



HHS Public Access

Author manuscript

J Clin Anesth. Author manuscript; available in PMC 2021 October 01.

Published in final edited form as:

J Clin Anesth. 2020 October ; 65: 109814. doi:10.1016/j.jclinane.2020.109814.

Patient Satisfaction Survey Scores are not an Appropriate Metric to Differentiate Performance Among Anesthesiologists

Robert E. Freundlich, M.D., M.S., M.S.C.I.,

Department of Anesthesiology, Department of Biomedical Informatics, Vanderbilt University Medical Center

Gen Li, M.Stat., M.Chem.,

Department of Anesthesiology, Vanderbilt University Medical Center

Brendan Grant, B.S. [Student],

Tulane University

Paul St. Jacques, M.D.,

Department of Anesthesiology, Department of Biomedical Informatics, Vanderbilt University Medical Center

Warren S. Sandberg, M.D., Ph.D.,

Department of Anesthesiology, Department of Biomedical Informatics, Department of Surgery, Vanderbilt University Medical Center

Jesse M. Ehrenfeld, M.D., M.P.H.,

Department of Anesthesiology, Medical College of Wisconsin

Matthew S. Shotwell, Ph.D.,

Department of Biostatistics, Department of Anesthesiology, Vanderbilt University Medical Center

Corresponding Author: Gen Li, M.Stat., M.Chem., Department of Anesthesiology, Vanderbilt University Medical Center, 1211 21st Avenue South Medical Arts Building #708, Nashville, TN 37212, Phone: (615) 875-2971, gen.li@vumc.org.

Author Contributions:

Robert E. Freundlich, M.D., M.S., M.S.C.I. contributed to the conception and design of this study and to the acquisition, and interpretation of data, drafted and revised the manuscript for important intellectual contents.

Gen Li, M.Stat., M.Chem. contributed to the study design and to the acquisition, analysis and interpretation of data for this study, drafted and revised the manuscript for important intellectual contents.

Brendan Grant contributed to the acquisition, analysis and interpretation of data for this study.

Paul St. Jacques, M.D. contributed to the acquisition of data, revised the manuscript for important intellectual contents.

Warren S. Sandberg M.D., Ph.D. contributed to the conception and design of this study, revised the manuscript for important intellectual contents.

Jesse M. Ehrenfeld M.D., M.P.H. contributed to the conception and design of this study, revised the manuscript for important intellectual contents.

Matthew S. Shotwell, Ph.D. contributed to the analysis, and interpretation of data for this study, revised the manuscript for important intellectual contents.

Jonathan P. Wanderer, M.D., M.Phil., FASA contributed to the conception and design of this study and to the acquisition of data for this study, revised the manuscript for important intellectual contents.

All authors have reviewed and approved the final version of the manuscript.

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

Prior Presentations: IARS 2018 annual meeting, Chicago, IL

Conflict of Interest: None

Jonathan P. Wanderer, M.D., M.Phil., FASA

Department of Anesthesiology, Department of Biomedical Informatics, Vanderbilt University Medical Center

Abstract

Study Objective: With the focus of patient-centered care in healthcare organizations, patient satisfaction plays an increasingly important role in healthcare quality measurement. We sought to determine whether an automated patient satisfaction survey could be effectively used to identify outlying anesthesiologists.

Design: Retrospective Observational Study

Setting: Vanderbilt University Medical Center (VUMC)

Measurements: Patient satisfaction data were obtained between October 24, 2016 and November 1, 2017. A multivariable ordered probit regression was conducted to evaluate the relationship between the mean scores of responses to Likert-scale questions on SurveyVitals' Anesthesia Patient Satisfaction Questionnaire 2. Fixed effects included demographics, clinical variables, providers and surgeons. Hypothesis tests to compare each individual anesthesiologist with the median-performing anesthesiologist were conducted.

Main Results: We analyzed 10,528 surveys, with a 49.5% overall response rate. Younger patient (odds ratio (OR) 1.011 [per year of age]; 95% confidence interval (CI) 1.008 to 1.014; $p < .001$), regional anesthesia (versus general anesthesia) (OR 1.695; 95% CI 1.186 to 2.422; $p = 0.004$) and daytime surgery (versus nighttime surgery) (OR 1.795; 95% CI 1.091 to 2.959; $p = 0.035$) were associated with higher satisfaction scores. Compared with the median-ranked anesthesiologist, we found the adjusted odds ratio for an increase in satisfaction score ranged from 0.346 (95% CI 0.158 to 0.762) to 1.649 (95% CI 0.687 to 3.956) for the lowest and highest scoring providers, respectively. Only 10.10% of anesthesiologists at our institution had an odds ratio for satisfaction with a 95% CI not inclusive of 1.

Conclusions: Patient satisfaction is impacted by multiple factors. There was very little information in patient satisfaction scores to discriminate the providers, after adjusting for confounding. While patient satisfaction scores may facilitate identification of extreme outliers among anesthesiologists, there is no evidence that this metric is useful for the routine evaluation of individual provider performance.

Keywords

Anesthesiologist; Patient Satisfaction; Survey Score; Performance Improvement

1. Introduction

The 2001 Institute of Medicine report, "Crossing the Quality Chasm", advocated for broad, sweeping changes in how healthcare is delivered in the United States, with a focus on implementing improved assessment of quality [1]. One of the six specific aims for improvement identified was patient-centered care -- "providing care that is respectful of and responsive to individual patient preferences, needs, and values, and ensuring that patient

values guide all clinical decisions.” However, evaluating the quality of anesthesia is particularly challenging. Proposed quality metrics such as avoidance of intraoperative hypotension [2–3], post-operative pain scores [4], and timely administration of perioperative antibiotics [5] have all been shown to have significant limitations and perform poorly as quality indicators.

