The Journal of Arthroplasty xxx (2017) 1–7

Contents lists available at ScienceDirect

The Journal of Arthroplasty

journal homepage: www.arthroplastyjournal.org

Treatments of Missing Values in Large National Data Affect Conclusions: The Impact of Multiple Imputation on Arthroplasty Research

Nathaniel T. Ondeck, BS [a], Michael C. Fu, MD, MHS [b], Laura A. Skrip, PhD [c], Ryan P. McLynn, BS [a], Edwin P. Su, MD [b], Jonathan N. Grauer, MD [a,*]

[a] Department of Orthopaedics and Rehabilitation, Yale School of Medicine, New Haven, Connecticut
[b] Department of Orthopaedic Surgery, Hospital for Special Surgery, New York, New York
[c] Center for Infectious Disease Modeling and Analysis, Yale School of Public Health, New Haven, Connecticut

ARTICLE INFO

Article history:
Received 10 September 2017
Received in revised form
11 October 2017
Accepted 11 October 2017
Available online xxx

Keywords:
missing values
large data
complete case analysis
multiple imputation
adverse outcomes
unicompartmental knee arthroplasty

ABSTRACT

Background: Despite the advantages of large, national datasets, one continuing concern is missing data values. Complete case analysis, where only cases with complete data are analyzed, is commonly used rather than more statistically rigorous approaches such as multiple imputation. This study characterizes the potential selection bias introduced using complete case analysis and compares the results of common regressions using both techniques following unicompartmental knee arthroplasty.
Methods: Patients undergoing unicompartmental knee arthroplasty were extracted from the 2005 to 2015 National Surgical Quality Improvement Program. As examples, the demographics of patients with and without missing preoperative albumin and hematocrit values were compared. Missing data were then treated with both complete case analysis and multiple imputation (an approach that reproduces the variation and associations that would have been present in a full dataset) and the conclusions of common regressions for adverse outcomes were compared.
Results: A total of 6117 patients were included, of which 56.7% were missing at least one value. Younger, female, and healthier patients were more likely to have missing preoperative albumin and hematocrit values. The use of complete case analysis removed 3467 patients from the study in comparison with multiple imputation which included all 6117 patients. The 2 methods of handling missing values led to differing associations of low preoperative laboratory values with commonly studied adverse outcomes.
Conclusion: The use of complete case analysis can introduce selection bias and may lead to different conclusions in comparison with the statistically rigorous multiple imputation approach. Joint surgeons should consider the methods of handling missing values when interpreting arthroplasty research.

The last decade has seen the rapid evolution and subsequent utilization of large, national datasets in orthopedics research [1,2]. Arthroplasty studies are no exception and numerous recent projects have utilized large data to characterize outcomes following hip and knee replacement procedures [3–9]. Without a doubt, the use of national data has been instrumental in many contemporary papers that have had widespread clinical implications.

The large sample size available in such datasets allows for the statistically significant study of both rare adverse outcomes of commonly performed procedures (eg, stroke following total hip

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.arth.2017.10.034.

Conflict of Interest Statement: One of the authors (JNG) reports the following financial activities outside the submitted work: current consultancy with Bioventus (Durham, NC), ISTO Technologies (St Louis, MO), Medtronic (Minneapolis, MN), Stryker (Mahwah, NJ), Andante Medical Devices (White Plains, NY), and Vertex (Minneapolis, MN); ongoing expert testimony with legal case reviews; and a current grant with the Orthopaedic Trauma Association (Rosemont, IL). All other authors certify that he or she has no commercial associations that might pose a conflict of interest in connection with the submitted article.

Ethical Review Committee Statement: This study has been given an exemption from the senior author's Institutional Review Board under federal regulation 45 CFR 46.101(b) (4).

* Reprint requests: Jonathan N. Grauer, MD, Department of Orthopaedics and Rehabilitation, Yale School of Medicine, 800 Howard Avenue, New Haven, CT 06510.

https://doi.org/10.1016/j.arth.2017.10.034
0883-5403/

arthroplasty and total knee arthroplasty [9]) as well as the more typical postoperative course of less frequently performed surgeries (computer assisted arthroplasty) [10]. Additional benefits of using large data include the geographic and institutional diversity of study cohorts, allowing for broad generalizability of conclusions. For example, patients in the National Inpatient Sample represent a 20% stratified sample of all hospital inpatient stays in the United States, regardless of insurance provider, and the National Surgical Quality Improvement Program (NSQIP) collects information from over 600 sites across the nation [11,12].

However, there are certainly limitations to the use of large data, many of which can be mitigated once identified. For example, administratively coded datasets based on International Classification of Diseases coding have been shown to have poor to moderate accuracy in identifying patient comorbidities [13,14]. One way this has been addressed is through the use of chart abstracted clinical registries, such as NSQIP [15]. However, another area of concern that has been identified but has yet to be resolved, is missing values. In fact, the percentages of patients with missing data in NSQIP can range from 0% to as high as 80%, depending on the variable [16].

