BMJ Open

Protocol for the Anesthesiology Control Tower: Forecasting Algorithms to Support Treatments (ACTFAST 2) Trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020124
Article Type:	Protocol
Date Submitted by the Author:	17-Oct-2017

Complete List of Authors:	Fritz, Bradley; Washington University School of Medicine in St. Louis, Anesthesiology Chen, Yixin; Washington University in Saint Louis, Computer Science and Engineering Murray-Torres, Teresa; Washington University School of Medicine in St. Louis, Anesthesiology Gregory, SH; Washington University School of Medicine in St. Louis, Anesthesiology Ben Abdallah, Arbi; Washington University School of Medicine in St. Louis, Anesthesiology Kronzer, Alex; Washington University School of Medicine in St. Louis, Anesthesiology McKinnon, Sherry; Washington University School of Medicine, Anesthesiology Budelier, Thaddeus; Washington University School of Medicine in St. Louis, Anesthesiology Helsten, Daniel; Washington University School of Medicine in St. Louis, Anesthesiology Wildes, Troy; Washington University in Saint Louis School of Medicine, Anesthesiology Sharma, Anshuman; Washington University School of Medicine in St. Louis, Anesthesiology Avidan, Michael; Washington University School of Medicine, Anesthesiology

Keywords:	Adult anaesthesia < ANAESTHETICS, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.

Data Dictionary v4 Tab1 - Preop.xls
Data Dictionary v4 Tab2 - Intraop.xls
Data Dictionary v4 Tab3 - SATISFY-SOS Outcomes.xls
Data Dictionary v4 Tab4 - Sunrise Outcomes.xls

Protocol for the Anesthesiology Control Tower: Forecasting Algorithms to Support Treatments (ACTFAST 2) Trial

Bradley A Fritz, MD, Resident Physician[1]

Yixin Chen, PhD, Professor[2]

Teresa M Murray-Torres, MD, Resident Physician[1]

Stephen Gregory, MD, Instructor[1]

Arbi Ben Abdallah, PhD, Assistant Professor[1]

Alex Kronzer, Database Analyst[1]

Sherry McKinnon, Project Manager[1]

Thaddeus Budelier, Research Coordinator[1]

Daniel L Helsten, MD, Associate Professor[1]

Troy S Wildes, MD, Associate Professor[1]

Anshuman Sharma, MD, Professor[1]

Michael S Avidan, MBBCh, Professor[1]

1 – Department of Anesthesiology, Washington University in St. Louis, St. Louis, MO, USA

2 – Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO, USA

Corresponding Author: Bradley Fritz

1 Barnes Jewish Hospital Plaza

St. Louis, MO, 63108

(314) 747-4155

bafritz@wustl.edu

ABSTRACT

Introduction – Mortality and morbidity following surgery are pressing public health concerns in the United States. Traditional prediction models for postoperative adverse outcomes demonstrate good discrimination at the population level, but the ability to forecast an individual patient's trajectory in real time remains poor. We propose to apply machine-learning techniques to perioperative time-series data to develop algorithms for predicting adverse perioperative outcomes.

Methods and Analysis – This study will include all adult patients who had surgery at our tertiary care hospital over a four-year period. Patient history, laboratory values, minute-by-minute intraoperative vital signs, and medications administered will be extracted from the electronic medical record. Outcomes will include in-hospital mortality, postoperative acute kidney injury, and postoperative respiratory failure. Forecasting algorithms for each of these outcomes will be constructed using density-based logistic regression after employing a Nadaraya-Watson kernel density estimator. Time-series variables will be analyzed using first- and second-order feature extraction, shapelet methods, and convolutional neural networks. The algorithms will be validated using bootstrap methods.

Ethics and Dissemination – The successful development of these forecasting algorithms will allow perioperative health care clinicians to predict more accurately an individual patient's risk for specific adverse perioperative outcomes in real time. Knowledge of a patient's dynamic risk profile may allow clinicians to make targeted changes in the care plan that will alter the patient's outcome trajectory. This hypothesis will be tested in a future randomized controlled trial.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- Will utilize modeling techniques that take advantage of the rich time-series data that are available, rather than data from a single time point
- Will utilize efficient modeling techniques that can process large amounts of data quickly
- Will utilize group-based learning to increase model accuracy by separating groups of patients who likely have different relationship between underlying features and predicted outcomes
- Dissemination to other health care facilities may be limited by the availability of high-quality preoperative and intraoperative input data in a usable format

INTRODUCTION

An estimated 40 million people undergo surgery every year in the United States. Postoperative mortality rate at one year for surgical inpatients is between 5 and 10 percent,[1 2] and an estimated 10 percent of surgical patients suffer major in-hospital morbidity.[3-8] Perioperative morbidity and mortality are therefore pressing public health concerns. Many patient characteristics, including comorbid medical conditions, associate strongly and independently with perioperative mortality and major morbidity.[1 2 9-11] While many of these characteristics are not modifiable, some perioperative risk factors, such as intraoperative blood pressures and anesthetic concentrations,[1 2 9 10] can be modified in real time. Although the association between perioperative variables and postoperative outcomes has been well established at the population level using approaches such as standard logistic regression, [1 2 9 10 12] the ability to utilize deviations in physiological parameters in real time to dynamically forecast the trajectory of each individual patient remains poor.

There is a gap in the field with an opportunity to assess the potential utility of machine learning-based forecasting algorithms to anticipate adverse perioperative outcomes, guide interventions, and improve overall quality of care. Standard forecasting models, such as logistic regression, linear regression, and other statistical modeling procedures, have long been used to identify and prioritize risk factors for adverse outcomes. Although most of these statistical techniques have been shown to have moderate predictive values, they are limited in their prognostic ability and practical use.[1 2 6 9 10] In contrast to standard forecasting models, we have demonstrated machine learning and data mining approaches for patients on intensive care units that generate markedly superior prediction for outcomes such as mortality.[13] Our methods differ from standard statistical techniques in their ability to effectively incorporate time-series data. Most standard modeling techniques for surgical patients are based on a snapshot scheme, which only considers the data values at a given moment. They are not competent in extracting features from time-series data, especially in real-time fashion, such as temporal trends and shapes. Therefore, the objective of this study is to utilize machine-learning techniques to build forecasting algorithms that use patient characteristics and high-fidelity intraoperative time-series data to predict adverse perioperative outcomes.

METHODS AND ANALYSIS

Study Design

Our central hypothesis is that with sufficient knowledge of patient characteristics coupled with repeated, high-fidelity time series data from the perioperative electronic medical record, advanced models can be constructed for individual patients that will forecast adverse perioperative outcomes. To test this hypothesis, we will conduct an observational cohort study of adult patients who undergo surgery at Barnes-Jewish Hospital in St. Louis, Missouri. First, we plan to develop forecasting algorithms for specific adverse perioperative outcomes using historical data. Next, we plan to validate these algorithms by determining whether they can be used to reliably forecast individual adverse perioperative outcomes.

