

OPEN

Development and Evaluation of a Risk-Adjusted Measure of Intraoperative Hypotension in Patients Having Nonemergent, Noncardiac Surgery

Anna L. Christensen, PhD,* Ethan Jacobs, MPP† Kamal Maheshwari, MD, MPH,‡
 Fei Xing, PhD,* Xiaohong Zhao, PhD,§ Samuel E. Simon, PhD,† Karen B. Domino, MD, MPH,||
 Karen L. Posner, PhD,|| Alvin F. Stewart, MD,¶ Joseph A. Sanford, MD,¶ and Daniel I. Sessler, MD#

BACKGROUND: Intraoperative hypotension is common and associated with organ injury and death, although randomized data showing a causal relationship remain sparse. A risk-adjusted measure of intraoperative hypotension may therefore contribute to quality improvement efforts.

METHODS: The measure we developed defines hypotension as a mean arterial pressure <65 mm Hg sustained for at least 15 cumulative minutes. Comparisons are based on whether clinicians have more or fewer cases of hypotension than expected over 12 months, given their patient mix. The measure was developed and evaluated with data from 225,389 surgeries in 5 hospitals. We assessed discrimination and calibration of the risk adjustment model, then calculated the distribution of clinician-level measure scores, and finally estimated the signal-to-noise reliability and predictive validity of the measure.

RESULTS: The risk adjustment model showed acceptable calibration and discrimination (area under the curve was 0.72 and 0.73 in different validation samples). Clinician-level, risk-adjusted scores varied widely, and 36% of clinicians had significantly more cases of intraoperative hypotension than predicted. Clinician-level score distributions differed across hospitals, indicating substantial hospital-level variation. The mean signal-to-noise reliability estimate was 0.87 among all clinicians and 0.94 among clinicians with >30 cases during the 12-month measurement period. Kidney injury and in-hospital mortality were most common in patients whose anesthesia providers had worse scores. However, a sensitivity analysis in 1 hospital showed that score distributions differed markedly between anesthesiology fellows and attending anesthesiologists or certified registered nurse anesthetists; score distributions also varied as a function of the fraction of cases that were inpatients.

CONCLUSIONS: Intraoperative hypotension was common and was associated with acute kidney injury and in-hospital mortality. There were substantial variations in clinician-level scores, and the measure score distribution suggests that there may be opportunity to reduce hypotension which may improve patient safety and outcomes. However, sensitivity analyses suggest that some portion of the variation results from limitations of risk adjustment. Future versions of the measure should risk adjust for important patient and procedural factors including comorbidities and surgical complexity, although this will require more consistent structured data capture in anesthesia information management systems. Including structured data on additional risk factors may improve hypotension risk prediction which is integral to the measure's validity. (*Anesth Analg* 2021;133:445–54)

KEY POINTS

- **Question:** Does a risk-adjusted measure of intraoperative hypotension vary across providers and have good reliability and validity, allowing anesthesia providers to use it for quality improvement?
- **Findings:** The risk-adjusted measure of intraoperative hypotension has high provider-level variation and evidence of reliability and validity, but risk adjustment remains suboptimal.
- **Meaning:** The measure may be a useful tool for anesthesia providers involved in quality improvement efforts.

From the *Mathematica, Washington, DC; †Mathematica, Cambridge, Massachusetts; ‡Departments of General Anesthesiology and Outcomes Research, Cleveland Clinic, Cleveland, Ohio; §Mathematica, Princeton, New Jersey; ||Department of Anesthesiology & Pain Medicine, University

Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the International Anesthesia Research Society. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

DOI: 10.1213/ANE.0000000000005287

of Washington, Seattle, Washington; ¶Department of Anesthesiology, University of Arkansas for Medical Sciences, Little Rock, Arkansas; and #Department of Outcomes Research, Cleveland Clinic, Cleveland, Ohio.

Accepted for publication October 6, 2020.

Funding: Edwards Lifesciences funded this study.

Conflicts of Interest: See Disclosures at the end of the article.

Reprints will not be available from the authors.

Address correspondence to Anna L. Christensen, PhD, Mathematica, 1100 1st St NE, 12th Floor, Washington, DC 20002. Address e-mail to achristensen@mathematica-mpr.com.

GLOSSARY

AKI = acute kidney injury; **ASA** = American Society of Anesthesiologists; **AUC** = area under the receiver operating characteristic curve; **BMI** = body mass index; **CI** = confidence interval; **CRNA** = certified registered nurse anesthetist; **DBP** = diastolic blood pressure; **IRB** = institutional review board; **MAP** = mean arterial pressure; **O:E** = observed to expected; **RCT** = randomized controlled trial; **SBP** = systolic blood pressure; **SD** = standard deviation

Intraoperative hypotension is associated with poor perioperative outcomes. A 2018 systematic review of 40 observational studies and 2 small randomized controlled trials (RCTs) found an association between low mean arterial pressure (MAP) and organ injury, with greater risk at lower pressures and prolonged exposures.¹ MAP <65 mm Hg for ≥5 minutes was considered moderate risk, and MAP <65 mm Hg for ≥20 minutes was considered high risk.¹ A 2019 meta-analysis of 15 observational studies^{2–16} reported an association between intraoperative hypotension (MAP < 60 mm Hg) and acute kidney injury (AKI), myocardial injury after noncardiac surgery, and 30-day mortality.¹⁷ While there is no standardized definition of intraoperative hypotension,¹⁸ a recent observational study considered multiple definitions and found that patients' risk of myocardial and kidney injury increased substantially after 13 minutes at an MAP <65 mm Hg.¹¹ The investigators also found that absolute intraoperative blood pressure predicted myocardial injury and kidney injury as well as reductions from clinic baseline pressures did.¹¹ Currently, there are sparse randomized data suggesting a causal relationship between intraoperative hypotension and organ injury and death.

