



TECHNICAL NOTES
for Common Procedures Report

September
2022

Pennsylvania Health Care Cost Containment Council

Report Period: October 1, 2020 through September 30, 2021

225 Market Street, Suite 400, Harrisburg, PA 17101

Phone: 717-232-6787

Fax: 717-232-3821

www.phc4.org

Barry D. Buckingham, Executive Director

Table of Contents

OVERVIEW 3

DATA COLLECTION AND VERIFICATION 4

Handling of Anomalous Laboratory Test Results..... 4

STUDY POPULATION 5

Inclusion Criteria 5

General Exclusion Criteria 5

Measure-Specific Exclusions 5

MEASURES REPORTED 6

Total Number of Cases 6

Measures with Risk-Adjusted/Statistical Ratings 6

Case-Mix Adjusted Average Hospital Charge 8

Average Medicare Payment 8

RISK-ADJUSTMENT AND STATISTICAL RATINGS 9

Model Development 9

Determining Expected (Predicted) Value at the Patient Level 11

Determining Actual and Expected (Predicted) Values at the Hospital Level 11

Determining Statistical Ratings 12

Assignment of Statistical Ratings 14

CASE-MIX ADJUSTMENT AND AVERAGE HOSPITAL CHARGE 15

Trim Methodology 15

Determining Actual Charges 15

Determining Expected Charges 15

Determining Case-Mix Adjusted Charges..... 16

DATA TABLES 17

Table 1. Statewide Utilization and Outcome Data, by Procedure 17

Table 2. Exclusions from Analyses, by Procedure and Measure 18

APPENDICES

Appendix A. Definitions—Study Populations, Code-based Exclusions and Complications A-1

Appendix B. Valid Discharge Status Codes..... A-2

Appendix C. Example of Logistic Regression A-3

Appendix D. Example of Linear Regression A-4

Appendix E. Example of Case-Mix Adjustment A-5

OVERVIEW

The Technical Notes serve as a technical supplement to the Pennsylvania Health Care Cost Containment Council's (PHC4) report on common surgical procedures. This document describes the methodology and development of the report and provides data on statewide results and cases excluded from analyses. Statewide utilization and outcome data are displayed in Table 1 and exclusion data are displayed in Table 2. See Appendix A for ICD-10-CM/PCS codes and MS-DRGs defining the study populations and code-based exclusions.

The procedures and measures outlined below are reported for Pennsylvania acute care hospitals that typically perform these procedures on adults ages 18 years and older.¹ In addition to the measures listed below, the total number of cases (after exclusions) is reported.

The following measures are reported for the **Spinal Fusion** procedure group when a hospital has five or more cases:

- Risk-adjusted in-hospital complication rating
- Risk-adjusted readmission for complication rating
- Risk-adjusted extended postoperative length of stay rating
- Average hospital charge (case-mix adjusted)

The following measures are reported for the **Total Hip Replacement** and **Total Knee Replacement** procedure groups when a hospital has five or more cases:

- Risk-adjusted complication rating
- Risk-adjusted extended postoperative length of stay rating
- Average hospital charge (case-mix adjusted)

The methodology described in this document was developed to account for the differences among individual patients that had the potential to influence the outcomes of the common procedures reported.

Special note regarding COVID-19: the COVID-19 pandemic has had a substantial impact on Pennsylvania hospitals since its onset in the early part of 2020. The first case of COVID-19 in Pennsylvania was reported by the Pennsylvania Department of Health on March 6, 2020. Since that time, hospitals have faced extraordinary challenges in care delivery, particularly for patients diagnosed with COVID-19. While the *Common Procedures Report* includes inpatient discharge data from the pandemic period, cases with a COVID-19 diagnosis are excluded from all analyses (see "General Exclusion Criteria" section for details) since these hospitalizations are clinically complex and measuring hospital outcomes related to these atypical cases is not the intent of this report.

¹ Results are not displayed for hospitals that closed or merged with other facilities.

DATA COLLECTION AND VERIFICATION

The data for the *Common Procedures Report*, obtained from the inpatient UB-04 (Uniform Billing) form, was submitted electronically to PHC4 by Pennsylvania acute care hospitals that performed the procedure of interest primarily on adults. Federal hospitals were not included. The data included demographic information, hospital charges, and International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM) and Procedure Coding System (ICD-10-PCS) diagnosis and procedure codes.

Hospitals submitted data to the Council on a quarterly basis (within 90 days from the last day of each quarter). Upon receipt of the data, verification was performed to assure data were submitted in a readable format. Extensive quality assurance checks were completed and laboratory data were matched to inpatient records. Error reports for UB-04 data were then generated and returned to each hospital with an opportunity to correct any problems. Similarly, laboratory test results were evaluated each quarter and summary reports indicating data anomalies were sent to each hospital, again with an opportunity to make corrections. Data accuracy and completeness were ultimately the responsibility of each individual hospital.

Laboratory test results were submitted by hospitals to the Council for a select group of acute care inpatient records, including those used for analysis for the procedures included in the *Common Procedures Report*. Hospitals were required to submit the highest and/or lowest result(s) for a maximum of 29 laboratory tests as collected from patients during the initial period of their hospitalization. The requirements for submitting this data are specified elsewhere (refer to PHC4's *Laboratory Data Reporting Manual*, accessible at www.phc4.org). In brief, for patients admitted prior to 6:00 p.m., only laboratory results collected on Day 1 of the admission were to be submitted. For patients admitted after 6:00 p.m., results were to be submitted for tests collected on the day of admission (Day 1) through the next calendar day (Day 2).

Handling of Anomalous Laboratory Test Results

Risk adjustment relied on the submission of valid and accurate laboratory test data. As noted, hospitals were given the opportunity to correct data anomalies (laboratory data that was so unreasonably high or low that it was most plausibly representative of a data error). Hospitals were notified of anomalous laboratory data submissions via specific feedback reports provided on a quarterly basis. Since anomalous data that was not corrected had the potential to inaccurately skew all hospitals' final statistical ratings, such extreme values were replaced with default (typical) values when building risk-adjustment models. In effect, such laboratory test results were treated as if they were missing, in which neither penalty nor credit relative to the implicated data was applied in the calculation of a patient's risk.

