Aus der Klinik für Anästhesiologie der Ludwig-Maximilians-Universität München

Direktor: Prof. Dr. med. Bernhard Zwißler

Utilization of the Surgical Apgar Score as a Continuous Measure of Intra-operative Risk

Dissertation

zum Erwerb des Doktorgrades der Medizin

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

vorgelegt von

Monika Zdenka Jering

aus Starnberg

2015

Mit Genehmigung der Medizinischen Fakultät

der Universität München

Berichterstatter: Priv. Doz. Dr. Florian Weis

Mitberichterstatter: Prof. Dr. Thorsten Annecke

Prof. Dr. Orsolya Genzel-Boroviczény

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 23.06.2016

Table of Contents

Abbreviation

AIMS	Anesthesia Information Management System
ASA	American Society of Anesthesiologists physical status classification system
bpm	beats per minute
CI	Confidence Interval
CPR	Cardiopulmonary Resuscitation
CV	Coefficient of variation
EBL	Estimated blood loss
ECG	Electrocardiography
EDW	Vanderbilt Enterprise Data Warehouse
HR	Heart Rate
ICU	Intensive Care Unit
ICD 9	International Classification of Diseases 9th Revision
MAP	Mean arterial pressure
NT-proBNP	N-terminal pro brain natriuretic peptide

PDW	Perioperative Data Warehouse
POSSUM	Physiologic and Operative Severity Score for the enUmeration of Mortality and Morbidity
RCRI	Revised Cardiac Risk Index
ROC curve	Receiver operating curve
sAs	surgical Apgar score
SQL	Structured Query Language
VPIMS	Vanderbilt perioperative Information Management System database
WHO	World Health Organization

1 Introduction

1.1 Significance of surgical risk scores

Concurrently, the number of surgeries performed worldwide has dramatically increased.[1]Not surprising, patients suffering from chronic disease are more likely to undergo surgery and yet they are also at highest risk for post-operative complications and death.[1] Given the increasing proportion of patients with significant comorbidities undergoing surgery, awareness of post-operative complications and mortality associated with their chronic disease states should be heightened and these concerns should be viewed as a global public health priority. [2] Over the past years much attention has been given to the human failure within the surgical team but most patients faced major complications or death due to their comorbidities.[1] An individual risk assessment would benefit the patient, as medical treatment plans and intensified medical care if necessary could be applied accordingly suited to the patient's individual needs. [3] So far most clinician's rely more often on their intuition than on objective risk assessment tool. A study conducted in 2005 showed that surgeons underestimated the risk of complications for emergency cases in general surgery. On the other hand they tend to over predict the mortality and morbidity rates for elective surgeries in general surgery. [4] A previous study has argued that the surgeon's "gut-feeling" is a good predictor of postoperative morbidity, especially when the patient was doing well.[5] Hartley *et al.* argued that surgeons are more likely to be more pessimistic about their patient's outcome, which leads to more caution in the postoperative care.[5] Nevertheless, the surgeon's prediction is influenced by many different variables, such as pre- and intraoperative factors and the physician's clinical experience.[5]

A combination of the doctors' clinical assessment and applying an objective risk score might be the best reliable predictor for a patient's outcome. Today risk scores are necessary more than ever, as more complicated procedures are performed on sicker patients. [(6)]

1.2 Surgical risk scores

Since early identification of high-risk patients and prompt, appropriate intervention aimed at improving patient outcome can reduce the length of hospital stay,[(7)] morbidity, and mortality, [(8-10)] multiple risk scores have been designed to identify vulnerable patient populations' pre-and post-operatively. [(8-11)]

In the following chapters the most commonly applied risk scores in surgical patients will be discussed, such as the American Society of Anesthesiologists physical status classification system[(12)] (ASA classification), the physiologic and Operative Severity Score for the enUmeration of Mortality and Morbidity (POSSUM)[(13)], the revised cardiac risk index (RCRI)[(14)], and the surgical Apgar score (sAs)[(8)]. The American Society of Anesthesiologists classification is commonly used as a preoperative risk assessment of the patients' current health status.[(15, 16)] The RCRI predicts cardiac risk in non-cardiac surgeries and is commonly used in various surgical fields.[(11)] The POSSUM score was intended for surgical audit purposes but is also widely applied as a surgical risk assessment.[(13, 17)] However, most of these risk scores are based on rather complicates algorithms consisting of numerous variables, and are not easily applicable at the patient's bedside.[(11)] Each risk score has many benefits but also several limitations which will be discussed in detail within each section.

In 2007 Gawande *et al.* developed the surgical Apgar score, a simple ten-point scoring system.[(8)] End-of-case surgical Apgar score provides clinicians with useful and important information about patients' postoperative risk for major complications.[(18)] The score distinguishes between patients with high and low risk for adverse events and serves as a decision-support tool.[(19)]

We hypothesize that continuous monitoring of sAs trends may provide additional information about changes in a patient's risk for complications and may exhibit greater predictive ability about a patient's post-operative morbidity and mortality.

1.2.1 American Society of Anesthesiologists physical status classification system

1.2.1.1 Background

The ASA classification of Physical Status was first introduced in 1941 by Saklad for retrospective analysis of hospital records, classifying patients into seven categories.[(12)] In 1961 an alteration of the classification was proposed by Dripps *et al.*[20] and the new classification was modified to five classes with the approval of the American Society of Anesthesiologists in 1963.[21] The classification is based on the past medical history and a preoperative examination of the patient[22] :

ASA Physical Status I - A normal healthy patient

ASA Physical Status II - A patient with mild systemic disease

ASA Physical Status III - A patient with severe systemic disease

ASA Physical Status IV - A patient with severe systemic disease that is a constant threat to life

ASA Physical Status V - A moribund patient who is not expected to survive without the operation

ASA Physical Status VI - A declared brain-dead patient whose organs are being removed for donor purposes[22]

The patient is allocated to one of the six categories. The letter E behind the Roman numbers classifies an emergency case and the patient is therefore considered to be in a poorer condition. [22]

1.2.1.2 Application

Initially the ASA classification served as a description of the patient's physical status. However it was widely adapted and used as an individual surgical risk predictor for morbidity and mortality. [15, 23]

A study which was conducted in five different teaching hospitals in the United States demonstrated that a higher class in ASA physical status is associated with prolonged length of hospital stay, complications, and more follow-up visits at physicians after discharge.[24] Cullen *et al.* proposed to combine the ASA classification with age in order to equally use resources among patients. The adapted ASA classification might predict follow-up visits, the patient's risk for adverse events, and hospital length of stay more accurately.[24] However age was not incorporated in the commonly used ASA classification. It was reported that the ASA classification reports long term mortality in patients undergoing orthopedic surgery.[25] In a further study four different classification systems were validated for their predictive ability and it was reported that the ASA classification is more predictive for major postoperative complications than for minor adverse events. [6]

1.2.1.3 Benefits and limitations

Several studies have demonstrated that a higher ASA class is associated with a higher perioperative mortality, morbidity and longer hospitalization. [21, 23, 26-30] Additionally, the ASA classification was found to be significantly associated with intraoperative blood loss, intensive care unit stay, duration of postoperative ventilation, and cardiopulmonary complications.[21] As the classification is based on physical examination and the evaluation of the patient's past medical history the assessment is inexpensive as no additional diagnostic testing is required. [21, 31] The classification system is applied in various medical settings, such as in hospitals, [32] outpatient clinics,[33] and researchers who investigated the severity of surgical procedures and its risk evaluation.[34] The common application of the ASA classification might be due to its simplicity and suitability for all surgical services.[35]

Nevertheless, the simplicity of the ASA classification is double edged. On the one hand it alleviates communication among health care providers and enables a fast and simple application.[31] On the other hand its simplicity leads to discrepancy in the allocation among different physicians and inaccurate clinical interpretation.[22, 36] Additionally, the ASA classification does not consider the severity of the surgical procedure and it lacks scientific precision, [15] as the assessment does not incorporate objective criteria to the classification. Furthermore the assignment of an ASA class varies among physicians and is therefore not entirely reliable.[15, 22, 37, 38]Moreover, the commonly used classification does not incorporate age, weight, sex, and pregnancy.[11, 39] Aplin et colleagues discovered many variation in the ASA classification in children, concluding that the ASA classification is not a reliable tool in predicting surgical outcome in a pediatric population. [40]

1.2.2 Revised Cardiac Risk Index

1.2.2.1 Background

The Revised Cardiac Risk Index (RCRI)[14], an alteration of the Goldman Index [14, 41], predicts the risk of major cardiac complications and cardiovascular mortality [42] in non-cardiac surgery. Major complications were defined as myocardial infarction, pulmonary edema, ventricular fibrillation or primary cardiac arrest, and complete heart block. [14]

The Revised Cardiac Risk Index was derived from analysis of 29 variables in a cohort of 4,315 patients undergoing major elective non-cardiac surgery in a tertiary-care teaching hospital. The Revised Cardiac Risk Index is based on the following six values:

- High-risk type of surgery (intraperitoneal, intrathoracic, or suprainguinal vascular surgery)
- History of ischemic heart disease (history of myocardial infarction, history of positive exercise test, current complain of chest pain considered secondary to myocardial ischemia, use of nitrate therapy, ECG with pathological Q waves)

- History of congestive heart failure (history of congestive heart failure, pulmonary edema, paroxysmal nocturnal dyspnea, bilateral rales or S3 gallop, chest radiograph showing pulmonary vascular redistribution)
- History of cerebrovascular disease (history of transient ischemic attack or stroke)
- Preoperative treatment with insulin
- Preoperative serum creatinine greater than 2.0 mg/dL[14]

If two of the six criteria are met, patients with intermediate and higher risk of cardiac risk can be identified.[14]

The study conducted by Lee *et al.* has several limitations: no neurosurgical cases and emergency cases were incorporated. Furthermore patients who had an anticipated hospital length of stay of less than 2 days were also excluded. The study was predominated by thoracic, vascular and orthopedic cases. [14, 43, 44]

1.2.2.2 Application

The Revised Cardiac Risk Index is the optimal cardiac risk score applied in non-cardiac surgery, according to the American College of Cardiology/American Heart Association 2007 guidelines [45] and European Society of Cardiology/European Society of Anesthesiology guidelines. [46]
Due to the good predictive ability for cardiac adverse events in non-cardiac surgeries, the RCRI is the most commonly used cardiac risk score.[47] The RCRI has been confirmed to be applicable in numerous surgical procedures such as in vascular surgery and also in lung resection.[48-52]
The revised cardiac risk index was modified and tested for its predictive ability in elective orthopedic surgery for non-cardiac adverse events. It was shown that the score is not a strong predictor for postoperative morbidity.[53] A recent study has shown the significant association between the RCRI, postoperative hospital length of stay and morbidity in patients undergoing elective orthopedic surgery. The risk for morbidity and prolonged hospitalization increased, with a modified revised cardiac risk index greater than three.[53]

1.2.2.3 Benefits and Limitations

The Revised Cardiac Risk Index is predictive for myocardial infarction and perioperative death in non-cardiac surgery.[11, 53]

The index has proven to predict short and long term cardiac complications within patients undergoing non-cardiac surgery.[54] The RCRI is an inexpensive predictive tool derived from the past medical history, physical examination, and serum creatinine.[55] The index is a simple tool to identify patients, over the age of 50 at risk of having a cardiac complication undergoing elective non-cardiac surgery.[56]

However, it was reported that the index is not an accurate cardiac risk predictor for patients undergoing vascular surgery.[57, 58] A study has shown that the predictive ability of the RCRI is decreased in patients over the age of 75, who undergo vascular surgery.[49] Furthermore it was described by Choi *et al.* that the Revised Cardiac Risk Index is not predicting acute myocardial infarction, pulmonary edema or cardiovascular death as accurately as biomarkers, such as preoperative NT-proBNP or cardiopulmonary resuscitation (CPR).[59] It is commonly known, that elderly patients are at an increased risk of suffering from cardiac adverse events and therefore cardiac risk assessment is of high importance.[49, 60, 61] The RCRI does not indicate which patient needs additional diagnostic investigation and therefore leads to extravagant cardiac testing. [62] Recent results of Lupei *et al.* indicated that the RCRI is not associated with postoperative intensive care outcomes.[63] Another limitation of the Revised Cardiac Risk Index is that it only predicts perioperative cardiac complications, as it is not designed to predict the overall mortality risk in surgical patients. [58, 64] Consequently, the score cannot predict any other surgical complication, such as surgical-site infection or pulmonary complications.[65] This limitation is relevant, as only 1% of all patients who undergo non-cardiac surgery every year suffer from cardiac complications.[44, 66]

1.2.3 Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM)

1.2.3.1 Background

The Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM) was introduced by Copeland *et al.* in 1991 and was originally used as an operative severity score for general surgery in comparative surgical audit.[13] The intention of the score was to compare individual surgical performance and different hospitals among each other. Surgical performance is measured by comparing predicted negative outcome with observed outcome. Copeland and colleagues incorporated 48 physiologic parameters and 14 operative and postoperative factors to assess the score.[13] By using multivariate analysis the POSSUM was simplified.[13] The final POSSUM score incorporates 12 physiological variables and in combination with 6 operative variables. (Table 1)[13, 67] The POSSUM score predicts the following complications described in Table 2.[13]

Table 1. POSSUM: physiological and operative parameters

Physiological parameters	Operative parameters
Age	Operative severity
Cardiac signs	Multiple procedures
Respiratory history (dyspnea)	Total blood loss
Blood pressure	Peritoneal soiling
Pulse rate	Presence of malignancy
Glasgow Coma Score	Mode of surgery
Hemoglobin level	
White cell count	
Urea concentration	
Sodium	
Potassium	
Electrocardiography	

Table 2. Definition of POSSUM score complications

Wound hemorrhage	local hematoma requiring evacuation
Deep hemorrhage	postoperative bleeding requiring re-exploration
Chest infection	production of purulent sputum with positive bacteriological cultures, with or without chest radiography changes or pyrexia, or consolidation seen on chest radiograph
Wound infection	wound cellulites or the discharge of purulent exudate
Urinary infection	the presence of $>10^5$ bacteria/ml with the presence of white cells in the urine, in previously clear urine
Deep infection	the presence of an intra-abdominal collection confirmed clinically or radiologically
Septicemia	positive blood culture
Pyrexia of unknown origin	any temperature above 37°C for more than 24 h occurring after the original pyrexia following surgery (if present) had settled, for which no obvious cause could be found
Wound dehiscence	superficial or deep wound breakdown
Deep venous thrombosis and pulmonary embolus	when suspected, confirmed radiologically by venography or ventilation/perfusion scanning, or diagnosed at post mortem
Cardiac failure	symptoms or signs of left ventricular or congestive cardiac failure that required an alteration from preoperative therapeutic measures
Impaired renal function	arbitrarily defined as an increase in blood urea of > 5 mmol/l from preoperative levels
Hypotension	a fall in systolic blood pressure below 90 mmHg for more than 2 hours as determined by sphygmomanometry or arterial pressure transducer measurement

In the original paper pediatric patients were excluded from the study, as the physiologic values are different than those in adults. Additionally all patients leaving the same day as their surgery, were excluded, as their mortality and morbidity rates were very low. [13] Cardiac signs are defined as no cardiac failure, diuretic, digoxin, antianginal, warfarin or antihypertensive therapy, peripheral edema, warfarin therapy, cardiomegaly and raised jugular venous pressure.[68]

The physiological and surgical parameters are divided into four categories with exponentially increasing scores of 1,2,4, and 8. If a variable is not available a score of 1 is assigned.[67] In order to calculate the risk of suffering from complication, physiological and surgical variables are summed and entered into the subsequent equations for mortality and morbidity. [11] For the prediction of a patient's mortality risk the following formula is used:

ln R/1 -R= -7.04+ (0.13 x physiological score)+ (0.16 x operative severity score)[13]

For the prediction of the morbidity rate the later equation is used:

ln R/1- R = - 5.91 + (0-16 x physiological score)+ (0.19 x operative severity score) [13]

The letter R represents the predictive risk of mortality or morbidity.[13] The POSSUM score ranges between 12 and 88 points for the physiologic parameters. Operative parameters are ranging between 6 and 44 points. Operative and physiologic values are added and represent the POSSUM score.[69] The lower the score the less likely is the chance to suffer from major complications.[69]

1.2.3.1.1 Application

Several studies have shown that the POSSUM score over predicts mortality especially in a low-risk population.[70-72] Therefore the POSSUM score was revised to the Portsmouth POSSUM (P-POSSUM).[72] The P-POSSUM uses the same parameters as the POSSUM score. Other than the original POSSUM score which uses logistic regression models, the P-POSSUM uses linear models in order to calculate the mortality risk.[67] The Portsmouth POSSUM equation was established for predicting mortality:

ln[R/(1-R)]= - 9.37 + (0.19x physiological score)+ (0.15 x operative severity score)[72]

The letter R represents the predicted risk for mortality. The alteration of the score predicts mortality more accurately in patients who are at low surgical risk.[72] The POSSUM and P-POSSUM was validated in different surgical subspecialties such as general, colorectal [17], upper gastrointestinal surgery[73, 74], and vascular surgery.[75, 76] However, it was reported that the POSSUM score has several limitations in different surgical subspecialties such as over predicting mortality rates in different subsets of patients.[77, 78] Therefore specialty specific POSSUM scores have been developed such as the V-POSSUM[79, 80] for elective vascular surgery, the CR-POSSUM[77] for colorectal surgery, and the O-POSSUM[81] for upper gastrointestinal surgery. The CR-POSSUM reduces the physiologic factors from original 12 values to 6 values and adjusts for age.[77]The O-POSSUM does not use multiple procedures, total blood loss, and peritoneal soiling as operative values. Thus the operative severity score is reduced to three values.[81]

Predicting morbidity and mortality with only the physiological parameters has proven to be applicable in vascular surgery based on the P-POSSUM equation.[79]

1.2.3.2 Benefits and Limitations

The POSSUM score is widely acknowledged as a surgical audit tool and surgical risk stratification.[17, 69, 82] The major advantages of the POSSUM score are that it incorporates intra-operative data, surgical parameters, and operative risk. Additionally it predicts thirty-day mortality accurately in an elderly population after an age adjustment.[83] The vascular POSSUM score is considered to be a good measurement for surgical performance in major vascular surgery.[76] It is argued that the values needed for the risk calculation are routinely captured before undergoing surgery and are easily obtained. [71, 72, 84] The POSSUM score seems to be applicable in different countries across the world with different accessibility to medical resources.[73, 85] The P-POSSUM is reported to predict mortality risk in patients undergoing emergency surgery accurately.[82]

Nevertheless the POSSUM score has several limitations. It is not applicable in children or in day patients.[86] It was reported that the P-POSSUM is a better predictor for in-hospital mortality than the POSSUM system.[17, 87, 88] However, the P-POSSUM only has a mathematical formula to predict mortality. The authors argued that P-POSSUM did not support its use as a reliable tool to predict morbidity.[70] Although it is argued that the values for the calculation of the score are easily obtained, most variables are not routinely measured and therefore assessing score values is more time consuming for the medical staff. [82, 86]The score cannot be easily calculated at the bedside and consists out of a complex formula.[11] Additionally the POSSUM score is more likely to over predict mortality especially in low risk surgical patients. [70-72]The POSSUM score predicts the overall risk of a patient population. Therefore it is not accurate enough to support the clinician's during the decision making progress for an individual patient.[75] For the application of the score, all preoperative data has to be gathered before the patient undergoes surgery and the calculation is rather complex. [89] It was reported that Electrocardiography (ECG) should only be conducted on elderly patients undergoing a higher risk surgery, [60, 90] as patients above 60 years are more likely to present an abnormal ECG.[90]There is no need for a preoperative ECG if the patient undergoes a minor surgery.[60, 90] This additional unnecessary screening that is needed for the POSSUM risk calculation would add to the already immense workload for ward staff. Another major limitation of the POSSUM is that the score is only available after the surgery is completed.[67]Adjusting the POSSUM score for different surgical specialties diminishes the overall practicability and the ability to compare surgical outcomes across all subspecialties.[91] Copeland also suggested that the POSSUM score should not be used as a risk prediction but only as a surgical audit tool.[67] In summary the POSSUM score is not suitable for risk assessment, as gathering data and calculating the score is not compatible with the daily routine in surgical care.

1.3 Surgical Apgar score

1.3.1 Development of a surgical risk score

In 1953 Virginia Apgar provided clinicians with a simple ten point scoring system to be used in newborn infants in order to assess their postnatal condition and to predict their 28 day survival following birth.[92] The Apgar score is determined at minute one, five, and ten after childbirth and gives the clinician immediate information on the patient's condition and on the likelihood of experiencing adverse outcomes following birth. Simple in design, the Apgar score is based on the following five factors: the infant's appearance, pulse, grimace, activity, and respiration. Based on the infant's condition, the young patient will be assigned a score from zero to two in each category. At the end of the assessment these points are summed together, giving the maximum of 10 points.[93] Ten points represent the highest score an infant can reach and corresponds to a condition of excellent health of a newborn.[94] The Apgar score is applicable in every hospital setting and since it was first risk stratification system to be applied in newborns, it revolutionized obstetrics.[95] Infants at high risk for postnatal complications can thus be identified early on, which initiated a cascade of appropriate work up for their poor health and interventions to optimize patient outcome.

