

An Electronic Medical Record-Derived Individualized Performance Metric to Measure Risk-Adjusted Adherence with Perioperative Prophylactic Bundles for Health Care Disparity Research and Implementation Science

Michael H. Andreae¹ Stephan R. Maman^{1,2} Abraham J. Behnam¹

¹Department of Anesthesiology, Penn State Milton S. Hershey Medical Center, Hershey, Pennsylvania, United States

²Penn State College of Medicine, Hershey, Pennsylvania, United States

Address for correspondence Michael H. Andreae, MD, Department of Anesthesiology, Penn State Milton S. Hershey Medical Center, 500 University Drive, Hershey, PA 17033, United States (e-mail: mandreae@pennstatehealth.psu.edu).

Appl Clin Inform 2020;11:497–514.

Abstract

Background Health care disparity persists despite vigorous countermeasures. Clinician performance is paramount for equitable care processes and outcomes. However, precise and valid individual performance measures remain elusive.

Objectives We sought to develop a generalizable, rigorous, risk-adjusted metric for individual clinician performance (MIP) derived directly from the electronic medical record (EMR) to provide visual, personalized feedback.

Methods We conceptualized MIP as risk responsiveness, i.e., administering an increasing number of interventions contingent on patient risk. We embedded MIP in a hierarchical statistical model, reflecting contemporary nested health care delivery. We tested MIP by investigating the adherence with prophylactic bundles to reduce the risk of postoperative nausea and vomiting (PONV), retrieving PONV risk factors and prophylactic antiemetic interventions from the EMR. We explored the impact of social determinants of health on MIP.

Results We extracted data from the EMR on 25,980 elective anesthesia cases performed at Penn State Milton S. Hershey Medical Center between June 3, 2018 and March 31, 2019. Limiting the data by anesthesia Current Procedural Terminology code and to complete cases with PONV risk and antiemetic interventions, we evaluated the performance of 83 anesthesia clinicians on 2,211 anesthesia cases. Our metric demonstrated considerable variance between clinicians in the adherence to risk-adjusted utilization of antiemetic interventions. Risk seemed to drive utilization only in few clinicians. We demonstrated the impact of social determinants of health on MIP, illustrating its utility for health science and disparity research.

Conclusion The strength of our novel measure of individual clinician performance is its generalizability, as well as its intuitive graphical representation of risk-adjusted individual performance. However, accuracy, precision and validity, stability over time, sensitivity to system perturbations, and acceptance among clinicians remain to be evaluated.

Keywords

- ▶ health care disparity
- ▶ perioperative care
- ▶ postoperative nausea and vomiting
- ▶ implementation science
- ▶ electronic medical record

received
March 24, 2020
accepted after revision
June 1, 2020

© 2020 Georg Thieme Verlag KG
Stuttgart · New York

DOI <https://doi.org/10.1055/s-0040-1714692>.
ISSN 1869-0327.

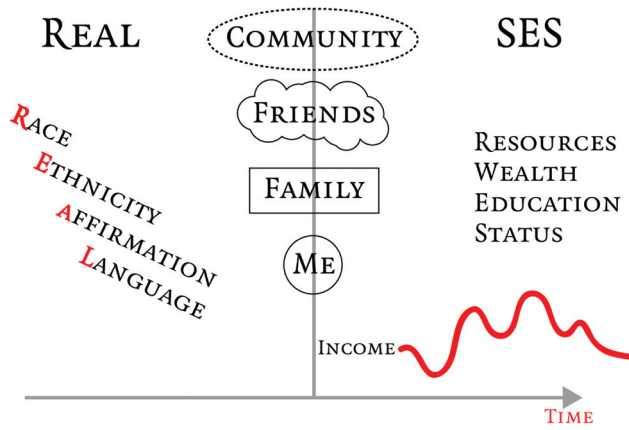


Fig. 1 Social determinants of health. Up to 80% of preventable morbidity and mortality in the United States may be attributable to social, economic, and behavioral factors,¹ including in perioperative medicine.^{5,7,67,68} The conditions in which people are born, live, learn, work, play, worship, and age are termed social determinants of health.³ (R)ace, (E)thnicity, (A)ffiliation, and (L)anguage (REAL) pertain to identify attributes. Socioeconomic status (SES) encompasses scholastic achievement, wealth, social status, and class but also access to resources or knowledge. REAL and SES both matter.¹⁴ Some determinants are more stable, e.g., gender identity, and others fluctuate, e.g., income. In their impact of health and outcomes, individual-level attributes are as important as family-level or community-level determinants of health.³

Background and Significance

Social determinants of health (SoDHs), e.g., socioeconomic status, race, ethnicity, and gender, mediated through individual clinician and/or systemic biases lead to drastically inferior perioperative outcomes.^{1,2} **Fig. 1** conceptualizes SoDHs, the conditions in which people are born, live, learn, work, play, worship, and age.³ With a concrete example, **Fig. 2** instantiates mechanisms through which SoDHs can lead to inferior perioperative outcomes, drawing on the concept of fundamental causes of disease¹; we also illustrated this in a figure elsewhere.⁴ Our goal is to offer individual clinicians, teams, and institutions meaningful feedback on those process disparities which lead to subpar perioperative outcomes.⁵⁻⁸

Process and performance are critically important for patient safety and equitable outcomes.^{5,9-11} Rigorous assessment of individual and team performance based on actual clinical care remains elusive.¹² Potential applications include individualized feedback, reporting for pay for performance (P4P),⁴ health systems science (HSS) implementation,¹³ and health care disparity^{1,4,6,7,14} and outcomes research. HSS investigates delivery of care and implementation of quality improvements, e.g., predictive clinical decision support (CDS).¹⁵

Process and outcome measures are already employed to evaluate patient care, processes, and CDS, although with mixed

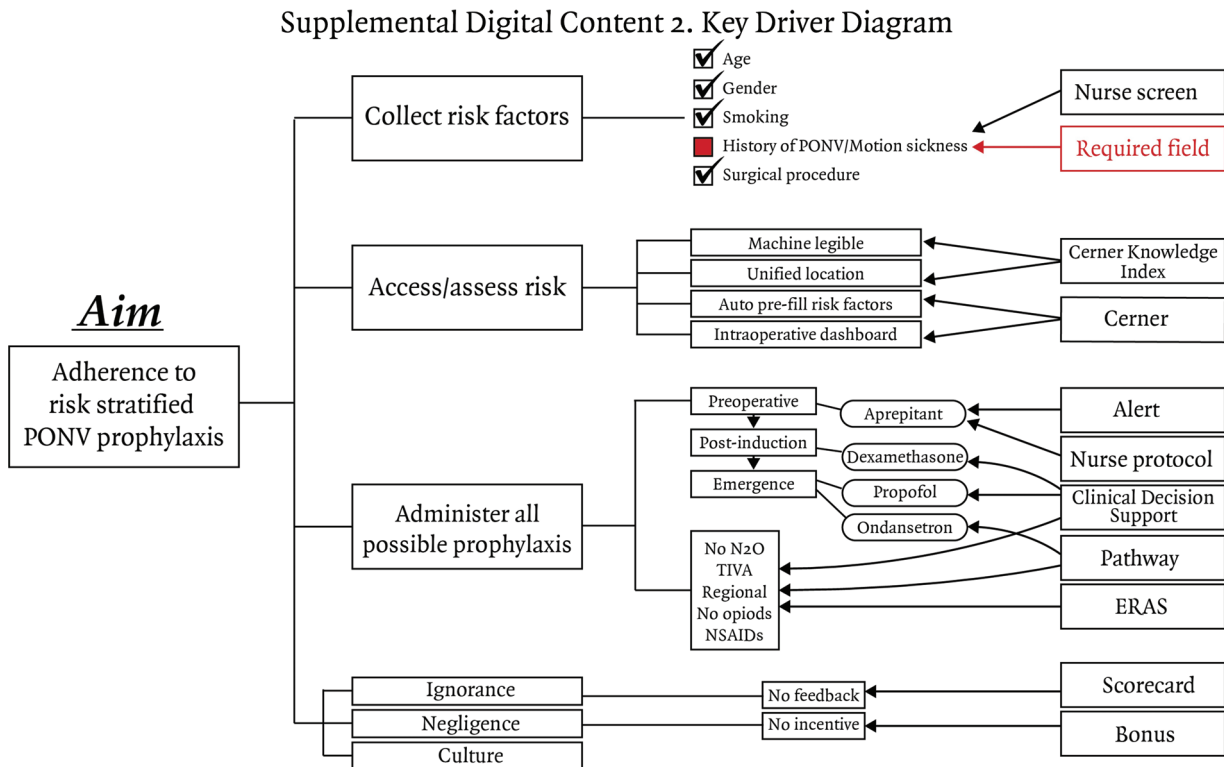


Fig. 2 Key drivers behind adherence to guidelines on postoperative nausea and vomiting (PONV) prophylaxis. Important mechanisms that may lead to the disparities in antiemetic prophylaxis are displayed in this key driver diagram. Collection of risk factors is a prerequisite for risk-adjusted utilization of interventions to prevent PONV. Risk elicitation can be hindered by individual or systemic bias, e.g., language barriers, clinician negligence, lack of insurance curtailing preoperative screening, etc. Uniform and standardized collection of risk provides reliable information for clinical decision making. Still, patient risk data can drive appropriate PONV prophylaxis only if individual risk is accessible in the electronic medical record and leveraged in automated clinical pathways, which mitigate potential individual clinician bias. Clinical decision support, standardized nursing protocols, and enhanced recovery programs are all suited to enhance comprehensive and equitable adherence to guidelines. Granular clinician- or team-level metrics of risk-adjusted adherence can identify localized under- and overutilization and focus on targeted feedback, training, and process redesign on objective identifiable process disparities.