Patient Population and Sample Size

This study will include all adult patients who had surgery in the 48 operating rooms at Barnes-Jewish Hospital in St. Louis, Missouri between June 1, 2012 and August 31, 2016. Patients who receive anesthesia care in areas outside the main operating rooms, such as the obstetric suite or the outpatient surgery suite, will not be included. Barnes-Jewish Hospital is a 1,252-bed academic university-affiliated adult tertiary care hospital, performing approximately 19,000 surgeries a year. On average, 125 surgeries take place in these operating rooms every business day. To be conservative, we estimate that information on 50 to 100 surgeries per day will be available for analysis. We therefore anticipate a minimum total sample size of 50,000 to 100,000 surgeries for algorithm development and validation.

The Human Research Protection Office at Washington University in St. Louis has granted a waiver of informed consent for all subjects enrolled in this study. This study has been determined to involve no more than minimal risk to participants, as no additional data will be collected beyond that already contained in the electronic record. For the same reason, the waiver of consent will not adversely affect the participants' rights and welfare. It is impracticable to conduct this research without a waiver of consent because 100% participation from the patients is imperative to obtain scientifically sound data.

Data Acquisition

For this project, we will use high-dimensional and complex data from a variety of electronic medical record sources to cover the entire perioperative period. Much of the relevant information will be imported from MetaVision®(iMDsoft, Wakefield, MA), an anesthesiology information management software system that is the perioperative electronic clinical documentation system currently utilized by the Department of Anesthesiology. MetaVision® captures comprehensive clinical data beginning with the preoperative assessment and continuing throughout the duration of the perioperative period. Information captured preoperatively includes patients' past medical and surgical histories, chronic medical issues, medications used, and functional capacity. Intraoperatively, minute-by-minute vital signs are captured, in addition to fluid balances, ventilator parameters, and anesthetic medications administered. All data fields are alphanumeric and are captured in a uniform and granular manner allowing for easy coding and data analysis. Reports from MetaVision® are commonly used to support many patient safety and quality improvement initiatives in addition to numerous research studies.

Postoperative outcome data will be obtained from Sunrise Clinical Manager (Allscripts, Chicago, IL), the electronic medical record currently used for inpatient care at Barnes-Jewish Hospital. Data will also be obtained from several registries, including the Systematic Assessment and Targeted Improvement of Services Following Yearlong Surgical Outcomes Surveys (SATISFY-SOS) patient-reported outcomes registry (NCT02032030), the National Surgical Quality Improvement Program (NSQIP) database, the Society of Thoracic Surgeons (STS) database. Preoperative and postoperative laboratory values will be obtained from the Center for Biomedical Informatics at Washington University, which hosts the data repository where these data are stored once they are processed by the laboratory. A data dictionary has been included as an online appendix detailing all the data elements that will be captured for this study.

The specific outcomes that will be predicted by the forecasting algorithms will include in-hospital mortality, postoperative acute kidney injury, and postoperative respiratory failure. In-hospital mortality will be ascertained from Sunrise Clinical Manager. Postoperative acute renal failure will be defined according to the KDIGO criteria[14]: an increase in serum creatinine of 0.3 mg/dL, increase in serum creatinine to 1.5 times the baseline value, or initiation of renal replacement therapy within 48 hours of surgery end time. Patients receiving renal replacement therapy prior to surgery, patients with no baseline creatinine available within 30 days prior to surgery, and patients undergoing kidney transplant or dialysis access procedures will be excluded from analysis of this outcome. Postoperative respiratory failure will be defined as mechanical ventilation for greater than 48 hours or unplanned postoperative intubation within 48 hours. These events will be extracted from clinical documentation recorded by respiratory therapists in Sunrise Clinical Manager.

Data Analysis, Part 1 – Forecasting Algorithm Development

We will develop hybrid learning techniques to combine the strength of nonparametric (generative) models such as histogram and kernel density estimation and parametric (discriminative) models such as support vector machines, logistic regressions, and kernel machines to improve predictions of adverse perioperative outcomes (in-hospital mortality, postoperative acute renal failure, postoperative respiratory failure). The goal is to deliver superior prediction quality with good interpretability and high computational efficiency that supports fast processing of big data. Based on our preliminary work using density-based logistic regression (DLR) to develop an early clinical deterioration warning system for patients in the general wards of Barnes-Jewish Hospital,[15 16] we propose to develop novel hybrid data mining/machine learning algorithms that exploit both non-parametric and parametric techniques. Also, the resulting algorithms can be viewed as hybrid generative/discriminative learning models.

DLR is a hybrid algorithm combining the distribution-free nonlinear separation ability of non-parametric (generative) models and the efficiency and interpretability of parametric (discriminative) models. It first applies a Nadaraya-Watson kernel density estimator, a non-parametric transformation, on the input data to extract features that conform best to the true distribution of data, and then applies the parametric logistic regression model on the transformed features. The resulting model exhibits five desirable properties: nonlinear separation ability, high efficiency, good interpretability, ability to handle mixed data types including numerical and categorical ones, and support for multi-way classification. Our previous results using Barnes-Jewish Hospital clinical data showed that DLR achieves better classification accuracy than state-of-the-art nonlinear classifiers such as support vector machines and KLR but is also much more efficient than nonlinear models.[17] In fact, DLR has the same asymptotic complexity as linear classifiers and can scale up to very large datasets in practice.[17]

To analyze the collected time-series data, we need to extract features that capture temporal patterns, such as a rapid temperature increases or abnormal heart rate fluctuations. We will first extract a large pool of time-series features including: first-order features such as variance, skewness, and kurtosis, and second-order features such as energy, entropy, correlation, inertia, and local homogeneity.[18 19] The second-order features are known to be robust under noises.[20 21] Self-similarity is widely observed in human physiological signs. Detrended fluctuation analysis [22] measures the degree of self-similarity in time series and has been applied to analyze heartbeat and oxygen levels.[23] Approximate entropy measures the degree of unpredictability in a time series.[24] Spectral analysis has also been used to analyze clinical time-series.[22] We will also consider cross-sign features including correlation,[25] coherence,[25] lagged regression, nonlinear regression,[19] and the synchronization index.[26] We will also extract features based on the bag-of-patterns approach[27-29] and autocorrelation.[30-32] In addition, we will also generate features based on shapelets.[33] A shapelet is a subseries that is used to compare against each time-series. For a shapelet with length l and a time series T, the shapelet gives a feature value which is the minimum Euclidean distance between the shapelet and any subseries of T with length l. Efficient methods have been developed to find good shapelets, based on length estimation and optimized search.[34-36]

We will also develop a novel deep learning method to extract more robust features from time-series. A leading method for feature selection from time series has been the shapelet method. However, we have shown that deep learning methods can significantly improve over shapelet. Deep learning methods, especially those using convolutional neural networks (CNNs),[37] have achieved great success in learning useful representations (features) from images.[38 39] However, its uses in time-series classification are very limited. We plan to apply CNNs to time-series data to generate good representations. We note that the convolutional layers in CNNs can be viewed as a collection of local filters over the input space; the filters' weights are learnt through back propagation. The filters in CNNs regulate the time series in different frequency bands, and the dot product operations in the CNNs measure distances between two subseries. Thus, CNNs can be viewed as a more general framework than shapelet learning which can adaptively find the suitable down-sampling rates and scales of the shapelets.

