2017 Knee Society Proceedings

Analysis of Outcomes After TKA: Do All Databases Produce Similar Findings?

Nicholas A. Bedard MD, Andrew J. Pugely MD, Michael McHugh MD, Nathan Lux, Jesse E. Otero MD, PhD, Kevin J. Bozic MD, MBA, Yubo Gao PhD, John J. Callaghan MD

Abstract

Background Use of large clinical and administrative databases for orthopaedic research has increased exponentially. Each database represents unique patient populations and varies in their methodology of data acquisition, which makes it possible that similar research questions posed to different databases might result in answers that differ in important ways.

Questions/purposes (1) What are the differences in reported demographics, comorbidities, and complications for patients undergoing primary TKA among four databases commonly used in orthopaedic research? (2) How does the difference in reported complication rates vary depending on whether only inpatient data or 30-day postoperative data are analyzed?

Methods Patients who underwent primary TKA during 2010 to 2012 were identified within the National Surgical Quality Improvement Programs (NSQIP), the Nationwide Inpatient Sample (NIS), the Medicare Standard Analytic Files (MED), and the Humana Administrative Claims database (HAC). NSQIP is a clinical registry that captures both inpatient and outpatient events up to 30 days after surgery using clinical reviewers and strict definitions for each variable. The other databases are administrative claims databases with their comorbidity and adverse event data defined by diagnosis and procedure codes used for reimbursement. NIS is limited to inpatient data only, whereas HAC and MED also have outpatient data. The number of patients undergoing primary TKA from each database was 48,248 in

N. A. Bedard, A. J. Pugely, M. McHugh, N. Lux, J. E. Otero, Y. Gao, J. J. Callaghan
Department of Orthopaedic Surgery and Rehabilitation, University of Iowa Hospitals and Clinics, Iowa City, IA, USA

K. J. Bozic, Department of Surgery & Perioperative Care, Dell Medical School at the University of Texas at Austin, Austin, TX, USA

J. J. Callaghan (✉), Department of Orthopaedic Surgery and Rehabilitation, University of Iowa Hospitals and Clinics, 01029 JPP, 200 Hawkins Drive, Iowa City, IA, 52245 USA, email: john-callaghan@uiowa.edu

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research®* editors and board members are on file with the publication and can be viewed on request.

HAC, 783,546 in MED, 393,050 in NIS, and 43,220 in NSQIP. NSQIP definitions for comorbidities and surgical complications were matched to corresponding International Classification of Diseases, 9th Revision/Current Procedural Terminology codes and these coding algorithms were used to query NIS, MED, and HAC. Age, sex, comorbidities, and inpatient versus 30-day postoperative complications were compared across the four databases. Given the large sample sizes, statistical significance was often detected for small, clinically unimportant differences; thus, the focus of comparisons was whether the difference reached an absolute difference of twofold to signify an important clinical difference.

Results Although there was a higher proportion of males in NIS and NSQIP and patients in NIS were younger, the difference was slight and well below our predefined threshold for a clinically important difference. There was variation in the prevalence of comorbidities and rates of postoperative complications among databases. The prevalence of chronic obstructive pulmonary disease (COPD) and coagulopathy in HAC and MED was more than twice that in NIS and NSQIP (relative risk [RR] for COPD: MED versus NIS 3.1, MED versus NSQIP 4.5, HAC versus NIS 3.6, HAC versus NSQIP 5.3; RR for coagulopathy: MED versus NIS 3.9, MED versus NSQIP 3.1, HAC versus NIS 3.3, HAC versus NSQIP 2.7; p < 0.001 for all comparisons). NSQIP had more than twice the obesity as NIS (RR 0.35). Rates of stroke within 30 days of TKA had more than a twofold difference among all databases (p < 0.001). HAC had more than twice the rates of 30-day complications at all endpoints compared with NSQIP and

more than twice the 30-day infections as MED. A comparison of inpatient and 30-day complications rates demonstrated more than twice the amount of wound infections and deep vein thromboses is captured when data are analyzed out to 30 days after TKA (p < 0.001 for all comparisons).

Conclusions When evaluating research utilizing large databases, one must pay particular attention to the type of database used (administrative claims, clinical registry, or other kinds of databases), time period included, definitions utilized for specific variables, and the population captured to ensure it is best suited for the specific research question. Furthermore, with the advent of bundled payments, policymakers must meticulously consider the data sources used to ensure the data analytics match historical sources.

Level of Evidence Level III, therapeutic study.

Introduction

The use of large clinical and administrative databases for orthopaedic research has increased exponentially over the last decade (Fig. 1) [11, 12]. Access to extremely large volumes of patient data has allowed researchers to answer questions previously difficult to evaluate using smaller single-institution cohort studies. Additionally, these large databases have allowed for more robust analysis of trends in procedures, complications, and outcomes after surgery. Furthermore, administrative claims databases are now being utilized for the public reporting of surgical outcomes with subsequent penalties for underperforming institutions [6, 14].

Although powerful in their ability to analyze large cohorts, each database summarizes the experiences of a unique patient population and varies in their methodology for data acquisition [11, 12]. These intrinsic differences may result in inconsistencies of reported comorbidities and surgical complications. Recently, studies have evaluated differences in reported comorbidities and surgical complications for multiple large databases across many surgical specialties, including orthopaedic surgery [2-4, 7-10]. For example, Bohl et al. demonstrated differences in inpatient adverse events after hip fracture surgery between the National Inpatient Sample (NIS) and National Surgical Quality Improvement Program (NSQIP) databases with frequencies of acute kidney injury and urinary tract infection in NIS being more than twice those in NSQIP [3]. Understanding the differences among these databases is important for appropriately evaluating research utilizing them.

Therefore, we asked: (1) What are the differences in reported demographics, comorbidities, and complications for patients undergoing primary TKA among four databases commonly used in orthopaedic research? (2) How does the difference in reported complication rates vary depending on whether only inpatient data or 30-day postoperative data are analyzed?

