Title:
Visualizing Opioid-Use Variation in a Pediatric Perioperative Dashboard

Authors:
Conrad Safranek, B.S.
Institution: Stanford University, Stanford, United States
Department: Department of Biology: Computational Biology

Lauren Feitzinger, B.S.
Institution: Stanford University, Stanford, United States
Department: Department of Management Science and Engineering

Alice Kate Cummings Joyner, M.S.
Institution: Stanford University, Stanford, United States
Department: Department of Management Science and Engineering

Nicole Woo, B.S.
Institution: Stanford University, Stanford, United States
Department: Department of Management Science and Engineering; Department of Computer Science

Virgil Smith, M.S.
Institution: Stanford University, Stanford, United States
Department: Department of Management Science and Engineering

Elizabeth De Souza, Ph.D.
Institution: Stanford University School of Medicine, Stanford, CA, United States
Department: Department of Anesthesiology, Perioperative and Pain Medicine

Christos Vasilakis, Ph.D.
Institution: University of Bath, Bath, United Kingdom
Department: School of Management, Centre for Healthcare Innovation and Improvement

T. Anthony Anderson, Ph.D., M.D.
Institution: Stanford University School of Medicine, Stanford, CA, United States
Department: Department of Anesthesiology, Perioperative and Pain Medicine

James Fehr, M.D.
Institution: Stanford University School of Medicine, Stanford, CA, United States
Department: Department of Anesthesiology, Perioperative and Pain Medicine

Andrew Y. Shin, M.D.
Institution: Stanford University School of Medicine, Stanford, CA, United States

Department: Department of Pediatrics - Cardiology

David Scheinker, Ph.D.
Institution: Stanford University, Stanford, United States
Department: Department of Management Science and Engineering

Ellen Wang, M.D.
Institution: Stanford University School of Medicine, Stanford, CA, United States
Department: Department of Anesthesiology, Perioperative and Pain Medicine

James Xie, M.D.
Institution: Stanford University School of Medicine, Stanford, CA, United States
Department: Department of Anesthesiology, Perioperative and Pain Medicine

Corresponding Author:
James Xie, M.D.
Clinical Assistant Professor
300 Pasteur Drive
Room H3580 MC 5640
Department of Anesthesiology, Perioperative and Pain Medicine
Stanford, CA 94305
Email: jxie@stanford.edu
Phone: 650-544-1567

Word Count: 3628 words

Human Subject Research Approval

This quality improvement initiative to reduce unnecessary variation in opioid use was reviewed by Stanford's Research Compliance Office and exempted from formal IRB review.

Conflict of Interests Statement:

All authors have no conflicts of interests to declare. Stanford University supplied a one-time research grant to Conrad Safranek during summer 2020 to continue development of the dashboard and to write a research manuscript. No other authors received any funding for this research.

ABSTRACT

Background: Anesthesiologists integrate numerous variables to determine an opioid dose that manages patient nociception and pain while minimizing adverse effects. Clinical dashboards that enable physicians to compare themselves to their peers can reduce unnecessary variation in patient care and improve outcomes. However, due to the complexity of anesthetic dosing decisions, comparative visualizations of opioid-use patterns are complicated by case-mix differences between providers.

Objectives: This single-institution case-study describes the development of a pediatric anesthesia dashboard and demonstrates how advanced computational techniques can facilitate nuanced normalization techniques, enabling meaningful comparisons of complex clinical data.

Methods: We engaged perioperative-care stakeholders at a tertiary care pediatric hospital to determine patient and surgical variables relevant to anesthesia decision-making and to identify end-user requirements for an opioid-use visualization tool. Case data were extracted, aggregated, and standardized. We performed multivariable machine learning to identify and understand key variables. We integrated interview findings and computational algorithms into an interactive dashboard with normalized comparisons, followed by an iterative process of improvement and implementation.

Results: The dashboard design process identified two mechanisms—interactive data filtration and machine-learning-based normalization—that enable rigorous monitoring of opioid utilization with meaningful case-mix adjustment. When deployed with real data encompassing 24,332 surgical cases, our dashboard identified both high and low opioid-use outliers with associated clinical outcomes data.

Conclusions: A tool that gives anesthesiologists timely data on their practice patterns while adjusting for case-mix differences empowers physicians to track changes and variation in opioid administration over time. Such a tool can successfully trigger conversation amongst stakeholders in support of continuous improvement efforts. Clinical analytics dashboards can enable physicians to better understand their practice and provide motivation to change behavior, ultimately addressing unnecessary variation in high impact medication use and minimizing adverse effects.

Keywords: Anesthesiology, Clinical Decision Support, Clinical Informatics Systems, Pediatrics, Machine Learning

1. BACKGROUND AND SIGNIFICANCE

Opioids are a mainstay of perioperative pain management in children.[1,2] Appropriate opioid utilization requires a balance between managing acute pain perioperatively while minimizing risk of opioid-related adverse effects.[3] This balance is particularly challenging for the pediatric population, which encompasses a range of weights and developmental stages.[4]

In addition to the immediate risks of postoperative nausea and vomiting, high-dose opioid regimens during surgery contribute to higher pain scores at 24 hours postoperatively and are associated with worse long-term outcomes for some surgeries.[3,5–7] Furthermore, minimizing unnecessary variation in perioperative opioid use may reduce the risk of persistent opioid use after surgery, which was recently estimated to have a 5% prevalence amongst pediatric surgical patients.[8,9]

Multimodal perioperative analgesia presents an opportunity to optimize pain management while maintaining or improving outcomes. Regional anesthesia (e.g. peripheral nerve blocks and neuraxial blocks) and non-opioid adjuncts such as acetaminophen and nonsteroidal anti-inflammatory drugs (NSAIDs) can improve analgesia by targeting different pain mechanisms, thus minimizing opioids' dose-dependent adverse effects.[10–12] Some hospitals have successfully promoted such alternatives to optimize opioid utilization without compromising quality of care.[13] While nerve-blocks and other analgesic alternatives are effective for many types of surgeries, integration of these alternatives into practice has been variable.[14] Moreover, even after accounting for demographic and surgical variables expected to influence opioid utilization (including differences in multimodal analgesia), recent evidence suggests there remains significant unexplained intra- and inter-institutional opioid-use variability. Moreover, even after accounting for demographic and surgical variables expected to influence opioid utilization (including differences in multimodal analgesia), a recent study of adult perioperative opioid utilization suggests there remains significant unexplained intra- and inter-institutional variability;[15] Given the aforementioned added complexity of pediatric patients' weight and developmental stages, this variability is likely present to an equal or greater degree for pediatric anesthesia.