Our preliminary work has shown that it is beneficial to use a large feature set: the modeling accuracy increases as more features are used and the top features in the final model include features from different categories.[23] With the above features, we will address overfitting. An overfit model will generally have poor predictive performance and interpretability. We will investigate three schemes to avoid over-fitting including: 1) using feature selection methods, such as forward feature selection based on F-score or area-under-curve score,[40] to find the most discriminative features; 2) adding regularization terms (such as L1,[41] L2,[42] Akaike information criterion, Bayesian information criterion,[43] minimum description length,[44] or a probabilistic prior) to the optimization objective; and 3) using meta-techniques such as bootstrap aggregation [45] and exploratory undersampling [46] to further address overfitting and class imbalance.

We plan to develop novel classification algorithms that best fit our data. In our preliminary work, we proposed DLR, a novel nonlinear hybrid classification algorithm that integrates kernel density estimation with logistic regression. DLR can achieve nonlinear separability by utilizing a nonlinear feature transformation, but is much more efficient than other nonlinear models since it fits a linear model. It can naturally handle mixed data types. It also offers good interpretability. In this task, we plan to develop more powerful algorithms on top of DLR.

A key area of improvement is feature transformation. In DLR, we use the Nadaraya-Watson kernel density estimator for each data point in each dimension, which has time complexity of $O(mN^2)$ where m is the number of dimensions and N is the number of data points. Therefore, it is still slow for big datasets with a large N. We propose to use bin-based kernel density estimation, another non-parametric technique, to process the input features in each dimension. The idea is to divide each dimension into equal-sized bins and estimate the density for each bin instead of each data point. This will reduce the time into $O(mB^2)$ where $B \ll N$ is the number of bins. Note that instead of using a simple histogram count for each bin, we will use a Gaussian kernel function to smooth the density estimation across bins. The time complexity can be further reduced to $O(mB)$ using techniques such as Gauss transformation.[47] Such dramatic reduction of computing time will enable us to process large datasets and perform quick model-building. We will also combine the kernel density estimator-based features with other parametric models such as Cox regression.

We will also develop efficient training algorithms. We will leverage a hierarchical optimization algorithm for training DLR,[17] which automatically learns free parameters in the model under a maximum likelihood framework. This optimization formulation not only learns the coefficients in the model, but also provides a way to automatically select the kernel bandwidth in the Nadaraya-Watson estimator or the bin size in the bin-based kernel density estimation, which is absent in previous work. We will also employ techniques including stochastic gradient descent [48] and its parallelized implementation [49] to further enhance the scalability of the training algorithm.

We will study another novel approach called group-based modeling. The idea is to first use a few key features to divide the patients into some major categories, and then train a separate classifier for each category. The intuition is that from clinical knowledge, we know that some different groups of patients have drastically different behaviors and should correspond to different statistical models. Mixing such vastly different groups together to train a single model may not give the best result. Therefore, it is instrumental to identify important sub-populations of patients, before we use sophisticated hybrid algorithms to accurately model the patients in each group. For a simple example, we can group the patients into a few age ranges, e.g., <45, 45-55, 56-65, etc. Although age can be used as a feature in a single classifier for all patients, such explicit division leads to multiple, more specific classifiers. It can be viewed as a hybrid algorithm combining a decision tree with other classifiers. We may also use metrics defined on multiple attributes to group the patients. Features that will be used as classifiers will include age, sex, and surgery type (cardiac versus non-cardiac). To systematically integrate such clinical knowledge into modeling, we plan to study hybrid models that are mixture of two or more classifiers. For example, we can construct a global decision tree whose nodes denote patient groups, where each group is modeled by a local classifier such as DLR. Different nodes may use different types of classifiers. Previous work on a similar idea has demonstrated improved performance [50] in an intensive care prognosis application.

Data Analysis, Part 2 – Forecasting Algorithm Validation

After algorithm development, the forecasting algorithms will be tested for accuracy of their predictive performances in two ways. First, algorithm validity will be tested within the training database using the bootstrap method. Second, the performance of the developed algorithms will be additionally validated prospectively (out-of-sample performance), using standard measures of model predictive accuracy, including measures of accuracy, precision and robustness.

The performance of any predictive modeling process is always evaluated by the accuracy of its predictions. Data mining techniques are used specifically to explain as well as forecast events. Their predictive accuracy needs therefore to be evaluated before they can be deployed and used for clinical decision-making. The proposed hybrid approach will be first tested using the bootstrap method.

In the bootstrap method, a large number of independent random samples are drawn with replacement from the entire database. These surrogate data sets are then used iteratively as the training sets for the development of the machine-learning algorithm, and the remaining data from the original sample are used for testing. The overall mean-squared prediction error and its variation is then used as an evaluation and test tools of stability of the algorithm development process.[51] We propose to draw 100 surrogate samples for this evaluation.

Additionally, we propose to perform a validation test of the predictive performance of the developed algorithms prospectively, using patient records that did not belong to the learning database. For this evaluation, we will apply the most commonly used criteria for predictive model performance, including accuracy (defined as the overall percentage of correct forecasts), precision (defined as the percentage of correctly forecasted events), and robustness (defined as predictive ability when data includes noise and missing values).

Prespecified Secondary Analyses

In addition to the primary algorithms described above (in-hospital mortality, postoperative acute kidney injury, and postoperative respiratory failure), we anticipate using the acquired data to develop prediction algorithms for additional outcomes. These outcomes are outlined in Table 1.

Table 1. Prespecified Secondary Outcomes

Data Source	**Outcome**
Sunrise Clinical Manager	- Thirty-day hospital readmission - Intensive care unit admission - Postoperative delirium
National Surgical Quality Improvement Program (NSQIP) database	- Thirty-day mortality - Thirty-day hospital readmission - Unplanned intubation - Postoperative sepsis - Postoperative myocardial infarction - Postoperative cerebrovascular accident - Postoperative pulmonary embolism - Postoperative deep vein thrombosis - Postoperative cardiac arrest requiring cardiopulmonary resuscitation

Society of Thoracic Surgeons database	- Thirty-day mortality - Thirty-day hospital readmission - Postoperative atrial fibrillation - Postoperative venous thromboembolism - Postoperative acute respiratory distress syndrome
SATISFY-SOS registry	- Patient-reported thirty-day readmission - Patient-reported postoperative myocardial infarction - Patient-reported postoperative cardiac arrest - Patient-reported postoperative heart failure - Patient-reported postoperative cerebrovascular accident - Patient-reported postoperative venous thromboembolism - Patient-reported postoperative respiratory arrest - Patient-reported postoperative pneumonia - Patient-reported severe postoperative pain lasting greater than one day - Patient-reported severe postoperative nausea and vomiting lasting greater than one day - Return to work 30 days after surgery - Quality of life 30 days after surgery - Ability to perform activities of daily living 30 days after surgery

DISCUSSION

Implications and Future Directions

We predict that the successful development of machine learning-based algorithms for predicting adverse postoperative outcomes will impact the perioperative care of surgical patients in important ways. Because our algorithms will utilize time-series data, we expect to be able to use them in real time to provide perioperative health care clinicians with dynamic predictions of their patients' risks for specific adverse outcomes. Because the features in our models will include modifiable risk factors such as blood pressure and concentrations of anesthetic agents, we believe clinicians will be able to make changes that may alter their patients' risk trajectories. To be feasible and efficient, we suggest that the forecasting algorithms could be incorporated into a telemedicine paradigm, such as an anesthesiology control tower for a perioperative suite. Once the forecasting algorithms are developed, we intend to conduct a randomized controlled trial to investigate whether implementation of the algorithms in the operating rooms leads to a reduction in the incidence of adverse postoperative outcomes. The incorporation of machine-learning forecasting algorithms into perioperative care will complement the expertise of clinicians, and has the potential to increase both safety and efficiency.

