January 2016

Using "big Data" To Study Orthopaedic Trauma Populations: Looking At Fractures From A Bird's Eye View As Illustrated By Three Studies

Andre Michael Samuel
Yale University, andre.samuel@yale.edu

Using "Big Data" to Study Orthopaedic Trauma Populations:
Looking at fractures from a bird's eye view as illustrated by three studies

A Thesis Submitted to the
Yale University School of Medicine
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Medicine

by

Andre Michael Samuel

2016

Abstract

Background

"Big data" is increasingly being used for orthopaedic research. However, the acute and complex nature of orthopaedic trauma makes data collection and data analysis difficult. This thesis presents three different clinical studies, which together illustrate how databases may best be used to answer clinical questions in orthopaedic trauma. Specifically the studies aim to show (1) how different databases capture trauma populations, (2) how databases may be used for hypothesis discovery studies, and (3) how databases may be used for hypothesis testing studies.

Study Questions

(1) How do populations of femoral shaft fracture patients differ in three commonly used national databases, specifically in regards to age and preexisting comorbidities? (2) What risk factors are associated with delayed surgery after elderly hip fractures in a national cohort and subsequently an institutional cohort? (3) Does hospital resource utilization differ between subpopulations of patients in Medicare Diagnosis Related Group 536 (fractures of the hip and pelvis), despite equal Medicare hospital reimbursement?

Methodology

(1) Patients with surgically managed femoral shaft fractures were identified in the Nationwide Inpatient Sample (NIS), National Surgical Quality Improvement Program (NSQIP) and National Trauma Data Bank (NTDB). The distributions of age and Charleston Comorbidity Index were compared between populations. (2) A retrospective cohort study was conducted of all elderly hip fracture patients receiving surgical management from 2011-2012 in the NTDB and from 2009-2015 at a single academic

trauma center. Multivariate analysis was used to identify the independent effect of various risk factors on surgical timing. (3) Patients with hip fractures, non-operative pelvic fractures, acetabulum fractures, and operative pelvic fractures were identified in the 2011 – 2012 NTDB. Total inpatient length of stay, intensive care unit (ICU) stay, and ventilator time were compared across groups using multivariate analysis that controlled for patient and hospital factors.

Results

(1) A predominantly older population with more preexisting comorbidities was found in NSQIP (age = 71.5, CCI = 4.9), while a substantially younger population with fewer preexisting comorbidities was fond in NTDB (age = 45.2, CCI = 2.1). Bimodal distributions in the NIS population indicate a more mixed population (age = 56.9, CCI = 3.2). Differences in age were all statistically significant ($p < 0.001$). (2) In the national cohort, mean time to surgery was 31.3 hours (standard deviation: 31.6 hours). The risk factors with largest association with delays were total arthroplasty surgery (coefficient, in hours [95% confidence interval]: 7.7 [6.1 – 9.3]) coagulopathy, including chronic anticoagulation (7.1 [6.1 – 8.0]), and congestive heart failure (6.9 [6.0 – 7.9]). In the institutional cohort, mean time to surgery was 32.4 hours (standard deviation: 29.0 hours). In this cohort, the only statistically significant risk factors associated with surgical timing were total arthroplasty surgery (24.5 [13.7 – 35.4]), transfer from outside hospital (22.1 [15.1 – 29.1]), warfarin anticoagulation (13.7 [8.5 – 18.8]), other anticoagulation (10.5 [2.4 – 18.5]), and preoperative hematocrit < 35% (5.5 [2.0 – 9.0]). (3) After controlling for patient and hospital factors, the difference in inpatient length of stay compared to hip fracture patients was -0.2 days (95% C.I.: -0.4 to -0.1 days; P = 0.001)

for non-operative pelvis fractures, 1.7 days (95% C.I.: 1.4 to 1.9 days; $P < 0.001$) for acetabulum fractures, and 7.7 days (95% C.I.: 7.0 to 8.4 days; $P < 0.001$) for operative pelvic fractures. Similar differences were also noted for IVU stay and ventilator time.

Conclusion

(1) While these three national databases have been commonly used for orthopaedic trauma research, differences in the populations they contain are not always readily apparent. Care must be taken to fully understand these populations before performing or evaluating database research, as these differences clearly affect observed outcomes. (2) Of all risk factors identified, access to arthroplasty and management of chronic anticoagulation may be the most modifiable in order to reduce delayed hip fracture surgeries. Physician call coverage and algorithms for more rapid reversal of anticoagulation, namely warfarin anticoagulation, warrant further investigation.

(3) Because hospitals are reimbursed equally for these subgroups of Medicare DRG 536, those centers that care for a greater proportion of more-complex pelvic trauma will experience lower financial margins per trauma patient, limiting their potential for growth and investment compared with competing institutions that may not routinely see high-energy trauma.

Acknowledgements

First, I would like to thank my thesis advisor, Dr. Jonathan N. Grauer, for the countless lab meetings, hours spent revising manuscripts, endless advocacy for his students, and constant support, both personal and financial. He has a tireless dedication to Yale medical student education and mentorship, and, without a doubt, my time spent working with Dr. Grauer will end up being one of the most formative parts of my orthopaedic education. Adam, Matt, and myself are so very fortunate to have had the opportunity!

I would also like to thank the other members of the Yale orthopaedic faculty who have supported be during my time at Yale, including, in alphabetical order, Dr. Michael Baumgaertner, Dr. Gary Friedlaender, Dr. Michael Leslie, Dr. Dieter Lindskog, and Dr. Brian Smith.

Finally, I would like to thank the Office of Student Research, which enabled my research through a One-Year Student Research Fellowship.

Table of Contents

Introduction to Thesis

The use of large national databases in orthopaedic research has grown substantially over the last 10 years.[1] While the randomized controlled trial (RCT) remains the gold standard for answering clinical questions, RCTs are often unfeasible, especially in orthopaedic settings. An adequately powered trial requires a large number of patients, often from multiple centers, with the logistics of this making RCTs very costly. The large Spine Patients Outcomes Research Trial (SPORT) was estimated to cost $30 million,[1, 2] while the smaller Bracing in Adolescent Idiopathic Scoliosis Trial (BrAIST) cost $7 million.[1]

In addition blinding to placebo or “sham surgery” treatments may be unethical in many cases, while this blinding has also proven to be essential to avoid bias based on patient expectations. For example Moseley et al found that after randomizing 180 patients to arthroscopic knee debridement, lavage, or sham surgery that patient reported outcomes were no different between groups.[3] Furthermore, when surgical treatment is not blinded there remains the potential for substantial cross over from placebo to treatment groups, as was seen in the SPORT trial with conservatively managed patients eventually wanting surgical treatment for non-resolving symptoms.[4]

As these issues make RCTs impractical in many cases, the observational cohort study has come to be the standard for clinical research in orthopaedic surgery. Indeed, a well-designed observational study, carefully controlling for potential sources of bias, may provide a strong level of evidence, which approaches that of an RCT. For example in the SPORT trial, while intention-to-treat data was found to be inconclusive due to high rates of cross-over,[4] a secondary observational study of the data based on per protocol analysis and carefully controlling for sources of bias, did demonstrate benefits with surgical treatment of herniated lumbar discs.[5]

Large national databases provide a unique platform for observational clinical studies providing large patient populations suitable for studying rare procedures, populations, and outcomes that would otherwise be difficult to study by a single provider, institution, or even group of institutions.[1, 6] As data collection is done preemptively, these studies can be completed at low cost, and with careful study design, several of the inherent biases of retrospective studies can be controlled for. The lack of orthopaedic-specific data elements in many of the presently general surgery-focused national databases does limit the ability to measure orthopaedic outcomes of interest, such as functionality measures, and control for factors that influence orthopaedic outcomes, such as implant use or fracture classification. Nevertheless, the wealth of data available in these national datasets has already allowed several impactful studies to be completed.[7-13]

Clinical studies utilizing "big data" in orthopaedic surgery can be classified into two broad categories: hypothesis-finding and hypothesis-testing studies. Hypothesis-finding studies leverage the large patient populations and numerous data elements contains in databases to identify risk factors associated with surgical outcomes. For example Basques et al. identified several risk factors for blood transfusion in over 4,000 patients undergoing primary posterior lumbar fusion, including a greater number of levels fused, longer operative time, and preoperative anemia.[14] This type of study opens the door for further investigation of targeted interventions that may prevent blood transfusions in this population. It also provides risk stratification information to providers for both patient counseling and surgical planning purposes. In contrast, hypothesis-testing studies aim to answer a narrowly defined questions using statistical testing controlling for other confounding factors. For example, Bohl et al determined that patients receiving intramedullary implants for intertrochanteric hip fractures had a shorter postoperative lengths of stay compared with those receiving extramedullary implants.[7] This type of study adds to the body of literature that may favor one method of treatment versus another, impacting clinical standards of care. In orthopaedic surgery, both hypothesis-finding and hypothesis-testing studies have utility in answering clinical questions and directing practice.

While many orthopaedic questions have been answered using "big data," one population that remains difficult to study are orthopaedic trauma populations. In the acute trauma setting, careful data collection is often difficult when urgent test and interventions are necessary. In addition there is considerable variability in injury patterns and associated injuries among fracture patients, all of which may not be well captured in databases but do still impact clinical management and patient outcomes. Nevertheless, "big data" has already been used extensively to study orthopaedic trauma populations. Therefore, it is important to identify effective methods for conducting clinical research on orthopaedic fracture populations, using large national databases.

The current work aims to address this goal in three sections, each an individual completed clinical study, either published or submitted for publication, illustrating a different aspect of research design. In Section I, entitled "Do we really know our patient population in database research: A comparison of the femoral shaft fracture patient populations in three commonly used national databases," the high-energy femoral shaft fracture populations in three databases are compared. By doing this, the importance of selection of an appropriate database for studying high-energy trauma populations is highlighted.

In Section II, entitled "Using a National Database and Subsequently an Institutional Cohort to Identify Potentially Modifiable Risk Factors for Delayed Hip Fracture Surgery: An example of using databases for hypothesis discovery," an example of a hypothesis-finding study is presented as risk factors for delayed surgery after elderly hip fractures are analyzed. In addition, these risk factors are further explored using a local institutional cohort of hip fracture patients to identify modifiable factors that may expedite surgery after hip fractures locally. In this way the utility of hypothesis finding for further investigative efforts is highlighted.

Finally in Section III, entitled "Variation in Resource Utilization for Hip and Pelvic Fracture Patients, Despite Equal Medicare Reimbursement," the specifically defined hypothesis of whether subgroups of patients in Medicare Diagnosis Related Group (DRG) 536 (fractures of the hip and pelvis) have different amounts of hospital resource utilization, is tested. Larger health policy implications of the findings are then discussed, illustrating the potential power of large well-powered databases studies for identifying clinically significant trends. Together the aim of the current thesis is to illustrate the utility of thoughtfully designed clinical research using "big data" to study the complex population of fracture patients.

Section I

Do we really know our patient population in database research: A comparison of the femoral shaft fracture patient populations in three commonly used national databases

Published as: Samuel AM, Lukasiewicz AM, Webb ML, Bohl DD, Basques BA, Varthi AG, Leslie MP, Grauer JN. "Do we really know our patient population in database research?: A comparison of the femoral shaft fracture patient populations in three commonly used national databases." Bone Joint J. (Accepted Sept 2015, in press)

Background

The use of national databases for clinical research has increased dramatically in the field of orthopaedic surgery over the past 5 years (Figure 1). Examples of databases currently being used in orthopaedic research include the Nationwide Inpatient Sample (NIS),[15], the American College of Surgeons National Surgical Quality Improvement Program database (NSQIP),[16] and the American College of Surgeons National Trauma Data Bank Research Data Set (NTDB RDS).[17] Use of these databases can expedite clinical research and allows study of rare injury patterns using large cohorts, not possible for the individual researcher or even individual institution. However with increased use of these databases, it is critical to understand how the database populations differ, and how patients are included and excluded from each, before conclusions can be made and generalized to external populations.

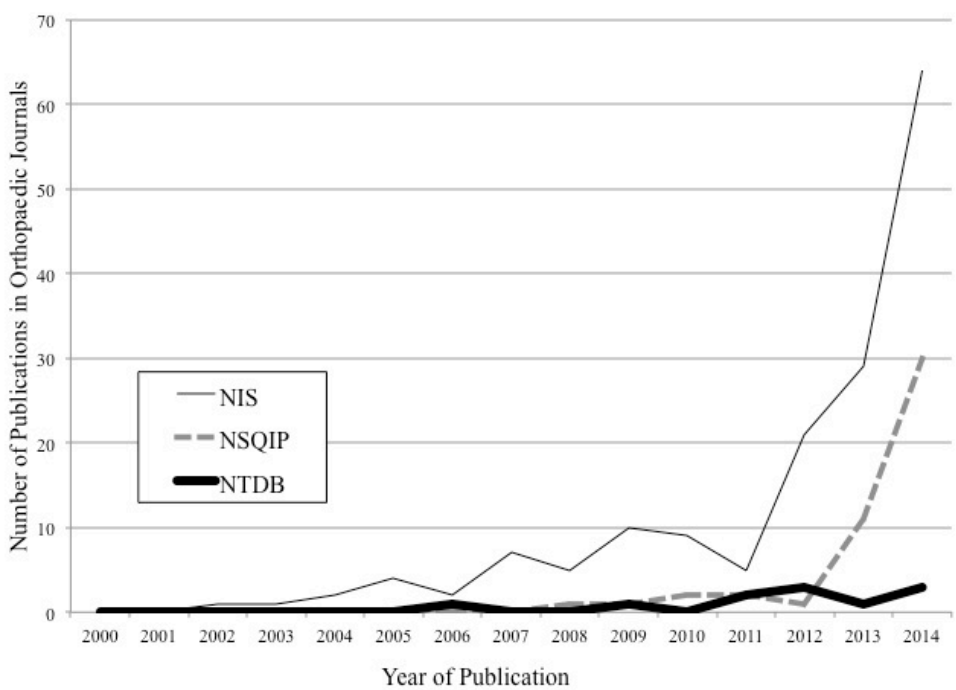

Figure 1: Database studies published in orthopaedic surgery journals have increased since the year 2000. Studies utilizing the Nationwide Inpatient Sample (NIS), the National Surgical Quality Improvement Program (NSQIP), or the National Trauma Data Bank (NTDB) were identified from a PubMed search of the top 41 orthopaedic journals as defined by Moverly, et al.1 Publications in 2014 include only journals published (electronically or in print) by September 9, 2014.

Previous comparisons of databases have focused on differences in quality of various data elements, namely patient comorbidities and adverse events, in similar patient populations. These studies of patients undergoing lumbar vertebral fusions,[18] hip fracture fixation,[19] pancreaticoduodenectomy,[20] and esophageal resection[21] identified considerable differences in the recording of comorbidities and adverse events between administratively-coded databases, such as NIS, and registry databases, such as NSQIP. These studies, however, focused on patient populations that were largely similar in terms of basic demographics between databases.

Regarding high-energy orthopaedic trauma, various databases may differ in what patients are included due to varying inclusion criteria. However of the databases cited above, all three have been used extensively for orthopaedic trauma research,[7, 22-30] with little prior discussion of how the populations may actually compare to the true general population of orthopaedic trauma patients. As orthopaedic trauma is an area where database research may provide particular insights, it is critical to understand the various differences between orthopaedic trauma populations in various national databases. The current retrospective cohort study aims to determine (1) how the populations of patients with femoral shaft fractures, a common high-energy injury, differ in these three commonly used national databases, specifically in regards to age and preexisting comorbidities, and (2) how these differences may affect studies of in-hospital adverse events observed in each population.

Patients and Methods

Study design and setting

A retrospective cohort study was performed using NIS, NSQIP, and NTDB RDS. For NSQIP and NTDB, data from 2011 and 2012 were used. Only 2011 data were used for NIS, as 2012 data were not available at the time of study. The study was approved by the local institutional human investigations committee.

Participants/study subjects

ICD-9 procedure codes were used to identify patients undergoing surgical treatment of femoral shaft fractures in NIS and NTDB (ICD-9 79.05, 79.15, 79.25, 79.35). Current Procedural Terminology (CPT) codes were used to identify the same population in NSQIP (CPT 27500, 27502, 27506, 27507). Differing methods for the identification of patients were necessary due to differing the data elements within each database.