Patients and Methods

A retrospective study of patients who had undergone primary TKA was performed with four databases commonly used in orthopaedic research: NSQIP, NIS, Medicare Standard
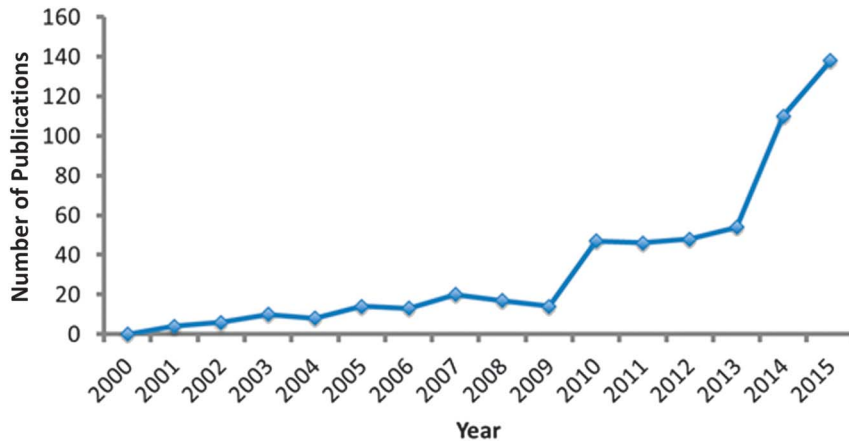
Fig. 1 Demonstrated is the exponential use of large databases for orthopaedic research in recent years. The following search query was used to identify the number of database TKA studies in PubMed each year: (("total knee replacement" or "total knee arthroplasty" or "knee replacement" or "TKA" or "total knee" or "knee replacement") AND ("NIS" or "nationwide inpatient sample" or "NSQIP" or "national surgical quality improvement program" or "PearlDiver" or "national database" or "insurance database" or "database")).

Analytic Files (MED), and the Humana Administrative Claims database (HAC). Only procedures that occurred between 2010 and 2012 were evaluated because these were the years of data available across all data sets. Patients undergoing primary TKA were identified using Current Procedural Terminology (CPT) code 27447 and International Classification of Diseases, 9th Revision (ICD-9) code 81.54. All data within these databases are deidentified and Health Insurance Portability and Accountability Act-compliant and were thus exempt from institutional review board approval.

The NSQIP database is maintained by the American College of Surgeons and uses trained clinical reviewers to perform data collection through chart review and patient/surgeon contact utilizing strict definitions for each comorbidity and complication variable catalogued [1]. This database captures both inpatient and outpatient events up to

30 days after surgery and there is roughly an equal mix of public and private hospitals in the cohort. Routine auditing of the NSQIP database is performed to ensure standardized data collection. Audits have demonstrated high data reliability with disagreement rates of < 2% [13]. Additionally, the database reports if any data are missing for a given variable to allow researchers to appropriately address the missing data. The three additional databases (NIS, HAC, and MED) were queried utilizing the PearlDiver Research Program (www.pearldiverinc.com; PearlDiver Inc, Fort Wayne, IN, USA). All three of these databases are administrative claims data sets with their comorbidity and adverse event data defined by reimbursement data in the form of ICD-9 and/or CPT codes. The HAC and MED databases utilize both ICD-9 and CPT codes and capture both inpatient and outpatient events. For the MED and HAC databases, there are no time

intervals after which outpatient events are no longer captured other than limitations by what years are included in the respective databases. In this study, outpatient events were only analyzed out to 30 days after TKA to allow for comparison to NSQIP. The NIS consists of a 20% sample of all inpatient discharges and includes only inpatient data. Similar to HAC and MED, NIS comorbidity and adverse event data are defined by reimbursement data for the inpatient admission of interest; however, only ICD-9 codes are supported by NIS. Although the absence of CPT codes in NIS does somewhat limit the accuracy of identifying complications that required a return to the operating room (ie, infection), this is unlikely to greatly limit the findings of this study because return to the operating room during the inpatient stay was overall a rare event, ICD-9 procedure codes do capture operations, and only the infection variables and cardiac arrest variable had CPT codes included (for HAC and MED) in their definitions along with ICD-9 diagnosis codes that are individually quite specific for these particular complications. The validity of administratively coded comorbidity and complication data in the total joint arthroplasty population has been studied by Bozic et al. [5]. They reported varied concordance between administrative claims and the clinical record, citing a high degree of specificity (> 92%) for all comorbidities and complications but a lower degree of sensitivity (29%-100%) suggesting comorbidities and complications in the administrative claims record are accurate but often incomplete [5]. The number of patients undergoing primary TKA analyzed in this study

2017 Knee Society Proceedings

Table 1. Definitions of comorbidity variables

Comorbidity	NSQIP	NIS, HAC, MED
Morbid obesity	BMI > 40 kg/m^2 based on "height" (height) and "weight" (weight)	ICD-9 278.01 (obesity [BMI from > 40 kg/m^2]) V85.3 (BMI > 40 kg/m^2, adult)
Obesity	BMI 30-40 kg/m^2 based on "height" (height) and "weight" (weight)	ICD-9 278.00 (obesity [BMI from 30-40 kg/m^2]) V85.3 (BMI between 30 and 39 kg/m^2, adult)
Coagulopathy	"bleeddis" (bleeding disorders)	ICD-9 286.x* (coagulation defect), 287.3x (primary thrombocytopenia), 287.4x (secondary thrombocytopenia), 287.5 (thrombocytopenia, unspecified), 269.0 (deficiency of vitamin K)
Diabetes	"diabetes" (diabetes mellitus with oral agents or insulin)	ICD-9 249.xx (secondary diabetes mellitus), 250.xx (diabetes mellitus)
Hypertension	"hypermed" (hypertension requiring medication)	ICD-9 401.xx (essential hypertension), 405.xx (secondary hypertension)
Chronic obstructive pulmonary disease (COPD)	"hxcopd" (history of severe COPD)	ICD-9 491.xx (chronic bronchitis), 492.xx (emphysema),496 (chronic airway obstruction, unspecified)
Smoking	"smoke" (current smoker within 1 year)	ICD-9 30.51 (tobacco use disorder, dependence)

*The letter "x" represents all fourth or fifth digits that could designate an ICD-9 code.
NSQIP = National Surgical Quality Improvement Program; NIS = National Inpatient Sample; HAC = Humana Administrative Claims database; MED = Medical Standard Analytic Files; BMI = body mass index; ICD-9 = International Classification of Diseases, 9th Revision.

was 48,248 in HAC, 783,546 in MED, 393,050 in NIS, and 43,220 in NSQIP.