Due to a lack of an efficient and simple objective assessment of the patient's postoperative prognosis, Gawande *et al.* developed a score similar to the Apgar score in 2007 and validated it in general and vascular surgery.[8] Previous studies had proven that perioperative tachycardia, low blood pressure and considerable blood loss would cause major complications following surgery.[21, 96, 97] By combining these three intraoperative values into one score calculated at the end of a surgical case, Gawande derived the 10-point surgical Apgar score, which is based on routinely measured values:

- **Lowest mean arterial pressure** (MAP)
- **Lowest heart rate** (HR)
- **Estimated blood loss** (EBL) (Table 3) [8]

Multivariable logistic regression was used to derive the sAs out of 49 preoperative variables, 28 intraoperative values, and 33 outcomes within a study population that totaled 1,172 patients. The aforementioned three independent values including lowest heart rate, mean arterial pressure and estimated blood loss were most predictive for postoperative mortality and major complications as defined by the

National Surgical Quality Improvement Program's (NSQIP).[98] These include the following major complications:

- Acute renal failure
- Bleeding requiring transfusion of ≥ 4 U red cells within 72 hours after operation
- Cardiac arrest requiring CPR
- Coma for 24 hours or longer
- Deep venous thrombosis
- Septic shock
- Myocardial Infarction
- Unplanned intubation
- Ventilator use for 48 hours or longer
- Pneumonia
- Pulmonary embolism
- Stroke
- Wound disruption
- Deep or organ-space surgical site infection
- Sepsis
- Systemic inflammatory response syndrome
- Vascular graft failure[98]

Gawande's study was first retrospectively validated in 303 patients undergoing vascular and general surgery at the Brigham and Women's Hospital in Boston, Massachusetts. Beta-coefficients were used to allocate appropriate weight to each of the chosen variables on a suitable 10-point scale. The intervals between the points were chosen such that a one-point increase in a patient's sAs would significantly raise the occurrence of postoperative complications (relative risk of 16.1 among patients with a surgical Apgar score below 4 points) and additionally showed clinical significance. Table 3 delineates the construction of the surgical Apgar score by component. A patient with a lowest heart rate of 54 beats/minute (4 points), lowest mean arterial blood pressure of 80 mmHg (3 points) and estimated blood loss of 50 ml (3 points) would receive the maximum sAs of ten. In contrast, a lowest intraoperative heart rate of 80 beats/minute (1 point), lowest mean arterial blood pressure of 50 mmHg (1 point) and an estimated blood loss of 700 ml (1 point) would equal a surgical Apgar score of three. The second patient with a sAs of three would have an increased risk (relative risk 15.4 [CI 7.2-33.1]) of suffering from postoperative complications compared to the first patient with a sAs of ten. As this example illustrates, the risk for mortality and morbidity increases significantly with decreasing sAs.[8]

Table 3. Ten-point Surgical Apgar Score

	0 Points	1 Point	2 Points	3 Points	4 Points
Estimated blood loss (mL)	> 1000	601-1,000	101-600	≤ 100	--
Lowest mean arterial pressure (mmHg)	< 40	40-54	55-69	≥ 70	--
Lowest mean heart rate (beats/minute)	>85	76-85	66-75	56-65	≤ 55*

Table 3. The surgical Apgar score = sum of the points for each category in the course of the procedure

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

Concurrently, Gawande *et al.* conducted a further prospective study with 869 patients undergoing vascular and general surgery at the same institution to validate the sAs and to evaluate the accuracy of the sAs in predicting postoperative morbidity and mortality. This study proved the surgical Apgar score to be adept in predicting the patient's risk for major complications and death within 30 day following surgery.[8]In 2009 the predictive ability of the surgical Apgar score was revalidated in 4,119 patients undergoing vascular and general surgery at the Massachusetts General Hospital in Boston. In this study electronic data records were used instead of handwritten intraoperative charts, which the score was initially derived from.[19]

1.3.2 Utility of the surgical Apgar score

Most risk scoring systems were developed for the use in an intensive care unit and there is still a lack of a rating which is applicable during surgical procedures.[99] Multiple studies have reported that intra-operative blood loss[13, 21], tachycardia[96, 100], bradycardia [100, 101], hypertension, and hypotension[102-104] are independently associated with patient's outcome. By combining these three predictive values has proven to predict a patient's postoperative risk of suffering from major complication.[8] Hypotension, tachycardia and extensive blood loss are very likely to be treated individually by clinicians. However each variation in vitals has an impact on the patient, which is captured by the surgical Apgar score. In several studies the sAs has proven to predict the morbidity and mortality rate adequately. [8, 18, 105-108] The 10-point scoring system is easily understandable and might help communication between the surgical team, the postoperative care team, the medical staff on the wards and the patient. A low sAs could support a clinician's decision to transfer the patient to an intensive care unit, where close monitoring and one to one nursing is possible. In recent studies a low surgical Apgar score has proven to be predictive for intensive care unit (ICU) admission[109, 110] and increased length of ICU stay in general surgery.[110] Furthermore Glass *et al*. described that the sAs is also associated with late transfer to the intensive care unit.[110] Patients with higher scores are more likely to be transferred to the ward where close monitoring is not necessary. [111] The surgical Apgar score immediately identifies patients who are at a higher risk of having major complications. Patients who were undergoing minor vascular or general surgery at the Massachusetts General Hospital and had a surgical Apgar score of 4 or less were 22.8 times more likely to develop major complications and 81.4 times more likely to die within 30 days after a surgical procedure.[19]Patients who underwent major or prolonged general or vascular surgical procedures at the same institution with a score of 4 or less were 6.5 times more likely to experience major complications and 112.0 times more likely to die within 30 days.[19] The surgical Apgar score can only predict the patient's risk for major complications and risk but it cannot measure the surgical team's intraoperative performance or can compare different clinicians and institutions, as the score does not take surgical complexity and patient's relevant risk factors into account. [18] Patients with greater blood loss have lower sAs and thus have an increased risk for major complications. Surgeons with greater EBL at the end of a surgical case would be therefore more likely to better predict patient's outcome. Consequently the score does not operate as an audit tool. [18]

However the surgical Apgar score is applicable in all different types of anesthesia with the exception of local anesthesia as no electronic records are available.[8] The surgical Apgar score could also be used as a tool to aid researchers worldwide to improve surgical procedures in their individual hospital setting and ensure patient's safety.[112, 113]

1.3.3 Validation of the surgical Apgar score

Since the development of the surgical Apgar score in patients undergoing vascular and general surgery, the risk assessment was validated in different surgical procedures and various surgical specialties, such as urological, colorectal, gynecological, pancreatic surgery and spine surgery. [105-107, 111, 114, 115] The score has also demonstrated usefulness in renal mass excision,[116] and predicting ICU admission after general surgery.[109] The sAs was validated in a wide range of international clinics and it was reported that the score is applicable in almost every hospital setting. Due to its simplicity in gathering the data for the score, resource poor hospitals can still apply the sAs. [112] Other risk scores, which rely on multiple variables and laboratory values, might not be calculated in different clinics throughout the world due to the expense factor. In some countries the vitals are being monitored manually in others electronically which leads to variability in obtaining the risk score. [108, 112] However, the sAs can also be derived from hand written anesthesia charts, as it was first validated form manually captured records.[8] The surgical Apgar score was validated in neurosurgical study population at the University of Michigan hospitals. In this study, the score was predictive for major complications, prolonged hospitalization, and intensive care unit stay.[117] Ohlsson *et al.* validated the association between perioperative complications and the surgical Apgar score, intensive care unit and hospital length in a Swedish setting. Lower sAs were strongly correlated with longer hospitalization and length of stay at an intensive care unit.[108] Melis *et al.* demonstrated the predictive ability of the surgical Apgar score in a veteran population undergoing general surgery.[118] In a recent study conducted in Japan, the surgical Apgar score was validated after gastrectomy for gastric cancer. Miki *et al.* modified the surgical Apgar score by using the quartile values of EBL as their median sAs was lower than in general surgery. This manuscript demonstrated that a modified sAs is an independent predictor for major complication in this particular study population.[119] Validating the sAs in hip and knee arthroplasty it was demonstrated that the score is not a reliable risk measurement. Only 6.1% of the patients with major complications had a score of 4 or less. 75.8% of patients with major complications had a score of 7 or higher.[113] Recently, it was reported that the surgical Apgar score is applicable in all major surgical subspecialties in predicting postoperative mortality.[120] However, the predictive ability of the sAs varies among different surgical subspecialties. The correlation between sAs and burn patients is very weak, whereas the relationship between the predictive ability of sAs in gynecology is very high. Different patient spectrum with various comorbidities might be accountable for the variation within the subspecialties.[115, 120] Nevertheless the surgical Apgar score still remains a valuable risk score in several validated subspecialties. [120]

1.3.4 Benefits of the surgical Apgar score

The surgical Apgar score is an objective measurement of the patient's condition and will provide the clinician with more information about the patient's postoperative well being without requesting additional diagnostic testing. As the values for the calculation of the score are already routinely captured, the sAs could be incorporated into routine clinical practice with minimum resource implications. Furthermore, no particular schooling or equipment is required to assess the score.[118] The variables used to calculate the surgical Apgar score are influenced by many different factors, such as anesthesia care and medication, the patient's prior condition, age, operative complexity, and the surgical team. [101, 105, 118, 121, 122] The estimation of the blood loss adds a subjective component to the score that incorporates the clinicians experience and the competent evaluation of the situation. By using intraoperative vital measurements, the score takes the magnitude of the operation performed, patient's intraoperative responsiveness to the procedure and postoperative condition into account. The calculation of the sAs at the end of a surgical case helps clinicians to identify patients who are at a higher risk of having major complications and immediate action, such as close monitoring on a step down unit or transfer to the intensive care unit, can be taken to prevent any incidents. [18] Patient's safety and better hospital outcomes might improve by applying the sAs as a routine measurement.

1.3.5 Limitations of the surgical Apgar score

There are several limitations to the surgical Apgar score. One major criticism of the score is the estimation of the blood loss. The original authors argue that the ranges set for the estimation of blood loss should provide an accurate estimate based on published studies. [8, 123, 124]

Some studies noted that the sAs might not comprehensively predict outcomes by itself: The surgical Apgar score did not support its use as a reliable tool to predict adverse events in hip and knee arthroplasty.[113] However, the score provided useful information to perioperative risk in this study population.[113] Recent results of Lau *et al.* indicated that sAs is not able to predict major complications following spinal metastasis surgery. Though this study only included a small study population of ninety-seven patients and therefore needs to be revalidated in a larger patient cohort. [125]A study, which was conducted in a district general Hospital in the United Kingdom, has shown that the surgical Apgar score has a low predictive ability for major complications and mortality in orthopedic cases such as elective major joint replacements and emergency femur fracture cases. In the same study it is supposed that the sAs does not improve individual postoperative care. [126] Urrutia *et al.* validated the sAs in general orthopedic surgery and the score failed to demonstrate usefulness in this study population. The score was only predictive in a subgroup of patients undergoing spine surgery for major complications.[127] Another frailty of the surgical Apgar score is that the anesthetic management and medication influence the variables used to calculate the score. [112] Hypotensive episodes may occur during induction due to various reasons and would subsequently lead to a lower sAs.[128, 129] Furthermore bradycardic episodes during the surgery cause a higher surgical Apgar score. However many studies have shown that intra-operative hypotension and tachycardia contribute significantly to a negative outcome, regardless of their cause.[96, 97, 121, 130, 131] Another of the sAs's limitations is that vitals must be captured automatically or manually at least every 5 minutes in order to obtain an appropriate score. [112]So far the sAs is only applicable in an adult population.[19] Nevertheless, the surgical Apgar score remains a solid predictor for postoperative negative outcome.[18, 19]

1.3.6 Continuous Evaluation of the Surgical Apgar Score

Measurement of risk for postoperative complications before, during, and after surgery is important in guiding medical decision making. Having a better understanding of when a patient's risk profile changes during a surgical procedure is an important goal which might guide more timely interventions, triage decisions, and enhance communication among the perioperative team. In spite of rapid technological advances, the state of the art with respect to perioperative risk measurement and appropriate real-time notification systems about perioperative risk and dynamic changes in operative risk is still quite limited.[132]

The surgical Apgar score was developed and validated using data available and applied at the end of the case. However monitoring the sAs trends throughout a case is possible and may provide additional information. Real-time assessment of the surgical Apgar score and notification systems apprising clinicians of rapid changes in a patient's surgical Apgar score may represent an objective tool to aid clinicians in their decision making process, allowing them to rely on objective data rather than on intuition and clinical experience in creating summative evaluations of patient condition. Additionally, the risk score may support the clinician's request for additional diagnostic testing, invasive patient monitoring, ordering a transfer to the intensive care unit and intensifying medical care. We hypothesize that continuous surgical Apgar score monitoring may be used intraoperatively to identify patients at high risk of postoperative complications.

1.4 Comparison of surgical risk scoring tools

Table 4. displays the different risk scores with components and main limitations which are discussed in this dissertation.

Table 4. Comparison of surgical risk scores

Risk score	components	main limitations
ASA classification	general medical history general examination	subjective assessment inconsistent rating
Revised Cardiac risk index	high-risk type of surgery	Only prediction of major cardiac complications
	History of ischemic heart disease	No mortality prediction
	History of congestive heart failure	
	History of cerebrovascular disease	
	Preoperative treatment with insulin	
	Preoperative serum creatinine >2.0 mg/dL	
POSSUM	**physiological parameters**	Only applicable in hospitalized patients
	Age	Many variables
	Cardiac signs	Complex calculation
	Respiratory history	
	Blood pressure	
	Pulse rate	
	Glasgow Coma Score	

	Hemoglobin level	
	White cell count	
	Urea concentration	
	Sodium	
	Potassium	
	Electrocardiography	
	Operative parameters	
	Operative severity	
	Multiple procedures	
	Total blood loss	
	Peritoneal soiling	
	Presence of malignancy	
	Mode of surgery	
Surgical Apgar score	Lowest heart rate	Estimation of blood loss imprecise
	Lowest mean arterial pressure	Variation in predictive ability
	Estimated blood loss	Variables influenced by cofactors

2 Specific Aims

The purpose of this study is to determine whether a continuous real-time assessment of the surgical Apgar score across the entire duration of a surgical case can provide additional information about the patient's postoperative condition when compared to a single sAs value determined at the very end of the procedure. By extending the utility of the sAs and transitioning from a single end-of-case value to a continuous trend that can be monitored and used throughout the perioperative process, we expect that more timely interventions in response to drastic fluctuations in surgical Apgar scores can reduce perioperative morbidity and mortality.

2.1 Revalidation of the surgical Apgar score in our study cohort

At first the correlation between the end-of-case surgical Apgar score and major postoperative complications will be evaluated to re-validate the sAs in our study cohort and prove its applicability for our model. In previous studies the surgical Apgar score has demonstrated predictive ability in terms of postoperative patient morbidity and mortality, where patients with lower scores had a significantly increased occurrence of major postoperative complications within thirty days of surgery. To re-validate the end-of-case sAs in our study cohort, the predictive ability of the sAs for postoperative adverse events will be examined.

2.2 Evaluation of surgical Apgar score trends in twenty different surgical subspecialties

We will investigate whether continuous monitoring of the sAs throughout the perioperative process and evaluation of sAs trends may allow for more optimal patient risk stratification. We will investigate the association between sAs trends and major perioperative complications. Subsequently these sAs trends will be analyzed for each surgical subspecialty included in this study. Trends in sAs may vary based on patient characteristics, surgical subspecialty investigated and duration of the procedure. The trend calculation will be based on mean sAs, coefficient of variation, and slope of sAs throughout the procedure. Lastly we will examine whether these trend characteristics are associated with postoperative morbidity or mortality.

2.3 Real-time assessment of surgical Apgar score in general, vascular, and general oncology surgery

To establish continuous assessment of the surgical Apgar scores, the sAs will be calculated at serial points in time throughout a case whenever new values required for the calculation are recorded in the patient's electronic anesthesia chart. Continuous sAs calculation will be conducted in general, vascular, and general oncology surgery. Hypothetical real-time assessment of the score will be established.

2.4 Comparison of the predictive ability between continuous assessment of surgical Apgar score and the American Society of Anesthesiologists physical status classification system

The ASA classification is widely used throughout the world as a risk stratification model for surgical patients. We will compare the predictive ability of the ASA classification system to that of continuous sAs monitoring in the form of sAs trend analysis. Additionally, the two scores will be combined into a single model and assessed for their predictive ability. For the combined model, analysis will be based on surgical Apgar scores calculated continuously throughout the surgery.

2.5 Establishing a model for a threshold probability

The surgical Apgar score serves as a clinical tool that helps predict the risk or probability for morbidity and mortality following surgery. We will design a model that raises a hypothetical alert when a patient that underwent general, vascular, or oncology surgery, exceeds a predetermined probability of postoperative complications based on his/her sAs trend. Since patients crossing this chosen threshold might exhibit higher admission rates to the intensive care unit, might require longer postoperative hospitalization, and might be more likely to suffer from postoperative morbidity and mortality, they might benefit from earlier and more aggressive interventions in the operating room.

2.6 Applying thresholds in hypothetical patient populations

All surgical procedures will be electronically re-analyzed to re-validate our model in a hypothetical patient population based on our original study cohort. Hypothetical alerts will be raised if a patient crosses a predetermined threshold probability for postoperative complications. As trends are able to capture acute changes in a patient's status, they may alert clinicians earlier about impending adverse events that can lead to postoperative morbidity and mortality. We will investigate at which point in the surgical procedure the first alert was raised. Lastly we will determine the predictive ability of the alert model.

3 Methods

3.1 Overview

We conducted a retrospective evaluation of patients who underwent non-cardiac surgery under general anesthesia performed at the Vanderbilt University Medical Center between January 1, 2009 and December 31, 2011. We divided the study in two parts:

In the first part of the study we investigated surgical Apgar score trends in twenty surgical subspecialties.

In the second part of the study we conducted a retrospective evaluation of patients who had general, vascular, or general oncology surgery.

We evaluated the continuous monitoring of the surgical Apgar score by developing a model of real-time assessment of the risk score during surgery performed in these three subspecialties. General surgery and vascular surgery were chosen for our study since the surgical Apgar score had originally been validated in these subspecialties.[8] General oncology was chosen due to the large sample size of the patient cohort giving a solid statistical basis. These three surgical subspecialties should serve as a model to verify or nullify the validity of the hypothesis stated in the introduction.

In addition we established alerting models based on the patient cohort on the second part of the study. Surgical subspecialties were defined by the primary service of the attending surgeon or the scrub nurse. If no service is assigned, the surgeon's department will be automatically pulled into the "Primary Service" field in the patient's chart. Intraoperative data were extracted from the Vanderbilt Medical center's perioperative data warehouse. The Vanderbilt University Human Research Protection Program, Nashville, TN, approved the study.

3.2 Patient population

Electronic records and Perioperative Information Management System databases were used to identify relevant patient characteristics, as defined below in data collection. The following measures were taken to refine our study population. Patients undergoing surgery over the age of 18 were eligible for inclusion. Characteristics of excluded patients are listed in Table 5.

Table 5. Excluded Patient Characteristic

Patient characteristic	Count n = 209,949
Non-operative cases	97,664
Age < 18	39,581
Single EBL recording >100 ml	39,048
Incomplete medical record	13,137
Other than general anesthesia	12,752
Cardiac patients	5,312
Community surgical patients	2,371
Age >100	41
Organ donors	21
Test patients	10
Date of Death>Surgery Start Date	12
All	209,949

Table 5. displays the number of patients excluded from the study for the reasons cited.

Since patients over the age of 100 years are more likely to die of causes other than the direct aftereffects of the surgery, they were excluded from the study. All organ donors, classified as ASA 6, were excluded. Patients cared for in off-campus surgical centers and non-operative cases (i.e. bronchoscopy, dental procedures, procedures in the intensive care unit, gastrointestinal, radiological and electrophysiology cases) were excluded, as were patients with incomplete patient data. We excluded patients with only a single documented blood loss recording at the end of the case and if the blood loss recording was greater than 100 ml. This allowed us to avoid sudden artificial changes of sAs trends at the end of a case in which patients' blood loss was not documented as it occurred throughout the procedure. Furthermore, all cardiac patients were excluded as most patients underwent a cardiopulmonary bypass procedure in which patients do not have a heart rate or a mean arterial pressure. In the absence of these recordings the sAs cannot be accurately calculated thereby mandating exclusion of these patients from the study. During investigation twelve patients were noted to have died before undergoing surgery, although they were, in fact, operated on. After further inquiry, the date of death had been incorrectly captured in the electronic database and consequently these patients were excluded. Lastly, all cancelled and test cases were excluded. Test patients are entered into the database for various learning purposes.

3.3 Data collection

Electronic records and Vanderbilt Perioperative Information Management System databases (VPIMS) were used to compute the surgical Apgar score and to identify relevant patient characteristics. All perioperative electronic data is stored in a data warehouse using Microsoft Structured Query Language (SQL) server technology (Microsoft Corporation, Redmond, WA). All pre- and postoperative data were excluded. The following variables were then extracted from the Vanderbilt Perioperative Data Warehouse (PDW):

- Patient demographics
- Medical record number
- Date of birth
- Gender
- Race
- Surgery date
- Surgery start
- Medical encounter number
- Hospital admission date
- Hospital discharge date
- Anesthesia patient number
- Anesthesia case number
- Indication (emergency or elective surgical procedure)
- Age on day of surgery
- Type of primary surgical procedure
- Surgery date and starting time
- Length of the surgical procedure
- American Society of Anesthesiologists Physical Status Classification

- Time of vital recording
- Heart rate
- Mean arterial blood pressure
- Estimated blood loss

Medical record number, surgery start date and time, and time of vital sign recordings make a case unique. All values for heart rate (HR), mean arterial blood pressure (MAP), and estimated blood loss (EBL) were extracted from the Enterprise Data Warehouse database, as time stamped data. Heart rate is recorded every 30 seconds, either electronically through the Plethysmogram, the electrocardiogram or manually when clinicians enter the heart rate into the patient's chart. Mean arterial blood pressure is captured every three minutes. Blood pressure was either measured noninvasively or invasively via an arterial line and sometimes required manual adjustment by clinicians. When both noninvasive and invasive blood pressure recordings were documented, we used the invasively monitored blood pressures since they are more accurate. Inaccuracies in the measurement of the heart rate and mean arterial pressure were noted when the catheter was flushed or the arm repositioned. To avoid artifacts, heart rate values outside of the range of 15 mmHg to 200 mmHg were discarded. MAP ranges were defined as 25-180 mmHg.

Our primary endpoint was death within 30 days of surgery or the occurrence of the following major postoperative complications within 30 days of the procedure:

- Acute renal failure
- Bleeding requiring transfusion of $\geq$ 4 units red cells within 72 hours after operation
- Cardiac arrest requiring CPR
- Coma for 24 hours or longer
- Deep venous thrombosis
- Septic shock
- Myocardial Infarction
- Unplanned intubation
- Ventilator use for 48 hours or longer
- Pneumonia
- Pulmonary embolism
- Stroke
- Wound disruption
- Deep or organ-space surgical site infection
- Sepsis

- Systemic inflammatory response syndrome
- Vascular graft failure[98]

The selection and definition of major complications was based on the National Surgical Quality Improvement Program (NSQIP) registry, which collects peri- and postoperative data from various institutions for a comparative analysis of complication rates and surgical outcomes.[133] [134]

ICD 9 codes, which are stored in the Perioperative Data Warehouse in association with each medical record, were used to identify peri- and postoperative complications. Data for wound disruption, ventilator use for 48 hours or longer, bleeding requiring $\geq$ 4 units red cell transfusion within 72 hours after operation, and unplanned intubation were extracted from the Vanderbilt Enterprise Data Warehouse (EDW; Oracle Corporation, Redwood Shores, CA) based on International Classification of Diseases ninth revision (ICD-9 codes) in the patient's chart. The medical record number was matched to ICD-9 codes corresponding to the postoperative complications defined above in order to identify study subjects who experienced adverse events prior to discharge or were readmitted to Vanderbilt University Hospital within 30 days. Patients' social security numbers were compared to the Social Security Death Index provided by the U.S. government in order to identify the date of death.