While anesthesiology has prided itself as a leader and innovator in the field patient safety, it has been criticized for being inadequately patient-centered by multiple leaders in the field [6,7]. In response to this concern, there has been a push for increased transparency on the part of providers and care groups about whether they are providing care that is perceived by patients to be of value. This perception of value, however, may correlate poorly with actual quality of care delivered [8].

One approach to improve patient-centered value has been through the measurement of patient satisfaction ratings. Satisfaction metrics have been proposed to evaluate individual providers. Although these ratings are widely-used and available, they have not been rigorously validated, particularly in the field of anesthesiology. Furthermore, it is unclear if these ratings can be used to compare providers or, perhaps more relevant to the patient centered quality imperative, to effectively identify underperforming providers within a cohort. We therefore sought to determine whether patient satisfaction data can be used to assess individual anesthesiologist performance at a large, academic medical center, after adjusting for confounding factors. We hypothesized that patient satisfaction scores are insufficient to rank-order individual provider performance.

2. Methods

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines were used in the preparation of this manuscript [9].

2.1. Human Subjects Protection

This study received approval from the Vanderbilt University Medical Center (VUMC) Institutional Review Board (#180688) with a waiver of informed consent.

2.2. Data Collection

Patient, provider, and procedural data were obtained from the VUMC Perioperative Data Warehouse, which are derived from data from the Anesthesia Information Management System (AIMS; VPIMS, internally-developed at VUMC, Nashville, TN). These data were merged with patient satisfaction data obtained by a third-party vendor (SurveyVitals, Dallas, TX). SurveyVitals is now used by 1 in 4 U.S. physician anesthesiologists across 2,691 facilities (<https://www.SurveyVitals.com/start/anesthesia>). The tool uses an automated approach to contacting patients after their procedure and administering the Anesthesia Patient Satisfaction Questionnaire 2.0 (APSQ2). The internal consistency of the overall methodology of APSQ2 has been calculated internally by SurveyVitals (Cronbach’s alpha 0.820, Appendix A.1) [10].

The APSQ2 is specifically designed to evaluate anesthesia care teams where an attending anesthesiologist is working with a Certified Registered Nurse Anesthetist (CRNA), and these two roles are individually assessed in the design of the questionnaire. Given this structure, we only included the surveys sent to patients whose anesthesia care was delivered by an anesthesia care team composed of one attending anesthesiologist and one CRNA in our primary analysis. A sensitivity analysis was performed as a secondary analysis to evaluate the effect of other care team models on satisfaction scores.

At our institution, contact is made within seven days after discharge or the conclusion of the procedure via email, telephone text message or an interactive voice response (IVR) phone call system. A patient can be contacted via several modalities as the contact campaign follows an established order: (1. SMS/Text message, two attempts; 2. Email, two attempts; 3. IVR call, three attempts). However, once the patient takes the survey, the campaign is stopped, and no more contact attempts are made. Patients who received anesthesiology services on an inpatient or outpatient basis, and for whom a valid email or telephone number is present in the hospital admission records are eligible to receive an APSQ2 survey, except for those with in-hospital mortality, or who had been surveyed within the past 60 days, or who had multiple anesthetics within their hospital stay, or whose anesthesia care team had multiple staff members per role (i.e. more than one attending).

As part of the APSQ2 (Q1, Table A), patients are shown a headshot picture of each member of their anesthesia care team, along with their name and role. These headshots were taken with each member of our department in clinical attire (i.e., hat, mask and scrubs), to improve recognition by patients. The APSQ2 only proceeds if the patient indicates that they remember each provider enough to answer questions about the care provided. Additionally, the patient is asked if they would like to be contacted to discuss their evaluation.

The APSQ2 consists of Likert scale questions, yes-no questions, and open-ended questions. Using an established methodology [10,11], responses from the seven Likert scale questions (Q3–Q8 and Q10) from the APSQ2 that focused on the patient satisfaction with their attending anesthesiologist were analyzed as the primary outcome. The remainder of the questions on the SurveyVitals questionnaire (Q12–15) pertain to other members of the anesthesia care team and were excluded. Three open-ended, unstructured questions (Q9, Q11 and Q18), and yes or no question focused on self-reported anesthesia outcomes, including pain, nausea/vomiting, and unpleasant memories, were similarly excluded.

All seven questions selected for analysis were posed as a typical Likert scale, with a 5-point ordinal scale used by patients to express their agreement or disagreement with a statement. “Strongly Agree” was assigned a value of 5 and “Strongly Disagree” a value of 1. The arithmetic mean scores of the responses to seven Likert-scale questions were also calculated as the overall satisfaction score in this study.

All anesthetics performed between October 24th, 2016 to November 1st, 2017 at VUMC, a large, tertiary academic medical center, were eligible for inclusion in the analysis. These dates were chosen based on the implementation of a new survey instrument (10/24/2016) and our go-live with a new electronic health record, which temporarily interrupted our

SurveyVitals data collection (11/1/2017). Based on a recent reliability study of APSQ2 conducted by SurveyVitals, we excluded attending anesthesiologists with fewer than 96 patient satisfaction ratings from analysis, which is estimated by SurveyVitals to provide a 95% confidence level for the accuracy of results (Appendix A.2).

2.3. Statistical Analysis

Descriptive statistics were calculated across patient encounters using the median and interquartile range for continuous variables and with percentages for categorical variables and stratified by overall satisfaction score. The raw response rates of the seven Likert scale questions centered on patient satisfaction were reported.