Definitions from the NSQIP user manual for seven comorbidities (morbid obesity, obesity, coagulopathy, diabetes, hypertension, chronic obstructive pulmonary disease [COPD], and smoking; Table 1) and nine postoperative complications (deep vein thrombosis [DVT], pneumonia, stroke, myocardial infarction, cardiac arrest, pulmonary embolism [PE], wound dehiscence, deep surgical site infection [SSI], and any SSI; Table 2) were matched to corresponding ICD-9 and CPT codes and compared for all patients across all databases. These particular comorbidities and complications were chosen because they were thought to be most relevant to patients undergoing TKA and they were felt to be most accurately matched to corresponding ICD-9 codes based on their NSQIP definitions. NSQIP variables not included were either unrelated to patients undergoing TKA or the definitions were extremely specific and did not have an adequate corresponding ICD-9 code. Additionally, patient demographics (age at the time of surgery and sex) were also compared across data sets.

Postoperative complications were evaluated and compared over two different time periods: those that occurred during the inpatient stay (NIS, NSQIP, HAC, MED) and those that occurred within 30 days of the primary TKA (NSQIP, HAC, MED).

Lastly, the differences in rates of complications occurring during the inpatient stay versus those occurring within 30 days after surgery were compared for databases with both of these times points available (NSQIP, HAC, MED).

Demographic characteristics were compared between databases with use of the Pearson chi-square test. Given the manner in which age data are provided by the PearlDiver research program for HAC, NIS, and MED, we compared age of patients between databases based on which age group the median age at the time of TKA fell within. A p value of < 0.05 was considered significant. Prevalence of comorbidities and complication rates was compared among databases with use of relative

Wolters Kluwer

Table 2. Definitions of complication variables

Complication	NSQIP	HAC and MED	NIS
Deep vein thrombosis	"nothdvt" (deep vein thrombosis, with or without inflammation, postoperative)	Acute venous embolism and thrombosis of: ICD-9 453.2 (inferior vena cava), 453.3 (renal vein), 453.40 (deep vessel unspecified), 453.41 (proximal lower extremity), 453.42 (distal lower extremity), 453.82 (deep veins of upper extremity), 453.83 (upper extremity unspecified), 453.84 (axillary vein), 453.85 (subclavian vein), 453.86 (internal jugular vein)	Acute venous embolism and thrombosis of: ICD-9 453.2 (inferior vena cava), 453.3 (renal vein), 453.40 (deep vessel unspecified), 453.41 (proximal lower extremity), 453.42 (distal lower extremity), 453.82 (deep veins of upper extremity), 453.83 (upper extremity unspecified), 453.84 (axillary vein), 453.85 (subclavian vein), 453.86 (internal jugular vein)
Pneumonia	"noupneumo" (pneumonia, postoperative)	ICD-9 480.x (viral pneumonia), 481 (pneumococcal pneumonia), 482.x (other bacterial pneumonia), 483.x (pneumonia resulting from other specified organism), 484.x (pneumonia in infectious diseases classified elsewhere), 485 (bronchopneumonia), 486 (pneumonia, organism unspecified)	ICD-9 480.x (viral pneumonia), 481 (pneumococcal pneumonia), 482.x (other bacterial pneumonia), 483.x (pneumonia resulting from other specified organism), 484.x (pneumonia in infectious diseases classified elsewhere), 485 (bronchopneumonia), 486 (pneumonia, organism unspecified)
Stroke	"ncnscva" (stroke or cerebrovascular accident, postoperative)	ICD-9 997.02 (iatrogenic cerebrovascular infarction or hemorrhage), 430 (SAH), 431 (intracerebral hemorrhage), 432.x (other and unspecified intracranial hemorrhage), 433.x1 (occlusion and stenosis of precerebral arteries with infarction), 434.x1 (occlusion of cerebral arteries, with cerebral infarction)	ICD-9 997.02 (iatrogenic cerebrovascular infarction or hemorrhage), 430 (SAH), 431 (intracerebral hemorrhage), 432.x (other and unspecified intracranial hemorrhage), 433.x1 (occlusion and stenosis of precerebral arteries with infarction), 434.x1 (occlusion of cerebral arteries, with cerebral infarction)
Myocardial infarction	"ncdmi" (myocardial infarction, postoperative) "typeintoc" (myocardial infarction, intraoperative)	ICD-9 410.x0, 410.x1 (acute myocardial infarction, initial and unspecified episode of care)	ICD-9 410.x0, 410.x1 (acute myocardial infarction, initial and unspecified episode of care)
Cardiac arrest	"ncdarrest" (cardiac arrest requiring cardiopulmonary resuscitation, postoperative) "typeintoc" (cardiac arrest requiring cardiopulmonary resuscitation, intraoperative)	ICD-9 427.5 (cardiac arrest), 427.41 (ventricular fibrillation), CPT-92950 (cardiopulmonary resuscitation; eg, in cardiac arrest)	ICD-9 427.5 (cardiac arrest), 427.41 (ventricular fibrillation)
Pulmonary embolism	"npulembol" (pulmonary embolism, postoperative)	ICD-9 415.11 (atrogenic pulmonary embolism and infarction), 415.12 (septic pulmonary embolism), 415.13 (saddle embolus of pulmonary artery), 415.19 (other pulmonary embolism)	ICD-9 415.11 (atrogenic pulmonary embolism and infarction), 415.12 (septic pulmonary embolism), 415.13 (saddle embolus of pulmonary artery), 415.19 (other pulmonary embolism)

