TECHNICAL NOTES
for Common Procedures Report

August 2019

Pennsylvania Health Care Cost Containment Council

Report Period:
April 1, 2016 through June 30, 2018
- CABG (without valve)

July 1, 2017 through June 30, 2018
- Knee Replacement
- Hip Replacement
- Spinal Fusion

225 Market Street, Suite 400, Harrisburg, PA 17101
Phone: 717-232-6787
Fax: 717-232-3821
www.phc4.org

Joe Martin, Executive Director

TABLE OF CONTENTS

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.[1]

The procedures and measures outlined below are reported for general acute care (GAC) hospitals that typically perform these procedures on adults. In addition to the measures listed below, the *total number of cases* (prior to exclusions) is reported for hospitals with at least one patient 18 years or older who underwent the procedure during the report period.[2]

- **CABG** *(coronary artery bypass graft without a valve procedure on the same day)*

  Report period: April 1, 2016 through June 30, 2018 discharges

  Measures reported for hospitals with five or more cases (includes patients 30 years and older):

  - Risk-adjusted in-hospital mortality rating
  - Risk-adjusted 30-day readmission for complication rating
  - Average hospital charge (case-mix adjusted)

- **Knee Replacement and Hip Replacement**

  Report period: July 1, 2017 through June 30, 2018 discharges

  Measures reported for hospitals with 5 or more cases (includes patients 18 years and older):

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

- **Spinal Fusion**

  Report period: July 1, 2017 through June 30, 2018 discharges

  Measures reported for hospitals with 5 or more cases (includes patients 18 years and older):

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

- **Average Medicare Payment**

  The overall statewide average payment for Medicare fee-for-service patients is reported for each procedure and each MS-DRG within a given procedure—to account for variations in case mix. The average payment is reported for July 1, 2017 through June 30, 2018 discharges.

The rigorous 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 surgical procedures reported.

[1] Statewide utilization and outcome data are displayed in Table 1 and exclusion data in Tables 2 to 4. See Appendix A for ICD-10-CM/PCS codes and MS-DRGs defining procedures, exclusions and complications.
[2] Results are not displayed for hospitals that closed or merged into 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 GAC 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 also submitted laboratory test results. In addition, hospitals that performed cardiac surgery submitted supplemental clinical data.

Laboratory test results were submitted by hospitals to the Council for a select group of acute care inpatient records, including those used for analysis of the hospitalizations 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).

For inpatient discharges of adult patients in which a CABG and/or valve surgery[1] was performed, hospitals submitted supplemental clinical data to the Council. Hospitals were required to submit the following clinical data elements related to the first CABG/valve surgery of the admission: anesthesia start date and start time, American Society of Anesthesiologists (ASA) class, the ASA emergency indicator, ejection fraction and percent stenosis in the coronary arteries and their branches. The requirements for submitting this data are specified elsewhere (refer to PHC4's *Cardiac Surgery Supplemental Clinical Data Reporting Manual*, accessible at www.phc4.org).

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 and supplemental clinical data submissions 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 the ultimate responsibility of each individual hospital.

Hospital and Cardiothoracic Surgeon Verification of Cardiac Surgery Data

Hospitals were asked to confirm the accuracy of discharge records, provide additional ICD-10-CM/PCS diagnosis and procedure codes as appropriate and confirm that cases had the correct surgeon assignment. Surgeons were asked to perform a patient level review of the submitted records and then attest to the accuracy of the data and the surgeon assignment. Hospitals and/or surgeons had the opportunity to request special exclusions for cases in which the patient's outcome was most directly associated with conditions unrelated to the CABG and/or valve surgical episode or the care received during that hospitalization that were not accounted for through risk adjustment. The medical record documentation was reviewed to determine whether special requests for exclusion would be granted. In addition, because of their importance as risk factors, hospitals and surgeons had the opportunity to submit medical record documentation for cases in which cardiogenic shock and/or acute renal failure were present at the time of or immediately prior to the first CABG/valve surgery. This record documentation was reviewed to verify that the criteria for preoperative cardiogenic shock and/or preoperative acute renal failure were met. The requirements for submitting cases for medical record

[1] The current version of the *Common Procedures Report* includes outcomes for adult patients who underwent CABG (without a valve procedure on the same day). Valve surgeries with and without CABG are not included in this report.

documentation review are specified elsewhere (refer to PHC4's *Guide for Review and Attestation of Cardiac Surgery Data*, accessible at www.phc4.org).