Given the drive to reduce unnecessary variation in opioid use and the opportunities presented by alternative analgesics, the perioperative care team is poised to develop clinical- and systems-based interventions; however, various barriers need to be addressed to understand and implement interventions effectively. Our ability to drive change in anesthesiologists' opioid utilization depends on awareness of current and historical usage trends. Many hospitals lack ready access to these objective measures. Currently for physicians, reviewing opioid use for historic cases involves a laborious process of manually accessing and analyzing patient charts. This is further complicated by differences in opioid requirements based on procedure type and patient demographics. To effectively guide clinical practice, anesthesiologists and managers need timely and accessible data that enable relevant comparisons while accounting for case-mix differences.

A perioperative opioid analytics dashboard is a promising solution to this challenge. Hospital dashboards are increasingly being adopted as a valuable addition to traditional electronic health record (EHR) systems, as they enable display of previously disparate data on a single platform, standardizing and summarizing clinical trends in easy-to-process visualizations.[16] As a result, dashboards can assist provider-level decision-making and manager-level continuous monitoring of quality of care. They also may be effective in reducing unnecessary variation in clinical practice and achieving sustainable quality improvement (QI) initiatives, which require real-time data collection and interpretable, timely feedback.[17–20]

2. OBJECTIVE

In this study, our team endeavored to design a dashboard with meaningful comparisons of a multivariable-dependent clinical decision-making metric: pediatric anesthetics. Through stakeholder interviews and an iterative development process, we identified and refined computational techniques to enable clinically meaningful, normalized comparisons of our anesthesiology division's system-wide and provider-level perioperative opioid-utilization patterns. This manuscript describes the design and development of this perioperative opioid analytics dashboard.

3. METHODS

To design an opioid-use dashboard with clinically meaningful comparisons, we followed four stages (Figure 1). First, we interviewed perioperative stakeholders to select relevant variables that may influence opioid dosing, identify key perioperative patient outcomes, and determine overall dashboard needs and design specifications. Second, we extracted, cleaned, aggregated, and standardized data for all outpatient surgeries between May 4, 2014, and August 31, 2019. Third, we implemented select univariable analyses and multivariable machine learning to identify and understand key variables and relationships. Finally, we integrated these findings and algorithms to design an interactive dashboard with normalized comparisons, followed by an iterative process of improvement and implementation.

This quality improvement initiative to reduce unnecessary variation in opioid use was reviewed by Stanford's Research Compliance Office and exempted from formal IRB review.

3.1 Interviews

Semi-structured interviews were conducted to inform dashboard design considerations and to identify variables relevant to perioperative opioid use. For interviewee selection, project sponsors identified perioperative leaders including practicing and managing anesthesiologists, surgeons, pharmacists, post-anesthesia care unit (PACU) nurses, hospital data analysts, and child life specialists.

One team member led the interviews while other members typed detailed notes. A team-designed script included standardized questions used across all interviews to prompt discussion. Interview notes were coded by the group to inductively identify project and dashboard goals. Qualitative interview data was aggregated and synthesized, and summary tables were used to identify project goals and their relative importance for each stakeholder group.

3.2 Data procurement, processing, and cohort selection

Our institution is a tertiary care academic pediatric hospital that performs ~5,000 outpatient surgeries per year. Data for all outpatient surgeries during the 5.33-year study period was extracted from the Epic EHR (Verona, WI). Both intraoperative data as well as post-anesthesia care unit (PACU) data were included. Queries from multiple sources were combined by unique case identification numbers to create a final data set that included case details (date, surgery department, American Society of Anesthesiologists [ASA] physical status classification, anesthesiologist, surgeon), patient information (age, weight), surgery details (primary procedure, anesthesia duration, surgery duration), PACU details (length of stay, pain scores), and intraoperative and PACU analgesia details (amount and type of opioid, absence or presence of regional anesthesia, absence or presence of opioid infusion, absence or presence of naloxone, amount of naloxone if used).

All perioperative opioid medications were converted to oral morphine equivalent units (MEU) using standard equivalency coefficients to determine total opioid received intraoperatively and in the PACU, respectively.[21] Cases with missing or invalid procedure duration or anesthesia duration due to missing or invalid start/stop times were excluded from analysis. Additionally, cases were excluded if they included the use of remifentanil, an ultra-short acting opioid with a high relative potency that would skew the total calculations of intraoperative MEU. Dataset merging, processing, and cleaning were completed with the dplyr package of R Version 1.2.5001.

Similar common primary procedures with differing names were grouped for future dashboard usability: We selected the top 20 procedures per service line to pare down 1,014 unique procedure types to 230. From there, procedures were manually grouped into a list of 167 procedures.

Anesthesiologist identities were anonymized using a unique and secure key for privacy and to minimize potential bias during analysis.

3.3 Data analysis and machine learning

Mean intraoperative and PACU MEU were compared by surgical service and primary procedure type. To further characterize the effects of different variables on opioid administration, machine learning was used to normalize for all available variables simultaneously. This was accomplished via comparison of MEU per case "observed" and the amount "expected," an algorithmic prediction for each case that integrates all available case data (e.g. patient weight, procedure duration, and service; see Table S1 for complete list).

Specifically, we implemented a random forest machine learning model using the randomForest R package. Of the two linear models (lasso and ridge regression) and six machine learning models (least angle regression, elastic net, regression trees, random forest, gradient boosting and XGBoost) considered and tested,[22] random forest and gradient boosting demonstrated the best performance. Gradient boosting had a marginally better R^2 but took more than 10x longer to train relative to random forest, hence our choice to proceed with the randomForest package for dashboard data visualizations.

Intraoperative and PACU MEU were combined to determine a total "observed" MEU for each case. We removed 13 high outliers (0.05% of all data) that were above 50 total MEU. Cases from May 2014 to December 2018 were designated as training data to optimally tune the random forest algorithm. The training/test sets were split chronologically because we wanted dashboard visualizations with untrained, machine learning predictions for the recent data most relevant to current clinicians.

The random forest was trained with 1000 trees to prevent data overfitting. Model optimization was achieved via minimization of the root mean squared error (RMSE) between the "expected" and "observed" MEU per case. Because the randomForest package can only process categorical variables with up to 53 unique values, for the primary procedure and surgeon variables the top 52 most common were kept while the remaining were labeled as "other" to create "reduced" features. The final formula for machine learning was total MEU for each case as predicted by patient age, patient weight, procedure duration, "reduced" procedure type, service (surgical department), "reduced" operating surgeon, ASA classification, and presence of intraoperative nerve-block. The model tracked the percent importance of each variable with respect to RMSE reduction for the training data. "Expected" dose predictions for cases in 2019 were used to calculate final R^2 and were used for subsequent dashboard visualizations.

3.4 Dashboard development

The iterative dashboard design process integrated continuous stakeholder feedback. Because the dashboard's end-users include practicing and managing anesthesiologists, we organized individual and group feedback sessions with four anesthesiologists from the interview process, including the chief of pediatric anesthesiology. Initial mock-ups were depicted with Microsoft PowerPoint, while all working dashboard versions were developed and iterated upon using Tableau Desktop 2020.1 (Seattle, WA, USA).