Table 2. continued

Complication	NSQIP	HAC and MED	NIS
Wound dehiscence	"ndehis" (wound disruption, postoperative)	ICD-9 998.30 (disruption of wound, unspecified), 998.31 (disruption of an internal operation [surgical] wound), 998.32 (disruption of an external operation [surgical] wound)	ICD-9 998.30 (disruption of wound, unspecified), 998.31 (disruption of an internal operation [surgical] wound), 998.32 (disruption of an external operation [surgical] wound)
Deep surgical site infection	"nwndinfd" (deep surgical site infection, postoperative); norgspcssi" (organ or space surgical site infection, postoperative)	CPT 27488 (removal of prosthesis, including total knee prosthesis, methylmethacrylate with or without insertion of spacer, knee), 10180 (incision and drainage, complex, postoperative wound infection), 27310 (arthrotomy, knee, with exploration, drainage, or removal of foreign body [eg, infection]), 11981 (insertion, nonbiodegradable drug delivery implant), 29871 (arthroscopy, knee, surgical; for infection, lavage and drainage), ICD-9 80.06 (arthrotomy for removal of prosthesis without replacement, knee)	NA
Any surgical site infection	"nsupinfec" (superficial surgical site infection, postoperative)"nwndinfd" (deep surgical site infection, postoperative); "norgspcssi" (organ or space surgical site infection, postoperative)	ICD-9 996.66 (infection and inflammatory reaction resulting from internal joint prosthesis), 998.83 (nonhealing surgical wound), 998.51 (infected postoperative seroma), 998.59 (other postoperative infection), 998.6 (persistent postoperative fistula) CPT 27488 (removal of prosthesis, including total knee prosthesis, methylmethacrylate with or without insertion of spacer, knee), 10180 (incision and drainage, complex, postoperative wound infection), 27310 (arthrotomy, knee, with exploration, drainage, or removal of foreign body [eg, infection]), 11981 (insertion, nonbiodegradable drug delivery implant), 29871 (arthroscopy, knee, surgical; for infection, lavage and drainage), ICD-9 80.06 (arthrotomy for removal of prosthesis without replacement, knee)	ICD-9 996.66 (infection and inflammatory reaction resulting from internal joint prosthesis), 998.83 (nonhealing surgical wound), 998.51 (infected postoperative seroma), 998.59 (other postoperative infection), 998.6 (persistent postoperative fistula), 80.06 (arthrotomy for removal of prosthesis without replacement, knee)

NSQIP = National Surgical Quality Improvement Program; HAC = Humana Administrative Claims database; MED = Medical Standard Analytic Files; NIS = National Inpatient Sample; ICD-9 = International Classification of Diseases, 9th Revision; SAH = subarachnoid hemorrhage; CPT = Current Procedural Terminology code; NA = the endpoint of deep surgical site infection was not compared for inpatients because we were unable to appropriately match procedure codes representing operative infection for NIS given this database does not include CPT codes.

58   Bedard et al.   Clinical Orthopaedics and Related Research®

2017 Knee Society Proceedings

risk (RR) and corresponding 95% confidence intervals. However, consistent with other database comparison studies, the large sample sizes and associated high power allowed for detection of statistical significance for small, clinically insignificant differences [3, 4]. Thus, the focus of these comparisons was on the magnitude of the differences, specifically utilizing an absolute difference threshold of greater than twofold to signify an important clinical difference.

Fig. 2 This graphic breakdown of age of the patients at the time of primary TKA demonstrates the relative similarity of age among the four databases.

Results

Demographic Comparisons Among Databases

Age distribution was clinically similar among databases with median age falling within the 70- to 74-year-old age group for NSQIP, MED, and HAC and 65 to 69 years for NIS (Fig. 2). The female-to-male ratio was 1.7 to one for NIS and NSQIP and 1.8 to one for HAC and MED. Despite observed statistical differences (p < 0.001 for age breakdown and sex ratios), these differences were small and well below our predetermined thresholds defining clinically important differences.

Differences in Comorbidities and Complications

Overall there was some variation in the prevalence of comorbidities (Fig. 3) and large variation in rates of inpatient complications (Fig. 4) and postoperative complications occurring within 30 days of TKA among databases compared (Fig. 5). However, given that many small and clinically unimportant differences were statistically significant in the
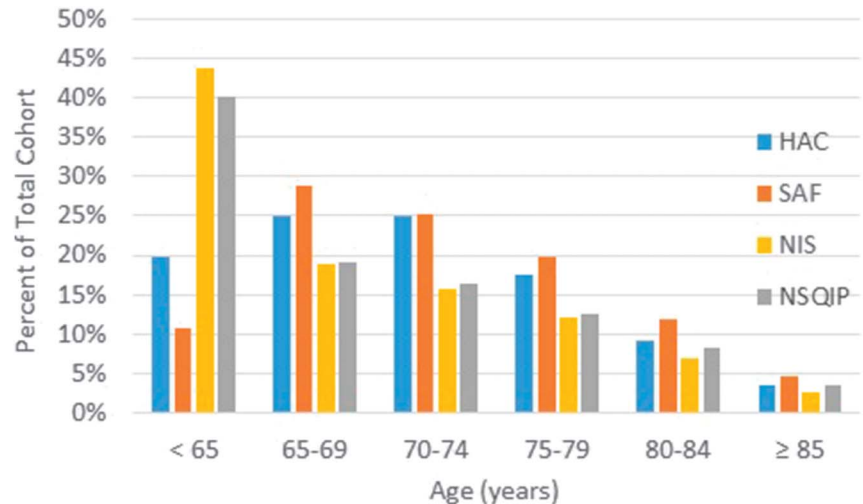
RR analysis, the focus of these comparisons was on the absolute difference in comorbidities and complication rates using a threshold of a twofold difference to define clinical importance. Comparison of comorbidities demonstrated a greater than twofold increase in RR of

COPD and coagulopathy in both HAC and MED compared with NIS and NSQIP (RR for COPD: MED versus NIS 3.1 [3.0-3.1], MED versus NSQIP 4.5 [4.3-4.7], HAC versus NIS 3.6 [3.6-3.7], HAC versus NSQIP 5.3 [5.0-5.6]; RR for coagulopathy: MED versus NIS
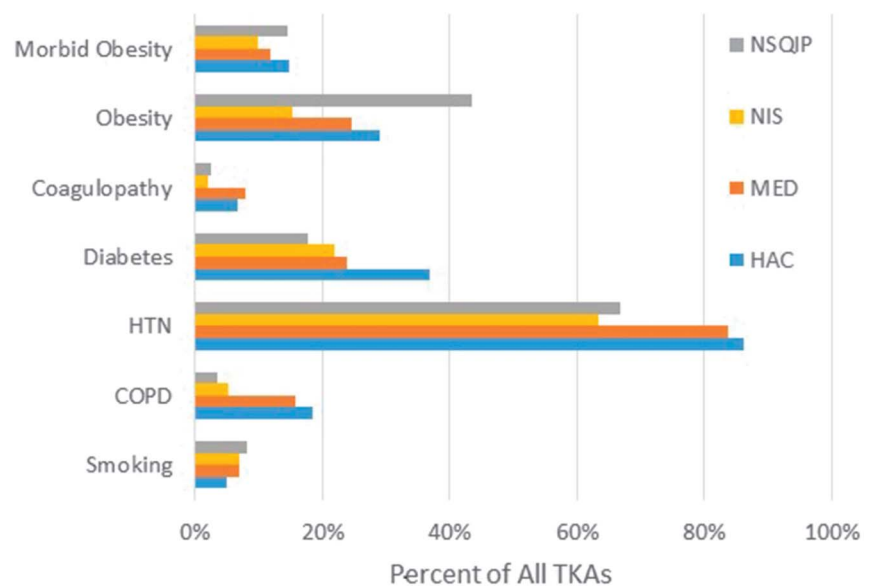