Strengths and Limitations

One of the greatest strengths of this project is the novel use of machine learning techniques to harness the abundant data in the perioperative electronic medical record. Unlike traditional risk prediction models, which utilize data from a single time point and therefore incorporate only a small fraction of the available information about the patient, our algorithms will take advantage of the rich time-series data generated in the operating rooms and, more broadly, in perioperative settings (e.g., preoperative assessment clinic, postoperative recovery area). Another strength is the efficiency of the proposed modeling techniques, which will need to quickly process large amounts of data. The use of group-based learning will increase the accuracy of the derived models by separating groups of patients who likely have different relationships between underlying features and the predicted outcomes.

This project does have limitations that should be noted. Because the forecasting algorithms will utilize large quantities of data, generalizability of the results and implementation of the algorithms at other health care facilities will depend upon the availability of high-quality input data. In particular, the preoperative evaluation and medical history may not be documented in an electronic format with discrete analyzable fields at some other institutions. Even when such data are available, differences in formatting will require caution during implementation at other hospitals.

Ethics and Dissemination

This study has been approved by the Human Research Protection Office at Washington University in St. Louis. As noted earlier in this document, a waiver of informed consent has been granted for all participants. This work will be funded largely by a grant from the National Science Foundation (award number 1622678) and from a grant from the Agency for Healthcare Research and Quality (R21 HS24581-01).

Once this investigation has been completed, we intend to publish the results in a peer-reviewed publication. We also intend to present the results of this work at professional conferences for both the anesthesiology and computer science communities. In accordance with the recent proposal from the International Committee of Medical Journal Editors, patient-level data will be made available within six months after publication of the primary manuscript.[52] Data will be provided to researchers who submit a methodologically sound research proposal including a protocol and statistical analysis plan.

REFERENCES

1. Kertai MD, Pal N, Palanca BJ, et al. Association of perioperative risk factors and cumulative duration of low bispectral index with intermediate-term mortality after cardiac surgery in the B-Unaware Trial. *Anesthesiology* 2010;112(5):1116-27. doi: 10.1097/ALN.0b013e3181d5e0a3 [published Online First: 2010/04/27]
2. Kertai MD, Palanca BJ, Pal N, et al. Bispectral index monitoring, duration of bispectral index below 45, patient risk factors, and intermediate-term mortality after noncardiac surgery in the B-Unaware Trial. *Anesthesiology* 2011;114(3):545-56. doi: 10.1097/ALN.0b013e31820c2b57 [published Online First: 2011/02/05]
3. Walsh M, Devereaux PJ, Garg AX, et al. Relationship between intraoperative mean arterial pressure and clinical outcomes after noncardiac surgery: toward an empirical definition of hypotension. *Anesthesiology* 2013;119(3):507-15. doi: 10.1097/ALN.0b013e3182a10e26 [published Online First: 2013/07/10]
4. Walsh M, Garg AX, Devereaux PJ, et al. The association between perioperative hemoglobin and acute kidney injury in patients having noncardiac surgery. *Anesthesia and analgesia* 2013;117(4):924-31. doi: 10.1213/ANE.0b013e3182a1ec84 [published Online First: 2013/09/12]
5. Devereaux PJ, Yang H, Yusuf S, 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. doi: 10.1016/s0140-6736(08)60601-7 [published Online First: 2008/05/16]
6. Kheterpal S, Tremper KK, Englesbe MJ, et al. Predictors of postoperative acute renal failure after noncardiac surgery in patients with previously normal renal function. *Anesthesiology* 2007;107(6):892-902. doi: 10.1097/01.anes.0000290588.29668.38 [published Online First: 2007/11/29]
7. Sharifpour M, Moore LE, Shanks AM, et al. Incidence, predictors, and outcomes of perioperative stroke in noncarotid major vascular surgery. *Anesthesia and analgesia* 2013;116(2):424-34. doi: 10.1213/ANE.0b013e31826a1a32 [published Online First: 2012/11/02]
8. Bhave PD, Goldman LE, Vittinghoff E, et al. Incidence, predictors, and outcomes associated with postoperative atrial fibrillation after major noncardiac surgery. *American heart journal* 2012;164(6):918-24. doi: 10.1016/j.ahj.2012.09.004 [published Online First: 2012/12/01]
9. Willingham MD, Karren E, Shanks AM, et al. Concurrence of Intraoperative Hypotension, Low Minimum Alveolar Concentration, and Low Bispectral Index Is Associated with Postoperative Death. *Anesthesiology* 2015;123(4):775-85. doi: 10.1097/aln.0000000000000822 [published Online First: 2015/08/13]
10. Willingham M, Ben Abdallah A, Gradwohl S, et al. Association between intraoperative electroencephalographic suppression and postoperative mortality. *British journal of anaesthesia* 2014;113(6):1001-8. doi: 10.1093/bja/aeu105 [published Online First: 2014/05/24]
11. Aranake A, Gradwohl S, Ben-Abdallah A, et al. Increased risk of intraoperative awareness in patients with a history of awareness. *Anesthesiology* 2013;119(6):1275-83. doi: 10.1097/aln.0000000000000023 [published Online First: 2013/10/12]
12. Sessler DI, Sigl JC, Kelley SD, et al. Hospital stay and mortality are increased in patients having a "triple low" of low blood pressure, low bispectral index, and low minimum alveolar concentration of volatile anesthesia. *Anesthesiology* 2012;116(6):1195-203. doi: 10.1097/ALN.0b013e31825683dc [published Online First: 2012/05/02]