Case-mix control with interactive data filtration

For interactive data filtration, the dashboard utilizes Tableau's built-in functionality. Drop-down menus enable users to visualize metrics from specified rows of the dataset by selecting specific values or a range of values for categorical and continuous metrics, respectively. Specific parameter options were selected based on anesthesiologists' needs, as identified during interviews. For comparisons between subsets of the data and the general data set, Tableau's data parameter feature enables overlay graphing.

Case-mix control with machine learning

Following regression analysis, the "expected" and "observed" MEUs could be compared for each case in 2019 to enable comparisons between providers. We applied two analytical techniques to the results: 1) An "index" for each provider was calculated by summing all of an individual provider's "observed" MEUs for their 2019 cases and dividing this value by the sum of all their "expected" MEUs, and 2) A residual for each case was calculated by taking the difference between "observed" and "expected" MEU.

4. RESULTS

4.1 Interviews

We conducted 16 semi-structured interviews with stakeholders, all of which indicated substantial interest in a tool to enhance understanding of perioperative opioid administration patterns. Inductive coding of interview notes enabled identification of key data visualization needs and corresponding dashboard design implications. The distinct needs and design specifications were cataloged for each primary stakeholder group, each of which emphasized specific aspects for a dashboard (Table 1).

The needs identified and associated design solutions were: 1) Understanding the distribution of opioid use with histograms, 2) Understanding historical trends of opioid use as a time series, 3) Understanding and visualizing outcomes (e.g. pain scores) in the PACU, in relation to opioid use, and 4) Ensuring long-term integration and sustainability of the dashboard into our institution's perioperative QI and analytics workflows.

4.2 Data overview

After exclusion criteria, 24,332 outpatient surgeries over 5.33 years (4,565 cases per year on average) were included in the dashboard. These surgeries encompassed 23 surgical departments, over 1,000 unique primary procedures, and 71 anesthesiologists. EHR data was validated with data visualizations in R and manual audits.

Median age and weight for patients receiving surgery were 7.56 (interquartile range [IQR]: 3.56 to 13.19 years) and 25.3 kilograms (IQR: 15.1 to 49.2 kg), respectively. There were 14,309 (58.8%) cases with a male patient. Median surgery duration was 0.55 hours (IQR: 0.28 to 1.03 hours), while median PACU length of stay was 1.52 hours (IQR: 1.15 to 1.97 hours) with a median maximum pain score on the numerical rating scale (NRS) of 0.00 (IQR: 0.00 to 5.00) (Table 2).

4.3 Data analysis and machine learning

Data analysis

When comparing MEU kg^{-1} $case^{-1}$, we found a wide range of opioid utilization by service line and procedure type. Across services, mean MEU kg^{-1} $case^{-1}$ ranged from 0.06 to 1.75 times the overall mean, with departments involving orthopedic or gynecologic operations generally receiving more opioids (Figure 2). The intraoperative to PACU utilization ratio also varied by service line, with some services such as cardiology and interventional radiology almost exclusively using intraoperative opioids (with no PACU opioid administration), and other services such as otolaryngology and orthopedics having ~15% of overall MEU kg^{-1} $case^{-1}$ administered in the PACU. We also observed variability among primary procedures within a department. These initial data visualizations emphasized the importance of case-mix control when comparing providers to their peers.

Machine learning

When trained on the 2014 to 2018 cases (18,793 observations, 85.6% of cohort) and then tested on the 2019 data (3,154 observations, 14.4% of cohort), the random forest model achieved an R^2 of 54.0%.

For the 2014-2018 data trained with the random forest model, the training process tracks which variables are most important for reducing the root mean squared error. Our model found that procedure duration, patient weight, and type of primary procedure were most predictive of total MEU (Table S1).

4.4 Dashboard development

One focus was determining what level of anonymity would best balance transparency and provider privacy. During the iterative design process, our team's managing anesthesiologists decided against implementing a name-based ranking system, in which each anesthesiologist would be identifiable to others in their department. To prevent dashboard users from determining opioid-use details of specific peers by using narrow date and case detail filtration, for some of the dashboard's data stratification features visualization was only available if at least 10 cases met filtration criteria. Furthermore, two versions—a Managerial Dashboard and a Provider-Level Dashboard—were created to provide granularity for managers while streamlining most-relevant data for individual anesthesiologists.

Case-mix control with interactive data filtration

For clinically relevant provider-level comparisons that account for variable case-mixes, the first mechanism implemented was interactive data filtration with drop-down menus. To address needs expressed in stakeholder interviews, these menus enable filtration of cases by service line, primary procedure, date range, and use of regional anesthesia.

This allows an anesthesiologist to compare his or her statistics to peers in their department, or more specifically to peers performing the same primary procedure (Figure 3). For granular comparison of cases with weight-normalization, the histogram of MEU kg^{-1} $case^{-1}$ for all of an individual's cases within the specific filtration criteria can be compared to the combined distribution of all other anesthesiologists' cases meeting the same criteria (Figure 3A). Additionally, this filtration offers individual-to-peer comparisons for relevant outcomes, such as PACU length of stay and PACU pain score distributions (Figure 3B and 3C).

By simultaneously accounting for two important variables in anesthetic decision-making (patient weight and procedure type), much of the expected and necessary opioid dose variation is normalized, allowing for more meaningful comparisons between providers. Figure 3 demonstrates this utility via a pair of two provider-to-peer comparisons: Provider A observes that, on average relative to their peers providing anesthesia for the same surgery, their patients receive more MEU per kilogram per tonsillectomy, have below average PACU pain scores, and have above average PACU length of stay; by comparison, Provider B's patients have similarly below average pain scores but average PACU length of stay, despite a much lower opioid utilization.

Case-mix control with machine learning

By comparing machine-learning-normalized "expected" opioid administration for a given case (accounting for all available variables) to the "observed" value for that case across all case data in 2019, we calculated an observed to expected "index" for each provider, thus showing providers' opioid utilization relative to their colleagues while accounting for variation in case-mix (Figure 4). For the Provider-Level Dashboard, we chose to present this machine-learning-normalized, department-wide comparison in quintiles, so that an individual provider can see which quintile they fall into relative to their peers.

The distribution of each provider index across the anesthesiology department enabled the identification of anesthesiologists who used more or less MEU than would be expected given their unique case mix (representing high and low outliers, respectively). Additionally, both individuals and managers can “drill down” into cases with high or low residual values, study case details, and look for patterns. This regression-based outlier identification system generated conversation amongst managing anesthesiologists and provided opportunities for learning from historical data to enhance continuous QI initiatives.