Variables, outcome measures, data sources, and bias

Patient age, comorbidities, and adverse events were analyzed. Within all three databases, binary variables were used to assess comorbidities. Seven comorbidities with equivalent definitions across the three databases were analyzed: alcoholism, coagulopathy, diabetes mellitus, cancer, hypertension, obesity, and current smoker status. Appendix 1 lists specific data elements used to identify comorbidities in each database. A modified Charlson comorbidity index (CCI) was then calculated using the seven comorbidities analyzed and patient age. Similar to the original index,[31] the modified CCI assigns one point for each comorbidity diagnosis (except for disseminated cancer, as in original CCI six points are assigned due to poor life expectancy). One point was then added for each decade greater than 40 years of age. A modified CCI has been shown to have comparable predictive value to the original CCI,[32] and has been used previously with national databases such as NSQIP.[8]

Adverse events were assessed using binary variables in NTDB and NSQIP and ICD-9 diagnosis codes in NIS. As NSQIP contains data on adverse events occurring up to 30 days postoperatively, while NTDB and NIS contain only inpatient data, adverse events occurring after discharge in NSQIP were excluded. Nine adverse events with equivalent definitions across databases were analyzed: acute kidney injury (AKI), cardiac arrest, cerebrovascular accident (CVA), death, deep vein thrombosis or pulmonary embolism (DVT/PE), myocardial infarction (MI), pneumonia, surgical site infection (SSI), and urinary tract infection (UTI). Mortality rate, serious adverse event (SAE) rate (death, cardiac arrest, MI, CVA, DVT/PE, and SSI), and all adverse event (AAE) rate were computed. Appendix 2 lists specific data elements used to identify adverse events in each database.

Statistical analysis, study size

Several different analyses were conducted in this study. First, age was compared between populations. Mean age was compared using analysis of variance (ANOVA) and age was plotted on a histogram. CCI was similarly compared between populations. Mean CCI was compared using ANOVA and CCI was plotted on a histogram.

Next, comorbidities were compared between databases. Rates of each of the seven comorbidities analyzed were compared between databases using Pearson's chi squared statistic. The relative risk of each comorbidity was then estimated for both NSQIP and NTDB compared to NIS as a reference, and then NSQIP compared to NTDB as a reference. This was done using Poisson regression with robust error variance using an indicator variable to designate the database. As Poisson regression overestimates the error for relative risk when using binomial data, robust error variance was used to directly estimate the error for the relative risk. NIS was chosen as the denominator for both comparisons as this is thought to be the most nationally representative patient population.

Concerning adverse events, rates of mortality, serious adverse events (SAE), and any adverse event (AAE), were computed for each population. Pearson's chi square statistic was used to compare rates between populations. Multivariate ANOVA was then used to compare the rates after adjusting for age and CCI.

Finally, a simulated theoretical analysis of risk factors associated with inpatient adverse events was conducted for each database population. Multivariate logistic regression was used to determine the association of the age and the seven individual comorbidities previously studied with the rate of AAEs.

All statistical analyses were performed using Stata® version 13.0 (StataCorp, LP, College Station, Texas, USA). All statistical tests were two-tailed, and the level of significance was set at $\alpha = 0.05$.

Results

A total of 25,121patients undergoing surgical repair of a femoral shaft fracture were identified and included in this study. There were 3,943 patients identified and included within NIS (2011 only, 2012 data was not available at time of study), 663 patients identified and included within NSQIP (2011-2012), and 20,515 identified and included within NTDB (2011-2012).

The mean age of patients in NIS was 56.9 years (standard deviation [SD]: 24.9 years). The mean age of patients in NSQIP was 71.5 years (SD: 15.6 years). The mean age of patients in NTDB was 45.2 (SD: 21.4 years). Differences in mean age between databases were statistically significant ($p < 0.001$). Both NSQIP and NTDB round ages greater than 90 to a value of 90 to prevent breach of privacy. The distribution of ages is markedly different between NSQIP (primarily older patients) and NTDB (primarily younger patients), with the bimodal distribution in NIS seemingly representing a more mixed population (Figure 2).

The mean CCI of patients in NIS was 3.2 (SD: 2.3). The mean CCI of patients in NSQIP was 4.9 (SD: 1.9). The mean CCI of patients in NTDB was 2.1 (SD: 2.0). Differences in mean CCI between databases were statistically significant ($p < 0.001$). Similar to age, the distribution of CCI is markedly different between patients in NSQIP (more comorbidities) and NTDB (primarily fewer comorbidities), with NIS seemingly representing a more mixed population with a bimodal distribution (Figure 3).

Rates of all individual comorbidities were greatest in the NSQIP population, except for alcoholism and current smoker status, which were greatest in the NTDB population (Figure 4). Incidence rates in the NIS population were in between the incidence rates in the NTDB and NSQIP populations for all seven comorbidities. The relative risks of each comorbidity in the NSQIP population was statistically significant when compared to NIS as a standard ($p < 0.05$), except for diabetes ($p = 0.173$) and current smoker status ($p = 0.184$). The relative risks of all comorbidities in the NTDB population were statistically significant when compared to NIS as a standard ($p < 0.05$).

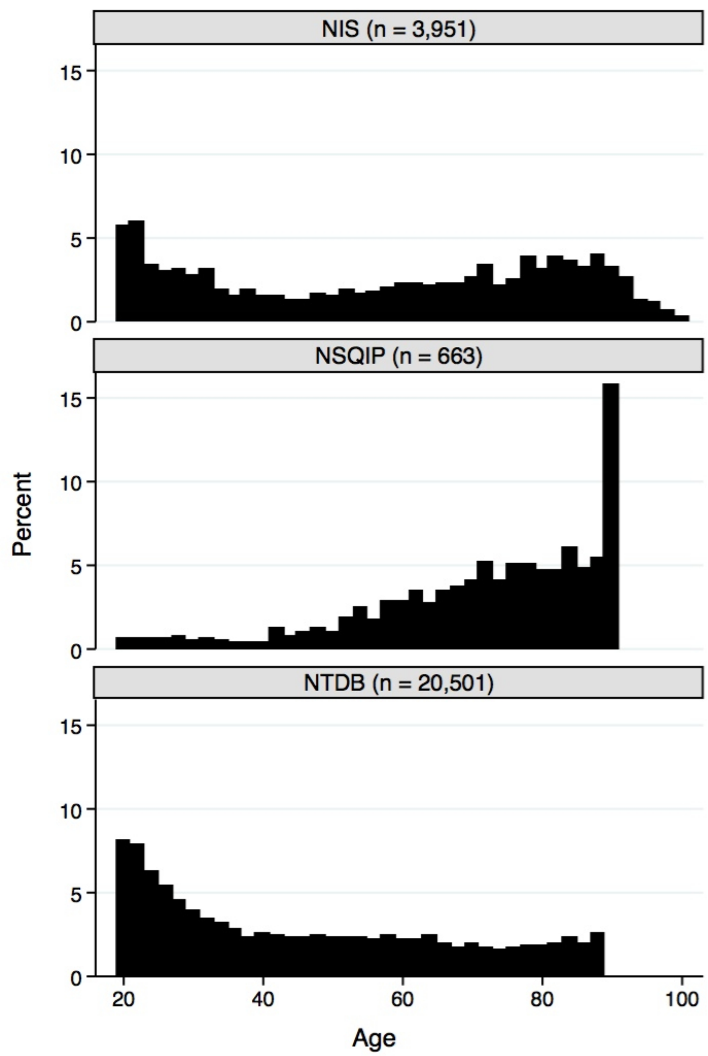

Figure 2: Age of femoral shaft fracture patients vary by database. Histograms are of age distributions within the Nationwide Inpatient Sample (NIS), the National Surgical Quality Improvement Program (NSQIP), and the National Trauma Data Bank (NTDB). Note that NSQIP and NTDB truncate ages greater than 90 to age = 90.

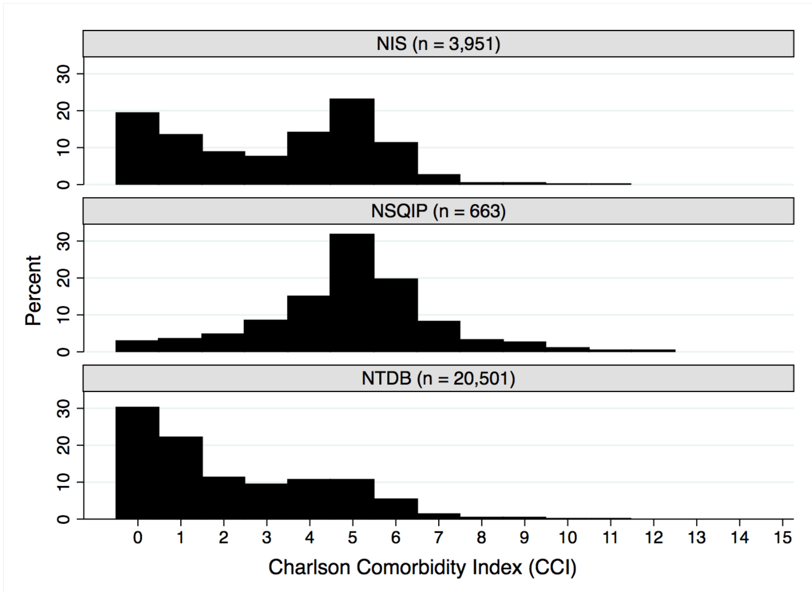

Figure 3: Charlson Comorbidity Index of femoral shaft fracture patients vary by database. Histograms are of CCI distributions within NIS, NSQIP, and NTDB.

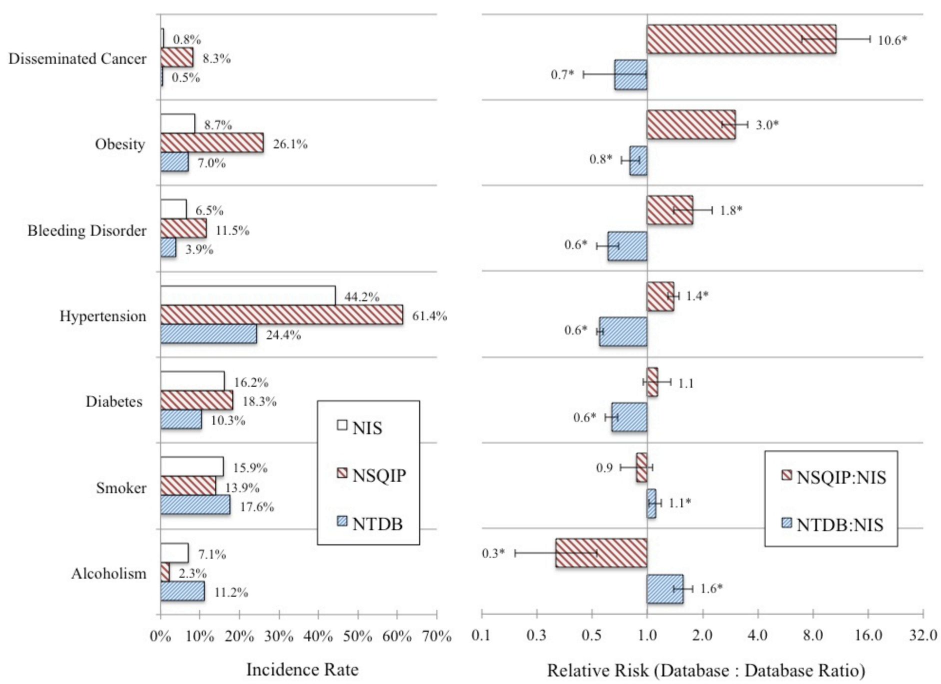

Figure 4: The rates of individual comorbidities vary by database. On left are bar graphs of incidence rates in the NIS, NSQIP, and NTDB populations. On right are bar graphs of relative risks (computed using Poisson regression with robust error variance) of comorbidities in NSQIP and NTDB compared to NIS as a reference. Asterisks (*) indicate statistically significant relative risks. Error bars represent 95% confidence interval of relative risks.

The differences in mortality rate between the different databases (range 1.1% [NIS] to 2.0% [NSQIP]; Figure 5) were not statistically significant before risk adjustment (P = 0.059), however after adjusting for age and CCI the differences became significant (P < 0.001). The differences in SAE rate between the different databases (range 5.1% [NIS] to 7.4% [NTDB]) were statistically significant before and after risk adjustment (P < 0.001, P < 0.001). The differences in AAE rate between the different databases (range 9.1 [NSQIP] to 21.6% [NIS]) were statistically significant before and after risk adjustment (P < 0.001, P < 0.001).

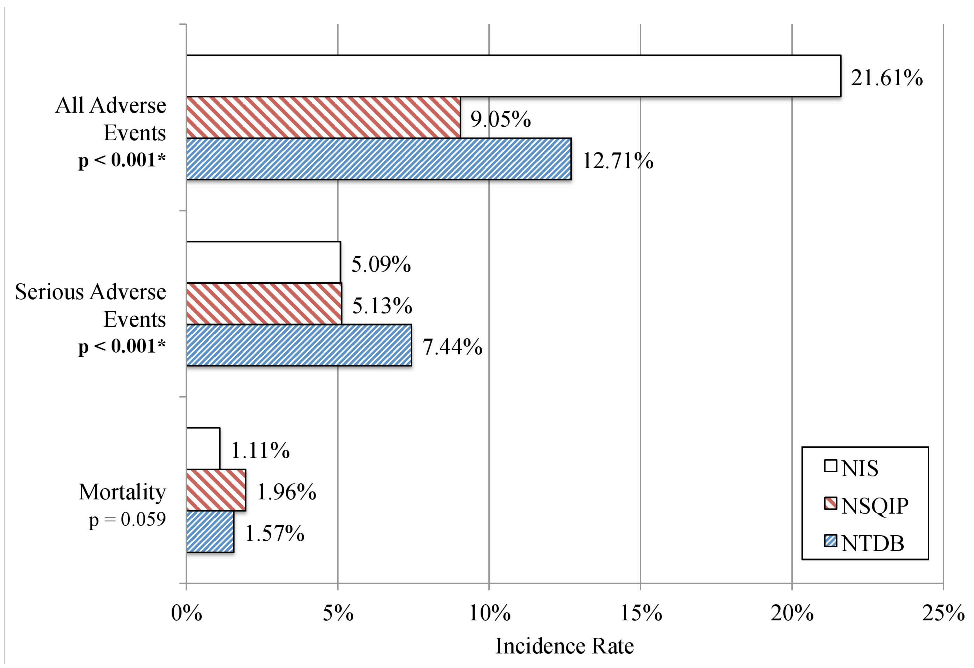

Figure 5: Adverse event rates in femoral shaft fracture patients vary by database. Bar graphs indicate incidence rates of mortality, serious adverse events, and all adverse events in the NIS, NSQIP, and NTDB populations. The p-values are given for Pearson's Chi Square statistic for the distribution of adverse events amongst databases. Statistically significant findings (p < 0.05) are indicated by asterisks and boldface.

In the simulated theoretical analysis there were several differences in the identified risk factors for inpatient adverse events (Table 1). In the NIS population older age was associated with increased AAEs (odds ratio [95% confidence interval]: 4.12

[2.82 - 6.01]), while younger age was associated with decreased AAEs (0.57 [0.39 - 0.84]). Old age is similarly associated with increased AAEs in the NSQIP population (5.56 [1.13 – 27.42]). However, in the NTDB population, older age (in the 60-69, 70-79, and 80+ age groups) is associated with fewer AAEs (0.78 [0.65 – 0.92]). Similarly, diabetes is associated with more AAEs in the NIS population (1.40 [1.14 – 1.72]) and in the NTDB population (1.41 [1.23 – 1.61]), and with fewer AAEs in the NSQIP population (0.41 [0.18 – 0.95]). Additionally, for 4 of the 6 additional comorbidities studied (alcoholism, smoking, obesity, and disseminated cancer), there was a change in the statistical significance of the association with AAEs.