Clinician-level assessment of process and outcome measures is a common strategy for improving quality of care. A measure for intraoperative hypotension is plausible because blood pressure is routinely recorded and modifiable.¹⁸ While patient characteristics affect presurgical risk of hypotension, anesthesia providers regularly monitor blood pressure and mitigate the risk of hypotension. The association between hypotension and adverse outcomes remains even after controlling for patient risk factors, including diagnoses, cardiac medication history, surgical procedure, American Society of Anesthesiologists (ASA) physical status, and age.¹¹ Also, hypotension can be objectively measured in a low-burden way by extracting intraoperative blood pressure readings directly from anesthesia information management systems. The Multicenter Perioperative Outcomes Group's Anesthesia Performance Improvement and Reporting Exchange includes non-risk-adjusted hypotension as a quality measure for noncardiac surgery¹⁹; however, risk adjustment is important to account for some of the observable characteristics of patients and procedures that affect blood pressure.

We describe development and evaluation of a risk-adjusted measure of intraoperative hypotension, defined as an MAP <65 mm Hg for at least 15 cumulative minutes. The measure assesses whether clinicians had more or fewer cases with hypotension than expected over a 12-month period, given their patient mix. We tested the hypothesis that our risk adjustment model has sufficient discrimination and calibration and that the measure has sufficient provider-level variation, reliability, and validity to support its use for quality improvement.

METHODS

Data Sources

The measure was developed and evaluated using data from 5 hospitals in 3 health systems that provided deidentified, patient-level data extracts from their electronic records. The data set included all nonemergent, noncardiac surgical cases from 2016 to 2017. One hospital also provided data from cases in 2015 and 2018. The institutional review boards (IRBs) from each of the 3 institutions assessed the work and determined that IRB review and approval was not required (2 determined that it did not constitute human subjects research; 1 determined that the institution was not engaged in the work, as defined by federal regulation); therefore, the IRBs did not require written informed consent from patients for this work.

Measure Development Process

We defined the measure population and numerator and then developed and validated a risk adjustment model. The development process was informed by a literature review, site visits to 4 hospitals to assess data systems and feasibility of reporting, a technical panel consisting of anesthesia providers and clinical researchers who are experts in hypotension, a second panel consisting of typical clinicians who might use the measure, a patient and caregiver advisory panel, and a 30-day public comment period.

Measure Specifications

We defined the denominator for the proposed measure as cases in which adults ≥18 years of age had nonemergency (including elective and urgent) noncardiac surgery that required general anesthesia or monitored anesthesia care. The following patients and surgeries were excluded from the denominator:

patients with ASA physical status V or VI; patients with a preinduction MAP <65 mm Hg; and patients having obstetric nonoperative procedures, liver transplants, lung transplants, and cataract surgeries. Supplemental Digital Content 1, Appendix 1, <http://links.lww.com/AA/D254>, contains detailed measure specifications and instructions for calculating.

Outcome

We defined the measure numerator—intraoperative hypotension—as MAP <65 mm Hg for a cumulative duration of at least 15 minutes, based on research from Salmasi et al.¹¹ To calculate cumulative duration, each MAP value <65 mm Hg was counted for the number of minutes between the index reading and the next reading (typically 1–5 minutes later), capped at a maximum of 5 minutes. If blood pressure was monitored using both invasive and noninvasive means, direct measurements were used to determine the outcome. Artifactual blood pressure values were removed from the data set before calculating the outcome. Values presumed to be artifacts included those documented as an artifact by the clinician, systolic blood pressure (SBP) ≤20 mm Hg or ≥300 mm Hg, diastolic blood pressure (DBP) ≤5 mm Hg or ≥225 mm Hg, readings with SBP and DBP within 5 mm Hg, or MAP ≤30 mm Hg or ≥250 mm Hg.

Covariates for Risk Adjustment

Risk adjustment variables included patient age (modeled as a continuous variable), ASA physical status (modeled categorically), body mass index (BMI; modeled as a continuous variable), duration of surgery (modeled categorically in 60-minute increments), and patient sex. Risk adjustment was limited to variables that are regularly recorded in structured data fields in the anesthesia information management system and can be automatically extracted to increase the feasibility and lower the burden of future use of the measure. The ASA physical status categories have standard definitions and are universally used for clinical care and billing. While ASA physical status assignments are subject to clinical interpretation of the physician or nurse assigning the status,²⁰ they nonetheless have acceptable interrater reliability and are statistically significantly correlated with patient age, Charlson comorbidity index, revised cardiac risk index, and length of stay.²¹ Also, ASA physical status is consistently recorded in the anesthesia record in a structured field that makes reporting the measure feasible and low burden.

Duration of surgery was included because longer surgeries are often complex and provide more opportunity for hypotension to accrue. Risk adjustment models typically control for risk factors that are present before the start of the surgery and are not affected by the clinician; while the length of the procedure can

be roughly estimated before the procedure begins, it is also affected by intraoperative events. To minimize the possibility that an episode of hypotension would affect the surgery length variable, our risk adjustment model considered time as a categorical variable, with each category representing 60-minute increments. It would be rare for an episode of hypotension to cause a case to move from one surgery length category to another. Therefore, even if hypotension lengthens the surgery slightly, it would rarely influence risk adjustment.

