.A. Kim from the National Heart, Lung, and Blood Institute, National Institutes of Health.

- D.H. Kim is supported by the Paul B. Beeson Clinical Scientist Development Award in Aging (K08AG051187) from the National Institute on Aging, American Federation for Aging Research, The John A. Hartford Foundation, and The Atlantic Philanthropies.
- K.F. Huybrechts is supported by grant K01MH099141 from the National Institute of Mental Health.
- B.T. Bateman is supported by grant K08HD075831 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development.
- E.R. Marcantonio is funded by grants R01AG030618, P01AG031720, and a Mid-Career Investigator Award (K24AG035075) from the National Institute on Aging.

Author Contributions:
Conception and design: D.H. Kim.
Collection of data: D.H. Kim, J. Lee, C.A. Kim.
Analysis and interpretation of data: All authors.
Drafting of the article: D.H. Kim.
Critical revision for intellectual content: All authors.
Final approval of the article: All authors.
Obtaining of funding: D.H. Kim, C.A. Kim.

Purpose—To evaluate the performance of delirium-identification algorithms in administrative claims and drug utilization data.

Methods—We used data from a prospective study of 184 older adults who underwent aortic valve replacement at a single academic medical center to evaluate the following delirium-identification algorithms: 1) International Classification of Diseases (ICD) diagnosis codes for delirium, 2) antipsychotics use, 3) either ICD diagnosis codes or antipsychotics use, and 4) both ICD diagnosis codes and antipsychotics use. These algorithms were evaluated against a validated bedside assessment, the Confusion Assessment Method, and a validated delirium severity scale, the CAM-S.

Results—Delirium occurred in 66 patients (36%), of which 14 (21%) had hyperactive or mixed features and 15 (23%) had severe delirium. ICD diagnosis codes for delirium were present in 15 patients (8%). Antipsychotics were used in 13 patients (7%). ICD diagnosis codes alone and antipsychotics use alone had comparable sensitivity (18% vs. 18%) and specificity (98% vs. 99%). Defining delirium using either ICD diagnosis codes or antipsychotics use, sensitivity improved to 30% with little change in specificity (97%). This algorithm showed higher sensitivity for hyperactive or mixed delirium (64%) and severe delirium (73%). Requiring both ICD diagnosis codes and antipsychotics use resulted in perfect specificity, but low sensitivity (6%).

Conclusion—Delirium-identification algorithms in claims data have low sensitivity and high specificity. Defining delirium using ICD diagnosis codes or antipsychotics use performs better than considering either type of information alone. This information should inform the design and interpretation of claims-based comparative effectiveness and safety research.

Keywords

Delirium; Case-Detection Algorithm; Claims Data; Antipsychotics

INTRODUCTION

Postoperative delirium is an acute brain dysfunction that affects 25–60% of older adults after abdominal surgery, 33–46% after cardiac and vascular surgery, and 24–61% after orthopedic surgery.[1,2] It is associated with prolonged hospitalization, death, institutional discharge, functional decline, and increased health care expenditure.[3–8] Considering this enormous public health and economic impact, it is essential to evaluate the burden, risk factors, and outcomes of postoperative delirium in health care utilization databases.

Identification of delirium requires bedside cognitive assessments and application of validated diagnostic methods, such as the Confusion Assessment Method (CAM)[9] or the Diagnostic and Statistical Manual of Mental Disorders criteria.[1] These assessments, however, are not routinely documented in the medical records.[10–12] Moreover, they are unavailable in health care utilization databases (e.g., claims data or hospital clinical data repository) that contain information on diagnoses, procedures, medications, and laboratory tests. The International Classification of Diseases (ICD) diagnosis codes that explicitly or implicitly indicate delirium variably under-diagnose the condition.[10–15] Some studies have attempted to use antipsychotic drug administration as an indicator of delirium,[16–18] but the accuracy of this approach alone or combining ICD diagnosis codes and antipsychotics use has not been well established. Evaluating the performance and improving the existing algorithms can enable measurement of delirium in health care utilization data and conduct of drug effectiveness and safety studies on delirium.

The objective of this study was to determine the performance of case detection algorithms of postoperative delirium based on ICD diagnosis codes and antipsychotics use in hospitalized older adults who had cardiac surgery.

METHODS

Study Design and Population

We conducted a prospective cohort study of older adults who were undergoing transcatheter or surgical aortic valve replacement at Beth Israel Deaconess Medical Center, Boston, MA, to determine the role of preoperative frailty assessment in predicting changes in functional status after aortic valve replacement. The enrollment began in February 2014 and ended in March 2016. In October 2014, we launched the Delirium Substudy to determine the incidence and risk factors of postoperative delirium. The Institutional Review Board of Beth Israel Deaconess Medical Center approved the protocol.

We screened 445 patients from the outpatient clinic schedule, inpatient consult list, and operating room schedule for eligibility. Individuals were eligible if all of the following criteria were met: 1) 70 years or older; 2) undergoing transcatheter or surgical aortic valve replacement for severe aortic stenosis; 3) able to provide informed consent. Those who met any of the following criteria were excluded: 1) undergoing emergent surgery, 2) having a surgery involving more than 1 valve or aorta; 3) clinically unstable (e.g., decompensated heart failure, active myocardial ischemia, or abnormal vital signs) and unable to participate in study assessment; 4) Mini-Mental State Examination score <15 or active psychosis; 5) not English-speaking. Since the Delirium Substudy started 8 months after the study enrollment began, 184 of 246 patients who met our selection criteria were included.