Table 1 Ideal metric characteristics

Aspects	Current metrics	Ideal metrics	Explanation	Example	Comment
Dimensionality	Unidimensional	Multidimensional	Measures just <i>one</i> aspect or action in the care process	Glucose level was measured in diabetic patients before the surgery (Y/N)	Glucose measurement is important, but only one small aspect of appropriate perioperative management of a diabetic patient
Range of instrument	Floor and ceiling effects	Relevant performance differences within the responsive range	The restricted range of the metric limits its utility to discriminate	Percentage of cases with preoperative antibiotic administered	With a target of 100% adherence, differences between 99 and 98% may not be informative
Risk adjustment	Not adjusted for risk	Risk adjusted	Care should be tailored to risk, with changing risk leading to more or different interventions	Contingent on patient PONV risk, the number of pertinent treatment options are considered	Responsiveness to risk is an important feature of process quality in medicine
Attribution	Outside scope of influence	Irrefutably attributable to individual clinician	Clinicians should not be evaluated on aspects of care they are not responsible for	On-time procedure start may be delayed due to events outside beyond one's control	Antiemetic prophylaxis as the sole domain of anesthesiologists, is attributable to the anesthesia clinician
Discrimination	Poor differentiator of "good" vs. "bad" care quality	Classifies care quality reliably and effectively	Precision, accuracy, and responsiveness influence the ability of a metric to classify	Postsurgical infection rates fluctuate and depend on multiple process factors	Improving discrimination often increases time, effort, and cost of the assessment
Cost	Manual or additional documentation or extraction drives cost of classical metrics	Automated extraction of processes already documented during routine delivery of care	Curated electronic clinical records with controlled vocabulary facilitate metrics	In most EMRs, PONV risk and antiemetics are documented routinely	Automated process evaluation starts with machine-legible process documentation

Abbreviations: EMR, electronic medical record; PONV, postoperative nausea and vomiting.

Note: This table enumerates and contrasts characteristics of current suboptimal versus ideal process metrics with examples and comments in the right two columns. Constraint by the cost and effort going into documentation, extraction and cleaning of process data, many currently employed perioperative metrics are not ideal. The Multicenter Perioperative Outcomes Group has a quality-improvement arm called ASPIRE, which offers individualized monthly feedback to clinicians on their actual performance with regards to several quality measures.⁶² For example, MPOG reports on the number of times the clinician treated high glucose with a time window, focused on *one* dimension of perioperative diabetic management, failed for example to consider if glucose levels were checked when they should be, or how high glucose levels were. With a target of 100% adherence, such a metric may lack the range and responsiveness to effectively distinguish quality care from subpar performance. Ideal metrics adjust for patient risk in considering the number or quality of interventions. Risk responsiveness may be one of the most important characteristics of good-quality metrics. Finally, clinicians can only be held accountable for processes they are controlling. Irrefutable attribution is hence likely important for clinical acceptance of a performance metric. Together with procedural justice, the aforementioned characteristics have been shown to enhance physician responses to feedback.⁶⁹

results.¹⁶ Despite these efforts, the performance of individual clinicians in the operating room is virtually nondescript to payers, patients, and even the clinicians themselves.^{17,18} **→Table 1** enumerates characteristics of ideal metrics and shortcomings of current unidimensional metrics for individual clinician performance (MIPs). Many suffer from floor and ceiling effects, lack responsiveness, and are not adjusted for risk.¹⁸ Ideal performance measures should be relevant, valid, multidimensional, risk-adjusted, irrefutable, and able to discriminate delivery of care by individual clinicians and across institutions contingent on relevant predictors, e.g., SoDHs.^{19,20} A metric derived from the electronic medical record (EMR) would be scalable to larger cohorts and national registries like the Multicenter Clinical Outcomes Group (MPOG).²¹ Finally, MIPs should focus on actions *individual* clinicians are accountable for, e.g., for anesthesiologists risk-adjusted prophylaxis of postoperative nausea and vomiting (PONV).⁵

To introduce a novel performance metric, we leverage antiemetic prophylaxis as both a reflective and a formative construct of perioperative clinician performance and process quality.^{5,22} If severe, PONV is associated with complications such as aspiration or wound dehiscence, and costly unplanned

hospital admissions.²³ Process metrics are reflective if they represent latent, underlying quality, e.g., risk-adjusted antiemetic prophylaxis may reflect overall anesthesia quality. Antiemetic prophylaxis is also a formative construct as it measures a phenomenon of interest, i.e., effective PONV prevention.²⁴

We discussed elsewhere additional aspects that make antiemetic prophylaxis suitable as one exemplary risk-adjusted metric of perioperative process quality and equity, demonstrating in particular that after controlling for PONV risk, fewer antiemetics are administered in patients of lower socioeconomic status.⁵ We hope that risk-adjusted personalized feedback might induce (anesthesia) clinicians to adhere better to prophylactic guidelines and especially provide more equitable care.^{12,25–31}

Objectives

Our objective is to demonstrate the utility and feasibility of a new risk-adjusted, multidimensional, visual, and intuitive performance metric, automatically derived from the EMR, to serve as a reflective and/or formative construct of perioperative process quality at the individual clinician, team, and institutional level.³²

Hypothesis

Our hypothesis is that risk-adjusted adherence to established PONV guidelines can be leveraged for personalized visual feedback on the (equitable) performance of individual clinicians and to investigate health care delivery processes quantitatively.

Methods

First in concept, we conceptualize and improve our MIP to demonstrate the utility for personalized feedback on (equitable) performance (→Fig. 3 and →Supplementary Fig. S1 [available in the online version]). Second in case study, we describe an observational prospective cohort case study on perioperative antiemetic prophylaxis to demonstrate the utility and feasibility of MIP for quantitative comparative HSS, following the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines.³³ In sta-

tistical modeling, we propose modeling approaches, but we did not perform a quantitative statistical analysis.

Concept

Risk Responsiveness as a Generalizable Individual Performance Metric

We developed a generalizable visual concept to measure individual provider performance (MIP) for personalized feedback and HSS. Individual clinician performance is conceptualized as risk responsiveness in utilizing (prophylactic) interventions, for example, anesthesia clinicians select and administer PONV medications according to the presence of PONV risk factors (young age, nonsmoker, previous PONV history, and female gender). This is assessed as a count of separable intervention items under the control of the individual clinicians administered contingent on the number of risk factors motivating the (prophylactic) intervention. As a concrete clinical example, we

