Accepted Manuscript

Development and Evaluation of the American College of Surgeons NSQIP Pediatric Surgical Risk Calculator

Kari Kraemer, PhD, Mark E. Cohen, PhD, Yaoming Liu, PhD, Douglas C. Barnhart, MD, Shawn J. Rangel, MD, MSCE, Jacqueline M. Saito, MD, MSCI, FACS, Karl Y. Bilimoria, MD, MS, FACS, Clifford Y. Ko, MD, MS, MSHS, FACS, Bruce L. Hall, MD, PhD, MBA, FACS

PII: S1072-7515(16)31285-6
DOI: 10.1016/j.jamcollsurg.2016.08.542
Reference: ACS 8478

To appear in: Journal of the American College of Surgeons

Received Date: 12 July 2016
Revised Date: 16 August 2016
Accepted Date: 16 August 2016


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Development and Evaluation of the American College of Surgeons NSQIP Pediatric Surgical Risk Calculator

Kari Kraemer, PhD¹, Mark E Cohen, PhD¹, Yaoming Liu, PhD¹, Douglas C Barnhart, MD², Shawn J Rangel, MD, MSCE³, Jacqueline M Saito, MD, MSCI, FACS⁴, Karl Y Bilimoria, MD, MS, FACS⁵,6, Clifford Y Ko, MD, MS, MSHS, FACS¹,7, Bruce L Hall, MD, PhD, MBA, FACS¹,4,8

1. Division of Research and Optimal Patient Care, American College of Surgeons, Chicago, IL
2. Division of Pediatric Surgery, Primary Children’s Medical Center, University of Utah, Salt Lake City, UT.
3. Associate in Surgery, Boston Children’s Hospital; Associate Professor of Surgery, Harvard Medical School, Boston, MA.
4. Department of Surgery, Washington University in St Louis, MO.
5. Surgical Outcomes and Quality Improvement Center (SOQIC), Department of Surgery and Center for Healthcare Studies, Feinberg School of Medicine, Chicago, IL.
6. Northwestern Medicine, Northwestern University, Chicago, IL.
7. Department of Surgery, University of California Los Angeles David Geffen School of Medicine and the VA Greater Los Angeles Healthcare System, Los Angeles, CA
8. Center for Health Policy and the Olin Business School at Washington University in St Louis; John Cochran Veterans Affairs Medical Center; and BJC Healthcare, St Louis, MO.

Disclosure Information: Nothing to disclose.
Correspondence address:
Kari Kraemer
American College of Surgeons
633 N. Saint Clair St.
22nd Floor
Chicago, IL 60611-3211
Phone: (312) 202-5570
Fax: (312) 202-5062
Email: kkraemer@facs.org

Running Title: NSQIP Pediatric Risk Calculator
ABSTRACT

Background: There is an increased desire of patients and families to be involved in the surgical decision making process. A surgeon’s ability to provide patients and families with patient-specific estimates of postoperative complications is critical for shared decision making and informed consent. Surgeons can also use patient-specific risk estimates to decide whether or not to operate and what options to give to patients. Our objective was to develop and evaluate a publicly available risk estimation tool that would cover many common pediatric surgical procedures across all specialties.

Study Design: ACS NSQIP Pediatric standardized data from 67 hospitals were used to develop a risk estimation tool. Surgeons enter 18 preoperative variables (demographics, comorbidities, procedure) that are used in a logistic regression model to predict nine postoperative outcomes. A surgeon adjustment score is also incorporated to adjust for any additional risk not accounted for in the 18 risk factors.

Results: A pediatric surgical risk calculator was developed based on 181,353 cases covering 382 CPT codes across all specialties. It had excellent discrimination for mortality (c-statistic=0.98), morbidity (c-statistic=0.81), and 7 additional complications (c-statistics>0.77). The HL statistic and graphical representations also showed excellent calibration.