For the current study, we obtained administrative data from the hospital clinical data repository that pooled data from hospital billing (with 39 diagnosis fields) and electronic medical record. We only considered discharge diagnosis codes and daily inpatient medication dispensing for 184 patients in the Delirium Substudy.

Assessment of Postoperative Delirium and Severity

Beginning on the first postoperative day, study geriatricians or trained research assistants interviewed the patients, their family when present, and nursing staff daily; and administered the CAM (or CAM for intensive care unit [CAM-ICU] if patients were on mechanical ventilation), the Delirium Symptom Interview,[19] after brief cognitive tests. Because the patients' mental status may fluctuate throughout the day, we attempted to standardize the timing of our daily assessment by interviewing patients between 12:00 PM and 6:00 PM. As discharges typically took place in the morning, most patients were not interviewed on the discharge day. Cognitive tests included the Mini-Mental State Examination (purchased from Psychological Assessment Resources, Inc.), digit span forward and backward, and backward citation of days of the week and months of the year. The diagnosis of delirium was made if the patient showed 1) acute onset or fluctuating course, 2) inattention, and either 3) disorganized thinking or 4) altered level of consciousness.[9] The diagnostic accuracy of CAM (sensitivity 94% and specificity 89%)[20,21] and CAM-ICU (sensitivity 93–100% and specificity 97–100%)[22,23] has been well established. Based on all available CAMs administered during the hospitalization, delirium was classified as normoactive if psychomotor agitation and retardation were never present, hypoactive if psychomotor retardation was present without agitation, hyperactive if psychomotor agitation was present without retardation, or mixed if both psychomotor agitation and retardation were present during the hospitalization.[24] The inter-rater reliability between geriatricians and trained research assistants was high for delirium diagnosis (agreement 95–100%, Cohen's kappa 0.90–1.00) and psychomotor items (agreement 92–100%, Cohen's kappa 0.83–1.00). Since only four patients developed hyperactive delirium, we combined hyperactive and mixed subtypes. We quantified the severity of delirium using a validated CAM-based severity scale, CAM-S.[25] The CAM-S assigns 0 (absent) or 1 (present) to acute onset or fluctuating course, and 0 (absent), 1 (mild), or 2 (marked) to the remaining CAM features: inattention, disorganized thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation, psychomotor retardation, and sleep-wake cycle disturbance. The total CAM-S score ranges from 0 to 19, with higher scores indicating greater severity. Based on the peak CAM-S score, we classified patients with delirium into mild (≤6), moderate (7–10), and severe delirium (11–19).[25]

Claims-Based Algorithms to Identify Delirium

We evaluated four different algorithms to identify delirium: 1) ICD diagnosis codes alone, 2) antipsychotics use alone, 3) either ICD diagnosis codes OR antipsychotics use, and 4) both ICD diagnosis codes AND antipsychotics use. For algorithms based on ICD diagnosis codes, delirium was considered present if any of the following 32 explicit (i.e., including the term "delirium") or implicit ICD-9 diagnosis codes (e.g., "encephalopathy") was used among the in-hospital discharge codes (see Supplementary Table S1).[14] In claims data from October 1, 2015, the corresponding 52 ICD-10 diagnosis codes were used.[14] For algorithms based on antipsychotics use, delirium was considered present if any antipsychotic drugs were used; in our study, we found that only haloperidol, olanzapine, and quetiapine had been prescribed. We did not examine other drugs that could influence the risk of delirium (e.g., benzodiazepine or dexmedetomidine).

Statistical Analysis

We summarized clinical characteristics in proportions, mean and standard deviation, or median and interquartile range. We used chi-square test to compare the proportion of patients with ICD diagnosis codes for delirium and of those who received antipsychotics by the delirium status, subtype, and severity. The performance of the four claims-based algorithms was assessed in terms of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio, and negative likelihood ratio. The exact binomial 95% confidence interval was computed. We also evaluated the performance of the algorithms separately in patients who had surgery before implementation of ICD-10 version on October 1, 2015 (N=128 and 50 with delirium) and after implementation (N=56 and 16 with delirium). The overall diagnostic accuracy was evaluated using Youden's Index, calculated as sensitivity + specificity − 1. To evaluate the potential impact of outcome misclassification due to variable accuracy of delirium-identification algorithms, we assessed the relative risk and 95% exact confidence interval estimates in a hypothetical claims-based study to evaluate the association of a drug on the delirium risk in 1000 exposed and 1000 unexposed patients. All statistical analyses were performed using Stata 14 (College Station, TX).

RESULTS

The study population had a mean age of 81.4 years (standard deviation [SD]: 6.4) and consisted of 52.7% male and 88.6% whites. They had, on average, 14 years of education (SD: 3.1) and scored 25.6 points on the preoperative Mini-Mental State Examination (SD: 3.2). After transcatheter aortic valve replacement (n=109) or surgical aortic valve replacement (n=75), they had a median of 3 CAM assessments (interquartile range [IQR]: 2 to 5). Postoperative delirium occurred in 66 patients (35.9%) and lasted a median of two days (IQR: 1 to 3). Hypoactive delirium (n=35) was more prevalent than normoactive type (n=17) or hyperactive or mixed type (n=14). The median peak CAM-S score in 66 delirious patients was 9 (IQR: 7 to 10). Thirty-five patients (53.0%) had moderate delirium and 15 (22.7%) had severe delirium.

