Does the Surgical Apgar Score Measure Intraoperative Performance?

Scott E. Regenbogen, MD, MPH[1,2], **R. Todd Lancaster, MD**[1,2], **Stuart R. Lipsitz, ScD**[3], **Caprice C. Greenberg, MD, MPH**[3], **Matthew M. Hutter, MD, MPH**[2], and **Atul A. Gawande, MD, MPH**[1,3]

1 *Department of Health Policy and Management, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115, USA*

2 *Department of Surgery, Massachusetts General Hospital, 55 Fruit Street, Boston, Massachusetts 02114, USA*

3 *Center for Surgery and Public Health, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115, USA*

Abstract

Objective—To evaluate whether Surgical Apgar Scores measure the relationship between intraoperative care and surgical outcomes.

Summary Background Data—With preoperative risk-adjustment now well-developed, the role of intraoperative performance in surgical outcomes may be considered. We previously derived and validated a ten-point Surgical Apgar Score—based on intraoperative blood loss, heart rate, and blood pressure—that effectively predicts major postoperative complications within 30 days of general and vascular surgery. This study evaluates whether the predictive value of this score comes solely from patients' preoperative risk, or also measures care in the operating room.

Methods—Among a systematic sample of 4,119 general and vascular surgery patients at a major academic hospital, we constructed a detailed risk-prediction model including 27 patient-comorbidity and procedure-complexity variables, and computed patients' propensity to suffer a major postoperative complication. We evaluated the prognostic value of patients' Surgical Apgar Scores before and after adjustment for this preoperative risk.

Results—After risk-adjustment, the Surgical Apgar Score remained strongly correlated with postoperative outcomes (p<0.0001). Odds of major complications among average-scoring patients (scores 7–8) were equivalent to preoperative predictions (likelihood ratio (LR) 1.05, 95%CI 0.78–1.41), significantly decreased for those who achieved the best scores of 9–10 (LR 0.52, 95%CI 0.35–0.78), and were significantly poorer for those with low scores—LRs 1.60 (1.12–2.28) for scores 5–6, and 2.80 (1.50–5.21) for scores 0–4.

Conclusions—Even after accounting for fixed preoperative risk—due to patients' acute condition, comorbidities and/or operative complexity—the Surgical Apgar Score appears to detect differences in intraoperative management that reduce odds of major complications by half, or increase them by nearly three-fold.

Correspondence to: Scott E. Regenbogen, MD, MPH, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, Email: sregenbogen@partners.org; Phone: 617-423-6137, Fax: 617-432-4494.

Introduction

Hospitals and surgical teams strive to provide a consistently low occurrence of major complications for patients undergoing any given operation. Marked variability in outcomes is inevitable, if only because of differences in patients' preoperative risks. However, the degree to which intraoperative performance further contributes to variation in patients' risk of complications remains unclear.[1]

Prevailing techniques of surgical quality assessment, such as the American College of Surgeons' National Surgical Quality Improvement Program (NSQIP),[2–4] evaluate surgical performance indirectly, using multivariable adjustment for preoperative risk, and attributing disparities between observed and expected complication rates to the care provided. In the operating room, surgeons have relied principally on "gut-feeling" clinical assessments of the operative course to inform postoperative prognostication, and guide clinical care.[5] Most believe that intraoperative management contributes importantly to overall outcomes, but quantitative metrics of operative care have not been available.[1] Among intraoperative factors, alterations of patient condition, including hypotension,[6–23] hypertension,[12, 15–18, 22, 24] hypothermia,[25–27] bradycardia,[20, 22] tachycardia,[11, 12, 20, 22, 24, 28–30] and blood loss[31–35] have been independently linked with adverse outcomes. And some risk prediction methods have integrated intraoperative variables,[32, 36–38] yet no consensus has been reached on how to directly evaluate performance and safety in the operating room.[39]

To provide surgeons with a simple, objective, and direct rating, we previously developed and validated a ten-point Surgical Apgar Score.[40] In deriving the score, we screened more than two dozen parameters collected in the operating room, and found that just three intraoperative variables remained independently predictive of major postoperative complications and death —the lowest heart rate, lowest mean arterial pressure, and estimated blood loss. A score built from these three predictors has proved strongly predictive of the risk of major postoperative complications and death in general and vascular surgery.[40] Yet, it remains simple enough for teams to collect immediately upon completion of an operation for patients in any setting, regardless of resource and technological capacity.

Like the obstetrical Apgar score,[41–44] however, it provides a measure only of the relative success of care. It cannot by itself assess the quality of care, as its three variables are influenced not only by the performance of medical teams, but also by the patients' prior condition and the magnitude of the operations they undergo.[18, 22, 45] For the score to be a clinically useful predictor of postoperative complications, it should inform operative teams about their contribution to surgical outcomes, even after accounting for fixed preoperative risk—an insight not previously available. In this study, we therefore evaluated the predictive ability of the score after application of a validated risk-adjustment method, incorporating both patient- and procedure-related risk characteristics.

Methods

Patient cohort

The Massachusetts General Hospital (MGH) Department of Surgery maintains an outcomes database on a systematic sample of patients undergoing general and vascular surgical procedures, for submission to the NSQIP. In this program,[2, 3] trained nurse-reviewers retrospectively collect 49 preoperative, 17 intraoperative, and 33 outcome variables on surgical patients, for the monitoring of risk-adjusted outcomes. Patients undergoing general or vascular surgery with general, epidural, or spinal anesthesia, or specified operations (carotid endarterectomy, inguinal herniorrhaphy, thyroidectomy, parathyroidectomy, breast biopsy, and endovascular repair of abdominal aortic aneurysm) regardless of anesthetic type, are

eligible for inclusion. Children under age 16 and patients undergoing trauma surgery, transplant surgery, vascular access surgery, or endoscopic-only procedures are excluded. At MGH, at least forty consecutive operations meeting inclusion criteria in each eight-day cycle are enrolled. No more than five inguinal herniorraphies and five breast biopsies are enrolled per eight-day cycle to ensure diversity of operations in the case mix.