Abnormal preoperative laboratory values are among a group of potentially modifiable variables that has received increased attention in orthopedic surgery due to their association with adverse outcomes. In fact, recent arthroplasty studies have utilized NSQIP to study the associations of preoperative albumin and hematocrit levels with adverse outcomes and have found strong correlations [8,17–19]. With respect to missing data, these studies have relied upon complete case analysis, an approach that removes patients with missing values. This method of treating missing values is subject to selection bias, and thus may be inferior to more rigorous statistical approaches such as multiple imputation, which approximates the associations that would have been present in a full dataset [20–22].

In the context of the significant value that arthroplasty studies utilizing large, national data can provide, contrasted with unresolved concerns regarding the handling of missing data, the present project serves to characterize the potential bias introduced using the commonly chosen complete case analysis technique in patients undergoing unicompartmental knee arthroplasty [UKA] (as an example arthroplasty population) in the NSQIP dataset. Additionally, this study compares the outcomes of common regressions after using both complete case analysis and multiple imputation for handling missing data.

Materials and Methods

Patient Sample

Patients with a Current Procedural Terminology code for UKA (27446) were extracted from NSQIP data years 2005 through 2015. This dataset collects patient information from over 600 centers in the United States to present over 300 patient variables with 30-day follow-up, regardless of admission status. The dataset undergoes rigorous inter-rater reliability audits, with the rate of disagreements being approximately 2% [12,23]. This clinical registry was chosen for this study as it has been utilized extensively in arthroplasty research and has existing concerns regarding the percentages of missing values [3–9,16].

Patient Variables

Patient demographic variables extracted from the dataset included age (integer variable), gender (dichotomous variable), height (continuous variable), and weight (continuous variable). Age was divided into commonly used categories (18-64, 65-74, and ≥75 years) and height and weight were utilized to calculate body mass index (BMI), which was further divided into commonly used categories (≤24, 25-29, 30-34, and ≥35 kg/m^2). Also, obtained through NSQIP was the American Society of Anesthesiologists (ASA) physical status classification system score, a measure of the totality of each patient's health. These scores were also categorized using common subgroups (Class 1/2 and Class ≥3).

Patient preoperative albumin values were divided into low (<3.5 g/dL), not low (≥3.5 g/dL), and missing. The preoperative hematocrit values were divided into low (female <36%, male <39%), not low (female ≥36%, male ≥39%), and missing. These groupings were chosen as they have been previously used in the arthroplasty literature [8,17,24].

Outcome Variables

The outcome variables studied in this analysis included the occurrence of severe adverse events (cardiac arrest, death, deep vein thrombosis, myocardial infarction, postoperative intubation, pulmonary embolism, return to the operating room, sepsis, or stroke) and minor adverse events (anemia requiring transfusion, pneumonia, surgical site infection, urinary tract infection, and wound dehiscence). A secondary outcome variable was discharged to a higher level of care (discharge to rehabilitation, a separate acute care, a skilled care facility that was not home, or an unskilled care facility that was not home). These groupings of adverse outcomes have been previously used in the arthroplasty literature [3,7,10]. All the aforementioned variables are reported in NSQIP and gathered for 30 days from surgery, regardless of admission status.

Missing Data Approach

The first approach for missing data was complete case analysis, which is currently commonly used in NSQIP studies. There are 2 steps involved in this technique. For the missing data step, patients missing any of the values that are to be used in the analysis are removed [25]. For this study, patients missing age, gender, height, weight, ASA, albumin, or hematocrit were excluded. For the analysis step, a single logistic regression for the outcome variable of choice was performed. In this study, 3 separate logistic regressions were performed with the outcome variables involving the occurrence of severe adverse events, minor adverse events, and discharge to a higher level of care.

The second method of handling missing data was multiple imputation using chained equations, which is a statistically rigorous method of utilizing existing data to reproduce the conclusions that would have been present in a complete dataset [20–22]. For the missing data step, missing values were imputed with predicted numbers from a logistic equation. Due to this process inherently leading to decreased variation in the imputed dataset, a randomly drawn residual value is added to the regression equation. Furthermore, in recognition of the inability to perfectly approximate the missing values, this is repeated M times, where M is the number of imputed datasets. In this study, M was chosen to be 60 in order to match or exceed the percentage of subjects with missing values, a practice that is endorsed in the imputation literature [26]. The analysis step involves M logistic regressions for the outcome variable of choice (in this case for severe adverse events, minor adverse events, and discharge to higher level of care), each one using one of the M imputed datasets. The third and final step, the pooling step, involves combining the results of each of the M imputed datasets to achieve the final conclusions.