3.4 Study procedure

The surgical Apgar score was determined based on the lowest intraoperative heart rate, the lowest mean arterial blood pressure and estimated blood loss, following the protocol proposed by Gawande *et al.* in 2007.[8] Vanderbilt Perioperative Data Warehouse records were used to retrieve information necessary to compute the surgical Apgar score during the maintenance phase of each surgical procedure. Maintenance phase is defined as the time interval between "Anesthesia Ready Time" and "Dressing on Time," both of which are manually entered by the operating room nurses.

3.5 Validation of surgical Apgar score in all surgical services

The first part of the study compromised all surgical patients who met our inclusion criteria. The end-of-case surgical Apgar score was revalidated in our study cohort. The frequency of major complications within 30 days of surgery was computed and the distribution of the ASA classification was recorded. We calculated the odds of experiencing a major complication for every one-unit increase in surgical Apgar score by using binary logistic regression.

3.5.1 Real-time assessment of surgical Apgar score

The Surgical Apgar score was calculated each time new information (i.e. vital signs or electronic documentation of blood loss) was documented in the patient's electronic record (typically every 30-60 seconds). Specifically, as time advanced, we used all data available up to each time-point in which new data appeared to calculate the surgical Apgar score utilizing the lowest intraoperative HR and MAP. Whenever new data were subsequently recorded, the lowest intraoperative HR and MAP values were used for the calculation of the surgical Apgar score. As in Regenbogen *et al.* heart rate values outside the range of 20 to 200 beats per minute and mean arterial pressures outside of the range of 25 to 180 mmHg were interpreted as artifact and were discarded.[19] Points assigned for estimated blood loss were renewed each time EBL was updated and then were added to the score determined for the patient's heart rate and mean arterial pressure to compute the overall surgical Apgar score. On average, blood loss was recorded in 15-minute intervals during standard surgical procedures without extensive blood loss. Data on resuscitation (with fluids and blood products) was not included in the analysis.

Estimated blood loss is a subjective measure, manually recorded into the patient record. The amount is determined by adding the estimated amount of blood within the collection container to that absorbed within used collecting gauzes. Finally the surgeon is consulted to attest the estimated amount of blood loss. The sAs was computed every minute new data were available by summing the points for lowest recorded MAP, heart rate value, and EBL at that point in time. The last lowest value for heart rate and mean arterial pressure were used throughout a case. Points given for the estimated blood loss are added to the points given for HR and MAP. (Table 3)

Figure 1. Calculation of surgical Apgar score

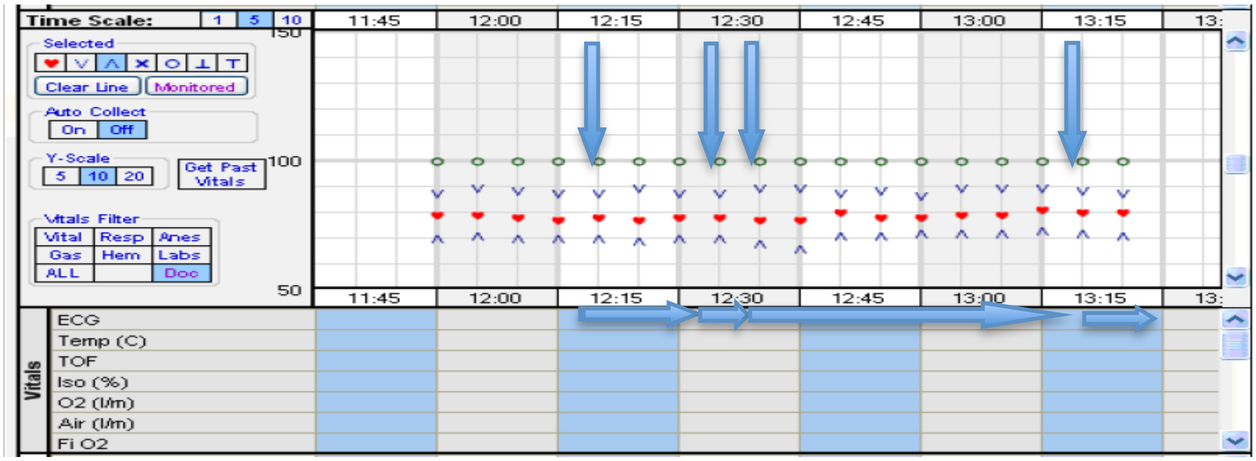


Figure 1 displays example intraoperative heart rates, blood pressure, and oxygen saturation at various time points within a surgical procedure. The arrows indicated what heart rate value was used to calculate the surgical Apgar score at certain time points.

3.6 Statistical Analysis

All statistical analyses were computed using SAS 9.3 statistical software package (SAS Institute, Cary, NC) and R version 3.0.3 (Vienna, Austria). Surgical Apgar scores were calculated for each patient as previously described[8] and examined for their association with the proposed postoperative complications. Univariate analysis was performed on age, gender, ASA class, surgical Apgar score, lowest heart rate, estimated blood loss, lowest mean arterial pressure, minutes spent in the operating room, complications, death and primary service. Given that the categorical nature of the independent variable was continuous and categorical, we chose to use logistic regression. In the second part of the study we categorized patients by their surgical Apgar score value 0 to 4, 5 to 6, 7 to 8, and 9 to 10. Scores ranging from 0 to 4 were merged due to small sample sizes in the lowest categories.

3.6.1 Bivariate statistical analysis

Pairwise analyses were performed to evaluate the association between demographic and operative patient characteristics and major postoperative complications within 30 days of surgery. P-values less than 0.05 were considered statistically significant. Chi^2 test was used to compare categorical data whereas one-sided t-test was used for continuous data. Chi^2 test was used to assess the association between each level of the surgical Apgar score and major postoperative complications within 30 days of surgery. As previously described[18, 109, 112] we chose patients with a surgical Apgar score of 7-8 as our comparison group.

3.6.2 Univariate logistic regression models

The study population was divided in two different cohorts. The first cohort consisted of patients who did not suffer from major complications or death while the second cohort was comprised of patients with complications or death within 30 days of surgery. Simple logistic regression was used to assess the association of end-of-case surgical Apgar score with investigated outcomes. P-values less than 0.05 were considered statistically significant. In order to determine if patients with low end-of-case surgical Apgar scores exhibited higher rates of major postoperative complications and mortality, univariate logistic regression was employed. Univariate logistic regression models were applied to determine the predictive ability of the sAs and the ASA classification.

3.6.3 Receiver operator characteristic curve

The Receiver operating curve (ROC curve) is an assessment of the predictive value of a test and is graphically represented by plotting sensitivity versus 1-specificity.(135) Sensitivity represents the proportion of true positives (i.e. patients who did indeed suffer from major complications) whereas the specificity is a measure of true negatives (i.e. patients who truly did not suffer from major complications).(136) The closer the Area under the curve (AUC) is to the value of one, the better is the discrimination.(136)An AUC of 0.5 cannot distinguish between patients with and patients without major complications and therefore marks a test as useless.(137) The AUC represents the probability of suffering from major complications for patients with lower surgical Apgar scores.(138) ROC curves were plotted and the area under the curve was calculated for sAs and the ASA classification.

3.6.4 Mulitvariate Logistic Regression

Multivariate logistic regression models predict outcome of a categorical variable that depends on multiple covariates or independent variables. We established logistic regression models to analyze the trends of surgical Apgar score throughout a surgical procedure in general, vascular and general oncology surgery.

3.6.4.1 Multivariate Logistic Regression for surgical Apgar score

We designed a multivariate logistic regression model to assess the relationship between low, intermediate, and high mean sAs, variation (high and low coefficient of variation) of sAs, and change (slope positive and negative) of sAs throughout the procedure and the likelihood of having a major complication.

A positive coefficient variation indicates that a high variability in surgical Apgar scores over time is significantly associated with mortality and morbidity. The more stable the sAs was during a surgery, the less likely was the patient to suffer from major complications. Patients with a mean sAs of four to seven were used as a reference. This statistical analysis was applied for all surgical patients included in this study.

3.6.4.2 Model for surgical Apgar score real-time assessment

Eight different models were constructed for general, vascular, and general oncology surgery, as described in Table 6. These models are denoted as follows: the ‘ASA,’ ‘sAs,’ ‘sAs and ASA,’ ‘HR,’ ‘HR and ASA,’ ‘MAP,’ ‘MAP and ASA,’ ‘HR, MAP, and ASA.’ The ‘ASA’ model uses the ASA classification assessed prior to the procedure, as well as the patient’s age, race, and gender to estimate the probability of major complications. The ‘sAs’ model uses continuously monitored surgical Apgar score and derived factors, age, gender, and race to predict adverse events. The ‘sAs and ASA’ model combines these two models. The ‘HR’ and ‘MAP’ models are similar to the ‘sAs’ model, substituting heart rate or mean arterial pressure for surgical Apgar score. These two models are combined with the ‘ASA’ model and displayed as the ‘HR and ASA,’ ‘MAP and ASA,’ and ‘HR, MAP, and ASA’ models. Interactions between the current value of continuously measured factors and each of the two associated derivates were also considered.

Table 6. Description of Prediction Models

Model	Description
‘ASA’	ASA class assessed prior to surgery adjusted for age, race, and gender
‘sAs’	Continuously monitored surgical Apgar score adjusted for age, race, and gender
‘sAs and ASA’	Combination of continuously monitored surgical Apgar score and ASA class assed prior to surgery adjusted for age, race, and gender
‘HR’	Continuously monitored heart rate, largest drop in heart rate from the baseline heart rate to the most recent recording, and the heart rate insult
‘HR and ASA’	Combination of continuously monitored heart rate, largest drop in heart rate from the baseline heart rate to the most recent recording, and the heart rate insult and ASA class assessed prior to surgery
‘MAP’	Continuously measured mean arterial pressure assessed during surgery, the largest drop in mean arterial pressure from baseline to the most recent recording, and the mean arterial pressure insult
‘MAP and ASA’	Combination of continuously measured mean arterial pressure assessed during surgery, the largest drop in mean arterial pressure from baseline to the most recent recording, and the mean arterial pressure insult and ASA class assessed prior to surgery
‘HR, MAP and ASA’	Combination of continuously monitored heart rate, largest drop in heart rate from the baseline heart rate to the most recent recording, and the heart rate insult, continuously measured mean arterial pressure assessed during surgery, the largest drop in mean arterial pressure from baseline to the most recent recording, and the mean arterial pressure insult, and ASA class assessed prior to surgery

Table 6 provides a detailed description of the eight risk models examined for their predictive ability in regard to postoperative complications.

3.6.4.3 Analysis of continuously monitored surgical Apgar score

Logistic regression models were used to intraoperatively update the risk of postoperative mortality or major complication, conditional on preoperative (i.e. baseline) and continuously monitored clinical factors, including heart rate, mean arterial pressure, blood loss, surgical Apgar score, and derived factors. At each intraoperative time-point, all continuously monitored factors were additionally summarized using two derived factors: the largest drop from baseline to the most current measurement, and a measure that we denote “insult.” Insult represents the cumulative drop in a continuously monitored factor from baseline to the current measurement (Figure 2). That is, when the current value of a continuously monitored factor is greater than or equal to the baseline value, the contribution to insult is zero. When the value is below baseline, the contribution to insult is positive. For example, if the heart rate falls 10 beats per minute (bpm) below the baseline value and remains there for 10 minutes, then the corresponding insult is 10 bpm × 10 minutes. An identical insult would occur if the decline in heart rate were 20 bpm and lasted for 5 minutes. The magnitude of insult may grow over the course of a procedure, but not shrink.

Logistic regression was implemented by assigning each patient’s outcome (one per patient) to every corresponding intraoperative record (multiple per patient).

Figure 2. Hypothetical Alert Model

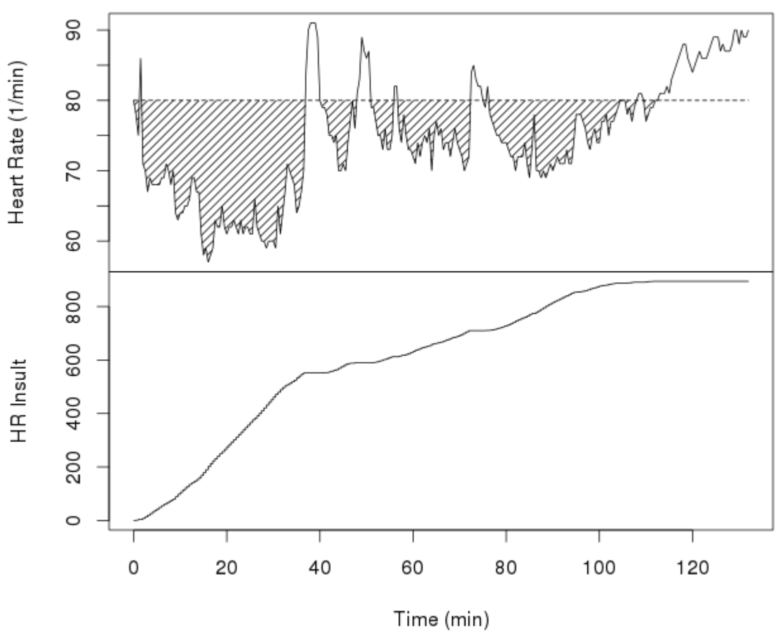


Figure 2. Illustration of an example heart rate (bpm) over procedure time (min) and the positive, cumulative contribution every drop in heart rate below the baseline value makes to HR insult. Shaded areas demark heart rate values below baseline during the procedure. The second graph illustrates how drops in heart rate cumulatively contribute to HR insult and thereby account for the depicted rise in HR insult over time.

3.6.5 Development of an Alert Model

Once the prediction models were fitted, we proposed a protocol for raising a hypothetical alert during a surgical procedure if a predetermined threshold probability for postoperative complications was reached.

Each risk model was used to evaluate a protocol for raising notifications intraoperatively. Based on the risk estimate at each intraoperative record, and for a sequence of threshold probabilities, we noted the procedure time at which the first notification would have been raised. Procedures where no notification was given were also noted. Box and whisker plots were used to display the times of first alert at various threshold probabilities. Since the 'ASA' model utilizes only preoperative information, an alert may only arise at the beginning of the surgical case. In contrast, the 'sAs and ASA' model may activate an alert at any time during the procedure. In particular, the current procedure duration, current surgical Apgar score and three derivatives of the surgical Apgar score history (baseline surgical Apgar score, maximum drop in score from baseline, and insult) were used for the prediction. Interactions between the current surgical Apgar score and each of the three derivatives were also considered.

The discriminative value of notification protocols was assessed for each model and surgical service by constructing receiver operating characteristic (ROC) curves and computing the area under the ROC curve (AUROC). AUROC values were compared among models in a pair wise manner by constructing a 95% confidence interval for the ratio (expressed as a percentage) of two AUROC values. The AUROC estimate in the general surgery cohort was internally validated using a bootstrap validation technique.[139]

We used this technique to test the “goodness-of prediction” of our model, the study population was replaced using 5,000 bootstrap replicates at the patient level, and the model was validated in this “new” cohort. In addition to these summaries of model discriminative value, calibration curves associated with the ‘sAs and ASA’ model are presented for each surgery service at procedure times 0, 60, and 120 minutes.

4 Results

4.1 First part: Surgical Apgar score in all surgical subspecialties

4.1.1 Patient Characteristics

Of 46,245 patients who fulfilled inclusion criteria, we obtained complete electronic intra-operative documentation on 33,108 patients (71.59%); this group comprised our final study cohort for the first part of the study. Demographic and intra-operative characteristics of the study population are displayed in Table 7. Patients with major complications are compared to patients without major complications. The overall incidence of major complications 30 days after undergoing a surgical procedure was 14.37%, corresponding to a total of 4,757 patients. Major complications included 342 (1.03%) deaths. Mean age of the study population was 51.0 years with patients suffering from major complications being on average one year older than patients without complications ($p<0.001$). Women represented 46.16% of the study population but made up 38.32% of patients with major complications in comparison to a 61.61% complication rate in the male population. Bivariate analysis showed that women were less likely to experience major complications ($p<0.001$). Patients with a higher American Society of Anesthesiologists' Physical Status Classification were associated with a significant increase of major complications ($p<0.001$). The lowest heart rate (65.84 versus 57.40) was significantly higher ($p< 0.001$) and the lowest mean arterial pressure (51.00 versus 52.27) was significantly lower among patients suffering from major complications ($p<0.001$). Patients with blood loss exceeding 950 ml were more likely to suffer from major complications ($p <0.001$). Longer operation duration (233.41 minutes versus 201.75 minutes) was significantly associated with major postoperative complications ($p<0.001$).

Table 7: Patient characterisitcs (n= 33,108) all surgical services

Variable	with major complications/death	without major complications/death	Total	P-value
Age: years ± SD	52 ± 16	51 ± 16	51 ± 16	< 0.001
Gender				< 0.001
Female	1,823 (38.32%)	13,461 (47.48%)	15,284 (46.16%)	
Male	2,931 (61.61%)	14,869 (52.45%)	17,800 (53.76%)	
Unknown	3 (0.06%)	21 (0.07%)	24 (0.07%)	
ASA Class				< 0.001
1	42 (0.88%)	1,323 (4.67%)	1,365 (4.12%)	
2	630 (13.24%)	12,629 (44.56%)	13,259 (40.05%)	
3	2,350 (49.40%)	13,028 (45.95%)	15,378 (46.45%)	
4	1,645 (34.58%)	1,344 (4.74%)	2,989 (9.03%)	
5	90 (1.89%)	27 (0.10%)	117 (0.35%)	
Lowest HR (beats/minute)	65.84	57.40	58.61	< 0.001
Lowest MAP (mmHg)	51.00	52.27	52.08	< 0.001
EBL	950.91 ml	306.76 ml	399.31 ml	< 0.001
Operation duration (minutes)	233.41 min	201.75 min	206.30 min	< 0.001
Number of patients (%)	4,757 (14.37%)	28,351 (85.63%)	33,108 (100%)	

Decreasing sAs were significantly associated with an increase of major complications (OR 0.62 (95 % CI: 0.61-0.63).

Table 8 displays the distribution of ASA class by gender. The majority of our study population was assigned to an ASA class of three. However, more male had an ASA class of three (47.46%) than the female patient population (45.33%). Most of the unknown gender had an ASA class of two (54.17%). Patients with an ASA class of five were less likely to undergo surgery and represented less than 1 % of the study population.

Table 8: Distribution of ASA classification by gender

ASA class	Female	Male	Unknown
1	717 (4.69%)	643 (3.61%)	5 (20.83%)
2	6,484 (42.42%)	6,762 (37.99%)	13 (54.17%)
3	6,928 (45.33%)	8,447 (47.46%)	3 (12.50%)
4	1,127 (7.37%)	1,859 (10.44%)	3 (12.50%)
5	28 (0.18%)	89 (0.50%)	0 (0.00%)

The five most common major complications in this study cohort were ventilator use for ≥48 hours (8.99%), wound disruption (2.28%), bleeding requiring ≥ 4 units of red blood cells (1.67%), organ space site infection (1.58%), and renal failure (1.35%) within 30 days of surgery. (Table 9)

Table 9. Frequency of major postoperative complications

	Frequency major complications n= 4,757 (14.37%)
Major complication	
Ventilator use ≥ 48h	2,977 (8.99%)
Wound disruption	755 (2.28%)
≥4 Units red blood cells	554 (1.67%)
Organ space side infection	524 (1.58%)
Renal failure	446 (1.35%)
Death within 30 days	342 (1.03%)
Sepsis	328 (0.99%)
Pneumonia	267 (0.81%)
Unplanned intubation	230 (0.69%)
SIRS	179 (0.54%)
Septic shock	165 (0.50%)
Graft failure	116 (0.35%)
Myocardial infarction	65 (0.20%)
Deep venous thrombosis	57 (0.17%)
Pulmonary embolism	56 (0.17%)
Stroke	55 (0.17%)
Cardiac arrest	21 (0.06%)
Coma ≥24h	21 (0.06%)

4.1.2 Association between surgical Apgar score and major complications

Table 10 displays the logistic regression results as well as test statistics that determine the overall significance of the logistic regression model. The odds of having a major complication are 0.62 times the odds of not experiencing a major complication for every one-unit increase in sAs. (OR 0.62 (95 % CI: 0.61-0.63).

Table 10. Binary Logistic Regression Analysis Results

Predictor	β	Standard Error	Wald χ^2	DF	p-value	OR (e^{β}, 95% CI)
Constant	1.16	0.06	386.96	1	<0.001	n/a
sAs	-0.48	0.01	2,379.41	1	<0.001	0.62 (0.61-0.63)

Model fit	χ^2	DF	p-value
Overall Model Evaluation Likelihood Ratio Test	2,572.00	1	<0.001
Score test	2,743.80	1	<0.001
Wald test	2,379.41	1	<0.001

DF= Degrees of Freedom

OR= Odds Ratio

4.1.3 Thirty-day Morbidity and Mortality by surgical Apgar score

In this study cohort patients were more likely to have an end of case sAs of 7 or 8. Among 2,640 patients with a score of 9 or 10, merely 153 patients (5.8%) suffered from major complications within 30 days. In comparison, 505 patients with a score of ≤2 279 patients (55.2%) experienced major complications. Patients who had a surgical Apgar score between 0-2 had 7.12 times (95% CI, 6.45-7.78) increase in risk to suffer from major complications than patients with a score between 7-8. Patients with an end sAs of 9 or 10 had on the other hand a 0.75 (95% CI, 0.64-0.88) chance of suffering from major complication than patients with a score of 7-8. (Table 11)

Table 11: Thirty-day morbidity and mortality by surgical Apgar score all surgical patients

	Surgical Apgar Score				
	0-2	3-4	5-6	7-8	9-10
Number of all Patients	505	3,201	10,644	16,118	2,640
Complications (%)	55.2	36.6	17.9	7.8	5.8
Relative Risk (95% CI)	7.12	4.72	2.30	Reference	0.75
	(6.45-7.78)	(4.40-5.10)	(2.16-2.47)		(0.64-0.88)

CI, Confidence Interval
Complications, number (%) of patients who suffered from major complications

4.1.4 Association between Surgical Apgar Score and Major Complications by Primary Service

Table 12 displays the distribution of patients within each primary service. For every unit increase of the sAs the odds of suffering from major complications decreased by almost 40% in General Surgery (OR, 95% CI: 0.571-0.679) but only decreased by less than 17% in Burn patients (OR, 95% CI: 0.756-0.920). In Ophthalmology only two deaths occurred during 30 days; thus there is a significant amount of uncertainty in the Odds Ratio.