Risk Adjustment Model Development and Validation

We developed and validated the risk adjustment model using data from 4 hospitals in 2 health systems (systems 1 and 2; initial data set; 178,343 cases after exclusions). We then further validated the model in 1 hospital in a third health system that provided data after measure development was complete (system 3; novel data set; 37,866 cases after exclusions). We randomly split the initial data set in half (development half and validation half) and estimated a logistic regression model in the development half. We tested different functional forms (categorical versus continuous) of each risk adjustment variable and tested interaction terms (including but not limited to interactions between ASA physical status and surgery length and between age and surgery length). Risk adjustment model coefficients are given in Supplemental Digital Content 2, Appendix, Table A, <http://links.lww.com/AA/D255>.

In the validation half, the model calibration was assessed using A/B testing and decile plots, and model discrimination was assessed by the area under the receiver operating characteristic curve (AUC) and its 95% confidence interval. Model calibration and discrimination were then similarly assessed in the novel data set.

Provider Attribution and Measure Calculation

Clinician-level unadjusted and risk-adjusted measure scores were calculated for anesthesiologists, anesthesia fellows, and certified registered nurse anesthetists (CRNAs). Unadjusted scores were the percentage of a clinician's cases during a 12-month measurement period that had hypotension (at least 15 minutes with an MAP <65 mm Hg). When multiple clinicians work on a case, either as a team or sequentially, the full case was attributed to each of the clinicians who signed into the case for at least 1 minute. Therefore, full cases could be attributed to multiple anesthesia providers.

Risk-adjusted scores were calculated as the observed to expected (O:E) ratio of the clinician's cases that had hypotension during the 12-month measurement period. To estimate the expected number of hypotension cases, the risk adjustment model calculated the

log odds that a given case would develop hypotension; log odds were then transformed into predicted probabilities. For each clinician, we then summed the predicted probabilities of all cases that met the measure inclusion criteria, resulting in an expected number of hypotension cases during the 12-month assessment period. The risk-adjusted score (O:E ratio) was calculated as the number of each clinician's cases that developed hypotension divided by the number of cases the risk adjustment model predicted that the clinician would have in the 12-month period.

We assessed the distribution of clinician-level O:E ratios overall and by hospital. For 1 large hospital that provided extra information on staff type, we did a sensitivity analysis to show the distribution of scores by staff type and percent of their cases during the year that are inpatient surgeries.

Statistical Analysis: Measure Reliability and Validity

We conducted reliability testing of the clinician-level, risk-adjusted scores to assess whether the scores were reflections of differences between clinicians rather than measurement error. Reliability was assessed using a signal-to-noise ratio analysis, similar to the beta-binomial method described by Adams,²² but adapted for the outcome being a ratio rather than a percentage. To calculate the signal-to-noise ratio of the risk-adjusted score (O:E ratio) for each clinician, we followed a multilevel hierarchical regression approach to estimate the signal and noise.

Specifically, we first estimated the within-clinician variability ("noise") by calculating the variance of the $\frac{\sum O_i}{\sum E_i} \cdot \bar{Y}$ within each clinician, where the randomness is contributed by the observed outcome (ie, MAP <65 mm Hg for cumulative total of at least 15 minutes) of each case within the clinician. In the context of the logistic regression of the risk adjustment model, the noise may be represented as follows:

$$\bar{Y}^2 \times \frac{\sum_{i=1}^{n_k} \text{Var}(Y_k^i)}{\left(\sum_{i=1}^{n_k} E(Y_k^i)\right)^2} = \bar{Y}^2 \times \frac{\sum_{i=1}^{n_k} p_i(1-p_i)}{\left(\sum_{i=1}^{n_k} p_i\right)^2},$$

where p_i is the estimated probability of surgery i with an event and n_k is the number of surgeries for clinician k . We next estimated the between-clinician variance ("signal") iteratively, using a maximum likelihood estimation approach described by Morris.²³ We computed the signal-to-noise ratio statistic, R , as the ratio of the signal variance (which is common across all entities) to the sum of the signal variance and the noise variance (which varies by entity): $R = \frac{\sigma_{\text{between}}^2}{\sigma_{\text{between}}^2 + \sigma_{\text{within}}^2}$.

A signal-to-noise ratio of ≥ 0.7 is considered acceptable for drawing conclusions about clinicians as a

group; a signal-to-noise ratio of ≥ 0.9 is considered acceptable for drawing conclusions about individual clinicians.²² We assessed reliability distributions for clinicians stratified by the number of surgical cases they had in the measurement period, because reliability estimates decrease with smaller denominator sizes. We present reliability for clinicians with 1–30 cases versus those with >30 cases during the measurement year.

We tested the predictive validity of the clinician-level, risk-adjusted measure (O:E ratio) to assess whether the measure was associated with adverse surgical outcomes that have been shown to be related to intraoperative hypotension in previous research. First, we classified clinicians by whether their risk-adjusted measure scores were significantly <1.0 (fewer cases of hypotension than expected), not statistically different from 1.0, and significantly >1.0 (more cases of hypotension than expected). To do so, we estimated each clinician's individual confidence interval based on his or her unadjusted measure rate and number of cases, and we determined whether his or her confidence interval included 1.0. Next, within each clinician group, we calculated the incidence of AKI and in-hospital mortality among their patients during the measurement period. To align with the AKI measure from the Multicenter Perioperative Outcomes Group, AKI was defined as an increase of 1.5 times the baseline serum creatinine level within 7 days after surgery or an increase in serum creatinine of >0.3 mg/dL with 48 hours after the end of anesthesia.¹⁹ AKI incidence was calculated only for the 41% of patients who had creatinine recorded.