Of 184 patients, ICD diagnosis codes were present in 15 patients (8.2%) and antipsychotics were used in 13 patients (7.1%) (Table 1). Patients with delirium were more likely than those without to have ICD diagnosis codes (algorithm 1) (18.2% [corresponding to sensitivity] vs. 2.5% [corresponding to false positive fraction], p<0.01), antipsychotics use (algorithm 2) (18.2% vs. 0.9%, p<0.01), either ICD diagnosis codes OR antipsychotics use (algorithm 3) (30.3% vs. 3.4%), p<0.01), or both ICD diagnosis codes AND antipsychotics use (algorithm 4) (6.1% vs. 0.0%, p<0.01). Of note, 46 of 66 patients with delirium (69.7%) had neither ICD diagnosis codes nor antipsychotics.

In general, all four claims-based algorithms showed low sensitivity and high specificity in identifying delirium (Table 2). The algorithms based on ICD diagnosis codes alone (algorithm 1) and based on antipsychotics use alone (algorithm 2) had comparable sensitivity (18% vs. 18%, respectively), specificity (98% vs. 99%), and NPV (68% vs. 68%), but the latter showed slightly higher PPV (80% vs. 92%). When delirium was defined using either ICD diagnosis codes OR antipsychotics use (algorithm 3), sensitivity improved to 30% with little change in specificity, PPV, or NPV. Requiring both ICD diagnosis codes AND antipsychotics use (algorithm 4) resulted in perfect specificity and PPV, but considerably lower sensitivity. The algorithm 2 showed the highest positive likelihood ratio of 21.5, whereas the algorithm 3 had the lowest negative likelihood ratio of 0.72 and the largest Youden's Index (0.27). No formal comparison between ICD-9 code-based algorithms and ICD-10 code-based algorithms could be made due to the limited number of assessments based on the ICD-10 version. Among those with delirium who received antipsychotics, initiation was a median of 2.5 days (IQR: 1 to 10) after the first day of a positive CAM.

In a hypothetical claims-based study to evaluate the association of a drug on the delirium risk in 1000 exposed and 1000 unexposed patients (Table 3), the algorithm with perfect specificity and low sensitivity (algorithm 4) gives the unbiased relative risk estimate with a wider confidence interval. Algorithms with near-perfect specificity and better sensitivity (algorithms 1, 2, and 3) underestimate the relative risk, but with a narrower confidence interval.

Among delirious patients (Figure, Panel A), ICD diagnosis codes (algorithm 1) appeared more often in patients with hyperactive or mixed delirium (42.9%) than in those with hypoactive (14.3%; p=0.03) or normoactive delirium (5.9%; p=0.01). No significant difference was found between hypoactive and normoactive subtypes (p=0.37). Antipsychotics (algorithm 2) were used more often in patients with hyperactive or mixed subtype (28.6%; p=0.02) or hypoactive subtype (22.9%; p=0.03) than in those with normoactive subtype (0%), with no difference between hyperactive or mixed subtype and hypoactive subtype (p=0.67). The presence of ICD codes OR antipsychotic use (algorithm 3) was more common with hyperactive or mixed delirium (64.3%) than with hypoactive (28.6%; p=0.02) or normoactive delirium (5.9%; p<0.01). The difference between hypoactive and normoactive delirium was not statistically significant (p=0.06). The presence of ICD codes AND antipsychotic use (algorithm 4) was not significantly more common with hyperactive or mixed subtype (7.1%; p=0.26) or hypoactive subtype (8.6%; p=0.21) than normoactive subtype (0%), with no difference between hyperactive or mixed subtype and hypoactive subtype (p=0.87).

None of the algorithms could detect mild delirium (Figure, Panel B). Except the algorithm based on both ICD diagnosis codes AND antipsychotics use (algorithm 4), the algorithms were more likely to detect more severe forms of delirium.

DISCUSSION

Identification of delirium in large-scale epidemiological studies using health care utilization data is essential to evaluate the population burden, risk factors, and health care utilization outcomes of delirium. In this study, we evaluated four claims-based algorithms based on the ICD diagnosis codes for delirium and antipsychotic use against the validated CAM in cardiac surgery patients. Although all four algorithms showed low sensitivity and high specificity, we found that defining delirium using the ICD diagnosis codes OR antipsychotics use had the highest sensitivity (30%) and maintained high specificity (97%). In particular, this algorithm showed higher sensitivity to detect hyperactive or mixed delirium (64.3%) and severe delirium (73.3%).

Our findings have important implications for conducting large-scale epidemiological studies on delirium using inpatient claims database, such as Premier Database (www.premierinc.com) or Vizient Database (www.vizientinc.com), or hospital clinical data repositories that contain ICD diagnosis codes and inpatient medication use. Low sensitivity of claims-based algorithms results in underestimation of the prevalence of delirium, in particular, delirium without hyperactive psychomotor symptoms or mild delirium. Previous studies reported that 1–22% of hospitalized patients had explicit or implicit ICD diagnosis codes indicating delirium[11,13–15,26,27] and that 4–9% of patients received antipsychotics during non-psychiatric hospitalization.[16–18,28] These estimates are substantially lower than the rates of delirium diagnosed according to the reference standard: 11–29% in medical wards, 11–51% in surgical wards, and 19–82% in the intensive care unit.[2] Accordingly, claims-based algorithms do not allow an accurate estimation of the delirium prevalence in health care utilization databases; yet claims-based algorithms could be used to estimate the prevalence of delirium with hyperactive psychomotor symptoms or severe delirium, given the reasonable sensitivities of 64.3% and 73.3%, respectively. In addition, high specificity, high PPV, and high positive likelihood ratio of claims-based algorithms ensure that the majority of patients detected by the algorithms have delirium. Thus, claims-based algorithms can be useful to select patients with delirium in health care utilization databases for comparative effectiveness and safety research. When these algorithms are used to identify delirium as a study outcome and the algorithm performance is non-differential between the exposed and unexposed groups, the relative risk estimates for a binary exposure are unbiased or minimally biased toward the null; low sensitivity decreases statistical power. Sensitivity analysis methods based on algorithm performance are available to assess the impact of outcome misclassification.[29–31]