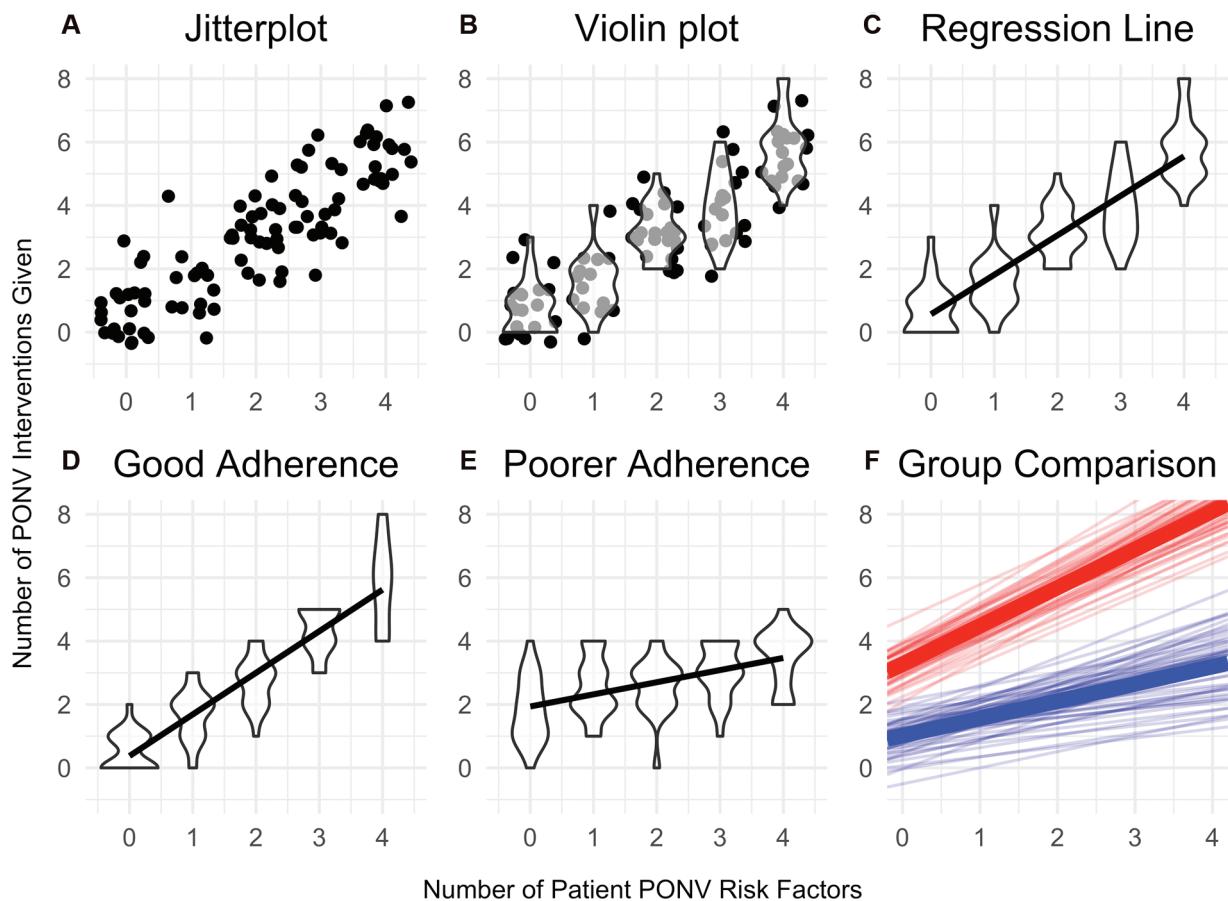


Fig. 3 Performance metric conceptualization. Panel (A) visualizes risk-adjusted adherence to guidelines for an individual clinician by plotting anesthesia cases with PONV risk in the x-axis and the number of interventions in the y-axis (simulated data). The dots represent individual anesthesia case and are “jittered” to expose overlapping data points. The distribution of the number of interventions administered in each risk category is better summarized in panel (B) by overlaying a violin plot onto the “jitter plot.” Finally, a fitted regression line provides a summary measure for the individual clinician’s risk-adjusted utilization of PONV prophylaxis in panel (C). The intercept and slope of the regression line can serve as quantitative measures of risk responsiveness of an individual clinician. Panel (D) demonstrates good adherence. The clinician responded to increased patient risk with additional interventions for high-risk patients, leading to a low intercept and a steep regression line. Less discriminate administration of prophylaxis is evident in panel (E), with a high intercept and an almost flat regression line indicating indiscriminate utilization of prophylaxis following a “one-size-fits-all” approach. The graphical representation of our proposed measure of individual clinician performance is both quantitative and intuitive, allowing us to compare adherence of two distinct groups (red vs. blue) of individual clinicians (*thin lines*) and institutional averages (*darker thicker lines*) visually and numerically, as shown in panel (F). PONV, postoperative nausea and vomiting.

consider the administration of antiemetic interventions contingent on known and accepted patient risk factors for PONV as a formative and reflective measure of clinician-specific anesthesia quality, visualized with simulated data in ▶Fig. 3. For quantitative comparison between clinicians, the slope and the intercept of the regression line across a scatter plot can serve as a numerical reflection on how the individual clinician on average adjusts the number of interventions to match the risk of the individual patient (▶Fig. 3).

Comparing Risk Responsiveness between Individual Clinicians

In ▶Fig. 3, we contrast the good adherence of one clinician (panel D) against the poorer adherence of another clinician in panel E (using simulated data to illustrate the concept). While the clinician in panel D on average increases the number of prophylactic interventions contingent on the risk profile of the patients, the clinician in panel E is treating high-risk cases with a similar number of interventions as low-risk cases, which is not responsive to the risk of the individual patient. For example, as illustrated in ▶Fig. 4, an older male smoker with no history of PONV (risk = 0) requires little or no antiemetic prophylaxis (recommended number of interventions = 0). On the other hand, a young female nonsmoker with a history of PONV (risk = 4) would benefit from receiving several interventions (ideal number of interventions = 6–10) to reduce her risk of PONV.^{22,30,34}

Our MIP concept allows us to investigate and contrast risk responsiveness of groups, (between institutions or before or after an intervention to enhance adherence), and to explore team dynamics. In ▶Fig. 3(F), we contrast two groups of clinicians from different institutions using simulated data. Local

culture and accessibility or availability of individual interventions can impact the observed performance (responsiveness to individual patient risk) of individual clinicians, leading to flatter regression curves in one institution and steeper response curves in others. Additionally, we might compare clinician performance before versus after an intervention, e.g., the implementation of CDS to improve adherence with guidelines or targeted feedback. Finally, we can explore team performance and dynamics in clinical dyads, in other words the adherence of clinical care team members (residents and nurse anesthetists) under the supervision of attending physicians (▶Fig. 5).

Comparing Process Equity between Clinicians

We leverage the new metric to investigate process disparity at the level of the individual clinician, who may treat different groups of patients differently contingent on patients' SoDHs. SoDHs are explained in ▶Fig. 1. A clinician may *on average* administer less prophylaxis to minority patients versus white patients or contingent on insurance status,⁵ as suggested based on clinical data in ▶Fig. 4.

Case Study

In our case study, we tested the MIP concept with clinical data on risk-adjusted perioperative antiemetic prophylaxis derived from the EMR.

Data Extraction

We tested the concept of MIP delineated above in clinical data obtained from the perioperative EMR (Cerner, Kansas City, Missouri, United States) and associated billing systems, using SAP BusinessObjects (SAP, Walldorf, Baden-Württemberg, Germany). To facilitate retrieval, data were assigned a Cerner

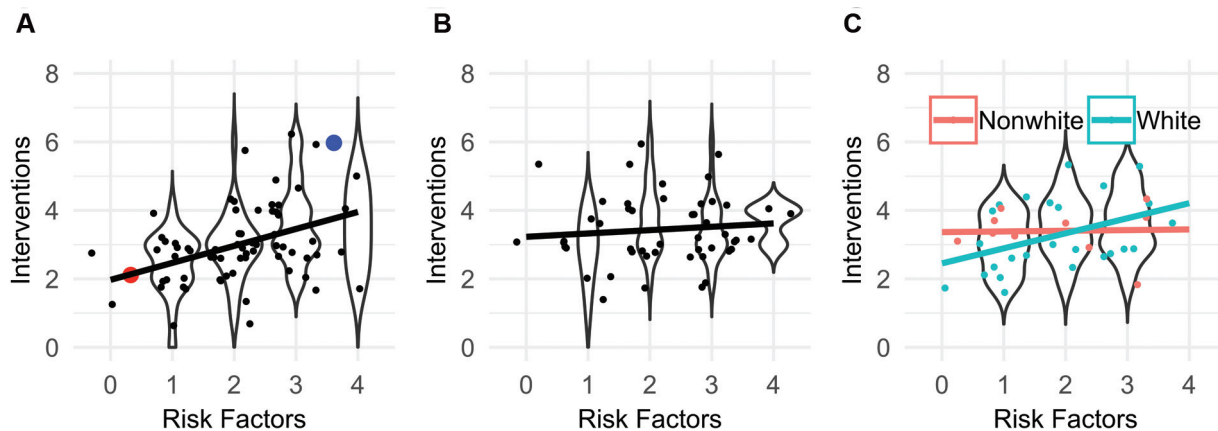


Fig. 4 Performance disparity between individual performances of several clinicians, employing the metric proposed. Extracting clinical data from the electronic medical record of Penn State Milton S. Hershey Medical Center, we compare the risk-adjusted adherence to PONV prophylaxis of our anesthesia clinicians using our individual clinician performance metric (MIP), as explained in ▶Fig. 3. The x-axis indicates PONV risk, the y-axis the number of antiemetic interventions administered for an individual patient, each case represented by a dot. Panel (A) shows a risk-responsive clinician with good adherence, as the clinician administers more interventions contingent on patient risk. For a low-risk case (light red dot at [0, 2]) fewer interventions are administered than for a high-risk patient (the dark blue dot at [4, 5]), leading to a steep regression line. Panel (B) shows a clinician less discriminate in the utilization of prophylaxis, summarized by a MIP regression line that is rather flat. In panel (C), we leverage the new metric to investigate process disparity at the level of the individual clinician. Panel (C) shows how a specific clinician treats different groups of patients differently when we grouped patients by social determinants of health (SoDHs), in this case minority patients versus white patients. SoDHs are explained in ▶Fig. 1. However, we urge caution in the *prima facie* interpretation of the suggested individual clinician bias leading to the observed disparity. Validity and stability of our model are still to be confirmed. Furthermore, local institutional effects may dominate the performance of an individual clinician, meaning the constraints of the health system may overwhelm or force the hand of the individual. PONV, postoperative nausea and vomiting.