RESULTS

The study cohort included 225,389 unique operations in 167,125 patients who met the criteria for the measure's initial population, presided over by 922 anesthesia clinicians (including anesthesiologists, anesthesiology fellows, and CRNAs) in 5 hospitals within 3 health systems (Table 1). Each case was fully attributed to every clinician who was signed into the case for at least 1 minute. Thirty-five percent of surgeries were in patients ≥ 65 years of age. Nearly half of surgical cases were classified as ASA physical status III, and another 13% were classified as physical status IV; only 4.4% were designated as physical status I.

Risk Adjustment Model Validation

Table 2 shows calibration and discrimination of the measure's risk adjustment model in the development and validation half-samples and in the novel validation sample. The AUC was 0.72 in the validation half-sample and 0.73 in the novel validation sample. Supplemental Digital Content 3, Appendix 2, Figure A.1, <http://links.lww.com/AA/D256>, shows the calibration plot.

Table 1. Study Population Overall and by Health System

	System 1		System 2		System 3		Total	
	N	Column %	N	Column %	N	Column %	n	Column %
No. of surgeries ^a	125,206		60,713		39,470		225,389	
No. of providers	629		205		88		922	
No. of hospitals	3		1		1		5	
Age								
18–39 y	19,834	16.2	13,108	22.0	8870	22.9	41,812	18.9
40–64 y	54,642	44.6	28,184	47.3	18,735	48.3	101,561	46.0
≥65 y	47,930	39.2	18,326	30.7	11,154	28.8	77,410	35.1
Sex: Female	64,874	53.0	32,998	55.3	20,529	53.0	118,401	53.6
BMI ^b (mean)	29.6	...	30.3	...	29.3	...	29.7	...
Race								
White	100,185	81.8	43,531	73.0	32,051	82.7	175,767	79.6
Black	15,081	12.3	13,348	22.4	2097	5.4	30,526	13.8
Asian	1136	0.9	316	0.5	2503	6.5	3955	1.8
Other	5981	4.9	2338	3.9	2108	5.4	10,427	4.7
Missing	23	0.0	85	0.1	0	0.0	108	0.0
ASA physical status								
I	4025	3.3	2320	3.9	3371	8.7	9716	4.4
II	31,935	26.1	25,309	42.5	15,699	40.5	72,943	33.0
III	62,942	51.4	28,831	48.4	17,027	43.9	108,800	49.3
IV	23,504	19.2	3158	5.3	2662	6.9	29,324	13.3
Surgery length								
<60 min	32,650	26.7	5988	10.0	7087	18.3	45,725	20.7
60–119 min	32,752	26.8	22,214	37.3	10,582	27.3	65,548	29.7
120–179 min	19,350	15.8	14,447	24.2	7788	20.1	41,585	18.8
180–239 min	12,915	10.6	7455	12.5	5167	13.3	25,537	11.6
240–299 min	9134	7.5	3917	6.6	3215	8.3	16,266	7.4
300+ min	15,605	12.7	5597	9.4	4920	12.7	26,122	11.8
Intraoperative hypotension								
MAP <65 mm Hg for 15+ minutes	30,436	24.9	28,061	47.1	17,446	45.0	75,943	34.4

Abbreviations: ASA, American Society of Anesthesiologists; BMI, body mass index; MAP, mean arterial pressure.

^aCases among adults ≥18 y of age who had noncardiac, nonemergency surgery under general anesthesia or monitored anesthesia care, before removing exclusions.

^bBMI values of <10 and >200 were considered implausible and were removed before measure calculation (n = 37).

Table 2. Calibration and Discrimination of the Hypotension Risk Adjustment Model in Different Samples

	Calibration (\bar{Y}_0, \bar{Y}_1)	Discrimination AUC (95% CI)
Development sample (1st half; n = 89,172 cases)	(0, 1)	0.714 (0.711-0.718)
Validation sample (2nd half; n = 89,171 cases)	(-0.006, 1.009)	0.719 (0.715-0.722)
Novel validation sample (n = 37,866 cases)	(0.048, 1.179)	0.727 (0.722-0.732)

Abbreviations: AUC, area under receiver operating characteristic curve; CI, confidence interval.

Measure Calculation

Table 3 shows the distribution of the clinician-level unadjusted and risk-adjusted measure scores. The mean clinician-level unadjusted measure score was 39% (standard deviation [SD] = 19%). The mean clinician-level, risk-adjusted score (O:E ratio) was 1.1. The mean value exceeds 1.0 because the expected amount of hypotension was estimated from 2 of the 3 health systems, including one with less hypotension than the others. In 2017, 42% of clinicians had a risk-adjusted score that was statistically significantly <1.0, meaning that they had fewer cases of hypotension than predicted; 22% of clinicians had scores not significantly different from 1.0, and 36% of clinicians had scores

Table 3. Distribution of Clinician-Level Unadjusted and Risk-Adjusted Measure Scores for Intraoperative Hypotension (n = 216,210 Cases; N = 921 Clinician-Level Scores)

	Unadjusted measure ^a	Risk-adjusted measure ^b (O:E ratio)
Mean (SD)	39% (19%)	1.1 (0.5)
Interquartile range	29%	0.7
Min	0%	0.0
10th percentile	16%	0.5
30th percentile	25%	0.7
50th percentile	40%	1.2
70th percentile	50%	1.3
90th percentile	65%	1.5
Max	100%	6.1

Abbreviations: O:E ratio, observed to expected ratio; SD, standard deviation.

^aPercent of clinician's cases in the 12-mo measurement period that have intraoperative hypotension, defined as MAP <65 mm Hg for at least 15 cumulative minutes.

^bNumber of cases of hypotension in the 12-mo measurement period divided by expected number of cases, as determined by the risk adjustment model.

that were statistically significantly >1.0, meaning that they had more cases of hypotension than predicted. Figure 1 shows a histogram of the distribution of O:E ratios; the distribution is bimodal with a peak <1.0 and a peak >1.0.