Two strategies were implemented to construct the primary outcome in this study. First, a specific multivariable ordered probit regression model was conducted using the natural overall satisfaction score on with ordered values between 1 and 5. In the second strategy, as a sensitivity analysis, the overall satisfaction score was dichotomized to a binary outcome, with overall satisfaction score of 4–5 were interpreted as “Satisfied” and 1–3 were interpreted as “Not Satisfied” for regular multivariable logistic regression [11].

2.3.1. Ordered Probit Regression—Data were analyzed using a multivariable ordered probit fixed-effects regression model with the ordered overall satisfaction scores of attending anesthesiologists as the outcome variable. CRNA performance was not assessed in this regression model given insufficient power. Covariates included attending anesthesiologists, CRNAs, patient age, patient gender, American Society of Anesthesiologists Physical Status Classification (ASA), anesthesia type, surgical service, surgery start time (as a categorical variable) [11], operating room location within the hospital, and surgeon [12,13]. After redundancy analysis, operating room location was removed from the model, given significant collinearity ($R^2 = 1.00$). The overall significance of the adjusted association between each covariate and the outcome was assessed using a Wald multiple degree of freedom Chi-squared test.

The odds ratios (ORs) with 95% CIs were reported for all significant factors to demonstrate the relative odds of the occurrence of the outcome of interest, given exposure to the other covariates for the ordered probit regression model. To be specific, for instance, let the ordinal patient satisfaction outcome be denoted by Y and one of its levels by y (e.g., 1, 2, 3, 4, or 5). Consider the probability that $Y = y$ for a patient seen by anesthesiologist A. The odds that $Y = y$ is the probability divided by one minus the probability. The odds ratio for provider A versus the median scoring provider is the ratio of the corresponding odds (provider A odds in the numerator). The ordered probit model assumptions require that the odds ratio is the same no matter which cutoff y is chosen. Thus, the OR should be interpreted as the fold-change in the odds of a higher patient satisfaction associated with a change in the corresponding covariate (e.g., provider, patient age), after controlling all the other covariates. For instance, a patient whose surgery was supervised by an anesthesiologist with corresponding odds ratio 1.25 would have 25% greater odds of higher satisfaction compared with the median scoring anesthesiologist. We also conducted a one-sided 0.05-

level Wald test for multiple comparisons for all pairs of anesthesiologists. No familywise hypothesis is considered, thus, no adjustment for multiple comparisons was made.

2.3.2. Logistic Regression—As a secondary analysis, a multivariable logistic regression was performed for the binary outcome to evaluate the covariates which would independently impact the patient satisfaction with their provider to receive a “Satisfied” response. The odds ratios (ORs) with 95% CIs were reported for all significant factors to demonstrate the relative odds of the occurrence of the outcome of interest, given exposure to the other covariates for the logistic regression model.

2.3.3. Sensitivity Analyses—A sensitivity analysis was conducted to discern the effect, if any, of different compositions of anesthesia care team on the responses from patient satisfaction surveys.

Additionally, we performed a multivariable ordered probit regression with two additional covariates. Surgery delay was included to examine how waiting affected patient satisfaction [14]. Moreover, the dosage of midazolam administered in the preoperative holding area was also included in the regression model to detect its association with patient satisfaction scores [15].

All statistical programming was conducted in SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

3. Results

During the one-year study period, 58,468 eligible patients received an APSQ2 survey and 28,832 surveys were returned (49.5%). Specifically, according to the aforementioned sequence, the response rate of SMS/Text message channel was 30.0%. Then among the patients who did not reply to SMS/Text message, 23.3% responded the email. Furthermore, 31.8% patients who did not previously respond to either the text message or email, answered IVR call. Of the returned surveys, the response rates varied from 71.6% to 78.7% for each individual Likert scale question (Table B). Of 17,002 surveys from patients whose anesthesia care was delivered by an anesthesia care team composed of one attending anesthesiologist and one CRNA, a total of 10,528 (61.9%) patient satisfaction surveys met inclusion criteria and were analyzed for the 55 attending anesthesiologists with at least 96 scores in the study period as the primary analysis (Table C). In the sensitivity analysis with all care team compositions, 15,889 surveys were analyzed for 79 attending anesthesiologists.

3.1. Ordered Probit Regression

From the results of ordered probit regression model, we found that a younger patient age increased the likelihood of a provider receiving higher (more favorable) score (OR 1.011 [per year of age]; 95% CI 1.008 to 1.014; $p < .001$). Patients undergoing general anesthetics were more likely to assign their provider a lower score, compared with the those undergoing regional anesthesia (OR 1.695; 95% CI 1.186 to 2.422; $p = 0.004$). Moreover, patients who underwent nighttime surgery, with a case start time between 6 PM and 6 AM, were less likely to give their provider a higher score versus those with a start time between 6 AM and 12 PM (OR 0.557; 95% CI 0.338 to 0.917; $p = 0.035$) (Table D). The anesthesiologist odds

ratios panel (Fig. A) shows the sorted odds ratios for each individual anesthesiologist versus the reference provider, defined as the provider with the median adjusted odds. Compared with the median scoring anesthesiologist, the adjusted odds ratio for a higher satisfaction score ranged from 0.346 (95% CI 0.158 to 0.762) to 1.649 (95% CI 0.687 to 3.956), with all but only one (98.1%) of 95% CIs inclusive of the null value 1. To further assess the sensitivity of the patient satisfaction survey, all pairs of anesthesiologists were compared by conducting Wald's pairwise comparisons. For each provider, Fig. B shows the percentage of other anesthesiologists that received significantly higher or lower scores, and those for which there was insufficient information in the patient satisfaction scores to distinguish provider performance. The odds for the lowest scoring anesthesiologist was significantly lower than 90.7% of the other providers. In contrast, the odds for the highest scoring anesthesiologist was significantly higher than 44.4% of other providers. Nevertheless, only 10.1% of all pairwise comparisons had a 95% confidence interval for the odds ratio for satisfaction not inclusive of 1.