13. Wang Y, Chen W, Heard K, et al. Mortality Prediction in ICUs Using A Novel Time-Slicing Cox Regression Method. *AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium* 2015;2015:1289-95. [published Online First: 2015/01/01]
14. Kellum JA, Lameire N. Diagnosis, evaluation, and management of acute kidney injury: a KDIGO summary (Part 1). *Critical care* 2013;17(1):1.
15. Hackmann G, Chen M, Chipara O, et al. Toward a two-tier clinical warning system for hospitalized patients. *AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium* 2011;2011:511-9. [published Online First: 2011/12/24]
16. Bailey TC, Chen Y, Mao Y, et al. A trial of a real-time alert for clinical deterioration in patients hospitalized on general medical wards. *Journal of hospital medicine : an official publication of the Society of Hospital Medicine* 2013;8(5):236-42. doi: 10.1002/jhm.2009 [published Online First: 2013/02/27]
17. Density-based logistic regression. Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining; 2013. ACM.
18. Dowdy S, Weardon S. Statistics for Research. Hoboken, NJ: Wiley 1983.
19. George A, Wild C. Nonlinear Regression. Hoboken, NJ: Wiley 2003.
20. Haralick RM, Shanmugam K. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics* 1973(6):610-21.
21. Arthanari T, Dodge Y. Mathematical Programming in Statistics. Hoboken, NJ: Wiley 1993.
22. Penzel T, Kantelhardt JW, Grote L, et al. Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. *IEEE transactions on bio-medical engineering* 2003;50(10):1143-51. doi: 10.1109/tbme.2003.817636 [published Online First: 2003/10/17]
23. An integrated data mining approach to real-time clinical monitoring and deterioration warning. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining; 2012. ACM.
24. Pincus SM. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences of the United States of America* 1991;88(6):2297-301. [published Online First: 1991/03/15]
25. Heart rate and respiration relationships as a diagnostic tool for late onset sepsis in sick preterm infants. 2006 Computers in Cardiology; 2006. IEEE.
26. Synchronization index for quantifying nonlinear causal coupling between RR interval and systolic arterial pressure after myocardial infarction. Computers in Cardiology 2000; 2000. IEEE.
27. Lin J, Khade R, Li Y. Rotation-invariant similarity in time series using bag-of-patterns representation. *Journal of Intelligent Information Systems* 2012;39(2):287-315.
28. Baydogan MG, Runger G, Tuv E. A bag-of-features framework to classify time series. *IEEE transactions on pattern analysis and machine intelligence* 2013;35(11):2796-802. doi: 10.1109/tpami.2013.72 [published Online First: 2013/09/21]
29. Deng H, Runger G, Tuv E, et al. A time series forest for classification and feature extraction. *Information Sciences* 2013;239:142-53.
30. Transformation Based Ensembles for Time Series Classification. SDM; 2012. SIAM.
31. Caiado J, Crato N, Peña D. A periodogram-based metric for time series classification. *Computational Statistics & Data Analysis* 2006;50(10):2668-84.
32. Bagnall A, Janacek G. A run length transformation for discriminating between auto regressive time series. *Journal of classification* 2014;31(2):154-78.
33. Time series shapelets: a new primitive for data mining. Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining; 2009. ACM.

34. A shapelet transform for time series classification. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining; 2012. ACM.
35. Learning time-series shapelets. Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining; 2014. ACM.
36. Fast shapelets: A scalable algorithm for discovering time series shapelets. Proceedings of the 13th SIAM international conference on data mining; 2013. SIAM.
37. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics* 1980;36(4):193-202.
38. Facenet: A unified embedding for face recognition and clustering. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015.
39. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Proceedings of the IEEE International Conference on Computer Vision; 2015.
40. An integrated machine learning approach to stroke prediction. Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining; 2010. ACM.
41. Shi J, Yin W, Osher S, et al. A fast hybrid algorithm for large-scale l1-regularized logistic regression. *Journal of Machine Learning Research* 2010;11(Feb):713-41.
42. Moore RC, DeNero J. L1 and L2 regularization for multiclass hinge loss models. 2011
43. Schwarz G. Estimating the dimension of a model. *The annals of statistics* 1978;6(2):461-64.
44. Grünwald PD. The minimum description length principle: MIT press 2007.
45. Breiman L. Bagging predictors. *Machine learning* 1996;24(2):123-40.
46. Liu X-Y, Wu J, Zhou Z-H. Exploratory undersampling for class-imbalance learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 2009;39(2):539-50.
47. Elgammal A, Duraiswami R, Davis LS. Efficient kernel density estimation using the fast gauss transform with applications to color modeling and tracking. *IEEE transactions on pattern analysis and machine intelligence* 2003;25(11):1499-504.
48. Ferguson TS. An inconsistent maximum likelihood estimate. *Journal of the American Statistical Association* 1982;77(380):831-34.
49. Parallelized stochastic gradient descent. Advances in neural information processing systems; 2010.
50. Abu-Hanna A, de Keizer N. Integrating classification trees with local logistic regression in Intensive Care prognosis. *Artificial Intelligence in Medicine* 2003;29(1):5-23.
51. Efron B, Tibshirani RJ. An introduction to the bootstrap Chapman & Hall. *New York* 1993;436
52. Taichman DB, Backus J, Baethge C, et al. Sharing Clinical Trial Data--A Proposal from the International Committee of Medical Journal Editors. *The New England journal of medicine* 2016;374(4):384-6. doi: 10.1056/NEJMe1515172 [published Online First: 2016/01/21]

AUTHORS' CONTRIBUTIONS

Bradley A Fritz, MD, contributed to overall study design, initial draft of protocol, and critical revision of protocol.

Yixin Chen, PhD, contributed to development of methods for creation of forecasting algorithms

Teresa M Murray-Torres, MD, contributed to study design and critical revision of protocol

Stephen Gregory, MD, contributed to study design and critical revision of protocol

Arbi Ben Abdallah, PhD, contributed to statistical methods for validation of forecasting algorithms and to critical revision of protocol

Alex Kronzer contributed to study design and critical revision of protocol

Sherry McKinnon contributed to study design and critical revision of protocol

Thaddeus Budelier contributed to critical revision of protocol

Daniel L Helsten, MD, contributed to critical revision of protocol

Troy S Wildes, MD, contributed to study design and critical revision of protocol

Anshuman Sharma, MD, contributed to study design and critical revision of protocol

Michael S Avidan, MBBCh, contributed to overall study design and critical revision of protocol

FUNDING STATEMENT

This work will be funded by a grant from the National Science Foundation (award number 1622678) and from a grant from the Agency for Healthcare Research and Quality (R21 HS24581-01).

COMPETING INTERESTS STATEMENT

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: all authors had financial support from the National Science Foundation and the Agency for Healthcare Research and Quality for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work."