Supplemental Digital Content 3, Appendix 2, Figure A.2, <http://links.lww.com/AA/D256>, shows

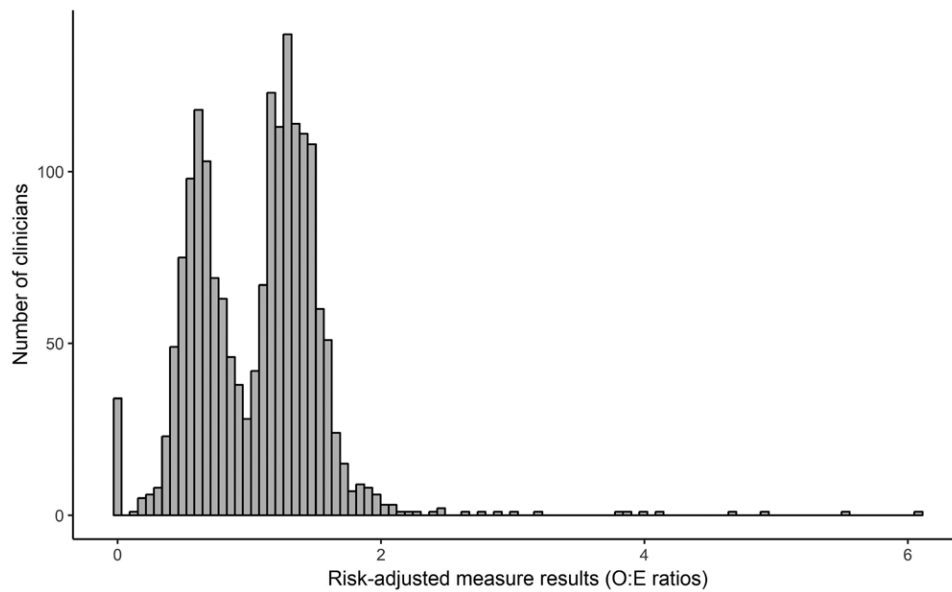


Figure 1. Distribution of clinician-level, risk-adjusted measure results (O:E ratios) for intraoperative hypotension. O:E ratio indicates observed to expected ratio.

the distribution of O:E ratios by hospital and reporting year; the bimodal distribution in Figure 1 is only present in 1 hospital. The hospital-level distributions varied, with some centered below an O:E ratio of 1.0 and some centered above an O:E ratio of 1.0. Supplemental Digital Content 3, Appendix 2, Figures A.3 and A.4, <http://links.lww.com/AA/D256>, presents sensitivity analyses in the hospital with a bimodal distribution (which is the largest hospital and provided additional information on staff type, not available from the others). Supplemental Digital Content 3, Appendix 2, Figure A.2, <http://links.lww.com/AA/D256> shows that this hospital’s score distributions were bimodal for both staff anesthesiologists and CRNAs, while score distributions for anesthesiology fellows, who anecdotally are assigned to longer, more complex cases, had O:E ratios predominantly >1.0. Supplemental Digital Content 3, Appendix 2, Figure A.3, <http://links.lww.com/AA/D256> shows the distribution of provider-level O:E ratios in hospital 1, stratified by the percent of their cases that were inpatient surgeries (the median is 76% inpatient). Providers who had the lowest percent inpatient (60% or lower) had a wide, flat distribution of O:E ratios spread predominantly between 0.25 and 1.75. Providers near the median percent inpatient (61%–80% inpatient) had O:E ratios primarily <1.0, with a long tail >1.0. Providers who nearly always worked on inpatient surgeries (over 80% inpatient) had O:E ratios primarily >1.0, with a long tail <1.0.

Measure Reliability and Validity

Table 4 shows the signal-to-noise reliability testing results for the risk-adjusted measure scores in 2017. The mean signal-to-noise ratio among all clinicians was 0.87. When stratified by denominator size,

Table 4. Distribution of Reliability Coefficients for Risk-Adjusted Hypotension Measure (O:E Ratio) in 2017, by Number of Denominator Anesthesia Cases

No. of clinicians	Signal-to-noise reliability ratio		
	All clinicians	Clinician subgroups, by number of cases	
		Clinicians with 1–30 cases	Clinicians with >30 cases
Mean	0.87	0.38	0.94
25th percentile	0.89	0.15	0.93
50th percentile	0.96	0.41	0.96
75th percentile	0.98	0.57	0.98

Abbreviation: O:E ratio, observed to expected ratio.

clinicians with 1–30 cases during the measurement period (12% of the clinician sample) had a median signal-to-noise ratio of 0.41, while clinicians with >30 cases had a median signal-to-noise ratio of 0.96. The median number of surgical cases per clinician during the measurement year was 288, and the reliability results were similar in 2016. Figure 2 shows that incidence rates of AKI and in-patient mortality were highest among patients whose anesthesia providers had significantly more cases of hypotension than expected over the 12-month period.