3.2. Logistic Regression

Similar results were observed from logistic regression model, an increasing patient age decreased the likelihood of a provider receiving a "Satisfied" feedback (OR 0.992 [per year of age]; 95% CI 0.986 to 0.998; $p = 0.013$), patients who underwent nighttime surgery between 6 PM and 6 AM were less likely to rank their provider "Satisfied" versus those with a 6 AM to 12 PM surgeries (OR 0.309; 95% CI 0.136 to 0.702; $p = 0.018$) (Table D). The adjusted odds ratio for a provider to receive a "Satisfied" ranged from 0.192 (95% CI 0.031 to 1.191) to 3.820 (95% CI 0.201 to 72.639), comparing with the median performing anesthesiologist (Supplementary A.1). The pairwise comparison panel displays the ranked ordered of comparisons between each individual anesthesiologist, in which only 4.8% of all pairwise comparisons were statistically significant (Supplementary A.2).

3.3. Sensitivity Analysis

Compared with the median performing anesthesiologist, the odds ratio for a provider to receive a higher satisfaction score ranged from 0.475 (95% CI 0.269 to 0.838) to 2.163 (95% CI 1.079 to 4.335), after including all different anesthesia care teams in the cohort. 14.3% of all pairwise comparisons had a 95% confidence interval for the odds ratio for satisfaction not inclusive of 1 (Supplementary B).

The effect estimates and the odds ratio comparisons did not substantially change after adjusting for additional confounding factors "surgery delay" and "midazolam dosage" (Supplementary C).

4. Discussion

There is substantial interest in incorporating patient-centered quality metrics, such as patient satisfaction, into the overall assessment of anesthesiology quality. The updated version of the American Society of Anesthesiologists (ASA) White Paper on patient satisfaction posits that the patient satisfaction measurement carries value even if the process is not statistically valid or clearly linked with patient outcomes [16]. Our goal in implementing a patient

satisfaction questionnaire was two-fold: to gather overall satisfaction data about the department's performance and to gather specific feedback from the open-ended questions. Both objectives facilitate the improvement of value delivered to patients by an anesthesia department. The presence of an ordinal scale in some questions invites the question of whether individual anesthesiologists can be ranked with respect patient satisfaction, though that was not our intent when implementing the system.

We present data demonstrating that patient satisfaction ratings may be used to rank attending anesthesiologists with confounder adjustment, but only to the extent of permitting outlier detection. In our study, elderly patients, general anesthetics and nighttime surgeries have been associated with lower survey scores confirming prior studies. After adjusting for confounders, the majority of providers could not be differentiated. While widely used, these scores seem to be unlikely to be helpful in evaluating individual performance for most providers.

In comparison to other patient satisfaction scores commonly employed in the perioperative environment, such as the Centers for Medicare and Medicaid Services-mandated Hospital Consumer Assessment of Healthcare Providers and Systems and Press-Ganey surveys, SurveyVitals is the most-widely used to assess anesthesia-specific data, with approximately 1 in 4 anesthesiologists currently using the metric to help guide their quality improvement efforts (Appendix A.1). This adoption has occurred in a broad mix of teaching and non-teaching hospitals. Yet despite its widespread adoption, it is not clear how the data obtained from these surveys should be used by institutions.

There are instances of patient satisfaction ratings falling short of expectations in other fields. Online ratings of neurosurgeons, for example, have been found to strongly correlate to where they trained and practice medicine [17]. And while patient satisfaction ratings may correlate with *US News and World Report* ratings in some fields, it is unclear if they correlate well with other metrics of the quality of care delivered, such as readmission rates [18]. Yet despite these issues, patient satisfaction may be used to guide physician compensation, even at the expense of decreasing physician job satisfaction [19].

Our work provides evidence that current implementations may be useful but appear to require further study on validity. Anesthesiology is practiced in a complex environment, in which patients may be meeting multiple physicians, nurse anesthetists, nurses, and care partners. In this context, trying to tie overall satisfaction to any one member of the team is inherently challenging, as a patient may attribute his or her satisfaction to a complex interplay of factors, as opposed to any one member of the care team.

There are limitations to this study. While data were obtained from questions asking broad questions about various aspects of the anesthetic care delivered, it is not clear that these are the best questions to assess provider quality. Other patient satisfaction questions could conceivably perform better and serve as a better metric of provider quality. The SurveyVitals questions were chosen because they are widely-used in a mix of practice settings and the internal consistency of the APSQ2 has been assessed by SurveyVitals [10] but, as with many other quality metrics, it is possible that they were inadequately validated and have not been

peer reviewed. Additional research is needed to better understand if other questions may better facilitate ranking and evaluating providers.

Even a tool as widely-used as a Likert scale may be challenging for patients [20]. In our analysis of responses, there were multiple instances of patients flipping the Likert scale, particularly in their first response - giving a provider a 1 but offering very positive written feedback. There is a large and robust field of literature around developing high-quality and reproducible survey instruments [21,22]. As designed, existing patient satisfaction surveys may be inadequately implementing these recommendations to facilitate meaningful ranking of anesthesiologists.