Fig. 3 Presented here is the prevalence of comorbidities in patients undergoing primary TKA for each database analyzed.
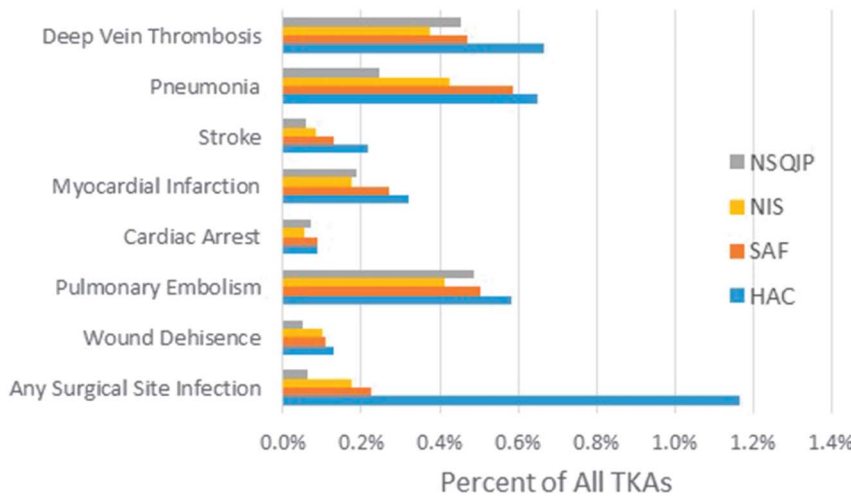
Fig. 4 Presented here is the prevalence of inpatient complications after TKA for each database analyzed.

3.9 [3.8-4.0], MED versus NSQIP 3.1 [2.9-3.2], HAC versus NIS 3.3 [3.2-3.4], HAC versus NSQIP 2.7 [2.5-2.8]; p < 0.001 for all). Additionally, NSQIP had twice the amount of obese patients as NIS (RR 0.4 [0.3-0.4], p < 0.001). The prevalence of all other comorbidities was not different among all databases (less than a twofold difference; Table 3).

The occurrence of inpatient complications was compared among all four databases included in this study (MED, NIS, HAC, NSQIP) and revealed HAC, MED, and NIS to have at least a twofold increase in RR of any SSI and wound dehiscence compared with NSQIP (RR any SSI: HAC: 17.9 [12.3-26.3], MED: 3.45[2.38-5.01],

NIS: 2.73 [1.88-3.99]; RR wound dehiscence: HAC: 2.64 [1.61-4.34], MED: 3.45 [2.38-5.01], NIS: 2.73 [1.88-3.99]; p < 0.001 for all). Additionally, there was a greater than fivefold increase in RR of any SSI for HAC compared with MED (RR 5.21 [4.74-5.73], p < 0.001) or NIS (RR 6.57 [5.88-7.34], p < 0.001). For the inpatient complications stroke and pneumonia, HAC and MED had more than a twofold RR of these complications relative to NSQIP (RR for stroke: HAC: 3.73 [2.41-5.77], MED: 2.23 [1.50-3.32]; RR for pneumonia: HAC: 2.62 (2.11-3.27), MED: 2.38 [1.96-2.88]; p < 0.001 for all). Lastly, HAC had a RR of 2.62 (2.10-3.27) for prevalence of stroke compared with NIS (p < 0.001). Prevalence of all other inpatient complications was not different among databases (Table 4).

Prevalence of complications occurring within 30 days after TKA was compared among HAC, MED, and NSQIP (NIS only included inpatient data) and varied greatly across databases with HAC having more than a twofold greater prevalence of every complication than NSQIP (p < 0.001 for all). Additionally, HAC had over twice the RR of stroke (2.62 [2.34-2.93]), deep SSI (10.72 [9.37-12.29]), and any SSI (2.04 [1.94-2.15]) than MED (p < 0.001 for all). Although MED had greater than a twofold prevalence of pneumonia (3.25 [2.77-3.81]), stroke (3.48 [2.48-4.89]), and wound dehiscence (2.66 [2.15-3.29]) than NSQIP, MED had less than half the prevalence of deep SSI relative to NSQIP (0.22 [0.18-0.27], p < 0.001 for all). Prevalence of all other complications occurring within 30 days of surgery was not different among databases (Table 5).

When looking at the number of complications captured by the databases relative to the time period
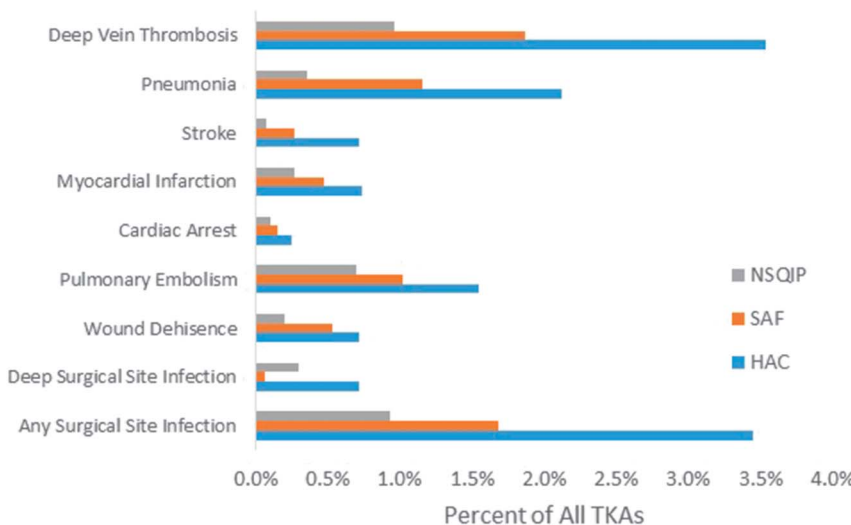

Fig. 5 Presented here is the prevalence of complications within 30 days after TKA for each database analyzed.