We conducted a sensitivity analysis in 1 health system to understand how the measure would change if cases were only attributed to providers signed in for at least 80% of the case. The median clinician-level denominator size decreased by over 250 cases in the year (about 65%), and the median surgery length of the remaining cases was around 70 minutes shorter than in the original sample because clinicians are less likely to be in signed in for 80% of a long operation than a short one. The unadjusted percent of cases with intraoperative hypotension was approximately

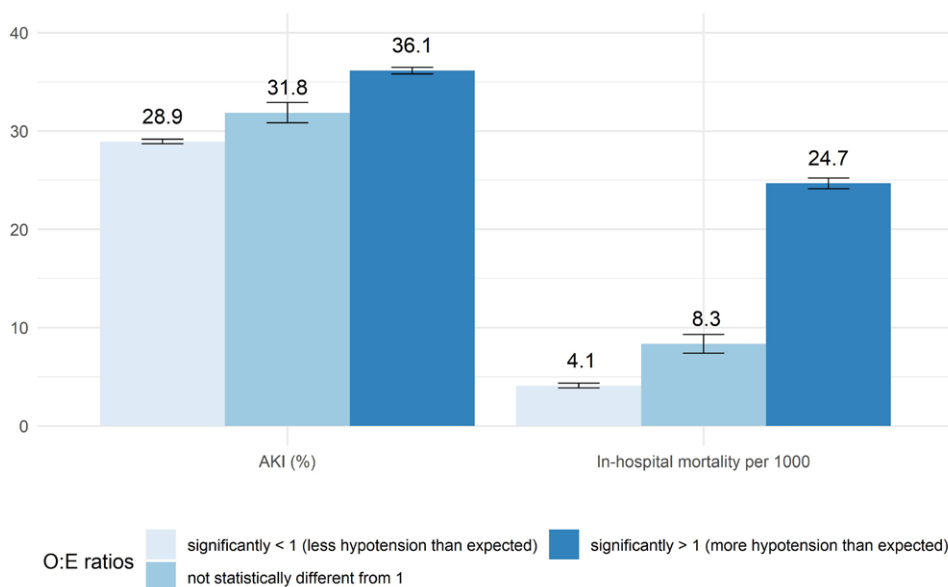


Figure 2. Clinician-level measure score validity: incidence of AKI and in-hospital mortality, by clinician-level, risk-adjusted measure score (N = 717 clinicians for AKI outcome; N = 921 clinicians for mortality outcome). AKI indicates acute kidney injury; O:E ratio, observed to expected ratio.

8% lower under this restricted attribution approach compared to the original attribution approach (27% vs 35% in 2016 and 24% vs 32% in 2017), suggesting that higher-risk cases were not included under the stricter attribution rule.

DISCUSSION

Risk-adjusted intraoperative hypotension may help anesthesia providers to assess hemodynamic control in patients having nonemergent, noncardiac surgery. The simple risk adjustment we propose improves on previous unadjusted measures.¹⁹ Across 5 hospitals, there was substantial variation in the clinician-level measure scores, and a third of anesthesia providers in this analysis had significantly more cases with hypotension than predicted. The results suggest that there may be room for improvement via patient safety initiatives. Because the measure is new, there are not yet benchmarks for the score; broader use in diverse settings could guide future benchmarks overall, and for specific subpopulations of providers or patients.

The measure had sound scientific properties including high signal-to-noise reliability, indicating that the difference between clinicians' scores was to some extent driven by differences in performance rather than statistical noise. However, reliability was low in the small group of anesthesia providers with 30 or fewer cases in the calendar year; consequently, the measure should not be used by clinicians with so few cases per year. Our analysis also suggests that the measure has predictive validity. High risk-adjusted measure scores (indicating more cases of hypotension than predicted) were associated with increased AKI and in-hospital mortality which is consistent with observational analyses.

The measure is not intended to substitute for the clinician's judgment about managing hypotension

for individual patients because clinicians may have sound reasons to allow or target lower pressures in specific patients.²⁴ For example, surgeons may request that the anesthesia clinicians deliberately induce hypotension during certain procedures to lower the risk of bleeding. Nevertheless, in most patients, an MAP <65 mm Hg represents a $\geq 25\%$ reduction from clinic baseline pressure, and clinicians routinely try to avoid such low blood pressures. The risk adjustment model gauges whether clinicians have more or fewer cases of hypotension than expected over a 12-month period, given their patient case mix. We assume that some cases will meet the numerator criteria for the measure for sound clinical reasons or because hypotension developed and persisted despite a clinician's best efforts. The measure does not prescribe specific interventions for stabilizing blood pressure, which is left to clinician judgment.

Our risk adjustment model was intentionally restricted to risk factors that affect hypotension, are outside clinician control, and—most importantly—are consistently recorded in electronic anesthesia records in structured data fields. Restricting required data minimizes the reporting burden of the measure (eg, no manual chart review is required) and facilitates calculating the score for routine use. Uncontrolled surgical and patient factors may nonetheless contribute to variation in clinician's measure scores. For example, our sensitivity analyses in hospital 1 show that anesthesiology fellows and clinicians with the highest fraction of inpatient surgeries generally have risk-adjusted hypotension scores on the higher end of the distribution, suggesting that additional risk adjustment for surgical complexity is needed. Specifically, ASA physical status and duration of surgery appear to be insufficient proxies for complexity. Presumably,

inclusion of additional variables, such as type of surgical procedure, patient history of comorbid conditions, Revised Cardiac Risk Index, cardiac medication history, preoperative renal function, use of arterial catheter, and intraoperative blood loss would improve the measure's risk adjustment. However, those variables are not routinely available in structured data for all patients. With improved, consistent capture of comorbidities and procedure details in structured/hard-coded fields in anesthesia information management systems, future versions of the model could and should risk adjust for other important confounders.

The measure's attribution approach implies team-based responsibility by assigning a case to each clinician who was signed in for any amount of time, an approach similar to other anesthesia and surgery measures. This approach oversimplifies attribution in favor of increasing the feasibility of calculating the measure, and it also allows attending anesthesiologists who are responsible for multiple simultaneous cases to use this measure. A sensitivity analysis of an alternative attribution approach, attributing cases to clinicians only if they were signed in for 80% of the surgery, resulted in fewer qualifying cases per clinician, and the remaining cases were shorter. The unadjusted percent of cases that met the measure's definition of hypotension decreased with the restrictive definition, suggesting that the riskier cases were less likely to meet the restrictive attribution threshold.