We evaluated all patients in the MGH-NSQIP database who underwent surgery between July 1, 2003, and June 30, 2005, and for whom complete 30-day follow-up was obtained. We excluded (i) carotid endarterectomies performed concurrently with coronary artery bypass grafting, because the score was not designed for application to patients on cardiopulmonary bypass; and (ii) operations performed with local anesthesia only, because no electronic anesthesia record is generated for these procedures.

The study protocol, including a waiver of informed consent from individual patients, was approved by the Human Subjects Research Committees of Massachusetts General Hospital and the Harvard School of Public Health.

Preoperative risk factors and postoperative outcomes

We collected all preoperative patient variables from the NSQIP database. All variables were either treated as dichotomous or categorized according the FY2005 NSQIP models.[46] Missing laboratory values were imputed with the overall sample median (because patients for whom preoperative laboratory data were not obtained were typically low-risk). Procedural work Relative Value Units were calculated by linkage of Current Procedural Terminology codes with listings from the 2005 Medicare Physician Fee Schedule (Centers for Medicare and Medicaid Services).

The primary endpoint was the occurrence of any major complication within 30 days after surgery, as recorded in the NSQIP database. The following NSQIP-defined[3] events were considered major complications: acute renal failure, bleeding requiring ≥4 units of red cell transfusion within 72 hours after surgery, cardiac arrest requiring CPR, coma for ≥24 hours, deep venous thrombosis, myocardial infarction, unplanned intubation, ventilator use for ≥48 hours, pneumonia, pulmonary embolism, stroke, wound disruption, deep or organ-space surgical site infection, sepsis, septic shock, systemic inflammatory response syndrome (SIRS), and vascular graft failure. All deaths were assumed to include a major complication. Superficial surgical site infection and urinary tract infection were not considered major complications. Patients having complications categorized in the database as "other occurrence" were reviewed individually and severity of the occurrence was evaluated according to the Clavien classification.[47] "Other occurrences" involving complications of Clavien Class III and greater (those that require surgical, endoscopic or radiologic intervention or intensive care admission, or are life-threatening) were considered major complications, in accordance with our previous methods.[40]

Preoperative risk stratification

To estimate each patient's preoperative likelihood of complications, we performed multivariable logistic regression using the variables included in the FY2005 NSQIP morbidity risk-adjustment model[46] as predictors, and the occurrence of any major complications as the outcome. We derived *de novo* regression coefficients from our dataset and computed the predicted likelihood of major complication from these regression parameters for each operation. These preoperative likelihoods were then stratified by quintiles for tabulation.[48]

Calculation of the intraoperative score

As described previously,[40] we originally devised the Surgical Apgar Score by using multivariable logistic regression to screen a collection of intraoperative measures. We found that only three intraoperative parameters remained independent predictors of 30-day major complications: the estimated blood loss (EBL), the lowest heart rate (HR), and the lowest mean arterial pressure (MAP) during the operation. The score was thus developed using these three variables, and their beta coefficients were used to weight the points allocated to each variable in a ten-point score (Table 1).

In this study, we extracted intraoperative hemodynamic data from the electronic Anesthesia Information Management System (Saturn, Dräger Medical, Telford, PA) database, using a Structured Query Language algorithm to filter out artifactual readings, using criteria developed through comparisons of electronic and hand-written intraoperative records.[49] For data quality assurance, we manually reviewed the printed electronic anesthesia record for 50 operations, and compared the results with those of the electronic data acquisition algorithm for these cases. The parameter values, as well as the total Score obtained, by each method were compared by computing kappa statistics for agreement, using Fleiss-Cohen weighting for ordered categorical data.[50]

Statistical analysis

All analyses were performed using the SAS 9.1 statistical software package (SAS Institute, Cary, N.C., 2003). We evaluated relationships between patient and procedure characteristics and levels of the Surgical Apgar Score using Spearman correlation coefficients for continuous variables, and Cochran-Armitage chi-square trend tests[51] for categorical predictors. Preoperative risk stratification was compared with the Surgical Apgar Score using Spearman correlation coefficients and kappa statistics, with Fleiss-Cohen weighting.[50] We calculated c-statistics for model discrimination (equivalent to the area under the ROC curve)[52] and compared models with the Hanley-McNeil z-test.[53] We used the Hosmer-Lemeshow goodness-of-fit test[54] to assess calibration.

Within quintiles of preoperative risk, the relationship between Surgical Apgar Score groups and postoperative occurrences was evaluated with the Cochran-Armitage chi-square trend test. [51] Controlling for preoperative risk predictions as a linear variable, we performed logistic regression with the Surgical Apgar Score as a categorical predictor and the incidence of major complications as the outcome, to compute adjusted effect sizes for each level of the Surgical Apgar Score. Adjusted likelihood ratios (LRs) were computed as the proportional change in odds of complications, comparing prior odds (preoperative prediction) with postoperative odds. We obtained confidence intervals for the LRs using the Bonferroni inequality.[55]

To evaluate the robustness of the risk-adjusted relationship between Surgical Apgar Scores and postoperative complications, we conducted sensitivity analyses, modeling the preoperative risk-adjustment parameters in a variety of ways. For simplicity, the primary analyses we present are stratification by quintiles[48] and logistic regression with the preoperative predictions treated linearly.[56]

Results

Data accrual and validation

Of 4,163 NSQIP cases that met inclusion criteria, 4,119 (98.9%) had complete electronic intraoperative records and comprised our final cohort. The automated data extraction algorithm achieved excellent agreement with manual record review, both for point values assigned to each variable ($\kappa = 0.97$ for HR; $\kappa = 0.75$ for MAP), and for the total Score ($\kappa = 0.94$).