Table 1: Theoretical analysis of risk factors associated with inpatient adverse events after femoral shaft fractures in three commonly used national clinical databases

	NIS (n = 3,951)		NSQIP (n = 663)		NTDB (n = 20,501)	
	Odds ratio (95% confidence interval)	P-value	Odds ratio (95% confidence interval)	P-value	Odds ratio (95% confidence interval)	P-value
Alcoholism	1.22 (0.84 - 1.77)	0.296	0.00 (omitted)*	-	1.39 (1.23 - 1.57)	< 0001
Smoking	0.60 (0.44 - 0.81)	0.001	1.68 (0.77 - 3.69)	0.196	0.86 (0.77 - 0.96)	0.010
Diabetes	1.40 (1.14 - 1.72)	0.001	0.41 (0.18 - 0.95)	0.038	1.41 (1.23 - 1.61)	< 0001
Hypertension	1.04 (0.86 - 1.26)	0.699	1.54 (0.82 - 2.89)	0.182	1.06 (0.95 - 1.19)	0.300
Bleeding	1.87 (1.41 - 2.48)	< 0.001	2.06 (1.03 - 4.12)	0.041	1.94 (1.62 - 2.31)	< 0001
Obesity	1.63 (1.25 - 2.14)	< 0.001	1.15 (0.60 - 2.19)	0.676	1.62 (1.41 - 1.87)	< 0001
Cancer	2.09 (1.00 - 4.34)	0.049	1.04 (0.38 - 2.85)	0.938	0.89 (0.50 - 1.61)	0.705
Age						
18 - 29	0.57 (0.39 - 0.84)	0.005	0.00 (omitted)*	-	0.69 (0.60 - 0.79)	< 0001
30 - 39	0.86 (0.56 - 1.31)	0.477	0.00 (omitted)*	-	0.86 (0.73 - 1.00)	0.051
40 - 49	1.26 (0.82 - 1.92)	0.29	2.52 (0.33 - 19.38)	0.374	1.04 (0.89 - 1.22)	0.581
50 - 59	Reference		Reference	-	Reference	-
60 - 69	1.38 (0.96 - 2.00)	0.086	3.53 (0.74 - 16.81)	0.113	0.72 (0.61 - 0.85)	< 0001
70 - 79	1.53 (1.07 - 2.18)	0.019	4.81 (1.06 - 21.74)	0.041	0.66 (0.55 - 0.79)	< 0001
80 - 89	2.84 (2.02 - 3.99)	< 0.001	3.82 (0.81 - 17.92)	0.09	0.78 (0.65 - 0.92)	0.004
90 +	4.12 (2.82 - 6.01)	< 0.001	5.56 (1.13 - 27.42)	0.035	-	-

Note: NIS = Nationwide Inpatient Sample, NSQIP = National Surgical Quality Improvement Program, NTDB = National Trauma Data Bank

*Categories were omitted from NSQIP analysis that perfectly predicted absence of adverse events in multivariate analysis

Light shading indicates statistically significant odds ratio with positive association with inpatient adverse events

Dark shading indicates statistically significant odds ratio with negative association with inpatient adverse events

Discussion

The field of orthopaedic surgery has seen increased use of large national databases for clinical research recently. Over the past three years there has been over a 200% increase in peer-reviewed publications utilizing NIS, NSQIP, and NTDB, in orthopaedic journals alone (Figure 1). In the field of orthopaedic trauma, national database use has increased as well, likely due to increased use of trauma-specific databases such as NTDB. However, as the current study indicates, various commonly used national clinical databases treat trauma patients very differently, and the resulting differences in database populations lead to significant differences in observed outcomes.

After comparing all patients with surgically treated femoral shaft fractures in NIS, NSQIP, and NTDB, the current study found that the NSQIP and NTDB populations are strikingly different, when considering age and comorbidities. The NTDB population is younger and has fewer preexisting comorbidities, while the NSQIP population is older and has more preexisting comorbidities. The NIS population appears to be bimodal in terms of age and comorbidities. While previous database comparisons have primarily focused on differences in specific data elements among demographically similar populations,[19-21] the current study is the first to highlight the significant demographic differences of trauma populations in national databases that are each commonly used for orthopaedic trauma research.

As expected, the observed differences in populations also lead to differences in outcomes. There were statistically significant differences in the rates of mortality, SAEs, and AAEs between the three database populations, after controlling for age and CCI. Interestingly mortality was not significantly different between populations, likely due to the low overall mortality rate among femoral shaft fracture patients (ranging from 1.1 - 2.0%). In our theoretical study of risk factors for inpatient adverse events, there were several differences in the simulated results. Most striking was the change in directionality of the association of certain risk factors (diabetes, older age) with adverse events, depending on the database studied. This illustrates the importance of understanding the database population being studied before attempting to generalize results of a study to other clinical populations.

The differences in database populations demonstrated in the current study highlight the relative strengths and weakness of each database (Table 2). Based on sampling methodology, NIS is the most nationally representative and comprehensive sample. NSQIP and NTDB include only data from voluntarily participating institutions, and are likely biased towards larger hospitals that have interests in quality monitoring and improvement. In contrast the Hospital Cost and Utilization Project takes data from all inpatient discharges from a nationally representative sample of community hospitals (approximately 20% of all U.S. community hospitals). While this is a relative strength of the NIS dataset, the quality of the billing code data it contains has previously been called into question.[33-35] NIS contains strictly ICD-9 billing code diagnoses compared to NSQIP and NTDB, which utilize chart-abstracted diagnoses for comorbidities and adverse events. ICD-9 coding inconsistencies have been noted in past studies comparing databases with administratively-coded versus chart-abstracted data. [19, 33-36] As administrative billing codes are subject to economic and political pressures, there is potential for underreporting or over-reporting certain diagnoses.

Table 2: Comparison of three commonly used national clinical databases

	Nationwide Inpatient Sample (NIS)	National Surgical Quality Improvement Program (NSQIP)	National Trauma Data Bank (NTDB) Research Data Set (RDS)
Overseeing organization:	Healthcare Cost and Utilization Project (HCUP)	American College of Surgeons	American College of Surgeons
Earliest year available:	1988	2005	2007
Participating hospitals:	Stratified systematic sample from all HCUP hospitals (equivalent to 20% of all discharges from U.S. community hospitals).	NSQIP participant institutions (requires annual membership renewal and NSQIP surgical clinical reviewer on staff)	Any voluntarily participating institution
Patient inclusion criteria:	All inpatient discharges	Systematic sampling (every 8 days) of hospital's daily surgical log. Hospital may include general/vascular cases only or multispecialty. Most acute trauma, transplant, and minor surgical cases excluded.	All trauma cases (based on ICD-9 diagnosis) resulting in inpatient admission or death.
Data collection methods:	Hospital billing codes	Chart review by SNIP-trained surgical clinical reviewer with regular inter-rater reliability audits	Institution dependent (chart abstraction/ billing code data)
Number of cases (2011):	8,023,590	442,149	773,299
Number of hospitals (2011):	1,049	315	744
Number of variables reported:	187	252	116
Diagnosis coding:	ICD-9	ICD-9	ICD-9
Number of diagnoses reported:	Multiple	Single preoperative diagnosis	Multiple
Procedure coding:	ICD-9	CPT	ICD-9
Number of procedures reported:	Multiple	Multiple	Multiple
Procedure timing:	By day	By day	By hour
Pre-hospital data?	Not available	Not available	From initial EMS dispatch
Post-discharge data?	Not available	Up to 30 postoperative days	No
Comorbidity data source:	ICD-9 diagnosis codes	Chart-review by surgical clinical reviewers using NSQIP comorbidity definitions	Institution dependent (Chart-review/ICD-9 codes). ND recommended comorbidity definitions.
Adverse events data source:	ICD-9 diagnosis codes	Chart-review by surgical clinical reviewers using NSQIP adverse event definitions	Institution dependent Chart-review/ICD-9 codes). NTDB recommended adverse event definitions.

The ACS NSQIP is primarily geared towards elective surgery, with some institutions only reporting elective general surgery and vascular surgery cases. Most acute trauma and transplant procedures are excluded. Review of NSQIP data collection guidelines for surgical clinical reviewers indicates that most high energy mechanisms of injury are excluded from the database, such as motor vehicle accidents and firearms.[37] Lower energy mechanisms that are included consist of falls from standing or from up to 3 steps resulting in non-penetrating, single-bone or single organ system injury.[37] Therefore, the fracture population contained in NSQIP likely consists largely of patients with preexisting comorbidities such as osteoporosis, preexisting prostheses, prior fractures or nonunions, or bony metastasis, all resulting in weaker bone stock and higher likelihood of fracture with lower energy mechanisms. Nevertheless, despite NSQIP including only a small subset of all fractures, studies have been published utilizing NSQIP to draw conclusions regarding orthopaedic trauma patient populations.[19, 28-30] While NSQIP does have the most reliable data collection system, including specially trained NSQIP surgical clinical reviewers and regular inter-rater reliability audits; trauma researchers should carefully consider whether NSQIP truly captures the intended population before using the database for clinical studies.

In contrast, NTDB likely contains the best representation of the acute high-energy fracture patient population. All patients admitted through the emergency department at participating institutions with a trauma-specific ICD-9 diagnosis code (800.0 – 959.9) meet NTDB inclusion criteria.[17] However, these criteria do not include pathologic fractures (ICD-9 733.1X), stress fractures (733.9X), or fracture nonunions and malunions (733.8X). These reasons help explain why the population of surgically treated femoral shaft fracture patients in NTDB is primarily younger patients, likely with higher-energy mechanisms of injury.

No database serves as a "gold standard" for study of trauma populations. Rather, study of each of these individual trauma populations does have merit. Femoral shaft fractures represent a spectrum of injuries resulting from both low-energy and high-energy mechanisms, as demonstrated by the bimodal age distribution seen in NIS. As a true national sampling of inpatients, NIS would be ideal for national-level demographic and incidence studies of all types of femoral shaft fracture patients. Studies intending to focus on osteoporotic type stress fractures would be well suited to the NSQIP population. On the other hand, NTDB is skewed toward younger, high-energy patients seen at Level I and II, and is best suited for study of this unique population. [17]

Furthermore, each of the databases offers unique data elements that allow for differing study opportunities. NSQIP has several unique intraoperative variables (i.e. operative time, anesthesia type, intraoperative transfusion) in addition to documentation of adverse events that occur up to 30 days after surgery (often after discharge). NTDB contains detailed pre-hospital variables, such as ambulance time and mechanism of injury, and emergency department variable. NIS, in contrast, largely contains only data derived from ICD-9 billing codes, but is useful due to its nationally weighted sampling from all U.S. inpatient admissions.

The primary limitation of this study is the lack of hospital specific identifiers that could be matched between databases. This would allow comparison of the patient samples obtained from the same institutions. However, in order to maintain anonymity of the participating institutions, these data are not present in any of the three databases.

While it is impossible to tell whether differences in populations are due to differences in the institutions included versus differences in how trauma patients are included within individual institutions, the differences observed are still significant and warrant careful consideration for future studies.

When planning database studies, careful consideration must be made to ensure researchers and those evaluating database research fully understand the populations being studied. Due to large sample sizes, statistically significant results may be drawn from each database. Therefore appreciation of population demographics and database inclusion criteria is imperative to ensure that the conclusions drawn are indeed valid and are generalized to the appropriate clinical populations.

Section II

Using a National Database and Subsequently an Institutional Cohort to Identify Potentially Modifiable Risk Factors for Delayed Hip Fracture Surgery: An example of using databases for hypothesis discovery

Submitted as: Samuel AM, Varthi AG, Fischer JM, Lukasiewicz AM, Webb ML, Bohl DD, Basques BA, Lane JM, Grauer JN. "Using a national database and subsequently an institutional cohort to identify potentially modifiable risk factors for delayed hip fracture surgery" Clin Orthop Relat Res (Submitted January 2016)

Background

There is considerable evidence associating earlier surgery after elderly hip fracture with reduced mortality [38-45], shorter length of stay [39, 44, 46], improved discharge disposition [44, 47], and fewer complications [48]. Currently, both in the United States and United Kingdom, national orthopaedic practice guidelines support early operative management and mobilization of these patients. In 2010, the England and Wales Department of Health instituted a Best Practice Tariff providing financial incentives for hospital to follow a specified care pathway for patients with hip fractures, including surgery within 36 hours of admission [49, 50]. In 2014 the American Academy of Orthopaedic Surgeons (AAOS) released a Clinical Practice Guideline, entitled Management of Hip Fractures in the Elderly, recommending surgery within 48 hours of admission for all elderly hip fracture patients [51]. In addition, a recent Danish cohort study demonstrated that surgical delay past 12 hours was associated with increased rates of 30-day mortality [52]. With the Canadian Hip Fracture Accelerated Surgical Treatment And Care Track Trial (HIP ATTACK) now ongoing [53], examining potential benefits of accelerated surgery within 6 hours, it is possible that clinical evidence and national bodies will further support even earlier surgical intervention in these patients.

With the growing literature supporting early and accelerated surgical management of elderly hip fractures, there is a need to identify means to reduce our current preoperative time in the most efficient means possible. A recent cohort study of 2 million patients in the National Inpatient Survey (NIS) database identified risk factors for delayed surgery: race, Medicaid insurance, Northeastern hospital location, and higher comorbidity scores were associated with delayed surgical management [54]. However, these factors are large non-modifiable. In addition, the authors did not comment on the marginal effect of various risk factors on overall surgical timing. The current analysis first utilizes a large national cohort to identify medical comorbidities that may predispose patients to surgical. Next, based on the results from the national cohort, a more focused analysis of more modifiable risk factor was conducted using a cohort from a single academic trauma center. The authors of the present study also hope to present a new paradigm for clinical research using national databases, utilizing a national cohort first to better focus the analysis of a local cohort.

The objectives of this study are to answer the following: (1) What risk factors are associated with delays to surgery in a national cohort? (2) What modifiable risk factors are associated with delays to surgery within in our institutional cohort?

Patients and Methods

Study Design and Setting

Two separate populations were studies. First a retrospective cohort study was conducted using the 2011-2012 American College of Surgeons (ACS) National Trauma Data Bank (NTDB). NTDB is the largest national database of trauma cases, including cases from over 900 voluntarily participating trauma centers across the United States[17]. Next, a retrospective cohort study was conduct at a single academic trauma center located in an urban setting.

Participants/Study Subjects

In both the national and institutional cohorts patients were identified using International Classification of Diseases, Ninth Revision (ICD-9) diagnosis codes All patients aged 65 years or greater with femoral neck fractures (820.0X, 820.1X), intertrochanteric fractures (820.2X, 820.3X), or unspecified hip fractures (820.8, 820.9) were included in the study. The NTDB cohort included patients from admission years 2011 and 2012, while the institutional cohort included patients from admission years 2009 to 2015. The following ICD-9 procedure codes were then used to identify patients that underwent surgical fixation: 79.15 (closed reduction and internal fixation of the femur), 79.35 (open reduction and internal fixation of the femur), 81.52 (hemiarthroplasty of the hip), 81.51 (total arthroplasty of the hip), and 81.40 (repair of hip, not elsewhere classified).

All patients without ICD-9 codes for surgical fixation or with missing data regarding time to surgical fixation were excluded. Furthermore, as the NTDB contains a disproportionate number of high-energy trauma patients from level-one trauma centers, only patients with Injury Severity Score < 9 were included in this population, to best represent the population of patients suffering typical elderly fragility fractures of the hip. An Injury Severity Score of 9 corresponds with a single, isolated hip fracture.

Description of experiment, treatment, or surgery

All patients underwent surgical treatment of their hip fracture during initial hospital admission in both cohorts. Patient were categorized based on the type of surgical procedures: closed reduction and internal fixation of the femur (CRIF), open reduction and internal fixation of the femur (ORIF), hemiarthroplasty of the hip, total arthroplasty of the hip, and repair of the hip not elsewhere classified.

Variables, outcome measures, data sources, and bias

In the national cohort, the following risk factors for delayed surgery were analyzed: gender, age, coagulopathy (including anticoagulation), congestive heart failure (CHF), coronary artery disease, obesity, chronic respiratory disease, functionally-dependent status, diabetes mellitus, type of fracture (femoral neck or intertrochanteric), type of surgery (CRIF, ORIF, hemiarthroplasty, total arthroplasty), and type of hospital (academic or non-academic). Age, gender, comorbidities, and hospital type are reported in NTDB as chart-abstracted data elements. Fracture classification and type of surgery are reported as ICD-9 codes.