Conclusions: The ACS NSQIP Pediatric Surgical Risk Calculator was developed using standardized and audited, multi-institutional data from the ACS NSQIP Pediatric and provides empirically derived, patient-specific postoperative risks. It can be used as a tool in the shared decision making process by providing clinicians, families, and patients with useful information for many of the most common operations performed on pediatric patients in the U.S.
Key words: ACS NSQIP; Pediatric; Surgical Risk Calculator; Calibration
INTRODUCTION

Informed consent in pediatric surgery involves a shared decision making process between the surgeons and the family of the child.\textsuperscript{1-4} Shared decision making has the potential to increase satisfaction with care, reduce decisional conflict and regret, improve understanding of and participation in care, and thereby improve health related quality of life.\textsuperscript{5} Families report having less decisional conflict when they are involved in the decision making process.\textsuperscript{6} For these reasons, and others, it is important that the surgeon be able to accurately estimate and describe the risks of the proposed surgery to the family. Risk assessment tools are one way to communicate these patient-specific risks to the patient and family in order to help them make decisions.\textsuperscript{7-11} Risk assessment tools can also guide the surgeon’s decision in whether or not operate, or what surgical options to present to the patient and family.

Until recently, there has not been sufficient pediatric data to support a risk calculator that could cover a broad selection of Current Procedural Terminology (CPT) codes\textsuperscript{6} and outcomes for children’s surgery. However, several risk calculators have been built for specific procedures and outcomes.\textsuperscript{7,12-13} The American College of Surgeons’ National Surgical Quality Improvement Program Pediatric (ACS NSQIP Pediatric) collects high quality, standardized clinical data including preoperative risk factors and postoperative outcomes from more than 70 pediatric hospitals.\textsuperscript{14-15} These data have been used to provide participating sites with risk adjusted 30-day outcome comparative reports that can aid in quality improvement efforts. The program has now collected three years of data that can be used to construct a risk calculator that covers the majority of ACS NSQIP Pediatric eligible CPT codes.

The pediatric surgical risk calculator described in this paper will satisfy a growing need for, and interest in, being able to estimate patient specific risk for a wide range of procedures and
outcomes within the pediatric patient population. Surgeons will be able to use it as a tool in the office or inpatient settings to inform patients and families of the estimated risks associated with both elective and emergent/urgent procedures. Our objective was to develop and evaluate a publicly available risk estimation tool that would cover many common pediatric surgical procedures across all specialties.

METHODS

Data Source and Patients

Data were obtained from the ACS NSQIP Pediatric. The ACS NSQIP Pediatric collects patient-level clinical data including demographics, comorbidities, laboratory values, and outcomes, and identifies cases by CPT. These data are reliable, as they are rigorously defined, and collected and recorded by trained Surgical Clinical Reviewers (SCRs) who undergo training and examination in variable definitions. In addition, the program performs random audits to check for data validity and definition compliance. Cases are systematically sampled across all specialties (general, neurosurgery, urology, otolaryngology, plastic, and orthopedic) at each participating hospital following an 8-day cycle. All cases are followed for 30 days using the medical record or patient outreach to verify the presence or absence of adverse events at the 30-day postoperative time point. In order to reliably estimate procedure risk, the dataset was restricted to only those cases whose CPT code is currently being collected by ACS NSQIP Pediatric and occurred at least 25 times in the dataset.

Predictors and Outcomes
The factors that were considered for calculating patient-specific risk of surgical outcomes were demographic and preoperative clinical variables. The final variables were chosen based on their predictive value, clinical face validity, and regular availability to the surgeon prior to the operation (Table 1). Predictive value was determined by examining past ACS NSQIP Pediatric Semi-Annual Reports (SARs). Predictors that had entered at least 20 models with a median step of fifth or earlier in the selection process were considered. In addition to these patient-specific risk factors, a procedure-specific risk variable based on CPT code was also included. This procedure-specific risk was created by taking a linear transformation of the predicted probabilities from preliminary hierarchical models where CPT was used as a random effect to predict each outcome.\textsuperscript{15}