2017 Knee Society Proceedings

analyzed (inpatient versus 30 days after surgery), it was found that over half of any SSI, deep SSI, wound dehiscence, and DVT was not captured by NSQIP if followup was limited to only inpatient events. Similarly, over half of all complication endpoints except deep SSI would not have been captured by HAC if limited to the inpatient time period. Additionally, MED would have missed over half of all SSI, deep SSI, wound dehiscence, PE, stroke, and DVT if only inpatient data were included.

Discussion

In an era of increasing use of large clinical registries and administrative claims databases for orthopaedic research and assessment of hospital quality, our study demonstrated that among four commonly utilized databases, there is considerable variation in prevalence of complication rates after primary TKA despite relatively similar demographic and comorbidity profiles across these data sets. Additionally, a large percentage of complications is not captured if only inpatient events are analyzed and included within the database. These findings highlight the importance of understanding the methodology utilized to create each respective database.

Limitations to this study include the inability to link specific patients across databases used in this analysis. If we were able to specifically identify the same patient within each database, we would have been able to provide a true measure of validity and determine the most accurate database for analyzing complication rates after TKA. It is likely the same patient may be represented in multiple databases because HAC does contain patients with Medicare advantage plans and both NSQIP and NIS

Table 3. Comparisons of comorbidity prevalence among patients undergoing TKA

Comorbidity	HAC versus MED	p value	HAC versus NIS	p value	HAC versus NSQIP	p value	MED versus NIS	p value	MED versus NSQIP	p value	NIS versus NSQIP	p value
Smoking	0.72 (0.69-0.74)	< 0.001	0.71 (0.68-0.74)	< 0.001	0.60 (0.57-0.63)	< 0.001	0.99 (0.98-1.02)	0.21	0.84 (0.81-0.87)	< 0.001	0.84 (0.81-0.88)	< 0.001
COPD	1.18 (1.16-1.20)	< 0.001	3.64 (3.56-3.73)	< 0.001	5.30 (5.02-5.59)	< 0.001	3.09 (3.04-3.13)	< 0.001	4.50 (4.28-4.73)	< 0.001	1.46 (1.38-1.55)	< 0.001
HTN	1.03 (1.03-1.03)	< 0.001	1.36 (1.36-1.37)	< 0.001	1.30 (1.30-1.31)	< 0.001	1.32 (1.32-1.33)	< 0.001	1.25 (1.24-1.26)	< 0.001	0.95 (0.94-0.95)	< 0.001
Diabetes	1.55 (1.53-1.57)	< 0.001	1.68 (1.65-1.70)	< 0.001	1.53 (1.50-1.57)	< 0.001	1.08 (1.08-1.09)	< 0.001	1.35 (1.33-1.38)	< 0.001	1.25 (1.22-1.28)	< 0.001
Coagulopathy	0.85 (0.92-0.88)	< 0.001	3.29 (3.17-3.43)	< 0.001	2.65 (2.48-2.84)	< 0.001	3.89 (3.08-3.98)	< 0.001	3.13 (2.95-3.22)	< 0.001	0.81 (0.76-0.86)	< 0.001
Obesity	1.18 (1.17-1.20)	< 0.001	1.96 (1.92-1.98)	< 0.001	0.67 (0.66-0.68)	< 0.001	1.60 (1.59-1.62)	< 0.001	0.56 (0.56-0.57)	< 0.001	0.35 (0.35-0.36)	< 0.001
Morbid obesity	1.25 (1.23-1.29)	< 0.001	1.49 (1.45-1.52)	< 0.001	1.01 (0.98-1.04)	0.35	1.19 (1.17-1.20)	< 0.001	0.81 (0.79-0.83)	< 0.001	0.68 (0.66-0.70)	< 0.001

Data presented as relative risk and corresponding 95% confidence interval.
HAC = Humana Administrative Claims database; MED = Medicare Standard Analytic Files; NIS = National Inpatient Sample; NSQIP = National Surgical Quality Improvement Program; COPD = chronic obstructive pulmonary disease; HTN = hypertension.

Table 4. Comparisons of inpatient complications after TKA

Complication	HAC versus MED	p value	HAC versus NIS	p value	HAC versus NSQIP	p value	MED versus NIS	p value	MED versus NSQIP	p value	NIS versus NSQIP	p value
DVT	1.41 (1.26-1.58)	< 0.001	1.78 (1.57-2.00)	< 0.001	1.47 (1.23-1.76)	< 0.001	1.23 (1.18-1.33)	< 0.001	1.04 (0.90-1.20)	0.59	0.83 (0.71-0.96)	< 0.001
Pneumonia	1.11 (0.99-1.24)	0.09	1.54 (1.36-1.73)	< 0.001	2.62 (2.11-3.27)	< 0.001	1.39 (1.32-1.47)	< 0.001	2.38 (1.96-2.88)	< 0.001	1.71 (1.40-2.01)	< 0.001
Stroke	1.67 (1.36-2.04)	< 0.001	2.62 (2.10-3.27)	< 0.001	3.73 (2.41-5.77)	< 0.001	1.57 (1.39-1.78)	< 0.001	2.23 (1.50-3.32)	< 0.001	1.43 (0.95-2.13)	0.09
MI	1.18 (1.01-1.39)	0.04	1.81 (1.52-2.16)	< 0.001	1.69 (1.29-2.21)	< 0.001	1.53 (1.41-1.68)	< 0.001	1.44 (1.15-1.79)	< 0.001	0.93 (0.74-1.17)	0.56
Cardiac arrest	0.98 (0.72-1.24)	0.89	1.55 (1.11-2.15)	0.01	1.25 (0.79-2.00)	0.34	1.58 (1.36-1.84)	< 0.001	1.28 (0.89-1.85)	0.18	0.81 (0.55-1.19)	0.29
PE	1.16 (1.02-1.3)	0.02	1.42 (1.25-1.61)	< 0.001	1.20 (1.02-1.44)	0.04	1.22 (1.15-1.30)	< 0.001	1.04 (0.91-1.20)	0.57	0.85 (0.73-0.97)	< 0.001
Wound dehiscence	1.19 (0.92-1.54)	0.19	1.29 (0.99-1.69)	0.06	2.64 (1.61-4.34)	< 0.001	1.09 (0.96-1.22)	0.18	2.22 (1.44-3.43)	< 0.001	2.05 (1.32-3.18)	< 0.001
Any SSI	5.21 (4.74-5.73)	< 0.001	6.57 (5.88-7.34)	< 0.001	17.9 (12.3-26.3)	< 0.001	1.26 (1.16-1.38)	< 0.001	3.45 (2.38-5.01)	< 0.001	2.73 (1.88-3.99)	< 0.001