STUDY POPULATION

The study population for each procedure reported is designed to represent a clinically cohesive group of patients. See Appendix A to access the ICD-10-PCS procedure codes, Medicare Severity Diagnosis Related Groups (MS-DRG), and Major Diagnostic Categories (MDC) associated with each study population.

Inclusion Criteria

The study populations included inpatient acute care records for adults (18 years and older) discharged from Pennsylvania acute care hospitals during the defined report period with an applicable ICD-10-PCS procedure code in either the principal or secondary procedure code positions in the discharge record.

General Exclusion Criteria

The number of cases included in any single type of analysis varied because each reported measure had its own unique set of exclusion criteria (see “Measure-Specific Exclusions” section). However, the following types of records were excluded from all measures for all reported procedures.

- Records with errors (e.g., systematic errors in coding of essential data fields such as discharge status, dates, charges, etc.)
- Duplicate records
- Records with discharge dates not in study period
- Records representing an inpatient rehabilitation hospitalization (not acute care) defined by rehabilitation revenue codes 0024, 0118, 0128, 0138, 0148 or 0158
- Non-adult records (< 18 years) or records with invalid age (e.g., records that did not have the necessary data for the calculation of age or for which age was ≥ 120 years)
- Patients treated in children’s hospitals
- ICD-10-CM/PCS code-based exclusions¹ (see Appendix A)
- Records not in MDC/MS-DRG of interest (records not assigned to MDC/MS-DRG specified in Appendix A).
- Records with missing or invalid discharge status (see Appendix B for valid discharge status codes)

Measure-Specific Exclusions

In addition to the cases excluded from the general study population (see “General Exclusion Criteria” section), individual hospitalizations were excluded from outcome analyses when the data in the record was insufficient or inappropriate to the measure of interest. For example, patients who died were excluded from the readmission for complication analysis but not the in-hospital complication analysis. See Table 2 for the exclusions applied for each measure and the number of cases excluded.

¹ Centers for Medicare and Medicaid Services. “2022 Procedure-Specific Complication Measure Updates and Specifications Report Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA) – Version 11.0.” April 2022. Available at <https://www.qualitynet.org/inpatient/measures/complication/methodology>.

MEASURES REPORTED

Total Number of Cases

For each procedure included in the report, the total number of hospitalizations for patients 18 years and older, after in-hospital complication exclusions, is reported. If two procedures from the same procedure group were performed during the same hospitalization, the case was only counted once.

If a total hip replacement and a total knee replacement were performed during the same hospitalization, the case was assigned to either the total hip or total knee replacement study population based on the principal diagnosis and procedure codes present in the patient record.

Measures with Risk-Adjusted/Statistical Ratings

Risk-adjusted ratings are reported for the complication measures and extended postoperative length of stay measure. The rating identifies whether the hospital's observed rate of a given outcome was significantly higher than, significantly lower than, or not significantly different than expected based on patient risk factors (see "Risk Adjustment and Statistical Ratings" for methodology details and Appendix C and D for examples). Ratings are reported for hospitals with five or more cases in the analysis.

In-Hospital Complication (Reported for Spinal Fusion)

In-hospital complications occurred during the hospitalization in which the procedure was performed (also referred to as the index hospitalization). A complication was counted when 1) an ICD-10-CM diagnosis code used to define a complication for spinal fusion was a secondary diagnosis and was not present on admission, as determined by the present on admission (POA) indicator (for certain complications the diagnosis code was paired with a procedure code) or 2) the patient died, as determined by a discharge status code of "20."¹

Readmission for Complication (Reported for Spinal Fusion)

A readmission for complication following spinal fusion is defined as a rehospitalization to a Pennsylvania acute care hospital within 7, 30, or 90 days (depending on the type of complication) from the date of discharge of the index hospitalization with a principal diagnosis that indicated a complication. For certain complications the principal diagnosis code is paired with a procedure code.¹

Readmission for complication is a dichotomous (yes/no) outcome; as such, it is counted only once when an index hospitalization results in multiple readmissions for complication. If, over the study period, a patient had multiple discharges for spinal fusion, each discharge was independently investigated to determine whether it had a readmission for complication with one exception. If a second hospitalization for spinal fusion occurred within 90 days of the first index hospitalization, the second hospitalization was excluded from the readmission for complication analysis.

¹ See Appendix A to access the ICD-10-CM/PCS diagnosis and procedure codes that define the complications for a particular procedure.

Complication (Reported for Total Hip Replacement and Total Knee Replacement)

The complication measure is based, in large part, on the Centers for Medicare and Medicaid Services measure designed for total hip and total knee replacements likely to be considered elective.¹

Complications included in the measure are those that:

- occurred during the hospitalization in which the procedure was performed (also referred to as the index hospitalization). A complication was counted when 1) an ICD-10-CM diagnosis code used to define a complication for a total hip or total knee replacement was a secondary diagnosis and was not present on admission, as determined by the present on admission (POA) indicator (for certain complications the diagnosis code was paired with a procedure code) or 2) the patient died, as determined by a discharge status code of “20.”²

or

- occurred as the principal diagnosis of a readmission to a Pennsylvania acute care hospital within 7, 30, or 90 days (depending on the type of complication) from the date of discharge of the index hospitalization. For certain complications the principal diagnosis code is paired with a procedure code.²

The complication measure is a dichotomous (yes/no) outcome; as such, it is counted only once when multiple complications occur. If, over the study period, a patient had multiple discharges for total hip replacement or total knee replacement, each discharge was independently investigated to determine whether it had a complication with one exception. If a second hospitalization for total hip replacement or total knee replacement occurred within 90 days of the first index hospitalization, the second hospitalization was excluded from the complication analysis.

Extended Postoperative Length of Stay

In general terms, an extended postoperative length of stay (PLOS) identifies that the actual length of time a patient remains in the hospital following the procedure is significantly longer than what would be expected, after accounting for patients’ risk. The development of this measure was guided, in part, by the approach used by Michael Pine and Associates.³

Details for determining the patient’s actual PLOS, expected (predicted) PLOS, and whether the hospital stay should be counted as an extended PLOS are outlined below (see Appendix D for example):

- The actual PLOS (in days) was calculated as the discharge date minus the date the procedure of interest was performed. Patients discharged on the same day the procedure was performed were assigned a PLOS (in days) of 0.5.
- The expected PLOS was determined using the risk-adjustment techniques described under “Model Development” and “Determining Expected Value at the Patient Level” in the “Risk-Adjustment and Statistical Ratings” section.