After review of initial results from the national cohort, more focused risk factors were identified for analysis in the institutional cohort. Gender was analyzed while age was not, as this was not significant even in the large national cohort. As coagulopathy was a significant risk factor in the national cohort, in the institutional cohort preoperative anemia (hematocrit < 30%), chronic warfarin anticoagulation therapy, other chronic anticoagulation therapy, and regular aspirin use were studied as risk factors. As CHF was a significant risk factor in the national cohort, recent exacerbation of CHF symptoms and inpatient furosemide therapy were studied as risk factors. Coronary artery disease, obesity, and transfer from an outside hospital (OSH) were studied as risk factors. In addition, type of hip fracture and type of surgery were studied as in the national cohort. All patient level data was determined based on chart-review or ICD-9 coding (for fracture type and surgery type).

The primary outcome measure in this study is surgical timing. Time from initial hospital admission to start of surgical intervention was calculated for each patient. Surgical start time was defined as time of initial incision.

Statistical analysis, study size

The mean (and standard deviation) time to surgery was computed in each cohort. In addition, the mean (and standard deviation) time to surgery was determined for patients with each risk factor studied. Multivariate regression analysis was used to determine the independent effect of each risk factor on surgical timing (in hours) in both cohorts. In the national cohort, 44,900 patients were analyzed. In the institutional cohort, 1,196 patients were analyzed.

All statistical analyses were performed using Stata® version 13.0 (StataCorp, LP, College Station, Texas, USA). All statistical tests were two-tailed and the level of statistical significance was set at α = 0.05. In order to reduce risk of a type 1 error, a Bonferroni correction was used to determine the threshold for statistical significance. Based on the 18 different hypotheses tested, the level of statistical significance was reduced to α = 0.003.

Demographic, description of study population

Of the 44,900 patients in the national cohort, 32,337 (72%) were female and 12,563 (28%) were male. The median age was 81 years (interquartile range [IQR]: 74 – 85; note that NTDB limits maximum age to 89 years). A total of 18,983 patients had intertrochanteric fracture (42%), while 7,390 had femoral neck fractures (16%), and 18,527 had unspecified hip fractures (41%). A total of 11,069 patients underwent CRIF surgery (25%), 19,030 underwent ORIF surgery (42%), 12,904 underwent hemiarthroplasty surgery (29%), and 1,897 underwent total arthroplasty surgery (4%).

Of the 1,196 patients in the institutional cohort, 866 (72%) were female and 330 (28%) were male. The median age was 85 years (interquartile range [IQR]: 78 – 90). A total of 626 patients had intertrochanteric fracture (52%), while 450 had femoral neck fractures (38%), and 120 had unspecified hip fractures (10%). A total of 115 patients underwent CRIF surgery (10%), while 699 underwent ORIF surgery (58%), 160 underwent hemiarthroplasty surgery (13%), 35 underwent total arthroplasty surgery (3%), and 187 had an unspecified repair of the hip (16%).

Results

National cohort

The median time to surgery in the national cohort was 24 hours (IQR: 17 – 39 hours). The distribution of time to surgery can be seen in Figure 1, with the number of patients who would have been considered as having had appropriate early surgery as indicated by the 2014 AAOS Guideline (86%), the 2010 UK NHS Best Practice Tariffs (74%), and the 2015 Danish Hip Fracture Registry Study (11%).

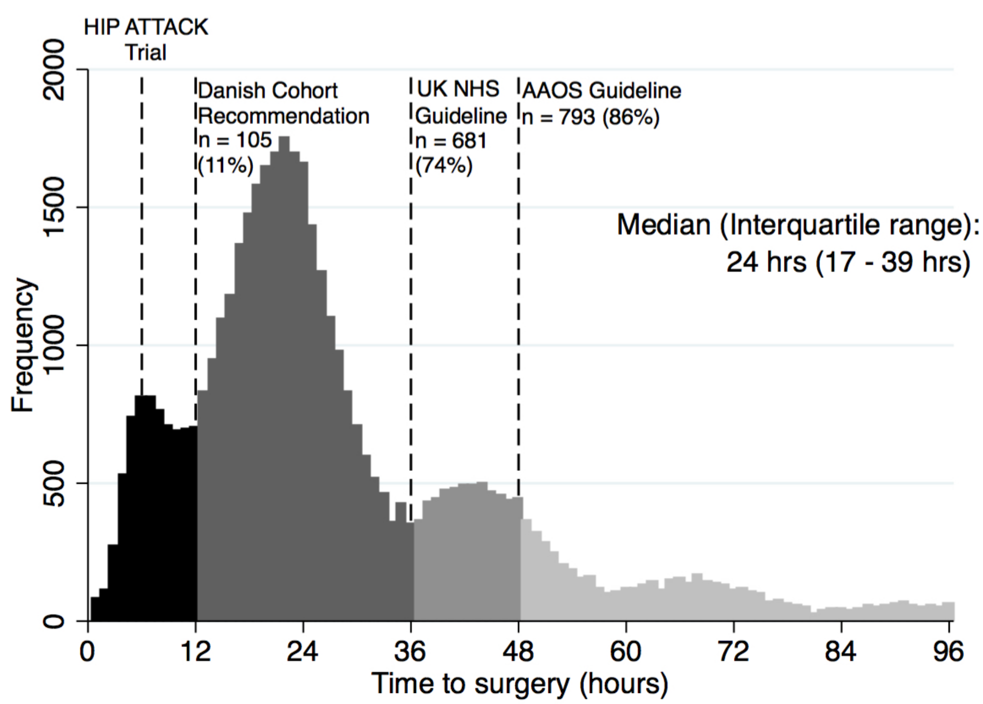

Figure 1: Distribution of surgical timing in the national elderly hip fracture cohort (2011 – 2012 NTDB, n = 44,900)

Patient / surgical variables for this cohort are presented in Table 1, with the number of patients and median time to surgery indicated in the first two data columns. Of all the risk factors for delay surgery that were studied, the highest median time to surgery was for patients with coagulopathies at 30 hours (IQR: 20 – 48 hours), congestive heart failure at 28 hours (IQR: 19 – 48 hours), coronary artery disease at 28 hours (IQR: 19 – 47 hours), and total arthroplasty at 27 hours (IQR: 19 – 45).

Multivariate analysis for factors significantly associated with time to surgery was then performed (Table 1, last two data columns). All risk factors had a statistically significant association with time to surgery except for age. The greatest independent effect sizes were for total arthroplasty at 7.7 additional hours (95% C.I.: 6.1 – 9.3 hours), coagulopathy at 7.1 additional hours (95% C.I.: 6.1 – 8.0 hours), and congestive heart failure at 6.9 additional hours (95% C.I.: 6.0 – 7.9 hours).

Institutional cohort

The median time to surgery in the institutional cohort was 25 hours (IQR: 18 – 39 hours). The distribution of time to surgery can be seen in Figure 2, with the number of patients who would have been considered as having had appropriate early surgery as indicated by the 2014 AAOS Guideline (85%), the 2010 UK NHS Best Practice Tariffs (73%), and the 2015 Danish Hip Fracture Registry Study (11%).

Patient / surgical variable for this cohort are presented in Table 2., with the number of patients and median time to surgery indicated in the first two columns. Of all the risk factors for delay surgery that were studied, the highest median time to surgery was for patients with total arthroplasty at 42 hours (IQR: 24 – 68 hours), inpatient furosemide therapy for CHF exacerbation at 41 hours (IQR: 30 – 53 hours), warfarin

Table 1: Multivariate of risk factors for surgical delay in the national cohort

Risk factor	n (% of 44,900 patients)	Median time to surgery, in hours (interquartile range)	Multivariate regression coefficient, additional hours to surgery (95% confidence interval)	P-value
Gender				
Female	32,337 (72%)	23 (17 - 38)	Reference	-
Male	**12,563 (28%)**	24 (17 - 41)	**2.1 (1.4 - 2.7)**	**< 0.001**
Age				
65- 69	5,211 (12%)	23 (16 - 38)	Reference	-
70 - 74	6,161 (14%)	23 (16 - 38)	-0.7 (-1.8 - 0.5)	0.245
75 - 79	8,364 (19%)	24 (17 - 39)	0.3 (-0.7 - 1.4)	0.583
80 - 84	12,131 (27%)	24 (17 - 39)	0.3 (-0.8 - 1.3)	0.622
85+	13,033 (29%)	24 (17 - 39)	-0.3 (-1.3 - 0.7)	0.556
Comorbidities				
Coagulopathy	**4,848 (11%)**	**30 (20 - 48)**	**7.1 (6.1 - 8.0)**	**< 0.001**
Congestive heart failure	**4,824 (11%)**	**28 (19 - 48)**	**6.9 (6.0 - 7.9)**	**< 0.001**
Coronary artery disease	**1,605 (4%)**	**28 (19 - 47)**	**3.8 (2.2 - 5.4)**	**< 0.001**
Obesity	**1,730 (4%)**	**26 (18 - 45)**	**3.7 (2.2 - 5.2)**	**< 0.001**
Chronic respiratory disease	**6,732 (15%)**	**25 (18 - 42)**	**2.9 (2.0 - 3.7)**	**< 0.001**
Functionally dependent	**2,386 (5%)**	**25 (18 - 41)**	**2.2 (0.9 - 3.5)**	**0.001**
Diabetes mellitus	**9,787 (22%)**	**25 (17 - 42)**	**2.0 (1.2 - 2.7)**	**< 0.001**
Fracture type (*)				
Femoral neck	**7,290 (16%)**	**25 (18 - 41)**	**2.3 (1.3 - 3.3)**	**< 0.001**
Pertrochanteric	18,983 (42%)	23 (16 - 36)	Reference	-
Procedure type (*)				
CRIF	11,069 (25%)	22 (15 - 33)	Reference	-
ORIF	**19,030 (42%)**	**24 (17 - 40)**	**3.3 (2.6 - 4.1)**	**< 0.001**
Hemi arthroplasty	**12,904 (29%)**	**24 (17 - 40)**	**2.9 (2.0 - 3.8)**	**< 0.001**
Total arthroplasty	**1,897 (4%)**	**27 (19 - 45)**	**7.7 (6.1 - 9.3)**	**< 0.001**
Non-academic hospital	31,688 (71%)	23 (16 - 37)	Reference	-
Academic hospital	**13,212 (29%)**	**24 (17 - 42)**	**4.1 (3.4 - 4.7)**	**< 0.001**

ORIF = open reduction and internal fixation, CRIF = closed reduction and internal fixation
Asterisks (*) indicate data elements with incomplete classification to due nonspecific coding of fracture types or procedure types.

anticoagulation at 40 hours (IQR: 27 – 60 hours), and other chronic anticoagulation at 36 hours (IQR: 24 – 61).

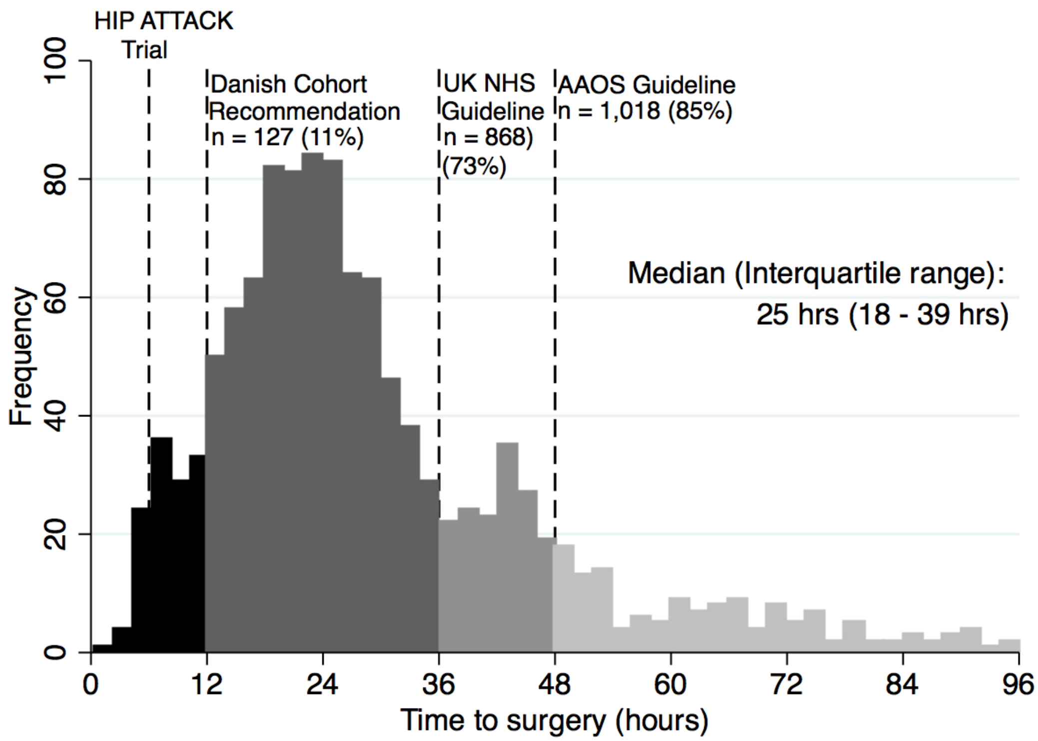

Figure 2: Distribution of surgical timing in the single-institution elderly hip fracture cohort (n = 1,196)

Multivariate analysis for factors significantly associated with time to surgery was again performed (Table 2, last two columns). Only 5 risk factors had a statistically significant association with time to surgery: total arthroplasty at 24.0 additional hours (95% C.I.: 13.2 – 34.8 hours), transfer from outside hospital at 22.0 additional hours (95% C.I.: 15.0 – 29.0 hours), warfarin anticoagulation at 13.4 additional hours (95% C.I.: 8.2 – 18.5 hours), other anticoagulation therapy at 10.0 additional hours (95% C.I.: 1.9 – 18.0 hours), and recent CHF exacerbation at 7.7 additional hours (95% C.I.: 1.8 – 13.6 hours).

Table 2: Multivariate of risk factors for surgical delay in the institutional cohort

Risk factor	n (% of 1,196 patients)	Mean time to surgery, in hours (standard deviation)	Multivariate regression coefficient, hours (95% confidence interval)	P-value
Gender				
Female	866 (72%)	25 (18 - 37)	Reference	-
Male	330 (28%)	26 (18 - 42)	0.9 (-2.8 - 4.5)	0.640
Medical risk factors				
Transfer from OSH	**66 (6%)**	**32 (17 - 60)**	**22.0 (15.0 - 29.0)**	**< 0.001**
Warfarin anticoagulation	**131 (11%)**	**40 (27 - 60)**	**13.4 (8.2 - 18.5)**	**< 0.001**
Other anticoagulation	**50 (4%)**	**36 (24 - 61)**	**10.0 (1.9 - 18.0)**	**0.015**
Recent CHF exacerbation	**99 (8%)**	**29 (23 - 43)**	**7.7 (1.8 - 13.6)**	**0.011**
Inpatient furosemide	16 (1%)	41 (30 - 53)	5.7 (-8.2 - 19.6)	0.421
Coronary artery disease	234 (20%)	28 (21 - 42)	3.0 (-1.2 - 7.2)	0.157
Hematocrit < 30	86 (7%)	28 (20 - 43)	2.3 (-3.9 - 8.5)	0.460
Obesity	25 (2%)	27 (21 - 41)	0.9 (-10.2 - 12.0)	0.879
Aspirin therapy	261 (22%)	26 (20 - 39)	0.7 (-3.2 - 4.6)	0.721
Fracture type (*)				
Femoral neck	450 (38%)	27 (20 - 42)	1.4 (-3.0 - 5.9)	0.529
Pertrochanteric	626 (52%)	24 (17 - 37)	Reference	-
Procedure type (*)				
CRIF	115 (10%)	24 (15 - 31)	Reference	-
ORIF	699 (58%)	24 (17 - 37)	4.6 (-1.0 - 10.2)	0.104
Hemi arthroplasty	160 (13%)	27 (20 - 39)	6.1 (-1.0 - 13.1)	0.095
Total arthroplasty	**35 (3%)**	**42 (24 - 68)**	**24.0 (13.2 - 34.8)**	**< 0.001**

OSH = outside hospital

CHF = congestive heart failure

CRIF = closed reduction and internal fixation
ORIF = open reduction and internal fixation
Asterisks (*) indicate data elements with incomplete classification to due nonspecific coding of fracture types or procedure types.