5. DISCUSSION

We describe the design and implementation of a clinical dashboard to visualize opioid administration in the pediatric perioperative setting. To realize the goal of a learning healthcare system, clinical dashboards can make data more accessible and easily interpretable for stakeholders, drive positive behavior changes in users, and ultimately improve patient outcomes.[16–18,23] Although significant amounts of clinical data have been collected in EHR systems, aggregation and visualization often require advanced data analytics skills and are complicated by case-mix differences between providers.

Other anesthesia-related clinical dashboards have demonstrated the feasibility and positive impact of tracking specific outcome metrics or indicators (e.g. opioid-related adverse drug events or substance documentation errors) for anesthesiology safety initiatives.[24–29] Few dashboards, however, have integrated quantitative analysis of standardized morphine equivalents,[17,29] and none have demonstrated the feasibility of quantitative, provider-level opioid-use comparisons.

This dashboard can more broadly serve as a framework for other institutions interested in comparative data visualizations for multivariable-dependent, clinical decision-making metrics.

5.1 Machine learning normalization

Although multivariable normalization techniques are not typically implemented in clinical support systems, such machine-learning models can reduce noise in the outcome of interest by accounting for expected variation, and thus highlight unnecessary or erroneous variation in that outcome.[23] We used machine learning to identify sources of expected opioid variation—such as patient age/weight or procedure type/length—and account for these factors in our dashboard; this assures anesthesiologists that differences shown in the dashboard are not due to their practice having systematically different patient demographics or procedures relative to their peers.

5.2 Understanding significant unexplained variation

Our machine learning model achieved an R^2 of 0.54, meaning that while 54% of the variation in opioid utilization can be accounted for by expected sources of variation reflected in the model, there is still unexplained variation. This R^2 represents a modest increase relative to a previous, multi-institution multivariable linear modeling study of intraoperative opioid use.[15]

Mathematical modeling is unable to determine what proportion of the remaining ~50% of unexplained variation is due to shortcomings of the multivariable modeling, factors influencing opioid dosing unaccounted for in our modeling, or truly due to unnecessary variation in opioid utilization by anesthesiologists. Other factors that could influence opioid administration, but that were not incorporated into this analysis, include intraoperative vital signs and hemodynamic changes, as these may prompt anesthesiologists to assess and modify anesthetics. A dashboard highlighting provider-level differences that remain after normalization for the 54% of predictable variability offers anesthesiologists a window into the unexplained variation and provides impetus to track and improve practices.

5.3 Importance of anonymization for physician privacy

Although other dashboards have leveraged the competitive nature of clinicians as a means to drive behavior change with non-anonymous ranking systems,[18] we chose not to disclose provider names. This was in part because the machine-learning normalization technique can detect high outliers, which may correspond to overt misuse of opioids or potential opioid diversion.[30] Thus, due to the sensitive potential of this dashboard, managing anesthesiologists felt it was in the best interest of confidentiality to provide practicing anesthesiologists with visualizations comparing their practice patterns only with anonymized, aggregated peer data. While managers can access the full data to review providers' practice patterns and trends, careful steps were taken for the anonymization of provider comparisons to their peers.

5.4 Limitations

A major limitation that may impede other hospitals from implementing such a dashboard is the requirement of a robust information technology infrastructure. While large hospitals generally have staff data analysts, smaller hospitals may not have such resources, thus limiting generalizability. Additionally, while the dashboard framework we developed in Tableau can be a starting point for other settings, it requires familiarity with the software.

Another limitation of our dashboard lies in the "expected" dose calculations as determined by machine-learning with historic, real-world data. Since dosing guidelines are typically empirically derived in pediatrics, we calculated an expected MEU dose based on historical trends at our institution. Thus, if providers were universally under or over utilizing opioids for a procedure, the expected dose would be skewed. However, the value generated by this technique can be a baseline for future change, and still assists in normalizing provider-level comparisons within our institution.

Moreover, our team noted that in the machine-learning normalized provider comparison (Figure 4), most provider quintiles fell below the "observed = expected" line. We hypothesize that this is primarily because the distribution of MEU per case is considerably right skewed and because our machine learning algorithm minimizes root mean squared error, which is sensitive to high outliers, ultimately leading to "expected" MEU predictions that are on average greater than the "observed" MEU values. We had also considered whether this represented an overall decreasing trend in opioid use over time between the training data prior to 2019 compared to the test data in 2019, but in analyses using different, non-chronological mixes of training sets, the majority of providers still fell below the "observed = expected" line.

5.5 Future directions

A dashboard which analyzes both individual physician and department-wide trends surrounding high-impact perioperative medication usage could be a valuable tool for implementing and assessing QI initiatives, and thus aid institutions in understanding and reducing unnecessary variation in clinical practice. For example, these differences could prompt systematically high opioid-use providers to learn from their peers’ effective use of alternative, non-opioid analgesics and thereby reduce variation across anesthesia practice. Stronger understanding of perioperative opioid utilization will allow better implementation of Enhanced Recovery after Surgery (ERAS) protocols, which seek to provide consistent and optimal perioperative care to improve quality of recovery, safety, and outcomes.[31] Sustainable implementation of ERAS requires tracking opioid usage, pain scores, length of stay, and adverse outcomes, so a dashboard that can analyze this aggregated data in real time is critical.[31]

This dashboard framework has the potential for expansion to other high impact and high-cost medications, such as cancer therapeutics or recombinant factor replacement agents. Medication doses and clinically relevant patient factors influencing dosing are generally more consistently recorded in EHRs for high impact medications, meaning data visualizations normalizing for expected variation could reduce noise to enable new clinical insights.

6. CONCLUSIONS

Design and implementation of a clinical dashboard to visualize variation in pediatric opioid administration is feasible. Moreover, interactive data filtration and machine learning techniques can be used to identify and adjust for factors that may influence opioid utilization, enabling meaningful clinical comparisons. The techniques implemented in this dashboard can serve as a framework for other institutions seeking comparative data visualizations for multivariable-dependent, clinical decision-making metrics. Such a tool enables physicians to compare their practice to their peers' and can thus motivate behavioral change, ultimately addressing unnecessary variation in clinical practice and minimizing medication-related adverse effects.

Clinical Relevance Statement:
Opioids are high impact medicines with significant side effects but are frequently used in the perioperative period. Clinicians lack ways to understand and receive feedback on their opioid-utilization patterns. By creating an opioid-utilization dashboard with normalized, provider-level comparisons, clinicians can work to reduce unnecessary variation in clinical practice, visualize usage trends, and ultimately improve patient outcomes.