Table 12. Odds Ratio and Major Complication by Primary Service

Primary Service	Total	number of complications	OR (95% CI)
Burn	708	235	0.834 (0.756-0.920)∗
Emergency	762	274	0.662 (0.601-0.730)∗
General oncology	1,795	255	0.584 (0.536-0.637)∗
General surgery	1,924	259	0.623 (0.571-0.679)∗
Gynecology	2,215	87	0.635 (0.555-0727)∗
Head and neck	1,453	197	0.719 (0.639-0.809)∗
Liver transplant	717	279	0.574 (0.519-0.635)∗
Neurosurgery	3,730	524	0.680 (0.638-0.724)∗
Ophthalmology	72	2	1.498 (0.448-5.006)
Oral	662	102	0.661 (0.571-0.766)∗
Orthopedic	3,570	316	0.703 (0.655-0.755)∗
Ortho sports/hand	1,109	84	0.616 (0.531-0.713)∗
Orthopedic trauma	4,056	726	0.652 (0.620-0.686)∗
Otolaryngology	955	25	0.509 (0.391-0.662)∗
Plastic Surgery	1,849	201	0.664 (0.601-0.733)∗
Renal	606	60	0.586 (0.496-0.693)∗
Thoracic	860	141	0.578 (0.511-0.653)∗
Trauma	1,271	591	0.771 (0.726-0.820)∗
Urology	3,785	151	0.588 (0.533-0.649)∗
Vascular	1,009	248	0.609 (0.554-0.670)∗

∗Odds Ratio are significant at the $\propto$ =0.05

4.1.5 Variation of surgical Apgar score throughout a surgical case in correlation to major complications

Table 13 shows the relative risk for major perioperative adverse events by categorized mean surgical Apgar score (sAs ≤ 4, 4<sAs ≤ 7, sAs>7). Out of all patients who had a complication, 11.50% had a mean sAs of ≤ 4 and 60.90% had an intermediate mean sAs (4<sAs ≤ 7). Patients who did not experience a complication, 56.35% had a high mean sAs (> 7). The risk of suffering from a complication among patients with a high mean sAs (sAs >7) is 61.52% lower than the risk of having an adverse event among patients with intermediate mean sAs (4<sAs≤7). The risk of experiencing a complication among patients with a low mean sAs (0≤sAs ≤4) is 58.87% higher than that of patients with an intermediate sAs. Patients with a lower mean sAs have an increased relative risk of 2.43 (95%, CI: 2.27-2.60) than patients with an intermediate sAs. On the other hand, patient with high mean sAs have a 0.38 (95%, CI: 0.36-0.41) increased risk in comparison to patients with an intermediate risk.

78.83% of patients who suffered from a major complication had a sAs decreased throughout the case (negative slope). 66.78 % of patients did not suffer from complication and less variation of sAs throughout the surgery (positive slope). The percentage of patients whose sAs got worse throughout their procedure and who had complications was significantly higher than to patients without outcome ($p<0.001$).

Patients who experienced postoperatively an adverse event, more than half (51.78%) had a coefficient of variation greater than 10 %. Out of all patients who did not suffer from any complication, only 25.97% had a greater than 10 % variation of their score during surgery. The greater variation in sAs is significantly associated with a higher risk of experiencing a complication.

Table 13: Relative Risk for Main Predictor

Variable	with major complications/death	without major complications/death	Total	Relative Risk
Mean sAs				
4	547 (11.5%)	593 (2.9%)	1,140 (3.4%)	2.43 (2.27-2.60)
$4 < sAs \leq 7$	2,897 (60.9%)	11,782 (41.6%)	14,679 (44.3%)	Reference
> 7	1,313 (27.6%)	15,976 (56.4%)	17,289 (52.2%)	0.38 (0.36-0.41)
Slope of sAs				
positive slope	1,007 (21.2%)	9,417 (33.2%)	10,424 (31.5%)	Reference
negative slope	3,750 (78.8%)	18,934 (66.8%)	22,684 (68.5%)	1.71 (1.60-1.83)
sAs CV				
high CV	2,463	7,362	9,825	2.54 (2.42-2.68)
low CV	2,294	20,989	23,283	Reference
Number of patients	4,757	28,351	33,108	

sAs = surgical Apgar scores
CV = coefficient of variation
95 % confidence intervals are provided below the relative risk values.

Table 14 displays a multivariate logistic regression model. In cases with a high coefficient of variation (CV >10%) throughout the procedure the odds of having a complication were 1.76 (95%, CI, 1.62-1.91) times the odds of patients who had a low coefficient of variation (CV≤ 10). Patients with mean sAs of less than four had 5.87 (3.63-9.50) the odds of having an adverse event than patient with an intermediate mean sAs. If the sAs decreased (negative slope) over the course of a surgery the patient had 1.47 times the odds of having a negative outcome than patients where the sAs increased (positive slope) over time. In our study the Hosmer-Lemeshow statistic is not significant, indicating that our model fits the data.

Table 14. Multivariate Logistic Regression Analysis Results

Predictor	β	Standard Error	Wald χ^2	DF	p-value	OR (e^{β}, 95% CI)
Constant	-1.96	0.04	2,358.34	1	<0.001	n/a
Slope	0.39	0.04	96.49	1	<0.001	1.47 (1.36-1.59)
CV	0.56	0.04	175.92	1	<0.001	1.76 (1.62-1.91)
Low mean sAs	1.77	0.25	51.88	1	<0.001	5.87 (3.63-9.50)
mean high sAs	-0.96	0.05	454.45	1	<0.001	0.38 (0.35-0.42)
cvhml	-0.83	0.25	10.52	1	0.001	0.44 (0.27-0.72)
cvhmh	0.22	0.08	7.86	1	0.005	1.25 (1.07-1.46)

Model fit	χ^2	DF	p-value
Overall Model Evaluation Likelihood Ratio Test	2,268.77	6	<0.001
Score test	2,514,70	6	<0.001
Wald test	2,106.86	6	<0.001
Hoshmer-Lemeshow	5.30	5	0.38

DF= Degrees of Freedom

OR= Odds Ratio

Cv= coefficient of variation with low cv (<10% as reference category)

χ^2= chi^2 test

Slope= reference category is positive slope

Interaction terms: **sncvh**= slope negative coefficient variation high; **cvhmh**=coefficient variation high mean high; **cvhml=** coefficient of variation high mean low; **sncvhml**=slope negative coefficient of variation high mean low; **sncvhmh**= slope negative coefficient of variation high mean high

Figure 3 demonstrates the ability of our logistic regression model to predict adverse events. Overall our model has an acceptable level of discrimination between patients with and without major postoperative complications with an area under the curve of 0.70.

Figure 3. ROC curve, assessing discriminatory performance using sensitivity and specifity parameters

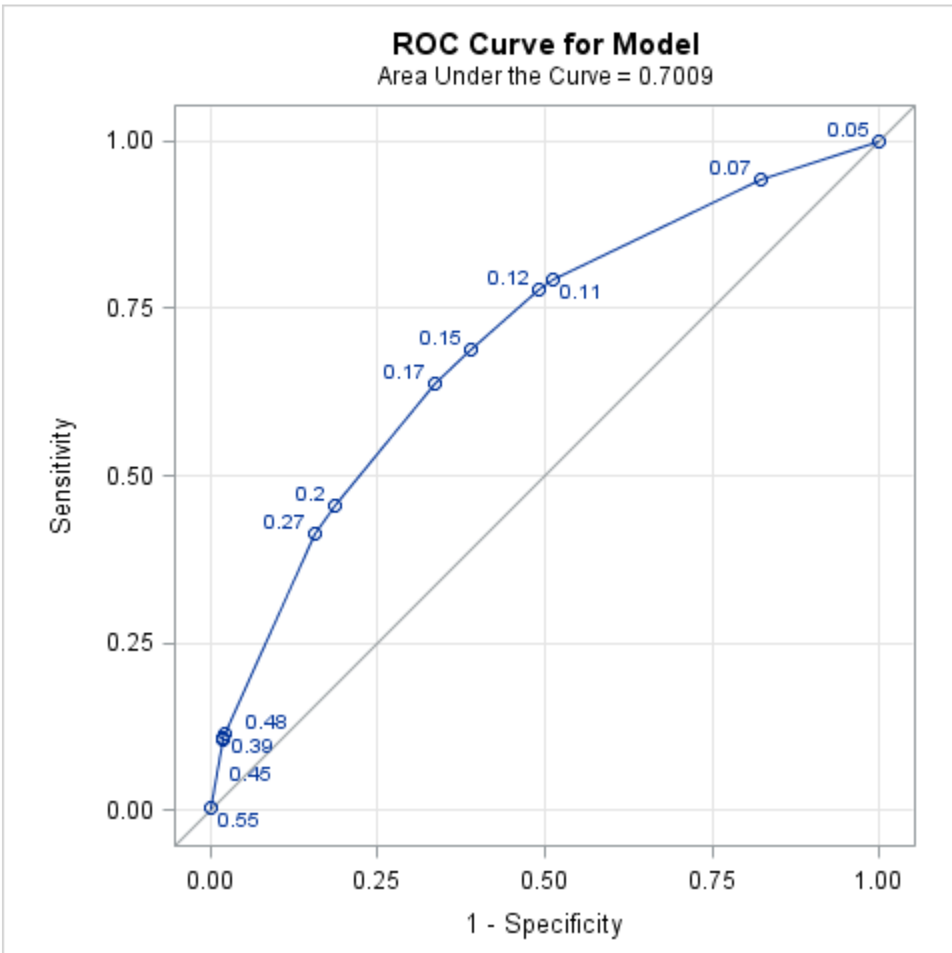


Table 15 summarizes odd ratios for different values of our interaction term variables. We performed affect modification to determine if the variables in our model are influenced by each other. Patients with an intermediate mean sAs and a low coefficient of variation were used as the reference group. Our initial model was the largest possible model incorporating sAs slope, mean sAs, and variation of sAs throughout the procedure and all possible interaction between each other.

We used backwards elimination to determine which interaction terms to keep in our final model. We dropped to least significant interaction term one at a time. Patients with sAs variation of >10% but a high mean sAs are less likely to experience major complications (OR 0.84, 95% CI, 0.74-0.96)

Table 15: Summary table of interaction assessment

	CV high	CV low
Mean low sAs	4.51 (3.94-5.17)*	5.68 (3.63-9.50)*
Mean intermediate sAs	1.76 (1.62-1.91)*	Reference
Mean high sAs	0.84 (0.74-0.96)*	0.38 (0.35-0.42)*

* $p < 0.001$
CV = coefficient of variation
sAs = surgical Apgar score

4.2 Second part: Surgical Apgar score in general, vascular, oncology surgery

4.2.1 Patient characteristic in general, vascular, and oncology surgery

In the second part of the study, we obtained complete electronic intraoperative data on 4,728 patients that fulfilled our inclusion criteria: 1,924 general surgery patients, 1,795 general oncology patients, and 1,009 vascular surgery patients. Demographic and intraoperative characteristics of the study population are displayed in Table 16, which compares patients with postoperative complications to patients without postoperative complications. The overall incidence of major complications within 30 days of surgery was 16.11% (95% CI, 15.08-17.20), corresponding to a total of 762 patients. Major complications included 71 deaths (1.50%, 95% CI, 1.17-1.89) within 30 days of surgery. Mean age of the study population was 55.0 years, with patients suffering from major complications being on average four years older than patients without complications (p<0.001). Women represented 53.39% of the study population and were less likely to experience major complications than men, 43.70% versus 56.04 %, p<0.001. Increased patient age, higher American Society of Anesthesiologists Physical Status Classification, and longer duration of the surgical procedure were associated with a statistically significant increase in adverse events following surgery (p<0.001). The lowest intraoperative heart rate was significantly higher (62 versus 56, p<0.001) and the lowest mean arterial pressure was significantly lower (49 versus 51, p<0.001) in patients with postoperative complications compared to patients without postoperative complications. Patients with blood loss exceeding 800 ml were significantly more likely to suffer from adverse events (p<0.001), as were patients with a lower surgical Apgar score (p<0.001). For every unit decrease in the surgical Apgar score, the univariate odds of having a major postoperative complication increase by 62% (OR 1.62; 95% CI, 1.59-1.65; p<0.001).

Table 16: Patient characteristics (n= 4,728) for general, vascular, and general oncology surgery

Variable	with major complications/ death	without major complications/ death	Total	P-value
Age, years mean ± SD	58 ±13.1	54 ± 14.7	55 ± 14.9	< 0.001
Gender				< 0.001
Female	333 (43.70%)	2,191 (55.24%)	2,524 (53.39%)	
Male	427 (56.04%)	1,775 (44.76%)	2,202 (46.57%)	
Unknown	2 (0.26%)	0 (0.00%)	2 (0.04%)	
Race				< 0.001
Caucasian	662 (86.88%)	3,491 (88.02%)	4,153 (87.84%)	
African American	64(8.40%)	335 (8.45%)	399 (8.44%)	
Other	10 (1.31%)	54 (1.36%)	64 (1.35%)	
Unknown	26 (3.41%)	86 (2.17%)	112 (1.89%)	
ASA Class				< 0.001
1	2 (0.26%)	58 (1.46%)	60 (1.27%)	
2	101 (13.25%)	1,485 (37.44%)	1,586 (33.54%)	
3	413 (54.20%)	2,217 (55.90%)	2,630 (55.63%)	
4	234 (30.71%)	205 (5.17%)	439 (9.29%)	
5	12 (1.57%)	1 (0.03%)	13 (0.27%)	
Lowest HR ±SD (beats/minute)	62 ± 17.6	56 ± 12.4	57 ± 13.3	< 0.001
Lowest MAP±SD (mmHg)	49± 12.3	51 ± 10.9	51 ± 11.1	< 0.001
EBL±SD	835 ml ± 1,608.0	240 ml ± 419.9	337 ml ± 782.2	< 0.001
Operation duration (minutes)	256.80 min	200.17 min	209.40 min	< 0.001
Number of patients (%)	762 (16.12%)	3,966 (83.88%)	4,728 (100%)	

Displayed are characteristics of the patient population with and without major postoperative complications within thirty days of surgery. Characteristics include mean age with standard deviation, number and percentage of patients of a specific gender, race and within each ASA class. Surgical data include average surgical procedure duration as well as mean lowest heart rate (HR), mean lowest mean arterial pressure (MAP) and mean estimated blood loss (EBL) with standard deviations. Pairwise analyses were performed to evaluate the associations between demographic and operative patient characteristics and major postoperative complications. SD, standard deviation.

4.2.2 Thirty-day Morbidity and Mortality by surgical Apgar score in general, vascular, and oncology Surgery

Out of 4,728 patients, 762 (16.11%) experienced one or more adverse events. The five most frequent major complications in this study population were ventilator use for more than 48h (27.86%), wound disruption (15.03%), deep or organ-space surgical site infection (12.58%), renal failure (10.13%), and sepsis (8.25%).

The association between major postoperative complications and various ranges of surgical Apgar scores in the second part of the study is illustrated in Table 17. Approximately half of the patients had an end-of-case surgical Apgar score of 7 or 8. Among 402 patients with a score of 9 or 10, only 32 patients (8.0%, 95% CI, 5.51-11.05) suffered from major complications within 30 days of surgery. In comparison, among 443 patients with a score of ≤ 4, 206 patients (46.5%, 95% CI, 41.78-51.27) experienced adverse events postoperatively. Patients who had an end-of-case surgical Apgar scores between 0-4 had almost a five times increased risk (Relative Risk [RR] 4.8, 95% CI, 4.4-5.6; $p<0.001$) of suffering from major postoperative complications compared to patients with a surgical Apgar score between 7-8. On the other hand, patients with an end-surgical Apgar score of 9 or 10 had a mildly decreased risk (RR 0.9, 95% CI, 0.6-1.3; $p<0.001$) of experiencing major complications compared to patients with a score of 7-8.

Table 17: Major complications within 30 days by end-of-case surgcial Apgar score for general, vascular, and oncology surgery

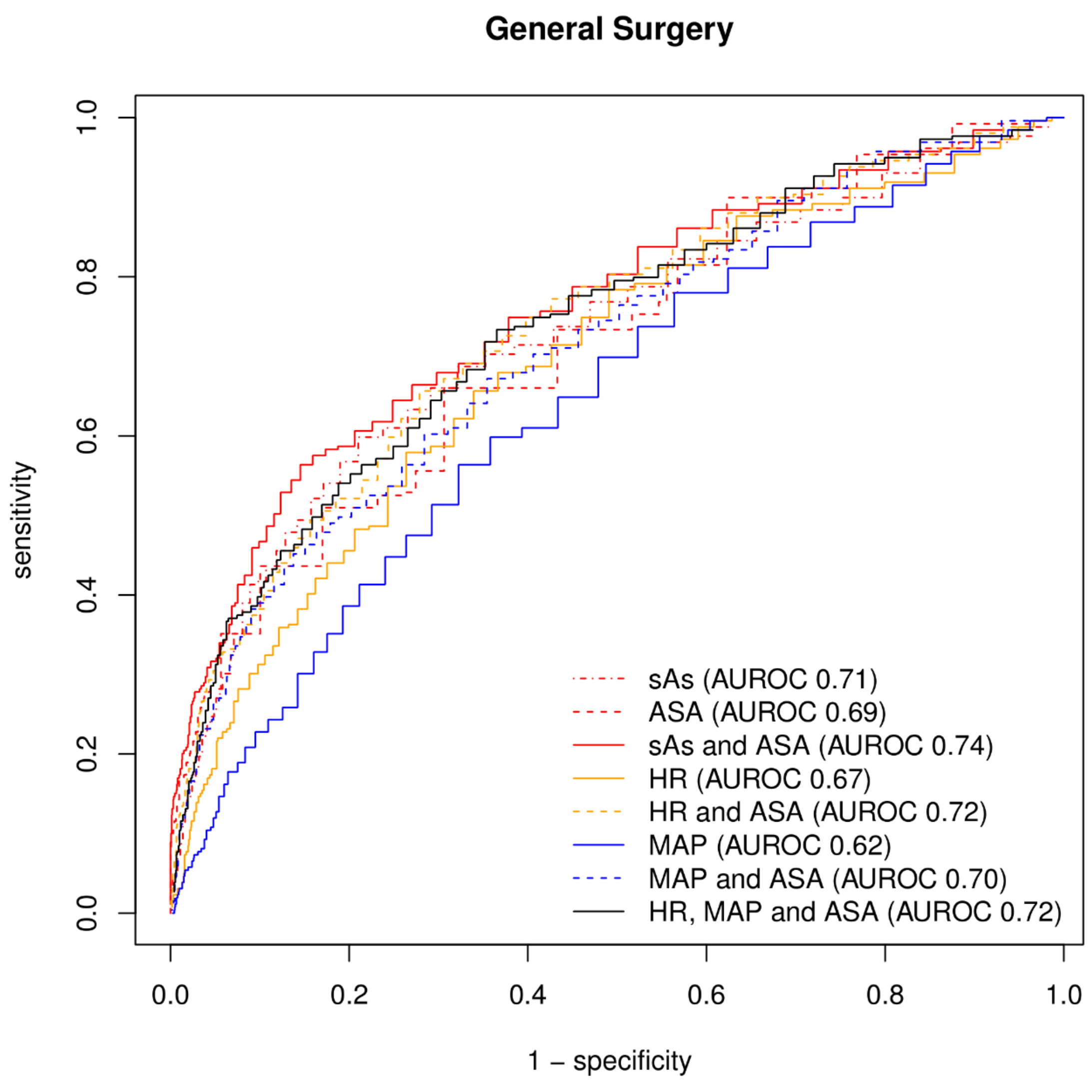
	surgical Apgar score			
	0-4	5-6	7-8	9-10
Number of all Patients	443 (9.4%)	1,422 (30.1%)	2,461 (52.0%)	402 (8.5%)
Complications (%)	206 (46.5%)	305 (21.4%)	219 (9.0%)	32 (8.0%)
Relative Risk (95% CI)	4.8 (4.1-5.6)	2.4 (2.1-2.8)	Reference	0.9 (0.6-1.3)