A few studies evaluated the performance of algorithms based on ICD diagnosis codes against a reference-standard diagnostic method (Table 4). Two studies that examined ICD diagnosis codes only in patients with confirmed delirium reported sensitivity of 9%[12] and 28%.[10] Other studies reported sensitivity of 3–36%, specificity of 85–99%, PPV of 38–83%, and NPV of 67–90%.[11,13–15] In our study, the same list of 32 ICD-9 codes and related ICD-10 codes used by Bui et al. showed somewhat lower sensitivity (18%), yet similar specificity (98%), PPV (80%), and NPV (68%). In general, studies that considered a larger number of diagnosis codes to define delirium tended to have higher sensitivity. Differences in patient population, clinical practice types, and documentation practice across hospitals may be responsible for choice of coding and inconsistencies across studies. The performance of antipsychotics use in identifying delirium has been evaluated by Herzig et al.,[16] Loh et al.,[17] and Rothberg et al.[18] In these studies, antipsychotics use had a PPV of 55–90% against the medical record-based definition of delirium. Since these investigators only reviewed medical records of patients who received antipsychotics, they were unable to estimate sensitivity, specificity, and NPV. Moreover, considering the incompleteness of documentation of delirium in medical records,[10–12] their PPV estimates may have been underestimated. We found that antipsychotics use had similar sensitivity and specificity to ICD diagnosis codes and, when combined, it could offer better performance than the algorithm based on either information alone. However, the performance of algorithms based on antipsychotics use may depend on antipsychotic prescribing practice in hospitalized patients. Use of antipsychotics for delirium is not approved nor supported by evidence.[32] The off-label use may be clinically justified if behavioral symptoms of delirium interfere with delivery of life-sustaining treatments or pose threats to safety of patients or others. More antipsychotic use in patients with hypoactive delirium (22.9%) than in those with normoactive delirium (0%) raises a suspicion that antipsychotic use might be implicated in conversion of normoactive subtype to hypoactive subtype. Another possible explanation is that some clinicians may use antipsychotics in all types of delirium. We expect that more judicious prescribing of antipsychotics may result in further decrease in sensitivity but increase in specificity for delirium with hyperactivity or severe delirium.

Our study has a few limitations. The algorithm performance was evaluated in a cohort of patients undergoing aortic valve replacement at a single academic medical center. Although inclusion of patients with Mini-Mental State Examination score ≥15 to ensure the quality of self-reported data may have excluded those at the highest risk for postoperative delirium, elective surgery for patients with severe cognitive impairment is not common. Therefore, our results may not be generalizable to patients on the medical service, those having non-cardiac or emergent surgery, or treated at community hospitals. Given the time and resource to conduct a detailed bedside assessment of delirium and its severity on a large scale, it will be useful to link the existing cohort studies designed to study delirium to electronic medical records or administrative claims database for additional algorithm validation. Another limitation is that the onset of delirium cannot be reliably determined using claims-based algorithms during the hospitalization. Although a previous study defined the onset of delirium using the date of antipsychotic initiation,[18] variable lag time between the first day of a positive CAM and initiation of antipsychotics makes this approach less reliable. In addition, the small sample size of our study did not allow an accurate assessment of algorithm performance based on ICD-10 diagnosis codes. The observed disparity in algorithm performance between ICD-9 and ICD-10 diagnosis codes could be due to the hospital's inexperience in coding delirium using the ICD-10 version, which may improve over time. Algorithms using ICD-10 diagnosis codes should be validated, once more ICD-10 data accumulate in the future.

In conclusion, the algorithm that defines delirium based on presence of one or more explicit or implicit ICD diagnosis codes or antipsychotic use can identify delirium with sensitivity 30%, specificity 97%, and PPV 83%. This algorithm performs modestly better than the existing algorithms based on either information alone. These results should be useful to conduct comparative effectiveness and safety research of drugs on delirium using administrative claims and drug utilization data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was funded by a KL2 Medical Research Investigator Training award (1KL2 TR001100-01) to D.H. Kim and an additional support (UL1 TR001102) from Harvard Catalyst / Harvard Clinical and Translational Science Center; the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health; and by the Training Program in Cardiovascular Research award (T32-HL007374) to C.A. Kim from the National Heart, Lung, and Blood Institute, National Institutes of Health.

D.H. Kim is supported by the Paul B. Beeson Clinical Scientist Development Award in Aging (K08AG051187) from the National Institute on Aging, American Federation for Aging Research, The John A. Hartford Foundation, and The Atlantic Philanthropies. K.F. Huybrechts is supported by grant K01MH099141 from the National Institute of Mental Health. B.T. Bateman is supported by grant K08HD075831 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development. E.R. Marcantonio is funded by grants R01AG030618, P01AG031720, and a Mid-Career Investigator Award (K24AG035075) from the National Institute on Aging.

The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic health care centers, or the National Institutes of Health.