Baseline patient and procedure characteristics

In univariate analyses, most demographic characteristics and preoperative risk factors varied significantly between levels of the Surgical Apgar Score (see Table 2). With decreasing scores, patients were increasingly older (p=0.009), and more likely to be male (p=0.0004) and of non-white race (p=0.04). Patients assigned higher American Society of Anesthesiologists' (ASA) Physical Status Classification had significantly lower Surgical Apgar Scores (Spearman's r = −0.24, p<0.0001). Two-thirds (85 of 128) of patients with scores ≤4 were ASA Class 3 or 4, whereas three quarters (1091 of 1441) of patients with scores of 9 or 10 were ASA Class 1 or 2 (p<0.0001).

Low-scoring patients were significantly more likely to be underweight (p=0.01), but not more likely to be obese (p=0.12). Among the 26 other preoperative comorbidity conditions, 22 of them were increasingly prevalent as patients' scores decreased. Only hypertension (p=0.06), coma (p=0.89), Do Not Resuscitate status (p=0.15), and alcohol use (p=0.53) were not significantly correlated with Surgical Apgar Scores. Abnormalities in all 12 preoperative laboratory measures were also increasingly more common as patients' scores decreased (all p<0.01). Operations with lower scores had increasing complexity (as measured by Work RVUs) and were more likely to be emergencies (both p<0.0001).

Surgical Apgar Scores were also predictive of postoperative outcomes. The incidence of major postoperative complications increased monotonically from 5% among patients with scores of 9–10, to 56% of those with scores ≤4 (p<0.0001). Patients with low scores were more likely to suffer multiple complications (p<0.0001), and had significantly longer median length of stay (p<0.0001). Among patients who experienced a complication, the likelihood of dying from that complication was nearly 20-fold greater for patients with scores 0–2 than for those with scores of 9–10 (p<0.0001).

Preoperative risk-adjustment

Logistic regression, using the 27 preoperative NSQIP variables[46] as predictors and the incidence of major postoperative complications as the outcome, generated a multivariable preoperative risk prediction model with a c-index of 0.820 (equivalent to that of the 34,000 patient FY2005 NSQIP cohort; p=0.23).[46] The Hosmer-Lemeshow chi-square statistic demonstrated adequate model calibration (p=0.49).

Forty percent of patients were missing at least one of the laboratory measures required for the model, ranging from 3.5% missing white blood cell count to 37% missing albumin. In sensitivity analyses, results from imputation with the sample median were not meaningfully different from those of multiple imputation, so median imputation was used for simplicity.

Patients were stratified into preoperative risk quintiles, based on their predicted likelihoods of major complication according to this model. Quintile 1 included patients with preoperative risk ≤3.8%; Quintile 2, 3.8–6.5%; Quintile 3, 6.5–10.6%; Quintile 4, 10.6–19.2%, Quintile 5, ≥19.2%. In logistic regression, discrimination by quintiles (c=0.795) was not significantly different from that of the saturated risk prediction model (p=0.12).

Risk-adjusted analysis of the Surgical Apgar Score

Patients' preoperative risk predictions and Surgical Apgar Scores were negatively correlated (r = −0.42, p<0.0001), confirming that the elements of the score are associated with preoperative risk factors. Accordingly, there was fair agreement between a patient's preoperative risk quintile and level of the score, with a weighted kappa of 0.24 (95% confidence interval 0.22–0.26).

The stratified data in Table 3 demonstrate this relationship. Of the 128 patients with scores ≤4, 112 (88%) came from the highest two risk quintiles—patients with preoperative likelihood of complication greater than 10.6%. In contrast, patients from the two lowest risk quintiles—with estimated preoperative risks less than 6.5%—comprised 59% of patients with scores of 9–10.

Despite concordance between preoperative factors and intraoperative metrics, after accounting for preoperative risk, the Surgical Apgar Score remained a significant predictor of postoperative complications. Within each quintile, patients with scores of 7–8 experienced complication rates similar to the expected mean rate for their stratum. Patients with scores of 9–10 had consistently lower incidence, and patients with scores <7 had consistently greater incidence of major postoperative complications than was expected preoperatively. Among the three highest risk strata, Surgical Apgar Scores remained significantly predictive of postoperative outcomes (each p<0.001). In the two lowest quintiles, however, we had limited power to detect a significant effect, because of the rarity of both complications and low scores among these low-risk patients. Nevertheless, a trend toward significant relationship was found (both p<0.10).

Adjusted and unadjusted likelihood ratios (LRs), representing the proportional change in odds of a major complication comparing preoperative expectations with postoperative predictions, are shown in Table 4. Unadjusted LRs compare posterior odds of a complication for patients within each score category against the average preoperative odds for the entire cohort. The adjusted LRs estimate the degree to which the Surgical Apgar Score alters any given patient's odds of complication, after accounting for the patient's fixed preoperative odds as measured by NSQIP risk factors.

Patients with scores of 7–8 had postoperative outcomes no different from preoperative predictions (LR 1.05, 95% confidence interval 0.78–1.41). Those with scores of 9–10, however, had significantly lower odds of complication than would be expected based on preoperative risk (LR 0.52, 95% confidence interval 0.35–0.78), and those with scores less ≤6 had significantly increased posterior odds, with LR of 1.60 (95% confidence interval 1.12–2.28) for scores 5–6, 2.79 (1.47–5.31) for scores 3–4. Confidence intervals around the LR for scores 0–2 (2.87, 95% confidence interval 0.49–16.76) crossed one due to small sample size.