While surveys were sent to all patients who underwent anesthesia care during the study period, there was likely response bias - patients at extremes of satisfaction were more likely to respond to the survey. This bias was present for all surveys and for all providers, however, and it is unlikely to have significantly impacted our results. Similarly, our exclusion of providers with < 96 completed surveys to improve power may have biased findings, as providers with less time in the operating room (such as providers with significant academic or non-anesthesia roles) were largely excluded from analysis.

The study was performed at VUMC, a large, academic medical center. The APSQ2 was designed to assess anesthesia care teams comprised of an attending anesthesiologist and CRNA. However, in our department, the surveys were sent to all eligible patients, regardless of the anesthesia care team composition. We focused on cases where an attending anesthesiologist worked with a nurse anesthetist in our primary analysis, as this team composition has been previously studied by SurveyVitals. We also performed a sensitivity analysis to address concerns about how different anesthesia care models may impact patient satisfaction. Further study is needed to determine how a patient satisfaction score reflects CRNA performance. Furthermore, patients may have interacted with multiple anesthesia providers and may not have been clear on which provider was the attending physician. SurveyVitals attempts to ensure whether patients are evaluating the correct person by including providing a picture of the attending physician along with the web survey, but it is possible that some biases or confusion about roles & identities may have persisted. Additionally, the providers' photos were not available to patients during an IVR call. While this limitation should have impacted all attending anesthesiologists equally, it is impossible to fully address concerns about attribution in this methodology. Finally, an alternative explanation for our results could be that anesthesiologists at our institution are working at an essentially equivalent level and there are no actual differences in performance among the group to detect. In this scenario, the survey instrument may be correctly registering that lack of difference and working as designed.

While patient satisfaction ratings permit some ranking of attending anesthesiologists, the value of these rankings is of unclear significance, given the limited ability to discriminate between most providers after adjusting for confounding factors. However, ranking individuals is not the purpose for using the instrument. Instead, the department is focused on ensuring that it is on par with high-performing anesthesia groups with respect to patient satisfaction, and on providing individual clinicians de-identified, patient level specific

feedback about positive interactions and opportunities to improve. Meanwhile, the aggregate data may be used to assess the overall satisfaction of patients with the care delivered by the department, and it could also be used to compare our department to other hospitals.

Additionally, members of the department (including the department chair, WSS) have found individual patient comments in SurveyVitals to be useful in identifying themes in the feedback and for developing personal interaction approaches to address such concerns. Members of the faculty have found free-text comments from patients particularly valuable and are able to review them on an ongoing basis. For example, a patient gave WSS a low score on the question related to discussing options for anesthetic care. The patient's comment paraphrases to: "I just don't recall that any options were discussed, but I was pleased with the overall care." When no options for anesthetic technique are practicably available, WSS had heretofore omitted discussion of options. In response to this comment, WSS has included in his preoperative conversation an explicit discussion of why there are no options for the primary anesthetic technique, and also makes a point of inviting patient participation in other planning, such as analgesic and antiemetic strategies.

Anecdotally, patient comments often help to clarify extreme Likert scores and identify patients in need of follow-up or those who have experienced adverse events. In the study period, 399 patients requested additional follow-up from our department (Appendix B). As these data are not structured, they were not totally analyzed in this study. However, a thematic analysis of free-text comments related to dissatisfaction with care is an important area for further research. A better understanding of themes in the comments that span some or all providers could help the department optimize patient care in the future by implementing systematic changes. Furthermore, we did not study the change in scores over time. Future studies will determine if the process of obtaining and reviewing feedback leads to improved patient satisfaction scores over time. Future studies should better evaluate how to merge patient satisfaction data with other important quality metrics, to better facilitate identifying outlying providers who need help with patient interaction, or, more likely, to identify care environments where the department needs to develop strategies to help providers mitigate challenges imposed by the environment. For example, many preoperative preparation areas are open rooms with bed spaces separated by curtains. This arrangement makes respecting privacy (a specific question in the survey) challenging, but can be mitigated by explicitly stating that the anesthesiologist is open to finding a private space to discuss sensitive issues, or even by asking permission to move closer to the patient so that discussions can occur quietly.

In summary, the SurveyVitals tool distinguishes poorly between most anesthesiologists with respect to patient satisfaction with anesthetic care, though it does allow identification of a few consistently high- and low-scoring clinicians with respect to the median performing clinician. Groups of anesthesiologists might systematically learn and adopt techniques from anesthesiologists who receive very high scores, and they might search for ways to help colleagues who receive very low satisfaction scores improve. Importantly, the SurveyVitals tool provides no information on why low or high scores are given, though the text comments might be useful. Consequently, the tool is better suited to identify opportunities to improve than to rank order individual anesthesiologists.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Disclosures:

Funding: Dr. Freundlich receives ongoing support from the NIH-National Center for Advancing Translational Sciences (NCATS) #1KL2TR002245.

APPENDIX

Appendix A.1.

The Current Scope of SurveyVitals in Anesthesia-Specific Patient Experience Solution.

1. 119 Organizations
2. 2,316 Locations/Facilities
3. 12,236 Providers
4. 969,237 Anesthesia Surveys Year-to-date ^a
5. 112,250 Anesthesia Surveys Avg/Month (last 4 months) ^b
6. 1,300,000 Anesthesia Surveys expected in 2019

^aFrom 1/1/2019–9/15/2019

^bFrom May/2019–Sep/2029

* Data source: SurveyVitals, received on September 16, 2019.

Appendix A.2.

The Internal Consistency Study of APSQ2 from SurveyVitals of 2019-Q3 for Divisions and Providers.

	Confidence Level (%)	Completed Surveys
Divisions (Facilities or Sites)	70	100
	80	145
	90	222
	95	294
	99	435
Anesthesiologists, CRNAs, AAs	70	59
	80	79
	90	87
	95	96
	99	107

* Data source: SurveyVitals, received on November 7, 2019.