The evidence linking hypotension to adverse patient outcomes that provides the rationale for the proposed measure is primarily based on observational studies that cannot determine causality. In the absence of robust RCTs, the associations between hypotension and AKI, myocardial injury, and mortality nonetheless satisfy at least 4 of the Bradford Hill criteria for causality, common in epidemiology: consistency, temporality, dose-response relationship, and biological plausibility of the association. While it is not ethical to randomly induce hypotension to assess its causal effects, it is possible to randomize treatment protocols that may result in different amounts of hypotension and lead to different outcomes. A recent small trial compared individualized perioperative blood pressure management versus standard blood pressure management in high-risk patients having abdominal surgery.²⁵ This study found that individualized blood pressure management using norepinephrine to raise blood pressure reduced the incidence of a composite outcome consisting of systemic inflammatory response syndrome and at least 1 organ system dysfunction, compared to the standard care. The average difference in blood pressure between the trial groups was only 6 mm Hg, and amount of hypotension (eg, minutes below a mean pressure of 65 mm Hg) have yet to be reported. Large, robust trials are needed

to demonstrate that avoiding hypotension actually improves perioperative patient outcomes.

Although our data set includes many surgeries and clinicians (over 225,000 surgeries and 922 clinicians), it is not nationally representative. The risk adjustment model was developed and validated using data from 4 hospitals and further validated using data from 1 additional hospital; all were teaching hospitals. It will be important to update the risk adjustment model and its coefficients when more data are available from a variety of clinical settings, including nonacademic community hospitals.

In summary, we developed and evaluated a risk-adjusted quality measure for intraoperative hypotension that assesses whether clinicians have more or fewer cases of hypotension than expected over a 12-month period, given their patient mix. Hypotension, defined by at least 15 minutes of MAP <65 mm Hg, was common during noncardiac surgical cases and was associated with AKI and in-hospital mortality. There were substantial variations in clinician-level scores, and the measure score distribution suggests that there may be opportunity to reduce hypotension which may improve patient safety and outcomes. The measure is predicated on a causal relationship between hypotension and adverse outcomes; we therefore caution that there is currently limited evidence for a causal relationship between intraoperative hypotension and organ injury. The measure is also predicated on accurately predicting amounts of hypotension over a 12-month period. While the model calibration statistics were acceptable, sensitivity analyses in 1 hospital showed that score distributions differed by type of provider and the fraction of cases that were inpatient. With improved data capture in electronic anesthesia records, future versions of the measure should risk adjust for other important patient and procedural factors. ■■

ACKNOWLEDGMENTS

The authors acknowledge the clinicians and patients who provided input into the measure development process via expert work group, measured entities panel, and patient and family advisory panel. Cindy Cullen, Noa Sager, Bob Dickerson, Anita Somplasky, and Jayanti Bandyopadhyay at Mathematica supported various aspects of this work.

DISCLOSURES

Name: Anna L. Christensen, PhD.

Contribution: This author helped develop the measure specifications, conduct the site visits, design the analysis, oversee the analysis, interpret the results, and draft the manuscript with input and final approval from all authors.

Conflicts of Interest: None.

Name: Ethan Jacobs, MPP.

Contribution: This author helped develop the measure specifications, conduct the site visits, lead the expert panels and patient panels, interpret the results, and draft the manuscript with input and final approval from all authors.

Conflicts of Interest: None.

Name: Kamal Maheshwari, MD, MPH.

Contribution: This author helped conceive the measure and inform the specifications, serve on expert panel, supply the data, interpret the results, and review the manuscript.

Conflicts of Interest: K. Maheshwari is a consultant for Edwards Lifesciences.

Name: Fei Xing, PhD.

Contribution: This author helped process the data, design the analysis, perform the analysis, interpret the results, and review the manuscript.

Conflicts of Interest: None.

Name: Xiaohong Zhao, PhD.

Contribution: This author helped process the data, design and perform the analysis, interpret the results, and review the manuscript.

Conflicts of Interest: None.

Name: Samuel E. Simon, PhD.

Contribution: This author helped develop the measure specifications, design the analysis, interpret the results, and review the manuscript.

Conflicts of Interest: None.

Name: Karen B. Domino, MD, MPH.

Contribution: This author helped host site visit, serve on expert panel, supply the data, interpret the results, and review and revise the manuscript.

Conflicts of Interest: K. B. Domino received research grant from the University of Washington funded by Edwards Lifesciences.

Name: Karen L. Posner, PhD.

Contribution: This author helped supply the data, interpret the results, and review and revise the manuscript.

Conflicts of Interest: K. L. Posner is a coinvestigator received grants from the Microsoft Corporation to the University of Washington related to predicting perioperative hypotension.

Name: Alvin F. Stewart, MD.

Contribution: This author helped host site visit, supply the data, interpret the results, and review the manuscript.

Conflicts of Interest: None.

Name: Joseph A. Sanford, MD.

Contribution: This author helped host site visit, supply the data, interpret the results, and review the manuscript.

Conflicts of Interest: J. A. Sanford has an equity position in SJ MedConnect Inc, a GME interview management, and scheduling services company with no relevant connection to this research.

Name: Daniel I. Sessler, MD.

Contribution: This author helped provide expert input on the evidence base, measure specifications, and analysis; interpret the data; and review and revise the manuscript.

Conflicts of Interest: D. I. Sessler is a consultant for Edwards Lifesciences and Sensifree.

This manuscript was handled by: Richard C. Prielipp, MD.