Acknowledgements:
We would like to thank the Systems Utilization Research for Stanford Medicine (SURF) group for facilitating the initial development of this project. JX and CS drafted the initial manuscript. CS, LF, AKCJ, NW, and VS were part of the original SURF project team that interviewed stakeholders, designed and iterated the visualization tool, and modeled the data under the intellectual guidance of CV, AYS, DS as the SURF course directors and JF, TAA, EW, and JX as the project sponsors and clinical content leads. EDS, EW, and JX were responsible for clinical data extraction and validation. All authors CS, LF, AKCJ, NW, VS, EDS, CV, TAA, JF, AYS, DS, EW, JX participated in data interpretation, reviewed the manuscript critically for important intellectual content, and gave final approval of the version to be published.

Protection of Human and Animal Subjects
This quality improvement initiative was reviewed by Stanford's Research Compliance Office and exempted from formal IRB review.

Multiple Choice Questions:

1. When considering whether a medication might be suitable for a data-visualization dashboard, which of the following is a consideration?

   A. Whether the medication and associated patient and clinical factors are well-documented in the electronic health record.
   B. Whether it is critical that physicians determine an optimal medication dose (i.e., when too high or too low of a dose leads to negative patient consequences).
   C. Whether the hospital department has clear goals to improve use-patterns of the medication.
   D. All of the above.

2. Who should primarily define design-specifications of a clinical dashboard?

   A. Data analysts.
   B. End users.
   C. Patients.
   D. Hospital executives.

3. When designing provider-level utilization comparisons regarding a high-impact perioperative medication, why is multi-variable normalization important?

   A. Physicians may have systematic differences in procedure types.
   B. Physicians may have systematic differences in patient age and weights.
   C. Physicians may have systematic differences in case complexity.
   D. All of the above.

References:

1. Morton NS, Errera A. APA national audit of pediatric opioid infusions. *Paediatr Anaesth*. 2010;20(2):119-125.

2. Howard RF, Lloyd-Thomas A, Thomas M, et al. Nurse-controlled analgesia (NCA) following major surgery in 10,000 patients in a children's hospital. *Paediatr Anaesth*. 2010;20(2):126-134.

3. Association of Paediatric Anaesthetists of Great Britain and Ireland. Good practice in postoperative and procedural pain management, 2nd edition. *Paediatr Anaesth*. 2012;22 Suppl 1:1-79.

4. Cravero JP, Agarwal R, Berde C, et al. The Society for Pediatric Anesthesia recommendations for the use of opioids in children during the perioperative period. *Paediatr Anaesth*. 2019;29(6):547-571.

5. Long DR, Friedrich S, Eikermann M. High intraoperative opioid dose increases readmission risk in patients undergoing ambulatory surgery. *Br J Anaesth*. 2018;121(5):1179-1180.

6. Kovac AL. Management of postoperative nausea and vomiting in children. *Paediatr Drugs*. 2007;9(1):47-69.

7. Lu Y, Beletsky A, Cohn MR, et al. Perioperative Opioid Use Predicts Postoperative Opioid Use and Inferior Outcomes after Shoulder Arthroscopy. *Arthroscopy*. Published online June 4, 2020. doi:10.1016/j.arthro.2020.05.044

8. Ward A, Jani T, De Souza E, Scheinker D, Bambos N, Anderson TA. Prediction of Prolonged Opioid Use After Surgery in Adolescents: Insights From Machine Learning. *Anesth Analg*. 2021;133(2):304-313.

9. Harbaugh CM, Lee JS, Hu HM, et al. Persistent Opioid Use Among Pediatric Patients After Surgery. *Pediatrics*. 2018;141(1). doi:10.1542/peds.2017-2439

10. Wong I, St John-Green C, Walker SM. Opioid-sparing effects of perioperative paracetamol and nonsteroidal anti-inflammatory drugs (NSAIDs) in children. *Paediatr Anaesth*. 2013;23(6):475-495.

11. Ong CKS, Seymour RA, Lirk P, Merry AF. Combining paracetamol (acetaminophen) with nonsteroidal antiinflammatory drugs: a qualitative systematic review of analgesic efficacy for acute postoperative pain. *Anesth Analg*. 2010;110(4):1170-1179.

12. McDaid C, Maund E, Rice S, Wright K, Jenkins B, Woolacott N. *Paracetamol and Selective and Non-Selective Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) for the Reduction of Morphine-Related Side Effects after Major Surgery: A Systematic Review*. NIHR Journals Library; 2010.

13. Franz AM, Martin LD, Liston DE, Latham GJ, Richards MJ, Low DK. In Pursuit of an

Opioid-Free Pediatric Ambulatory Surgery Center: A Quality Improvement Initiative. *Anesth Analg*. Published online April 10, 2020. doi:10.1213/ANE.0000000000004774

14. Kendall MC, Alves LJC, Suh EI, McCormick ZL, De Oliveira GS. Regional anesthesia to ameliorate postoperative analgesia outcomes in pediatric surgical patients: an updated systematic review of randomized controlled trials. *Local Reg Anesth*. 2018;11:91-109.

15. Naik BI, Kuck K, Saager L, et al. Practice Patterns and Variability in Intraoperative Opioid Utilization: A Report From the Multicenter Perioperative Outcomes Group. *Anesth Analg*. Published online July 21, 2021. doi:10.1213/ANE.0000000000005663

16. Dowding D, Randell R, Gardner P, et al. Dashboards for improving patient care: review of the literature. *Int J Med Inform*. 2015;84(2):87-100.

17. Wanderer JP, Lasko TA, Coco JR, et al. Visualization of aggregate perioperative data improves anesthesia case planning: A randomized, cross-over trial. *J Clin Anesth*. 2021;68:110114.

18. Frank SM, Thakkar RN, Podlasek SJ, et al. Implementing a Health System-wide Patient Blood Management Program with a Clinical Community Approach. *Anesthesiology*. 2017;127(5):754-764.

19. Mortimer F, Isherwood J, Pearce M, Kenward C, Vaux E. Sustainability in quality improvement: measuring impact. *Future Healthc J*. 2018;5(2):94-97.

20. Belostotsky V, Laing C, White DE. The sustainability of a quality improvement initiative. *Healthc Manage Forum*. Published online April 6, 2020:840470420913055.

21. *Opioid Analgesics Comparison (Lexicomp LPCH Housestaff Manual)*. Stanford Health Care http://www.crlonline.com.laneproxy.stanford.edu/lco/action/doc/retrieve/docid/chistn_f/67895?cesid=azCZIE3m6gG&searchUrl=%2Flco%2Faction%2Fsearch%3Fq%3Dopioid%26t%3Dname%26va%3Dopioid

22. McCarthy RV, McCarthy MM, Ceccucci W, Halawi L. *Applying Predictive Analytics: Finding Value in Data*. Springer; 2019.

23. Corny J, Rajkumar A, Martin O, et al. A machine learning–based clinical decision support system to identify prescriptions with a high risk of medication error. *J Am Med Inform Assoc*. 2020;27(11):1688-1694.

