Observational Study

Medicine®

OPEN

Assessment of obesity prevalence and validity of obesity diagnoses coded in claims data for selected surgical populations

A retrospective, observational study

Eric M. Ammann, PhD[a], Iftekhar Kalsekar, PhD[a], Andrew Yoo, MD, MS[a], Robin Scamuffa, MS[b], Chia-Wen Hsiao, MHA[c], Andrew C. Stokes, PhD[d], John M. Morton, MD, MPH, MHA[e], Stephen S. Johnston, MA[a,*]

Abstract

In many types of surgery, obesity may influence patient selection, prognosis, and/or management. Quantifying the accuracy of the coding of obesity and other prognostic factors is important for the design and interpretation of studies of surgical outcomes based on administrative healthcare data. This study assessed the validity of obesity diagnoses recorded in insurance claims data in selected surgical populations.

This was a retrospective, observational study. Deidentified electronic health record (EHR) and linked administrative claims data were obtained for US patients age ≥20 years who underwent a qualifying surgical procedure (bariatric surgery, total knee arthroplasty [TKA], cardiac ablation, or hernia repair) in 2014Q1–2017Q1 (first = index). Patients' body mass index (BMI) as coded in the claims data (error-prone measure) during the index procedure or 180d pre-index was compared with their measured BMI as recorded in the EHR (criterion standard) to estimate the sensitivity and positive predictive value (PPV) of obesity diagnosis codes.

Among patients who underwent bariatric surgery (N = 1422), TKA (N = 8670), cardiac ablation (N = 167), or hernia repair (N = 5450), obesity was present in 98%, 63%, 52%, and 54%, respectively, based on measured BMI. PPVs of obesity diagnosis codes were high: 99.3%, 96.0%, 92.8%, and 94.1% in bariatric surgery, TKA, cardiac ablation, and hernia repair, respectively. The sensitivity of obesity diagnoses was: 99.8%, 46.2%, 41.3%, and 42.3% in bariatric surgery, TKA, cardiac ablation, and hernia repair, respectively. Among false-positive patients diagnosed as obese but with measured BMI <30, the proportion with a BMI ≥28 was 40.0%, 67.6%, 60.7%, and 65.8% for bariatric surgery, TKA, cardiac ablation, and hernia repair, respectively.

Our data indicate that obesity is highly prevalent in many surgical populations, obesity diagnosis codes have high PPVs, but also obesity is generally undercoded in claims data. Quantifying the validity of diagnosis codes for obesity and other important prognostic factors is important for the design and interpretation of studies of surgical outcomes based on administrative data. Further research is needed to determine the extent to which undercoding of BMI and obesity can be addressed through the use of proxies that may be better documented in claims data.

Abbreviations: BMI = body mass index, EHR = electronic helath record, HCPCS = healthcare common procedure coding system, *ICD-10-CM = International Classification of Diseases, 10th Revision, Clinical Modification*, *ICD-9-CM = International Classification of Diseases, 9th Revision, Clinical Modification*, IRB = institutional review board, NPV = negative predictive value, PPV = positive predictive value, TKA = total knee arthroplasty.

Keywords: healthcare administrative data, obesity, operative, surgical procedures, validation

Editor: Amede Gogovor.

Ethics approval and consent to participate: The Optum Integrated Claims-Clinical Database consists of de-identified healthcare records. Use of the database was reviewed by the New England Institutional Review Board (IRB), which determined that it did not constitute human subjects research, and thus was exempt from IRB review and registration requirements. Throughout the research project, the study data remained de-identified and stored on encrypted, password-protected servers to maintain patient confidentiality.

Availability of data and materials: The data that support the findings of this study are available from Optum but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Optum.

Competing interests: EA, IK, AY, and SJ are employees of Johnson & Johnson; RS and CH are employees of Ethicon, a Johnson & Johnson company. AS reports grants from Johnson & Johnson, outside the submitted work. JM has nothing to disclose.

Supplemental Digital Content is available for this article.

The study was funded by Johnson & Johnson, Inc.

The authors report no conflicts of interest.

[a] Epidemiology, Medical Devices, Johnson & Johnson, New Brunswick, NJ, [b] Clinical Research, Ethicon, [c] Health Economics & Market Access, Ethicon, Cincinnati, OH, [d] Department of Global Health, School of Public Health, Boston University, Boston, MA, [e] Department of Surgery, Stanford University Medical Center, Stanford, CA.

[*] Correspondence: Stephen S. Johnston, Johnson & Johnson, 410 George St., New Brunswick, NJ 08901 (e-mail: sjohn147@its.jnj.com).

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc.
This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

Medicine (2019) 98:29(e16438)

Received: 14 November 2018 / Received in final form: 25 April 2019 / Accepted: 16 June 2019

http://dx.doi.org/10.1097/MD.0000000000016438

1. Introduction

In many types of surgery, obesity is an important factor that is relevant for patient selection, prognosis, and/or management.[1] For example, obesity is a strong risk factor for the development of knee osteoarthritis, and patients with severe obesity who undergo total knee arthroplasty (TKA) are at higher risk of surgical complications and may experience worse long-term outcomes relative to patients without obesity who undergo TKA.[2,3]

Coding of obesity and BMI, which is based on *International Classification of Diseases, 9th and 10th Revision, Clinical Modification* (*ICD-9-CM/ICD-10-CM*) diagnosis codes, is known to be imperfect in healthcare administrative data. Recently, the US Food & Drug Administration issued guidance on the use of real-world data in regulatory decision-making for medical devices, endorsing its use for a variety of purposes, including product label extensions, enhanced post-market safety surveillance, and for identifying historical controls for single-arm interventional trials.[4] Quantifying the validity of diagnosis codes for obesity and other important prognostic factors is important for the design and interpretation of studies of surgical outcomes based on administrative data.