Discussion

Substantial existing evidence supports earlier surgery after elderly hip fractures [38-48, 52]. Current national guidelines in the United States and United Kingdom recommend surgery within 48 hours and 36 hours of admission, respectively [49-51]. In addition, a recent study using the Danish Hip Fracture Registry has shown surgery delayed later than 12 hours after admission was associated with increased 30-day mortality [52]. With more ongoing research evaluating potential benefits with even earlier surgery after hip fracture, we may very soon see recommendations for surgery within even shorter time intervals. These recommendations for early surgery after elderly hip fracture may even become tied to hospital or physician reimbursement as in the U.K. In order to eventually reduce the overall time to surgery, changes in staffing of orthopaedic surgeons will be a necessary step.

The current study aims to identify potentially modifiable factors that may allow us provide earlier surgery for hip fracture patients. First, patient factors were analyzed in a national cohort of patients to identify broad factors associated with surgical delays. Second, the findings from the national-level analysis were used to identify modifiable factors associated with delays in an institutional cohort. While a number of factors had significant associations with later surgery in the national cohort, focused analysis in the institutional cohort demonstrated that management of chronic anticoagulation, recent CHF exacerbations, patient's requiring arthroplasty surgery, and patients transferred from an OSH are all factors to be addressed in order to reduce overall time to surgery.

Limitations

The current study has a number of limitations. First, as with any study using a large clinical database, the quality of results is dependent on the quality of data contained in the database. Fortunately, NTDB does include specific data elements for various comorbidity diagnoses that have been shown to be more accurate than the ICD-9 billing codes utilized in other databases [55]. Nevertheless, the accuracy of data entry in NTDB is institution-dependent and there may be inconsistencies and systematic biases. For this reason, the current study takes the further step of verifying the results of the national cohort analysis using an institutional cohort and conducting a more focused analysis.

A second limitation of the current study is the inability to characterize all factors that may be related to delays in surgery. Individual cases of logistical, staffing, or clinical delays likely occur with most patients, and cannot be well delineated even with chart-review. As a result, this study attempts to take a bird's eye view and identify potentially modifiable factors that may have the biggest impact on overall surgical timing. Any large-scale attempt to reduce the overall time to surgery after trauma would likely benefit from standardized protocols that eliminate systemic logistical and staffing delays.

National Cohort

Based on the results of our national cohort analysis, gender, all medical comorbidities, type of fracture, type of surgery, and hospital type had statistically significant associations with surgical timing. The largest effect sizes were for total arthroplasty (7.7 hours), coagulopathy (7.1 hours) and congestive heart failure (6.9 hours), all of which we chose to analyze in greater detail in our institutional cohort (Figure 3).

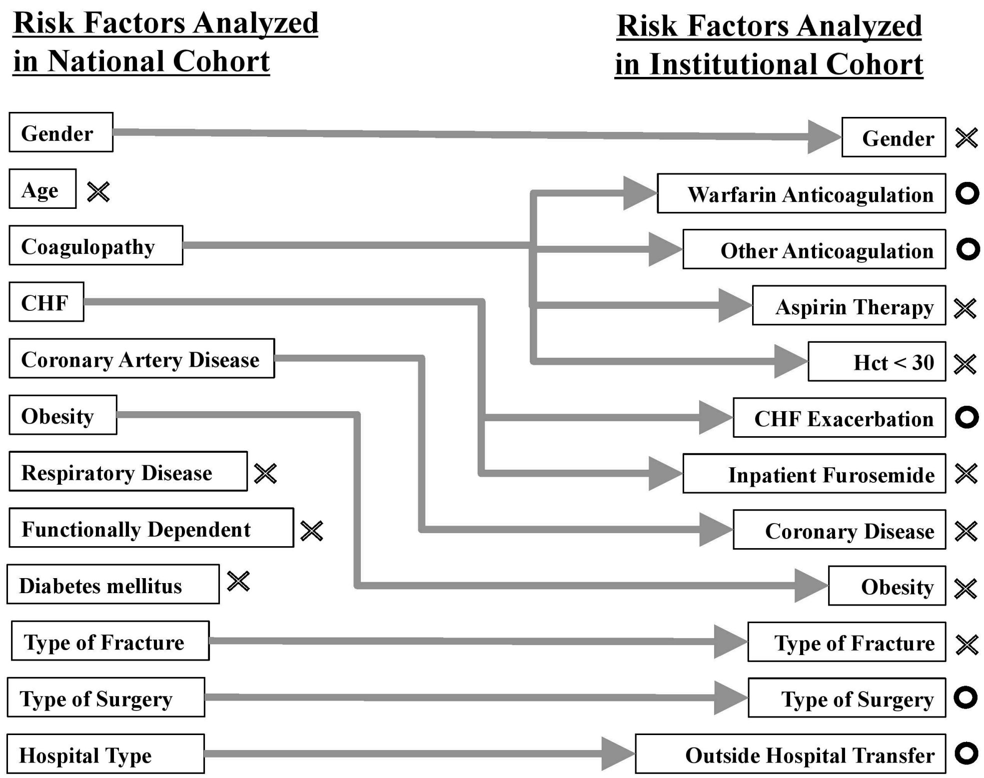

Figure 3: Diagram of step-wise selection of focused, modifiable risk factors for analysis between the national cohort and the institutional cohort. Circles indicate risk factors that were statistically significant (light green in the national cohort represents regression coefficient s< 3.0). X's indicate risk factors that were not statistically significant or did not have associations with an effect size that met a predetermined threshold. Note: CHF = congestive heart failure, Hct = hematocrit.

Coronary artery disease and obesity were two factors found to be associated with independent delays of over 3 hours in the national cohort. One previous analysis of over 600 patients at one level 1 trauma center found that patient requiring preoperative cardiac testing were delayed 1.5 more days prior to surgery, compared to those who are not [56], while longer time to surgery associated with obesity has not previously been shown. Both obesity and coronary disease are known to predispose patients to a number of preoperative risk factors for postoperative complications, risk factors that may necessitate preoperative cardiac, respiratory, endocrine, and volume status screening [57, 58]. While this screening is no doubt necessary, in order to reduce surgical timing after hip fractures in the future this screening also must be expedited for patients with body mass indices greater than determined thresholds.

Institutional Cohort

In the institutional cohort, total arthroplasty surgery was associated with the greatest delay (24.0 additional hours) and presumably was largely due to the lack of availability of specialized arthroplasty surgeons who may not be available for urgent surgery at time of patient admission. Additional risk factors that were identified were transfers from outside hospitals, warfarin anticoagulation, other chronic anticoagulation, and recent CHF exacerbation.

Based on these findings two modifiable aspects of hip fracture management may be appropriate aims for future investigation. First, timely availability of the correct surgeon and surgical staff is essential for early operative management, especially if hip fracture guidelines are revised with earlier goals. Longer time to hip fracture surgery for patients requiring total hip arthroplasty suggests that staffing of arthroplasty surgeons may be lacking compared to surgeons able to perform an ORIF of a hip fracture. Urgent total hip arthroplasty has previously been shown to be associated with higher total hospital costs and perioperative complications compared to elective arthroplasty [59]. Nevertheless, long-term outcomes after total hip arthroplasty for elderly hip fractures are good [60]. The ongoing Hip fracture Evaluation with ALternatives of THA versus Hemiarthroplasty (HEALTH) trial is investigating the superiority of hemi- versus total-arthroplasty after elderly hip fractures [61]. If the HEALTH trial does demonstrate better long-term outcomes with THA, then THA may become the new standard of care for elderly femoral neck fractures, further necessitating increased staffing of arthroplasty surgeons. However, if the benefits of THA over hemiarthroplasty are negligible then the large delays associated with THA may outweigh these benefits. This would then suggest that hemiarthroplasty, performed in a timely manner, be the treatment of choice. The growing concept of specialized hip fracture services, with expedited pathways to appropriate surgical treatment may be a suitable solution [62-64].

A second aspect of care that may be modified is medical co-managemnt of hip fracture patients. Both warfarin and alternative therapies for chronic anticoagulation were associated with large delays to surgery (13.4 and 10.0 additional hours, respectively). This accounted for 15% of our institutional cohort (177 patients). Reversal of chronic anticoagulation prior to surgery is clearly necessary to avoid excessive bleeding. However, with more pressure to operate on hip fracture earlier, more rapid medical strategies for reversal may be warranted. This may be optimally managed by an inpatient medical service that is co-managing hip fracture patients in concert with the surgical team. Previous studies of prothrombin complex concentrate (PCC) for reversal of vitamin K antagonist anticoagulation, have shown good safety and effectiveness when used prior to acute surgery [65], even when compared to fresh frozen plasma [66]. While this relatively new treatment is expensive, several analyses have demonstrated that use of PCC for rapid warfarin reversal after hip fractures is actually cost effective, due to reduce preoperative hospital time and shorter postoperative length of stay [67-69]. Adoption of rapid warfarin reversal protocols by medical teams may be a very effective method to reduce time to surgery for a large subset of hip fracture patients. For patients taking new oral anticoagulants (NOACs), PCC has also been shown to effectively reverse both the factor

Xa inhibitor rivaroxaban and the thrombin inhibitor dabigatran in one prospective randomized controlled trial [70]. Recent CHF exacerbations were also associated with a delay to surgery of 7.7 hours in the institutional cohort. This was largely due to need for preoperative medical screening and the rare circumstance of a patient presenting acute with decompensated CHF. Co-management by a medical team with standardized preoperative screening algorithms for patients with preexisting CHF may help to more rapidly identify patients who are too high-risk for surgical intervention.

Conclusion

In conclusion, the current study utilizes a large national cohort and an institutional cohort to identify modifiable risk factors for delayed hip fracture surgery. Based on the current findings, more aggressive reversal of chronic anticoagulation would be an effective mean to reduce delays in a large subset of the population. Standardized preoperative screening algorithms, which focus on identifying patients who are poor surgical candidates, would also help reduce preoperative delays. Additionally, if use of total hip arthroplasty to manage hip fractures increases, bases on current ongoing clinical trials, improved staffing of arthroplasty surgeons may also be warranted. Ultimately, with the growing body of research showing benefits with hip fracture fixation within 24 hours or even earlier, strategies for reducing our current preoperative interval are necessary. While standardized admission protocols and order sets will be necessary to reduce logistical and administrative delays, the current study identifies certain patient factors that providers should also be prepared to manage aggressively in order to achieve optimal surgical timing.

Section III

Variation in Resource Utilization for Hip and Pelvic Fracture Patients, Despite Equal Medicare Reimbursement: An example of using databases for hypothesis testing

Accepted as: Samuel AM, Webb ML, Lukasiewicz AM, Basques BA, Bohl DD, Lane JM, Grauer JN. "Differing resource utilization for hip and pelvic fracture patients, despite equal Medicare reimbursement." Clin Orthop Relat Res (Provisionally Accepted Dec 2015)

Background

The current Medicare Acute Inpatient Prospective Payment System determines hospital reimbursement for inpatient admissions by categorizing patients into Diagnosis-related Groups (DRGs), originally meant to create patient classes that are "clinically consistent and that have similar patterns of output utilization" [71]. When first implemented in 1983, the DRG model represented a substantial shift away from traditional cost-based billing, transferring both risk and potential cost-savings to hospitals [72]. Regardless of true hospital costs, all inpatient admissions under the same DRG are reimbursed using the same "bundled" payment, the only exception being a reimbursement modifier for either major or minor patient complications and/or comorbidities.

Although the DRG system has no doubt resulted in Medicare cost savings, proper division of patients into DRGs is critical to prevent under-reimbursement to hospitals or wasted Medicare payments. Under the current Medicare Severity (MS)-DRG system, MS-DRG 536 (fractures of the hip and pelvis) encompasses a large number of elderly orthopaedic trauma patients. While the age-adjusted incidence of elderly fractures has declined since 2005, the absolute incidence has increased dramatically due to the growing elderly population and is expected to continue to grow [73, 74]. The U.S. population older than age 85 is expected to increase 3-fold between 2010 and 2050 [73]. In addition, elderly fractures can be costly to manage, as they often require surgical intervention. However, the average low-energy hip fracture and high-energy pelvic fracture, both included in DRG 536, are considerably different injuries with differing management courses and, likely, differing associated inpatient costs [75]. Due to the common nature of these injuries, these differences in costs may account for substantial sums. Classification of other Medicare DRGs for procedures such as total joint arthroplasty, spinal fusion, and coronary artery bypass grafting has been modified to take into account cost variations [76, 77]. For example due to higher hospital costs associated with revision arthroplasty, versus primary arthroplasty, these procedures were separated into a separate DRG with greater reimbursement rates [76-78]. As the complexities of our current US healthcare environment make estimation of true healthcare costs difficult, several measures of hospital resource utilization were used as proxies for inpatient costs. In this way, differences in utilization can also be translated across health systems, regardless how cost accounting practices differ between individual institutions.

The purpose of this study was to determine whether (1) inpatient length of stay; (2) intensive care unit (ICU) stay; and (3) ventilator time differ between subpopulations of Medicare DRG 536, even after controlling for hospital factors.

Patients and Methods

A retrospective study was performed using the 2011 and 2012 American College of Surgeons National Trauma Data Bank Research Data Set. The National Trauma Data Bank is the largest national database of trauma cases including patients from over 900 trauma centers annually. While other commonly studied national databases, such as the National Surgical Quality Improvement Program database, do not include non-operative patients or high-energy trauma patients, the National Trauma Data Bank allows analysis of both operative and non-operative patients sustaining both high-energy and low-energy injuries in centers across the U.S, or the full spectrum of patients in DRG 536. In addition, there are currently a number of methodologies used by healthcare economists and hospital administrators to calculate costs, with no consensus between institutions or researchers. The National Trauma Data Bank offers data on various measure of healthcare resource utilization, namely inpatient length of stay, ICU stay, and mechanical ventilation time, which may be used as proxies for inpatient costs. In this way, differences in utilization can also be translated across health systems, regardless of the differing cost accounting practices of individual institutions Human investigations committee approval was acquired before initiating the study.

International Classification of Diseases, 9th Revision (ICD-9) diagnosis codes were used to identify patients in MS-DRG 536 based on diagnoses of acetabulum fractures, including both operative and non-operative, (ICD-9 808.0–808.1), other pelvic fractures (ICD-9 808.2–808.59), and hip fractures (ICD-9 820.20–820.32). Patients with other pelvic fractures were further subdivided into operative and non-operative pelvic fracture-based ICD-9 procedure codes (ICD-9 79.19, 79.29, 79.39, or 79.49). Note that patients managed operatively versus non-operatively, or with internal fixation versus arthroplasty, would similarly be grouped into MS-DRG 536, regardless of ultimate treatment modality. Only patients within the Medicare age group (65 years or older) were included.

A retrospective study was performed comparing inpatient length of stay, intensive care unit stay, and ventilator time among the four subpopulations of Medicare DRG 536. The three resource utilization measures of interest were reported as continuous variable in the National Trauma Data Bank. Multivariate analysis controlled for possible confounding patient and hospital factors including Charlson Comorbidity Index, hospital teaching status (university, community, or nonteaching), hospital size (≤ 200 beds, 201–400 beds, 401–600 beds, or > 600 beds), and geographic region (by US Census regions: Northeast, Midwest, South, and West). The modified Charlson Comorbidity Index has been shown to have comparable predictive value to the original Charlson Comorbidity Index and have been used previously in studies of national databases with national databases studies [7, 8, 32]. Computation of the modified Charlson Comorbidity Index was based on National Trauma Data Bank comorbidity and age data (Appendix 1; Supplemental materials are available with the online version of CORR®.). The hospital factors are reported in National Trauma Data Bank by participating institutions. Two additional possible confounding variables are month of admission and American College of Surgeons trauma center level. Previous studies have demonstrated that length of stay may vary according to a seasonal or "July" effect and by American College of Surgeons level. Unfortunately data on date or month of admission in unavailable in the database and data on American College of Surgeons level contains a large amount of missing data

(36%). However, as contributing trauma centers submit a full year of patient data to the National Trauma Data Bank, most seasonal variations can be expected to wash out. In addition, controlling for teaching status (non-teaching, community, or university) and hospital size should address much of the bias between American College of Surgeons levels.

Statistical Analysis

Multivariate linear regression was used to test the differences in outcome measures between subgroups after controlling for hospital factors. This was followed by Tukey's post hoc pairwise comparisons to test the significance of individual comparisons between the four subpopulations.