24. Stone AB, Jones MR, Rao N, Urman RD. A Dashboard for Monitoring Opioid-Related Adverse Drug Events Following Surgery Using a National Administrative Database. *Am J Med Qual*. 2019;34(1):45-52.

25. Dolan JE, Lonsdale H, Ahumada LM, et al. Quality Initiative Using Theory of Change and Visual Analytics to Improve Controlled Substance Documentation Discrepancies in the Operating Room. *Appl Clin Inform*. 2019;10(3):543-551.

26. Laurent G, Moussa MD, Cirenei C, Tavernier B, Marcilly R, Lamer A. Development, implementation and preliminary evaluation of clinical dashboards in a department of anesthesia. *Journal of Clinical Monitoring and Computing*. 2021;35(3):617-626. doi:10.1007/s10877-020-00522-x

27. Bersani K, Fuller TE, Garabedian P, et al. Use, Perceived Usability, and Barriers to Implementation of a Patient Safety Dashboard Integrated within a Vendor EHR. *Appl Clin Inform*. 2020;11(1):34-45.

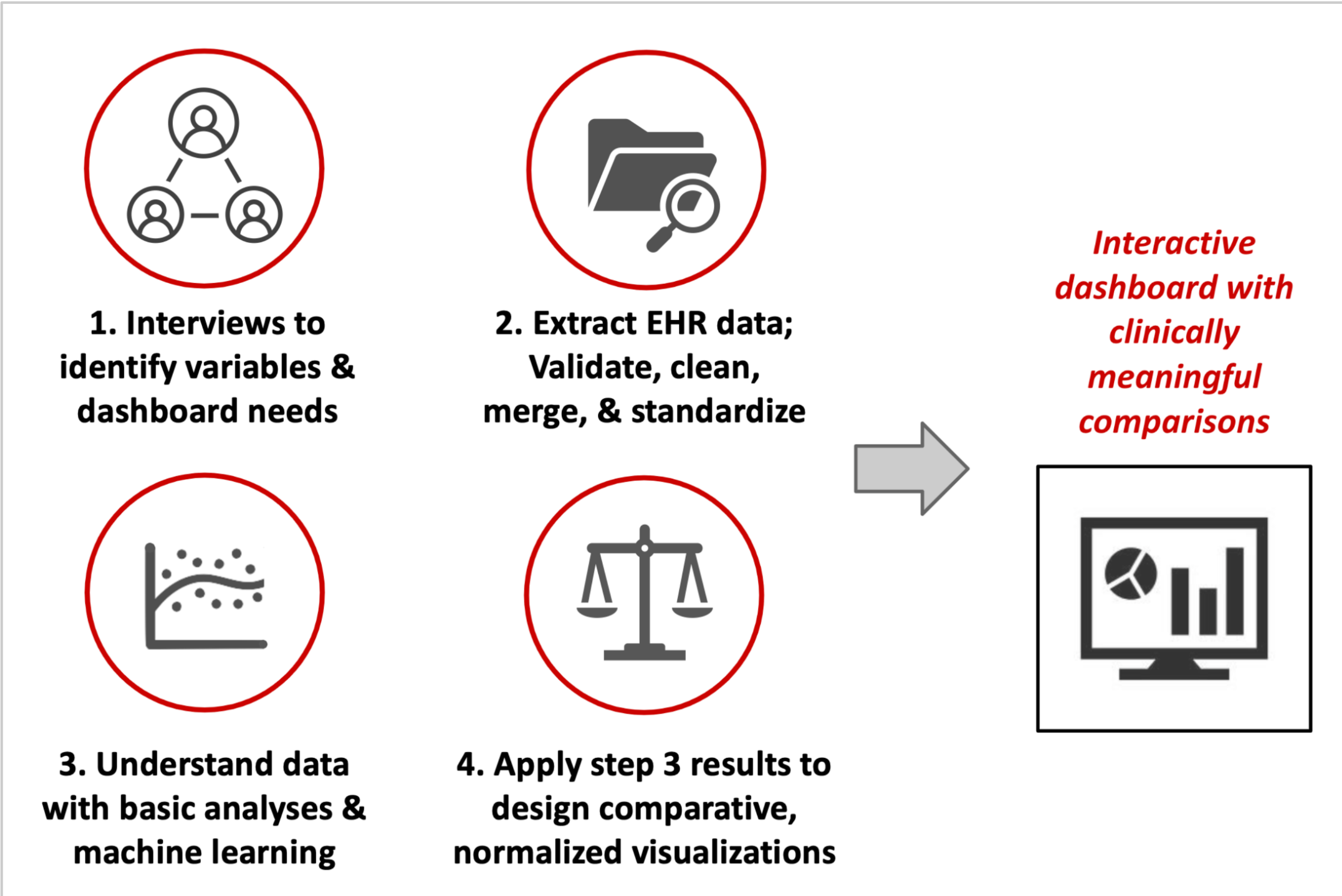
28. Mlaver E, Schnipper JL, Boxer RB, et al. User-Centered Collaborative Design and Development of an Inpatient Safety Dashboard. *Jt Comm J Qual Patient Saf*. 2017;43(12):676-685.

29. Nelson O, Sturgis B, Gilbert K, et al. A Visual Analytics Dashboard to Summarize Serial Anesthesia Records in Pediatric Radiation Treatment. *Appl Clin Inform*. 2019;10(4):563-569.

30. Epstein RH, Gratch DM, McNulty S, Grunwald Z. Validation of a system to detect scheduled drug diversion by anesthesia care providers. *Anesth Analg*. 2011;113(1):160-164.