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 lab 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 POPULATIONS

The study population for each procedure reported is designed to represent a clinically cohesive group of patients. See Appendix A for ICD-10-CM/PCS diagnosis and 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 GAC hospitals during the defined report period with an applicable ICD-10-PCS procedure code(s) in either the principal or secondary procedure code positions of the discharge record.

Exclusion Criteria

Clinically complex cases were excluded from all outcome analyses. These atypical cases were defined by ICD-10-CM/PCS diagnosis and procedure codes and MS-DRG and MDC combinations *not* in the study population definition. For the CABG population, clinically complex cases also included cases granted special request for exclusion.

Additional exclusions were measure-specific, such as cases with insufficient data for the measure analyzed or cases not applicable to the measure analyzed. For example, patients that died during the hospitalization in which a given procedure was performed were excluded from the readmission for complication analysis but not the in-hospital mortality analysis.

See Tables 2 to 4 for the types of exclusions applied for each measure and the number of cases excluded.

MEASURES REPORTED

Total Number of Cases

Reported for each procedure included in the report, the total number of cases included all hospitalizations for patients 18 years and older (prior to clinically complex and other case exclusions). If two surgeries were performed for a given procedure during the same hospitalization, the case was only counted once.

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

Measures with Risk-Adjusted/Statistical Ratings

Risk-adjusted ratings are reported for in-hospital mortality and the complication measures. For extended postoperative length of stay, the postoperative length of stay was risk adjusted. 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 B and C for examples). Ratings are reported for hospitals with five or more cases in the analysis.

In-Hospital Mortality *(Reported for CABG)*. The in-hospital mortality analysis included patients 30 years and older. The mortalities were identified in the patient discharge record as a discharge status of "20."

30-Day Readmission for Complication *(Reported for CABG)*. A readmission for complication following CABG surgery is defined as a rehospitalization to a Pennsylvania GAC hospital *within 1 to 30 days* of discharge from the hospitalization in which the CABG surgery was performed (also referred to as the index hospitalization) *with a principal diagnosis* that indicated a complication following CABG surgery.[1]

The 30-day readmission for complication measure is a dichotomous (yes/no) outcome; as such, it is counted only once when an index hospitalization results in multiple 30-day readmissions for complication. If, over the study period, a patient had multiple discharges for CABG surgery, each discharge was independently investigated to determine whether it had a 30-day readmission for complication.

Complication *(Reported for Knee Replacement and Hip Replacement)*. The complication measure is based, in large part, on the Centers for Medicare and Medicaid Services measure designed for total knee and hip replacements likely to be considered elective.[2] 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) one of the ICD-10-CM complication codes listed for knee or hip replacement was a secondary diagnosis that 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 of "20." [1]

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

[2] Centers for Medicare and Medicaid Services. "2017 Procedure-Specific Measure Updates and Specifications Report: Hospital-Level Risk-Standardized Complication Measure, Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA) – Version 6.0." March 2017. Available at http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Measure-Methodology.html.

or

- *occurred as the principal diagnosis of a readmission* to a Pennsylvania GAC hospital *within 7, 30, or 90 days* (depending on the type of complication) of discharge from the index hospitalization. For certain complications the principal diagnosis code is paired with a procedure code. [1]

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 knee or hip replacement, each discharge was independently investigated to determine whether it had a complication with one exception. If a second hospitalization for knee or hip replacement occurred within 90 days of the first index hospitalization, the second hospitalization was excluded from the complication 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) one of the ICD-10-CM complication codes listed for spinal fusion was a secondary diagnosis that 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 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 GAC hospital *within 7, 30, or 90 days* (depending on the type of complication) of the index hospitalization *with a principal diagnosis* that indicated a complication following spinal fusion. 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.

Extended Postoperative Length of Stay *(Reported for Knee Replacement, Hip Replacement and Spinal Fusion)*. In general terms, an extended postoperative length of stay (PLOS) evaluates whether the actual length of time patients remain in the hospital following the procedure was significantly longer than what would be expected, after accounting for patients' risk. The development of this new measure was guided, in part, by the approach used by Michael Pine and Associates. [2]

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 C for example):

- The actual PLOS (in days) is 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.