Previous studies in the United States and Canada indicate that diagnosis codes for obesity recorded in administrative data have a high positive predictive value (PPV; 66%–91%) but low sensitivity (8%–19%).[5–9] The present study builds on previous work by using contemporary insurance claims data and linked electronic health record (EHR) data to assess the PPV and sensitivity of obesity diagnosis codes in the following settings where obesity is an important clinical factor: bariatric surgery,[10,11] TKA for knee osteoarthritis,[2,3] cardiac ablation for atrial fibrillation,[12] and hernia repair.[13]

2. Methods

2.1. Data source

Data from 2013 to 2017 were obtained from the Optum Integrated Claims-Clinical Database, which includes linked and deidentified insurance claims and EHR data for commercially insured and Medicare Advantage health plan members.[14,15] The insurance claims files include facility, physician, and pharmacy claims submitted for reimbursement on behalf of covered health plan members. The EHR data include records of clinical diagnoses, procedures, body measurements, vital signs, laboratory results, prescriptions, and notes recorded as part of routine clinical practice in EHR systems that contribute data to the Optum healthcare analytics platform. Claims and EHR data were available for a geographically diverse population of 5.5 million health plan members in the United States who have been successfully linked to the Optum EHR database.

2.2. Protection of human subjects

The Optum Integrated Claims-Clinical Database consists of deidentified healthcare records. Use of the database was reviewed by the New England Institutional Review Board (IRB), which determined that it did not constitute human subjects research, and thus was exempt from IRB review and registration requirements. Throughout the research project, the study data remained deidentified and stored on encrypted, password-protected servers to maintain patient confidentiality.

2.3. Study population and design

The study population consisted of patients who underwent any of the following surgical procedures between January 1, 2014 and March 1, 2017: bariatric surgery (specifically, laparoscopic sleeve gastrectomy or Roux-en-Y gastric bypass; inpatient only); unilateral TKA for knee osteoarthritis (inpatient only); cardiac ablation for atrial fibrillation (inpatient or outpatient); or ventral/incisional, umbilical, or epigastric hernia repair (inpatient or outpatient). See Supplementary Table 1, Additional File 1, http://links.lww.com/MD/D125, for the codes used to identify these surgical procedures. For each patient, we restricted to the first surgery of each type that was observed during the study period, and refer to it below as the index procedure and the associated admission date (service date for outpatient procedures) as the index date.

Surgical patients were eligible for inclusion if the following criteria were met: continuous health plan enrollment from the index date minus 180 days through discharge; ≥1 BMI measure recorded in the EHR from the index date minus 180 days through discharge; measured BMI within 15 to 90 kg/m^2 (imposed as a quality control measure to guard against data entry errors); nonmissing values for demographic variables (age, sex, census region, type of insurance); and age ≥20 years. Patients younger than 20 years at index were excluded because clinical classifications and administrative coding of BMI differ for patients age <20 years in the *ICD-9-CM/ICD-10-CM*. Each surgical cohort was defined independently of the others, and an individual patient could contribute data to multiple cohorts provided that the criteria described above were satisfied.

2.4. BMI measures

The error-prone measures of BMI evaluated in this study were the administrative *ICD-9-CM/ICD-10-CM* diagnosis codes for overweight and obesity recorded in the Optum claims data during the index procedure or the 180 days before. In *ICD-9-CM* and *ICD-10-CM*, overweight and obesity may be coded as a metabolic disorder (eg, *ICD-9-CM* diagnosis code 278.00 "Obesity, unspecified") and specific BMI ranges can be coded as a health status diagnosis (eg, *ICD-9-CM* diagnosis code V85.30 "Body Mass Index 30.0–30.9, adult"). Both types of codes were used in this study (see Supplementary Table 2, Additional File 1, http://links.lww.com/MD/D125). If diagnoses for both overweight and obesity were coded for a patient, then the one recorded last was used for analysis.

This study did not seek to evaluate the accuracy of Healthcare Common Procedure Coding System (HCPCS) G-codes (used for reporting quality measures in fee-for-service Medicare Part B) for BMI screening. These codes (included at the end of Supplementary Table 2, Additional File 1, http://links.lww.com/MD/D125) do not provide specific information concerning a patient's BMI (eg, "BMI above normal parameters"), and were recorded for <0.1% of our study sample, which did not include any fee-for-service Medicare beneficiaries.

The criterion standard BMI measure in this study was BMI as recorded in the Optum EHR data during the index procedure or 180 days before. For patients with multiple BMI measurements recorded during this time period, the last BMI measurement was used for analysis. Measured BMI was used to classify patients as underweight (<18.5), normal weight (≥18.5 and <25), overweight (≥25 and <30), or obese (≥30) according to standard

clinical cutpoints.[16] In secondary analyses, we also classified patients as having "morbid obesity" (a term used but not explicitly defined in *ICD-9-CM* and *ICD-10-CM*) using 2 definitions (BMI ≥35 and ≥40) to reflect varied usage.