Sensitivity analyses

We tested eight additional methods of modeling patients' preoperative risk: (i) adding the Surgical Apgar Score to the 27-variable NSQIP morbidity model; (ii) modeling predictions as a linear plus quadratic and/or (iii) cubic term; (iv) stratifying by deciles; (v) stratifying by quintiles or (vi) deciles of patients with complications only; (vii) one-to-one matching[57] of patients with and without complications by their preoperative risk score; and (viii) construction of our own risk-adjustment model, using all available patient-comorbidity and procedure-complexity variables available to us, from either the NSQIP record, or other data collection, and including all significant interaction terms. Regardless of the risk-adjustment technique, the Surgical Apgar Score remained a strong predictor of postoperative outcomes (p values all <0.0001), and all point estimates for the odds ratios at each score level remained within the 95% confidence bounds of the primary analysis.

Conclusion

We find that even after detailed adjustment for comorbidity and procedure-specific risk factors, the amount of blood loss, lowest heart rate and lowest blood pressure were still important predictors of the risk of a major complication. The Surgical Apgar Score, therefore, conveyed useful prognostic information, either in isolation or in combination with assessments of the risks that patients brought to the operating room. It also may provide an immediate assessment

of how well or poorly the operation has gone for a patient. In this cohort, surgical teams could cut a patient's risk-adjusted odds of major complications nearly in half with a score of 9–10, or conversely, nearly triple the risk-adjusted odds with scores ≤4.

This finding, that intraoperative blood loss, heart rate, and blood pressure are critical predictors of postoperative risk, is consistent with a variety of previous observations. Hemodynamic stability[6–23, 29, 30] and the amount of blood loss[31–35] during surgery have long been recognized as important independent factors in patient outcomes. What had not been recognized were the collective importance of these variables, and their potential contribution to an easily-implemented intraoperative performance metric.[40]

As an adjunct to surgeons' subjective impressions of the operation,[5] the score may thus aid decision-making about unplanned admission after outpatient surgery, admission to the intensive care unit, or frequency of postoperative examinations by physicians and nurses, with the goal of preventing poor outcomes among low-scoring patients. More broadly, the Surgical Apgar Score provides a novel metric for evaluating the efficacy of safety interventions in the operating room—a much-needed tool for surgical safety initiatives, because more than two-thirds of surgical adverse events involve complications in the operating room.[58–60]

In this study, we use likelihood ratios to quantify the degree to which a patient's Surgical Apgar Score alters his or her likelihood of complications. The likelihood ratio describes the proportional change in odds of complication, comparing information available before the operation with what is available after the operation. The unadjusted likelihood ratios associated with levels of the score give a measure of the change in postoperative odds of complications, compared with the baseline rate of complications for the sample as a whole. This type of comparison could be used as part of a broad-based public health audit for surgical safety. Targeting low scores would allow surgeons and administrators to focus on patients coming out of surgery who are at highest risk of major complications or death. Routine surveillance and case-review for patients with low surgical scores (e.g., a score of 4 or less), even when no complications result, may enable early identification of latent safety problems. The score could also provide a target for surgical teams and researchers aiming to improve outcomes, and a measure for quality monitoring and improvement programs, even in resource-poor settings. The ultimate goal would be to encourage development and implementation of practices that reduce the proportion of patients with low scores and increase the proportion with the highest scores.

The risk-adjusted likelihood ratios provide different, yet complementary information, allowing individual surgeons to objectively discern whether, and how much, their operation increased or decreased a patient's predicted risk of major complications. As Table 4 indicates, an operation with a score of 7 or 8 has not altered the expected risk; one with a score of 9 or 10 has reduced it by half; one with a score of 5–6 has increased the odds of complications by approximately 60%; and one with a score of 4 or less has increased it by almost 200%.

Even in hospitals such as ours, where preoperative risk information is avidly collected for outcomes monitoring in the NSQIP, detailed preoperative risk predictions are rarely, if ever, available at the time of surgery. Missing data (especially laboratory data) and the computational complexity of multivariable prediction models preclude their routine use.[37, 45, 61] The majority of surgeons in our institution and elsewhere depend instead on their subjective impressions,[5, 62] and rate patients in broader categories, similar to the risk quintiles we analyzed. For these surgeons, the data in Table 3 provide a means for using the Surgical Apgar Score as additional prognostic information, beyond their preoperative expectations.

There are, however, other possible explanations for these findings. Because this was an observational study, it is possible that alterations in intraoperative hemodynamics simply

represent residual confounding—still unmeasured aspects of patient- or procedure-related risk. The sensitivity analyses argue against this being the only explanation, however. Relationships were similar across nine different modeling techniques, and effect sizes for the score were not attenuated by the inclusion of additional predictors to the model. Another explanation could be that the stress of surgery unmasks debility or risk not otherwise measured among preoperative variables. As seen in Table 2, the Surgical Apgar Score is closely associated with many of the same variables that predict postoperative complications. Still, the critical variables for the Surgical Apgar Score are measures that have been consistently recognized as important independent contributors to surgical morbidity.[15–17, 21–24, 30, 33, 35] It is reasonable to believe, therefore, that interventions that produce measurable improvement in Surgical Apgar Scores will also improve intraoperative safety and reduce postoperative complications.

Other important limitations remain. The score has not been evaluated beyond major academic medical centers, adult patients, or general and vascular surgery, due to a lack of reliable and comprehensive outcomes assessment against which the measures could be validated in these areas. Whether the score will be effective at grading risk in trauma, pediatric surgery or other surgical specialties remains uncertain.

In summary, we have found that a simple clinimetric surgical outcome score can provide both clinical surgeons and surgical safety researchers with useful and important information. The Surgical Apgar Score integrates components of patient susceptibility, procedure complexity and operative performance, providing a measure of immediate postoperative condition and prognostication, beyond standard risk-adjustment. As a decision-support tool, the score can inform postoperative prognostication, communication and triage, regardless of the sophistication of preoperative risk stratification available. And as a simple intraoperative outcome measure and safety improvement metric, it may prove useful as an indicator of surgical performance.