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

[2] 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.

- 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]
- An extended PLOS was counted when the difference between the actual and expected PLOS for a particular patient was significantly higher than the average difference (between actual and expected PLOS) for all patients in the analysis.

  In statistical terms, an extended PLOS is counted when the residual log transformed PLOS (i.e., the actual log transformed PLOS minus the predicted 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

Reported for each procedure included in the report, the charge is the amount a hospital bills for a patient's care and includes hospital charges for the entire hospitalization during which a given procedure was performed (not just the treatment associated with surgery). 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

Reported for each procedure included in the report, the statewide average payment is reported for Medicare fee-for-service (FFS) patients (Pennsylvania residents). The Medicare payment data was provided to PHC4 by the Centers for Medicare and Medicaid Services (CMS) and then matched by PHC4 to cases meeting study population criteria for a given procedure and discharged during state fiscal year 2018 (July 1, 2017 through June 30, 2018). Average payments were calculated using the claim payment amounts obtained from the CMS payment data file. Patient liabilities (e.g., coinsurance and deductible dollar amounts) were not included. Payments from Medicare Advantage plans (e.g., Medicare HMOs) were not included.

For each procedure, the statewide average payment is reported overall and by MS-DRG—to account for variations in case mix. The number of cases included in the average payment is also displayed. Average payment is not reported for hospitals. To meet current CMS privacy guidelines, average payments (and the number of cases included in the average payment) are only displayed for MS-DRGs with 11 or more cases.

[1] A natural log transformation of each PLOS was performed to account for skewness in the PLOS distribution. Determination of expected PLOS and calculation of differences between actual and expected were calculated using the actual log PLOS and the expected log PLOS.

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., in-hospital mortality and complication) occurring, or 2) for, the extended postoperative length of stay (PLOS) measure, predicting the log transformed PLOS[1]. 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, taking into account the risk factors that were included in the risk-adjustment models.

Model Development

The reference datasets used to build models for procedure/measure combinations included two years of data, discharges from Quarter 4, 2015 through Quarter 3, 2017 meeting the study population/measure inclusion and exclusion criteria.

Identifying potential risk factors. The first step in building the models was to identify risk factors that potentially contributed to the event or outcome (i.e., in-hospital mortality, 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, supplemental clinical data for CABG, 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.

[1] A natural log transformation of each PLOS was performed to account for skewness in the PLOS distribution. Determination of expected PLOS and calculation of differences between actual and expected were calculated using the actual log PLOS and the expected log PLOS.

- 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: Variables were required to have significance ($p<0.10$) and, for categorical variables, meet the Schwarz criterion. Exceptions were made to these criteria when a variable had particular clinical relevance to the outcome.

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 knee and hip replacement, which were modeled together. Binary logistic regression was used for analyses of the in-hospital mortality and complication measures. Linear regression was used for PLOS analysis.

Model selection. Risk factors selected for testing were added to the model in the following order: 1) procedure group (knee or hip replacement models only), 2) patient characteristics (gender, race/ethnicity, age) and socioeconomic factors (poverty rate, education, percent not speaking English very well), 3) supplemental clinical data (for CABG models only), 4) laboratory test results, 5) ICD-10-CM/PCS code-based variables, then 5) other UB-04-derived data elements (e.g., 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, one hundred 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 (reflective of an increased risk) in at least 75% of the sample models.

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:

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

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 and 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., mortality and 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 B for an example of logistic regression. See Appendix C 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:

CABG	In-hospital mortality percent
	30-day readmission for complication percent
Knee Replacement	Complication percent
	Extended postoperative length of stay percent
Hip Replacement	Complication percent
	Extended postoperative length of stay percent
Spinal Fusion	In-hospital complication percent
	Readmission for complication percent
	Extended postoperative length of stay percent

Determining Expected (Predicted) Values

For the mortality and 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., mortality or 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, $0 \le a \le N$).

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 three PA regions (Western, Central and Northeastern, and Southeastern) and also the nine smaller PA regions depending on MS-DRG grouping. When the study population for a procedure included more than four quarters of data, cases were assigned to a time period/PA region/MS-DRG group combination.