Sponsor's Role: The funding sources did not have any role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5. Arlington, VA: 2013.
2. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. Lancet. 2014; 383(9920): 911–922. [PubMed: 23992774]
3. Marcantonio ER, Goldman L, Mangione CM, et al. A clinical prediction rule for delirium after elective noncardiac surgery. Jama. 1994; 271(2):134–139. [PubMed: 8264068]
4. Robinson TN, Raeburn CD, Tran ZV, Angles EM, Brenner LA, Moss M. Postoperative delirium in the elderly: risk factors and outcomes. Ann Surg. 2009; 249(1):173–178. [PubMed: 19106695]
5. Gleason LJ, Schmitt EM, Kosar CM, et al. Effect of Delirium and Other Major Complications on Outcomes After Elective Surgery in Older Adults. JAMA Surg. 2015; 150(12):1134–1140. [PubMed: 26352694]
6. Saczynski JS, Marcantonio ER, Quach L, et al. Cognitive trajectories after postoperative delirium. N Engl J Med. 2012; 367(1):30–39. [PubMed: 22762316]
7. Brummel NE, Jackson JC, Pandharipande PP, et al. Delirium in the ICU and subsequent long-term disability among survivors of mechanical ventilation. Crit Care Med. 2014; 42(2):369–377. [PubMed: 24158172]
8. Leslie DL, Marcantonio ER, Zhang Y, Leo-Summers L, Inouye SK. One-year health care costs associated with delirium in the elderly population. Arch Intern Med. 2008; 168(1):27–32. [PubMed: 18195192]
9. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med. 1990; 113(12):941–948. [PubMed: 2240918]
10. Hope C, Estrada N, Weir C, Teng CC, Damal K, Sauer BC. Documentation of delirium in the VA electronic health record. BMC Res Notes. 2014; 7:208. [PubMed: 24708799]
11. Inouye SK, Leo-Summers L, Zhang Y, Bogardus ST Jr, Leslie DL, Agostini JV. A chart-based method for identification of delirium: validation compared with interviewer ratings using the confusion assessment method. J Am Geriatr Soc. 2005; 53(2):312–318. [PubMed: 15673358]
12. Johnson JC, Kerse NM, Gottlieb G, Wanich C, Sullivan E, Chen K. Prospective versus retrospective methods of identifying patients with delirium. J Am Geriatr Soc. 1992; 40(4):316–319. [PubMed: 1556357]
13. Katznelson R, Djaiani G, Tait G, et al. Hospital administrative database underestimates delirium rate after cardiac surgery. Can J Anaesth. 2010; 57(10):898–902. [PubMed: 20645040]
14. Bui LN, Pham VP, Shirkey BA, Swan JT. Effect of delirium motoric subtypes on administrative documentation of delirium in the surgical intensive care unit. J Clin Monit Comput. 2016
15. Campbell NL, Cantor BB, Hui SL, et al. Race and documentation of cognitive impairment in hospitalized older adults. J Am Geriatr Soc. 2014; 62(3):506–511. [PubMed: 24576177]
16. Herzig SJ, Rothberg MB, Guess JR, et al. Antipsychotic Use in Hospitalized Adults: Rates, Indications, and Predictors. J Am Geriatr Soc. 2016; 64(2):299–305. [PubMed: 26889839]
17. Loh KP, Ramdass S, Garb JL, Brennan MJ, Lindenauer PK, Lagu T. From hospital to community: use of antipsychotics in hospitalized elders. J Hosp Med. 2014; 9(12):802–804. [PubMed: 25345721]
18. Rothberg MB, Herzig SJ, Pekow PS, Avrunin J, Lagu T, Lindenauer PK. Association between sedating medications and delirium in older inpatients. J Am Geriatr Soc. 2013; 61(6):923–930. [PubMed: 23631415]
19. Albert MS, Levkoff SE, Reilly C, et al. The delirium symptom interview: an interview for the detection of delirium symptoms in hospitalized patients. J Geriatr Psychiatry Neurol. 1992; 5(1): 14–21. [PubMed: 1571069]
20. Wong CL, Holroyd-Leduc J, Simel DL, Straus SE. Does this patient have delirium?: value of bedside instruments. Jama. 2010; 304(7):779–786. [PubMed: 20716741]

21. Wei LA, Fearing MA, Sternberg EJ, Inouye SK. The Confusion Assessment Method: a systematic review of current usage. J Am Geriatr Soc. 2008; 56(5):823–830. [PubMed: 18384586]
22. Ely EW, Inouye SK, Bernard GR, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). Jama. 2001; 286(21):2703–2710. [PubMed: 11730446]
23. Ely EW, Margolin R, Francis J, et al. Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). Crit Care Med. 2001; 29(7):1370–1379. [PubMed: 11445689]
24. Albrecht JS, Marcantonio ER, Roffey DM, et al. Stability of postoperative delirium psychomotor subtypes in individuals with hip fracture. J Am Geriatr Soc. 2015; 63(5):970–976. [PubMed: 25943948]
25. Inouye SK, Kosar CM, Tommet D, et al. The CAM-S: development and validation of a new scoring system for delirium severity in 2 cohorts. Ann Intern Med. 2014; 160(8):526–533. [PubMed: 24733193]
26. Redelmeier DA, Thiruchelvam D, Daneman N. Delirium after elective surgery among elderly patients taking statins. Cmaj. 2008; 179(7):645–652. [PubMed: 18809895]
27. Swan JT, Fitousis K, Hall JB, Todd SR, Turner KL. Antipsychotic use and diagnosis of delirium in the intensive care unit. Crit Care. 2012; 16(3):R84. [PubMed: 22591601]
28. Herzig SJ, Rothberg MB, Guess JR, Gurwitz JH, Marcantonio ER. Antipsychotic medication utilization in nonpsychiatric hospitalizations. J Hosp Med. 2016; 11(8):543–549. [PubMed: 27130311]
29. Lanes S, Brown JS, Haynes K, Pollack MF, Walker AM. Identifying health outcomes in healthcare databases. Pharmacoepidemiol Drug Saf. 2015; 24(10):1009–1016. [PubMed: 26282185]
30. Greenland S. Statistical uncertainty due to misclassification: implications for validation substudies. J Clin Epidemiol. 1988; 41(12):1167–1174. [PubMed: 3062138]
31. Brenner H, Gefeller O. Use of the positive predictive value to correct for disease misclassification in epidemiologic studies. Am J Epidemiol. 1993; 138(11):1007–1015. [PubMed: 8256775]
32. Neufeld KJ, Yue J, Robinson TN, Inouye SK, Needham DM. Antipsychotic Medication for Prevention and Treatment of Delirium in Hospitalized Adults: A Systematic Review and Meta-Analysis. J Am Geriatr Soc. 2016; 64(4):705–714. [PubMed: 27004732]