BMJ Open

Protocol for a Retrospective Study Using Machine Learning Techniques to Develop Forecasting Algorithms for Postoperative Complications: The ACTFAST-2 Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020124.R1
Article Type:	Protocol
Date Submitted by the Author:	18-Dec-2017
Complete List of Authors:	Fritz, Bradley; Washington University School of Medicine in St. Louis, Anesthesiology Chen, Yixin; Washington University in Saint Louis, Computer Science and Engineering Murray-Torres, Teresa; Washington University School of Medicine in St. Louis, Anesthesiology Gregory, SH; Washington University School of Medicine in St. Louis, Anesthesiology Ben Abdallah, Arbi; Washington University School of Medicine in St. Louis, Anesthesiology Kronzer, Alex; Washington University School of Medicine in St. Louis, Anesthesiology McKinnon, Sherry; Washington University School of Medicine, Anesthesiology Budelier, Thaddeus; Washington University School of Medicine in St. Louis, Anesthesiology Helsten, Daniel; Washington University School of Medicine in St. Louis, Anesthesiology Wildes, Troy; Washington University in Saint Louis School of Medicine, Anesthesiology Sharma, Anshuman; Washington University School of Medicine in St. Louis, Anesthesiology Avidan, Michael; Washington University School of Medicine, Anesthesiology
Primary Subject Heading:	Anaesthesia
Secondary Subject Heading:	Health informatics
Keywords:	Adult anaesthesia < ANAESTHETICS, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.
Data Dictionary v4 Tab1 - Preop.xls Data Dictionary v4 Tab2 - Intraop.xls

Data Dictionary v4 Tab3 - SATISFY-SOS Outcomes.xls
Data Dictionary v4 Tab4 - Sunrise Outcomes.xls

Protocol for a Retrospective Study Using Machine Learning Techniques to Develop Forecasting Algorithms for Postoperative Complications: The ACTFAST-2 Study

Bradley A Fritz, MD, Resident Physician[1]

Yixin Chen, PhD, Professor[2]

Teresa M Murray-Torres, MD, Resident Physician[1]

Stephen Gregory, MD, Instructor[1]

Arbi Ben Abdallah, PhD, Assistant Professor[1]

Alex Kronzer, Database Analyst[1]

Sherry McKinnon, Project Manager[1]

Thaddeus Budelier, Research Coordinator[1]

Daniel L Helsten, MD, Associate Professor[1]

Troy S Wildes, MD, Associate Professor[1]

Anshuman Sharma, MD, Professor[1]

Michael S Avidan, MBBCh, Professor[1]

1 – Department of Anesthesiology, Washington University in St. Louis, St. Louis, MO, USA

2 – Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO, USA

Corresponding Author: Bradley Fritz

1 Barnes Jewish Hospital Plaza

St. Louis, MO, 63108

(314) 747-4155

bafritz@wustl.edu

ABSTRACT

Introduction – Mortality and morbidity following surgery are pressing public health concerns in the United States. Traditional prediction models for postoperative adverse outcomes demonstrate good discrimination at the population level, but the ability to forecast an individual patient's trajectory in real time remains poor. We propose to apply machine-learning techniques to perioperative time-series data to develop algorithms for predicting adverse perioperative outcomes.

Methods and Analysis – This study will include all adult patients who had surgery at our tertiary care hospital over a four-year period. Patient history, laboratory values, minute-by-minute intraoperative vital signs, and medications administered will be extracted from the electronic medical record. Outcomes will include in-hospital mortality, postoperative acute kidney injury, and postoperative respiratory failure. Forecasting algorithms for each of these outcomes will be constructed using density-based logistic regression after employing a Nadaraya-Watson kernel density estimator. Time-series variables will be analyzed using first- and second-order feature extraction, shapelet methods, and convolutional neural networks. The algorithms will be validated through measurement of precision and recall.

Ethics and Dissemination – This study has been approved by the Human Research Protection Office at Washington University in St. Louis. The successful development of these forecasting algorithms will allow perioperative health care clinicians to predict more accurately an individual patient's risk for specific adverse perioperative outcomes in real time. Knowledge of a patient's dynamic risk profile may allow clinicians to make targeted changes in the care plan that will alter the patient's outcome trajectory. This hypothesis will be tested in a future randomized controlled trial.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- Will utilize modeling techniques that take advantage of the rich time-series data that are available, rather than data from a single time point
- Will utilize efficient modeling techniques that can process large amounts of data quickly
- Will utilize group-based learning to increase model accuracy by separating groups of patients who likely have different relationship between underlying features and predicted outcomes
- Dissemination to other health care facilities may be limited by the availability of high-quality preoperative and intraoperative input data in a usable format

INTRODUCTION

An estimated 40 million people undergo surgery every year in the United States. Postoperative mortality rate at one year for surgical inpatients is between 5 and 10 percent,(1, 2) and an estimated 10 percent of surgical patients suffer major in-hospital morbidity.(3-8) Perioperative morbidity and mortality are therefore pressing public health concerns. Many patient characteristics, including comorbid medical conditions, associate strongly and independently with perioperative mortality and major morbidity.(1, 2, 9-11) While many of these characteristics are not modifiable, some perioperative risk factors, such as intraoperative blood pressures and anesthetic concentrations,(1, 2, 9, 10) can be modified in real time. Although the association between perioperative variables and postoperative outcomes has been well established at the population level using approaches such as standard logistic regression, (1, 2, 9, 10, 12) the ability to utilize deviations in physiological parameters in real time to dynamically forecast the trajectory of each individual patient remains poor.

There is a gap in the field with an opportunity to assess the potential utility of machine learning-based forecasting algorithms to anticipate adverse perioperative outcomes, guide interventions, and improve overall quality of care. Standard forecasting models, such as logistic regression, linear regression, and other statistical modeling procedures, have long been used to identify and prioritize risk factors for adverse outcomes. Although most of these statistical techniques have been shown to have moderate predictive values, they are limited in their prognostic ability and practical use.(1, 2, 6, 9, 10) In contrast to standard forecasting models, we have demonstrated machine learning and data mining approaches for patients on intensive care units that generate markedly superior prediction for outcomes such as mortality.(13) Our methods differ from standard statistical techniques in their ability to effectively incorporate time-series data. Most standard modeling techniques for surgical patients are based on a snapshot scheme, which only considers the data values at a given moment. They are not competent in extracting features from time-series data, especially in real-time fashion, such as temporal trends and shapes. Therefore, the objective of this study is to utilize machine-learning techniques to build forecasting algorithms that use patient characteristics and high-fidelity intraoperative time-series data to predict adverse perioperative outcomes.

METHODS AND ANALYSIS

Study Design

Our central hypothesis is that with sufficient knowledge of patient characteristics coupled with repeated, high-fidelity time series data from the perioperative electronic medical record, advanced models can be constructed for individual patients that will forecast adverse perioperative outcomes. To test this hypothesis, we will conduct an observational cohort study of adult patients who undergo surgery at Barnes-Jewish Hospital in St. Louis, Missouri. First, we plan to develop forecasting algorithms for specific adverse perioperative outcomes using historical data. Next, we plan to validate these algorithms by determining whether they can be used to reliably forecast individual adverse perioperative outcomes.

Patient Population and Sample Size

This study will include all adult patients who had surgery in the 48 operating rooms at Barnes-Jewish Hospital in St. Louis, Missouri between June 1, 2012 and August 31, 2016. Patients who receive anesthesia care in areas outside the main operating rooms, such as the obstetric suite or the outpatient surgery suite, will not be included. Barnes-Jewish Hospital is a 1,252-bed academic university-affiliated adult tertiary care hospital, performing approximately 19,000 surgeries a year. We therefore anticipate that gathering data from a 4.25-year period will lead to a total sample size of approximately 80,000-90,000 surgeries for algorithm development and validation.

The Human Research Protection Office at Washington University in St. Louis has granted a waiver of informed consent for all subjects enrolled in this study. This study has been determined to involve no more than minimal risk to participants, as no additional data will be collected beyond that already contained in the electronic record. For the same reason, the waiver of consent will not adversely affect the participants' rights and welfare. It is impracticable to conduct this research without a waiver of consent because 100% participation from the patients is imperative to obtain scientifically sound data.