For example, the reference database to adjust charges for CABG cases was constructed with cases assigned to one of two time periods based on discharge date, Quarter 2, 2016 to Quarter 1, 2017 *or* Quarter 2, 2017 to Quarter 2, 2018. Each time period was comprised of two subsets of PA region/MS-DRG group combinations:

- MS-DRGs 228, 229/230, 231 and 232 assigned to one of the large PA regions (Western, Central and Northeastern, and Southeastern).
- MS-DRGs 233, 234, 235 and 236 assigned to one of the nine PA regions.

Low volume (<10 cases) PA region/MS-DRG group combinations were excluded from the database. Trimming was then performed to remove outliers from each combination.

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 group combination 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 (75^{th} percentile total charge) of all patient records from the comparative database in a particular category*

IQR = *Q3 – Q1*

Lower Trim Point = *Q1 – (3.0 x IQR)*

Upper Trim Point = *Q3 + (3.0 x 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:

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

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 D for an example of how case-mix adjusted charges were computed.

DATA TABLES

Table 1. Statewide Utilization and Outcome Data, by Procedure

CABG	
Total Number of Cases	18,379
In-Hospital Mortality	1.5%
30-Day Readmission for Complication	3.6%
Average Hospital Charge	$197,460

Knee Replacement	
Total Number of Cases	39,269
Complication	1.5%
Extended PLOS	1.9%
Average Hospital Charge	$52,660

Hip Replacement	
Total Number of Cases	24,645
Complication	1.9%
Extended PLOS	2.2%
Average Hospital Charge	$55,444

Spinal Fusion	
Total Number of Cases	19,858
In-Hospital Complication	1.8%
Readmission for Complication	2.0%
Extended PLOS	2.5%
Average Hospital Charge	$118,163

Note: The total number of cases (prior to exclusions) is reported for hospitals with one or more patients 18 years and older who underwent a given procedure during the report period.

Tables 2 to 4. Exclusions from Analyses, by Procedure and Measure

Table 2. Exclusions for CABG

	# Cases	% Cases
Total Number of Cases (18 years and older)	**18,379**	**100.0%**
Exclusions from In-Hospital Mortality Analysis		
Patients < 30 years	8	<0.1%
Left against medical advice	12	0.1%
Clinically complex cases[1]	1,537	8.4%
Included in-hospital mortality analysis	**16,822**	**91.5%**
Exclusions from 30-Day Readmission for Complication Analysis		
Patients < 30 years	8	<0.1%
Left against medical advice	12	0.1%
Clinically complex cases[1]	1,537	8.4%
Patients who died	248	1.3%
Missing or invalid social security number	378	2.1%
Out-of-state residents[2]	1,338	7.3%
Included in 30-day readmission for complication analysis	**14,858**	**80.8%**
Exclusions from Average Hospital Charge Analysis		
Patients < 30 years	8	<0.1%
Left against medical advice	12	0.1%
Clinically complex cases[1]	1,537	8.4%
Invalid charges	0	0.0%
ECMO/Tracheostomy (MS-DRG 003 and MDC 5)[3]	255	1.4%
Low volume MS-DRGs[4]	153	0.8%
Charge outliers[5]	223	1.2%
Included in average charge analysis	**16,191**	**88.1%**

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

[1] Clinically complex cases included cases with an ICD-10-CM/PCS code found in the Definitions link in Appendix A. Additional exclusions included cases *not* in the study MS-DRGs (See Appendix A: Definitions—Study Populations, Exclusions and Complications) and cases granted special request for exclusion.

[2] Out-of-state residents were excluded because such patients could undergo a a given procedure in a Pennsylvania hospital, return to their state of residence and be readmitted to a hospital in their home state. Therefore, readmission data would not be available for these patients.

[3] Cases 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 5 (Diseases and Disorders of the Circulatory System) were excluded.

[4] MS-DRGs with low volume, including MS-DRG groups when a particular combination of timeframe/PA region/MS-DRG group had fewer than 10 cases.

[5] Charge outliers were determined using the "+/- 3.0 interquartile range" method—after accounting for differences in charges by timeframe/PA region/ MS-DRG group.