Take-Home Points

- Algorithms based on the International Classification of Diseases diagnosis codes and antipsychotic use in the inpatient administrative data have low sensitivity (6–30%), high specificity (97–100%), and high positive predictive value (80–100%) against the Confusion Assessment Method in detecting postoperative delirium.
- These algorithms are more sensitive in detection of hyperactive or mixed delirium (64%) or severe delirium (73%).
- Given high specificity and high positive predictive values, delirium-identification algorithms can be used to study delirium as an outcome in claims-based comparative effectiveness and safety studies of drugs.

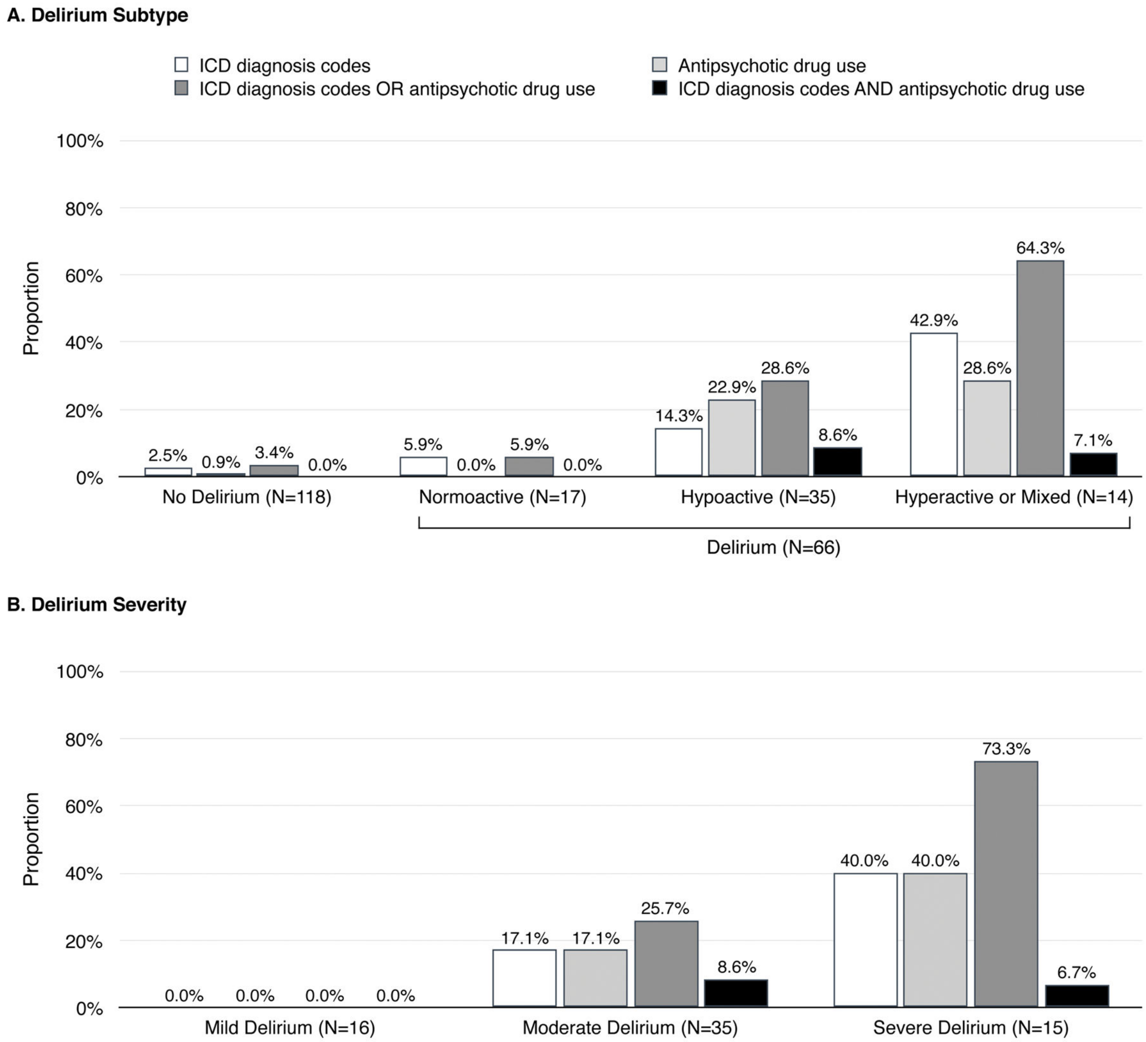

Figure. Delirium Subtype and Frequency of Delirium Diagnosis Codes and Antipsychotic Medication Use in Administrative Claims and Drug Utilization Data[a]

Abbreviation: ICD, International Classification of Diseases.