The outcomes that were modeled were: mortality, composite morbidity (any defined morbidity), surgical site infection (SSI), venous thromboembolism (VTE), reintubation, cardiac event, pneumonia, renal insufficiency/failure, and urinary tract infection (UTI), all defined as binary variables indicating the absence or presence of the event within 30 days of the operation. The composite morbidity outcome was defined as having one or more of the following intra-operative or post-operative adverse events: pneumonia, reintubation, pulmonary embolism, renal insufficiency, renal failure, urinary tract infection, central line associated blood stream infection, coma >24 hours, seizure, peripheral nerve injury, any cerebral intra-ventricular hemorrhage, CVA/Stroke or intracranial hemorrhage, cardiac arrest requiring CPR, venous thrombosis requiring therapy, graft/prosthesis/flap failure, sepsis, surgical site infection (superficial, deep, or organ/space), blood transfusion $\geq 25\text{ml/kg}$, deep wound disruption, or superficial wound disruption. Some of these morbidity events and their corresponding outcomes were adjusted for, or negated by, pre-existing conditions: pneumonia with preoperative pneumonia, reintubation
with ventilator dependence, renal insufficiency or renal failure with preoperative renal failure or
dialysis. The SSI outcome includes superficial, deep, and organ space SSIs.

Statistical Analysis

Hierarchical, random effects modeling, which accounts for clustering of cases within
hospitals and imposes an empirical Bayes type shrinkage estimate to adjust for small sample
sizes, was used throughout (SAS PROC GLIMMIX).\textsuperscript{15} Each model included the demographic
and preoperative clinical variables listed in Table 1, as well as the CPT linear risk variable and a
random effect for hospital. Although hospital was used in the modeling, only the fixed effects
estimates, patient level factors, were used for risk prediction.

Model Performance

Two metrics were used to evaluate model performance: the c-statistic and the Hosmer-
Lemeshow (HL) statistic. First, the models’ discriminative ability was measured using the c-
statistic, also referred to as the area under the receiver operating characteristic (AUROC) curve
(the plot of sensitivity vs. 1-specificity). The c-statistic indicates how well the model does in
predicting membership in one of two groups (the binary response). It takes values from 0.5 to
1.0, with 0.5 indicating the model is no better than chance and 1.0 indicating perfect prediction.
In general, models with a c-statistic greater than 0.7 are considered reasonable and models are
considered strong when it exceeds 0.8. Although the c-statistic is useful in evaluating model
performance, it may not be the best metric, as it does not directly evaluate the accuracy of the
prediction and it suffers when a population is homogenous.
Second, the Hosmer-Lemeshow (HL) statistic was used as a measure of calibration. It attempts to detect bias in predicting risk over the range of risk. The data are ordered by predicted risk and then divided into ten risk groups of equal size. If there is a tendency for the model to over or under estimate risk based on the risk group, the HL statistic will be statistically significant. Unfortunately, the HL statistic varies depending on the (arbitrary) number of groups the data is divided into and finds smaller deviations to be statistically significant when the sample size becomes large due to its inherent reliance on the chi-square distribution. To avoid these issues, graphical representations of fit are provided, in addition to the statistic, in which the mean predicted probability is plotted versus the mean observed probability for each of the ten risk groups.

Surgeon risk adjustment

The Pediatric Surgical Risk Calculator does not capture all possible comorbidities and there is uncertainty surrounding any estimate. To address this issue, Bilimoria et al. incorporated a Surgeon Adjustment Score (SAS) for the Universal ACS NSQIP Surgical Risk Calculator. The SAS provides clinicians with the opportunity to increase a patient’s risk of surgery within a reasonable range for each specific CPT code. The default estimated risk from the model is assigned a SAS of 1. The surgeon can increase this to a SAS of 2 (median + 1 standard deviation of predicted risk) or SAS of 3 (median + 2 standard deviations). If an individual patient’s estimated risk is greater than the SAS, then the estimated risk from the model remains unmodified, as the patient’s predicted risk is already higher than the SAS modification. This adjustment is to be used only if the surgeon feels the patient has significant risk factors for surgery that are not already captured by the predictors that have been entered.
RESULTS

From January 1, 2012 through December 31, 2014, there were 183,233 patients entered into the database that underwent procedures at 67 participating ACS NSQIP Pediatric hospitals. All patients were under 18 years of age. Only CPT codes with at least 25 cases were utilized, resulting in 382 CPT codes, or 63% of the CPT codes in the full dataset, being used in modeling and available in the risk calculator. The final dataset consisted of 181,353 cases spanning these 382 CPT codes for the three-year period, representing these sub-specialties: general surgery, cardiothoracic, neurosurgery, orthopedics, otolaryngology, gynecology, urology, and plastic surgery.