Table 3. Exclusions for Knee Replacement and Hip Replacement

	Knee		Hip	
	# Cases	% Cases	# Cases	% Cases
Total Number of Cases (18 years and older)	**39,269**	**100.0%**	**24,645**	**100.0%**
Exclusions from Complication Analysis				
Clinically complex cases[1]	2,766	7.0%	2,912	11.8%
Invalid discharge status	22	0.1%	9	<0.1%
Left against medical advice	7	<0.1%	8	<0.1%
Missing or invalid social security number	666	1.7%	442	1.8%
Out-of-state residents[2]	1,757	4.5%	1,528	6.2%
Subsequent index within 90 days	528	1.3%	335	1.4%
Included in complication analysis	**33,523**	**85.4%**	**19,411**	**78.8%**
Exclusions from Extended Postoperative Length Of Stay (EPLOS) Analysis				
Clinically complex cases[1]	2,766	7.0%	2,912	11.8%
Invalid discharge status	22	0.1%	9	<0.1%
Patients who died/discharged to hospice	19	<0.1%	3	<0.1%
Left against medical advice	7	<0.1%	8	<0.1%
Transferred to acute care	67	0.2%	31	0.1%
Invalid date needed for calculation	273	0.7%	30	0.1%
Included in EPLOS analysis	**36,115**	**92.0%**	**21,652**	**87.9%**
Exclusions from Average Hospital Charge Analysis				
Clinically complex cases[1]	2,766	7.0%	2,912	11.8%
Invalid charges	1	<0.1%	0	0.0%
ECMO/Tracheostomy (MS-DRG 003 and MDC 8)[3]	1	<0.1%	1	<0.1%
Charge outliers[4]	820	2.1%	529	2.1%
Low volume MS-DRGs[5]	23	0.1%	14	0.1%
Included in average charge analysis	**35,658**	**90.8%**	**21,189**	**86.0%**

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

[1] Clinically complex cases included cases with an ICD-10-CM/PCS code found in the Definitions link in Appendix A and not in MDC 8. Additional exclusions included cases *not* in the study MS-DRGs (See Appendix A: Definitions—Study Populations, Exclusions and Complications).

[2] Out-of-state residents were excluded because such patients could undergo a given procedure in a Pennsylvania hospital, return to their state of residence and be readmitted to a hospital in their home state. Therefore, readmission data would not be available for these patients.

[3] Cases 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 and Disorders of the Musculoskeletal System and Connective Tissue) were excluded.

[4] Charge outliers were determined using the "+/- 3.0 interquartile range" method—after accounting for differences in charges by procedure/PA region/MS-DRG group.

[5] MS-DRGs with low volume, including MS-DRG groups when a particular combination of procedure/PA region/MS-DRG group had fewer than 10 cases.

Table 4. Exclusions for Spinal Fusion

	# Cases	% Cases
Total Number of Cases (18 years and older)	**19,858**	**100.0%**
Exclusions from In-Hospital Complication Analysis		
Clinically complex cases[1]	4,485	22.6%
Invalid discharge status	3	<0.1%
Included in in-hospital complication analysis	**15,370**	**77.4%**
Exclusions from Readmission for Complication Analysis		
Clinically complex cases[1]	4,485	22.6%
Invalid discharge status	3	<0.1%
Patients who died	14	0.1%
Left against medical advice	14	0.1%
Missing or invalid social security number	324	1.6%
Out-of-state residents[2]	1,275	6.4%
Subsequent index within 90 days	57	0.3%
Included in readmission for complication analysis	**13,686**	**68.9%**
Exclusions from Extended Postoperative Length Of Stay (EPLOS) Analysis		
Clinically complex cases[1]	4,485	22.6%
Invalid discharge status	3	<0.1%
Left against medical advice	14	0.1%
Patients who died/discharged to hospice	19	0.1%
Transferred to acute care	30	0.2%
Invalid date needed for calculation	176	0.9%
Included in EPLOS analysis	**15,131**	**76.2%**
Exclusions from Average Hospital Charge Analysis		
Clinically complex cases[1]	4,485	22.6%
Invalid charges	1	<0.1%
Charge outliers[3]	254	1.3%
Low volume MS-DRGs[4]	3	<0.1%
Included in average charge analysis	**15,115**	**76.1%**

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

[1] Clinically complex cases included cases with an ICD-10-CM/PCS code found in the Definitions link in Appendix A and not in MDC 8. Additional exclusions included cases *not* in the study MS-DRGs (See Appendix A: Definitions—Study Populations, Exclusions and Complications).