¹ Centers for Medicare and Medicaid Services. “2022 Procedure-Specific Complication Measure Updates and Specifications Report Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA) – Version 11.0.” April 2022. Available at <https://www.qualitynet.org/inpatient/measures/complication/methodology>.

² See Appendix A to access the ICD-10-CM/PCS diagnosis and procedure codes that define the complications for a particular procedure.

³ Fry DE, Pine M, Jones BL, Meinban RJ. Adverse outcomes in surgery: redefinition of postoperative complications. *The American Journal of Surgery*. 2009;197:479-484.

- A natural log transformation of each PLOS (actual and expected values) was performed to account for skewness in the PLOS distribution.
- An extended PLOS was identified when the difference between the actual and expected log transformed PLOS values—the residual log PLOS—for a particular patient was significantly higher than the average residual log PLOS for all patients in the analysis. In statistical terms, an extended PLOS is identified when the residual log transformed PLOS is greater than two standard deviations above the average residual log transformed PLOS for a given procedure.

Case-Mix Adjusted Average Hospital Charge

The hospital charge is the total amount charged to the patient for the entire hospitalization during which the procedure of interest was performed. It does not include professional fees (e.g., physician fees) or other additional post-discharge costs, such as rehabilitation treatment, long-term care and/or home health care. The average charges reported were trimmed and case-mix adjusted as described in the “Case-Mix Adjustment and Average Hospital Charge” section. Average charges are reported for each hospital with five or more cases.

Average Medicare Payment

The statewide average payment is reported for Medicare fee-for-service (FFS) patients (Pennsylvania residents only). Claim payment amounts were obtained from the Centers for Medicare and Medicaid Services (CMS) and matched to cases meeting the study population criteria for a given procedure. The payment data displayed in this report corresponds to hospitalizations from calendar year (CY) 2020. Patient liabilities (e.g., coinsurance and deductible dollar amounts) and payments from Medicare Advantage plans (e.g., Medicare HMOs) were not included.

For each procedure, the overall statewide average payment is reported by MS-DRG to account for variations in case mix. The number of cases included in the average payment is also displayed. Average payments are reported at the statewide level only and are displayed for MS-DRGs with 11 or more cases.

RISK-ADJUSTMENT AND STATISTICAL RATINGS

In order to report fair comparisons among hospitals for a given procedure and measure, regression techniques were used to construct “risk models” 1) for predicting the risk of a particular event (e.g., complication) occurring, or 2) for the extended postoperative length of stay (PLOS) measure, predicting the log transformed PLOS. Each model was a mathematical formula used to ultimately predict a patient’s probability of the event occurring or log PLOS based on relevant risk factors. Cases with these risk factors were given more “credit” in the calculation, leading to a higher predicted probability of the event or, in effect, a longer PLOS. The ratings indicate whether the hospital’s event rates were within the expected range or higher or lower than the expected range, after taking into account the risk factors that were included in the risk-adjustment models.

Model Development

The first step in building the risk adjustment models was to prepare a reference database. UB-04 data and laboratory test results from adult (age ≥ 18 years) discharges from Pennsylvania acute care hospitals were used. The reference database for each procedure and measure-specific model was based on two years of data (after exclusions).

Identifying Potential Risk Factors

The next step in building the models was to identify risk factors that potentially contributed to the event or outcome (i.e., complication or longer PLOS). These factors were identified through their importance in past models, review of scientific literature and consideration of high-risk populations. Types of risk factors included patient characteristics, socioeconomic factors, laboratory test results, diagnoses and procedures identified by ICD-10-CM/PCS codes, and other UB-04-derived factors.

Using the reference database, potential risk factors were subject to univariate analysis to determine which, because of their potential to predict the event of interest, should be tested for inclusion in the model for a given procedure and measure. Variables were constructed and analyzed as linear (continuous), categorical and binary as appropriate. For some factors, multiple forms of variable construction were analyzed to determine which approach best fit the data. For example, patient age was tested as a linear or linear spline with up to two knots to determine which approach best fit the data.

When constructing categorical variables, data was partitioned into a maximum of five categories as appropriate:

- For variables with continuous data (e.g., laboratory test results) one category represented “typical” results with additional categories representative of abnormal results generally associated with increased risk. (In the final model, all records in a specified abnormal category would receive the same amount of credit, regardless of the value within the category.) Records with missing values were combined with records in the typical category.
- For ICD-10-CM/PCS code-based categorical variables, one category represented the absence of the risk factor and additional categories represented the presence of diagnosis or procedure codes indicating increased risk for that particular condition (e.g., no cancer, primary cancer and metastatic cancer).

Categorical and binary variables were selected for testing in the model based on the following criteria:

- Minimum volume: For categorical variables, each category represented at least one percent of the total cases in the study. For binary variables, cases with the risk factor were required to represent at least one percent of the total cases in the study. Exceptions were made to this criterion when a variable had particular clinical relevance to the outcome.
- Order of risk: For categorical variables, categories farther away from the “typical” category were required to have rates of increasing risk (e.g., when the typical category was defined as level A, categories B, C, D and E were required to have increasingly higher rates of risk). For binary variables, cases with the risk factor were required to have a higher rate of risk than cases without the risk factor.
- Significance: Categories were required to have significantly different rates of risk.

To avoid developing models that were “overfitted” (i.e., unnecessarily complex models with factors that may be insignificant when applied to a different dataset), a statistical criterion called the Schwarz criterion was used. This application avoided the problem of overfitting by including a penalty value for each factor as it was added to the model. In this way, the best end point for the model build (i.e., the point in which no more factors should be added to the model) could be determined. In some instances, exceptions were made to the Schwarz criterion for factors identified in the research literature as clinically important.

Each procedure and measure combination was modeled separately, with the exception of total hip and knee replacement, which were modeled together. Binary logistic regression was used for analyses of the complication measures. Linear regression was used for postoperative length of stay (PLOS) analysis.