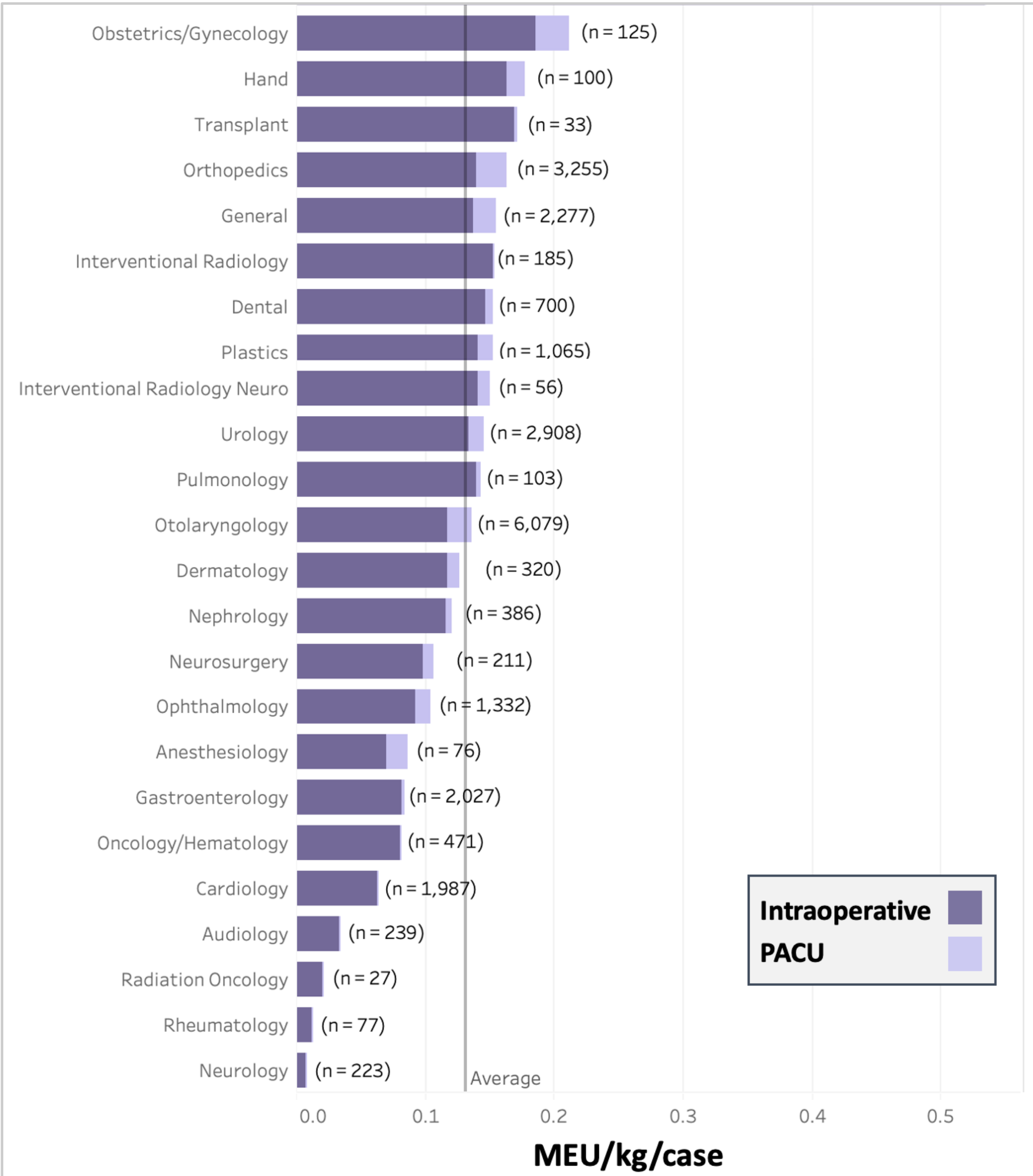
31. Ljungqvist O, Scott M, Fearon KC. Enhanced Recovery After Surgery: A Review. *JAMA Surg*. 2017;152(3):292-298.

Figure 1: Overview of the stages of development in building an opioid-utilization dashboard with clinically meaningful comparisons.



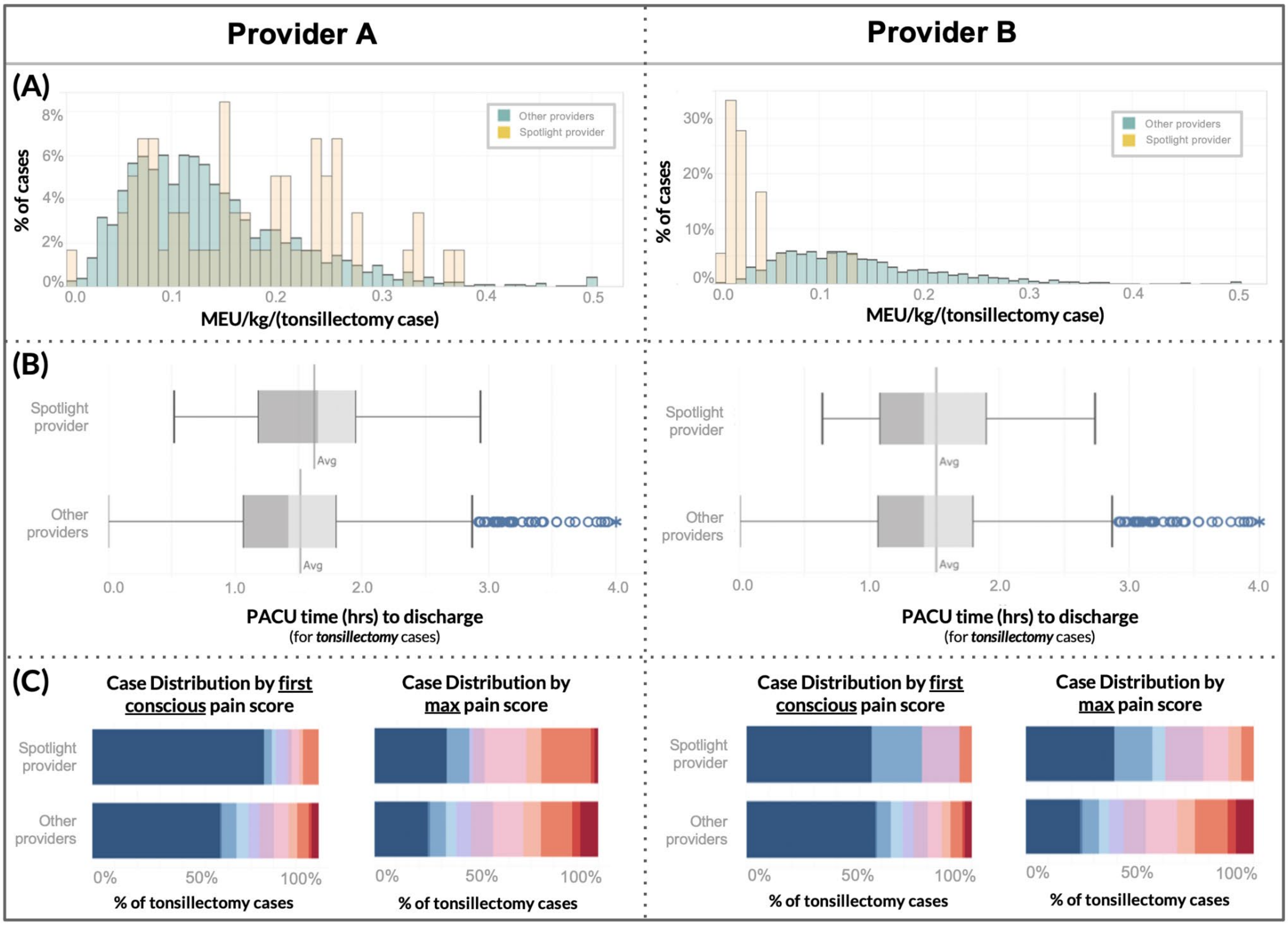
EHR, electronic health record.