2.5. Estimation of validation statistics

In evaluating the validity of weight-related diagnosis codes from administrative data, the primary statistics of interest were sensitivity and PPV:

- Sensitivity—The sensitivity of administrative diagnosis codes for the detection overweight or obesity was calculated as the proportion of health plan members in the weight category based on measured BMI who had an accurate weight-related diagnosis coded. For example, in evaluating the sensitivity of administrative diagnosis codes for obesity, a patient with a BMI of 34 would be classified as a true-positive if any code for obesity was present, and as a false-negative otherwise.
- PPV—The PPV of administrative diagnosis codes for overweight and obesity was calculated as the proportion of health plan members coded as being in a weight category who were correctly classified as such. For example, a patient with an administrative diagnosis code for obesity (eg, *ICD-10* diagnosis code E66.9, "Obesity, unspecified") would be considered accurately classified if the patient's measured BMI was ≥30.

Sensitivity and PPV were the primary validation statistics of interest for this study; however, for completeness the specificity and negative predictive value (NPV) of administrative diagnosis codes for obesity and overweight were also estimated, as described below.

- Specificity—Specificity was calculated as the proportion of health plan members not in a specific weight category based on measured BMI (eg, not obese) who were not coded as such.
- NPV—NPV was calculated as the proportion of health plan members not coded as being in a particular weight category (eg, not coded as obese) who were in fact not in that weight category based on measured BMI.

Data management and preparation of the analytic dataset were done using the Instant Health Data platform (Boston Health Economics, Boston, MA). Statistical analyses were performed with SAS Enterprise Guide 7.1 for Windows (SAS Institute, Cary, NC).

2.6. Secondary analyses

To assess whether the validity of overweight/obesity diagnosis codes differed in the *ICD-9-CM* and *ICD-10-CM* coding eras, we classified surgical cases and their preceding 180-day lookback periods as being in the *ICD-9-CM* era (January 1, 2014, through August 31, 2015) or *ICD-10-CM* era (April 1, 2016, through March 1, 2017) based on index date. Sensitivity and PPV were calculated separately for the *ICD-9-CM* and *ICD-10-CM* eras as a secondary analysis. The interval from September 1, 2015 through March 30, 2016 was classified as a transition and washout period, and patients with an index date during this period were excluded from this secondary analysis.

An individual's BMI may change during the approximately 6-month cross-sectional assessment period used in this study. To evaluate whether a recorded weight-related diagnosis is more accurate if compared against a proximate BMI measurement, we recalculated our PPV estimates after restricting to health plan members wherein the weight-related diagnosis code was recorded within 60 days, 45 days, 30 days, or 15 days of the BMI measurement in the EHR. We hypothesized that PPVs would increase as the number of days between the diagnosis code and BMI measurement in EHR decreased.

Finally, to further characterize the nature of obesity false-positives, we characterized the distribution of BMI among patients who were diagnosed as obese in the administrative data but ultimately did not have BMI ≥30 in the EHR. We also calculated the proportion of such patients with a BMI ≥28 to examine how many may have been classified as true positives when allowing for deviation by up to 2 BMI points below the traditional obesity threshold of BMI ≥30.

2.7. Covariates

To characterize the study population, the following demographic and clinical characteristics were assessed: age at index procedure, sex, census region, type of health insurance (commercial or Medicare Advantage), Charlson Comorbidity Index using Quan's coding algorithms,[17] and whether the index procedure occurred in an inpatient or outpatient setting. The Charlson Comorbidity Index was assessed using administrative diagnoses coded during the index procedure or the 180 days before.

3. Results

3.1. Descriptive information

The study population consisted of 17,217 surgical cases from 2014 to 2017Q1 who met eligibility criteria, as shown in Table 1. Of this total, 1422 patients underwent bariatric surgery (median age = 48 years; 75% female; 100% inpatient surgery; 66% Roux-en-Y gastric bypass and 34% sleeve gastrectomy), 8670 unilateral TKA for knee osteoarthritis (median age = 68 years; 61% female; 100% inpatient), 1675 cardiac ablation for atrial fibrillation (N = 1675; median age = 65 years; 34% female; 59%

Table 1

Identification of eligible surgical cases with BMI recorded in the EHR.

Eligibility criterion	Bariatric surgery (N)	Total knee arthroplasty (N)	Cardiac ablation (N)	Hernia repair (N)
1. Index surgery from January 1, 2014 through March 1, 2017, with health plan enrollment and ≥1 BMI measurement recorded in EHR from 180 days before index surgery admission through discharge	1477	8823	1705	5684
2. BMI 15 –90 kg/m^2	1475	8819	1704	5667
3. Nonmissing values for demographic variables (age, sex, census region, type of insurance)	1430	8670	1677	5531
4. Age ≥20 y at index surgery	1422	8670	1675	5450

BMI = body mass index, EHR = electronic health record.

Table 2

Demographic and clinical characteristics of study population*.