REFERENCES

- Wesselink EM, Kappen TH, Torn HM, Slooter AJC, van Klei WA. Intraoperative hypotension and the risk of postoperative adverse outcomes: a systematic review. *Br J Anaesth*. 2018;121:706–721.
- Alcock RF, Kouzios D, Naoum C, Hillis GS, Brieger DB. Perioperative myocardial necrosis in patients at high cardiovascular risk undergoing elective non-cardiac surgery. *Heart*. 2012;98:792–798.
- Bijker JB, van Klei WA, Vergouwe Y, et al. Intraoperative hypotension and 1-year mortality after noncardiac surgery. *Anesthesiology*. 2009;111:1217–1226.
- Hallqvist L, Mårtensson J, Granath F, Sahlén A, Bell M. Intraoperative hypotension is associated with myocardial damage in noncardiac surgery: an observational study. *Eur J Anaesthesiol*. 2016;33:450–456.
- Kertai MD, White WD, Gan TJ. Cumulative duration of “triple low” state of low blood pressure, low bispectral index, and low minimum alveolar concentration of volatile anesthesia is not associated with increased mortality. *Anesthesiology*. 2014;121:18–28.
- Mascha EJ, Yang D, Weiss S, et al. Intraoperative mean arterial pressure variability and 30-day mortality in patients having noncardiac surgery. *Anesthesiology*. 2015;123:79–91.
- Monk TG, Saini V, Weldon BC, Sigl JC. Anesthetic management and one-year mortality after noncardiac surgery. *Anesth Analg*. 2005;100:4–10.
- Monk TG, Bronsert MR, Henderson WG, et al. Association between intraoperative hypotension and hypertension and 30-day postoperative mortality in noncardiac surgery. *Anesthesiology*. 2015;123:207–319.
- Ning YJ, Luo Z, Guo DQ, et al. Evaluation of acute kidney injury as defined by the risk, injury, failure, loss, and end-stage criteria in critically ill patients undergoing abdominal aortic aneurysm repair. *Chin Med J*. 2013;126:431–436.
- Oh NH, Kim KJ. Outcomes and risk factors affecting mortality in patients who underwent colorectal emergency surgery. *Ann Coloproctol*. 2016;32:133–138.
- Salmasi V, Maheshwari K, Yang D, et al. Relationship between intraoperative hypotension, defined by either reduction from baseline or absolute thresholds, and acute kidney and myocardial injury after noncardiac surgery: a retrospective cohort analysis. *Anesthesiology*. 2017;126:47–65.
- Sun LY, Wijeyesunderan DN, Tait GA, et al. Association of intraoperative hypotension with acute kidney injury after elective noncardiac surgery. *Anesthesiology*. 2015;123:515–523.
- Taffé P, Sicard N, Pittet V, et al. The occurrence of intraoperative hypotension varies between hospitals: observational analysis of more than 147,000 anaesthesia. *Acta Anaesthesiol Scand*. 2009;53:995–1005.
- Tallgren M, Niemi T, Pöyhiä R, et al. Acute renal injury and dysfunction following elective abdominal aortic surgery. *Eur J Vasc Endovasc Surg*. 2007;33:550–555.
- van Waes JA, van Klei WA, Wijeyesundera DN, van Wolfswinkel L, Lindsay TF, Beattie WS. Association between intraoperative hypotension and myocardial injury after vascular surgery. *Anesthesiology*. 2016;124:35–44.
- Walsh M, Devereaux PJ, Garg AX, et al. Relationship between intraoperative mean arterial pressure and clinical outcomes after noncardiac surgery: toward an empirical definition of hypotension. *Anesthesiology*. 2013;119:507–515.
- An R, Pang QY, Liu HL. Association of intra-operative hypotension with acute kidney injury, myocardial injury and mortality in non-cardiac surgery: a meta-analysis. *Int J Clin Pract*. 2019;73:e13394.
- Bijker JB, van Klei WA, Kappen TH, van Wolfswinkel L, Moons KG, Kalkman CJ. Incidence of intraoperative hypotension as a function of the chosen definition: literature definitions applied to a retrospective cohort using automated data collection. *Anesthesiology*. 2007;107:213–220.
- Multicenter Perioperative Outcomes Group (MPOG). ASPIRE Measures. 2020. Available at: <https://spec.mpopg.org/Measures/Public>. Accessed June 5, 2020.
- American Society of Anesthesiologists (ASA). ASA Physical Status Classification System. October 2019. Available at: <https://www.asahq.org/standards-and-guidelines/asa-physical-status-classification-system>. Accessed June 5, 2020.
- Sankar A, Johnson SR, Beattie WS, Tait G, Wijeyesundera DN. Reliability of the American Society of Anesthesiologists

- physical status scale in clinical practice. *Br J Anaesth*. 2014;113:424–432.
22. Adams JL. *The Reliability of Provider Profiling: A Tutorial. TR-653-National Committee for Quality Assurance*. RAND Corporation, 2009. Available at https://www.rand.org/content/dam/rand/pubs/technical_reports/2009/RAND_TR653.pdf. Accessed March 10, 2018.
 23. Morris CN. Parametric empirical Bayes inference: theory and applications. *J Am Stat Assoc*. 1983;78:47–55.
 24. Brady KM, Hudson A, Hood R, DeCaria B, Lewis C, Hogue CW. Personalizing the definition of hypotension to protect the brain. *Anesthesiology*. 2020;132:170–179.
 25. Futier E, Lefrant JY, Guinot PG, et al; INPRESS Study Group. Effect of individualized vs standard blood pressure management strategies on postoperative organ dysfunction among high-risk patients undergoing major surgery: a randomized clinical trial. *JAMA*. 2017;318:1346–1357.