Figure 2: Stacked bar graph of oral morphine equivalent units (MEU) per kg per case by service line in the operating room (intraoperatively) and post anesthesia recovery unit (PACU).



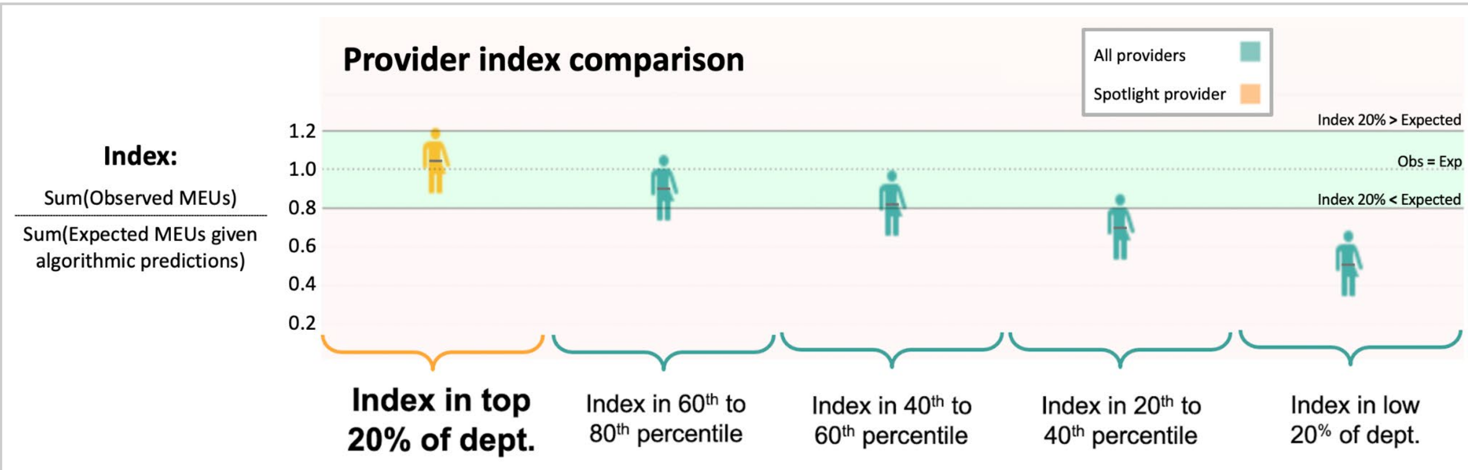
MEU, oral morphine equivalent units; kg, kilogram; PACU, post anesthesia care unit.

Figure 3: In the Provider Tab, we implement case-mix adjustment via interactive data filtration by any combination of anesthesiologist, primary procedure, service, date, and regional anesthesia. This mechanism allows clinically meaningful comparisons of individual providers to their peers: 3A) Histograms compare spotlighted provider (yellow) to peers' (green) morphine equivalent units (MEU) per kilogram per tonsillectomy cases; 3B) Box and whisker plots compare tonsillectomy post anesthesia recovery unit (PACU) discharge time for spotlighted provider's patients (above) to peers' patients ; 3C) Heat maps (blue = 0/10 pain, red = 10/10 pain) compare PACU first conscious and max tonsillectomy pain scores of spotlighted provider's patients (above) to their peers' patients.



MEU, oral morphine equivalent units; kg, kilogram; PACU, post anesthesia care unit; hrs, hours.

Figure 4: Algorithmic case-mix adjustment enables visualization and comparison of an individual provider's normalized opioid-utilization "index," which is equal to the total sum of observed utilization divided by the total sum of machine-learning expected predictions given case details.



MEU, oral morphine equivalent units; Obs, observed; Exp, expected.

Table 1: Findings from qualitative, inductive analysis of semi-structured interviews with 16 stakeholders directly or indirectly involved with pediatric perioperative care.

Stakeholder	Needs	Design implications
Managing anesthesiologists	*– To compare variation between anesthesiologists* *– To quantify impact of interventions*	– Comparisons that account for differences in case mix – Track practice relative to outcomes over time
Practicing anesthesiologists	*– To compare their practice to general distribution* *– To connect practice to outcomes*	– Comparisons that account for differences in case mix – Anonymization – Graph axes relating opioid-use patterns to patient outcomes
Surgeons	*– To understand patient outcome patterns for nausea and vomiting*	– Analysis and visualization of anti-nausea and naloxone medication patterns
Data analysts for anesthesia department	*– A tool that is easy to update and maintain*	– Code that is streamlined and well commented
PACU nurses	*– Visualization of first-conscious pain score (FCPS) distributions*	– Compare FCPS by provider

PACU, post-anesthesia care unit; FCPS, first-conscious pain score.

Table 2: Overview of outpatient surgical cases in the final case cohort.

Data Summary	
Hospital overview	
Total outpatient procedures	**24,332**
Total years encompassed	**5.33**
Anesthesiologists	**71**
Services (surgery dept.)	**23**
Primary procedure types (before grouping)	**1,014**
Primary procedure types (post-grouping)	**160**
Patient Characteristics for the included cases	
Patient age (years; median and IQR)	**7.56 [3.56-13.19]**
Patient weight (kilograms; median and IQR)	**25.3 [15.1-49.2]**
Patient gender (# male cases, percent male cases)	**14,309 [58.8%]**
Case Overview	
Duration of surgery (hours; median & IQR)	**0.55 [0.28-1.03]**
ASA class distribution (class; percent)	**I: 32.5%, II: 45.8%, III: 21.2%, IV: 0.5%**
PACU length of stay (hours; median & IQR)	**1.52 [1.15 - 1.97]**
Max PACU pain score (NRS; median & IQR)	**0.00 [0.00 - 5.00]**

IQR, interquartile range; ASA rating, American Society of Anesthesiologists (ASA) physical status classification system rating; PACU, post-anesthesia care unit; NRS, numerical rating scale.

Table S1: For case-mix normalization with machine learning, a multivariable random forest model was trained to generate "expected" MEU predictions for each surgical case. Table S1 presents all parameters included in this modeling and compares each parameter's relative influence upon intraoperative opioid utilization, as determined by the algorithm's optimal minimization of root mean squared error (RMSE). Other factors that could influence opioid administration, but that were not incorporated into this analysis, include intraoperative vital signs and hemodynamic changes.

Parameter	Data type	Relative importance with respect to RMSE reduction
Procedure duration (hours)	Numerical	124.4
Patient weight (kilograms)	Numerical	90.1
Primary procedure (52 categories + "other")	Categorical	74.8
Primary surgeon (52 most common + "other")	Categorical	55.7
Patient age (years)	Numerical	55.3
Surgical service	Categorical	55.1
Presence of intraoperative nerve block	Categorical	46.7
ASA classification	Categorical	24.0

RMSE, root mean squared error; ASA classification, American Society of Anesthesiologists physical status classification system rating.