Appendix B.

Examples of the Follow-ups of Contact Requests from Free-Text Comments.

- | |
|---|
| 1. Accidentally requested to be contacted. No issues or concerns. |
|---|

2. Accidentally requested to be contacted. Spouse did have a question concerning change in lower extremity function. Spouse encouraged to contact patient's oncologist and let them know.

3. Everything ok now. Pt instructed to notify anesthesia team of previous experience if having another procedure. Pt instructed to follow up with primary care MD.

4. Everything was fine but complains about throat discomfort. Explained sore throat can occur with general anesthesia. Encouraged to follow up with physician.

5. Accidentally requested to be contacted. However, did request to have same in room provider for next procedure. Request submitted to AIC and Lead CRNAs.

6. Very appreciative of the care received. Pt didn't know what expect from anesthesia and was a little scared, but doing much better. No issues now.

7. Spoke with spouse, patient unavailable at this time. Encourage patient follow up with surgeon regarding concerns he had expressed in the survey. Wife indicated patient has a follow up appointment with surgeon next week.

8. Pt took suboxone prior to procedure so didn't feel fentanyl worked, also felt needle for block. Pt instructed to let providers know during next procedures.

9. Pt was a little upset about the postoperative hallucinations. Other than that, he was pleased with care

References

- [1]. Institute of Medicine Committee on Quality of Health Care in A. Crossing the Quality Chasm: A New Health System for the 21st Century Washington (DC): National Academies Press (US) Copyright 2001 by the National Academy of Sciences All rights reserved., 2001.
- [2]. Epstein RH, Dexter F, Schwenk ES. Hypotension during induction of anaesthesia is neither a reliable nor a useful quality measure for comparison of anaesthetists' performance. *Br J Anaesth* 2017;119:106–14. [PubMed: 28974070]
- [3]. Freundlich RE, Ehrenfeld JM. Quality metrics: hard to develop, hard to validate. *Br J Anaesth* 2017;119:10–1. [PubMed: 28974060]
- [4]. Wanderer JP, Shi Y, Schildcrout JS, Ehrenfeld JM, Epstein RH. Supervising anesthesiologists cannot be effectively compared according to their patients' postanesthesia care unit admission pain scores. *Anesth Analg* 2015;120:923–32. [PubMed: 25325754]
- [5]. Schonberger RB, Barash PG, Lagasse RS. The Surgical Care Improvement Project Antibiotic Guidelines: Should We Expect More Than Good Intentions? *Anesth Analg* 2015;121:397–403. [PubMed: 26197373]
- [6]. Fleisher LA. Quality AnesthesiaMedicine Measures, Patients Decide. *Anesthesiology* 2018;129:1063–9. [PubMed: 30273268]
- [7]. Barnett SF, Alagar RK, Grocott MP, Giannaris S, Dick JR, Moonesinghe SR. Patient-satisfaction measures in anesthesia: qualitative systematic review. *Anesthesiology* 2013;119:452–78. [PubMed: 23669268]
- [8]. Daskivich TJ, Houman J, Fuller G, Black JT, Kim HL, Spiegel B. Online physician ratings fail to predict actual performance on measures of quality, value, and peer review. *J Am Med Inform Assoc* 2018;25:401–7. [PubMed: 29025145]
- [9]. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;370:1453–7. [PubMed: 18064739]
- [10]. SurveyVitals. SurveyVitals Anesthesia Survey Undergoes Cronbach's Alpha Analysis, <https://www.surveyvitals.com/start/surveyvitals-anesthesia-survey-undergoes-chronbachs-alpha-analysis/>; 2015 [accessed 7 August 2015].
- [11]. Pozdnyakova A, Tung A, Dutton R, Wazir A, Glick DB. Factors Affecting Patient Satisfaction With Their Anesthesiologist: An Analysis of 51,676 Surveys From a Large Multihospital Practice. *Anesth Analg* 2019.
- [12]. Fortney JT, Gan TJ, Graczyk S, Wetchler B, Melson T, Khalil S, et al. A Comparison of the Efficacy, Safety, and Patient Satisfaction of Ondansetron Versus Droperidol as Antiemetics for Elective Outpatient Surgical Procedures. *Anesth Analg* 1998;86:731–8. [PubMed: 9539593]

- [13]. Kynes JM, Schildcrout JS, Hickson GB, Pichert JW, Han X, Ehrenfeld JM, et al. An analysis of risk factors for patient complaints about ambulatory anesthesiology care. *Anesth Analg* 2013;116:1325–32. [PubMed: 23385054]
- [14]. Tiwari V, Queenan C, St. Jacques P. Impact of waiting and provider behavior on surgical outpatients' perception of care. *PCORM*. 2017;7:7–11.
- [15]. Chen Y, Cai A, Fritz BA, Dexter F, Pryor KO, Jacobsohn E, et al. Amnesia of the operating room in the B-Unaware and BAG-RECALL clinical trials. *Anesth Analg* 2016;122:1158–68. [PubMed: 26859877]
- [16]. Mesrobian J, Barnett SR, Domino KB, Mackey D, Pease S, Urman R. Patient Satisfaction with Anesthesia White Paper. ASA Committee on Performance and Outcomes Measurement (CPOM) Patient Satisfaction workgroup. 2019:1–15.
- [17]. Cloney M, Hopkins B, Shlobin N, Dahdaleh NS. Online Ratings of Neurosurgeons: An Examination of Web Data and its Implications. *Neurosurgery* 2018;83:1143–52. [PubMed: 29618127]
- [18]. Wang DE, Wadhera RK, Bhatt DL. Association of Rankings With Cardiovascular Outcomes at Top-Ranked Hospitals vs Nonranked Hospitals in the United States. *JAMA Cardiol* 2018.
- [19]. Zgierska A, Rabago D, Miller MM. Impact of patient satisfaction ratings on physicians and clinical care. *Patient preference and adherence* 2014;8:437–46. [PubMed: 24729691]
- [20]. Sullivan GM, Artino AR Jr. Analyzing and interpreting data from likert-type scales. *Journal of graduate medical education* 2013;5:541–2. [PubMed: 24454995]
- [21]. Rickards G, Magee C, Artino AR Jr. You Can't Fix by Analysis What You've Spoiled by Design: Developing Survey Instruments and Collecting Validity Evidence. *Journal of graduate medical education* 2012;4:407–10. [PubMed: 24294413]
- [22]. Artino AR Jr., Gehlbach H, Durning SJ. AM last page: Avoiding five common pitfalls of survey design. *Acad Med* 2011;86:1327. [PubMed: 21955721]