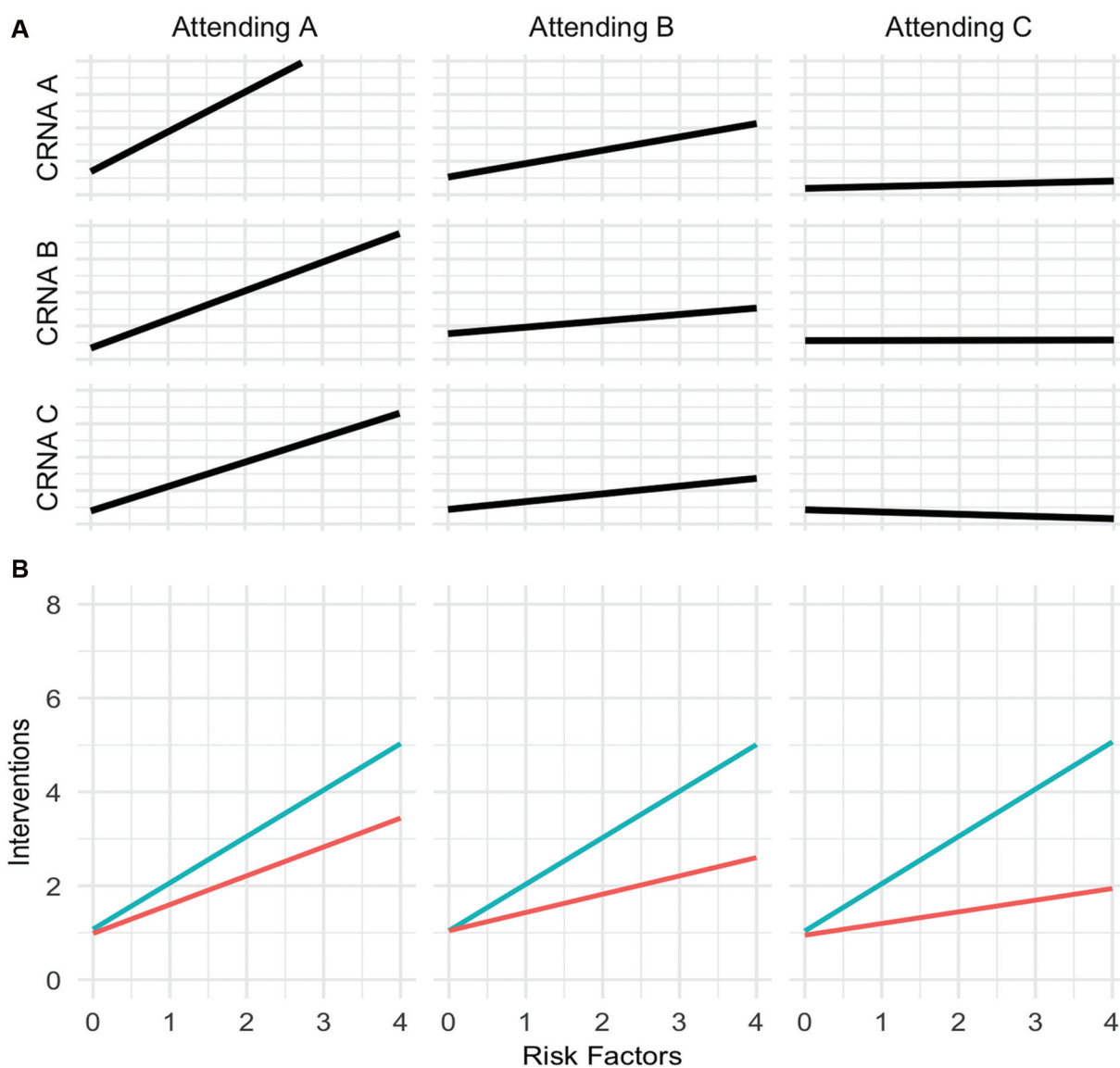


Fig. 5 Effect of supervision on performance. Panel (A) demonstrates how different attending physicians may influence the performance of those working under their direct supervision, such as nurse anesthetists (or residents) or vice versa. For each clinician dyad (attending/certified registered nurse anesthetist [CRNA]), risk is in the x-axis against the number of interventions in the y-axis. Better adherence to risk-adjusted prophylaxis is represented by a steeper regression line. The supervising attending (A, B, or C) may be exerting a heavy influence on the performance of CRNA A, significantly improving adherence compared with attending C, but attending B seems less influential (simulated data). Panel (B) (similar axes) illustrates how a supervising attending may attenuate or accentuate the disparity in adherence to risk-adjusted utilization of interventions in a given CRNA, contingent on social determinants of health (simulated data). Our proposed metric could be used in team systems to investigate the impact of system perturbations or patient characteristics on perioperative clinician dyads. But the question remains: who is influencing whom? May be the nurse anesthetists and residents are in fact “managing up,” i.e., exerting an influence on the attendings they work with? (The figure is specifically created to suit the above explanation by making the slope of the regression line of the supervised provider contingent on the supervising attending physician.)

Knowledge Index. Clinical data were documented during the routine clinical care of adult patients undergoing elective surgery at the Penn State Milton S. Hershey Medical Center (PSHMC) through various Cerner “PowerCharts” and “PowerForms” capturing most patients scheduled for elective surgery (→ **Supplementary Fig. S2**, available in the online version). We extracted antiemetic risk and antiemetic interventions from anesthesia records for the 25,980 patients who underwent elective surgery in the period between June 3, 2018 and March 31, 2019.

The unit of analysis for the performance metric is the individual anesthesia clinician (an attending physician or nurse anesthetist). We identified individual anesthesia clinicians from the billing record. We de-identified clinicians at data extraction by assigning them a random integer to protect their privacy. For cases where more than one clinician was providing the anesthesia service, we attributed the case to the first anesthesiologist attending on the case. On the same or subsequent visits, a patient may undergo anesthesia repeatedly. We counted these as separate observations,

without statistical adjustment for repeated observations. This simplification reflects the regulatory constraints of working with de-identified data.

We counted the pre- and intraoperative administration of dexamethasone, ondansetron, aprepitant, and scopolamine as pharmacological prophylaxis; we counted the pre- and intraoperative use of acetaminophen and ketorolac, regardless of the time or route of administration as adjuvant medications to reduce the opioid requirements. We counted the use of total intravenous anesthesia, the complete avoidance of nitrous oxide, and the use of regional blocks or neuraxial anesthesia (detected as intraoperative use of bupivacaine, chloroprocaine, or ropivacaine), among the intraoperative anesthesia choices. The count of prophylactic antiemetic interventions ranged from 0 to 10. We extracted four patient-specific risk factors for PONV from the EMR: patient younger than 50 years of age, female sex, nonsmoking status, and positive history for PONV were counted as risk factors, resulting in a final risk score for each patient ranging from 0 to 4. The MIP visualizes the individual performance of each clinician in **Supplementary Fig. S3** (available in the online version).

Exploring Social Determinants of Health to Elicit Variability in Health System Processes

As an exploratory analysis, patient-reported billing-derived SoDHs were obtained to differentiate how clinicians may adjust their actions to patient risk differently in patient groups defined by social (as opposed to medical) indicators. We extracted self-reported race from billing records, despite recognized limitations of billing records on race.^{3,35} We applied our MIP concept to visually inspect the effect of race on the risk responsiveness of individual clinicians (**Fig. 4** and **Supplementary Fig. S4**, available in the online version).⁵

Statistical Modeling

The complexities of modeling the proposed MIP are sketched briefly in **Supplementary Appendix A** (available in the online version), but attrition limited our available data. Instead we focus on demonstrating feasibility and utility of the conceptual approach.

While we considered the Conway–Maxwell–Poisson (CMP) distribution, a generalization of the Poisson distribution, hierarchical modeling (with several levels) of a positive integer outcome using discrete distributions like the CMP can become complicated with the available software. Instead we employed a simplex vector and the concept of majorization (an ordering on nonnegative vectors), to model the MIP in the typical *hierarchical* context of contemporary medical care in the United States, as described in greater detail and more formally in **Supplementary Appendix A** (available in the online version), developed in cooperation with Jonah Gabry and Benjamin Goodrich of GG services, New York, United States.

Software Used

We used R/Rstudio, and packages knitr, readxl, dplyr, and ggplot2 for data importing, wrangling, and visualization, respectively.^{36–40}

Results

Data Extraction and Cleaning

We extracted data from the EMR on risk-adjusted utilization of antiemetic prophylaxis from 25,980 elective anesthesia cases performed at PSHMC in the time between June 3, 2018 and March 31, 2019. The flow diagram in **Fig. 6** depicts the number of cases with complete data for the various subsets in the data cleaning and filtering process of our clinical cohort. Of the total 25,980 cases, only 3,969 cases had complete data regarding all patient-risk factors for PONV. Of these, all had complete data detailing the interventions considered in our analysis. We limited the data to cases with anesthesia Current Procedural Terminology (CPT) codes as defined by MPOG for their performance metric on PONV, leaving 2,211 cases to implement and test our MIP. All of these cases had an individual (first) anesthesia clinician identified. We enumerated the number of cases by clinician, dividing the population into three groups according to the number of cases with complete data for an individual clinician: 8 anesthesia clinicians had complete data on over 50 unique anesthesia cases, 43 anesthesia clinicians on over 25 cases, and 61 on over 10 cases. For our exploratory analysis, we further limited the data to those 2,210 cases with complete data on race and focused on the 26 anesthesia clinicians providing anesthesia on greater than or equal to 5 minority patients. Typically, clinicians administer approximately 500 cases per year.