Acknowledgements

References

1. Vincent C, Moorthy K, Sarker SK, et al. Systems approaches to surgical quality and safety: from concept to measurement. Ann Surg 2004;239(4):475–82. [PubMed: 15024308]

2. Fink AS, Campbell DA Jr, Mentzer RM Jr, et al. The National Surgical Quality Improvement Program in non-veterans administration hospitals: initial demonstration of feasibility. Ann Surg 2002;236(3): 344–53. [PubMed: 12192321]discussion 353–4

3. Khuri SF, Daley J, Henderson W, et al. The National Veterans Administration Surgical Risk Study: risk adjustment for the comparative assessment of the quality of surgical care. J Am Coll Surg 1995;180 (5):519–31. [PubMed: 7749526]

4. Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185(4):315–27. [PubMed: 9328380]

5. Hartley MN, Sagar PM. The surgeon's 'gut feeling' as a predictor of post-operative outcome. Ann R Coll Surg Engl 1994;76(6 Suppl):277–8. [PubMed: 7598397]

6. Mauney FM Jr, Ebert PA, Sabiston DC Jr. Postoperative myocardial infarction: a study of predisposing factors, diagnosis and mortality in a high risk group of surgical patients. Ann Surg 1970;172(3):497–503. [PubMed: 4248189]

7. Plumlee JE, Boettner RB. Myocardial infarction during and following anesthesia and operation. South Med J 1972;65(7):886–9. [PubMed: 5038186]

8. Steen PA, Tinker JH, Tarhan S. Myocardial reinfarction after anesthesia and surgery. Jama 1978;239 (24):2566–70. [PubMed: 660789]

9. Goldman L, Caldera DL. Risks of general anesthesia and elective operation in the hypertensive patient. Anesthesiology 1979;50(4):285–92. [PubMed: 434530]

10. von Knorring J. Postoperative myocardial infarction: a prospective study in a risk group of surgical patients. Surgery 1981;90(1):55–60. [PubMed: 7245051]

11. Lieberman RW, Orkin FK, Jobes DR, et al. Hemodynamic predictors of myocardial ischemia during halothane anesthesia for coronary-artery revascularization. Anesthesiology 1983;59(1):36–41. [PubMed: 6859610]

12. Rao TL, Jacobs KH, El-Etr AA. Reinfarction following anesthesia in patients with myocardial infarction. Anesthesiology 1983;59(6):499–505. [PubMed: 6650905]

13. Kotter GS, Kotrly KJ, Kalbfleisch JH, et al. Myocardial ischemia during cardiovascular surgery as detected by an ST segment trend monitoring system. J Cardiothorac Anesth 1987;1(3):190–9. [PubMed: 2979094]

14. Knight AA, Hollenberg M, London MJ, et al. Perioperative myocardial ischemia: importance of the preoperative ischemic pattern. Anesthesiology 1988;68(5):681–8. [PubMed: 3259409]

15. Charlson ME, MacKenzie CR, Gold JP, et al. The preoperative and intraoperative hemodynamic predictors of postoperative myocardial infarction or ischemia in patients undergoing noncardiac surgery. Ann Surg 1989;210(5):637–48. [PubMed: 2530940]

16. Charlson ME, MacKenzie CR, Gold JP, et al. Postoperative renal dysfunction can be predicted. Surg Gynecol Obstet 1989;169(4):303–9. [PubMed: 2781448]

17. Charlson ME, MacKenzie CR, Gold JP, et al. Intraoperative blood pressure. What patterns identify patients at risk for postoperative complications? Ann Surg 1990;212(5):567–80. [PubMed: 2241312]

18. Charlson ME, MacKenzie CR, Gold JP, et al. Preoperative characteristics predicting intraoperative hypotension and hypertension among hypertensives and diabetics undergoing noncardiac surgery. Ann Surg 1990;212(1):66–81. [PubMed: 2363606]

19. Ashton CM, Petersen NJ, Wray NP, et al. The incidence of perioperative myocardial infarction in men undergoing noncardiac surgery. Ann Intern Med 1993;118(7):504–10. [PubMed: 8442621]

20. Reich DL, Bodian CA, Krol M, et al. Intraoperative hemodynamic predictors of mortality, stroke, and myocardial infarction after coronary artery bypass surgery. Anesth Analg 1999;89(4):814–22. [PubMed: 10512249]

21. Barone JE, Bull MB, Cussatti EH, et al. Perioperative myocardial infarction in low-risk patients undergoing noncardiac surgery is associated with intraoperative hypotension. J Intensive Care Med 2002;17(5):250–255.

22. Rohrig R, Junger A, Hartmann B, et al. The incidence and prediction of automatically detected intraoperative cardiovascular events in noncardiac surgery. Anesth Analg 2004;98(3):569–77. [PubMed: 14980900]table of contents

23. Monk TG, Saini V, Weldon BC, et al. Anesthetic management and one-year mortality after noncardiac surgery. Anesth Analg 2005;100(1):4–10. [PubMed: 15616043]

24. Reich DL, Bennett-Guerrero E, Bodian CA, et al. Intraoperative tachycardia and hypertension are independently associated with adverse outcome in noncardiac surgery of long duration. Anesth Analg 2002;95(2):273–7. [PubMed: 12145033]table of contents

25. Kurz A, Sessler DI, Lenhardt R. Perioperative normothermia to reduce the incidence of surgical-wound infection and shorten hospitalization. Study of Wound Infection and Temperature Group. N Engl J Med 1996;334(19):1209–15. [PubMed: 8606715]

26. Frank SM, Fleisher LA, Breslow MJ, et al. Perioperative maintenance of normothermia reduces the incidence of morbid cardiac events. A randomized clinical trial. Jama 1997;277(14):1127–34. [PubMed: 9087467]

27. Bock M, Muller J, Bach A, et al. Effects of preinduction and intraoperative warming during major laparotomy. Br J Anaesth 1998;80(2):159–63. [PubMed: 9602578]