Characteristic	Bariatric surgery (N=1422)	Total knee arthroplasty (N=8670)	Cardiac ablation (N=1675)	Hernia repair (N=5450)
Age, y, N (%)				
20–44	579 (41%)	82 (1%)	68 (4%)	1142 (21%)
45–64	695 (49%)	3087 (36%)	749 (45%)	2528 (46%)
65+	148 (10%)	5501 (63%)	858 (51%)	1780 (33%)
Sex, N (%)				
Female	1060 (75%)	5283 (61%)	568 (34%)	2366 (43%)
Male	362 (25%)	3387 (39%)	1107 (66%)	3084 (57%)
Calendar year of index admission				
2014	422 (30%)	2522 (29%)	489 (29%)	1633 (30%)
2015	455 (32%)	2651 (31%)	536 (32%)	1668 (31%)
2016–2017Q1	545 (38%)	3497 (40%)	650 (39%)	2149 (39%)
Census region, N (%)				
Midwest	423 (30%)	4445 (51%)	566 (34%)	2409 (44%)
Northeast	313 (22%)	650 (7%)	273 (16%)	559 (10%)
South	528 (37%)	2665 (31%)	670 (40%)	1869 (34%)
West	158 (11%)	910 (10%)	166 (10%)	613 (11%)
Insurance type, N (%)				
Commercial	1125 (79%)	3443 (40%)	919 (55%)	3664 (67%)
Medicare Advantage	297 (21%)	5227 (60%)	756 (45%)	1786 (33%)
CCI†, N (%)				
CCI score: 0	617 (43%)	4677 (54%)	653 (39%)	2922 (54%)
CCI score: 1	272 (19%)	1883 (22%)	205 (12%)	829 (15%)
CCI score: 2	286 (20%)	1135 (13%)	362 (22%)	833 (15%)
CCI score: 3+	247 (17%)	975 (11%)	455 (27%)	866 (16%)
Index procedure inpatient setting, N (%)	1422 (100%)	8670 (100%)	981 (59%)	1813 (33%)

CCI=Charlson Comorbidity Index.

* Some % sum to <100% across categories because of rounding error.

† CCI scores based on diagnoses recorded during the index procedure and the 180 days before.

inpatient), and 5450 hernia repair (median age=58 years; 43% female; 33% inpatient). Of the hernia repair patients, 53%, 45%, and 2% had ventral/incisional, umbilical, and epigastric hernias, respectively. Additional cohort characteristics are provided in Table 2.

Based on measured BMI recorded in the EHR, obesity was present in 98%, 63%, 52%, and 54%, of patients who underwent bariatric surgery, TKA, cardiac ablation, hernia repair, respectively. BMI distributions by surgery type are shown in Figure 1.

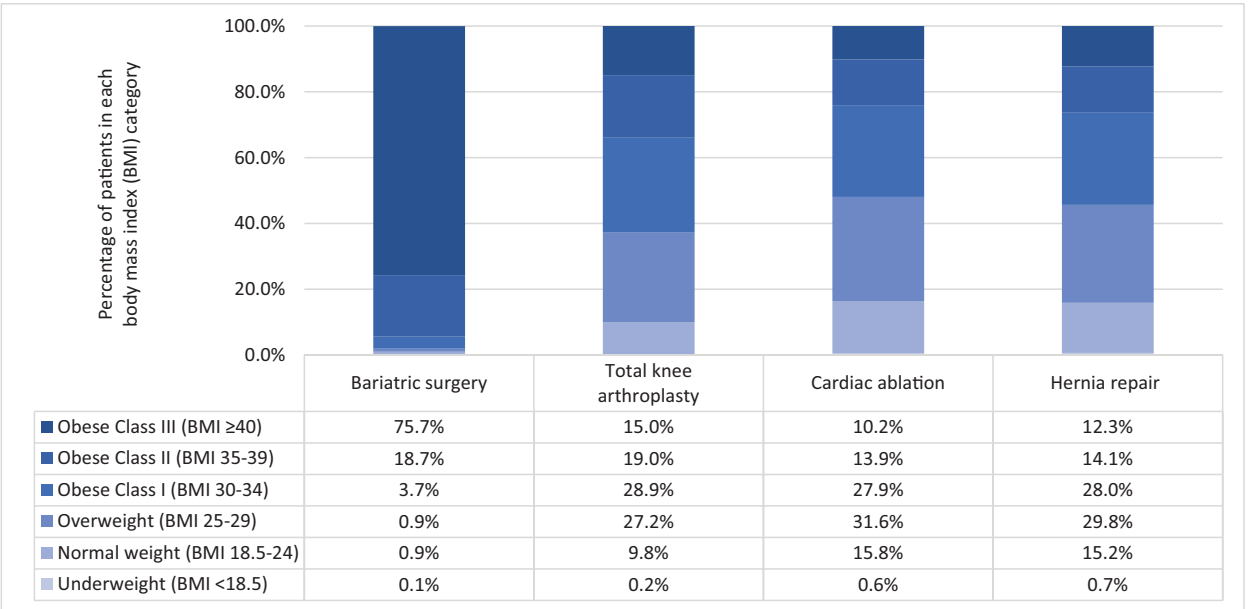

Figure 1. Distribution of measured BMI as recorded in electronic health record by surgery type. BMI=body mass index.

Table 3

Sensitivity of administrative diagnosis codes for detection of overweight, obesity and morbid obesity.

	Bariatric surgery		Total knee arthroplasty		Cardiac ablation		Hernia repair	
Measured BMI as recorded in HER	N	Sens.	N	Sens.	N	Sens.	N	Sens.
Overweight (BMI 25–29 kg/m^2)	13	0.0%	2,356	7.6%	530	6.4%	1622	6.4%
Obese (BMI ≥30 kg/m^2)	1394	99.8%	5445	46.2%	870	41.3%	2962	42.3%
Morbid obesity definition 1 (BMI ≥35 kg/m^2)	1342	99.3%	2943	48.5%	403	37.7%	1437	46.7%
Morbid obesity definition 2 (BMI ≥40 kg/m^2)	1076	95.6%	1297	74.3%	170	58.2%	671	62.6%

BMI = body mass index, EHR = electronic health record, Sens. = sensitivity.