Patient Demographics and Case Characteristics

Patient demographics and case characteristics are tabulated in **Table 2**. The proportion of male patients was 40.8%. The ages of patients ranged from 18 to 90, with a median of 48 and interquartile range of 27. We excluded patients younger than 18 years of age. The most commonly reported race was white (86.6%). In total, 1% identified as Asian American, 5.7% as African American, and 6% as other minority or of mixed descentance. By the anesthesia CPT code, the highest fraction of case was lower or upper abdominal procedures (19.9%). By primary surgical CTP codes, the most frequent procedures in our cohort were surgery of the lower abdomen 00840 (13.9%), surgery of the upper abdomen 00790 (6%), and surgery on the chest wall and shoulder girdle 00400 (5.5%).

Patient Risk Factors

The individual and aggregate patient risk factors and antiemetic interventions administered are summarized in **Table 3**. The patient risk factors for PONV ranged from 0 to 4. After calculating the risk scores for patients, the median number of risks factors was 3 with an interquartile range of 2. 57 patients reported no risk factors, 565 reported one, 827 two, 648 three, and 114 four risk factors, which is representative of the general population undergoing anesthesia in the United States. In total, 85.3% of the patients were identified as nonsmokers, 12.8% reported a history of PONV, 59.2% had female gender as a risk factor for PONV, and 51.7% were under 50 years of age, increasing their risk of PONV, as well.

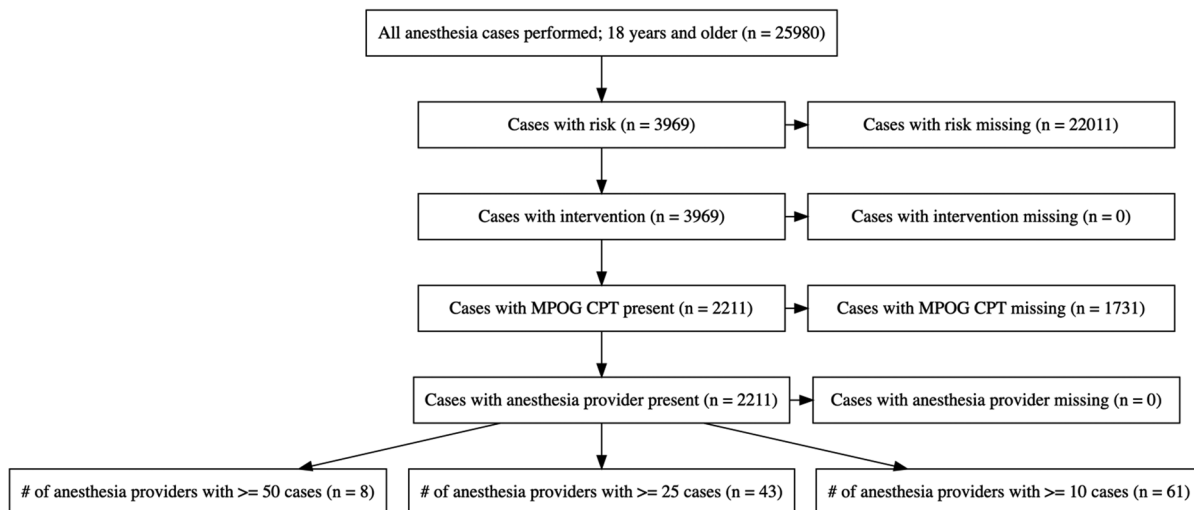


Fig. 6 Flow diagram. The flow diagram depicts the data-cleaning steps involved in filtering the cohort of Penn State Milton S. Hershey Medical Center anesthesia cases. Missing data hindered our analysis and only 2,211 cases of the total 3,969 anesthetics delivered had complete data regarding all patient-risk factors for PONV and antiemetic interventions administered. MPOG defined anesthesia CPT codes for their PONV performance metric and limiting our data to these left 2,211 cases to implement and test our individual clinician metric. The individual (first) anesthesia clinician was identified in all cases. Dividing the population into three groups according to the number of cases with complete data for an individual clinician, we enumerated the number of cases by clinician: 8 anesthesiologists had complete data on over 50 unique anesthesia cases, 43 anesthesia clinicians on over 25 cases, and 61 on over 10 cases. For our exploratory analysis, we further limited the data to those 2,210 cases with complete data on race and focused on the 26 anesthesia clinicians providing anesthesia on more than 5 patients. CPT, Current Procedural Terminology; MPOG, Multicenter Clinical Outcomes Group; PONV, postoperative nausea and vomiting.

Utilization of Antiemetic Prophylaxis and Anesthesia Choices

Patient risk factors, antiemetic interventions, and anesthesia choices are tabulated in detail in [Table 3](#). Regional anesthesia was administered in 3.1% of the cases. Preoperative antiemetic prophylaxis with aprepitant was administered in 109 cases (4.9%). Scopolamine was provided in 57 cases (2.6%). Intraoperatively, 1,602 cases received dexamethasone and 1,880 cases received ondansetron. In 12.3% of cases, no volatile anesthetic (i.e., sevoflurane, the commonly used agent at our institution) was administered and in 12.9% of cases no nitrous oxide was administered, both choices reducing the risk of PONV. Ketorolac was administered in 477 cases (21.6%), while acetaminophen was administered in 1,649 cases (74.6%), with their opioid sparing effects reducing the risk of PONV. The median of the cumulative number of interventions given was 4, with an interquartile range of 2.

[Supplementary Fig. S3](#) (available in the online version) summarizes clinician risk-adjusted adherence to antiemetic prophylaxis in our clinical cohort with a summary line in red for all clinicians. Uneven distribution across PONV risk categories is evident as well as increasing utilization of prophylaxis with risk. We also observed considerable variation among clinicians as to what extent they respond to PONV risk in their individual patient with additional prophylactic interventions.³⁴ The overall slope of responsiveness to risk is disappointingly flat, suggesting a somewhat undiscerning utilization of PONV prophylaxis, which we discuss critically further below.

Individual Performance Variability

In [Fig. 4](#), panels A and B provide exemplary plots of the proposed MIP, displaying the variable risk-adjusted utili-

zation of antiemetic prophylaxis for two individual clinicians. In [Supplementary Fig. S1](#) (available in the online version), a faceted plot by individual attending anesthesiologist, we report the MIP graphically for those anesthesia clinicians with complete information on greater than or equal to five cases, again showing tremendous variability in how individual clinicians respond to the same patient risk. Each panel represents a unique clinician. Panels for anesthesia clinicians are organized in a descending order from highest to lowest number of anesthesia cases per attending anesthesiologist.

Exploratory Analysis on the Impact of Social Determinants of Health on Individual Clinician Performance

In panel C of [Fig. 4](#), we color-coded cases by the self-reported race of the patient for an individual clinician, as an example of the possible utility of our performance metric in investigating the impact of patient SoDHs on perioperative processes, with details down to the granular level of individual clinician on the patient-specific risk-adjusted utilization of a prophylactic bundle to prevent PONV.⁴¹ It appears that the clinician in panel C is *systematically* administering more interventions in cases with white patients compared with cases with minority patients. Additional examples for individual clinician performance, differentiated by patients' SoDHs, are shown in [Supplementary Fig. S4](#) (available in the online version).

Quantitative Analysis

Bayesian hierarchical models, described in [Supplementary Appendix A](#) (available in the online version) and ran on a personal laptop in parallel with four processors, were converging effectively in minutes. However, we chose not to report the result of our quantitative analysis. We had

Table 2 Patient demographics and case characteristics

	Total (%)
Total anesthetics	2,211
Male gender (%)	901 (40.8)
Age (%)	
18–49	1,142 (51.7)
50–64	607 (27.5)
65–79	360 (16.3)
80+	102 (4.6)
Patient race (%)	
American Indian/Alaskan Native	6 (0.3)
Asian	22 (1.0)
Black or African American	126 (5.7)
Native Hawaiian or Pacific Islander	1 (0.0)
Other race	105 (4.7)
Two or more races	28 (1.3)
Unavailable	8 (0.4)
White	1,914 (86.6)
NA	1 (0.0)
Anesthesia CPT codes (top 10) (%)	
00840 [surgery lower abdomen]	307 (13.9)
00790 [surgery upper abdomen]	132 (6.0)
00400 [skin, ext/per/atrunk]	122 (5.5)
01480 [open procedures on bones of lower leg, ankle, foot]	91 (4.1)
00910 [bladder surgery]	87 (3.9)
00300 [head/neck/ptrunk]	75 (3.4)
01961 [cesarean delivery]	69 (3.1)
00320 [neck organ, 1 and over 0]	62 (2.8)
00520 [closed chest procedures]	56 (2.5)
01230 [surgery of femur upper two-thirds]	56 (2.5)
Other	1,154 (52.2)
Surgical procedures (top 10) (%)	
ANESTH SURG LOWER ABDOMEN	307 (13.9)
ANESTH SURG UPPER ABDOMEN	140 (6.3)
ANESTH SPINE CORD SURGERY	132 (6.0)
ANESTH SKIN EXT/PER/ATRUNK	122 (5.5)
ANESTH LOWER LEG BONE SURG	91 (4.1)
ANESTH BLADDER SURGERY	87 (3.9)
ANESTH HEAD/NECK/PTRUNK	75 (3.4)
ANESTH CS DELIVERY	69 (3.1)
ANESTH NECK ORGAN 1YR/>	60 (2.7)
ANESTH LOWER ARM SURGERY	56 (2.5)
Other	1,072 (48.5)

Abbreviations: CPT, Current Procedural Terminology; NA, not available. Note: This table summarizes the demographics of the patients who received anesthesia, tabulating their age, race, ethnicity, and gender. The unit of analysis is the anesthesia record and de-identification of patients meant that individual patients who received several anesthetics might have been double counted in our study. The table also summarizes anesthesia CPT codes and the surgical procedures to describe the most frequent surgeries these patients underwent.