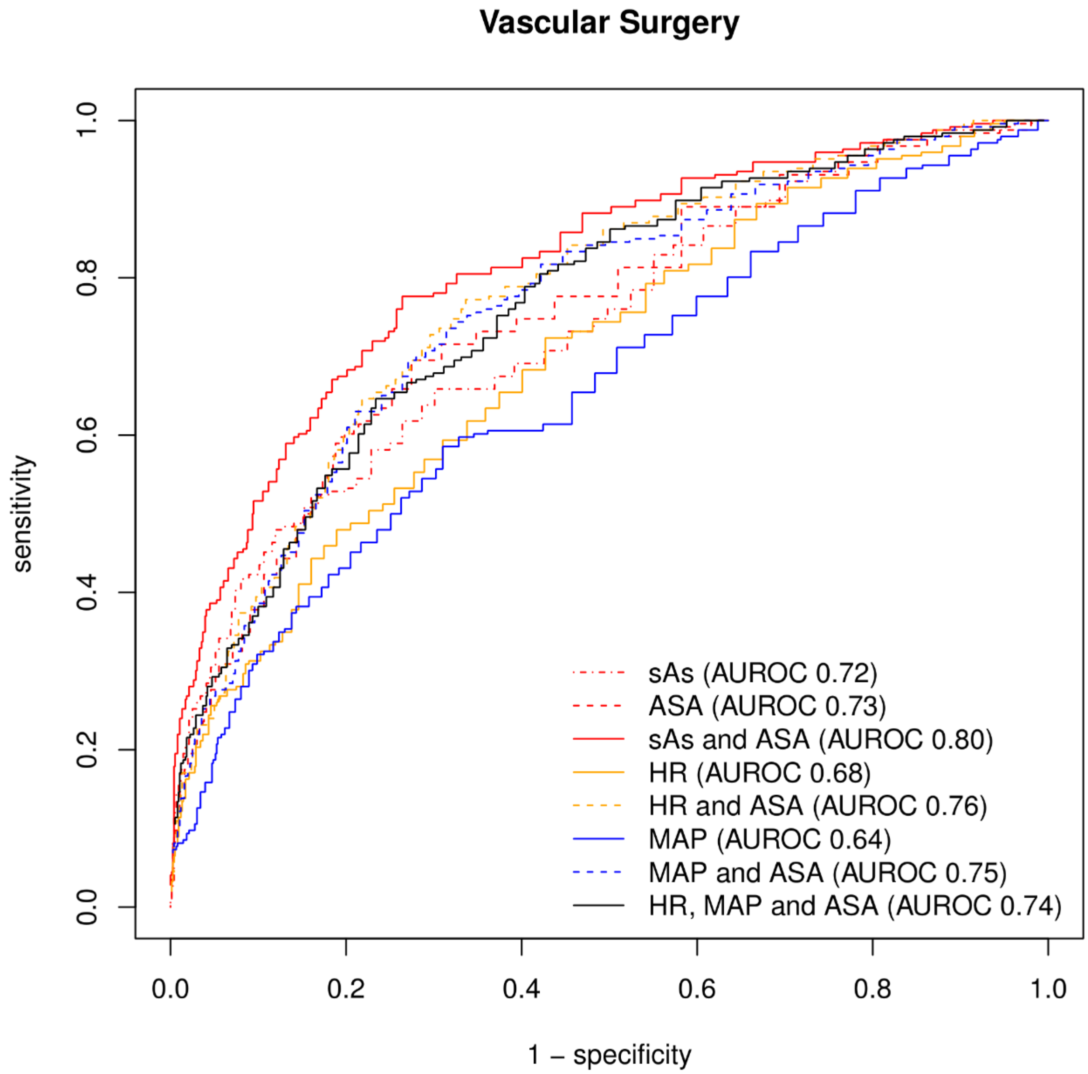
Table 17. lists the number and percentage of patients who experienced major complications as defined by the National Surgical Quality Improvement Program (NSQIP) registry within thirty days of surgery. Patients were grouped into four classes based on their end-of-case surgical Apgar score. Simple logistic logistic regression was used to assess the association of end-of-case surgical Apgar score with investigated outcomes. The relative risk of suffering from major complications was referenced to patients with an end-of-case score of 7-8, with 95 % confidence intervals.
CI, Confidence Interval

4.2.3 Correlation between the surgical Apgar score and American Society of Anesthesiologists physical status classification system

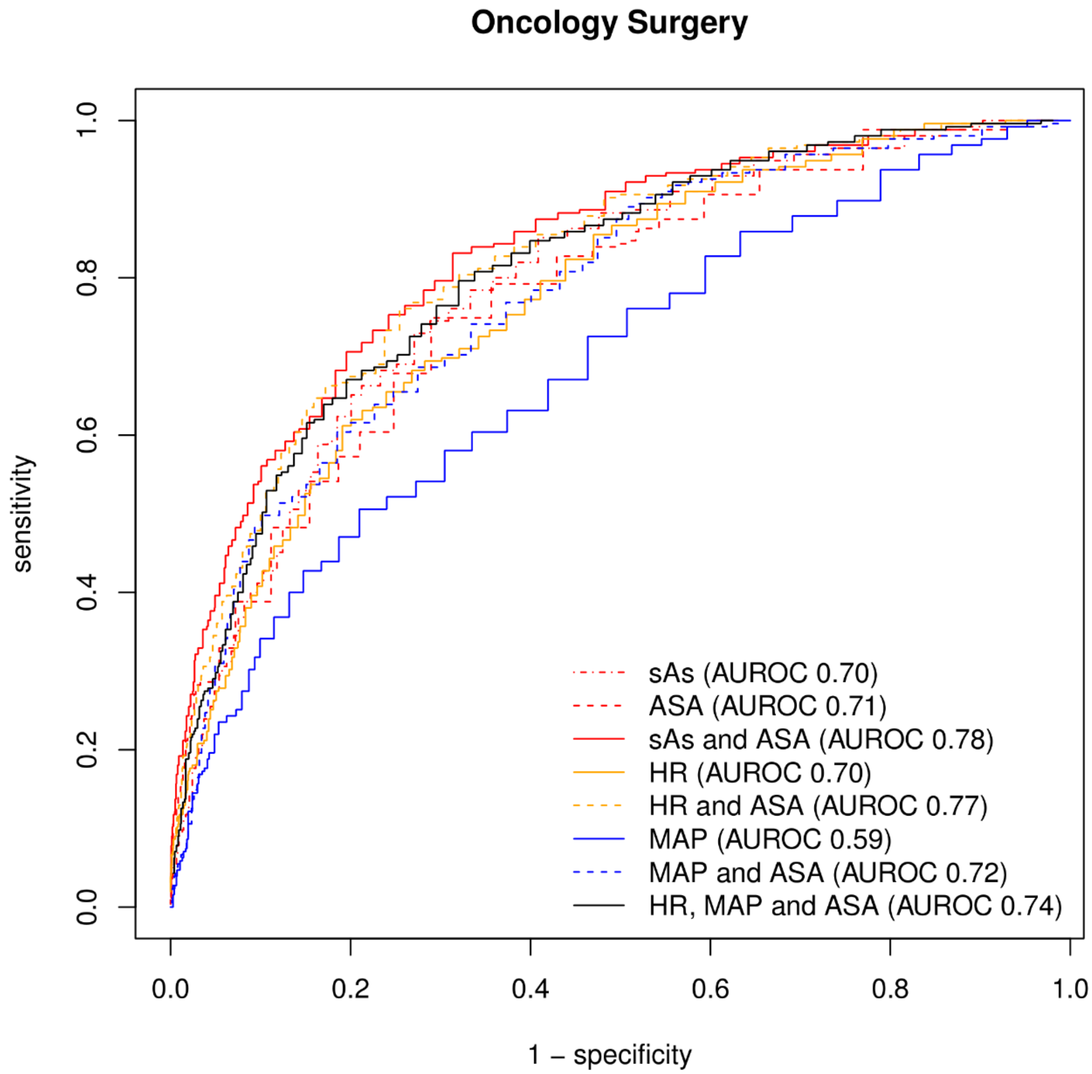
Figure 4 illustrates the ROC curves for each model and surgical service.

Figure 4. Receiver Operating Characteristic Curves for Three Surgical Subspecialties





Vascular Surgery
sensitivity
1 – specificity
0.0
0.2
0.4
0.6
0.8
1.0
sAs (AUROC 0.72)
ASA (AUROC 0.73)
sAs and ASA (AUROC 0.80)
HR (AUROC 0.68)
HR and ASA (AUROC 0.76)
MAP (AUROC 0.64)
MAP and ASA (AUROC 0.75)
HR, MAP and ASA (AUROC 0.74)


Oncology Surgery
sensitivity
1 − specificity
0.0
0.2
0.4
0.6
0.8
1.0
sAs (AUROC 0.70)
ASA (AUROC 0.71)
sAs and ASA (AUROC 0.78)
HR (AUROC 0.70)
HR and ASA (AUROC 0.77)
MAP (AUROC 0.59)
MAP and ASA (AUROC 0.72)
HR, MAP and ASA (AUROC 0.74)

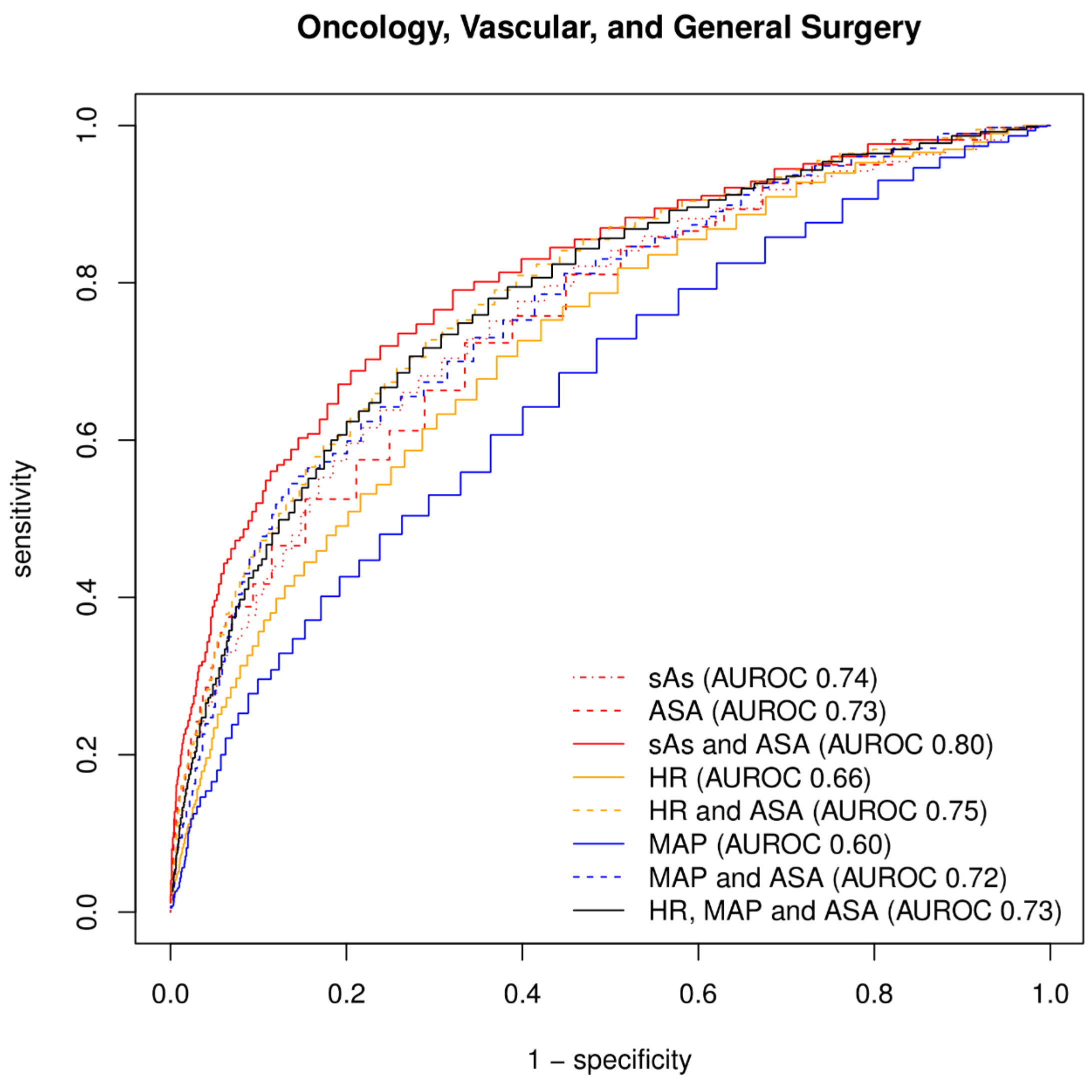


Figure 4. Receiver operating characteristic curves for each surgical service and for the pooled cohort were established for each risk model investigated. The discriminative value for each model is illustrated by the area under the receiver operating characteristic curve (AUROC) cited in the figure.

Table 18 summarizes the discriminative value of each model using the area under the receiver operating characteristic curve (AUROC), with 95% bootstrap confidence intervals. In general surgery, the AUROC for the 'ASA' model is 0.69 compared to 0.71 for the 'sAs' model. The 'HR and ASA' and the 'HR, MAP, and ASA' models demonstrated a slightly better predictive ability than the 'sAs' model, with an AUROC of 0.72 in both models, respectively. Nevertheless, the 'sAs and ASA' model results in a superior predictive ability, with an AUROC of 0.74. The bootstrap validated estimate of AUROC for the 'sAs and ASA' model in general surgery cases was 0.74, indicating very little 'optimism' due to model overfitting (Table 18).

Table 18. Summary of Areas under the Receiver Operating Characteristic Curve

	General	Vascular	Oncology	Combined
sAs and ASA	0.74 (0.73; 0.75)	0.80 (0.78; 0.81)	0.78 (0.76; 0.80)	0.80 (0.79; 0.80)
ASA	0.69 (0.68; 0.69)	0.73 (0.72; 0.74)	0.71 (0.69; 0.74)	0.73 (0.73; 0.74)
sAs	0.71 (0.69; 0.72)	0.72 (0.70; 0.74)	0.70 (0.65; 0.74)	0.74 (0.74; 0.75)
HR	0.67 (0.65; 0.69)	0.68 (0.65; 0.70)	0.70 (0.64; 0.74)	0.66 (0.62; 0.69)
HR and ASA	0.72 (0.70; 0.73)	0.76 (0.74; 0.77)	0.77 (0.73; 0.79)	0.75 (0.72; 0.77)
MAP	0.62 (0.59; 0.64)	0.64 (0.62; 0.66)	0.59 (0.55; 0.63)	0.60 (0.57; 0.63)
MAP and ASA	0.70 (0.68; 0.71)	0.75 (0.72; 0.76)	0.72 (0.69; 0.74)	0.72 (0.70; 0.74)
HR, MAP, and ASA	0.72 (0.70; 0.73)	0.74 (0.72; 0.76)	0.74 (0.71; 0.77)	0.73 (0.70; 0.75)

Table 18. summarizes the area under the receiver operating characteristic curve (AUROC) statistics for each model and surgical service, with bootstrap 95% confidence interval, demonstrating the discriminative value of each model.

In vascular surgery, the AUROCs for the 'ASA' and 'sAs' models are similar, with an AUROC of 0.73 and 0.72, respectively. Even in this specialty the combined 'sAs and ASA' model exhibits a better predictive ability, with an AUROC of 0.80 (95% CI: 0.78-0.81), which is superior to that of all other models. In general oncology the 'ASA' model predicts postoperative complications slightly more accurately (AUROC 0.71) than the 'sAs' model (AUROC 0.70), 'HR' model (AUROC 0.70), and the 'MAP' model (AUROC 0.59). Among all three subspecialities investigated in the study, the discriminative value of the 'sAs and ASA' model is the strongest in vascular surgery with an AUROC of 0.80. When all subspecialties were merged, the predictive ability of the 'sAs' model (AUROC 0.74) was slightly better than that of the 'ASA' model (AUROC 0.73). The 'sAs and ASA' model (AUROC 0.80) remained superior to the 'HR' model (AUROC 0.66), the 'HR and ASA' model (AUROC 0.75), the 'MAP' model (AUROC 0.60), the 'MAP and ASA' model (AUROC 0.72), and the 'HR, MAP, and ASA' model (AUROC 0.73). Calibration curves associated with the 'sAs and ASA' model for each surgical service at procedure times 0, 60, and 120 minutes (Figure 5) indicate acceptable calibration, especially for the most common risk values.

Figure 5. Calibration curves

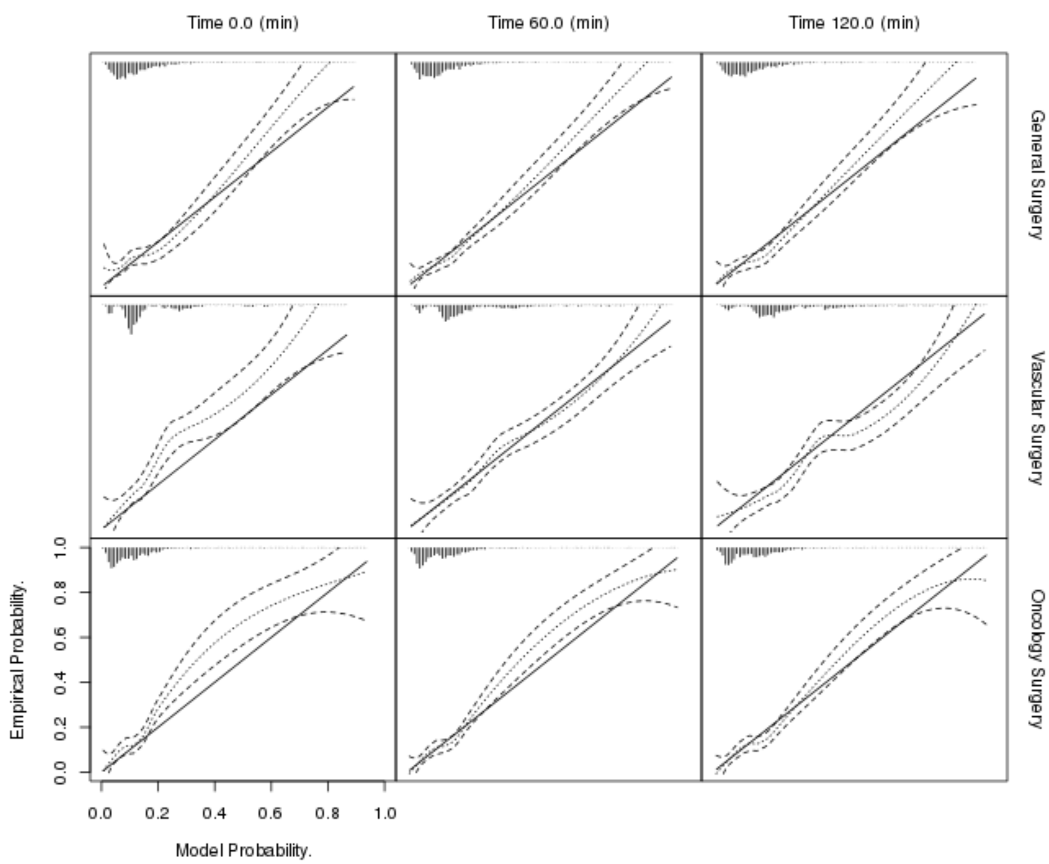


Figure 5. The calibration plot compares the model's predictive ability. The diagonal line represents the ideal situation. The curve corresponds to the relation nonparametrically. Calibration curves were constructed for each of the three surgical subspecialties at the beginning of each procedure, at 60 minutes, and at 120 minutes.

Table 19 describes the pairwise comparison of the eight risk models, using ratios (expressed as percentages) of the corresponding AUROCs. In particular, the AUROC for the combined 'sAs and ASA' model was improved by 7.9% (95% CI, 6.3-9.6) relative to the 'ASA' model in general surgery cases, 9.1% (95% CI, 7.1-11.0) in vascular surgery, 10.1% (95% CI, 6.4-13.5) in oncology surgery, and 8.5% (95% CI, 7.8-9.4) in the combined cohort. The AUROC for the combined 'sAs and ASA' model, relative to the 'sAs' model, was improved by 4.5% (95% CI, 3.6-5.5) in general surgery, 11.3% (95% CI, 9.4-

13.7) in vascular surgery, 12.0 % (95% CI, 7.9-17.2) in oncology surgery, and 7.2% (95% CI, 6.8-7.8) in the pooled cohort. In summary, in all three specialties included in this study and in the pooled cohort, combining ASA classification with surgical Apgar score trend analysis showed a superior discriminative value in comparison to either risk score alone. (Table 19)

Table 19. Model AUROC Comparison

General Surgery

	ASA	sAs	HR	HR and ASA	MAP	MAP and ASA	HR, MAP, and ASA
sAs and ASA	7.9 (6.3; 9.6)	4.5 (3.6; 5.5)	10.1 (7.5; 13.2)	3.0 (1.3; 5.0)	20.4 (16.4; 25.5)	6.0 (3.7; 8.6)	3.0 (0.7; 5.8)
ASA		-3.2 (-4.6; -1.2)	2.0 (-0.5; 5.0)	-4.5 (-6.3; -2.5)	11.5 (8.3; 16.0)	-1.8 (-3.8; 0.5)	-4.6 (-6.8; -2.0)
sAs			5.4 (2.9; 8.3)	-1.4 (-3.4; 0.7)	15.2 (11.3; 20.0)	1.4 (-1.0; 3.9)	-1.4 (-3.8; 1.3)
HR				-6.3 (-7.7; -5.3)	9.3 (5.4; 13.9)	-3.7 (-6.4; 0.8)	-6.5 (-8.8; -4.0)
HR and ASA					16.9 (12.7; 22.0)	2.9 (0.4; 5.6)	0.0 (-2.1; 2.4)
MAP						-11.9 (-14.3; -10.3)	-14.4 (-17.1; -12.2)
MAP and ASA							-2.8 (-4.4; -1.3)

Vascular Surgery

	ASA	sAs	HR	HR and ASA	MAP	MAP and ASA	HR, MAP, and ASA
sAs and ASA	9.1 (7.1; 11.0)	11.3 (9.4; 13.7)	18.6 (14.9; 23.0)	5.9 (3.6; 8.3)	24.5 (20.4; 29.8)	7.5 (4.5; 10.5)	8.0 (4.6; 11.3)

ASA		2.1 (-1.2; 5.6)	8.8 (5.5; 12.9)	-2.9 (-5.0; -0.42)	14.1 (10.3; 18.6)	-1.4 (-4.2; 1.4)	-1.0 (-4.0; 2.3)
sAs			6.6 (2.7; 11.0)	-4.8 (-8.1; -1.7)	11.9 (7.7; 17.1)	-3.4 (-7.1; -0.1)	-3.0 (-6.8; 0.5)
HR				-10.7 (-12.8; -9.0)	5.0 (-0.1; 10.0)	-9.3 (-13.2; -5.6)	-9.0 (-12.2; -6.0)
HR and ASA					17.5 (13.3; 22.3)	1.5 (-1.4; 4.5)	1.9 (-0.6; 4.6)
MAP						-13.6 (-15.9; -11.4)	-13.3 (-15.8; -10.7)
MAP and ASA							0.4 (-1.8; 2.7)

Oncology Surgery

	ASA	sAs	HR	HR and ASA	MAP	MAP and ASA	HR, MAP, and ASA
sAs and ASA	10.1 (6.4; 13.5)	12.0 (7.9; 17.2)	11.6 (6.2; 19.9)	2.0 (-1.3; 6.5)	32.6 (25.5; 40.4)	8.8 (4.7; 13.2)	5.2 (1.5; 9.7)
ASA		1.7 (-3.2; 8.2)	1.3 (-4.4; 10.0)	-7.4 (-11.0; -2.6)	20.4 (15.1; 26.8)	-1.2 (-4.7; 3.2)	-4.4 (-8.5; 1.1)
sAs			-0.3 (-6.0; 6.0)	-8.8 (-14.3; -3.3)	18.5 (10.0; 25.9)	-2.7 (-9.7; 3.1)	-6.0 (-12.2; -0.9)
HR				-8.5 (-13.0; -5.4)	19.0 (9.6; 27.4)	-2.4 (-9.9; 4.2)	-5.6 (-11.5; -0.8)
HR and ASA					30.1 (21.8; 38.2)	6.7 (1.9; 11.7)	3.2 (-0.1; 6.5)
MAP						-17.9	-20.6