Highlights:

- A large sample of patient satisfaction survey data were analyzed.
- Nighttime surgery was associated with a lower satisfaction score.
- Decreasing age of patient was associated with a higher satisfaction score.
- Regional anesthesia was associated with a higher satisfaction score.
- While extreme outliers were identified, most providers could not be distinguished.

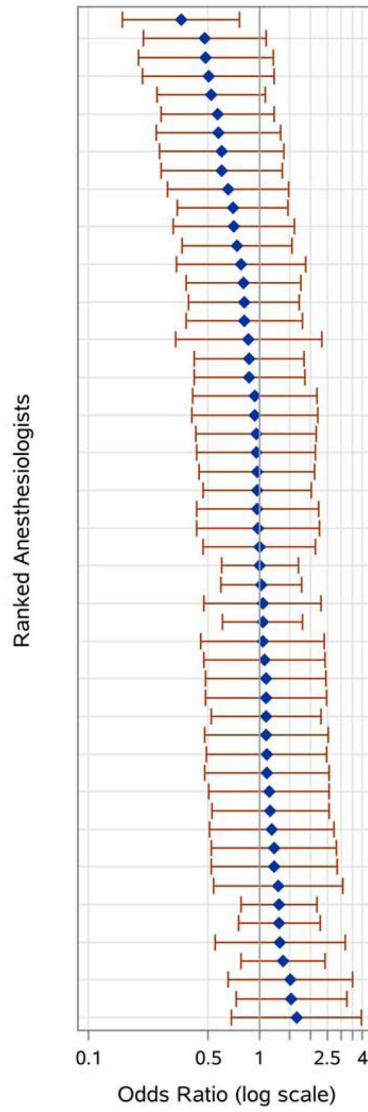


Fig. A. Visualization of the distribution of comparisons to the median-ranked individual anesthesiologist, that derived from the multivariable ordered probit regression model (10,528 surveys, 55 attending anesthesiologists). The ORs panel shows the OR (95% CI) of receiving a higher satisfaction score compared to the median-ranked anesthesiologist.

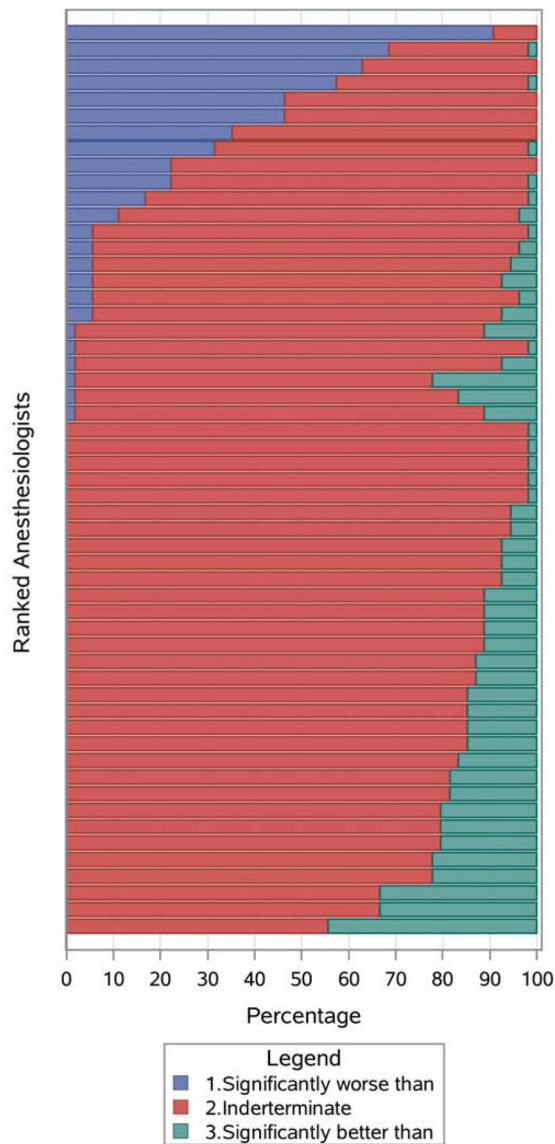


Fig. B. Visualization of the pairwise comparison distribution that derived from the multivariable ordered probit regression model (10,528 surveys, 55 attending anesthesiologists). The percentage panel reveals the proportion of other individual anesthesiologists for whom each individual was statistically significantly worse than (blue), indeterminate (red) and better than (green).

Table A.