Table 3 Patient risk factors, antiemetic interventions, and anesthesia choices

	Overall
N	2,211
Patients younger than 50 (%)	1,142 (51.7)
History of PONV (%)	282 (12.8)
No history of smoking (%)	1,885 (85.3)
Female gender (%)	1,310 (59.2)
Risk levels (%)	
0	57 (2.6)
1	565 (25.6)
2	827 (37.4)
3	648 (29.3)
4	114 (5.2)
Sevoflurane not used (%)	272 (12.3)
N ₂ O not used (%)	285 (12.9)
Dexamethasone (%)	1,602 (72.5)
Ondansetron (%)	1,880 (85.0)
Scopolamine (%)	57 (2.6)
Regional anesthesia used (%)	68 (3.1)
Aprepitant (%)	109 (4.9)
Propofol (%)	506 (22.9)
Bupivacaine (%)	13 (0.6)
Ropivacaine (%)	2 (0.1)
Acetaminophen (%)	1,649 (74.6)
Ketorolac (%)	477 (21.6)
Interventions used (%)	
0	18 (0.8)
1	102 (4.6)
2	438 (19.8)
3	909 (41.1)
4	541 (24.5)
5	170 (7.7)
6	30 (1.4)
7	3 (0.1)

Abbreviation: PONV, postoperative nausea and vomiting. Note: This table summarizes risk factors and corresponding antiemetic interventions chosen by the clinicians. The table details the elicited patient risk factors, the antiemetic interventions and anesthesia choices (including the use of volatile anesthesia gases like sevoflurane (the commonly used volatile anesthetic at our institution) and nitrous oxide (N₂O), including interventions to mitigate the risk of PONV with medications (Aprepitant, Ondansetron, Dexamethasone), or adjuvant pain modalities (acetaminophen, ketorolac, or local anesthetics) to reduce the amount of opioids administered, to reduce the risk of PONV.

concerns about attrition and selection bias and the limited number of data points for individual clinicians. This may lead to instability in analytic modeling, in particular for the investigation of process disparities driven by SoDHs, leading us to abstain from reporting the results of a formal quantitative statistical analysis.

Discussion

Leveraging data derived directly from EMRs, we constructed a novel intuitive, visual, perioperative MIP (→Figs. 3 and 4), grounded in a realistic hierarchical model of contemporary health care (→Supplementary Appendix A, available in the online version).⁴²

We choose to instantiate our model with perioperative antiemetic prophylaxis^{12,25,43} as a prime example where individual clinicians can enhance value-based care.⁴⁴ PONV is a still highly common occurrence,^{22,45} and cited by patients as among the most undesirable outcomes after surgery.⁴⁶ The costs to the patient and to the health system from PONV, such as a longer length of stay, have been consistently demonstrated.⁴⁷ We hypothesize that our model could be leveraged to improve value and reduce cost by providing clinicians, teams, and institutions individual feedback about their risk-adjusted adherence to PONV guidelines, and by encouraging discussion of shortcomings in PONV prophylaxis.

MIP summarizes individual performance as a regression line of risk against interventions for a set of patients under the care of the clinician. The example of antiemetic prophylaxis revealed that clinicians in our sample are underutilizing and/or overutilizing the PONV interventions in the patients they treat. As an exploratory analysis, we examined the utility of the metric to investigate how SoDHs might impact clinician performance and disparity in health care processes (→Fig. 4).^{48,49} Our MIP is novel in that it (1) considers a bundle of interventions under the control of the individual clinician, (2) adjusts for patient risk and allows for clinician discretion, (3) is completely automated, extracted from the EMR, and (4) measures the performance of individual clinicians during routine clinical care.

We are unaware of any perioperative MIP with these features to date.^{17,42,50,51} The MPOG has developed specific process measures to quantify perioperative performance. Specifically, MPOG can report on individual anesthesia clinician utilizing PONV prophylaxis.⁵² They fall short of estimating the (1) overutilization of PONV medications, (2) the gradient of risk responsiveness for a clinician, or (3) many additional possible choices open to the anesthesiologist to mitigate PONV in high-risk individuals.

Overutilization of antiemetic prophylaxis is wasteful and costly, given that the cost of strong newer antiemetics is around US\$100 per treatment.⁵³ Overutilization may have adverse effects as their indiscriminate use carries risks, including cardiac arrhythmias, impaired glycemic control, tumor lysis, dysphoria, other psychiatric effects, unplanned pregnancies, etc.^{5,54,55}

The need for specific, intuitive, visual, and individual MIP has become apparent as the health care systems shift from volume-based reimbursement to value-based care and P4P.^{4,17,18,44} Objective and fair assessment of clinical performance in the perioperative setting is critical for focused education, meaningful feedback and incentives, and for HSS. Both under- and overutilization can be evaluated with our metric, for example to (1) investigate perioperative clinician dyads (attending/resident, attending/nurse anesthetist),⁵⁶ (2) explore the impact of

CDS on risk-adjusted utilization of resources,^{57–59} or (3) evaluate clinician, team, and/or institutional improvements in enhancing value or adherence to guidelines,²² and (4) report quality measures to P4P programs that are relevant to the scope of practice of anesthesiologists and linked to outcomes.¹⁷

Strengths and Limitations

Our model is able to differentiate risk-adjusted over- and underutilization of PONV prophylaxis comparing clinicians, teams, and institutions. Some clinicians in our cohort administer the same low number of antiemetic interventions to all patients, disregarding individual patient risk. Such therapeutic nihilism is misplaced and negates the impact of diligent clinicians on perioperative outcomes. We posit that indiscriminate use (or avoidance) of antiemetic interventions indicates subpar perioperative care.

We concede that attributing the case to the first attending clinician, if several clinicians are involved, is a limitation, as antiemetics are also administered at the end. However, the initial clinician is responsible to ensure adequate care transition at hand-over.

Our exploratory analysis of the utility of our metric to investigate perioperative process disparities down to the granular level of the individual clinician level is innovative. Previous studies demonstrated disparities between populations, (e.g., in the risk-adjusted utilization of PONV prophylaxis),⁵ but failed to measure and attribute disparities at the individual clinician level. Still, did we find a disconcerting case of health care disparity attributable to an individual clinician? We caution against drawing premature conclusions from our observations, which we discuss and illustrate in →Fig. 2.

Process measures like our proposed metric focus on “what the doctor does” rather than on “how the patient does,” an important criticism in comparison to outcome measures.¹⁷ However, P4P incentives may inadvertently worsen health care disparity for example through cherry-picking, i.e., selecting patients with SoDHs (→Fig. 1) that predict favorable outcomes.^{5,49,60} Risk-adjusted process measures, like our MIP, may be used to prevent or monitor unintended exacerbations of health system disparity through P4P programs.⁴⁸

The clinical implementation of our metric raised further questions. Extracting clearly defined operational measures on a larger scale from our EMR proved challenging. We noted that key risk factors were inconsistently elicited and documented by clinicians, e.g., history of motion sickness. Lack of salient data precludes adherence with risk-adjusted utilization and may induce clinician bias (→Supplementary Fig. S2, available in the online version). PONV prophylaxis administration is documented in many locations by various professionals, making it difficult to effectively utilize or extract the information.

Test R-test reliability would translate to seeing a consistent result when measuring the performance of an individual (or a group of clinicians) over repeated periods, i.e., obtaining a similar estimate of responsiveness to risk for the months of November and December. For example, our metric would demonstrate that a given provider’s regression line is similar over time (in slope and intercept). In this study, we simply

had too few data points for individual providers to demonstrate this but shall investigate reliability as we collect more data over a longer time period.

Given the lack of any accepted gold standard performance metric for anesthesiologists,^{61–64} establishing the construct validity of our metric to assess individual clinician performance will be challenging (as we discussed this also elsewhere).⁵ Reviewers and colleagues questioned if adherence to PONV guidelines (alone) reflects on underlying anesthesia quality.⁵ Face validity of the formative construct is more credible, as both risk factors and appropriate antiemetic prophylaxis are uncontroversial.⁶⁵ Both reliability and construct validity remain to be evaluated in future research.