In the claims data, obesity was coded for 98.5%, 30.3%, 23.1%, and 24.4% of bariatric surgery, TKA, cardiac ablation, and hernia repair patients, respectively; an additional 0.1%, 2.8%, 3.3%, and 2.6 were coded as being overweight. Of patients coded as obese or overweight, 33.9% had a metabolic disorder code (eg, *ICD-9-CM* diagnosis code 278.00 "Obesity, unspecified"), 14.3% had a health status code for a specific BMI range (eg, *ICD-9-CM* diagnosis code V85.30 "Body Mass Index 30.0–30.9, adult"), and 51.8% had both.

3.2. Validation statistics

The sensitivity of diagnoses recorded in the claims data for the detection of obesity was 99.8%, 46.2%, 41.3%, and 42.2% in patients who underwent bariatric surgery, TKA, cardiac ablation, and hernia repair, respectively. Sensitivity of diagnosis codes for overweight was <10% across all types of surgery. The sensitivity of diagnosis codes for morbid obesity was >95% in bariatric surgery but ranged from 37.7% to 74.3% in other surgeries depending on the surgery type and definition used for morbid obesity (≥35 or ≥40 kg/m^2). Detailed sensitivity statistics are provided in Table 3.

The PPV of administrative diagnosis codes for obesity was 99.3% in bariatric surgery patients and >90% for the other surgical cohorts (Table 4). PPVs for overweight ranged from 61% to 73% in TKA, cardiac ablation, and hernia repair patients. The PPV for overweight was 0% in bariatric surgery; however, only 1 patient met the BMI criteria for overweight in this cohort. PPVs for morbid obesity ranged from 63.5% to 96.9% depending on the surgery type and definition used for morbid obesity. Detailed PPV estimates are provided in Table 4. Continuous BMI distributions associated with individual *ICD-9-CM* and *ICD-10-CM* diagnosis codes are shown in Supplementary Figure 1 and Supplementary Figure 2, http://links.lww.com/MD/D125, Additional File 1, http://links.lww.com/MD/D125.

Specificity was high (≥98%) for overweight diagnosis codes in each surgical cohort. In patients who underwent TKA, cardiac ablation, or hernia repair, specificity was high (≥95%) for diagnosis codes corresponding to obesity and morbid obesity, but was lower among bariatric surgery patients (46%–69%). NPVs for obesity diagnosis codes ranged from 52% in TKA patients to 86% in bariatric surgery patients; NPVs were generally higher for diagnosis codes corresponding to overweight and morbid obesity. Details on the specificity and NPV of overweight/obesity diagnosis codes in each surgical cohort are provided in Supplementary Table 3 and Supplementary Table 4, Additional File 1, http://links.lww.com/MD/D125.

3.3. Secondary analyses

As described in the methods section, sensitivity and PPV were recalculated separately for patients in the *ICD-9-CM* and *ICD-10-CM* coding eras. In general, the sensitivity of administrative diagnosis codes for overweight/obesity was somewhat higher in the *ICD-10-CM* era; PPVs were similar across the 2 coding eras. Detailed validation statistics are provided in Supplementary Table 5 and Supplementary Table 6, Additional File 1, http://links.lww.com/MD/D125, and continuous BMI distributions associated with individual *ICD-9-CM* and *ICD-10-CM* diagnosis codes are shown in Supplementary Figure 1 and Supplementary Figure 2, Additional File 1, http://links.lww.com/MD/D125.

Among patients with a coded diagnosis of overweight or obesity, the median time interval between the diagnosis and the BMI measurement recorded in the EHR was 8 days (interquartile range: 0–51). Supplementary Table 7, Additional File 1, http://links.lww.com/MD/D125 shows PPVs recalculated when restricting to patients where this time interval was 60, 45, 30, or ≤15. Except for the bariatric surgery group, the recalculated PPVs for obesity diagnosis codes generally increased slightly as the number of days between the diagnosis code and BMI measurement in the EHR decreased. For example, among patients undergoing cardiac ablation, PPVs increased from 94.5% when restricting the time interval to ≤60 days to 96.9% when restricting the time interval to ≤45 days; such increases were attained at the expense of sample size. PPVs for

Table 4

PPVs of administrative diagnosis codes for overweight, obesity, and morbid obesity.

	Bariatric surgery		Total knee arthroplasty		Cardiac ablation		Hernia repair	
BMI category as coded in claims data	N	PPV	N	PPV	N	PPV	N	PPV
Overweight (BMI 25–29 kg/m^2)	1	0.0%	247	72.9%	56	60.7%	142	73.2%
Obese (BMI ≥30 kg/m^2)	1401	99.3%	2623	96.0%	387	92.8%	1329	94.1%
Morbid obesity definition 1 (BMI ≥35 kg/m^2)	1376	96.9%	1534	93.1%	171	88.9%	732	91.7%
Morbid obesity definition 2 (BMI ≥40 kg/m^2)	1137	90.5%	1271	75.8%	156	63.5%	600	70.0%

BMI = body mass index, PPV = positive predictive value.

overweight diagnosis codes changed nonmonitonically across the time intervals.