The pediatric risk calculator reports on nine 30-day post-operative outcomes using a fixed set of 18 predictors (Table 1). Event rates ranged from 0.08% for renal failure to 5.91% for composite morbidity (Table 2). The c-statistics ranged from 0.77 for SSI to 0.98 for mortality, with eight of the nine models having c-statistics greater than 0.8 (Table 2). HL statistics were variable – several reached significance due to sensitivity to sample size (Table 2). However, fit was excellent when portrayed graphically with mean predicted probability plotted versus mean observed probability. Figure 1 displays fit for data grouped by CPT codes for three outcomes (mortality, morbidity, SSI). The diagonal line indicates perfect prediction (mean predicted rate = mean observed rate). Lines denoting ±25% are also shown. Points lie close to the diagonal line with very few outside the ±25% error lines suggesting that the model provides predictions that are consistent with the observed values. Figure 2 displays fit for all 9 outcomes across the observed risk spectrum. For most outcomes, there is a small tendency to under predict risk in the highest risk groups.
While all the data were used to create the calculator’s prediction equations, we also split the data into two mutually exclusive datasets for model validation. A simple random sample of two-thirds of the three years of data was used as the training dataset with the remaining one-third being the test set. A random allocation was considered superior to splitting the data by year, since there could be time trends in the data and minor changes in variable definitions over time. The average c-statistic in the training and test datasets were 0.912 and 0.869, respectively, though one model (reintubation) showed a more dramatic reduction of 0.949 to 0.652. These results confirm expectations for stable discrimination across training/test datasets typical of large sample sizes.

**DISCUSSION**

The desire of patients and families to be more involved in the surgical decision making process has led to a need for surgeons to be able to provide more information and information that is tailored to an individual patient’s circumstances. There is currently no tool available to help surgeons estimate the surgical risk for a pediatric patient across multiple specialties. The ACS NSQIP Pediatric now has sufficient standardized, high-quality data to develop a risk calculator with excellent calibration and discrimination. This tool will assist clinicians, patients, and families in estimating the risks of surgery on an individual basis and making informed decisions. The pediatric risk calculator was made publicly available as a web application at [http://riskcalculator.facs.org/peds](http://riskcalculator.facs.org/peds) in July of 2016 (Figure 3).

Other pediatric risk calculators have been developed, but they have several limitations. First, they are often for a single surgical procedure, complication, or specialty. Second, they use only one to two years of data, or data from a single institution, which can limit accuracy and
generlizability.\textsuperscript{12-13} Third, they have not typically been made publicly available in a user-friendly format. In order to aid in the decision-making process, clinicians and families need the risk estimation available during their surgical consultations.

There are also several limitations to this study and the pediatric risk estimation effort. First, the data utilized come from only 67 ACS NSQIP Pediatric hospitals. This could lead to bias, as participation in the program is voluntary and although ACS NSQIP Pediatric currently has a mix of both large research hospitals and smaller hospitals, the hospitals that choose to participate may share certain other characteristics, including a focus on quality and safety, increased resources, and different case mix. This bias may become less of an issue as the program continues to grow, and the pediatric risk calculator is updated. Second, only the preoperative variables collected by ACS NSQIP Pediatric were available to estimate postoperative risk. Although it may seem that additional variables are clinically important, it has been shown that the majority of risk adjustment can be done with a limited set of variables.\textsuperscript{18} This work focused on variables that would be known to surgeons and families prior to intervention to truly assist in decision making, and also focused on controlling the data entry burden to optimize usability. Third, some of the variable definitions have been evolving since the beginning of the ACS NSQIP Pediatric. However, we recognized this and since most of the major changes were made in the first program year, the first year of program data (2011) was not included in the development of the pediatric risk calculator. The majority of the definition changes that have been made since then have been clarifications and exclusions. These types of changes would not cause significantly more variation in the data than the variation that we would expect to see between data abstractors. Fourth, as with any statistical model, the results of the pediatric risk calculator are estimates and have some amount of error. The calculator should be
used by clinicians in combination with their clinical expertise to aid all parties involved in the process of making informed decisions.