All statistical analyses were performed using Stata® Version 13.0 (StataCorp, LP, College Station, TX, USA). All statistical tests were two-tailed, and the level of significance was set at $\alpha = 0.05$. All analyses were adjusted for clustering by institution.

Results

Demographics, Description of Study Population

A total of 58,977 patients met inclusion criteria. Of those, 35,119 patients had hip fractures, 15,506 had nonoperative pelvic fractures, 7670 had acetabulum fractures, and 682 had operative pelvic fractures. Gender, age, Charlson Comorbidity Index, and Injury Severity Score varied predictably between groups (Table 1). Of note, 6,180 hip fracture patients (17.6% of all hip fractures) and 307 acetabular fracture patients (4.0% of all acetabular fractures) were treated with arthroplasty surgery (total hip arthroplasty or hemiarthroplasty).

Inpatient length of stay

Patients with operative pelvic fractures had the longest mean length of stay at 15.2 days (SD: 13.6 days; Figure 1). Mean length of stay for hip fracture patients was 6.6 days (SD: 5.1 days), and this difference from operative pelvic fracture patients was statistically significant after controlling for patient and hospital factors (difference: 7.7 days [95% confidence interval: 7.0 to 8.4]; $P < 0.001$; Table 2). Non-operative pelvic fracture patients had a mean length of stay of 6.5 days (SD: 8.2 days), and this difference from operative pelvic fracture patients was statistically significant after controlling for patient and hospital factors (difference: 8.0 days [95% confidence interval: 7.3 to 8.6]; $P < 0.001$). Acetabulum fracture patients had a mean length of stay of 8.7 days (SD: 9.5 days), and this difference from operative pelvic fracture patients was statistically significant after controlling for patient and hospital factors (difference: 6.0 days [95% confidence interval: 5.4 to 6.8]; $P < 0.001$).

Table 1. Patient summary statistics

Patient factors	Hip fracture (n = 35,119)		Nonoperative pelvic fracture (n = 15,506)		Acetabulum fracture (n = 7670)		Operative pelvic fracture (n = 682)	
Gender								
Female	24,046	68%	11,685	75%	3377	44%	252	37%
Male	11,073	32%	3821	25%	4293	56%	430	63%
Age (years)	(mean = 80)		(mean = 79)		(mean = 77)		(mean = 73)	
65-69	3702	10%	2002	13%	1804	24%	272	40%
70-74	4573	13%	2105	14%	1437	19%	150	22%
75-79	6472	18%	2773	18%	1352	18%	124	18%
80-84	9670	28%	4172	27%	1626	21%	84	12%
85+	10,702	30%	4454	29%	1451	19%	52	8%
Charlson Comorbidity Index	(mean = 5)		(mean = 5)		(mean = 4)		(mean = 4)	
3	2219	6%	1553	10%	1304	17%	223	33%
4	16,579	47%	8155	53%	3498	46%	303	44%
5	9265	26%	3524	23%	1739	23%	101	15%
6+	7056	20%	2274	15%	1129	15%	55	8%
Injury Severity Score	(mean = 10)		(mean = 11)		(mean = 12)		(mean = 20)	
0-9	29,769	85%	10,052	65%	4445	58%	133	20%
10-14	3778	11%	1996	13%	1099	14%	125	18%
15-19	637	2%	952	6%	636	8%	117	17%
20+	935	3%	2506	16%	1490	19%	307	45%

Note: Underline indicates median values.

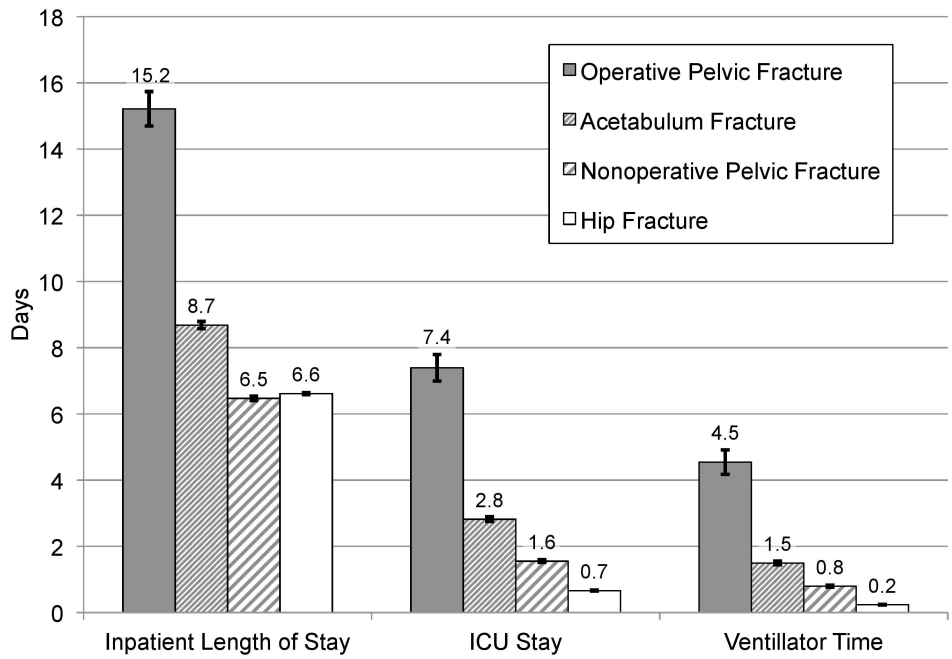

Figure 1: Mean length of stay, inpatient care unit (ICU) stay, and mechanical ventilation time vary between patients in each subgroup of DRG 536. Error bars represent standard error of the mean.

ICU stay

Patients with operative pelvic fractures had the longest mean ICU stay at 7.4 days (SD: 10.5 days; Figure 1). Mean ICU stay for hip fracture patients was 0.7 days (SD: 2.7 days), and this difference from operative pelvic fracture patients was statistically significant after controlling for patient and hospital factors (difference: 6.3 days [95% confidence interval: 5.9 to 6.7]; $P < 0.001$; Table 3). Non-operative pelvic fracture patients had a mean ICU stay of 1.6 days (SD: 4.7 days), and this difference from operative pelvic fracture patients was statistically significant after controlling for patient and hospital factors (difference: 5.4 days [95% confidence interval: 5.0 to 5.8]; P <

0.001). Acetabulum fracture patients had a mean ICU stay of 2.8 days (SD: 6.4 days), and this difference from operative pelvic fracture patients was statistically significant after controlling for patient and hospital factors (difference: 4.3 days [95% confidence interval: 3.9 to 4.8]; P < 0.001).

Ventilator Time

Patients with operative pelvic fractures had the longest mean ventilator time at 4.5 days (SD: 9.7 days; Figure 1). Mean ventilator time for hip fracture patients was 0.2 days (SD: 1.9 days), and this difference from operative pelvic fracture patients was statistically significant after controlling for patient and hospital factors (difference: 3.9 days [95% confidence interval: 3.6 to 4.2]; P < 0.001; Table 4). Non-operative pelvic fracture patients had a mean ventilator time of 0.8 days (SD: 3.9 days), and this difference from operative pelvic fracture patients was statistically significant after controlling for patient and hospital factors (difference: 3.4 days [95% confidence interval: 3.1 to 3.7]; P < 0.001). Acetabulum fracture patients had a mean ventilator time of 1.5 days (SD: 5.1 days), and this difference from operative pelvic fracture patients was statistically significant after controlling for patient and hospital factors (difference: 2.8 days [95% confidence interval: 2.5 to 3.2]; P < 0.001).

Table 2. Multivariate analysis of length of stay

Outcome: inpatient length of stay			
Subpopulations of DRG 536	Mean length of stay, in days (standard deviation)	Difference from operative pelvic fractures, in days (95% confidence interval)	Multivariate P-value
Hip fracture	6.6 (5.1)	7.7 (7.0-8.4)	< 0.001
Non-operative pelvic fracture	6.5 (8.2)	8.0 (7.3-8.6)	< 0.001
Acetabulum fracture	8.7 (9.5)	6.0 (5.4-6.8)	< 0.001
Operative pelvic fracture	15.2 (13.6)	Reference	-
Patient and hospital factors (controlled for in multivariate analysis)		Regression coefficient, in days (95% confidence interval)	Multivariate P-value
Charlson Comorbidity Index			
3		Reference	
4		-0.5 (-0.7 to -0.3)	< 0.001
5		-0.1 (-0.4 to 0.1)	0.190
6+		0.4 (0.2-0.6)	0.001
Hospital teaching status			
Nonteaching		Reference	
University		1.4 (1.3-1.6)	< 0.001
Community		0.4 (0.2-0.5)	< 0.001
Hospital size (beds)			
≤ 200		Reference	
201-400		0.5 (0.3-0.7)	< 0.001
401-600		1.0 (0.8-1.2)	< 0.001
> 600		1.4 (1.2-1.7)	< 0.001
Geographic region			
Northeast		Reference	
Midwest		0.39 (0.2-0.6)	< 0.001
South		0.49 (0.4-0.6)	< 0.001
West		0.21 (0.0-0.4)	0.016

Table 3. Multivariate analysis of ICU length of stay

Outcome: intensive care unit days			
Subpopulations of DRG 536	Mean intensive care unit stay, in days (standard deviation)	Difference from operative pelvic fractures, in days (95% confidence interval)	Multivariate P-value
Hip fracture	0.7 (2.7)	6.3 (5.9-6.7)	< 0.001
Non-operative pelvic fracture	1.6 (4.7)	5.4 (5.0-5.8)	< 0.001
Acetabulum fracture	2.8 (6.4)	4.3 (3.9-4.8)	< 0.001
Operative pelvic fracture	7.4 (10.5)	Reference	-
Patient and hospital factors (controlled for in multivariate analysis)		Regression coefficient, in days (95% confidence interval)	Multivariate P-value
Charlson Comorbidity Index			
3		Reference	
4		-0.4 (-0.6 to -0.3)	< 0.001
5		-0.4 (-0.5 to -0.2)	< 0.001
6+		-0.3 (-0.4 to -0.3)	< 0.001
Hospital teaching status			
Nonteaching		Reference	
University		0.9 (0.8-1.0)	< 0.001
Community		0.1 (0.0-0.2)	0.012
Hospital size (beds)			
≤ 200		Reference	
201-400		0.2 (0.0-0.3)	0.005
401-600		0.4 (0.2-0.5)	< 0.001
> 600		0.7 (0.5-0.8)	< 0.001
Geographic region			
Northeast		Reference	
Midwest		-0.3 (-0.4 to -0.2)	< 0.001
South		0.1 (0.02-0.2)	0.017
West		0.3 (0.2-0.4)	< 0.001

ICU = intensive care unit.

Table 4. Multivariate analysis of ventilator days

Outcome: mechanical ventilation days			
Subpopulations of DRG 536	Mean mechanical ventilation time, in days (standard deviation)	Difference from operative pelvic fractures, in days (95% confidence interval)	Multivariate P-value
Hip fracture	0.2 (1.9)	3.9 (3.6-4.2)	< 0.001
Non-operative pelvic fracture	0.8 (3.9)	3.4 (3.1-3.7)	< 0.001
Acetabulum fracture	1.5 (5.1)	2.8 (2.5-3.2)	< 0.001
Operative pelvic fracture	4.5 (9.7)	Reference	-
Patient and hospital factors (controlled for in multivariate analysis)		Regression coefficient, in days (95% confidence interval)	Multivariate P-value
Charlson Comorbidity Index			
3		Reference	
4		-0.3 (-0.4 to -0.2)	< 0.001
5		-0.3 (-0.4 to -0.2)	< 0.001
6+		-0.3 (-0.4 to -0.2)	< 0.001
Hospital teaching status			
Nonteaching		Reference	
University		0.6 (0.5-0.7)	< 0.001
Community		0.1 (0.1-0.2)	< 0.001
Hospital size (beds)			
≤ 200		Reference	
201-400		0.0 (0.0 to 0.1)	0.346
401-600		0.2 (0.0-0.2)	0.004
> 600		0.4 (0.2-0.5)	< 0.001
Geographic region			
Northeast		Reference	
Midwest		-0.2 (-0.2 to -0.1)	< 0.001
South		0.1 (0.0-0.2)	0.002
West		0.1 (0.1-0.2)	0.001

Discussion

The current Medicare Acute Inpatient Prospective Payment System determines reimbursement for inpatient admissions by categorizing patients into 746 MS-DRGs that are annually revised. Inpatient admissions under the same DRG are reimbursed with the same prospective, "bundled" payment, regardless of actual hospital costs. Currently, DRG 536 (fractures of the hip and pelvis) includes a broad spectrum of orthopaedic injuries, from low-energy fragility fractures of the hip to high-energy acetabulum and pelvic fractures. With the absolute incidence of elderly fractures expected to increase dramatically due to the growing elderly population [73, 74] and the high costs of managing these injuries,, any substantial variation in costs may result in over- or under-reimbursement to certain hospitals, based on their percentage of high-energy versus low-energy elderly fractures treated. Excess financial strain on these hospitals treating high-energy trauma patients may even result trauma center closures or shifts away from managing high-energy trauma, resulting in reduced access for these patients. Similar discrepancies between hospital costs and Medicare reimbursements have also been seen within other Medicare DRGs in the past [76, 77]. The results of the current study demonstrate that, even after controlling for patient and hospital factors, there were important differences in various measures of resource utilization between the four subgroups of Medicare DRG 536. These differences indicate that hospitals are currently receiving under- or over-reimbursement for certain subgroups, depending on their baseline reimbursement for DRG 536. In particular, these increases in resource utilization were most striking for patients with the fractures more likely to be associated with higher energy mechanisms: operative pelvic fractures and acetabulum fractures.

The primary limitation of this study stems from this exclusion of true cost figures. Standard methods of healthcare cost estimation have important limitations. Cost measurement using charges or cost-to-charge ratios estimates only total costs, and not the marginal costs of services, and ignores cost shifting that occurs regularly in hospital accounting [79, 80]. However, the current study uses various measures of resource utilization as a proxy for costs. With the estimated costs of an inpatient day at USD 1236, an ICU day at USD 2278 to USD 5973, and a day on mechanical ventilation at USD 10,299 [81, 82], the results presented in the current study do represent significant differences between subpopulations of DRG 536. An additional limitation of this study is the inability to identify patients with major complications or comorbidities (MCCs), who might be classified into DRG 535 (fractures of the hip or pelvis with MCCs) based on the ICD-9 diagnosis data given in the National Trauma Data Bank. The authors acknowledge that inability to identify patients with major complications or comorbidities may result in overestimation of resource utilization for the subgroups of DRG 536 that are associated with higher rates of major comorbidities and complications. Higher-energy fractures, associated with greater total burden of bodily injury, may very likely also be associated with greater inpatient complications due to more complex inpatient course and longer length of stay. In addition, lower energy fractures associated with osteoporosis may also be associated with higher rates of comorbidities and complications, due to poorer overall health status. However, by controlling for Charlson Comorbidity Index, the current study shows differences in resource utilization that are independent of age and comorbidities. These differences likely indicate disparities in resource utilization for patients in both DRG 535 and DRG 536. Nevertheless, further study using true cost figures and DGR classifications at a single-institution is warranted to estimate the actual financial burden caused by the large spectrum of patients included in DRG 536.

Another limitation of the current study results from the inherent biases associated with using a national dataset such as the National Trauma Data Bank. While the NTDB includes data from over 900 trauma centers, this data is self-reported by hospitals and accuracy is dependent on the accuracy of data submitted from individual hospitals. Hospital may submit incomplete datasets resulting in over- or under representation of certain fracture types. Nevertheless, the large overall size of the dataset allowed statistically significant differences to be identified, despite large disparities in sample size between subpopulations (there were far more hip fractures than operative pelvic fractures). Selection bias is also possible as the dataset, while including data from over 900 U.S. trauma centers, is not necessarily nationally representative. There may certainly be biases toward level I and II trauma centers which see greater numbers of high-energy trauma patients. Nevertheless, the aim of the current study was not to measure the relative incidences of the different subpopulations included in MS-DRG 536, but rather to compare how resource utilization varies between patients with these different subpopulations. By controlling for hospital factors, such as hospital size, teaching status, and geographic region, the current study attempts to control for local variations in resource utilization and isolate the differences in utilization attributable to the various injury types. Another type of bias that may be present is transfer bias, as a percentage of patients in the National Trauma Data Bank have incomplete data regarding inpatient length of stay, ICU stay, or ventilator time. These patients were excluded from the analysis. Of a total of 69,055 patients in MS-DRG 536 and aged 65 or older in the

National Trauma Data Bank, 10,078 patients were dropped due to incomplete data in regards to inpatient length of stay (n = 478), ICU stay (n = 7,484), or ventilator time (n = 8,534) resulting in the 58,977 patients ultimately analyzed. It may be that patients with more severe acute injuries and complicated inpatient courses represent a higher percentage of these patients with incomplete data, biasing the current study results. However, in a post hoc analysis the mean Injury Severity Score for these dropped patients was 9.9, compared to 10.6 in the analyzed patients. This difference of 0.7 in Injury Severity Score, while statistically significant ($P < 000.1$) due to large sample size, was not clinically significant.