Patient Satisfaction Survey Questionnaires Sample

Questionnaires	Questionnaires Text	Included in Analysis
Q1	Please select the provider(s) for whom you remember enough to answer questions about the care provided.	No
Q2	Were you able to spend time with your anesthesia provider before surgery?	No
Q3	Your anesthesia provider did his or her best to respect your privacy.	Yes
Q4	Your options for anesthesia were explained before your surgery.	Yes
Q5	Your questions about anesthesia, the process, risks, and possible after effects were answered.	Yes
Q6	You were well prepared to make informed decisions.	Yes
Q7	Your anesthesiologist helped ease any anxiety you were feeling.	Yes
Q8	Your anesthesiologist ensured your comfort during the surgical experience.	Yes
Q9	Please share any thoughts or concerns from your visit to the operating suite.	No
Q10	Using a number from 5 to 1, where 5 is the best anesthesiologist possible and 1 is the worst, please rate your anesthesiologist.	Yes
Q11	Please share any additional comments about your anesthesiologist.	No
Q12–15	<i>Nurse Anesthetist Questions</i>	No
Q16	Was your anesthesia provider available to answer questions after surgery?	No
Q17	Did you experience nausea or vomiting after surgery?	No
Q18	Please add any comments you would like to make about your experience immediately after surgery.	No

Table B.

The Distribution of Patient Responses to the APSQ2 Survey.

	Strongly Disagree, Disagree, or Neutral	Strongly Agree or Agree	Total Cases	Response Rate
Q3	1,151 (5.4%)	20,347 (94.6%)	21,498	75.0%
Q4	1,690 (7.5%)	20,888 (92.5%)	22,578	78.7%
Q5	973 (4.3%)	21,421 (95.7%)	22,394	78.0%
Q6	712 (3.2%)	21,819 (96.8%)	22,531	78.5%
Q7	996 (4.8%)	19,647 (95.2%)	20,643	71.9%
Q8	692 (3.1%)	21,678 (96.9%)	22,370	78.0%
Q10	582 (2.8%)	19,949 (97.2%)	20,531	71.6%

* APSQ2: Anesthesia Patient Satisfaction Questionnaire 2.0.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table C.

Demographic Information of Patients Stratified by Overall Patient Satisfaction Score

	Overall Patient Satisfaction Score				
	Score 1 (Strongly Disagree)	Score 2 (Disagree)	Score 3 (Neutral)	Score 4 (Agree)	Score 5 (Strongly Agree)
Cases (N)	10	49	494	3,669	6,306
Age in Years, mean (SD)	33.7 (23.6)	48.1 (20.7)	50.7 (22.2)	48.2 (23.6)	42.4 (24.6)
BMI in kg/m ² , mean (SD)	25.8 (10.2)	27.8 (7.0)	27.9 (7.8)	27.9 (9.4)	27.1 (9.9)
Gender (%)					
Female	60.0%	59.2%	54.9%	56.7%	55.4%
Race (%)					
Caucasian	80.0%	85.7%	82.6%	81.7%	81.6%
Black	10.0%	10.2%	8.7%	10.0%	9.1%
Asian	0.0%	0.0%	1.2%	1.2%	1.1%
American Indian	0.0%	0.0%	0.4%	0.1%	0.1%
Unknown	10.0%	4.1%	7.1%	7.0%	8.1%
ASA (%)					
I	0.0%	2.0%	5.0%	7.1%	9.5%
II	70.0%	44.9%	43.1%	45.5%	49.6%
III	30.0%	49.0%	48.0%	43.8%	37.7%
IV&V	0.0%	4.1%	3.9%	3.6%	3.2%
Patient Status (%)					
Same Day Surgery	80.0%	89.8%	85.4%	82.2%	82.2%
Observation Patient	10.0%	8.2%	5.7%	7.0%	7.2%
Inpatient	0.0%	2.0%	2.8%	3.1%	3.2%
Others	10.0%	0.0%	6.1%	7.7%	7.4%
Anesthesia Type (%)					
General	100.0%	81.6%	84.6%	82.9%	82.3%
MAC	0.0%	18.4%	14.8%	15.7%	15.6%
Regional	0.0%	0.0%	0.6%	1.4%	2.1%

Overall Patient Satisfaction Score					
	Score 1 (Strongly Disagree)	Score 2 (Disagree)	Score 3 (Neutral)	Score 4 (Agree)	Score 5 (Strongly Agree)
Surgery Start Time (%)					
6 AM to 12 PM	70.0%	61.2%	66.2%	67.0%	68.2%
12 PM to 6 PM	30.0%	38.8%	32.4%	32.4%	31.3%
6 PM to 6 AM	0.0%	0.0%	1.4%	0.6%	0.5%

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table D.

Multivariable Fixed Effects Modeling Analysis of the Response of APSQ2.

	Covariates	Wald Test χ^2	p-value
Ordered Probit Regression	Anesthesiologist	163.36	<.001*
	Patient Age	61.89	<.001*
	Patient Gender	0.86	0.650
	CRNA	171.46	0.146
	Surgery Start Time	6.72	0.035*
	ASA Class	0.01	0.809
	Anesthesia Type	11.18	0.004*
	Surgery Service	39.56	0.490
	Surgeon	309.50	1.000
	Logistic Regression	Anesthesiologist	127.35
Patient Age		6.21	0.013*
Patient Gender		2.49	0.289
CRNA		155.98	0.418
Surgery Start Time		7.99	0.018*
ASA Class		2.68	0.101
Anesthesia Type		2.47	0.291
Surgery Service		30.77	0.853
Surgeon		442.31	0.198

* Level of significance p = 0.05.

CRNA: Certified registered nurse anesthetists; ASA: American Society of Anesthesiologists Physical Status Classification.