Supplementary Table 8, Additional File 1, http://links.lww.com/MD/D125, shows the distribution of BMI among false-positive patients who were diagnosed as obese in the administrative data but ultimately did not have BMI ≥30 in the EHR. Among these patients, BMIs recorded in the EHR tended to be near the threshold of BMI ≥30 for obesity. The proportion of such patients with a BMI ≥28 was 40.0%, 67.6%, 60.7%, and 65.8% for bariatric surgery, TKA, cardiac ablation, and hernia repair, respectively.

3.4. Comparison of BMI distributions in patients with actual versus coded obesity

To evaluate how representative patients coded as obese in administrative data are of the total population of obese patients, we compared BMI distributions of patients who were classified as obese in the administrative data or were in fact obese based on measured BMI. Little difference was found in the setting of bariatric surgery; however, for the other surgical cohorts, median BMI in patients coded as obese was 6.4% to 14.9% higher than median BMI among all patients who were in fact obese according to the EHR (Fig. 2).

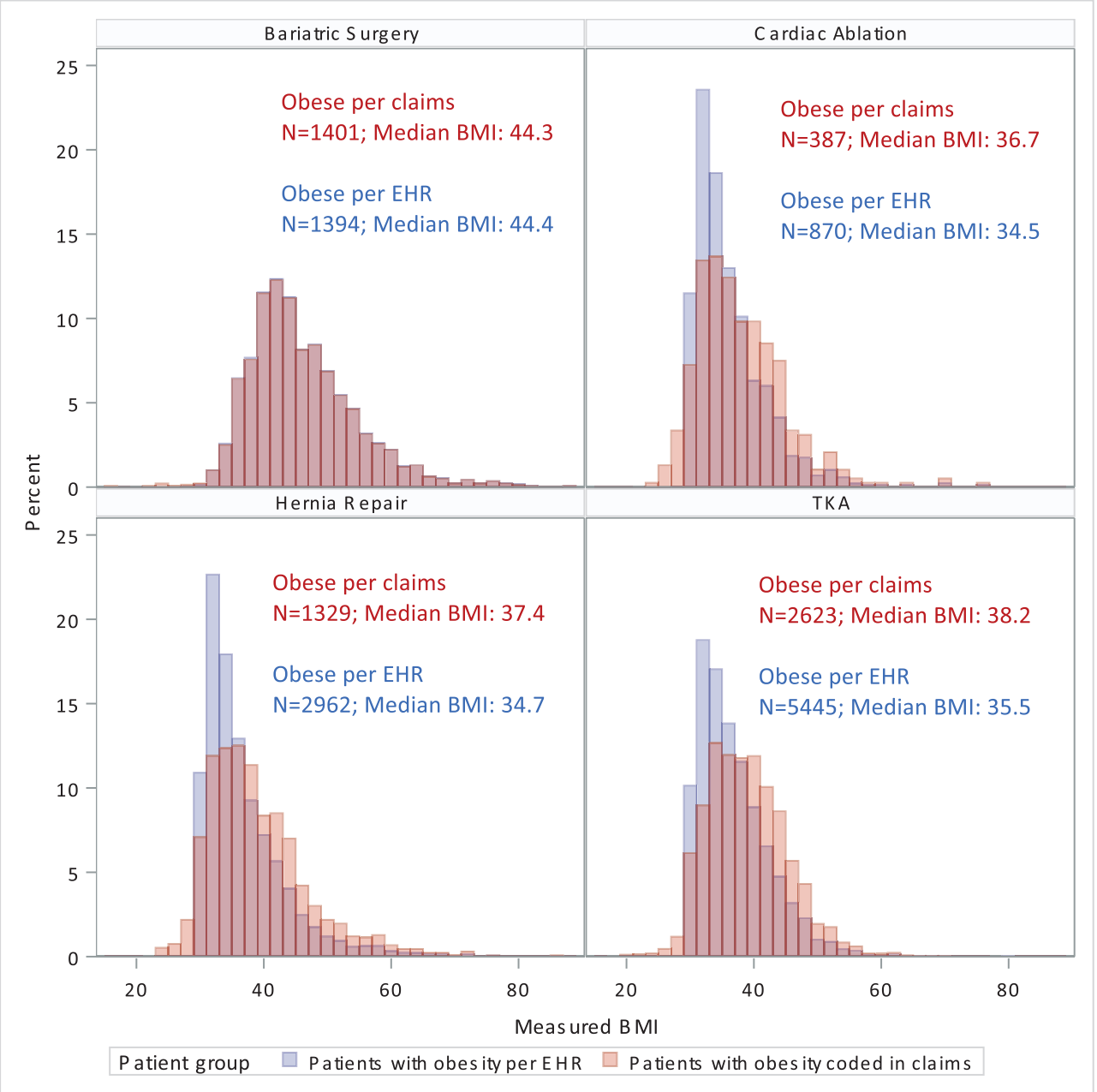

Figure 2. Continuous distribution of measured BMI in patients who were coded as obese in the claims data vs. all patients who were in fact obese. BMI = body mass index, EHR = electronic health record.

4. Discussion

In this large population-based study of US commercial and Medicare Advantage health plan members who underwent selected surgeries in 2014Q1–2017Q1, we found a high prevalence of overweight and obesity (>80% overweight or obese in each type of surgery) based on measured BMI, and high PPV but only moderate sensitivity for obesity diagnosis codes recorded in the claims data. The PPV and sensitivity for overweight diagnosis codes was lower compared to those for obesity. In general, patients with more severe obesity (BMI ≥35) were more likely to have a weight-related diagnosis recorded than patients who were overweight or less severely obese.