28. Coriat P, Harari A, Daloz M, et al. Clinical predictors of intraoperative myocardial ischemia in patients with coronary artery disease undergoing non-cardiac surgery. Acta Anaesthesiol Scand 1982;26(4): 287–90. [PubMed: 7124301]

29. Fillinger MP, Surgenor SD, Hartman GS, et al. The association between heart rate and inhospital mortality after coronary artery bypass graft surgery. Anesth Analg 2002;95(6):1483–8. [PubMed: 12456405]table of contents

30. Hartmann B, Junger A, Rohrig R, et al. Intra-operative tachycardia and peri-operative outcome. Langenbecks Arch Surg 2003;388(4):255–60. [PubMed: 12920601]

31. Gatch WD, Little WD. Amount of blood lost during some of the more common operations. JAMA 1924;83:1075–1076.

32. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. Br J Surg 1991;78 (3):355–60. [PubMed: 2021856]

33. Wolters U, Wolf T, Stutzer H, et al. ASA classification and perioperative variables as predictors of postoperative outcome. Br J Anaesth 1996;77(2):217–22. [PubMed: 8881629]

34. Pasternak LR. Risk assessment in ambulatory surgery: challenges and new trends. Canadian Journal of Anesthesiology 2004;51(6):R1–R5.

35. Wolters U, Mannheim S, Wassmer G, et al. What is the value of available risk-scores in predicting postoperative complications after aorto-iliac surgery? A prospective non randomized study. J Cardiovasc Surg (Torino) 2006;47(2):177–85.

36. Pasternak LR. Preanesthesia evaluation of the surgical patient. ASA Refresher Course 1996;24:205–219.

37. Sutton R, Bann S, Brooks M, et al. The Surgical Risk Scale as an improved tool for risk-adjusted analysis in comparative surgical audit. Br J Surg 2002;89(6):763–8. [PubMed: 12027988]

38. Aust JB, Henderson W, Khuri S, et al. The impact of operative complexity on patient risk factors. Ann Surg 2005;241(6):1024–7. [PubMed: 15912052]discussion 1027–8

39. Greenberg CC, Roth EM, Sheridan TB, et al. Making the operating room of the future safer. Am Surg 2006;72(11):1102–8. [PubMed: 17120955]discussion 1126–48

40. Gawande AA, Kwaan MR, Regenbogen SE, et al. An Apgar Score for Surgery. J Am Coll Surg 2007;204(2):201–208. [PubMed: 17254923]

41. Apgar V. A proposal for a new method of evaluation of the newborn infant. Curr Res Anesth Analg 1953;32(4):260–7. [PubMed: 13083014]

42. Apgar V, Holaday DA, James LS, et al. Evaluation of the newborn infant; second report. J Am Med Assoc 1958;168(15):1985–8. [PubMed: 13598635]

43. Casey BM, McIntire DD, Leveno KJ. The continuing value of the Apgar score for the assessment of newborn infants. N Engl J Med 2001;344(7):467–71. [PubMed: 11172187]

44. Finster M, Wood M. The Apgar score has survived the test of time. Anesthesiology 2005;102(4):855–7. [PubMed: 15791116]

45. Bann SD, Sarin S. Comparative audit: the trouble with POSSUM. J R Soc Med 2001;94(12):632–4. [PubMed: 11733590]

46. ACS NSQIP Advisory Committee. Semi-annual report. 2006 [Accessed: November 8, 2006]. https://acsnsqip.org/main/resources_semi_annual_report.pdf

47. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240(2):205–13. [PubMed: 15273542]

48. Cochran WG. The effectiveness of adjustment by subclassification in removing bias in observational studies. Biometrics 1968;24(2):295–313. [PubMed: 5683871]

49. Reich DL, Wood RK Jr, Mattar R, et al. Arterial blood pressure and heart rate discrepancies between handwritten and computerized anesthesia records. Anesth Analg 2000;91(3):612–6. [PubMed: 10960387]

50. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1973;33:613–619.

51. Armitage P. Test for linear trends in proportions and frequencies. Biometrics 1955;11:375–386.

52. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143(1):29–36. [PubMed: 7063747]

53. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839–43. [PubMed: 6878708]

54. Hosmer, DW.; Lemeshow, S. Applied Logistic Regression. 2. New York: John Wiley and Sons; 2000.

55. Natarajan S, Lipsitz SR, Rimm E. A simple method of determining confidence intervals for population attributable risk from complex surveys. Stat Med 2007;26(17):3229–39. [PubMed: 17309113]

56. Brenner H, Blettner M. Controlling for continuous confounders in epidemiologic research. Epidemiology 1997;8(4):429–34. [PubMed: 9209859]

57. Rubin DB, Thomas N. Matching using estimated propensity scores: relating theory to practice. Biometrics 1996;52(1):249–64. [PubMed: 8934595]

58. Gawande AA, Thomas EJ, Zinner MJ, et al. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. Surgery 1999;126(1):66–75. [PubMed: 10418594]

59. Gawande AA, Zinner MJ, Studdert DM, et al. Analysis of errors reported by surgeons at three teaching hospitals. Surgery 2003;133(6):614–21. [PubMed: 12796727]

60. Rogers SO Jr, Gawande AA, Kwaan M, et al. Analysis of surgical errors in closed malpractice claims at 4 liability insurers. Surgery 2006;140(1):25–33. [PubMed: 16857439]

61. Jones HJ, de Cossart L. Risk scoring in surgical patients. Br J Surg 1999;86(2):149–57. [PubMed: 10100780]

62. Markus PM, Martell J, Leister I, et al. Predicting postoperative morbidity by clinical assessment. Br J Surg 2005;92(1):101–6. [PubMed: 15635697]

Table 1

The Ten-Point Surgical Apgar Score

The Surgical Apgar Score is calculated at the end of any general or vascular surgery operation, from the estimated blood loss, lowest mean arterial pressure and lowest heart rate entered in the anaesthesia record during the operation. The score is the sum of the points from each category.