Data Acquisition

For this project, we will use a variety of electronic medical record sources to cover the entire perioperative period. Much of the relevant information will be imported from MetaVision®(iMDsoft, Wakefield, MA), an anesthesiology information management software system that is the perioperative electronic clinical documentation system currently utilized by the Department of Anesthesiology. MetaVision® captures comprehensive clinical data beginning with the preoperative assessment and continuing throughout the duration of the perioperative period. Information captured preoperatively includes patients' past medical and surgical histories, chronic medical issues, medications used, and functional capacity. Intraoperatively, minute-by-minute vital signs are captured, in addition to fluid balances, ventilator parameters, and anesthetic medications administered. Blood pressure measurements are available at intervals ranging from once per minute to once every five minutes, while other vital signs are captured once per minute. Thus, a three-hour procedure would have about 180 measurements for each vital sign. All data fields are alphanumeric and are captured in a uniform and granular manner allowing for easy coding and data analysis. Reports from MetaVision® are commonly used to support many patient safety and quality improvement initiatives in addition to numerous research studies.

Postoperative outcome data will be obtained from Sunrise Clinical Manager (Allscripts, Chicago, IL), the electronic medical record currently used for inpatient care at Barnes-Jewish Hospital. Data will also be obtained from several registries, including the Systematic Assessment and Targeted Improvement of Services Following Yearlong Surgical Outcomes Surveys (SATISFY-SOS) patient-reported outcomes registry (NCT02032030), the National Surgical Quality Improvement Program (NSQIP) database, the Society of Thoracic Surgeons (STS) database. Preoperative and postoperative laboratory values will be obtained from the Center for Biomedical Informatics at Washington University, which hosts the data repository where these data are stored once they are processed by the laboratory. In general, a preoperative complete blood count is available if the patient is undergoing major surgery with potential significant blood loss or if other clinical reasons are present. Electrolytes and renal function are available if there is clinical reason to suspect an abnormality (including, but not limited to, patients with hypertension, diabetes mellitus, or chronic kidney disease). Additional tests, such as hepatic function and coagulation studies, are available on smaller sets of patients in whom the tests are clinically indicated. A data dictionary has been included as supplementary files (Data Dictionary v4 Tab1, Tab2, Tab3, and Tab4) detailing all the data elements that will be captured for this study.

The specific outcomes that will be predicted by the forecasting algorithms will include in-hospital mortality, postoperative acute kidney injury, and postoperative respiratory failure. In-hospital mortality will be ascertained from Sunrise Clinical Manager. Postoperative acute renal failure will be defined according to the KDIGO criteria(14): an increase in serum creatinine of 0.3 mg/dL, increase in serum creatinine to 1.5 times the baseline value, or initiation of renal replacement therapy within 48 hours of surgery end time. Patients receiving renal replacement therapy prior to surgery, patients with no baseline creatinine available within 30 days prior to surgery, and patients undergoing kidney transplant or dialysis access procedures will be excluded from analysis of this outcome. Postoperative respiratory failure will be defined as mechanical ventilation for greater than 48 hours or unplanned postoperative intubation within 48 hours. These events will be extracted from clinical documentation recorded by respiratory therapists in Sunrise Clinical Manager. Patients receiving mechanical ventilation prior to surgery will be excluded from analysis of this outcome.

Data Analysis, Part 1 – Forecasting Algorithm Development

We will develop hybrid learning techniques to combine the strength of generative models such as histogram and kernel density estimation and discriminative models such as support vector machines, logistic regressions, and kernel machines to improve predictions of adverse perioperative outcomes (in-hospital mortality, postoperative acute renal failure, postoperative respiratory failure). The goal is to deliver superior prediction quality with good interpretability and high computational efficiency that supports fast processing of big data. Based on our preliminary work using density-based logistic regression (DLR) to develop an early clinical deterioration warning system for patients in the general wards of Barnes-Jewish Hospital,(15, 16) we propose to develop novel hybrid data mining/machine learning algorithms that exploit both non-parametric and parametric techniques. For each target outcome, we plan to develop a model that will predict the likelihood of the postoperative outcome in real time using preoperative features and time-series data from the preceding 60 minutes.

DLR first applies a Nadaraya-Watson kernel density estimator, a non-parametric transformation, on the input data to extract features that conform best to the true distribution of data, and then applies the parametric logistic regression model on the transformed features. The resulting model exhibits five desirable properties: nonlinear separation ability, high efficiency, good interpretability, ability to handle mixed data types including numerical and categorical ones, and support for multi-way classification. Our previous results using Barnes-Jewish Hospital clinical data showed that DLR achieves better classification accuracy than state-of-the-art nonlinear classifiers such as support vector machines and KLR but is also much more efficient than nonlinear models.(17) In fact, DLR has the same asymptotic complexity as linear classifiers and can scale up to very large datasets in practice.(17)

To analyze the collected time-series data, we need to extract features that capture temporal patterns, such as a rapid temperature increases or abnormal heart rate fluctuations. To make predictions at a given point in time, time-series values from the preceding 60 minutes will be used. Missing values will be handled using linear interpolation. We will first extract a large pool of time-series features including: first-order features such as variance, skewness, and kurtosis, and second-order features such as energy, entropy, correlation, inertia, and local homogeneity.(18, 19) The second-order features are known to be robust under noises.(20, 21) Self-similarity is widely observed in human physiological signs. Detrended fluctuation analysis (22) measures the degree of self-similarity in time series and has been applied to analyze heartbeat and oxygen levels.(23) Approximate entropy measures the degree of unpredictability in a time series.(24) Spectral analysis has also been used to analyze clinical time-series.(22) We will also consider cross-sign features including correlation,(25) coherence,(25) lagged regression, nonlinear regression,(19) and the synchronization index.(26) We will also extract features based on the bag-of-patterns approach(27-29) and autocorrelation.(30-32) In addition, we will also generate features based on shapelets.(33) A shapelet is a subseries that is used to compare against each time-series. For a shapelet with length l and a time series T, the shapelet gives a feature value which is the minimum Euclidean distance between the shapelet and any subseries of T with length l. Efficient methods have been developed to find good shapelets, based on length estimation and optimized search.(34-36)

We will also develop a novel deep learning method to extract more robust features from time-series. A leading method for feature selection from time series has been the shapelet method. However, we have shown that deep learning methods can significantly improve over shapelet. Deep learning methods, especially those using convolutional neural networks (CNNs),(37) have achieved great success in learning useful representations (features) from images.(38, 39) However, its uses in time-series classification are very limited. We plan to apply CNNs to time-series data to generate good representations. We note that the convolutional layers in CNNs can be viewed as a collection of local filters over the input space; the filters' weights are learnt through back propagation. The filters in CNNs regulate the time series in different frequency bands, and the dot product operations in the CNNs measure distances between two subseries. Thus, CNNs can be viewed as a more general framework than shapelet learning which can adaptively find the suitable down-sampling rates and scales of the shapelets.