Across surgeries, it was notable that the prevalence of overweight/obesity was high relative to the US general population.[18] In each surgical cohort, >80% of cases were overweight/obese (BMI ≥25), >50% were obese (BMI ≥30), and >10% had class III obesity (BMI ≥40). The high prevalence of obesity has important clinical and economic implications. Previous studies have found that obesity—particularly severe obesity—is independently associated with higher complication rates, longer length of stay, greater post-discharge healthcare utilization, and worse patient outcomes following many surgical procedures.[2,3,12,13,19,20]

Our results indicate that the use of administrative diagnoses coded in claims data to classify patients as overweight or obese may meaningfully underestimate the prevalence of those conditions. For example, among TKA patients the prevalence of obesity was 63% according to the EHR but only 30% according to the claims data, and the sensitivity of obesity diagnosis codes was 46%. In claims-based patient outcomes research where obesity is an important prognostic factor, risk-adjustment based on coded obesity diagnoses may be vulnerable to residual confounding. On the contrary, for false-positive patients who were diagnosed as obese in the administrative data but ultimately did not have BMI ≥30 in the EHR, BMIs recorded in the EHR tended to be near the threshold of BMI ≥30 for obesity. Thus, some of the identified false-positive misclassifications were relatively minor in absolute BMI unit terms, which may be tolerable for certain research applications.

Despite these limitations, the high PPV of obesity diagnosis codes (>90%) indicates that it is feasible to define a cohort of patients with obesity using diagnoses coded in claims data. Similarly, morbid obesity diagnoses—with a PPV >88% when a BMI ≥35 is used as the validation criterion—may be used to define a surgical cohort with severe obesity. With respect to generalizability, in all surgery types except for bariatric surgery, patients coded as obese had BMIs that on average were 2 to 5 kg/m^2 higher than the larger population of patients who were in fact obese according to the EHR. Stated differently, outside the setting of bariatric surgery, patients with severe obesity were more likely to be coded as obese than patients with less severe obesity.

Strengths of our study include its use of recent US claims data from a large and geographically diverse sample of surgical patients. By including several surgical groups, we could assess the extent to which coding of overweight/obesity is consistent across procedures and age groups. In addition, the span of included years allowed us to evaluate the validity of both *ICD-9-CM* and *ICD-10-CM* administrative diagnosis codes for overweight/obesity. In general, the sensitivity of administrative diagnosis codes for overweight/obesity was somewhat higher in the *ICD-10-CM* era, and PPV estimates were similar in *ICD-9-CM* and *ICD-10-CM*.

Limitations of our study include the possibility of measurement error. A patient's BMI may have changed during the course of the 6-month assessment period. To explore this possibility, we recalculated PPVs after restricting to patients where the BMI measurement and overweight/obesity diagnosis date occurred within 60, 45, 30, or 15 days. As described in the results section, PPVs for obesity improved modestly with these restrictions, but such improvements were attained at the expense of sample size. Although we treated measured BMI as recorded in the EHR data as the criterion standard, such data are also subject to potential measurement error arising from inaccurate recording within the EHR system. We also did not have access to information on race/ethnicity; future research is needed to determine whether race/ethnicity may play a role in the under- or misdiagnosis of obesity in relation to actual BMI.

The present study adds to an existing literature on the validity of administrative diagnosis codes for obesity. Our study of obesity diagnosis codes found somewhat higher sensitivity (≥30%) and PPV (≥93%) than previous studies, which reported sensitivity estimates of 7% to 19% and PPV estimates of 66% to 91%.[5–9,21,22] The higher sensitivity observed in our study population may reflect the fact that obesity is likely to affect patient care in the types of surgery studied, and thus according to coding guidelines would qualify as a relevant comorbidity that should be coded.[23] In addition, our data came from the more recent past, and coding of overweight/obesity may have increased over time. Finally, the lowest published sensitivity estimates for obesity administrative diagnosis come from studies of hospital discharge abstract databases, wherein administrative diagnosis come from a single encounter; in contrast, our study also included diagnoses recorded during the 180 days before index.

5. Conclusions

Our data indicate that obesity is highly prevalent in many surgical populations, that obesity diagnosis codes have high PPVs, but also that obesity is generally undercoded in claims data.[4] Quantifying the validity of diagnosis codes for obesity and other important prognostic factors is important for the design and interpretation of studies of surgical outcomes based on administrative data. Further research is needed to determine the extent to which undercoding of BMI and obesity can be addressed through the use of proxies that may be better documented in claims data. For example, in a comparative effectiveness study, an imbalance on BMI between 2 comparison groups would likely be reduced by matching on a propensity score that incorporated information on obesity-related health conditions such as type 2 diabetes, hypertension, hyperlipidemia, and osteoarthritis, and medications used to treat these conditions.

Acknowledgments

The authors thank Chantal Holy, PhD, and Paul Coplan, PhD, of Johnson & Johnson for providing feedback on an earlier version of this manuscript.

Author contributions

Conceptualization: Eric M. Ammann, Iftekhar Kalsekar, Andrew Yoo, Stephen Johnston.
Data curation: Stephen Johnston.