						(-21.4; -14.5)	(-24.8; -16.2)
MAP and ASA							-3.3 (-6.7; 0.5)

Combined Surgery

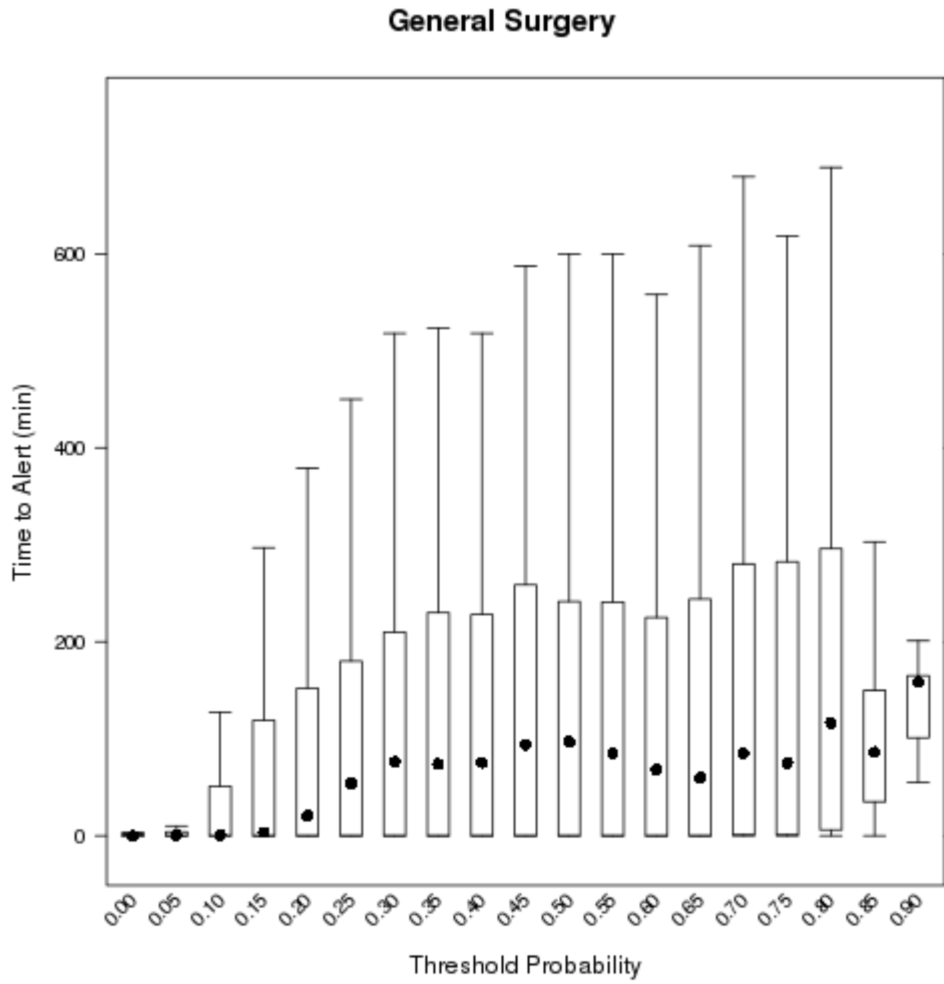
	ASA	sAs	HR	HR and ASA	MAP	MAP and ASA	HR, MAP, and ASA
sAs and ASA	8.5 (7.8; 9.4)	7.2 (6.8; 7.8)	20.0 (14.4; 28.1)	6.9 (4.0; 10.8)	33.2 (27.7; 40.5)	10.8 (7.9; 14.5)	9.1 (5.7; 13.0)
ASA		-1.2 (-2.2; -0.3)	10.5 (5.4; 17.8)	-1.6 (-4.3; 2.4)	22.8 (17.7; 29.5)	2.0 (-0.6; 5.6)	0.5 (-2.7; 4.2)
sAs			11.9 (6.4; 19.3)	-0.3 (-3.2; 3.1)	24.3 (18.9; 31.1)	3.3 (0.5; 6.8)	1.8 (-1.5; 5.6)
HR				-10.9 (-14.3; -8.5)	11.1 (4.7; 17.3)	-7.6 (-13.3; -3.0)	-9.0 (-13.7; -5.2)
HR and ASA					24.7 (19.1; 30.6)	3.7 (0.2; 7.6)	2.1 (-0.6; 5.0)
MAP						-16.9 (-19.8; -14.2)	-18.1 (-21.1; -14.9)
MAP and ASA							-1.5 (-3.6; 1.2)

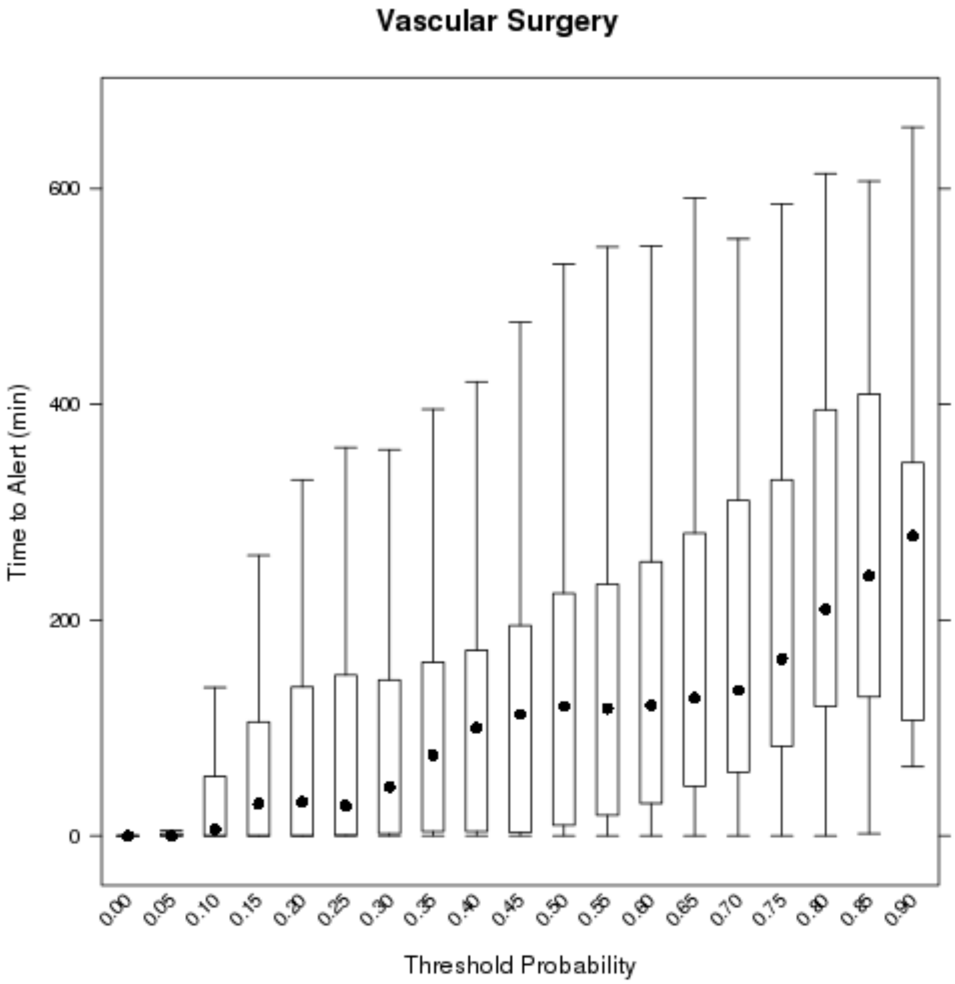
Table 19 presents the ratio of the corresponding AUROC values (row/column) as a percentage, with bootstrap 95% confidence interval. Positive values indicate that the area under the receiver operating curve corresponding to the model listed in the row header was larger than that in the column header.

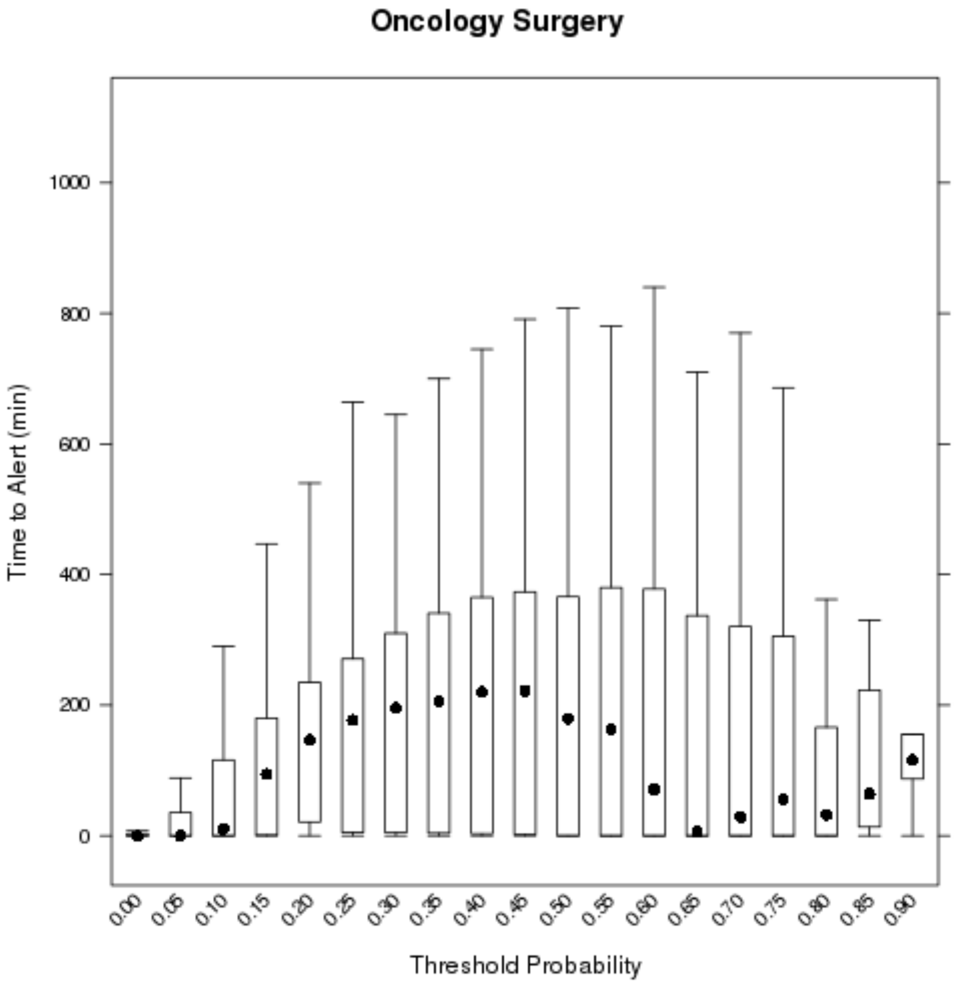
4.2.4 Notification model

In order to determine when real-time monitoring surgical Apgar score trends might raise an alarm in response to acute changes in a patient’s surgical Apgar score at various threshold probabilities for postoperative complications, we calculated the first time when a notification would have been raised for each case in our patient population. Figure 6 displays box and whisker plots for the times to first hypothetical intraoperative notifications, at increasing threshold probabilities. At high threshold probabilities, alerts are raised later during the surgical procedure. Conversely, alerts are activated earlier in the case when the threshold probability is set low since with this low tolerance for postoperative complications the threshold is exceeded more readily. A notification criterion based exclusively on the ASA classification would have raised a notification either at the very beginning of the procedure or not at all since the ASA classification utilizes only preoperative information that does not change throughout a procedure. In contrast, utilizing continuously monitored intraoperative factors in addition to ASA classification enables notification at any time during the surgical procedure. We believe that this feature is largely responsible for gains in sensitivity and specificity. Outliers in Figure 6 may be attributed to a paucity of patients who had a high probability of experiencing adverse events. Although threshold levels for notification should be optimized and validated in a prospective manner, our findings indicate that a threshold probability of 0.24 would exhibit acceptable specificity (0.85) and sensitivity (0.53) while maximizing clinical utility and avoiding premature activation.

Figure 6. Whisker Plots for Hypothetical Notification Models




Vascular Surgery
Time to Alert (min)
0
200
400
600
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
0.75
0.80
0.85
0.90
Threshold Probability



Oncology Surgery
Time to Alert (min)
0
200
400
600
800
1000
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65
0.70
0.75
0.80
0.85
0.90
Threshold Probability

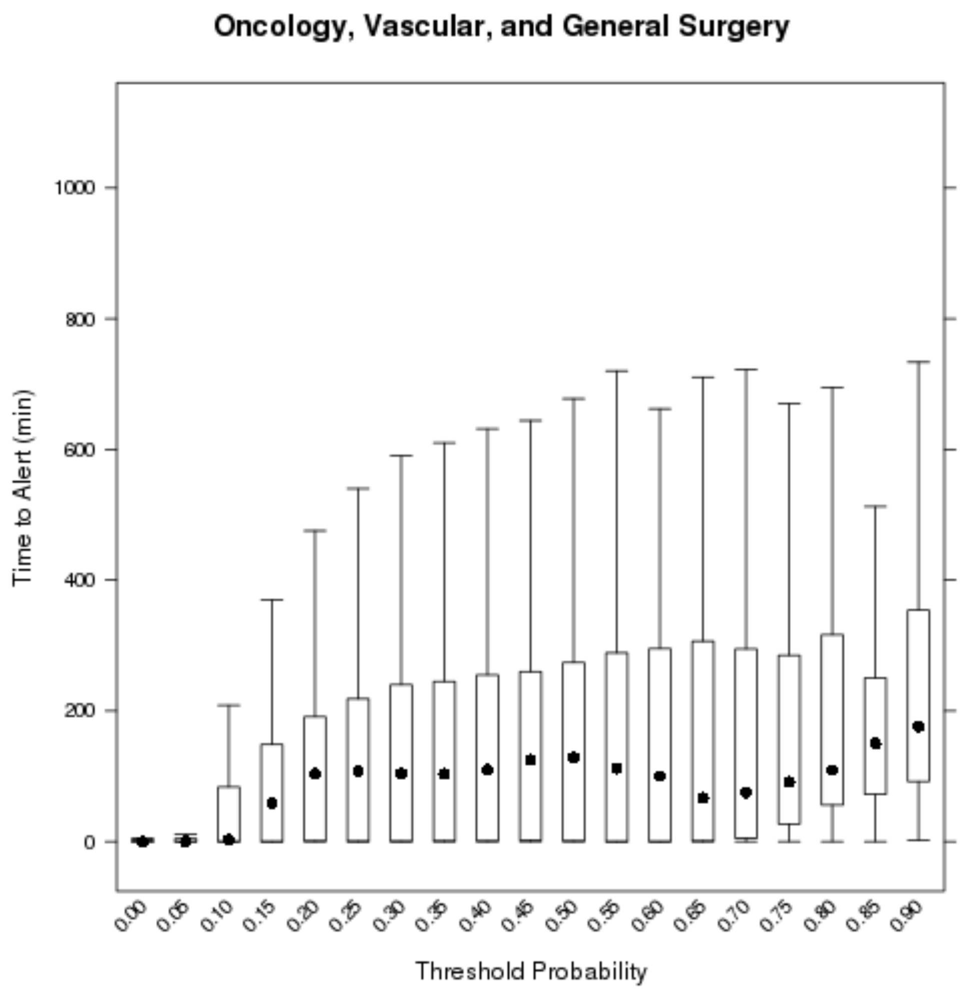


Figure 6. Displayed are Box and whisker plots for the times to first hypothetical intraoperative notification, at increasing threshold probabilities for each surgical service and the pooled cohort.

5 Discussion

5.1 Implications

In this study we found that continuous surgical Apgar score measurement combined with static demographic and clinical patient data can provide useful information about acute changes in a patient's status and his/her risk for postoperative complications in general, vascular, and general oncology surgery. In our study this form of postoperative risk assessment was superior to conventional methods of risk stratification. The model was generated in only three surgical services in order to investigate the predictive ability of the continuous assessment of sAs for adverse events. General surgery and vascular surgery were chosen for our study since the surgical Apgar score had originally been validated in these subspecialties. General oncology was chosen due to the large sample size of this patient cohort.

Our findings were consistent with previous studies, as we observed that a lower surgical Apgar score was associated with a patient's risk of suffering from major postoperative complications within 30 days following surgery.[8, 19, 106] The score had originally been derived from general and vascular surgery cases,[8] and here we have demonstrated that the predictive ability of both morbidity and mortality of the surgical Apgar score remains valid in an unselected surgical population.

In 2011 Reynolds *et al.* demonstrated that the predictive ability of the surgical Apgar score varies among different surgical subspecialties.[120] The score has proven to be applicable in all different specialties.[120] However, by calculating the end-of-case sAs in 20 different surgical subspecialties, we demonstrated that the association between surgical Apgar score and major complications varies among the subspecialties, with Ophthalmology being the only not significant primary service. Due to only two incidences of major complication within this service, this leads to a significant amount of vagueness in the Odds Ratio. Odds Ratios are ranging from 0.812 to 0.513. The difference of the predictive ability of the sAs between the subspecialties might occur due to various reasons, such as diverse comorbidities and different causes of death within each surgical service.[115, 120]

High intra-operative heart rate values, lower mean arterial pressure, and large amount of intra-operative blood loss lead to a lower surgical Apgar score. In our study population 55.2% of high-risk patients with a low end of case surgical Apgar score of less than 2, a total of 279 patients suffered from major complications. In contrast patient with a sAs greater than 9 only 5.8% of patients would suffer from major complications. With increasing level of sAs the percentage of major complication would decrease. Likewise with decreasing level of surgical Apgar score the rate of major complication would increase.

After the validation of the end-of-case sAs in this study population, we investigated if surgical Apgar score changes throughout a surgical case. At first, we calculated the mean sAs and determined its association to major complications. Unexpectedly, the majority of patients who encountered a major complication had a mean sAs between 4 and 7. This result demonstrates that the mean sAs might not be an accurate predictor of major complication. Subsequently, we evaluated how changing sAs over time influences the risk of suffering from major complications. Surprisingly, the Odds Ratio for patients with less variation in their sAs and with a low mean sAs was greater than in patients with a great variation and low mean sAs. This finding was interesting as the risk for major complication increased among patients with a low mean sAs. The results show that patients with great variation or drastic drop in their surgical Apgar scores have a greater risk of suffering from major complication. On the other hand, patients with relatively stable sAs are less likely to experience adverse events. Some patients with lower end-of-case surgical Apgar scores did not have a major complication. This evaluation of sAs trends does not raise a notification or suggest when a clinician should intervene. Systematic changes in surgical Apgar score throughout a surgical procedure, as opposed to random or periodic fluctuations, were inversely associated with major complications.

In order to improve the predictive ability of the continuously measured surgical Apgar score, we decided to calculate the sAs, whenever new data is available and investigate the predictive ability of the continuous surgical Apgar score. We have demonstrated that calculating the surgical Apgar score at various time points throughout the procedure, determining the magnitude of drop in surgical Apgar score from the previous measurement, and quantifying the time spent at that lower level are all predictive of postoperative morbidity and mortality within 30 days after surgery. Furthermore our study extends previous work by suggesting that continuous monitoring of surgical Apgar score trends in the surgical Apgar score model is useful in predicting postoperative adverse events and might allow clinicians to intervene early when significant changes in a patient's surgical Apgar score are noted intraoperatively. We validated continuous surgical Apgar score assessment in general, vascular, and general oncology surgery and compared it to the widely used ASA classification. Although we found ASA classification to accurately predict major postoperative complications, continuous surgical Apgar score monitoring by itself had similar predictive ability. There may be some advantage to the surgical Apgar score if one had to use only one metric, in that it is assigned objectively and can be tracked dynamically, as the case progresses.

We have shown that a continuously monitored surgical Apgar score and its associated derivative factors improve on the discriminative value of ASA classification alone, in regard to postoperative morbidity and mortality within 30 days after surgery. Combining the ASA classification with the surgical Apgar score into the 'sAs and ASA' model improves predictive ability in all three surgical specialties, demonstrating the highest AUROC of 0.80 in vascular surgery. Even when data from all three subspecialties were pooled, the predictive ability was comparable to that of each subspecialty examined separately. Additionally we compared the discriminative value of intraoperatively monitored HR and MAP, both components of the surgical Apgar score, to that associated with ASA classification alone. The 'sAs and ASA' model predicts major complications slightly more accurately than the 'HR and ASA,' the 'MAP and ASA,' and the 'HR, MAP, and ASA' models. We used bootstrap validation to verify that there was little 'optimism' in our estimate of the AUROC associated with the 'sAs and ASA' model in general surgery. Lastly, we identified various time-points throughout the procedure at which the first notification would be raised for a sequence of threshold probabilities.

A variety of measures and intraoperative variables are used to assess a patient's condition during a surgical case, providing clinicians with an overabundance of perioperative data elements.[140] Owing to this large amount of information, real-time notifications have been introduced to analyze data elements, enable trend detection, and alert clinicians to abnormal values or a patient's deteriorating condition.[140] Alerting systems using anesthesia information management system data have been implemented into hospital settings. These systems have been validated to improve compliance with prophylactic antibiotic administration,[141] support clinicians decision on suitable dispensation of antiemetic medication,[142] and the improvement of electronic anesthesia documentation.[143]

Real-time notifications draw the clinicians' attention to acute changes that may warrant timely intervention to improve patient outcome. Our altering model could be used to identify patients who are at a higher risk of experiencing adverse events and are in the need of more postoperative monitoring by experienced clinicians. Providing real-time notifications in anesthesia has proven to reduce hospital costs, improve patient care, and prevent postoperative complications.[140, 144] Real-time notifications about changes in a patient's surgical Apgar score trend will motivate providers to better allocate postoperative resources driven by the patient's tailored, acute physiology. Consequently appropriate resources can be used to improve patient care and this might lower complication rates.

Ghaferi *et al.* investigated the variation of mortality rates of patients who experience postoperative complications across different hospitals. This study suggests that hospitals with higher mortality rates were unable to rapidly detect and manage complications once they occurred.[145] This study exemplifies the fallibility of clinicians' decision-making. This study shows the need of an objective risk assessment that could be used as a communication aid among medical staff regarding the patient's postoperative level of care and severity of illness. The surgical Apgar score could support clinicians in their decision-making process and allowing a better use of resources for early identification of postoperative complications.