Data presented as relative risk and corresponding 95% confidence interval.
HAC = Humana Administrative Claims database; MED = Medicare Standard Analytic Files; NIS = National Inpatient Sample; NSQIP = National Surgical Quality Improvement Program; DVT = deep vein thrombosis; MI = myocardial infarction; PE = pulmonary embolism; SSI = surgical site infection.

represent a sample of the population not defined by an insurance provider. However, the degree to which there is overlap between these databases is not known. Although we are unable to determine the effect of overlap between databases from this study, we would assume a lesser degree of overlap would result in a lower degree of concordance among the findings in each database. Additionally, there was the potential for bias resulting from limitations of ICD-9/CPT coding and the strict nature of NSQIP definitions of comorbidities and complications possibly resulting in imperfect matching of some variables. For example, in the case of DVT, the NSQIP variable requires diagnosis and treatment, but for the administrative claims data, the treatment threshold is not a necessity for a diagnosis. Although matched as closely as possible, the differences in NSQIP definitions and ICD-9/CPT coding for a given variable likely played a large role in reported differences in NSQIP and administrative claims databases. Lastly, a twofold cutoff was used as a threshold for signifying excessive differences in the prevalence of comorbidities and complications. This threshold may have not signified clinically important differences of more common comorbidities/complications that had higher baseline prevalence because it would have still needed to double or halve to meet this cutoff. For example, 66% of NSQIP patients undergoing TKA had hypertension (HTN) and thus as few as 33% of patients could only have HTN in the other databases to meet this threshold. However, we believed that having a higher cutoff threshold was imperative to highlight the large clinically important differences within these data sets.

Despite differences in age and sex observed between databases, the magnitude of these differences was quite small, and all of them fell well below our a priori definitions of what would

Wolters Kluwer

Table 5. Comparisons of complications within 30 days after TKA

Complication	HAC versus MED	p value	HAC versus NSQIP	p value	MED versus NSQIP	p value
DVT	1.89 (1.80-1.99)	< 0.001	3.69 (3.32-4.10)	< 0.001	1.95 (1.77-2.15)	< 0.001
Pneumonia	1.83 (1.72-1.95)	< 0.001	5.94 (5.02-7.04)	< 0.001	3.25 (2.77-3.81)	< 0.001
Stroke	2.62 (2.34-2.93)	< 0.001	9.12 (6.41-12.96)	< 0.001	3.48 (2.48-4.89)	< 0.001
MI	1.55 (1.39-1.73)	< 0.001	2.70 (2.19-3.32)	< 0.001	1.74 (1.45-2.09)	< 0.001
Cardiac arrest	1.60 (1.33-1.93)	< 0.001	2.48 (1.75-3.51)	< 0.001	1.55 (1.14-2.10)	< 0.001
PE	1.53 (1.42-1.65)	< 0.001	2.23 (1.95-2.55)	< 0.001	1.46 (1.30-1.64)	< 0.001
Wound dehiscence	1.35 (1.21-1.51)	< 0.001	3.59 (2.84-4.55)	< 0.001	2.66 (2.15-3.29)	< 0.001
Deep SSI	10.72 (9.37-12.29)	< 0.001	2.36 (1.93-2.88)	< 0.001	0.22 (0.18-0.27)	< 0.001
Any SSI	2.04 (1.94-2.15)	< 0.001	3.69 (3.32-4.12)	< 0.001	1.81 (1.64-1.99)	< 0.001

Data presented as relative risk and corresponding 95% confidence interval.
HAC = Humana Administrative Claims database; MED = Medicare Standard Analytic Files; NSQIP = National Surgical Quality Improvement Program; DVT = deep vein thrombosis; MI = myocardial infarction; PE = pulmonary embolism; SSI = surgical site infection.

represent a clinically important difference. Similar findings have been reported in other database comparison studies with small, likely clinically unimportant differences found when comparing both patients with hip fracture and those undergoing lumbar spine fusion in NIS and NSQIP [3, 4]. These findings are important because they demonstrate that despite similar demographic profiles, patients undergoing TKA have wide variability in the prevalence of comorbidities and postoperative complications depending on the specific database utilized for analysis.

The differences in prevalence of complications for the various databases in this analysis are likely a result of the differences in methodology used to categorize complications for each respective data set. Additional reasons for differences in complication rates may be the manner in which the data are collected (trained clinical reviewers in NSQIP versus coders for administrative claims databases) and the primary reason for the data collection (quality assessment for NSQIP versus billing for administrative claims databases). These differences seem to apply to both inpatient and outpatient complications. We suspect that

NSQIP, which utilized chart abstraction by trained clinical reviewers with strict definitions for complications, is likely more accurate than administrative claims databases given limitations in ICD-9/CPT codes and the need for accurate documentation. The concern for the accuracy of administrative claims data relative to clinical registries such as NSQIP has recently emerged in the literature across multiple surgical specialties [2-4, 7-10]. Two of these studies performed by Bohl et al. compared NSQIP with NIS for patients with hip fracture and patients who underwent lumbar spine surgery utilizing similar methodology as the present study [3, 4]. Both studies demonstrated variation in rates of inpatient adverse events depending on the database after either hip fracture or lumbar spine surgery despite having similar patient populations [3, 4]. Additionally, similar to the current study, limitations of data sets that only included inpatient data were highlighted in these articles because over half of SSIs, DVTs, and urinary tract infections after hip fracture surgery and over half of SSIs and mortalities after lumbar spine surgery are not captured by NSQIP if the time period analyzed is

limited to only inpatient data [3, 4]. However, despite the data cited and the results of this study, the differences in prevalence of complications among databases cannot be entirely attributed to methods of data acquisition (specifically, administrative claims versus clinical registry) because even comparisons among multiple administrative claims databases in this study had drastically different results for the complications of stroke and SSI (Figs. 4, 5) despite querying the database with the exact same methodology.