	0 points	1 point	2 points	3 points	4 points
Estimated blood loss (mL)	>1000	601–1000	101–600	≤100	
Lowest mean arterial pressure (mm Hg)	<40*	40–54	55–69	>70	
Lowest heart rate (beats per min)	>85*	76–85	66–75	56–65	≤55*

*
Occurrence of pathologic bradyarrhythmia, including sinus arrest, atrioventricular block or dissociation, junctional or ventricular escape rhythms, and asystole also receive 0 pts for lowest heart rate

Table 2

Characteristics of patients, procedures, and outcomes, by Surgical Apgar Score

Score	0–2 N=16	3–4 N=112	5–6 N=720	7–8 N=1830	9–10 N=1441	Total N=4119	p value*
Patient Characteristics							
Age (years) (mean ± sd)	62 ± 13	60 ± 18	59 ± 17	57 ± 17	58 ± 16	58 ± 17	0.01
Male sex	11 (69)	64 (57)	343 (48)	692 (38)	599 (38)	1710 (42)	0.0004
Non-white race	2 (13)	12 (11)	69 (10)	171 (9)	108 (7)	362 (9)	0.04
ASA Class							<0.0001
1	0 (0)	1 (0.9)	27 (4)	181 (10)	216 (15)	425 (10)	
2	3 (19)	39 (35)	334 (46)	1053 (58)	875 (61)	2304 (56)	
3	7 (44)	50 (45)	304 (42)	558 (30)	334 (23)	1253 (30)	
4	6 (38)	22 (20)	55 (8)	38 (2)	16 (1)	137 (3)	
Body mass index (mean ± sd)	27 ± 12	27 ± 9	29 ± 10	29 ± 10	28 ± 9	29 ± 10	0.79
Obese (BMI>35)	4 (25)	11 (10)	122 (17)	350 (19)	201 (14)	688 (17)	0.12
Underweight (BMI<18.5)	2 (13)	8 (7)	30 (4)	60 (3)	43 (3)	143 (3)	0.01
Cardiac disease (MI, CHF, angina, coronary revascularization)	3 (19)	25 (22)	126 (18)	229 (13)	173 (12)	556 (14)	<0.0001
Hypertension	8 (50)	52 (46)	328 (46)	744 (41)	597 (41)	1729 (42)	0.06
Pneumonia	1 (6)	10 (9)	22 (3)	25 (1)	3 (0.2)	61 (1)	<0.0001
COPD	1 (6)	13 (13)	42 (6)	66 (4)	39 (3)	162 (4)	<0.0001
Ventilator dependence	7 (44)	15 (13)	31 (4)	22 (1)	2 (0.1)	77 (2)	<0.0001
Dyspnea (at rest or with exertion)	2 (13)	17 (15)	78 (11)	141 (8)	77 (5)	315 (8)	<0.0001
Diabetes mellitus	3 (19)	19 (17)	109 (15)	251 (14)	111 (8)	493 (12)	<0.0001
Renal failure	2 (13)	14 (13)	27 (4)	40 (2)	10 (0.7)	93 (2)	<0.0001
Sepsis	9 (56)	24 (21)	74 (10)	54 (3)	11 (0.8)	172 (4)	<0.0001
Open wound	5 (31)	19 (17)	79 (11)	104 (6)	48 (3)	255 (6)	<0.0001
Bleeding disorder	8 (50)	22 (20)	44 (6)	55 (3)	29 (2)	153 (4)	<0.0001
History of stroke or TIA	2 (13)	3 (3)	37 (5)	120 (7)	110 (8)	272 (7)	0.02
Current smoker	5 (31)	30 (27)	138 (19)	325 (18)	222 (15)	720 (17)	0.0004
Disseminated cancer	3 (19)	15 (13)	59 (8)	78 (4)	38 (3)	193 (5)	<0.0001
Weight loss >10% in 6 months	3 (19)	19 (17)	80 (11)	140 (8)	62 (4)	304 (7)	<0.0001
Steroid use (oral or parenteral)	1 (6)	12 (11)	46 (6)	80 (4)	41 (3)	180 (4)	<0.0001
Ascites	1 (6)	13 (12)	29 (4)	29 (2)	10 (0.7)	82 (2)	<0.0001
Oesophageal varices	1 (6)	1 (0.9)	3 (0.4)	5 (0.3)	3 (0.2)	13 (0.3)	0.03
Rest pain or gangrene	2 (13)	5 (4)	38 (5)	56 (3)	13 (0.9)	114 (3)	<0.0001
Coma	0 (0)	0 (0)	0 (0)	1 (0.05)	0 (0)	1 (0.02)	0.89
Do Not Resuscitate Status	0 (0)	2 (2)	5 (0.7)	9 (0.5)	6 (0.4)	22 (0.5)	0.15
Alcohol use >2 drinks per day	1 (6)	5 (4)	36 (5)	67 (4)	61 (4)	170 (4)	0.53
Impaired Sensorium	3 (19)	9 (8)	26 (4)	13 (0.7)	5 (0.4)	56 (1)	<0.0001
Chemotherapy within 30 days	1 (6)	6 (5)	21 (3)	38 (2)	25 (2)	91 (2)	0.006
Radiation therapy within 90 days	1 (6)	1 (0.9)	19 (3)	33 (2)	11 (0.8)	65 (2)	0.002
Previous angioplasty, revascularization or amputation for peripheral vascular disease	2 (13)	15 (13)	82 (11)	126 (7)	70 (5)	295 (7)	<0.0001
Laboratory data							
White blood cell count >11,000/mm³	9 (56)	37 (33)	142 (20)	230 (13)	127 (9)	545 (13)	<0.0001
Hematocrit <38%	11 (69)	85 (66)	360 (50)	716 (39)	455 (32)	1616 (39)	<0.0001
Platelet count <150,000/mm³	4 (25)	21 (19)	60 (8)	102 (6)	62 (4)	249 (6)	<0.0001
Platelet count >400,000/mm³	2 (13)	19 (17)	66 (9)	143 (8)	57 (4)	287 (7)	<0.0001
Partial thromboplastin time >35 seconds	6 (38)	28 (25)	61 (8)	100 (5)	55 (4)	250 (6)	<0.0001
Prothrombin time >13.67 seconds	12 (75)	42 (38)	120 (17)	190 (10)	106 (7)	470 (11)	<0.0001
Sodium <135 mEq/L	5 (31)	22 (20)	70 (10)	96 (5)	42 (3)	235 (6)	<0.0001
Sodium >145 mEq/L	1 (6)	4 (4)	14 (2)	27 (1)	12 (0.8)	58 (1)	0.002
BUN >40 mg/dL	5 (31)	16 (14)	25 (4)	55 (3)	26 (2)	127 (3)	<0.0001
Creatinine >1.2 mg/dL	6 (38)	32 (29)	127 (18)	229 (13)	189 (13)	583 (14)	<0.0001