Risk Factor Selection

Risk factors selected for testing were added to the model in the following order: 1) procedure group (total hip replacement or total knee replacement models only), 2) patient characteristics (patient age and sex), 3) laboratory test results, 4) ICD-10-CM/PCS code-based variables, 5) UB-04-derived data elements (e.g., race/ethnicity, socioeconomic status), then 6) insurance type. All factors within a risk factor type were evaluated before considering factors from the next type.

Risk factors were considered statistically significant in a model if they met the $p < 0.10$ significance criterion and the Schwartz criterion and indicated an increase in the risk of the event. However, risk factors were evaluated for relevance by considering both mathematical (statistical significance) and clinical perspectives (clinical importance).

Bootstrap validation

Once the model variables were chosen, the bootstrap technique was used to identify and eliminate factors that were unstable and unlikely to predict the same level of risk when applied to other (future) datasets. Using this technique, two hundred fifty sample datasets were randomly generated from the reference database. Records were allowed to appear multiple times in the sample datasets if they were selected repeatedly. The prepared model was then fit to each sample dataset to determine the percent of sample models in which each factor maintained significance ($p < 0.10$). Risk factors at or above a 75% cutoff and those with particular clinical relevance to the outcome (even if below the 75% cutoff) were retained in the final model. This same approach was used to eliminate any factor that did not have a consistently expected direction of the numeric value/coefficient in at least 75% of the sample models. Finally, factors in the model were investigated to be sure that they were not overly influenced by the effects of a few hospitals. This could be a special concern for factors that may be concentrated in a few

hospitals. A hierarchical model was run with a nested random intercept unique for each hospital to assess the power of the factors after accounting for hospital differences. Factors no longer significant or with a changed sign in this hierarchical model were eliminated.

Determining Expected (Predicted) Value at the Patient Level

The final risk models estimated the relative effects (β_n) that each of the risk factors had on the relevant outcome value for each hospitalization. The model equations took the following form:

$$\beta X = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n$$

where:

β_n = the relevant model coefficient (β_0 is the intercept)

X_n = the value of the risk factor for a hospitalization

These models were then used to calculate the predicted values (e.g., predicted probability of an event occurring or predicted log transformed PLOS) for each individual hospitalization (after exclusions). The risk factor values (X) were multiplied by the model coefficients (β) and summed to determine the value βX for each hospitalization.

Using logistic regression modeling, the predicted value for a patient's probability of the event (i.e., complication) occurring was calculated as

$$p = \frac{e^{\beta X}}{1 + e^{\beta X}}$$

where $e \approx 2.7182818285$

Using linear regression modeling, a patient's predicted log transformed PLOS was calculated as βX . This value was then used in calculations to identify hospitalizations with an extended PLOS as described in the "Measures Reported" section.

To account for changes in the statewide rates over time, the intercept (β_0) of the models were adjusted so that the statewide expected rate, or average log transformed PLOS, for the current study period was equal to the actual statewide rate for this same period.

See Appendix C for an example of logistic regression. See Appendix D for an example of linear regression and the calculation to determine if a hospitalization had an extended PLOS.

Determining Actual and Expected (Predicted) Values at the Hospital Level

Separate analyses were performed to determine, for each hospital, the actual and expected percent of hospitalizations with a given outcome. Significance tests were conducted to determine whether the difference between a hospital's actual and expected values was too large to be attributed solely to chance. These results were displayed as ratings.

Determining Actual (Observed) Values

Outcome Percent This percent was determined by dividing the total number of hospitalizations with an event by the number of hospitalizations in the analysis for a given procedure.

The actual value was calculated for each procedure and outcome combination as shown below:

Spinal Fusion	In-hospital complication percent Readmission for complication percent Extended postoperative length of stay percent
Total Hip Replacement	Complication percent Extended postoperative length of stay percent
Total Knee Replacement	Complication percent Extended postoperative length of stay percent

Determining Expected (Predicted) Values

For the complication outcomes, the expected value for a particular hospital was calculated as the average of the predicted probabilities for each patient (i.e., hospitalization) in the analysis. This was done to adjust for the risk inherent to each particular hospital's patient population.

For the extended PLOS outcome, the expected value for a particular hospital was the percent of hospitalizations statewide with an extended PLOS.

Determining Statistical Ratings

Significance tests (using the binomial distribution, see below) were performed for each outcome measure. To account for random variation, statistical evaluation was used to determine whether the difference between a hospital's observed and expected values was too large to be attributed solely to chance.

Binomial Distribution

The use of the binomial distribution required the following assumptions:

- Each observation included in the study had one of two observable events (e.g., complication vs. no complication). In other words, the response was dichotomous.
- The probability of the event (i.e., complication) for each observation studied was equal to the probability provided by the associated logistic risk model. For extended PLOS, it is assumed the probability of an extended PLOS occurring is equal to the statewide extended PLOS rate.
- The result for any one observation in the analyses had no impact on the result of another observation. In other words, the observations were independent.

The probability distribution for a specific hospital's outcome in one area of analysis was based on the hospital's predicted or expected values. Using the probability distribution, a p-value was calculated for

each observed value. This p-value was the probability, or likelihood, that the value could have occurred by chance. If it was very unlikely ($p < 0.05$; see “Inferential Error” section below) that the observed or actual value could have occurred only by chance, it was concluded that the observed value was “significantly different” from the expected value.

Calculation of p-values

The binomial distribution defined a probability of each potential outcome (e.g., the probability of observing exactly 3 complications out of 40) according to the binomial formula:

$$P(a) = \left[\frac{N!}{a!(N-a)!} \right] p^a (1 - p)^{N-a}$$

where:

- a was the number of events (e.g., complications) that were observed (i.e., a = 1 complication, a = 2 complications, etc.) in N hospitalizations. The value of “a” ranged from 0 through N (in other words, $0 \leq a \leq N$).
- P(a) was the probability that exactly “a” events would be observed.
- N was the number of hospitalizations for a particular hospital.
- p was the overall expected rate (e.g., expected percent of complication) for a particular hospital.

The rating process evaluated both fewer than expected as well as greater than expected complications. Thus, a two-tailed test was used. In the example above (3 complications out of 40), the probability associated with the left-hand tail was the sum of the probability for 0, 1, 2, or 3 complications out of 40. The probability of the right-hand tail was the sum of the probabilities at the upper end of the range (40, 39, 38...) until that sum was as close as possible to (but still less than) the probability associated with the left-hand tail. The two-tailed p-value was the sum of the probability of the left-hand and right-hand tails.