Although we found a number of clinically important differences across the four databases we evaluated, we still believe each of these databases has important roles to play in orthopaedic research. Clinical registries such as NSQIP likely would be best utilized to evaluate the impact of comorbidities on specific complications after surgical procedures given their prospective collection of data, robust review of multiple sources to obtain clinical information (chart abstraction, medical providers, and patients), and their well-defined definitions for each included variable. However, NSQIP is limited to a short followup period (30 days) and is just a small sample of voluntarily participating hospitals, which

limits the ability of this type of database to estimate disease prevalence or to examine trends. Prevalence of disease or trends over time are probably best studied with administrative claims databases such as NIS, MED, or HAC because they are more representative of the general population and generally have more years of data included. Additionally, administrative claims databases appear to be better suited for financial analysis, analysis of questions outside the scope of strict clinical definitions often associated with clinical registries, and to evaluate complications that occur further in time after the index procedure (ie, revision of a TKA) because specific patients can be followed for the entire time period they remain within a given claims databases. Although in this study, inpatient-only analysis missed a number of clinically important complications that occurs in the 30 days after discharge, inpatient-only databases such as NIS can still be very useful because they often allow for better analysis of length of stay issues and are usually less geographically limited compared with other databases. The strengths and weaknesses of these databases must be considered when evaluating literature that utilizes a large administrative claims database or clinical registry.

In conclusion, this study demonstrates that among clinical and administrative databases commonly used in orthopaedic research, there is considerable variation in the prevalence of comorbidities and rates of complications after primary TKA depending on the database and postoperative time period used for analysis. The drivers of this variation are likely the result of major differences in the variable definitions, collections methods, and patient cohorts. When evaluating research utilizing large databases, one must pay particular attention to the type of database

used (administrative claims, clinical registry, or other kinds of databases), the duration of followup, and the population captured in that data set to ensure it is best suited for the specific research question. Furthermore, attention must also be paid to definitions utilized to define comorbidities and complications to ensure they accurately represent the variable of interest. Lastly, in the era of value-based health care, these differences must be considered when developing risk adjustment models for initiatives such as bundled payments. In the development of these programs, policymakers must carefully consider the data sources used to ensure the data analytics match historical sources. For example, if administrative claims data will be used actively within a bundle, then expected rates and risk adjustment models should be built off of those data.

References

1. American College of Surgeons National Surgical Quality Improvement Program. Data collection, analysis, and reporting. Available at: http://site.acsnsqip.org/program-specifics/data-collection-analysis-and-reporting/. Accessed May 15, 2016.
2. Awad MI, Shuman AG, Montero PH, Palmer FL, Shah JP, Patel SG. Accuracy of administrative and clinical registry data in reporting postoperative complications after surgery for oral cavity squamous cell carcinoma. *Head Neck.* 2015;37:851–861.
3. Bohl DD, Basques BA, Golinvaux NS, Baumgaertner MR, Grauer JN. Nationwide Inpatient Sample and National Surgical Quality Improvement Program give different results in hip fracture studies. *Clin Orthop Relat Res.* 2014;472:1672–1680.
4. Bohl DD, Russo GS, Basques BA, Golinvaux NS, Fu MC, Long WD 3rd, Grauer JN. Variations in data collection methods between national databases affect study results: a comparison of the Nationwide Inpatient Sample and National Surgical Quality Improvement Program databases for lumbar spine fusion procedures. *J Bone Joint Surg Am.* 2014;96:e193.
5. Bozic KJ, Bashyal RK, Anthony SG, Chiu V, Shulman B, Rubash HE. Is administratively coded comorbidity and complication data in total joint arthroplasty valid? *Clin Orthop Relat Res.* 2013;471:201–205.
6. Centers for Medicare & Medicaid Services. CMS dry run hospital-specific report for hospital-wide all-cause unplanned readmission (HWR) measure. 2012 Available at: http://www.qualitynet.org. Accessed August 15, 2016.
7. Enomoto LM, Hollenbeak CS, Bhayani NH, Dillon PW, Gusani NJ. Measuring surgical quality: a national clinical registry versus administrative claims data. *J Gastrointest Surg.* 2014;18:1416–1422.
8. Kulaylat AN, Engbrecht BW, Rocourt DV, Rinaldi JM, Santos MC, Cilley RE, Hollenbeak CS, Dillon PW. Measuring surgical site infections in children: comparing clinical, electronic, and administrative data. *J Am Coll Surg.* 2016;222:823–830.
9. Lawson EH, Louie R, Zingmond DS, Brook RH, Hall BL, Han L, Rapp M, Ko CY. A comparison of clinical registry versus administrative claims data for reporting of 30-day surgical complications. *Ann Surg.* 2012;256:973–981.
10. Lawson EH, Louie R, Zingmond DS, Sacks GD, Brook RH, Hall BL, Ko CY. Using both clinical registry and administrative claims data to measure risk-adjusted surgical outcomes. *Ann Surg.* 2016;263:50–57.
11. Pugely AJ, Martin CT, Harwood J, Ong KL, Bozic KJ, Callaghan JJ. Database and registry research in orthopaedic surgery: part I: claims-based data. *J Bone Joint Surg Am.* 2015;97:1278–1287.
12. Pugely AJ, Martin CT, Harwood J, Ong KL, Bozic KJ, Callaghan JJ. Database and registry research in orthopaedic surgery: part 2: clinical registry data. *J Bone Joint Surg Am.* 2015;97:1799–1808.
13. Shiloach M, Frencher SK Jr, Steeger JE, Rowell KS, Bartzokis K, Tomeh MG, Richards KE, Ko CY, Hall BL. Toward robust information: data quality and inter-rater reliability in the American College of Surgeons National Surgical Quality Improvement Program. *J Am Coll Surg.* 2010;210:6–16.
14. US Senate. HR 3590: The Patient Protection and Affordable Care Act. 2009. Available at: https://www.govtrack.us/congress/bills/111/hr3590/text. Accessed August 25, 2016.