Previous work in the field of orthopaedics showed similar differences in inpatient costs and resource utilization between primary and revision total joint arthroplasty [77, 78], at the time both grouped under Medicare DRG 209 (lower extremity arthroplasty). An initial multicenter cohort study demonstrated differences in operative time, use of allograft, length of stay, and costs, between almost 8,000 patients undergoing either primary or revision hip and knee arthroplasty [78]. A subsequent study at a single institution demonstrated mean differences in hospital costs of over $7,000 per patient between revision and primary arthroplasty patients [77]. This work ultimately supported efforts by the American Association of Orthopaedic Surgeons, the American Association of Hip and Knee Surgeons, and the Hip Society to propose modification of the DRG to the Centers for Medicare & Medicaid Services (CMS). This was ultimately accepted in 2005 with creation of two new DRGs: DRG 544 (primary hip and knee replacement) and DRG 545 (revision hip and knee replacement) [76]. Similar efforts with DRG 536 should be pursued, next with analysis of true cost figures. This should be conducted within a single institution because comparing hospital costs between multiple institutions becomes complicated due to varying cost accounting practices. In addition, further investigation of other MS-DRGs that include a broad spectrum of orthopaedic patients may be warranted. One such example is DRG 563, which includes all fractures, sprains, strains, and dislocations, excluding those of the femur, hip, and pelvis.

A secondary finding of the current study was the associations of patient and hospital factors with resource utilization. As expected, increased Charleston Comorbidity Index is associated with increased costs, likely due to increased resources necessary to treat patients with worse overall health status. In addition, it was found that larger hospitals and University hospitals were associated with increased utilization. This would be expected as more critically injured patients are often seen in larger, university-associated trauma centers [83]. Previous studies have demonstrated worse outcomes for hip fractures managed at teaching hospitals compared to non-teaching hospitals, suggesting more severely injured patients. In addition, a previous study specifically demonstrated increased resource utilization in university hospitals compared to community teaching hospitals for tPA treated stroke patients [84]. The geographic differences observed have not been previously described but likely results from regional variations in hospital practices and physician customs for inpatient discharge and ICU transfer.

Now may be an opportune time for CMS to reevaluate classification of DRG 536 because the Medicare DRG system may soon see more-expanded use. As part of the Patient Protection and Affordable Care Act (PPACA), the CMS Bundled Payments for Care Improvement Initiative will test new bundled payment delivery models aimed at reducing costs and improving quality of care [85]. Three of the four proposed models, all currently being tested, expand bundling of payments to include services outside of inpatient hospital care including physician services and postacute care services. All four models use the current MS-DRG system to categorize patients. The American Association of Hip and Knee Surgeons has already created a Bundled Payment Task Force to identify issues in the system specific to hip and knee surgeons and to aid in the eventual transition [86, 87]. In addition, many private health plans, which cover patients of all ages, use a MS-DRG based model for reimbursement. However, as low-energy fragility fractures are less common in younger patients we would expect less variation in resource utilization between patients with hip or pelvis fractures, which would all more likely result from higher energy mechanisms. Similar measures should be taken throughout the field of orthopaedic surgery, including in the field of orthopaedic trauma, which sees over two million elderly patients with fragility fractures annually [88].

In conclusion, the current study demonstrates considerable variation in resource utilization for hospital admissions due to hip or pelvis fractures, which are all reimbursed equally under DRG 536. With the current multitiered US trauma system, it is very likely that some trauma centers treat a higher proportion of the more the resource-intensive injuries, resulting in higher costs and lower financial margins. Appropriate reimbursement of hospital services is essential for maintaining patient access to inpatient care in the currently shifting US healthcare landscape. Therefore, creating a separate MS-DRG for operative pelvic fractures and acetabulum fractures would likely help reduce this financial strain. Further comparison of true cost figures at a single institution may be a suitable next step to demonstrate clear financial benefits of regulatory changes. Indeed, just as new healthcare standards and quality benchmarks are currently being developed in the field of orthopaedics to meet PPACA guidelines, the classification of Medicare DRGs for orthopaedic diagnoses such as hip and pelvic fractures also presently warrant reevaluation.

Conclusion of Thesis

The preceding three studies illustrate the utility of "big data" for studying orthopaedic trauma patients, a unique patient population that may require special considerations when designing clinical analyses. In Section I, we demonstrated that not all national databases capture the typical high-energy fracture populations well, including the commonly used NSQIP database that primarily includes elderly fragility type fracture patients. As a result, fully understanding the populations contained in various national databases is critical before designing or evaluating database research. Trauma specific databases such as NTDB are best suited to study high-energy trauma patients, however they are not as nationally representative of fracture patients as databases with systematic sampling techniques, such as NIS.

In Section II the NTDB was utilized to identify a number of risk factors associated with delayed hip fracture surgery, including coagulopathy, congestive heart failure, and arthroplasty surgery. These findings were then used to conduct a more in depth analysis using a local cohort of hip fracture patients. It was found that total hip arthroplasty was associated with the greatest delays of over 24 hours, while warfarin anticoagulation, other chronic anticoagulation, and recent CHF exacerbations were additional risk factors. The importance of appropriate surgeon staffing and medical co-management of patients were highlighted as possible areas for improvement with increasing evidence and guidance for earlier surgical intervention after hip fractures.

Finally in Section III the NTDB was again used to demonstrate that management of both operative pelvic fractures and acetabulum fractures require substantially greater hospital resource utilization, compared to hip fractures or non-operative pelvic fractures.

This finding is interesting as hospitals are reimbursed equally for treatment of each of these types of fractures under Medicare DRG 536. For hospitals that manage a relatively greater number of higher energy injuries, such as acetabulum fractures or operative pelvic fractures, financial returns per patient would be lower relative to other hospitals. As a result, refinement of this Medicare DRG to separate these higher cost injuries was recommended, just as has been done for primary and revision arthroplasty patients, who were also previously grouped into the same Medicare DRG.

In conclusion, thoughtfully designed clinical studies using “big data” will continue to be used to study orthopaedic trauma populations, and these studies will be both hypothesis-finding and hypothesis-testing. While there remain limitations to national databases that are currently in use, many of these limitations can be addressed with thoughtful design of orthopaedic trauma-focused databases in the future. A hip fracture-specific NSQIP dataset, currently in data collection phase, is a first step, while development of robust institutional orthopaedic trauma registries at busy trauma centers, which may even leverage electronic health records for automated data collection, may provide a wealth of clinical data for future clinical study.

References

1. 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 Aug 5;97(15):1278-87.

2. Levin PE. The impact of the Spine Patient Outcome Research Trial (SPORT) results on orthopaedic practice. J Am Acad Orthop Surg. 2012 Jun;20(6):331; author reply -2.

3. Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. N Engl J Med. 2002 Jul 11;347(2):81-8.

4. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. JAMA. 2006 Nov 22;296(20):2441-50.

5. Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. JAMA. 2006 Nov 22;296(20):2451-9.

6. 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 Nov 4;97(21):1799-808.

7. Bohl DD, Basques BA, Golinvaux NS, Miller CP, Baumgaertner MR, Grauer JN. Extramedullary compared with intramedullary implants for intertrochanteric hip

fractures: thirty-day outcomes of 4432 procedures from the ACS NSQIP database. J Bone Joint Surg Am. 2014 Nov 19;96(22):1871-7.

8. Bohl DD, Fu MC, Golinvaux NS, Basques BA, Gruskay JA, Grauer JN. The "July Effect" in Primary Total Hip and Knee Arthroplasty: Analysis of 21,434 Cases From the ACS-NSQIP Database. J Arthroplasty. 2014 Jul;29(7):1332-8. Epub 2014/03/19.

9. Helwani MA, Avidan MS, Ben Abdallah A, Kaiser DJ, Clohisy JC, Hall BL, et al. Effects of regional versus general anesthesia on outcomes after total hip arthroplasty: a retrospective propensity-matched cohort study. J Bone Joint Surg Am. 2015 Feb 4;97(3):186-93.

10. Kester BS, Merkow RP, Ju MH, Peabody TD, Bentrem DJ, Ko CY, et al. Effect of post-discharge venous thromboembolism on hospital quality comparisons following hip and knee arthroplasty. J Bone Joint Surg Am. 2014 Sep 3;96(17):1476-84.

11. Pugely AJ, Martin CT, Gao Y, Mendoza-Lattes S, Callaghan JJ. Differences in short-term complications between spinal and general anesthesia for primary total knee arthroplasty. J Bone Joint Surg Am. 2013 Feb 6;95(3):193-9.

12. Nwachukwu BU, McLawhorn AS, Simon MS, Hamid KS, Demetracopoulos CA, Deland JT, et al. Management of End-Stage Ankle Arthritis: Cost-Utility Analysis Using Direct and Indirect Costs. J Bone Joint Surg Am. 2015 Jul 15;97(14):1159-72.

13. Odum SM, Springer BD. In-Hospital Complication Rates and Associated Factors After Simultaneous Bilateral Versus Unilateral Total Knee Arthroplasty. J Bone Joint Surg Am. 2014 Jul 2;96(13):1058-65.

14. Basques BA, Anandasivam NS, Webb ML, Samuel AM, Lukasiewicz AM, Bohl DD, et al. Risk Factors for Blood Transfusion with Primary Posterior Lumbar Fusion. Spine (Phila Pa 1976). 2015 Jul 10.

15. Agency for Healthcare Research and Quality Healthcare Cost and Utilization Project (HCUP). Introduction to the HCUP Nationwide Inpatient Sample. 2013.

16. American College of Surgeons National Surgical Quality Improvement Program. User Guide for the 2012 ACS NSQIP Participant Use Data. 2013.

17. American College of Surgeons National Trauma Data Bank. NTDB Research Data Set Admission Year 2012 User Manual. 2013.

18. Bohl DD, Russo GS, Basques BA, Golinvaux NS, Fu MC, Long WD, 3rd, et al. 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. The Journal of bone and joint surgery American volume. 2014 Dec 3;96(23):e193. Epub 2014/12/05.

19. 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. Clinical orthopaedics and related research. 2014 Jun;472(6):1672-80. Epub 2014/03/13.

20. Enomoto LM, Hollenbeak CS, Bhayani NH, Dillon PW, Gusani NJ. Measuring surgical quality: a national clinical registry versus administrative claims data. Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract. 2014 Aug;18(8):1416-22. Epub 2014/06/15.

21. Lapar DJ, Stukenborg GJ, Lau CL, Jones DR, Kozower BD. Differences in reported esophageal cancer resection outcomes between national clinical and administrative databases. The Journal of thoracic and cardiovascular surgery. 2012 Nov;144(5):1152-7. Epub 2012/09/04.

22. Kukreja S, Kalakoti P, Murray R, Nixon M, Missios S, Guthikonda B, et al. National trends of incidence, treatment, and hospital charges of isolated C-2 fractures in three different age groups. Neurosurgical focus. 2015 Apr;38(4):E19. Epub 2015/04/02.

23. Menendez ME, Ring D. Does the timing of surgery for proximal humeral fracture affect inpatient outcomes? Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al]. 2014 Sep;23(9):1257-62. Epub 2014/06/14.

24. Ryan DJ, Yoshihara H, Yoneoka D, Egol KA, Zuckerman JD. Delay in hip fracture surgery: an analysis of patient- and hospital-specific risk factors. Journal of orthopaedic trauma. 2015 Feb 19. Epub 2015/02/26.

25. Belmont PJ, Jr., Garcia EJ, Romano D, Bader JO, Nelson KJ, Schoenfeld AJ. Risk factors for complications and in-hospital mortality following hip fractures: a study using the National Trauma Data Bank. Archives of orthopaedic and trauma surgery. 2014 May;134(5):597-604. Epub 2014/02/27.

26. Bolorunduro OB, Haider AH, Oyetunji TA, Khoury A, Cubangbang M, Haut ER, et al. Disparities in trauma care: are fewer diagnostic tests conducted for uninsured patients with pelvic fracture? American journal of surgery. 2013 Apr;205(4):365-70. Epub 2013/02/05.

27. Cantu RV, Graves SC, Spratt KF. In-hospital mortality from femoral shaft fracture depends on the initial delay to fracture fixation and Injury Severity Score: a

retrospective cohort study from the NTDB 2002-2006. The journal of trauma and acute care surgery. 2014 Jun;76(6):1433-40. Epub 2014/05/24.

28. Basques BA, Webb ML, Bohl DD, Golinvaux NS, Grauer JN. Adverse Events, Length of Stay, and Readmission Following Surgery for Tibial Plateau Fractures. Journal of orthopaedic trauma. 2014 Aug 26. Epub 2014/08/28.

29. Molina CS, Thakore RV, Blumer A, Obremskey WT, Sethi MK. Use of the National Surgical Quality Improvement Program in Orthopaedic Surgery. Clinical orthopaedics and related research. 2014 Apr 5. Epub 2014/04/08.

30. Sathiyakumar V, Molina CS, Thakore RV, Obremskey WT, Sethi MK. ASA Score as a Predictor of 30-Day Perioperative Readmission in Patients with Orthopaedic Trauma Injuries: A NSQIP Analysis. Journal of orthopaedic trauma. 2014 Jul 25. Epub 2014/07/30.

31. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987;40(5):373-83. Epub 1987/01/01.

32. D'Hoore W, Bouckaert A, Tilquin C. Practical considerations on the use of the Charlson comorbidity index with administrative data bases. Journal of clinical epidemiology. 1996 Dec;49(12):1429-33. Epub 1996/12/01.

33. Faciszewski T, Broste SK, Fardon D. Quality of data regarding diagnoses of spinal disorders in administrative databases. A multicenter study. The Journal of bone and joint surgery American volume. 1997 Oct;79(10):1481-8. Epub 1997/10/31.

34. Faciszewski T, Jensen R, Berg RL. Procedural coding of spinal surgeries (CPT-4 versus ICD-9-CM) and decisions regarding standards: a multicenter study. Spine. 2003 Mar 1;28(5):502-7. Epub 2003/03/05.

35. Golinvaux NS, Bohl DD, Basques BA, Fu MC, Gardner EC, Grauer JN. Limitations of administrative databases in spine research: a study in obesity. The spine journal : official journal of the North American Spine Society. 2014 Apr 26. Epub 2014/05/02.

36. Bohl DD, Grauer JN, Leopold SS. Editor's spotlight/Take 5: nationwide inpatient sample and national surgical quality improvement program give different results in hip fracture studies. Clinical orthopaedics and related research. 2014 Jun;472(6):1667-71. Epub 2014/04/09.

37. Chapter 2: The Data Collection Process. ACS NSQIP: American College of Surgeons; 2014.

38. Elliott J, Beringer T, Kee F, Marsh D, Willis C, Stevenson M. Predicting survival after treatment for fracture of the proximal femur and the effect of delays to surgery. Journal of clinical epidemiology. 2003 Aug;56(8):788-95. Epub 2003/09/05.

39. Fox HJ, Pooler J, Prothero D, Bannister GC. Factors affecting the outcome after proximal femoral fractures. Injury. 1994 Jul;25(5):297-300. Epub 1994/07/01.

40. McGuire KJ, Bernstein J, Polsky D, Silber JH. The 2004 Marshall Urist award: delays until surgery after hip fracture increases mortality. Clinical orthopaedics and related research. 2004 Nov;428:294-301. Epub 2004/11/10.