The two-tailed p-value was calculated for each hospital.

Inferential Error

A type of inferential error that can be made in statistics is called a Type I error or “false positive.” The probability of committing a Type I error is equal to the level of significance established by the researcher. For the current analyses, the level of significance was set to 0.05.

In the context of the *Common Procedures Report*, a Type I error would have occurred when the difference between the actual complication percent and the expected complication percent was declared statistically significant, when in fact, the difference was due to chance. That is, the hospital was declared to be statistically higher or lower than expected when in reality the hospital’s level of performance was comparable to its expected performance, as determined by its risk profile. Since the level of significance was set to 0.05, there was a 5% chance (or 1 in 20) of committing this type of error.

Assignment of Statistical Ratings

A statistical rating of higher than expected or lower than expected was assigned to each hospital if the difference between what was observed and what was expected was statistically significant (p-value <0.05). The p-value, calculated in terms of a “two-tailed” test, was compared to the level of significance.

For example, in determining the complication rating for each hospital:

- If the calculated p-value was less than 0.05, then the conclusion was made that the difference between what was expected and what was observed was statistically significant.
 - If the actual percent of events was less than expected, the hospital was assigned the symbol “○” (as shown in the *Common Procedures Report*) to indicate that the percent of events was significantly less than expected.
 - If the actual percent was higher than expected, the hospital was assigned the symbol “●” (as shown in the *Common Procedures Report*) to indicate that the percent of events was significantly greater than expected.
- If the calculated p-value was greater than or equal to 0.05, then the conclusion was made that the difference between the expected and the actual percent of events was not statistically significant. It cannot be concluded that the actual percent for that particular hospital was different from the expected percent derived from the particular hospital's risk profile. In this case the hospital was assigned the symbol “◉” (as shown in the *Common Procedures Report*).

CASE-MIX ADJUSTMENT AND AVERAGE HOSPITAL CHARGE

Hospital charges were adjusted separately for each procedure in the report to account for differences in the charges across Pennsylvania geographical regions and hospital variation in the mix of cases across MS-DRGs for a given procedure. This adjustment was made at the level of the nine PA regions (see Appendix E for regions by county).

Cases from PA region/MS-DRG group combinations with low volume (< 20 cases after trimming of outliers) were excluded from the analysis.

Trim Methodology

Trimming was used to remove outlier charges from the study population for a given procedure in order to eliminate extreme values that may have a significant and unrepresentative impact on the average. Since charges varied dramatically among regions, upper and lower trim points were calculated at the regional level for each MS-DRG for the procedure. Cases with charges that were below the lower trim point or above the upper trim point were excluded from further analysis.

Upper and lower trim points were calculated using the “+/- 3.0 interquartile range” method. This non-parametric methodology was used because, historically, the distribution for charges does not follow a normal “bell-shaped” pattern.

Trim points were determined as follows:

$Q1$ = *the first quartile (25th percentile total charge) of all patient records from the comparative database in a particular category*

$Q3$ = *the third quartile (75th percentile total charge) of all patient records from the comparative database in a particular category*

IQR = $Q3 - Q1$

Lower Trim Point = $Q1 - (3.0 \times IQR)$

Upper Trim Point = $Q3 + (3.0 \times IQR)$

Determining Actual Charges

The actual average charge (Average ActChg) was determined as the average (arithmetic mean) charge for the hospitalizations included in the hospital's charge analysis for the procedure analyzed.

Determining Expected Charges

The expected charge (ExpChg) for a hospitalization was equal to the average charge for all hospitalizations in that particular PA region/MS-DRG group combination for the procedure analyzed. The hospital's expected charge was determined as the average (arithmetic mean) of the expected charges for the hospitalizations included in the hospital's charge analysis:

$$\text{Average ExpChg} = \frac{\sum \text{ExpChg}}{n}$$

Determining Case-Mix Adjusted Charges

The case-mix adjusted charge was calculated by dividing the average actual charge (Average ActChg) by the average expected charge (Average ExpChg) for the hospital and then multiplying this quantity by the average charge for the hospital's region for the a given procedure:

$$\frac{\text{Average ActChg}}{\text{Average ExpChg}} \text{ (Average Actual Charge for a particular region)}$$

See Appendix E for an example of how case-mix adjusted charges were computed.

DATA TABLES

Table 1. Statewide Utilization and Outcome Data, by Procedure

Spinal Fusion	
Total Number of Cases	13,296
In-Hospital Complication	1.9%
Readmission for Complication	2.0%
Extended PLOS	2.4%
Average Hospital Charge	\$139,340
Total Hip Replacement	
Total Number of Cases	12,365
Complication	2.0%
Extended PLOS	3.8%
Average Hospital Charge	\$60,542
Total Knee Replacement	
Total Number of Cases	17,986
Complication	1.5%
Extended PLOS	3.1%
Average Hospital Charge	\$56,517

Table 2. Exclusions from Analyses, by Procedure and Measure**Table 2A. Exclusions¹ for Spinal Fusion**

	Spinal Fusion	
	# Cases	% Cases
Total Cases Before Exclusions	18,315	100.0%
Universal Exclusions		
Records from hospitals with data problems	0	0.0%
Duplicate records	4	<0.1%
Discharge date not in study period	0	0.0%
Rehabilitation revenue code in record	6	<0.1%
Invalid age, patients <18 or ≥ 120 years	639	3.5%
Patients treated in children's hospitals	61	0.3%
Total Universal Exclusions	710	3.9%
In-Hospital Complication Exclusions		
Universal exclusions	710	3.9%
ICD-10-CM/PCS code-based exclusions	3,780	20.6%
Not MS-DRG or MDC of interest	527	2.9%
Missing or invalid discharge status	2	<0.1%
Included in In-Hospital Complication Measure	13,296	72.6%
Readmission for Complication Exclusions		
Universal exclusions	710	3.9%
ICD-10-CM/PCS code-based exclusions	3,780	20.6%
Not MS-DRG or MDC of interest	527	2.9%
Missing or invalid discharge status	2	<0.1%
Patients who left against medical advice	23	0.1%
In-hospital mortality	13	0.1%
Out-of-state residents	1,245	6.8%
Missing or invalid patient identifier	11	0.1%
Subsequent index hospitalization within 90 days	48	0.3%
Included in Readmission for Complication Measure	11,956	65.3%
Extended Postoperative Length of Stay Exclusions		
Universal exclusions	710	3.9%
ICD-10-CM/PCS code-based exclusions	3,780	20.6%
Not MS-DRG or MDC of interest	527	2.9%
Missing or invalid discharge status	2	<0.1%
Patients who left against medical advice	23	0.1%
Patients transferred to another acute care hospital	48	0.3%
In-hospital mortality	13	0.1%
Discharge to hospice (records with a discharge status code of 50 or 51)	3	<0.1%
Postoperative length of stay inconsistencies ²	352	1.9%
Included in Extended PLOS Measure	12,857	70.2%