The overall goal of this study was to extend the utility of the Surgical Apgar score by developing a new, easily computable tool that can support clinical decision making in the operating room based on real-time changes in patient status during surgery. By extending utility of the surgical Apgar score from a single value computed at the conclusion of a surgical procedure to an objective tool that monitors sAs throughout the perioperative process, we expect that this implementation will affect clinical decision-making and behavior in the operating room. A continuous assessment of the surgical Apgar score may represent an objective tool to aid clinicians to support their request for additional diagnostic testing, invasive patient monitoring, and ordering transfer to the intensive care unit and modifying medical care.

We hypothesize that ultimately a number of more timely interventions in response to changing sAs over time may reduce perioperative morbidity and mortality.

5.2 Comparison to other work

Our findings were consistent with previous studies, as we observed that surgical Apgar score could accurately predict a patient's risk of suffering from major postoperative complications within thirty days following surgery.[8, 19, 106] The score had originally been derived from general and vascular surgery cases [8] and here we have demonstrated that the predictive ability of both morbidity and mortality of the surgical Apgar score remains valid in an unselected surgical population. The score is adept in identifying those patients who are more likely to experience major complications in the postoperative period and may thus represent an excellent tool to help guide clinical decisions.

In a recent study, Hyder *et al.* examined the predictive ability for major complications of surgical Apgar score assessed at various time points within general and vascular surgery. [146] Additionally, the sAs was compared to the ASA classification. Five models were established for the calculation of the sAs: The score was determined instantaneously, at an interval of 5 and 10 minutes (0-5 minutes, 6-10 minutes, etc., or 0-10 minutes, 11-20 minutes, etc.), and overlapping windows of 5 and 10 minutes (0-5 minutes, 1-6 minutes, etc., or 0-10 minutes, 1-11 minutes, etc.). The result was an increase in sampling intervals of sAs calculation improved the predictive ability of the score for postoperative complications and caused a better model discrimination. The analysis was not applicable to mortality, as the number of deaths was insufficient.[146]

Our approach differs somewhat from the work of Hyder *et al*. We investigated the direction of changing sAs during a surgical procedure by determining the slope of mean surgical Apgar score. The result of this investigation indicated that the direction of the slope was significantly associated with perioperative complications. We included more than 4,000 patients and also validated continuous assessment of the sAs in general oncology surgery. Thus, our analysis demonstrated statistical significance in predicting major complications, including death. One strength of our analysis is the combination of continuously monitored sAs with the ASA classification, which demonstrates superior discriminative value than either score alone. We assessed the predicative ability for major complications of continuously measured sAs combined with the ASA classification by using AUROCs and concluded that our model performs well. Lastly, we extended previous studies by establishing a notification model based on risk estimates.

5.3 Study limitations

Our study is limited by the following factors: First, this is a single center study conducted at a major academic institution and is restricted to an adult patient population undergoing surgery under general anesthesia. The notification protocol was only based on smaller patient cohort who underwent general, vascular, and general oncology surgery. Considering that the Vanderbilt University Medical Center draws patients from a wide geographic radius, postoperative complications of some patients who sought postoperative care elsewhere would not have been captured in the study if these patients were lost to follow up with our clinicians. However, 95.2% of the patient cohort included in the study had a follow-up visit at this hospital.

Extensive intra-operative blood loss is known to be associated with major postoperative complications.[21] Therefore an accurate measurement of blood loss should be granted within each surgical procedure. Until today there is no gold standard in assessing intra-operative blood loss although there are many different objective and subjective methods. Mostly the visual estimation of intra-operative blood loss underestimates the actual amount.[147] However, the visual estimation is free of cost, fast, and possible at any time during the procedure.[123] The estimation of intraoperative blood loss has often been criticized to be imprecise and may present a major limitation to the surgical Apgar score.[8, 19] However, the original authors argue that the amount of blood loss needed to calculate the score falls into a wide enough range based on the surgical Apgar score definition to render accurate assessment of intraoperative blood loss possible.[8] The broad grouping of EBL to calculate the sAs (0-100ml, 101-600 ml, 601-1,000ml, >1,000 ml) are simply within the observer's range of estimation.[123, 124] Another limitation of the study is the exclusion of all patients with only one estimated blood loss recording greater than 100 ml at the end of a surgical procedure. However, this allowed us to avoid sudden artificial changes of the surgical Apgar score trends at the end of a case where blood loss was not documented as it occurred throughout the procedure. With improved documentation practices and real-time recording of blood loss, the lack of EBL documentation will become a self-correcting problem.

Electronic monitoring plays a major role in patient anesthesia care and treatment decision-making process. Basic electronic monitoring of a patient has proven to reduce intra-operative mortality and adverse events.[148] However, intra-operative data artifacts might occur due to various reasons. Data artifacts in heart rate and blood pressure readings may occur due to external interventions. If an arterial line is flushed, blood is drawn from the access or during pressure line zeroing falsely high values might appear. Moving a patient may also influence heart rate recordings. Non-invasively measured blood pressure might also be influenced by a not appropriate size cuff, false positioning of the cuff, leaking cuff, and many more factors.[149]

Although clinicians are required to correct artificial recordings, flawless documentation is not guaranteed. Momentarily clinicians determine whether a vital recording is artificial or correct. Much research has to be done to ensure appropriate automated artifact detection filters and signal processing techniques in order to ensure an accurate surgical Apgar score calculation.[149] Although extreme values in heart rate and mean arterial pressure were excluded, artifactual measurements within these limits might be present.

Finally, no external validation of our model that we developed was performed. In order to verify the benefits of continuous surgical Apgar score assessment further studies have to be performed in a different patient population.

6 Conclusion and future work

6.1 Summary of findings

As patients move through the perioperative process, their overall risk for complications may change depending on certain events. A number of risk scores have been developed to facilitate rapid assessment of patients near and long term outcomes, including the surgical Apgar score. The surgical Apgar score is a ten-point risk score assessed at the end of a surgical case. The score predicts major complications within thirty days of surgery. The surgical Apgar score is based on intraoperative lowest heart rate, lowest mean arterial blood pressure, and estimated blood loss. The risk score was designed to be applied at the end of a surgical case. The predictive ability of the end-of-case surgical Apgar score for adverse events was previously evaluated in different surgical services and international hospital settings. This study assessed if a model based on continuously measured surgical Apgar score can provide clinicians with more information about a patient's changing surgical risk throughout the procedure. We hypothesized that evaluating the surgical Apgar score continuously during surgery may identify patients at higher risk for complication.

Surgical patients were retrospectively identified at Vanderbilt University Medical Center. We revalidated the surgical Apgar score in our study population and evaluated surgical Apgar score trends in all surgical services. Moreover, we identified general, vascular, and general oncology surgical patients who constituted the study population for further investigation of continuously monitored surgical Apgar score. For our study general surgery and vascular surgery were chosen since the surgical Apgar score had originally been validated in these subspecialties. General oncology was chosen due to the large sample size of this patient cohort. Logistic regression methods were used to construct a series of predictive models in order to continuously estimate the risk of major postoperative complications, and to alert care providers during surgery in case the risk exceeds a predefined threshold. Area under the receiver operating characteristic curve was used to evaluate the discriminative ability of a model utilizing a continuously measured surgical Apgar score relative to models that use only preoperative clinical factors or continuously monitored individual constituents of the surgical Apgar score (i.e. heart rate, blood pressure, and blood loss). Area under the receiver operating characteristic curve estimates were validated internally using a bootstrap method.

In our study, we found that continuous surgical Apgar score measurement combined with the established ASA classification can provide useful information about acute changes in a patient's status and his/her risk for postoperative complications in general, as well as vascular and general oncology surgery. In our study population, the continuously measured surgical Apgar score was predictive for major postoperative complications. Hypothetical real-time notifications revealed changes in patient's surgical Apgar score and helped to understand how surgical risk for adverse events changed over the course of a surgical case. This form of perioperative risk assessment was superior to conventional methods of risk stratification in our study.

In conclusion, this study will guide the prospective analysis to evaluate the utility of providing a real-time assessment of patients' surgical Apgar score. In the future, real-time notifications might allow for detection and mitigation of changes in a patient´s accumulating risk of complications during a surgical procedure.

6.2 Future avenues

This study sought to lay the groundwork to extend the utility of the surgical Apgar score by developing a new, easily computable tool that can support clinical decision-making in the operating room based on real-time changes in patient status during a surgical procedure. In a future prospective study we intend to investigate the utility of providing the perioperative team with a real-time assessment of the patient's surgical Apgar score. The surgical Apgar score for each patient will be displayed on a screen in the operation room and will continuously be recalculated once new data (i.e. heart rate, mean arterial blood pressure, blood loss) are documented in the electronic anesthesia record. When the patient's sAs drops precipitously, an automatic alert will inform the clinician about the changing acuity level of the patient.

With such a study we hope to elicit how continuous assessment of the surgical Apgar score over the course of a surgical case will affect clinical decision-making, patient care, and postoperative patient outcomes. Transitioning from a calculating a single end-of-case surgical Apgar score value to using an objective tool that monitors real-time surgical Apgar score throughout the procedure may enable clinicians to understand how surgical risk accumulates and changes over time. Once integrated into the operating room setting, such a tool might help with clinical decision-making. Real-time display of surgical Apgar score trends might also lead to a new approach to anesthetic and operative management, with earlier medical interventions aimed at stabilizing the patient intraoperatively such as modifying ventilation or changing clinicians' medication and fluids usage. The evaluation of real-time surgical Apgar score trends and knowledge of their trajectory may support clinicians' request for additional diagnostic testing, invasive patient monitoring, ordering transfer to the intensive care unit, and modifying medical care. We do not expect the changing score itself to influence the clinician's decision-making process, as the score serves as a surrogate of the patient's intra- and postoperative condition but does not mandate certain changes in treatment. Thus, we hypothesize that more timely interventions in response to changing surgical Apgar score over time may reduce perioperative morbidity and mortality.

7 Zusammenfassung

Die vorliegende Doktorarbeit befasst sich mit der kontinuierlichen Erfassung des Surgical Apgar Scores während Operationen. Der Surgical Apgar Score wurde 2007 von Gawande *et al.* in der Allgemein- und Gefäßchirurgie entwickelt, um das Risiko für schwerwiegende Komplikation innerhalb der ersten 30 Tagen nach dem Eingriff vorauszusagen. Der Score ist ein 10-Punkte-System, basierend auf folgenden intraoperativ gemessenen Parametern: niedrigste Herzfrequenz, niedrigster mittlerer Blutdruck und geschätzter Blutverlust.

Der Surgical Apgar Score wird dementsprechend am Ende einer Operation erhoben. Dieser Risikoscore wurde bereits in mehreren chirurgischen Fachrichtungen validiert und zeigte unterschiedliche, teilweise sehr gute Vorhersagefähigkeit für postoperative Komplikationen.

Diese Doktorarbeit beschäftigt sich mit der Fragestellung, ob eine kontinuierliche Erfassung des Surgical Apgar Scores mehr Informationen über das Risikoprofil eines Patienten bietet. Die Daten, die dieser Doktorarbeit zugrunde liegen stammen aus einem Datensatz aller Patienten des Universitätskrankenhauses der Vanderbilt Universität, die in den chirurgischen Abteilungen behandelt wurden. Insgesamt wurden die perioperativen Daten von 33,108 Patienten untersucht.

Zuerst wurde der Score in einer retrospektiven Studie in allen chirurgischen Fachrichtungen revalidiert und die Trends des wechselnden Scores während einer Operation erstellt. Diese Trendmodelle zeigten, dass eine starke Variation und ein starker Abfall des Surgical Apgar Scores statistisch signifikant für schwerwiegende Komplikationen während eines Zeitraumes von 30 Tagen sind.

Der zweite Teil der Studie bestand darin, dass Patienten, die sich einer allgemein-, gefäßchirurgischen oder einer tumorchirurgischen Operation unterzogen hatten, für weitere Analysen des kontinuierlich gemessenen Surgical Apgar Scores herangezogen wurden. Für dieses Patientenkollektiv wurden mit Hilfe von logistischen Regressionsanalysen drei Hauptmodelle entwickelt wurden. Das erste Model (ASA Model) war statisch und beinhaltete den ASA Status, die ethnische Zugehörigkeit, Patientenalter und das Geschlecht des Patienten. Das zweite Model (sAs Model) beinhaltete die ethnische Zugehörigkeit, Patientenalter und das Geschlecht des Patienten sowie Parameter aus dem kontinuierlich bestimmten Surgcial Apgar Score („Baseline", „maximum drop" und „insult"). Der Insult stellt eine Multiplikation aus dem Integral der maximalen Abfällen des sAs und der Zeit dar. Das dritte Model ist eine Kombination des ASA und sAs Modells. Die daraufhin durchgeführte ROC-Analyse bestimmte das „ASA+sAs" Model als das genaueste in der Vorhersage schwerer Komplikationen mit einer AUROC von 0,8 in den untersuchten Kollektiven.

Die vorliegende Arbeit konnte das „ASA+sAs" Model als gute Methode in der Prädiktion postoperativer Komplikationen bestätigen.

Anhand dieser Modelle wurde ein hypothetisches Benachrichtigungssytem für die Allgemein-, Gefäß- und Tumorchirurgie entwickelt, welches anhand von Wahrscheinlichkeiten einen hypothetischen Alarm auslöste, sobald der Surgical Apgar Score einen zuvor bestimmten Schwellenwert unterschreitet. Das Beste Benachrichtigungssystem arbeitet mit einer Komplikationswahrscheinlichkeit von 24%. Dieses System soll eine Vorlage für eine prospektive Studie bieten, in der Warnsysteme während einer Operation eingesetzt werden könnten.

8 References

1. Pearse RM, Holt PJ, Grocott MP. Managing perioperative risk in patients undergoing elective non-cardiac surgery. Bmj. 2011;343:d5759.

2. Weiser TG, Regenbogen SE, Thompson KD, Haynes AB, Lipsitz SR, Berry WR, et al. An estimation of the global volume of surgery: a modelling strategy based on available data. Lancet. 2008;372(9633):139-44.

3. Ridley S. Cardiac scoring systems--what is their value? Anaesthesia. 2003;58(10):985-91.

4. Markus PM, Martell J, Leister I, Horstmann O, Brinker J, Becker H. Predicting postoperative morbidity by clinical assessment. The British journal of surgery. 2005;92(1):101-6.

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.

6. Arvidsson S, Ouchterlony J, Sjostedt L, Svardsudd K. Predicting postoperative adverse events. Clinical efficiency of four general classification systems. The project perioperative risk. Acta anaesthesiologica Scandinavica. 1996;40(7):783-91.

7. Ivanov J, Borger MA, Rao V, David TE. The Toronto Risk Score for adverse events following cardiac surgery. The Canadian journal of cardiology. 2006;22(3):221-7.

8. Gawande AA, Kwaan MR, Regenbogen SE, Lipsitz SA, Zinner MJ. An Apgar score for surgery. Journal of the American College of Surgeons. 2007;204(2):201-8.

9. Dalton JE, Kurz A, Turan A, Mascha EJ, Sessler DI, Saager L. Development and validation of a risk quantification index for 30-day postoperative mortality and morbidity in noncardiac surgical patients. Anesthesiology. 2011;114(6):1336-44.

10. Sessler DI, Sigl JC, Manberg PJ, Kelley SD, Schubert A, Chamoun NG. Broadly applicable risk stratification system for predicting duration of hospitalization and mortality. Anesthesiology. 2010;113(5):1026-37.

11. Barnett S, Moonesinghe SR. Clinical risk scores to guide perioperative management. Postgraduate medical journal. 2011;87(1030):535-41.

12. Saklad MMD. Grading of Patients for Surgical Procedures. Anesthesiology. 1941;2:281-4.

13. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. The British journal of surgery. 1991;78(3):355-60.

14. Lee TH, Marcantonio ER, Mangione CM, Thomas EJ, Polanczyk CA, Cook EF, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. Circulation. 1999;100(10):1043-9.

15. Haynes SR, Lawler PG. An assessment of the consistency of ASA physical status classification allocation. Anaesthesia. 1995;50(3):195-9.

16. Rix TE, Bates T. Pre-operative risk scores for the prediction of outcome in elderly people who require emergency surgery. World journal of emergency surgery : WJES. 2007;2:16.

17. Tekkis PP, Kessaris N, Kocher HM, Poloniecki JD, Lyttle J, Windsor AC. Evaluation of POSSUM and P-POSSUM scoring systems in patients undergoing colorectal surgery. The British journal of surgery. 2003;90(3):340-5.

18. Regenbogen SE, Lancaster RT, Lipsitz SR, Greenberg CC, Hutter MM, Gawande AA. Does the Surgical Apgar Score measure intraoperative performance? Annals of surgery. 2008;248(2):320-8.

19. Regenbogen SE, Ehrenfeld JM, Lipsitz SR, Greenberg CC, Hutter MM, Gawande AA. Utility of the surgical apgar score: validation in 4119 patients. Arch Surg. 2009;144(1):30-6; discussion 7.

20. Dripps RD, Lamont A, Eckenhoff JE. The role of anesthesia in surgical mortality. JAMA : the journal of the American Medical Association. 1961;178:261-6.

21. Wolters U, Wolf T, Stutzer H, Schroder T. ASA classification and perioperative variables as predictors of postoperative outcome. British journal of anaesthesia. 1996;77(2):217-22.

22. Owens WD, Felts JA, Spitznagel EL, Jr. ASA physical status classifications: a study of consistency of ratings. Anesthesiology. 1978;49(4):239-43.

23. Vacanti CJ, VanHouten RJ, Hill RC. A statistical analysis of the relationship of physical status to postoperative mortality in 68,388 cases. Anesthesia and analgesia. 1970;49(4):564-6.

24. Cullen DJ, Apolone G, Greenfield S, Guadagnoli E, Cleary P. ASA Physical Status and age predict morbidity after three surgical procedures. Annals of surgery. 1994;220(1):3-9.

25. Bjorgul K, Novicoff WM, Saleh KJ. American Society of Anesthesiologist Physical Status score may be used as a comorbidity index in hip fracture surgery. The Journal of arthroplasty. 2010;25(6 Suppl):134-7.

26. Prause G, Offner A, Ratzenhofer-Komenda B, Vicenzi M, Smolle J, Smolle-Juttner F. Comparison of two preoperative indices to predict perioperative mortality in non-cardiac thoracic surgery. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 1997;11(4):670-5.

27. Fowkes FG, Lunn JN, Farrow SC, Robertson IB, Samuel P. Epidemiology in anaesthesia. III: Mortality risk in patients with coexisting physical disease. British journal of anaesthesia. 1982;54(8):819-25.

28. Hennein HA, Mendeloff EN, Cilley RE, Bove EL, Coran AG. Predictors of postoperative outcome after general surgical procedures in patients with congenital heart disease. Journal of pediatric surgery. 1994;29(7):866-70.

29. Menke H, John KD, Klein A, Lorenz W, Junginger T. [Preoperative risk assessment with the ASA classification. A prospective study of morbidity and mortality in various ASA classes in 2,937 patients in general surgery]. Der Chirurg; Zeitschrift fur alle Gebiete der operativen Medizen. 1992;63(12):1029-34.

30. Menke H, Klein A, John KD, Junginger T. Predictive value of ASA classification for the assessment of the perioperative risk. International surgery. 1993;78(3):266-70.

31. Davenport DL, Bowe EA, Henderson WG, Khuri SF, Mentzer RM, Jr. National Surgical Quality Improvement Program (NSQIP) risk factors can be used to validate American Society of Anesthesiologists Physical Status Classification (ASA PS) levels. Annals of surgery. 2006;243(5):636-41; discussion 41-4.

32. Mantilla CB, Horlocker TT, Schroeder DR, Berry DJ, Brown DL. Risk factors for clinically relevant pulmonary embolism and deep venous thrombosis in patients undergoing primary hip or knee arthroplasty. Anesthesiology. 2003;99(3):552-60; discussion 5A.

33. Pasternak LR, Johns A. Ambulatory gynaecological surgery: risk and assessment. Best practice & research Clinical obstetrics & gynaecology. 2005;19(5):663-79.

34. Han KR, Kim HL, Pantuck AJ, Dorey FJ, Figlin RA, Belldegrun AS. Use of American Society of Anesthesiologists physical status classification to assess perioperative risk in patients undergoing radical nephrectomy for renal cell carcinoma. Urology. 2004;63(5):841-6; discussion 6-7.

35. Chand M, Armstrong T, Britton G, Nash GF. How and why do we measure surgical risk? J R Soc Med. 2007;100(11):508-12.

36. Mak PH, Campbell RC, Irwin MG, American Society of A. The ASA Physical Status Classification: inter-observer consistency. American Society of Anesthesiologists. Anaesthesia and intensive care. 2002;30(5):633-40.

37. Ranta S, Hynynen M, Tammisto T. A survey of the ASA physical status classification: significant variation in allocation among Finnish anaesthesiologists. Acta anaesthesiologica Scandinavica. 1997;41(5):629-32.

38. Aronson WL, McAuliffe MS, Miller K. Variability in the American Society of Anesthesiologists Physical Status Classification Scale. AANA journal. 2003;71(4):265-74.

39. Barbeito A, Muir HA, Gan TJ, Reynolds JD, Spahn T, White WD, et al. Use of a modifier reduces inconsistency in the American Society of Anesthesiologists Physical Status Classification in parturients. Anesthesia and analgesia. 2006;102(4):1231-3.

40. Aplin S, Baines D, J DEL. Use of the ASA Physical Status Grading System in pediatric practice. Paediatric anaesthesia. 2007;17(3):216-22.

41. Goldman L, Caldera DL, Nussbaum SR, Southwick FS, Krogstad D, Murray B, et al. Multifactorial index of cardiac risk in noncardiac surgical procedures. The New England journal of medicine. 1977;297(16):845-50.

42. Boersma E, Kertai MD, Schouten O, Bax JJ, Noordzij P, Steyerberg EW, et al. Perioperative cardiovascular mortality in noncardiac surgery: validation of the Lee cardiac risk index. The American journal of medicine. 2005;118(10):1134-41.

43. Rao JY, Yeriswamy MC, Santhosh MJ, Shetty GG, Varghese K, Patil CB, et al. A look into Lee's score: peri-operative cardiovascular risk assessment in non-cardiac surgeries-usefulness of revised cardiac risk index. Indian heart journal. 2012;64(2):134-8.