[2] Out-of-state residents were excluded because such patients could undergo a given procedure in a Pennsylvania hospital, return to their state of residence and be readmitted to a hospital in their home state. Therefore, readmission data would not be available for these patients.

[3] Charge outliers were determined using the "+/- 3.0 interquartile range" method—after accounting for differences in charges by PA region/ MS-DRG group.

[4] MS-DRGs with low volume, including MS-DRG groups when a particular combination of PA region/MS-DRG group had fewer than 10 cases.

APPENDICES

Appendix A. Definitions—Study Populations, Exclusions and Complications

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

CABG *(coronary artery bypass graft without a valve procedure on the same day)*

http://www.phc4.org/reports/commonprocedures/18/data/DefinitionsCABG-2016Q2_2018Q2.xlsx

Knee Replacement

http://www.phc4.org/reports/commonprocedures/18/data/DefinitionsKneeReplacement-2017Q3_2018Q2.xlsx

Hip Replacement

http://www.phc4.org/reports/commonprocedures/18/data/DefinitionsHipReplacement-2017Q3_2018Q2.xlsx

Spinal Fusion

http://www.phc4.org/reports/commonprocedures/18/data/DefinitionsSpinalFusion-2017Q3_2018Q2.xlsx

Appendix B. Example of Logistic Regression

Calculations Used in Determining Expected In-Hospital Mortality Rates for a Given Hospital
CABG

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

Actual Percent Mortality: Total number of deaths (death is a discharge status equal to 20)/ total number of hospitalizations.

Expected Percent Mortality: Mean of the predicted probability of death for each hospitalization (PDeath).

Step 1: Calculate the predicted probability of death for each hospitalization (PDeath):

$$\beta X = (\beta_0 + \text{TimeFactor}) + \beta_1 X_1 + \ldots + \beta_{10} X_{10} + \ldots + \beta_{13} X_{13} + \ldots + \beta_{17} X_{17}$$
$$= (-10.1015) + (0.0583)(X_1) + \ldots + (0.7286)(X_{10}) + \ldots + (0.5008)(X_{13}) + \ldots + (0.6905)(X_{17}).$$

Where:

X_1 = Age

…

X_{10} = Preoperative Cardiogenic Shock (1 if true, 0 if false)

…

X_{13} = Heart Failure (1 if true, 0 if false)

…

X_{17} = Morbid Obesity (1 if true, 0 if false)

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

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

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

where $e \approx 2.7182818285$

Step 2: Calculate the mean PDeath for a hospital (expected percent of deaths):

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

Appendix C. 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):

$$\beta X = (\beta_0 + \text{TimeFactor}) + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots \beta_{43} X_{43}$$
$$= (0.7849) + (0.0361)(X_1) + (0.1643)(X_2) + (0.7486)(X_3) + \ldots (0.0762)(X_{43})$$

Where:
X_1 = Female (1 if true, 0 if false)
X_2 = Black (non-Hispanic) (1 if true, 0 if false)
X_3 = % Population w/in Patient's Zip Code Not Speaking English Well
…
X_{43} = Medicaid is Anticipated Primary Payer (1 if true, 0 if false)

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

A time factor (TimeFactor) is used to get the adjusted intercept so that the statewide expected average log PLOS for the current study period was 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.

$$Residual = (APLOS - PPLOS)$$

$$\text{Average } Residual \left(\overline{Residual}\right) = \frac{\sum_{i=1}^{n} (Residual_i)^2}{n}$$

$$\text{Standard Deviation of } Residual \ (SDResidual) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} (Residual_i - \overline{Residual})^2}$$

Step 3: An *EPLOS* is counted when the *Residual* exceeds two standard deviations above the mean *Residual*.

$$EPLOS \text{ if } Residual > (\overline{Residual} + 2 * SDResidual)$$

Expected Percent EPLOS Total number of hospitalizations with EPLOS statewide / total number of hospitalizations statewide

Appendix D. Example of Case-Mix Adjustment

Region 1 – Southwestern PA
Knee Replacement

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

Actual Average Charge, Hospital: Mean of the charges among all hospitalizations for a hospital (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 a hospital (Average ExpChg).

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

ExpChg is the expected charge for a 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 461 or 462: $66,061
or
MS-DRG 469: $73,794
or
MS-DRG 470: $44,476

Step 2: Calculate the average expected charge for a hospital (ExpChg):

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

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