(Continued on next page)

	Spinal Fusion	
	# Cases	% Cases
<i>Continued from previous page</i>		
Charges Exclusions		
Universal exclusions	710	3.9%
ICD-10-CM/PCS code-based exclusions	3,780	20.6%
Not MS-DRG or MDC of interest	527	2.9%
Missing or invalid discharge status	2	<0.1%
Patients who left against medical advice	23	0.1%
Patients transferred to another acute care hospital	48	0.3%
Invalid charges	6	<0.1%
Charge outliers	140	0.8%
No reference data	258	1.4%
Included in Average Charge Calculation	12,821	70.0%
Medicare Payment Exclusions ³		
Records from hospitals with data problems	0	0.0%
Duplicate records	1	<0.1%
Discharge date not in study period	0	0.0%
Rehabilitation revenue code in record	3	<0.1%
Invalid age, patients <18 or ≥ 120 years	642	3.6%
Patients treated in children's hospitals	45	0.3%
ICD-10-CM/PCS code-based exclusions	3,685	20.7%
Not MS-DRG or MDC of interest	498	2.8%
Missing or invalid discharge status	1	<0.1%
Patients who left against medical advice	14	0.1%
Patients transferred to another acute care hospital	43	0.2%
Out-of-state residents	1,224	6.9%
Not a CMS beneficiary	5,152	28.9%
Not Medicare entitled at time of discharge	161	0.9%
Not enrolled in Medicare FFS during month of discharge	2,725	15.3%
No matching CMS FFS payment	524	2.9%
CMS not the primary payer	164	0.9%
CMS FFS Payment <\$1,408 ⁴	68	0.4%
Included in Average Medicare Payment Calculation	2,847	16.0%

¹ The exclusions are listed in the order in which they were removed from the reference database.

² Records with the date of the first procedure of interest missing or occurs before admit date or after discharge date were excluded.

³ The Medicare Payments exclusions are based on CY 2020 data.

⁴ The Medicare Part A inpatient hospital deductible beneficiaries paid when admitted to the hospital in 2020, <https://www.cms.gov/newsroom/fact-sheets/2020-medicare-parts-b-premiums-and-deductibles>.

Table 2B. Exclusions¹ for Total Hip Replacement and Total Knee Replacement

	Total Hip Replacement		Total Knee Replacement	
	# Cases	% Cases	# Cases	% Cases
Total Cases Before Exclusions	15,186	100.0%	20,334	100.0%
Universal Exclusions				
Records from hospitals with data problems	0	0.0%	0	0.0%
Duplicate records	1	<0.1%	2	<0.1%
Discharge date not in study period	0	0.0%	0	0.0%
Rehabilitation revenue code in record	2	<0.1%	1	<0.1%
Invalid age, patients <18 or ≥ 120 years	13	0.1%	2	<0.1%
Patients treated in children's hospitals	0	0.0%	0	0.0%
Total Universal Exclusions	16	0.1%	5	<0.1%
Complication Exclusions				
Universal exclusions	16	0.1%	5	<0.1%
ICD-10-CM/PCS code-based exclusions	2,754	18.1%	2,252	11.1%
Not MS-DRG or MDC of interest	45	0.3%	72	0.4%
Missing or invalid discharge status	6	<0.1%	19	0.1%
Patients who left against medical advice	5	<0.1%	14	0.1%
Out-of-state residents	935	6.2%	1,137	5.6%
Missing or invalid patient identifier	5	<0.1%	8	<0.1%
Subsequent index hospitalization within 90 days	151	1.0%	241	1.2%
Included in Complication Measure	11,269	74.2%	16,586	81.6%
Extended Postoperative Length of Stay Exclusions				
Universal exclusions	16	0.1%	5	<0.1%
ICD-10-CM/PCS code-based exclusions	2,754	18.1%	2,252	11.1%
Not MS-DRG or MDC of interest	45	0.3%	72	0.4%
Missing or invalid discharge status	6	<0.1%	19	0.1%
Patients who left against medical advice	5	<0.1%	14	0.1%
Patients transferred to another acute care hospital	20	0.1%	27	0.1%
In-hospital mortality	2	<0.1%	6	<0.1%
Discharge to hospice (records with a discharge status code of 50 or 51)	2	<0.1%	3	<0.1%
Postoperative length of stay inconsistencies ²	552	3.6%	918	4.5%
Included in Extended PLOS Measure	11,784	77.6%	17,018	83.7%
Charge Exclusions				
Universal exclusions	16	0.1%	5	<0.1%
ICD-10-CM/PCS code-based exclusions	2,754	18.1%	2,252	11.1%
Not MS-DRG or MDC of interest	45	0.3%	72	0.4%
Missing or invalid discharge status	6	<0.1%	19	0.1%
Patients who left against medical advice	5	<0.1%	14	0.1%
Patients transferred to another acute care hospital	20	0.1%	27	0.1%
Invalid charges	12	0.1%	18	0.1%

(Continued on next page)