44. Devereaux PJ, Goldman L, Cook DJ, Gilbert K, Leslie K, Guyatt GH. Perioperative cardiac events in patients undergoing noncardiac surgery: a review of the magnitude of the problem, the

pathophysiology of the events and methods to estimate and communicate risk. CMAJ. 2005;173(6):627-34.

45. Fleisher LA, American College of Cardiology/American Heart A. Cardiac risk stratification for noncardiac surgery: update from the American College of Cardiology/American Heart Association 2007 guidelines. Cleveland Clinic journal of medicine. 2009;76 Suppl 4:S9-15.

46. Poldermans D, Bax JJ, Boersma E, De Hert S, Eeckhout E, Fowkes G, et al. Guidelines for pre-operative cardiac risk assessment and perioperative cardiac management in non-cardiac surgery. European heart journal. 2009;30(22):2769-812.

47. Poldermans D, Hoeks SE, Feringa HH. Pre-operative risk assessment and risk reduction before surgery. Journal of the American College of Cardiology. 2008;51(20):1913-24.

48. Karkos CD, Thomson GJ, Hughes R, Hollis S, Hill JC, Mukhopadhyay US. Prediction of cardiac risk before abdominal aortic reconstruction: comparison of a revised Goldman Cardiac Risk Index and radioisotope ejection fraction. Journal of vascular surgery : official publication, the Society for Vascular Surgery [and] International Society for Cardiovascular Surgery, North American Chapter. 2002;35(5):943-9.

49. Welten GM, Schouten O, van Domburg RT, Feringa HH, Hoeks SE, Dunkelgrun M, et al. The influence of aging on the prognostic value of the revised cardiac risk index for postoperative cardiac complications in vascular surgery patients. European journal of vascular and endovascular surgery : the official journal of the European Society for Vascular Surgery. 2007;34(6):632-8.

50. Archan S, Roscher CR, Fairman RM, Fleisher LA. Revised Cardiac Risk Index (Lee) and perioperative cardiac events as predictors of long-term mortality in patients undergoing endovascular abdominal aortic aneurysm repair. Journal of cardiothoracic and vascular anesthesia. 2010;24(1):84-90.

51. Brunelli A, Varela G, Salati M, Jimenez MF, Pompili C, Novoa N, et al. Recalibration of the revised cardiac risk index in lung resection candidates. The Annals of thoracic surgery. 2010;90(1):199-203.

52. Brunelli A, Cassivi SD, Fibla J, Halgren LA, Wigle DA, Allen MS, et al. External validation of the recalibrated thoracic revised cardiac risk index for predicting the risk of major cardiac complications after lung resection. The Annals of thoracic surgery. 2011;92(2):445-8.

53. Ackland GL, Harris S, Ziabari Y, Grocott M, Mythen M, Investigators SO. Revised cardiac risk index and postoperative morbidity after elective orthopaedic surgery: a prospective cohort study. British journal of anaesthesia. 2010;105(6):744-52.

54. Moitra VK, Flynn BC, Mazzeffi M, Bodian C, Bronheim D, Ellis JE. Indication for surgery, the revised cardiac risk index, and 1-year mortality. Annals of vascular surgery. 2011; 25(7):902-8.

55. Grayburn PA, Hillis LD. Cardiac events in patients undergoing noncardiac surgery: shifting the paradigm from noninvasive risk stratification to therapy. Annals of internal medicine. 2003;138(6):506-11.

56. Moran PJ, Ghidella T, Power G, Jenkins AS, Whittle D. The use of Lee and co-workers' index to assist a risk adjusted audit of perioperative cardiac outcome. Anaesthesia and intensive care. 2008;36(2):167-73.

57. Bertges DJ, Goodney PP, Zhao Y, Schanzer A, Nolan BW, Likosky DS, et al. The Vascular Study Group of New England Cardiac Risk Index (VSG-CRI) predicts cardiac complications more accurately than the Revised Cardiac Risk Index in vascular surgery patients. Journal of vascular surgery : official publication, the Society for Vascular Surgery [and] International Society for Cardiovascular Surgery, North American Chapter. 2010;52(3):674-83, 83 e1-83 e3.

58. Goldman L. The revised cardiac risk index delivers what it promised. Annals of internal medicine. 2010;152(1):57-8.

59. Choi JH, Cho DK, Song YB, Hahn JY, Choi S, Gwon HC, et al. Preoperative NT-proBNP and CRP predict perioperative major cardiovascular events in non-cardiac surgery. Heart. 2010;96(1):56-62.

60. Bader AM. Advances in preoperative risk assessment and management. Current problems in surgery. 2012;49(1):11-40.

61. Mangano DT, Browner WS, Hollenberg M, London MJ, Tubau JF, Tateo IM. Association of perioperative myocardial ischemia with cardiac morbidity and mortality in men undergoing noncardiac surgery. The Study of Perioperative Ischemia Research Group. The New England journal of medicine. 1990;323(26):1781-8.

62. Mauck KF, Manjarrez EC, Cohn SL. Perioperative cardiac evaluation: assessment, risk reduction, and complication management. Clinics in geriatric medicine. 2008;24(4):585-605, vii.

63. Lupei MI, Chipman JG, Beilman GJ, Oancea SC, Konia MR. The association between ASA status and other risk stratification models on postoperative intensive care unit outcomes. Anesthesia and analgesia. 2014;118(5):989-94.

64. Ford MK, Beattie WS, Wijeysundera DN. Systematic review: prediction of perioperative cardiac complications and mortality by the revised cardiac risk index. Ann Intern Med. 2010;152(1):26-35.

65. Fleischmann KE, Goldman L, Young B, Lee TH. Association between cardiac and noncardiac complications in patients undergoing noncardiac surgery: outcomes and effects on length of stay. Am J Med. 2003;115(7):515-20.

66. Group PS, Devereaux PJ, Yang H, Yusuf S, Guyatt G, Leslie K, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. Lancet. 2008;371(9627):1839-47.

67. Copeland GP. The POSSUM system of surgical audit. Arch Surg. 2002;137(1):15-9.

68. Horzic M, Kopljar M, Cupurdija K, Bielen DV, Vergles D, Lackovic Z. Comparison of P-POSSUM and Cr-POSSUM scores in patients undergoing colorectal cancer resection. Archives of surgery. 2007;142(11):1043-8.

69. Neary WD, Heather BP, Earnshaw JJ. The Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM). The British journal of surgery. 2003;90(2):157-65.

70. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. POSSUM and Portsmouth POSSUM for predicting mortality. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity. The British journal of surgery. 1998;85(9):1217-20.

71. Yii MK, Ng KJ. Risk-adjusted surgical audit with the POSSUM scoring system in a developing country. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity. The British journal of surgery. 2002;89(1):110-3.

72. Whiteley MS, Prytherch DR, Higgins B, Weaver PC, Prout WG. An evaluation of the POSSUM surgical scoring system. The British journal of surgery. 1996;83(6):812-5.

73. Gocmen E, Koc M, Tez M, Keskek M, Kilic M, Ertan T. Evaluation of P-POSSUM and O-POSSUM scores in patients with gastric cancer undergoing resection. Hepato-gastroenterology. 2004;51(60):1864-6.

74. Luna A, Rebasa P, Navarro S, Montmany S, Coroleu D, Cabrol J, et al. An evaluation of morbidity and mortality in oncologic gastric surgery with the application of POSSUM, P-POSSUM, and O-POSSUM. World journal of surgery. 2009;33(9):1889-94.

75. Byrne JS, Condon ET, Ahmed M, Conroy R, Mehigan D, Sheehan SJ, et al. Surgical audit using the POSSUM scoring tool in vascular surgery patients. Irish journal of medical science. 2009;178(4):453-6.

76. Mosquera D, Chiang N, Gibberd R. Evaluation of surgical performance using V-POSSUM risk-adjusted mortality rates. ANZ journal of surgery. 2008;78(7):535-9.

77. Tekkis PP, Prytherch DR, Kocher HM, Senapati A, Poloniecki JD, Stamatakis JD, et al. Development of a dedicated risk-adjustment scoring system for colorectal surgery (colorectal POSSUM). The British journal of surgery. 2004;91(9):1174-82.

78. Zafirellis KD, Fountoulakis A, Dolan K, Dexter SP, Martin IG, Sue-Ling HM. Evaluation of POSSUM in patients with oesophageal cancer undergoing resection. The British journal of surgery. 2002;89(9):1150-5.

79. Prytherch DR, Ridler BM, Beard JD, Earnshaw JJ. A model for national outcome audit in vascular surgery. European journal of vascular and endovascular surgery : the official journal of the European Society for Vascular Surgery. 2001;21(6):477-83.

80. Prytherch DR, Sutton GL, Boyle JR. Portsmouth POSSUM models for abdominal aortic aneurysm surgery. The British journal of surgery. 2001;88(7):958-63.

81. Tekkis PP, McCulloch P, Poloniecki JD, Prytherch DR, Kessaris N, Steger AC. Risk-adjusted prediction of operative mortality in oesophagogastric surgery with O-POSSUM. The British journal of surgery. 2004;91(3):288-95.

82. Neary WD, Prytherch D, Foy C, Heather BP, Earnshaw JJ. Comparison of different methods of risk stratification in urgent and emergency surgery. The British journal of surgery. 2007;94(10):1300-5.

83. Tran Ba Loc P, du Montcel ST, Duron JJ, Levard H, Suc B, Descottes B, et al. Elderly POSSUM, a dedicated score for prediction of mortality and morbidity after major colorectal surgery in older patients. The British journal of surgery. 2010;97(3):396-403.

84. Mohil RS, Bhatnagar D, Bahadur L, Rajneesh, Dev DK, Magan M. POSSUM and P-POSSUM for risk-adjusted audit of patients undergoing emergency laparotomy. The British journal of surgery. 2004;91(4):500-3.

85. Kumar S, Gupta A, Chaudhary S, Agrawal N. Validation of the use of POSSUM score in enteric perforation peritonitis - results of a prospective study. The Pan African medical journal. 2011;9:22.

86. Bann SD, Sarin S. Comparative audit: the trouble with POSSUM. Journal of the Royal Society of Medicine. 2001;94(12):632-4.

87. Brooks MJ, Sutton R, Sarin S. Comparison of Surgical Risk Score, POSSUM and p-POSSUM in higher-risk surgical patients. The British journal of surgery. 2005;92(10):1288-92.

88. Wijesinghe LD, Mahmood T, Scott DJ, Berridge DC, Kent PJ, Kester RC. Comparison of POSSUM and the Portsmouth predictor equation for predicting death following vascular surgery. The British journal of surgery. 1998;85(2):209-12.

89. Chen FG, Khoo ST. Critical care medicine--a review of the outcome prediction in critical care. Annals of the Academy of Medicine, Singapore. 1993;22(3):360-4.

90. Goldberger AL, O'Konski M. Utility of the routine electrocardiogram before surgery and on general hospital admission. Critical review and new guidelines. Annals of internal medicine. 1986;105(4):552-7.

91. Sutton R, Bann S, Brooks M, Sarin S. The Surgical Risk Scale as an improved tool for risk-adjusted analysis in comparative surgical audit. The British journal of surgery. 2002;89(6):763-8.

92. Apgar V. A proposal for a new method of evaluation of the newborn infant. Current researches in anesthesia & analgesia. 1953;32(4):260-7.

93. 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.

94. Apgar V, Holaday DA, James LS, Weisbrot IM, Berrien C. Evaluation of the newborn infant; second report. J Am Med Assoc. 1958;168(15):1985-8.

95. Apgar V. The newborn (Apgar) scoring system. Reflections and advice. Pediatric clinics of North America. 1966;13(3):645-50.

96. Hartmann B, Junger A, Rohrig R, Klasen J, Jost A, Benson M, et al. Intra-operative tachycardia and peri-operative outcome. Langenbeck's archives of surgery / Deutsche Gesellschaft fur Chirurgie. 2003;388(4):255-60.

97. Charlson ME, MacKenzie CR, Gold JP, Ales KL, Topkins M, Shires GT. Intraoperative blood pressure. What patterns identify patients at risk for postoperative complications? Annals of surgery. 1990;212(5):567-80.

98. Khuri SF, Daley J, Henderson W, Barbour G, Lowry P, Irvin G, et al. The National Veterans Administration Surgical Risk Study: risk adjustment for the comparative assessment of the quality of surgical care. Journal of the American College of Surgeons. 1995;180(5):519-31.

99. Jones HJ, de Cossart L. Risk scoring in surgical patients. The British journal of surgery. 1999;86(2):149-57.

100. Reich DL, Bodian CA, Krol M, Kuroda M, Osinski T, Thys DM. Intraoperative hemodynamic predictors of mortality, stroke, and myocardial infarction after coronary artery bypass surgery. Anesthesia and analgesia. 1999;89(4):814-22.

101. Rohrig R, Junger A, Hartmann B, Klasen J, Quinzio L, Jost A, et al. The incidence and prediction of automatically detected intraoperative cardiovascular events in noncardiac surgery. Anesthesia and analgesia. 2004;98(3):569-77, table of contents.

102. Charlson ME, MacKenzie CR, Gold JP, Ales KL, Topkins M, Fairclough GP, Jr., et al. The preoperative and intraoperative hemodynamic predictors of postoperative myocardial infarction or ischemia in patients undergoing noncardiac surgery. Annals of surgery. 1989;210(5):637-48.

103. Monk TG, Saini V, Weldon BC, Sigl JC. Anesthetic management and one-year mortality after noncardiac surgery. Anesthesia and analgesia. 2005;100(1):4-10.

104. Chang HS, Hongo K, Nakagawa H. Adverse effects of limited hypotensive anesthesia on the outcome of patients with subarachnoid hemorrhage. Journal of neurosurgery. 2000;92(6):971-5.

105. Prasad SM, Ferreria M, Berry AM, Lipsitz SR, Richie JP, Gawande AA, et al. Surgical apgar outcome score: perioperative risk assessment for radical cystectomy. The Journal of urology. 2009;181(3):1046-52; discussion 52-3.

106. Regenbogen SE, Bordeianou L, Hutter MM, Gawande AA. The intraoperative Surgical Apgar Score predicts postdischarge complications after colon and rectal resection. Surgery. 2010;148(3):559-66.

107. Zighelboim I, Kizer N, Taylor NP, Case AS, Gao F, Thaker PH, et al. "Surgical Apgar Score" predicts postoperative complications after cytoreduction for advanced ovarian cancer. Gynecologic oncology. 2010;116(3):370-3.

108. Ohlsson H, Winso O. Assessment of the Surgical Apgar Score in a Swedish setting. Acta anaesthesiologica Scandinavica. 2011;55(5):524-9.

109. Sobol JB, Gershengorn HB, Wunsch H, Li G. The surgical Apgar score is strongly associated with intensive care unit admission after high-risk intraabdominal surgery. Anesthesia and analgesia. 2013;117(2):438-46.

110. Glass NE, Pinna A, Masi A, Rosman AS, Neihaus D, Okochi S, et al. The surgical apgar score predicts postoperative ICU admission. Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract. 2015;19(3):445-50.

111. Assifi MM, Lindenmeyer J, Leiby BE, Grunwald Z, Rosato EL, Kennedy EP, et al. Surgical Apgar score predicts perioperative morbidity in patients undergoing pancreaticoduodenectomy at a high-volume center. Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract. 2012;16(2):275-81.

112. Haynes AB, Regenbogen SE, Weiser TG, Lipsitz SR, Dziekan G, Berry WR, et al. Surgical outcome measurement for a global patient population: validation of the Surgical Apgar Score in 8 countries. Surgery. 2011;149(4):519-24.

113. Wuerz TH, Regenbogen SE, Ehrenfeld JM, Malchau H, Rubash HE, Gawande AA, et al. The Surgical Apgar Score in hip and knee arthroplasty. Clinical orthopaedics and related research. 2011;469(4):1119-26.

114. Urrutia J, Valdes M, Zamora T, Canessa V, Briceno J. An assessment of the Surgical Apgar Score in spine surgery. Spine J. 2015;15(1):105-9.

115. La Torre M, Ramacciato G, Nigri G, Balducci G, Cavallini M, Rossi M, et al. Post-operative morbidity and mortality in pancreatic surgery. The role of surgical Apgar score. Pancreatology. 2013;13(2):175-9.

116. Ito T, Abbosh PH, Mehrazin R, Tomaszewski JJ, Li T, Ginzburg S, et al. Surgical Apgar Score Predicts an Increased Risk of Major Complications and Death after Renal Mass Excision. The Journal of urology. 2014.

117. Ziewacz JE, Davis MC, Lau D, El-Sayed AM, Regenbogen SE, Sullivan SE, et al. Validation of the surgical Apgar score in a neurosurgical patient population. Journal of neurosurgery. 2013;118(2):270-9.

118. Melis M, Pinna A, Okochi S, Masi A, Rosman AS, Neihaus D, et al. Validation of the Surgical Apgar Score in a veteran population undergoing general surgery. Journal of the American College of Surgeons. 2014;218(2):218-25.

119. Miki Y, Tokunaga M, Tanizawa Y, Bando E, Kawamura T, Terashima M. Perioperative risk assessment for gastrectomy by surgical apgar score. Ann Surg Oncol. 2014;21(8):2601-7.

120. Reynolds PQ, Sanders NW, Schildcrout JS, Mercaldo ND, St Jacques PJ. Expansion of the surgical Apgar score across all surgical subspecialties as a means to predict postoperative mortality. Anesthesiology. 2011;114(6):1305-12.

121. Charlson ME, MacKenzie CR, Gold JP, Ales KL, Topkins M, Shires GT. Preoperative characteristics predicting intraoperative hypotension and hypertension among hypertensives and diabetics undergoing noncardiac surgery. Annals of surgery. 1990;212(1):66-81.

122. Mangano DT. Perioperative cardiac morbidity. Anesthesiology. 1990;72(1):153-84.

123. Delilkan AE. Comparison of subjective estimates by surgeons and anaesthetists of operative blood loss. Br Med J. 1972;2(5814):619-21.

124. Gardiner AJ, Dudley HA. The measurement of blood loss at operation. British journal of anaesthesia. 1962;34:653-6.

125. Lau D, Yee TJ, La Marca F, Patel R, Park P. Utility of the Surgical Apgar Score for Patients who Undergo Surgery for Spinal Metastasis. J Spinal Disord Tech. 2014.

126. Thorn CC, Chan M, Sinha N, Harrison RA. Utility of the Surgical Apgar Score in a District General Hospital. World journal of surgery. 2012.

127. Urrutia J, Valdes M, Zamora T, Canessa V, Briceno J. Can the Surgical Apgar Score predict morbidity and mortality in general orthopaedic surgery? Int Orthop. 2012;36(12):2571-6.

128. da Costa VV, Caldas AC, Nunes LG, Beraldo PS, Saraiva RA. Influence of angiotensin-converting enzyme inhibitors on hypotension after anesthetic induction: is the preoperative discontinuation of this drug necessary? Revista brasileira de anestesiologia. 2009;59(6):704-15.

129. Ebert TJ, Muzi M, Berens R, Goff D, Kampine JP. Sympathetic responses to induction of anesthesia in humans with propofol or etomidate. Anesthesiology. 1992;76(5):725-33.

130. 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. Annals of surgery. 1970;172(3):497-503.

131. Reich DL, Bennett-Guerrero E, Bodian CA, Hossain S, Winfree W, Krol M. Intraoperative tachycardia and hypertension are independently associated with adverse outcome in noncardiac surgery of long duration. Anesthesia and analgesia. 2002;95(2):273-7, table of contents.

132. Lawrence JP. Advances and new insights in monitoring. Thoracic surgery clinics. 2005;15(1):55-70.

133. Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, et al. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. Annals of surgery. 1998;228(4):491-507.

134. Khuri SF. The NSQIP: a new frontier in surgery. Surgery. 2005;138(5):837-43.

135. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. Intensive care medicine. 2003;29(7):1043-51.

136. Szmukler G, Everitt B, Leese M. Risk assessment and receiver operating characteristic curves. Psychological medicine. 2012;42(5):895-8.

137. Akobeng AK. Understanding diagnostic tests 3: Receiver operating characteristic curves. Acta paediatrica. 2007;96(5):644-7.

138. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29-36.

139. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in medicine. 1996;15(4):361-87.

140. Wanderer JP, Sandberg WS, Ehrenfeld JM. Real-time alerts and reminders using information systems. Anesthesiology clinics. 2011;29(3):389-96.

141. Wax DB, Beilin Y, Levin M, Chadha N, Krol M, Reich DL. The effect of an interactive visual reminder in an anesthesia information management system on timeliness of prophylactic antibiotic administration. Anesthesia and analgesia. 2007;104(6):1462-6, table of contents.

142. Kooij FO, Klok T, Hollmann MW, Kal JE. Decision support increases guideline adherence for prescribing postoperative nausea and vomiting prophylaxis. Anesthesia and analgesia. 2008;106(3):893-8, table of contents.

143. Sandberg WS, Sandberg EH, Seim AR, Anupama S, Ehrenfeld JM, Spring SF, et al. Real-time checking of electronic anesthesia records for documentation errors and automatically text messaging clinicians improves quality of documentation. Anesthesia and analgesia. 2008;106(1):192-201, table of contents.

144. Nair BG, Newman SF, Peterson GN, Wu WY, Schwid HA. Feedback mechanisms including real-time electronic alerts to achieve near 100% timely prophylactic antibiotic administration in surgical cases. Anesthesia and analgesia. 2010;111(5):1293-300.

145. Ghaferi AA, Birkmeyer JD, Dimick JB. Complications, failure to rescue, and mortality with major inpatient surgery in medicare patients. Annals of surgery. 2009;250(6):1029-34.

146. Hyder JA, Kor DJ, Cima RR, Subramanian A. How to improve the performance of intraoperative risk models: an example with vital signs using the surgical apgar score. Anesthesia and analgesia. 2013;117(6):1338-46.

147. Brecher ME, Monk T, Goodnough LT. A standardized method for calculating blood loss. Transfusion. 1997;37(10):1070-4.

148. Buhre W, Rossaint R. Perioperative management and monitoring in anaesthesia. Lancet. 2003;362(9398):1839-46.

149. Takla G, Petre JH, Doyle DJ, Horibe M, Gopakumaran B. The problem of artifacts in patient monitor data during surgery: a clinical and methodological review. Anesthesia and analgesia. 2006;103(5):1196-204.

Eidesstattliche Versicherung

Jering, Monika

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Thema

Utilization of the Surgical Apgar Score as a Continuous Measure of Intra-operative Risk

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 01.08.2016

Ort, Datum Unterschrift Doktorandin/Doktorand