Attribution to Individual Clinicians

Our performance metric will only be valid if it focuses on interventions under individual clinician control. The interventions considered are under the influence of clinicians, but also contingent on team performance.⁵ System factors may impact the individual clinician utilization of interventions. A case in point is the process disparity observed in **Fig. 4(C)**. While the apparent underutilization of risk-adjusted PONV prophylaxis in minority patients should raise concerns, the data are no *proof* that the clinician in panel C is driven by implicit or explicit bias, or guilty of providing inferior care to minority patients (deliberately or driven through implicit or explicit bias).

Instead of targeting individuals for perceived negligence, systems-based interventions should seek to re-engineer care processes to prevent patient harm in the face of misaligned resources and fallible human clinicians. We propose to develop MIP into an additional quality metric, which, with perioperative qualified data registries like ASPIRE/MPOG, could feed back to individual clinicians to inform them on their equitable care.⁶² We discuss key drivers of appropriate antiemetic prophylaxis in **Fig. 2**. One of the many possible explanations for the systematic underutilization of risk-adjusted prophylaxis is the unavailability of medication in care settings where minority patients are treated, which might lead to underutilization. The proposed precise and granular process performance metric, fine-tuned with SoDHs, (**Fig. 1**) may help us define key drivers and confront disparities effectively.⁶⁶

Conclusion

Our MIP may serve as an objective, precise, intuitive, visual, and personalized assessment of risk-adjusted adherence to care bundles, allowing clinicians to compare themselves with their peers, in or outside their own institution, or to explore differences in institutional performances down to the individual clinician level (**Fig. 3**). Our model is generalizable, yet exemplified in a domain under the sole control of individual anesthesia clinicians, risk-adjusted antiemetic prophylaxis. Utilizing EMR data at PSHMC, our model was sensitive to treatment disparities at the clinician level between patients (**Fig. 4**), e.g., contingent on SoDHs (**Fig. 1**) and to both underutilization and overutilization of interventions. Our metric is risk-adjusted, yet leaves room

for professional discretion. Our hierarchical model is scalable to national registries, but granular to the individual clinician. Importantly, our metric can give clinicians personalized intuitive visual feedback to boost adherence. We admit that clinical implementation implies complex changes to EMRs and workflow. The metric requires great caution in implementation to avoid capricious inferences. We still need to investigate if the model is robust, accurate, unbiased, and stable over time. However, our model is an important step toward measuring individual clinician performance for HSS.

Clinical Relevance Statement

EMR-derived, automated, risk-adjusted perioperative *individual clinician* performance assessment remains elusive, but would be valuable for HSS, quality improvement, and health care disparity research. We tested a multidimensional risk-adjusted metric to assess individual clinician adherence to guidelines on comprehensive prophylactic bundles to reduce the risk of PONV. Our metric is an example of a generalizable approach allowing investigation of individual and team performance based on data automatically documented during the routine delivery of clinical care.

Multiple Choice Questions

1. A reflective construct of clinician performance would be construed to measure:
 - a. A phenomenon of interest
 - b. A latent quality
 - c. The effect of interventions
 - d. The clinician perspective of events

Correct Answer: The correct answer is option b.
2. Ideal clinician performance measures should be:
 - a. Risk-adjusted
 - b. Risk-indifferent
 - c. Indiscriminate
 - d. Refutable

Correct Answer: The correct answer is option a.
3. Multidimensional performance metrics typically:
 - a. Suffer from floor and ceiling effects
 - b. Are risk-indifferent
 - c. Are more responsive
 - d. Are irrefutable

Correct Answer: The correct answer is option c.

Protection of Human and Animal Subjects

The Penn State Health Institutional Review Board determined that the described work met the criteria for exempt research according to the policies of this institution and the provisions of applicable federal regulations (STUDY 00007035).

Funding

Research reported in this publication was supported by the National Center for Advancing Translational Sciences

(NIH grants TL1 TR002016 and UL1 TR002014). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The research reported was furthermore supported through the Donald E. Martin Professorship for Anesthesia and Pain Medicine, a career development endowment of the Department of Anesthesiology at Penn State Milton S. Hershey Medical Center.

Conflict of Interest

None declared.

Acknowledgments

We would like to thank our statistical consultants, GG Statistics LLC, for the development of the statistical modeling proposal through the concept of majorization. We would like to acknowledge Shalini Maruthoti for her advice on data sources and for extracting data from our electronic health records.

References

- Link BG, Phelan J. Understanding racial and ethnic disparities in postsurgical complications occurring in U.S. hospitals. *J Health Soc Behav* 1995(Spec No):80–94
- Witt WP, Coffey RM, Lopez-Gonzalez L, et al. Understanding racial and ethnic disparities in postsurgical complications occurring in U.S. hospitals. *Health Serv Res* 2017;52(01):220–243
- Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health* 1997;18:341–378
- Andreae MH, Andreae DA, Maman S. The ethics of universal health insurance. *Anesth Analg* 2020;130(03):e85–e87
- Andreae MH, Gabry JS, Goodrich B, White RS, Hall C. Antiemetic prophylaxis as a marker of health care disparities in the national anesthesia clinical outcomes registry. *Anesth Analg* 2018;126(02):588–599
- Thorlby R, Jorgensen S, Siegel B, Ayanian JZ. How health care organizations are using data on patients' race and ethnicity to improve quality of care. *Milbank Q* 2011;89(02):226–255
- Andreae MH, White RS, Chen KY, Nair S, Hall C, Shaparin N. The effect of initiatives to overcome language barriers and improve attendance: a cross-sectional analysis of adherence in an inner city chronic pain clinic. *Pain Med* 2017;18(02):265–274
- Sequist TD, Fitzmaurice GM, Marshall R, Shaykevich S, Safran DG, Ayanian JZ. Physician performance and racial disparities in diabetes mellitus care. *Arch Intern Med* 2008;168(11):1145–1151
- Silber JH, Kennedy SK, Even-Shoshan O, et al. Anesthesiologist direction and patient outcomes. *Anesthesiology* 2000;93(01):152–163
- Gaba DM. Anaesthesiology as a model for patient safety in health care. *BMJ* 2000;320(7237):785–788
- Atchabahian A, Andreae M. Long-term functional outcomes after regional anesthesia: a summary of the published evidence and a recent cochrane review. *Refresh Courses Anesthesiol* 2015;43(01):15–26
- Maman S, Andreae M. 3025 individual anesthesia provider performance assessment. *J Clin Transl Sci* 2019;3(S1):147. Doi: 10.1017/cts.2019.334
- Skochelek SE, Hammoud MM, Lomis KD, Borkan JM, Gonzalo JD, Lawson LE, Starr SR, eds. *Health Systems Science*, 2nd ed. Philadelphia, PA: Elsevier; 2020
- Kawachi I, Daniels N, Robinson DE. Health disparities by race and class: why both matter. *Health Aff (Millwood)* 2005;24(02):343–352
- Epstein RH, Dexter F, Patel N. Influencing anesthesia provider behavior using anesthesia information management system data for near real-time alerts and post hoc reports. *Anesth Analg* 2015;121(03):678–692
- Chau A, Ehrenfeld JM. Using real-time clinical decision support to improve performance on perioperative quality and process measures. *Anesthesiol Clin* 2011;29(01):57–69
- Hyder JA, Hebl JR. Performance measurement to demonstrate value. *Anesthesiol Clin* 2015;33(04):679–696
- Hyder JA, Niconchuk J, Glance LG, et al. What can the national quality forum tell us about performance measurement in anesthesiology? *Anesth Analg* 2015;120(02):440–448
- Palmer RC, Ismond D, Rodriguez EJ, Kaufman JS. Social determinants of health: Future directions for health disparities research. *Am J Public Health* 2019;109(S1):S70–S71
- Bazemore AW, Cottrell EK, Gold R, et al. "Community vital signs": incorporating geocoded social determinants into electronic records to promote patient and population health. *J Am Med Inform Assoc* 2016;23(02):407–412
- Kheterpal S. Clinical research using an information system: the multicenter perioperative outcomes group. *Anesthesiol Clin* 2011;29(03):377–388
- Gan TJ, Diemunsch P, Habib AS, et al; Society for Ambulatory Anesthesia. Consensus guidelines for the management of postoperative nausea and vomiting. *Anesth Analg* 2014;118(01):85–113
- Junger A, Hartmann B, Benson M, et al. The use of an anesthesia information management system for prediction of antiemetic rescue treatment at the postanesthesia care unit. *Anesth Analg* 2001;92(05):1203–1209
- Shwartz M, Restuccia JD, Rosen AK. Composite measures of health care provider performance: a description of approaches. *Milbank Q* 2015;93(04):788–825
- Maman S, Andreae M. An EMR-derived individual anesthesia provider metric to measure risk-adjusted adherence with perioperative prophylactic bundles for health systems science and disparity research. Paper presented at: Annual Meeting of the Association of University Anesthesiologists; May 16–20, 2019; Montreal, Canada
- Kappen TH, Vergouwe Y, van Wolfswinkel L, Kalkman CJ, Moons KGM, van Klei WA. Impact of adding therapeutic recommendations to risk assessments from a prediction model for postoperative nausea and vomiting. *Br J Anaesth* 2015;114(02):252–260
- Biedler A, Wermelt J, Kunitz O, et al. A risk adapted approach reduces the overall institutional incidence of postoperative nausea and vomiting. *Can J Anaesth* 2004;51(01):13–19
- Kumar A, Brampton W, Watson S, Reid VL, Neilly D. Postoperative nausea and vomiting: simple risk scoring does work. *Eur J Anaesthesiol* 2012;29(01):57–59
- Rüsch D, Eberhart L, Biedler A, Dethling J, Apfel CC. Prospective application of a simplified risk score to prevent postoperative nausea and vomiting. *Can J Anaesth* 2005;52(05):478–484
- Frenzel JC, Kee SS, Ensor JE, Riedel BJ, Ruiz JR. Ongoing provision of individual clinician performance data improves practice behavior. *Anesth Analg* 2010;111(02):515–519
- Kooij FO, Klok T, Hollmann MW, Kal JE. Automated reminders increase adherence to guidelines for administration of prophylaxis for postoperative nausea and vomiting. *Eur J Anaesthesiol* 2010;27(02):187–191
- Bowen DJ, Kreuter M, Spring B, et al. How we design feasibility studies. *Am J Prev Med* 2009;36(05):452–457
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;335(7624):806–808
- Macario A, Chung A, Weinger MB. Variation in practice patterns of anesthesiologists in California for prophylaxis of postoperative nausea and vomiting. *J Clin Anesth* 2001;13(05):353–360