41. Novack V, Jotkowitz A, Etzion O, Porath A. Does delay in surgery after hip fracture lead to worse outcomes? A multicenter survey. International journal for quality

in health care : journal of the International Society for Quality in Health Care / ISQua. 2007 Jun;19(3):170-6. Epub 2007/02/21.

42. Parker MJ, Pryor GA. The timing of surgery for proximal femoral fractures. The Journal of bone and joint surgery British volume. 1992 Mar;74(2):203-5. Epub 1992/03/01.

43. Radcliff TA, Henderson WG, Stoner TJ, Khuri SF, Dohm M, Hutt E. Patient risk factors, operative care, and outcomes among older community-dwelling male veterans with hip fracture. The Journal of bone and joint surgery American volume. 2008 Jan;90(1):34-42. Epub 2008/01/04.

44. Siegmeth AW, Gurusamy K, Parker MJ. Delay to surgery prolongs hospital stay in patients with fractures of the proximal femur. The Journal of bone and joint surgery British volume. 2005 Aug;87(8):1123-6. Epub 2005/07/29.

45. Moja L, Piatti A, Pecoraro V, Ricci C, Virgili G, Salanti G, et al. Timing matters in hip fracture surgery: patients operated within 48 hours have better outcomes. A meta-analysis and meta-regression of over 190,000 patients. PloS one. 2012;7(10):e46175. Epub 2012/10/12.

46. Orosz GM, Magaziner J, Hannan EL, Morrison RS, Koval K, Gilbert M, et al. Association of timing of surgery for hip fracture and patient outcomes. Jama. 2004 Apr 14;291(14):1738-43. Epub 2004/04/15.

47. Al-Ani AN, Samuelsson B, Tidermark J, Norling A, Ekstrom W, Cederholm T, et al. Early operation on patients with a hip fracture improved the ability to return to independent living. A prospective study of 850 patients. The Journal of bone and joint surgery American volume. 2008 Jul;90(7):1436-42. Epub 2008/07/03.

48. Smektala R, Endres HG, Dasch B, Maier C, Trampisch HJ, Bonnaire F, et al. The effect of time-to-surgery on outcome in elderly patients with proximal femoral fractures. BMC musculoskeletal disorders. 2008;9(1):171. Epub 2008/12/31.

49. Uzoigwe CE, Burnand HG, Cheesman CL, Aghedo DO, Faizi M, Middleton RG. Early and ultra-early surgery in hip fracture patients improves survival. Injury. 2013 Jun;44(6):726-9. Epub 2012/09/27.

50. Khan SK, Weusten A, Bonczek S, Tate A, Port A. The Best Practice Tariff helps improve management of neck of femur fractures: a completed audit loop. British journal of hospital medicine. 2013 Nov;74(11):644-7.

51. AAOS. Management of Hip Fractures in the Elderly. 2014.

52. Nyholm AM, Gromov K, Palm H, Brix M, Kallemose T, Troelsen A, et al. Time to Surgery Is Associated with Thirty-Day and Ninety-Day Mortality After Proximal Femoral Fracture: A Retrospective Observational Study on Prospectively Collected Data from the Danish Fracture Database Collaborators. The Journal of bone and joint surgery American volume. 2015 Aug 19;97(16):1333-9.

53. Hip Fracture Accelerated Surgical T, Care Track I. Accelerated care versus standard care among patients with hip fracture: the HIP ATTACK pilot trial. CMAJ. 2014 Jan 7;186(1):E52-60.

54. Ryan DJ, Yoshihara H, Yoneoka D, Egol KA, Zuckerman JD. Delay in hip fracture surgery: an analysis of patient- and hospital-specific risk factors. Journal of orthopaedic trauma. 2015 Feb 19;29(8):343-8. Epub 2015/02/26.

55. Samuel AM, Lukasiewicz AM, Webb ML, Bohl DD, Basques BA, Davis KA, et al. ICD-9 diagnosis codes have poor sensitivity for identification of preexisting

comorbidities in traumatic fracture patients: A study of the National Trauma Data Bank. J Trauma Acute Care Surg. 2015 Oct;79(4):622-30.

56. Ricci WM, Brandt A, McAndrew C, Gardner MJ. Factors affecting delay to surgery and length of stay for patients with hip fracture. Journal of orthopaedic trauma. 2015 Mar;29(3):e109-14.

57. Poso T, Kesek D, Aroch R, Winso O. Morbid obesity and optimization of preoperative fluid therapy. Obes Surg. 2013 Nov;23(11):1799-805.

58. Ortiz VE, Kwo J. Obesity: physiologic changes and implications for preoperative management. BMC Anesthesiol. 2015;15:97.

59. Kamath AF, Austin DC, Derman PB, Israelite CL. Unplanned hip arthroplasty imposes clinical and cost burdens on treating institutions. Clinical orthopaedics and related research. 2013 Dec;471(12):4012-9.

60. Godoy Monzon D, Iserson KV, Jauregui J, Musso C, Piccaluga F, Buttaro M. Total hip arthroplasty for hip fractures: 5-year follow-up of functional outcomes in the oldest independent old and very old patients. Geriatric orthopaedic surgery & rehabilitation. 2014 Mar;5(1):3-8.

61. Bhandari M, Devereaux PJ, Einhorn TA, Thabane L, Schemitsch EH, Koval KJ, et al. Hip fracture evaluation with alternatives of total hip arthroplasty versus hemiarthroplasty (HEALTH): protocol for a multicentre randomised trial. BMJ Open. 2015;5(2):e006263.

62. Egol KA, Marcano AI, Lewis L, Tejwani NC, McLaurin TM, Davidovitch RI. Can the use of an evidence-based algorithm for the treatment of intertrochanteric

fractures of the hip maintain quality at a reduced cost? The bone & joint journal. 2014 Sep;96-b(9):1192-7. Epub 2014/09/04.

63. Dy CJ, Dossous PM, Ton QV, Hollenberg JP, Lorich DG, Lane JM. Does a multidisciplinary team decrease complications in male patients with hip fractures? Clinical orthopaedics and related research. 2011 Jul;469(7):1919-24. Epub 2011/02/26.

64. Sivakumar BS, McDermott LM, Bell JJ, Pulle CR, Jayamaha S, Ottley MC. Dedicated hip fracture service: implementing a novel model of care. ANZ journal of surgery. 2013 Jul;83(7-8):559-63. Epub 2012/09/05.

65. Cruz JL, Moss MC, Chen SL, Hansen KM, Amerine LB. Retrospective evaluation of the clinical use of prothrombin complex concentrate for the reversal of anticoagulation with vitamin K antagonists. Blood Coagul Fibrinolysis. 2015 Jun;26(4):378-82.

66. Hickey M, Gatien M, Taljaard M, Aujnarain A, Giulivi A, Perry JJ. Outcomes of urgent warfarin reversal with frozen plasma versus prothrombin complex concentrate in the emergency department. Circulation. 2013 Jul 23;128(4):360-4.

67. Francis J, Girish G, Jones R. Prothrombin complex concentrate reversal of warfarin in patients with hip fracture. Anaesthesia. 2014 May;69(5):520-1. Epub 2014/04/18.

68. Purmonen T, Tormalehto S, Saavuori N, Kokki H. Budget impact analysis of warfarin reversal therapies among hip fracture patients in Finland. Drugs in R&D. 2015 Mar;15(1):155-62. Epub 2015/03/10.

69. Guest JF, Watson HG, Limaye S. Modeling the cost-effectiveness of prothrombin complex concentrate compared with fresh frozen plasma in emergency warfarin reversal in the United kingdom. Clin Ther. 2010 Dec;32(14):2478-93.

70. Eerenberg ES, Kamphuisen PW, Sijpkens MK, Meijers JC, Buller HR, Levi M. Reversal of rivaroxaban and dabigatran by prothrombin complex concentrate: a randomized, placebo-controlled, crossover study in healthy subjects. Circulation. 2011 Oct 4;124(14):1573-9.

71. Fetter RB, Shin Y, Freeman JL, Averill RF, Thompson JD. Case mix definition by diagnosis-related groups. Medical care. 1980 Feb;18(2 Suppl):iii, 1-53. Epub 1980/02/01.

72. Mayes R. The origins, development, and passage of Medicare's revolutionary prospective payment system. Journal of the history of medicine and allied sciences. 2007 Jan;62(1):21-55. Epub 2006/02/10.

73. Friedman SM, Mendelson DA. Epidemiology of fragility fractures. Clin Geriatr Med. 2014 May;30(2):175-81.

74. Brauer CA, Coca-Perraillon M, Cutler DM, Rosen AB. Incidence and mortality of hip fractures in the United States. Jama. 2009 Oct 14;302(14):1573-9. Epub 2009/10/15.

75. Browner BD, Jupiter JB, Krettek C, Anderson PA. Skeletal Trauma: Basic Science, Management and Reconstruction. 5th ed. Philadelphia, PA: Saunders; 2009.

76. Bozic K. CMS changes ICD-9 and DRG codes for revision TJA. AAOS Bulletin. 2005:17-21.

77. Bozic KJ, Katz P, Cisternas M, Ono L, Ries MD, Showstack J. Hospital resource utilization for primary and revision total hip arthroplasty. The Journal of bone and joint surgery American volume. 2005 Mar;87(3):570-6. Epub 2005/03/03.

78. Bozic KJ, Durbhakula S, Berry DJ, Naessens JM, Rappaport K, Cisternas M, et al. Differences in patient and procedure characteristics and hospital resource use in

primary and revision total joint arthroplasty: a multicenter study. The Journal of arthroplasty. 2005 Oct;20(7 Suppl 3):17-25. Epub 2005/10/11.

79. Alvin MD, Miller JA, Lubelski D, Rosenbaum BP, Abdullah KG, Whitmore RG, et al. Variations in cost calculations in spine surgery cost-effectiveness research. Neurosurgical focus. 2014 Jun;36(6):E1. Epub 2014/06/03.

80. Haley RW. Measuring the costs of nosocomial infections: methods for estimating economic burden on the hospital. The American journal of medicine. 1991 Sep 16;91(3B):32S-8S. Epub 1991/09/16.

81. Cooper LM, Linde-Zwirble WT. Medicare intensive care unit use: analysis of incidence, cost, and payment. Critical care medicine. 2004 Nov;32(11):2247-53. Epub 2005/01/11.

82. Dasta JF, McLaughlin TP, Mody SH, Piech CT. Daily cost of an intensive care unit day: the contribution of mechanical ventilation. Critical care medicine. 2005 Jun;33(6):1266-71. Epub 2005/06/09.

83. Anderson KL, Koval KJ, Spratt KF. Hip fracture outcome: is there a "July effect"? American journal of orthopedics (Belle Mead, NJ). 2009 Dec;38(12):606-11. Epub 2010/02/11.

84. Caveney AF, Silbergleit R, Frederiksen S, Meurer WJ, Hickenbottom SL, Smith RW, et al. Resource utilization and outcome at a university versus a community teaching hospital in tPA treated stroke patients: a retrospective cohort study. BMC health services research. 2010;10:44. Epub 2010/02/23.

85. Rana AJ, Bozic KJ. Bundled payments in orthopaedics. Clinical orthopaedics and related research. 2015 Feb;473(2):422-5. Epub 2014/02/21.

86. Froimson MI, Rana A, White RE, Jr., Marshall A, Schutzer SF, Healy WL, et al. Bundled payments for care improvement initiative: the next evolution of payment formulations: AAHKS Bundled Payment Task Force. The Journal of arthroplasty. 2013 Sep;28(8 Suppl):157-65. Epub 2013/09/27.

87. Bozic KJ, Ward L, Vail TP, Maze M. Bundled payments in total joint arthroplasty: targeting opportunities for quality improvement and cost reduction. Clinical orthopaedics and related research. 2014 Jan;472(1):188-93. Epub 2013/05/08.

88. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research. 2007 Mar;22(3):465-75. Epub 2006/12/06.

Appendices

Appendix 1: Data elements used to identify comorbidities in NIS, NSQIP, and NTDB

Comorbidity	NIS data element	NSQIP data element	NTDB data element
Alcoholism	□ "cm_alcohol" (Alcohol abuse)	□ "etoh" (EtOH > drinks/day in 2 wks before admission)	□ "comorkey" = 2 (Alcoholism)
Bleeding disorder	□ "cm_coag" (Coagulopathy)	□ "bleeddis" (Bleeding disorders)	□ "comorkey" = 4 (Bleeding disorder)
Diabetes mellitus	□ "cm_dm" (Diabetes, uncomplicated) □ "cm_dmcx" (Diabetes with chronic complications)	□ "diabetes" (Diabetes mellitus with oral agents or insulin)	□ "comorkey" = 11 (Diabetes mellitus)
Disseminated cancer	□ "cm_mets" (Metastatic cancer)	□ "discancr" (Disseminated cacner)	"□ comorkey" = 12 (Disseminated cancer)
Hypertension	□ "cm_htn_c" (Hypertension, uncomplicated and complicated)	□ "hypermed" (Hypertension requiring medication)	□ "comorkey" = 19 (Hypertension requiring medication)
Obesity	□ "cm_obese" (Obesity)	□ "height" (Height) □ "weight" (Weight)[†]	□ "comorkey" = 22 (Obesity)
Current smoker	□ICD-9 diagnosis 305.1 (Tobacco Use Disorder)	□ "smoke" (Current smoker within one year)	□ "comorkey" = 8 (Current smoker)

[†]In NSQIP obesity was defined as 703*(weight/(height*height)) ≥ 30 [BMI ≥ 30]

Appendix 2: Data elements used to identify inpatient adverse events in NIS, NSQIP, and NTDB

Adverse event	**NIS data element or ICD-9 diagnosis code**	**NSQIP data element**	**NTDB data element**
Acute kidney injury	• 584.X (Acute kidney injury)	• "noprenafl" (Acute renal failure, postoperative) • "nreninsf" (Progressive renal insufficiency, postoperative)	• "complkey" = 4 (Acute kidney injury)
Cardiac arrest	• 427.41 (Ventricular fibrillation) • 427.5 (Cardiac arrest)	• "ncdarrest" (Cardiac arrest requiring cardiopulmonary resuscitation, postoperative)	• "complkey" = 8 (Cardiac arrest with CPR)
Cerebrovascular accident (CVA)	• 430 (Subarachnoid hemorrhage) • 431 (Intracerebral hemorrhage) • 433.X1 (Precerebral artery occlusion and stenosis with cerebral infarction) • 434.X1 (Cerebral artery occlusion with cerebral infarction) • 997.02 (Iatrogenic cerebrovascular infarction or hemorrhage)	• "ncnscva" (Stroke/CVA, postoperative)	• "complkey" = 22 (Stroke / CVA)
Death	• "died" (Death)	• "dopertod" ≤ "doptodis" (Date of death ≤ Date of discharge)	• "hospdisp" = "Expired" (Death) • "eddisp" = "Died" (Death)
Deep vein thrombosis (DVT)/ Pulmonary embolism	• 415.1 (Pulmonary embolism) • 452.X-454.X, 453.82, 453.84, 453.85, 453.86 (Acute embolism and thrombosis of deep vein)	• "nothdvt" (DVT, with or without inflammation, postoperative) • "npulembol" (Pulmonary embolism, postoperative)	• "complkey" = 14 (DVT/ thrombophlebitis) • "complkey" = 21 (Pulmonary embolism)
Myocardial infarction	• 410.X (Acute myocardial infarction)	• "ncdmi" (Myocardial infarction, postoperative)	• "complkey" = 18 (Myocardial infarction)
Pneumonia	• 480-486 • (Pneumonia/ Bronchopneumonia)	• "noupneumo" (Pneumonia, postoperative)	• "complkey" = 20 (Pneumonia)
Surgical site infection (SSI)	• 998.5X (Postoperative infection, not elsewhere classified) • 998.67 (Infection and inflammatory reaction due to other internal orthopedic device, implant, or graft)	• "nsupinfec" (Superficial SSI, postoperative) • "nwndinfd" (Deep SSI, postoperative) • "norgspcssi" (Organ/space SSI, postoperative)	• "complkey" = 12 (Deep SSI) • "complkey" = 19 (Organ/space SSI) • "complkey" = 23 (Superficial SSI)
Urinary tract infection	• 599.0 (Urinary tract infection)	• "nurninfec" (Urinary tract infection, postoperative)	• "complkey" = 27 (Urinary tract infection)