	Total Hip Replacement		Total Knee Replacement	
	# Cases	% Cases	# Cases	% Cases
Charge Exclusions (continued)				
ECMO/Tracheostomy MS-DRG 003 with a principal diagnosis from MDC 8 ³	0	0.0%	0	0.0%
Charge outliers	240	1.6%	294	1.4%
No reference data	84	0.6%	67	0.3%
Included in Average Charges Calculation	12,004	79.0%	17,566	86.4%
Medicare Payment Exclusions⁴				
Records from hospitals with data problems	0	0.0%	0	0.0%
Duplicate records	0	0.0%	3	<0.1%
Discharge date not in study period	0	0.0%	0	0.0%
Rehabilitation revenue code in record	2	<0.1%	2	<0.1%
Invalid age, patients <18 or ≥ 120 years	18	0.1%	4	<0.1%
Patients treated in children's hospitals	1	<0.1%	1	<0.1%
ICD-10-CM/PCS code-based exclusions	2,898	15.9%	2,485	10.4%
Not MS-DRG or MDC of interest	37	0.2%	47	0.2%
Missing or invalid discharge status	6	<0.1%	22	0.1%
Patients who left against medical advice	5	<0.1%	11	<0.1%
Patients transferred to another acute care hospital	23	0.1%	33	0.1%
Out-of-state residents	1,106	6.1%	1,273	5.3%
Not a CMS beneficiary	5,599	30.7%	6,680	28.0%
Not Medicare entitled at time of discharge	213	1.2%	324	1.4%
Not enrolled in Medicare FFS during month of discharge	3,743	20.5%	5,592	23.4%
No matching CMS FFS payment	596	3.3%	1,010	4.2%
CMS not the primary payer	236	1.3%	397	1.7%
CMS FFS Payment <\$1,408 ⁵	524	2.9%	756	3.2%
Included in Average Medicare Payment Calculation	3,248	17.8%	5,237	21.9%

¹ The exclusions are listed in the order in which they were removed from the reference database.

² Records with the date of the first procedure of interest missing or occurs before admit date or after discharge date were excluded.

³ Records assigned to MS-DRG 003 (ECMO or Tracheostomy with Mechanical Ventilation >96 Hours or Principal Diagnosis Except Face, Mouth, and Neck with Major O.R.) and MDC 8 (Diseases & Disorders of the Musculoskeletal System & Connective Tissue) were excluded.

⁴ The Medicare Payments exclusions are based on CY 2020 data.

⁵ The Medicare Part A inpatient hospital deductible beneficiaries paid when admitted to the hospital in 2020,

<https://www.cms.gov/newsroom/fact-sheets/2020-medicare-parts-b-premiums-and-deductibles>.

APPENDICES

Appendix A. Definitions—Study Populations, Code-based Exclusions and Complications

For each procedure included in the *Common Procedures Report*, the ICD-10-CM/PCS codes and MDC/MS-DRGs used to define study populations, code-based exclusions, and complications can be downloaded using the links below.

Spinal Fusion

<https://www.phc4.org/reports/commonprocedures/21/data/definitionsspinalfusion2021.xlsx>

Total Hip Replacement

<https://www.phc4.org/reports/commonprocedures/21/data/definitionshipreplacement2021.xlsx>

Total Knee Replacement

<https://www.phc4.org/reports/commonprocedures/21/data/definitionskneereplacement2021.xlsx>

Appendix B. Valid Discharge Status Codes

Code	Description
01	Discharged to home or self-care (routine discharge)
02	Discharged/transferred to a short-term general hospital for inpatient care
03	Discharged/transferred to skilled nursing facility (SNF) with Medicare certification in anticipation of skilled care
04	Discharged/transferred to a facility that provides custodial or supportive care
05	Discharged/transferred to a designated cancer center or children's hospital
06	Discharged/transferred to home under care of organized home health service organization in anticipation of covered skilled care
07	Left against medical advice (AMA) or discontinued care
20	Expired
21	Discharged/transferred to court/law enforcement
43	Discharged/transferred to a federal health care facility
50	Discharged to hospice—home
51	Discharged to hospice—medical facility (certified) providing hospice level of care
61	Discharged/transferred to a hospital-based Medicare approved swing bed
62	Discharged/transferred to an inpatient rehabilitation facility (IRF) including rehabilitation distinct part units of a hospital
63	Discharged/transferred to a Medicare certified long term care hospital (LTCH)
64	Discharged/transferred to a nursing facility certified under Medicaid but not certified under Medicare
65	Discharged/transferred to a psychiatric hospital or psychiatric distinct part unit of a hospital
66	Discharged/transferred to a critical access hospital (CAH)
69	Discharged/transferred to a designated disaster alternative care site
70	Discharged/transferred to another type of health care institution not defined elsewhere in this code list
81	Discharged to home or self care (routine discharge) with a planned acute care hospital inpatient readmission
82	Discharged/transferred to a short term general hospital for inpatient care with a planned acute care hospital inpatient readmission
83	Discharged/transferred to a skilled nursing facility with Medicare certification in anticipation of skilled care with a planned acute care hospital inpatient readmission
84	Discharged/transferred to a facility that provides custodial or supportive care with a planned acute care hospital inpatient readmission
85	Discharged/transferred to a designated cancer center or children's hospital with a planned acute care hospital inpatient readmission
86	Discharged/transferred to home under care of organized home health service organization in anticipation of covered skilled care with a planned acute care hospital inpatient readmission
87	Discharged/transferred to court/law enforcement with a planned acute care hospital inpatient readmission
88	Discharged/transferred to a federal health care facility with a planned acute care hospital inpatient readmission
89	Discharged/transferred to a hospital-based Medicare approved swing bed with a planned acute care hospital inpatient readmission
90	Discharged/transferred to an inpatient rehabilitation facility including rehabilitation distinct part units of a hospital with a planned acute care hospital inpatient readmission
91	Discharged/transferred to a Medicare certified long term care hospital with a planned acute care hospital inpatient readmission
92	Discharged/transferred to a nursing facility certified under Medicaid but not certified under Medicare with a planned acute care hospital inpatient readmission
93	Discharged/transferred to a psychiatric hospital or psychiatric distinct part unit of a hospital with a planned acute care hospital inpatient readmission
94	Discharged/transferred to a critical access hospital with a planned acute care hospital inpatient readmission
95	Discharged/transferred to another type of health care institution not defined elsewhere in this code list with a planned acute care hospital inpatient readmission

Appendix C. Example of Logistic Regression

Calculations Used in Determining Expected In-Hospital Complication Rates for a Given Hospital Spinal Fusion

Total Cases:	Number of hospitalizations for a hospital after exclusions (equal to n).
Actual Percent Complication:	Total number of in-hospital complications / total number of hospitalizations.
Expected Percent Complication:	Mean of the predicted probability of in-hospital complication for each hospitalization (PComp).