- 35 Duran D, Asada Y, Millum J, Gezmu M. Harmonizing health disparities measurement. *Am J Public Health* 2019;109(S1):S25–S27
- 36 Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. Vienna, Austria: R Foundation for Statistical Computing; New York, NY: Springer; 2016
- 37 Wickham H, François R, Henry L, Müller K. *Dplyr: A Grammar of Data Manipulation*. Vienna, Austria: R Foundation for Statistical Computing; 2019
- 38 Wickham H, Bryan J. *Readxl: Read Excel Files*. Vienna, Austria: R Foundation for Statistical Computing; New York, NY: Springer; 2019
- 39 Xie Y. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. Vienna, Austria: R Foundation for Statistical Computing; New York, NY: Springer; 2019
- 40 Xu HF, White RS, Sastow DL, Andreae MH, Gaber-Baylis LK, Turnbull ZA. Medicaid insurance as primary payer predicts increased mortality after total hip replacement in the state inpatient databases of California, Florida and New York. *J Clin Anesth* 2017;43:24–32
- 41 Cooper LA, Hill MN, Powe NR. Designing and evaluating interventions to eliminate racial and ethnic disparities in health care. *J Gen Intern Med* 2002;17(06):477–486
- 42 Kaye AD, Okanlawon OJ, Urman RD. Clinical performance feedback and quality improvement opportunities for perioperative physicians. *Adv Med Educ Pract* 2014;5:115–123
- 43 Maman S, Andreae M. Quantifying the impact of physician supervision on nurse anesthetists' risk-adjusted antiemetic prophylaxis. Paper presented at: Annual Meeting of the International Anesthesia Research Society; May 16–20, 2019; Montreal, Canada
- 44 Kolarczyk LM, Arora H, Manning MW, Zvara DA, Isaak RS. Defining value-based care in cardiac and vascular anesthesiology: the past, present, and future of perioperative cardiovascular care. *J Cardiothorac Vasc Anesth* 2018;32(01):512–521
- 45 Apfel CC, Bacher A, Biedler A, et al. A factorial trial of six interventions for the prevention of postoperative nausea and vomiting [in German]. *Anaesthesist* 2005;54(03):201–209
- 46 Macario A, Weinger M, Carney S, Kim A. Which clinical anesthesia outcomes are important to avoid? The perspective of patients. *Anesth Analg* 1999;89(03):652–658
- 47 Chung F, Mezei G. Factors contributing to a prolonged stay after ambulatory surgery. *Anesth Analg* 1999;89(06):1352–1359
- 48 Eggleton K, Liaw W, Bazemore A. Impact of gaps in merit-based incentive payment system measures on marginalized populations. *Ann Fam Med* 2017;15(03):255–257
- 49 Anderson AC, O'Rourke E, Chin MH, Ponce NA, Bernheim SM, Burstin H. Promoting health equity and eliminating disparities through performance measurement and payment. *Health Aff (Millwood)* 2018;37(03):371–377
- 50 Glance LG, Neuman M, Martinez EA, Pauker KY, Dutton RP. Performance measurement at a "tipping point". *Anesth Analg* 2011;112(04):958–966
- 51 Peccora CD, Gimlich R, Cornell RP, Vacanti CA, Ehrenfeld JM, Urman RD. Anesthesia report card - a customizable tool for performance improvement. *J Med Syst* 2014;38(09):105
- 52 Multicenter Perioperative Outcomes Group. (2020, July 14). Post-operative Nausea and Vomiting PONV-01. ASPIRE Measures. Available at: <https://spec.mpog.org/Spec/Public/24>
- 53 Jeyabalan S, Thampi SM, Karuppusami R, Samuel K. Comparing the efficacy of aprepitant and ondansetron for the prevention of postoperative nausea and vomiting (PONV): a double blinded, randomised control trial in patients undergoing breast and thyroid surgeries. *Indian J Anaesth* 2019;63(04):289–294
- 54 Bartlett R, Hartle AJ. Routine use of dexamethasone for postoperative nausea and vomiting: the case against. *Anaesthesia* 2013;68(09):892–896
- 55 McDonnell C, Barlow R, Campisi P, Grant R, Malkin D. Fatal perioperative acute tumour lysis syndrome precipitated by dexamethasone. *Anaesthesia* 2008;63(06):652–655
- 56 Frasier LL, Pavuluri Quamme SR, Ma Y, et al. Familiarity and communication in the operating room. *J Surg Res* 2019;235:395–403
- 57 Sintchenko V, Magrabi F, Tipper S. Are we measuring the right end-points? Variables that affect the impact of computerised decision support on patient outcomes: a systematic review. *Med Inform Internet Med* 2007;32(03):225–240
- 58 Kooij FO, Klok T, Hollmann MW, Kal JE. Decision support increases guideline adherence for prescribing postoperative nausea and vomiting prophylaxis. *Anesth Analg* 2008;106(03):893–898
- 59 Simpaio AF, Tan JM, Lingappan AM, Gálvez JA, Morgan SE, Krall MA. A systematic review of near real-time and point-of-care clinical decision support in anesthesia information management systems. *J Clin Monit Comput* 2017;31(05):885–894
- 60 Chien AT, Chin MH, Davis AM, Casalino LP. Pay for performance, public reporting, and racial disparities in health care: how are programs being designed? *Med Care Res Rev* 2007;64(5, Suppl):283S–304S
- 61 Warner DO, Isaak RS, Peterson-Layne C, et al. Development of an objective structured clinical examination as a component of assessment for initial board certification in anesthesiology. *Anesth Analg* 2020;130(01):258–264
- 62 Colquhoun DA, Shanks AM, Kapeles SR, et al. Considerations for integration of perioperative electronic health records across institutions for research and quality improvement: the approach taken by the multicenter perioperative outcomes group. *Anesth Analg* 2020;130(05):1133–1146
- 63 McCormick PJ, Yeoh C, Vicario-Feliciano RM, et al. Improved compliance with anesthesia quality measures after implementation of automated monthly feedback. *J Oncol Pract* 2019;15(06):e583–e592
- 64 Tetzlaff JE. Assessment of competency in anesthesiology. *Anesthesiology* 2007;106(04):812–825
- 65 Kooij FO, Klok T, Preckel B, Hollmann MW, Kal JE. The effect of requesting a reason for non-adherence to a guideline in a long running automated reminder system for PONV prophylaxis. *Appl Clin Inform* 2017;8(01):313–321
- 66 Hoberman J. Medical racism and the rhetoric of exculpation: how do physicians think about race? *New Lit Hist* 2007;38(03):505–525
- 67 Andreae MH, Nair S, Gabry JS, Goodrich B, Hall C, Shaparin N. A pragmatic trial to improve adherence with scheduled appointments in an inner-city pain clinic by human phone calls in the patient's preferred language. *J Clin Anesth* 2017;42:77–83
- 68 Shaparin N, White R, Andreae M, Hall C, Kaufman A. A longitudinal linear model of patient characteristics to predict failure to attend an inner-city chronic pain clinic. *J Pain* 2014;15(07):704–711
- 69 Payne VL, Hysong SJ. Model depicting aspects of audit and feedback that impact physicians' acceptance of clinical performance feedback. *BMC Health Serv Res* 2016;16:260
- 70 Carpenter B, Gelman A, Hoffman M, et al. Stan: a probabilistic programming language. *J Stat Softw* 2017;76(01):1–32
- 71 Bürkner P-C. Advanced Bayesian multilevel modeling with the R package brms. *R J* 2018;10(01):395–411