[a]ICD-9 diagnosis codes (before 10/1/2015) and ICD-10 diagnosis codes (since 10/1/2015) are listed in Supplementary Table S1.

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Probability of Delirium According to the International Classification of Diseases Diagnosis Codes and Antipsychotic Medication Use[a]

	ICD Diagnosis Codes		
APM Use	**Present**	**Absent**	**Total**
Yes	4 / 4 (100)	8 / 9 (88.9)	12 / 13 (92.3)
No	8 / 11 (72.7)	46 / 160 (28.8)	54 / 171 (31.6)
Total	12 / 15 (80.0)	54 / 169 (32.0)	66 / 184 (35.9)

Abbreviations: APM, antipsychotic medication; ICD, International Classification of Diseases.

[a]N with delirium / N total (%) is presented for each combination.

Table 2

Accuracy of Algorithms To Identify Delirium in Administrative Claims and Drug Utilization Data

Algorithms	Delirium $N_{positive}$ / N_{total}	No Delirium $N_{positive}$ / N_{total}	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)	LR− (95% CI)
1. ICD diagnosis codes	12 / 66	3 / 118	0.18 (0.10, 0.30)	0.98 (0.93, 1.00)	0.80 (0.52, 0.96)	0.68 (0.60, 0.75)	7.15 (2.09, 24.4)	0.84 (0.75, 0.94)
1) ICD-9 codes[a]	10 / 50	1 / 78	0.20 (0.10, 0.34)	0.99 (0.93, 1.00)	0.91 (0.59, 1.00)	0.66 (0.57, 0.74)	15.6 (2.06, 118)	0.81 (0.70, 0.93)
2) ICD-10 codes[a]	2 / 16	2 / 40	0.13 (0.02, 0.38)	0.95 (0.83, 0.99)	0.50 (0.07, 0.93)	0.73 (0.59, 0.84)	2.50 (0.38, 16.3)	0.92 (0.76, 1.12)
2. APM use	12 / 66	1 / 118	0.18 (0.10, 0.30)	0.99 (0.95, 1.00)	0.92 (0.64, 1.00)	0.68 (0.61, 0.75)	21.5 (2.85, 161)	0.83 (0.74, 0.93)
3. ICD diagnosis codes OR APM use	20 / 66	4 / 118	0.30 (0.20, 0.43)	0.97 (0.92, 0.99)	0.83 (0.63, 0.95)	0.71 (0.64, 0.78)	8.94 (3.19, 25.0)	0.72 (0.61, 0.85)
1) ICD-9 codes[a] OR APM use	16 / 50	2 / 78	0.32 (0.20, 0.47)	0.97 (0.91, 1.00)	0.89 (0.65, 0.99)	0.69 (0.60, 0.78)	12.5 (3.00, 52.0)	0.70 (0.58, 0.72)
2) ICD-10 codes[a] OR APM use	4 / 16	2 / 40	0.25 (0.07, 0.52)	0.95 (0.83, 0.99)	0.67 (0.22, 0.96)	0.76 (0.62, 0.87)	5.00 (1.01, 24.6)	0.79 (0.59, 1.06)
4. ICD diagnosis codes AND APM use	4 / 66	0 / 118	0.06 (0.02, 0.15)	1.00 (0.97, 1.00)[b]	1.00 (0.40, 1.00)[b]	0.66 (0.58, 0.73)	Not estimated[d]	0.94 (0.88, 1.00)
1) ICD-9 codes[a] AND APM use	4 / 50	0 / 78	0.08 (0.02, 0.19)	1.00 (0.95, 1.00)[b]	1.00 (0.40, 1.00)[b]	0.63 (0.54, 0.71)	Not estimated[d]	0.92 (0.85, 1.00)
2) ICD-10 codes[a] AND APM use	0 / 16	0 / 40	0.00 (0.00, 0.21)[b]	1.00 (0.91, 1.00)[b]	Not estimated[c]	0.71 (0.58, 0.83)	Not estimated[d]	1.00 (1.00, 1.00)

Abbreviations: APM, antipsychotic medication; CI, confidence interval; ICD, International Classification of Diseases; LR, likelihood ratio; NPV, negative predictive value; PPV, positive predictive value.

[a] ICD-9 diagnosis codes (before 10/1/2015) and ICD-10 diagnosis codes (since 10/1/2015) are listed in Supplementary Table S1.

[b] One-sided 97.5% confidence interval was estimated.

[c] Not estimated because all patients were classified as negative according to the algorithm.

[d] Not estimated due to perfect specificity.