Step 1: Calculate the predicted probability of in-hospital complication for each hospitalization (PComp):

$$\begin{aligned} \beta X &= (\beta_0 + \text{TimeFactor}) + \beta_1 X_1 + \dots + \beta_{10} X_{10} + \dots + \beta_{13} X_{13} \\ &= (-6.2973) + (0.0225)(X_1) + \dots + (0.6037)(X_{10}) + \dots + (1.1317)(X_{13}). \end{aligned}$$

Where:

X_1 = Age

...

X_{10} = Chronic Lung Disease (1 if true, 0 if false)

...

X_{13} = Alcohol-related Disorder: Dependence or Abuse (1 if true, 0 if false)

β 's are the regression coefficients that correspond to each risk factor (X).

A time factor (TimeFactor) is added to the intercept so the statewide expected rate for the current study period was equal to the actual statewide rate for this same period.

$$PComp = \frac{e^{\beta X}}{1 + e^{\beta X}}$$

where $e \approx 2.7182818285$

Step 2: Calculate the mean PComp for a hospital (expected percent of in-hospital complications):

$$\text{Mean PComp} = \frac{\sum PComp}{n}$$

Appendix D. Example of Linear Regression

Calculations Used in Determining Extended Postoperative Length of Stay (EPLOS) Rates for a Given Hospital Spinal Fusion

Total Cases: Number of hospitalizations for a hospital after exclusions (equal to n).

Actual Percent EPLOS: Total number of hospitalizations with EPLOS / total number of hospitalizations.

Step 1: Calculate the predicted log transformed postoperative length of stay for each hospitalization (PPLOS):

$$\beta X = (\beta_0 + \text{TimeFactor}) + \beta_1 X_1 + \dots + \beta_{15} X_{15} + \dots + \beta_{42} X_{42}$$

$$= (0.4595) + (0.0232)(X_1) + \dots + (0.4843)(X_{15}) + \dots + (0.0522)(X_{43})$$

Where:

X_1 = Female (1 if true, 0 if false)

...

X_{15} = Malnutrition (1 if true, 0 if false)

...

X_{43} = Anticipated Primary Payer: Medicare, Medicaid, Self-Pay (1 if true, 0 if false)

β 's are the regression coefficients that correspond to each risk factor (X).

A time factor (TimeFactor) is added to the intercept so the statewide expected average log PLOS for the current study period is equal to the actual statewide average log PLOS for this same period.

Step 2: Calculate the residual log transformed postoperative length of stay (*Residual*) as actual log transformed postoperative length of stay (*APLOS*) minus *PPLOS* for each hospitalization.

$$\text{Residual} = (\text{APLOS} - \text{PPLOS})$$

Step 3: Calculate statewide statistics.

Step 3a: Calculate the average *Residual* using statewide records.

$$\text{AverageResidual}(\overline{\text{Residual}}) = \frac{\sum_{i=1}^n (\text{Residual}_i)^2}{n}$$

Step 3b: Calculate the standard deviation of *Residual* using statewide records.

$$\text{StandardDeviationofResidual}(SD\text{Residual}) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\text{Residual}_i - \overline{\text{Residual}})^2}$$

Step 4: An *EPLOS* is counted when the *Residual* exceeds two standard deviations above the mean *Residual*.

$$\text{EPLOSifResidual} > (\overline{\text{Residual}} + 2 * SD\text{Residual})$$

Expected Percent EPLOS Total number of hospitalizations with EPLOS statewide / total number of hospitalizations statewide. The expected percent EPLOS for each hospital is the same as the statewide rate of EPLOS.

Appendix E. Example of Case-Mix Adjustment

**Calculations Used in Determining Average Charge for a Hospital
 Example Hospital: Hospital "A" in Southwestern PA, Region 1
 Total Knee Replacement**

- Total Cases:** Number of hospitalizations for Hospital A after exclusions (equal to n).
- Actual Average Charge, Hospital:** Mean of the charges among all hospitalizations for Hospital A (Average ActChg).
- Actual Average Charge, Region:** Mean of the charges among all hospitalizations for the hospital region (Region 1).
- Expected Average Charge, Hospital:** Mean of the expected charges among all hospitalizations for Hospital A (Average ExpChg).

Step 1: Calculate each hospitalization's expected charge (ExpChg):

ExpChg is based on the MS-DRG of the hospitalization and is equal to the average charge for all hospitalizations (after exclusion) in the hospital's same region and MS-DRG group.

Region 1 – Southwestern PA:

- MS-DRG 462: \$87,002
- MS-DRG 469: \$90,960
- MS-DRG 470: \$50,027

Step 2: Calculate the average expected charge for Hospital A (ExpChg):

$$\text{Average ExpChg} = \frac{\sum \text{ExpChg}}{n}$$

Case-Mix Adjusted Charge: $\frac{\text{Average ActChg}}{\text{Average ExpChg}}$ (Region 1 Actual Average Charge)

Nine Pennsylvania Regions:

1. Southwestern – Allegheny, Armstrong, Beaver, Butler, Fayette, Greene, Washington, and Westmoreland counties
2. Northwestern – Cameron, Clarion, Clearfield, Crawford, Elk, Erie, Forest, Jefferson, Lawrence, McKean, Mercer, Potter, Venango, and Warren counties
3. Southern Allegheny – Bedford, Blair, Cambria, Indiana, and Somerset counties
4. Northcentral – Centre, Clinton, Columbia, Lycoming, Mifflin, Montour, Northumberland, Snyder, Tioga, and Union counties
5. Southcentral – Adams, Cumberland, Dauphin, Franklin, Fulton, Huntingdon, Juniata, Lancaster, Lebanon, Perry, and York counties
6. Northeastern – Bradford, Lackawanna, Luzerne, Monroe, Pike, Sullivan, Susquehanna, Wayne, and Wyoming counties
7. Eastcentral – Berks, Carbon, Lehigh, Northampton, and Schuylkill counties
8. Southeastern – Bucks, Chester, Delaware, and Montgomery counties
9. Philadelphia – Philadelphia County