The calculator, as has been described, is being released to the general public, but will be subject to continual modifications based on feedback from users. In addition, we anticipate that later versions of the calculator will include more granular composite measures (e.g., infection rather than overall morbidity), which should offer advantages for clinical risk evaluation and decision-making, and therefore be more informative and useful for providers and families.

CONCLUSIONS

There is an increased desire for patients and families to be more involved in the surgical decision making process. Surgeons need a quick and easy way to provide individualized information on surgical risk during consultations in the office and hospital settings. The ACS NSQIP Pediatric Surgical Risk Calculator will equip surgeons with an easy-to-use tool for providing patients and parents with empirically-derived, patient-specific estimates of post-operative complication risks. This additional information will enhance shared decision making, which should lead to better patient outcomes and patient care.
References:


Table 1. American College of Surgeons NSQIP Pediatric Variables Utilized in the Pediatric Surgical Risk Calculators

<table>
<thead>
<tr>
<th>Variable</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age group</td>
<td>0-28 d, 29 d-&lt;1 y, 1-&lt;2 y, 3-5 y, 6-7 y, 8-12 y, 13-&lt;18 y</td>
</tr>
<tr>
<td>Sex</td>
<td>Male, female</td>
</tr>
<tr>
<td>Transfer status</td>
<td>Admitted home, admitted ER, chronic care, transferred ICU, other</td>
</tr>
<tr>
<td>Case type</td>
<td>Elective, emergent, urgent</td>
</tr>
<tr>
<td>American Society of Anesthesiologists class</td>
<td>1, 2, 3, 4, or 5</td>
</tr>
<tr>
<td>Wound class</td>
<td>Clean, clean/contaminated, contaminated, dirty/infected</td>
</tr>
<tr>
<td>Inpatient/outpatient</td>
<td>Inpatient, outpatient</td>
</tr>
<tr>
<td>Systemic sepsis within 48 h prior to operation</td>
<td>None, systemic inflammatory response syndrome, sepsis, septic shock</td>
</tr>
<tr>
<td>Ventilator dependent</td>
<td>Yes, no</td>
</tr>
<tr>
<td>Neuromuscular disorder</td>
<td>Yes, no</td>
</tr>
<tr>
<td>Hematological disorder</td>
<td>Yes, no</td>
</tr>
<tr>
<td>Oxygen support</td>
<td>Yes, no</td>
</tr>
<tr>
<td>Cardiac risk factor</td>
<td>Yes, no</td>
</tr>
<tr>
<td>Structural CNS abnormality</td>
<td>Yes, no</td>
</tr>
<tr>
<td>Nutritional support</td>
<td>Yes, no</td>
</tr>
<tr>
<td>Developmental delay</td>
<td>Yes, no</td>
</tr>
<tr>
<td>Neonate</td>
<td>Yes, no</td>
</tr>
<tr>
<td>CPT-specific linear risk</td>
<td>382 values</td>
</tr>
</tbody>
</table>
Table 2. Complication Rates and Model Statistics for the Pediatric Surgical Risk Calculator Models