Table 3

Outcome Misclassification in a Hypothetical Study of Drug Exposure on Delirium[a]

Delirium-Identification Algorithm	Drug Exposure	Delirium Status: Present	Delirium Status: Absent	Delirium Risk	RR (95% CI)
1. Perfect delirium assessment	Exposed	300	700	30.0%	3.00 (2.44, 3.70)
Sensitivity 100%	Unexposed	100	900	10.0%	
Specificity 100%	Total	400	1600	2000	
2. Delirium assessment using Algorithm 1	Exposed	68[b]	932[c]	6.8%	1.89 (1.27, 2.80)
Sensitivity 18%	Unexposed	36[b]	964[c]	3.6%	
Specificity 98%	Total	104	1896	2000	
3. Delirium assessment using Algorithm 2	Exposed	61	939	6.1%	2.26 (1.45, 3.52)
Sensitivity 18%	Unexposed	27	973	2.7%	
Specificity 99%	Total	88	1912	2000	
4. Delirium assessment using Algorithm 3	Exposed	111	889	11.1%	1.95 (1.43, 2.65)
Sensitivity 30%	Unexposed	57	943	5.7%	
Specificity 97%	Total	168	1832	2000	
5. Delirium assessment using Algorithm 4	Exposed	18	982	1.8%	3.00 (1.20, 7.53)
Sensitivity 6%	Unexposed	6	994	0.6%	
Specificity 100%	Total	24	1976	2000	

Abbreviations: CI, confidence interval; RR, relative risk.

[a]A hypothetical study involved 1000 exposed and 1000 unexposed patients. No selection bias, no misclassification of drug exposure status, and no confounding are assumed. Outcome misclassification is non-differential with regards to the exposure status.

[b]Exposed delirium cases are computed from true positive cases (300 × 0.18 = 54) and false positive cases (700 × 0.02 = 14). Unexposed delirium cases are computed from true positive cases (100 × 0.18 = 18) and false positive cases (900 × 0.02 = 18).

[c]Exposed non-delirium cases are computed from true negative cases (700 × 0.98 = 686) and false negative cases (300 × 0.82 = 246). Unexposed non-delirium cases are computed from true negative cases (900 × 0.98 = 882) and false negative cases (100 × 0.82 = 82).

Table 4

Delirium-Identification Algorithms in Administrative Claims and Drug Utilization Database in the Literature

Algorithms (Reference standard)	Population (N delirium / N total)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Johnson 1992[12] 14 ICD-9 diagnosis codes[a] (DSM III)	Medical (47/235)	0.09 (0.02, 0.20)	NR	NR	NR
Inouye 2005[11] 8 ICD-9 diagnosis codes[b] (CAM)	Medical (115/919)	0.03 (0.01, 0.07)	0.99 (0.99, 1.00)	0.38 (0.09, 0.76)	0.88 (0.85, 0.90)
Katznelson 2010[13] ICD-10 diagnosis codes[c] (CAM-ICU)	Cardiac surgery (182/1528)	0.18 (0.13, 0.25)	0.99 (0.98, 1.00)	0.72 (0.56, 0.84)	0.90 (0.88, 0.91)
Campbell 2014[15] ICD-9 diagnosis codes[c] (CAM)	Medical (163/424)	0.32 (0.25, 0.40)	0.85 (0.80, 0.89)	0.57 (0.46, 0.67)	0.67 (0.61, 0.72)
Hope 2014[10] 8 ICD-9 diagnosis codes[d] (expert consensus)	Medical, surgical, ICU (25/NR)	0.28 (0.12, 0.49)	NR	NR	NR
Bui 2016[14] 32 ICD-9 diagnosis codes[e] (CAM-ICU)	Surgical ICU (423/1055)	0.36 (0.31, 0.40)	0.95 (0.93, 0.97)	0.83 (0.76, 0.88)	0.69 (0.66, 0.72)
Rothberg 2013[18] Use of any of 4 APMs[f] (medical record)	Medical (NR)	NR	NR	0.55 (0.39, 0.70)	NR
Loh 2014[17] Use of any of 5 APMs[g] (medical record)	Medical, surgical, ICU (NR)	NR	NR	0.83 (0.78, 0.87)	NR
Herzig 2016[26] Use of any of 20 APMs[h] (medical record)	Medical, surgical, ICU (NR)	NR	NR	0.65 (0.55, 0.74)	NR

Abbreviations: APM, antipsychotic medications; CAM, Confusion Assessment Method; CI, confidence interval; DSM, Diagnostic and Statistical Manual of Mental Disorders; ICD, International Classification of Diseases; ICU, intensive care unit; NPV, negative predictive value; NR, not reported; PPV, positive predictive value.

[a]ICD-9 diagnosis codes include 290.11, 290.3, 290.41, 291.0, 292.0, 292.1, 292.12, 292.2, 292.81, 292.9, 293.0, 293.1, 293.9, and 780.0.

[b]ICD-9 diagnosis codes include 290.11, 290.3, 290.41, 291.0, 292.81, 293.0, 293.1, and 780.09.

[c]ICD diagnosis codes were not specified.

[d]ICD-9 diagnosis codes include 290.3, 291, 291.1, 292, 292.81, 293, 293.1, and 780.09.

[e]ICD-9 diagnosis codes include 290.11–290.13, 290.20, 290.3, 290.41–290.43, 290.8, 290.9, 291.0, 292.0–292.2, 292.81, 292.82, 293.0–293.9, 348.30, 348.31, 348.39, 349.82, 780.02, 780.09, and 780.97.

[f]Antipsychotic medications included aripiprazole, haloperidol, risperidone, and ziprasidone, and excluded olanzapine.

[g]Antipsychotic medications included haloperidol, thorazine, olanzapine, quetiapine, and risperidone.

[h] Antipsychotic medications included aripiprazole, asenapine, chlorpromazine, clozapine, fluphenazine, haloperidol, iloperidone, loxapine, lurasidone, molindone, olanzapine, paliperidone, perphenazine, pimozide, quetiapine, risperidone, thioridazine, thiothixene, trifluoperazine, and ziprasidone.