Score	0–2 N=16	3–4 N=112	5–6 N=720	7–8 N=1830	9–10 N=1441	Total N=4119	p value*
Albumin g/dL (mean ± sd)	2.7 ± 0.9	2.9 ± 0.9	3.4 ± 0.8	3.8 ± 0.7	4.0 ± 0.5	3.7 ± 0.7	<0.0001
SGOT >40 units/L	6 (38)	22 (20)	107 (15)	173 (9)	91 (6)	399 (10)	<0.0001
Bilirubin >1 g/dL	7 (44)	27 (24)	82 (11)	117 (6)	70 (5)	303 (7)	<0.0001
Alkaline phosphatase >125 units/L	9 (56)	26 (23)	114 (16)	164 (9)	79 (5)	392 (10)	<0.0001
Procedure Characteristics							
Emergency operation	10 (63)	29 (26)	89 (12)	128 (7)	57 (4)	313 (8)	<0.0001
Work Relative Value Units (mean ± sd)	22 ± 15	25 ± 12	24 ± 13	18 ± 10	14 ± 7	18 ± 11	<0.0001
Postoperative Outcomes							
Major complication(s)	12 (75)	60 (54)	201 (28)	236 (13)	72 (5)	581 (14)	<0.0001
Number of major complications							<0.0001
1	2 (13)	21 (19)	83 (12)	132 (7)	49 (3)	287 (7)	
2–3	8 (50)	20 (18)	75 (10)	76 (4)	18 (1)	197 (5)	
>3	2 (13)	19 (17)	43 (6)	28 (2)	5 (0.3)	97 (2)	
Deaths	7 (44)	18 (16)	33 (5)	34 (2)	2 (0.1)	94 (2)	<0.0001
Mortality rate among patients with major complications	58%	30%	16%	14%	3%	16%	<0.0001
Length of stay (days; median, IQR)	15 (5–25)	8 (6–17)	6 (3–9)	3 (1–5)	1 (0–2)	2 (1–5)	<0.0001

*
Hypothesis testing employed Cochrane-Armitage chi-square trend tests for categorical variables, and Spearman correlation coefficients for continuous variables, except for length of stay, for which we used the Kruskal-Wallis test.

Table 3

Postoperative outcomes, by preoperative risk stratum and Surgical Apgar Score

Using a detailed preoperative risk prediction model, patients were stratified by quintiles of likelihood of a major postoperative complication. The predicted and observed incidence of major postoperative complications are presented.

| Preoperative Risk Predictions | | | Surgical Apgar Score | | | | | | | | | |
| | | | 0–2 | | 3–4 | | 5–6 | | 7–8 | | 9–10 | | |
Quintile	Mean Risk	Range	N	% Major Complication	N	% Major Complication	N	% Major Complication	N	% Major Complication	N	% Major Complication	p value*
1	1.9%	0.5–3.8%	0	—	0	—	45	6.7%	338	1.5%	440	1.4%	0.09
2	5.2%	3.8–6.5%	0	—	2	0%	73	9.6%	346	4.0%	403	3.5%	0.07
3	8.4%	6.5–10.6%	0	—	14	28.6%	115	12.2%	408	9.3%	287	4.5%	0.0003
4	14.1%	10.6–19.2%	2	100%	19	31.6%	181	21.5%	410	15.6%	212	7.1%	<0.0001
5	41.0%	19.2–99.7%	14	71.4%	77	64.9%	306	45.1%	328	35.1%	99	24.2%	<0.0001
Total	14.1%	0.5–99.7%	16	75.0%	112	53.6%	720	27.9%	1830	12.9%	1441	5.0%	<0.0001

*
Cochrane-Armitage chi-square trend test

Table 4

Complication rates and adjusted and unadjusted likelihood ratios for levels of the Surgical Apgar Score

Model		SURGICAL APGAR SCORE										Total	
		0–2		3–4		5–6		7–8		9–10			
		N	% Major Complication	N	% Major Complication	N	% Major Complication	N	% Major Complication	N	% Major Complication	N	% Major Complication
Complication Rate		16	75.0%	112	53.6%	720	27.9%	1830	12.9%	1441	5.0%	4119	14.1%
Unadjusted Likelihood Ratio (95% CI)			18.27 (5.33–107.40)		7.03 (5.01–10.97)		2.36 (2.15–2.84)		0.90 (0.85–1.05)		0.32 (0.27–0.42)		—
Adjusted Likelihood Ratio (95% CI)			2.87 (0.49–16.76)		2.79 (1.47–5.31)		1.60 (1.12–2.28)		1.05 (0.78–1.41)		0.52 (0.35–0.78)		—