Formal analysis: Eric M. Ammann, Stephen Johnston.
Investigation: Iftekhar Kalsekar, Robin Scamuffa, Andrew C Stokes, John M. Morton, Chia Wen Hsiao, Stephen Johnston.
Methodology: Iftekhar Kalsekar, Andrew Yoo, Stephen Johnston.
Supervision: Andrew Yoo, Stephen Johnston.
Validation: Iftekhar Kalsekar, Andrew Yoo, Robin Scamuffa, Andrew C Stokes, John M. Morton, Chia Wen Hsiao, Stephen Johnston.
Visualization: Eric M. Ammann, Stephen Johnston.
Writing – original draft: Eric M. Ammann, Stephen Johnston.
Writing – review & editing: Iftekhar Kalsekar, Andrew Yoo, Robin Scamuffa, Andrew C Stokes, John M. Morton, Chia Wen Hsiao, Stephen Johnston.

References

[1] Sood A, Abdollah F, Sammon JD, et al. The effect of body mass index on perioperative outcomes after major surgery: results from the National Surgical Quality Improvement Program (ACS-NSQIP) 2005-2011. World J Surg 2015;39:2376–85.
[2] McElroy MJ, Pivec R, Issa K, et al. The effects of obesity and morbid obesity on outcomes in TKA. J Knee Surg 2013;26:83–8.
[3] Vaishya R, Vijay V, Wamae D, et al. Is total knee replacement justified in the morbidly obese? a systematic review. Cureus 2016;8:e804.
[4] U.S. Food & Drug Administration. Use of real-world evidence to support regulatory decision-making for medical devices: Guidance for industry and Food and Drug Administration Staff. 2017; available at: https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm513027.pdf. Accessed March 14, 2018.
[5] Lloyd JT, Blackwell SA, Wei II, et al. Validity of a claims-based diagnosis of obesity among medicare beneficiaries. Eval Health Prof 2015;38:508–17.
[6] Martin BJ, Chen G, Graham M, et al. Coding of obesity in administrative hospital discharge abstract data: accuracy and impact for future research studies. BMC Health Serv Res 2014;14:70.
[7] Nickel KB, Wallace AE, Warren DK, et al. Modification of claims-based measures improves identification of comorbidities in non-elderly women undergoing mastectomy for breast cancer: a retrospective cohort study. BMC Health Serv Res 2016;16(a):388.
[8] Peng M, Southern DA, Williamson T, et al. Under-coding of secondary conditions in coded hospital health data: Impact of co-existing conditions, death status and number of codes in a record. Health Informatics J 2017;23:260–7.
[9] Ammann EM, Kalsekar I, Yoo A, et al. Validation of body mass index (BMI)-related ICD-9-CM and ICD-10-CM administrative diagnosis codes recorded in US claims data. Pharmacoepidemiol Drug Saf 2018;27:1092–100.
[10] Mechanick JI, Youdim A, Jones DB, et al. Clinical practice guidelines for the perioperative nutritional, metabolic, and nonsurgical support of the bariatric surgery patient—2013 update: cosponsored by American Association of Clinical Endocrinologists, The Obesity Society, and American Society for Metabolic & Bariatric Surgery. Obesity (Silver Spring) 2013;21(suppl 1):S1–s27.
[11] Schauer DP, Arterburn DE, Livingston EH, et al. Impact of bariatric surgery on life expectancy in severely obese patients with diabetes: a decision analysis. Ann Surg 2015;261:914–9.
[12] Chilukuri K, Dalal D, Gadrey S, et al. A prospective study evaluating the role of obesity and obstructive sleep apnea for outcomes after catheter ablation of atrial fibrillation. J Cardiovasc Electrophysiol 2010;21:521–5.
[13] Owei L, Swendiman RA, Kelz RR, et al. Impact of body mass index on open ventral hernia repair: a retrospective review. Surgery 2017;162:1320–9.
[14] Optum. Retrospective database analysis. 2019; available at: https://www.optum.com/content/dam/optum/resources/productSheets/Retrospective-Database-Analysis.pdf. Accessed April 10, 2019.
[15] Optum. Optum Integrated Data. 2019; available at: https://www.optum.com/content/dam/optum3/optum/en/resources/fact-sheets/Integrated-Data-product-sheet.pdf. Accessed April 10, 2019.
[16] Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults–The Evidence Report. National Institutes of Health. Obes Res 1998;6(suppl 2):51S–209S.
[17] Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43:1130–9.
[18] Flegal KM, Kruszon-Moran D, Carroll MD, et al. Trends in obesity among adults in the United States, 2005 to 2014. JAMA 2016;315:2284–91.
[19] Perlow JH, Morgan MA. Massive maternal obesity and perioperative cesarean morbidity. Am J Obstet Gynecol 1994;170:560–5.
[20] Chen CL, Shore AD, Johns R, et al. The impact of obesity on breast surgery complications. Plast Reconstr Surg 2011;128:395e–402e.
[21] Goff SL, Pekow PS, Markenson G, et al. Validity of using ICD-9-CM codes to identify selected categories of obstetric complications, procedures and co-morbidities. Paediatr Perinat Epidemiol 2012;26:421–9.
[22] Golinvaux NS, Bohl DD, Basques BA, et al. Limitations of administrative databases in spine research: a study in obesity. Spine J 2014;14:2923–8.
[23] National Center for Health Statistics. ICD-10-CM Official Guidelines for Coding and Reporting FY2018. 2018. Available at: https://www.cdc.gov/nchs/data/icd/10cmguidelines_fy2018_final.pdf. Accessed January 26, 2018.