(N=181,353)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No. events and rate (%)</th>
<th>C-statistic</th>
<th>HL (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>62 (0.34)</td>
<td>0.9780</td>
<td>0.4821</td>
</tr>
<tr>
<td>Morbidity</td>
<td>10,834 (5.91)</td>
<td>0.8063</td>
<td>0.0000</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>652 (0.36)</td>
<td>0.8984</td>
<td>0.0000</td>
</tr>
<tr>
<td>Cardiac</td>
<td>264 (0.14)</td>
<td>0.9606</td>
<td>0.4860</td>
</tr>
<tr>
<td>Surgical site infection</td>
<td>3,598 (1.96)</td>
<td>0.7714</td>
<td>0.0004</td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>955 (0.52)</td>
<td>0.8818</td>
<td>0.0001</td>
</tr>
<tr>
<td>Venous thromboembolism</td>
<td>205 (0.11)</td>
<td>0.9424</td>
<td>0.1844</td>
</tr>
<tr>
<td>Renal Failure</td>
<td>138 (0.08)</td>
<td>0.9502</td>
<td>0.7243</td>
</tr>
<tr>
<td>Reintubation</td>
<td>827 (0.45)</td>
<td>0.9386</td>
<td>0.0121</td>
</tr>
</tbody>
</table>

Higher C-statistics are more predictive; non-significant Hosmer-Lemeshow (HL) P-values (>0.05) are more predictive.
FIGURE LEGENDS

Figure 1. Plots of mean observed rates vs mean predicted rates for Current Procedure Terminology (CPT) codes. The diagonal line represents perfect prediction (mean observed = mean predicted). SSI, surgical site infection.

Figure 2. Mean observed vs mean predicted probabilities for 10 risk deciles. SSI, surgical site infection; UTI, urinary tract infection; VTE, venous thromboembolism.

Figure 3. Screenshots of the American College of Surgeons NSQIP pediatric surgical risk calculator (http://riskcalculator.facs.org/peds). (A) Risk factor entry screen; (B) report screen.
Precis

The pediatric risk calculator can be used as a tool in the shared decision-making process by providing clinicians, families, and patients with useful information about common surgeries.
### Enter Patient and Surgical Information

Start by entering the procedure name or CPT code. One or more procedures will appear below the procedure box. You will need to click on the desired procedure to properly select it. You may also search using two words (or two partial words) by placing a `*` in between, for example: "cholecystectomy + cholangiography".

**Reset All Selections**

### Are there other potential appropriate treatment options?  
- [ ] Other Surgical Options
- [ ] Other Non-operative Options
- [x] None

Please enter as much of the following information as you can to receive the best risk estimates. A rough estimate will still be generated if you cannot provide all of the information below.

<table>
<thead>
<tr>
<th><strong>Age Group</strong></th>
<th>6-7 years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td>Female</td>
</tr>
<tr>
<td><strong>Transfer Status</strong></td>
<td>Admitted from ER</td>
</tr>
<tr>
<td><strong>Case Status</strong></td>
<td>Urgent</td>
</tr>
<tr>
<td><strong>ASA Class</strong></td>
<td>Mid systemic disease</td>
</tr>
<tr>
<td><strong>Wound Classification</strong></td>
<td>Clean/Contaminated</td>
</tr>
<tr>
<td><strong>Inpatient/Outpatient</strong></td>
<td>Inpatient</td>
</tr>
<tr>
<td><strong>Systemic Sepsis within 48 hours prior to surgery</strong></td>
<td>Sepsis</td>
</tr>
<tr>
<td></td>
<td>Ventilator Dependent</td>
</tr>
<tr>
<td><strong>Neuromuscular Disorder</strong></td>
<td>No</td>
</tr>
<tr>
<td><strong>Hematologic Disorder</strong></td>
<td>No</td>
</tr>
<tr>
<td><strong>Oxygen Support</strong></td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Cardiac Risk Factors</strong></td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Structural CNS Abnormality</strong></td>
<td>No</td>
</tr>
<tr>
<td><strong>Nutritional Support</strong></td>
<td>No</td>
</tr>
<tr>
<td><strong>Neonatal Status</strong></td>
<td>Pediatric</td>
</tr>
<tr>
<td><strong>Developmental Delay/Impaired Cognitive Status</strong></td>
<td>No</td>
</tr>
</tbody>
</table>