Our preliminary work has shown that it is beneficial to use a large feature set: the modeling accuracy increases as more features are used and the top features in the final model include features from different categories.(23) With the above features, we will address overfitting. An overfit model will generally have poor predictive performance and interpretability. We will investigate three schemes to avoid over-fitting including: 1) using feature selection methods, such as forward feature selection based on F-score or area-under-curve score,(40) to find the most discriminative features; 2) adding regularization terms (such as L1,(41) L2,(42) Akaike information criterion, Bayesian information criterion,(43) minimum description length,(44) or a probabilistic prior) to the optimization objective; and 3) using meta-techniques such as bootstrap aggregation (45) and exploratory undersampling (46) to further address overfitting and class imbalance.

We plan to use bin-based kernel density estimation, another non-parametric technique, to process the input features in each dimension. In previously described DLR, we use the Nadaraya-Watson kernel density estimator for each data point in each dimension, which has time complexity of $O(mN^2)$ where m is the number of dimensions and N is the number of data points. Therefore, it is still slow for big datasets with a large N. Bin-based kernel density estimation differs from the Nadaraya-Watson kernel density estimator in that we divide each dimension into equal-sized bins and estimate the density for each bin instead of each data point. This will reduce the time into $O(mB^2)$ where $B<<N$ is the number of bins. Note that instead of using a simple histogram count for each bin, we will use a Gaussian kernel function to smooth the density estimation across bins. The time complexity can be

further reduced to O(*mB*) using techniques such as Gauss transformation.(47) Such dramatic reduction of computing time will enable us to process large datasets and perform quick model-building. We will also combine the kernel density estimator-based features with other parametric models such as Cox regression.

We will leverage a hierarchical optimization algorithm for training DLR,(17) which automatically learns free parameters in the model under a maximum likelihood framework. This optimization formulation not only learns the coefficients in the model, but also provides a way to automatically select the kernel bandwidth in the Nadaraya-Watson estimator or the bin size in the bin-based kernel density estimation, which is absent in previous work. We will also employ techniques including stochastic gradient descent (48) and its parallelized implementation (49) to further enhance the scalability of the training algorithm.

Our algorithm will utilize group-based modeling. The idea is to first use a few key features to divide the patients into some major categories, and then train a separate classifier for each category. The intuition is that from clinical knowledge, we know that some different groups of patients have drastically different behaviors and should correspond to different statistical models. Mixing such vastly different groups together to train a single model may not give the best result. Therefore, it is instrumental to identify important sub-populations of patients, before we use sophisticated hybrid algorithms to accurately model the patients in each group. For a simple example, we can group the patients into a few age ranges, e.g., <45, 45-55, 56-65, etc. Although age can be used as a feature in a single classifier for all patients, such explicit division leads to multiple, more specific classifiers. It can be viewed as a hybrid algorithm combining a decision tree with other classifiers. We may also use metrics defined on multiple attributes to group the patients. Features that will be used as classifiers will include age, sex, and surgery type (cardiac versus non-cardiac). To systematically integrate such clinical knowledge into modeling, we plan to study hybrid models that are mixture of two or more classifiers. For example, we can construct a global decision tree whose nodes denote patient groups, where each group is modeled by a local classifier such as DLR. Different nodes may use different types of classifiers. Previous work on a similar idea has demonstrated improved performance (50) in an intensive care prognosis application.

Data Analysis, Part 2 – Forecasting Algorithm Validation

After algorithm development, the forecasting algorithms will be tested for accuracy of their predictive performances in two ways. First, algorithm validity will be tested within the historical database by dividing the database into training, validation, and testing datasets. Second, the performance of the developed algorithms will be additionally validated prospectively (out-of-sample performance), using precision and recall.

For initial model training and validation, the historical database will be divided into a training dataset (60% of the database), a validation dataset (20% of the database), and a testing dataset (20% of the database). Because we expect that our target outcomes will be relatively rare events, overall classification accuracy is not likely to be a useful measure of model performance. Instead, we will use precision (true positives/[true positives + false positives]) and recall (true positives/[true positives + false negatives]). We will optimize model parameters using the training dataset. Then we will pre-specify our desired recall and use the validation dataset to select the decision threshold that leads to the highest precision without sacrificing our desired recall. Then we will apply our model to the testing dataset and report the observed precision and recall.

Additionally, we propose to perform a validation test of the predictive performance of the developed algorithms prospectively, using patient records that did not belong to the learning database.

For this evaluation, we will apply our model to the prospectively-collected data. We will report the observed precision and recall as measures of model performance.

Prespecified Secondary Analyses

In addition to the primary algorithms described above (in-hospital mortality, postoperative acute kidney injury, and postoperative respiratory failure), we anticipate using the acquired data to develop prediction algorithms for additional outcomes. These outcomes are outlined in Table 1.

Table 1. Prespecified Secondary Outcomes

Data Source	**Outcome**
Sunrise Clinical Manager	- Thirty-day hospital readmission - Intensive care unit admission - Postoperative delirium
National Surgical Quality Improvement Program (NSQIP) database	- Thirty-day mortality - Thirty-day hospital readmission - Unplanned intubation - Postoperative sepsis - Postoperative myocardial infarction - Postoperative cerebrovascular accident - Postoperative pulmonary embolism - Postoperative deep vein thrombosis - Postoperative cardiac arrest requiring cardiopulmonary resuscitation
Society of Thoracic	- Thirty-day mortality

Surgeons database	- Thirty-day hospital readmission - Postoperative atrial fibrillation - Postoperative venous thromboembolism - Postoperative acute respiratory distress syndrome
SATISFY-SOS registry	- Patient-reported thirty-day readmission - Patient-reported postoperative myocardial infarction - Patient-reported postoperative cardiac arrest - Patient-reported postoperative heart failure - Patient-reported postoperative cerebrovascular accident - Patient-reported postoperative venous thromboembolism - Patient-reported postoperative respiratory arrest - Patient-reported postoperative pneumonia - Patient-reported severe postoperative pain lasting greater than one day - Patient-reported severe postoperative nausea and vomiting lasting greater than one day - Return to work 30 days after surgery - Quality of life 30 days after surgery - Ability to perform activities of daily living 30 days after surgery

DISCUSSION

Implications and Future Directions

We anticipate that the successful development of machine learning-based algorithms for predicting adverse postoperative outcomes will impact the perioperative care of surgical patients in important ways. Because our algorithms will utilize time-series data, we expect to be able to use them in real time to provide perioperative health care clinicians with dynamic predictions of their patients' risks for specific adverse outcomes. Because the features in our models will include modifiable risk factors such as blood pressure and concentrations of anesthetic agents, we believe clinicians will be able to make changes that may alter their patients' risk trajectories. The models may also help clinicians make decisions regarding their patients' postoperative disposition (intensive care unit versus hospital ward; inpatient admission versus discharge). To be feasible and efficient, we suggest that the forecasting algorithms could be incorporated into a telemedicine paradigm, such as an anesthesiology control tower for a perioperative suite. Once the forecasting algorithms are developed, we intend to conduct a randomized controlled trial to investigate whether implementation of the algorithms in the operating rooms leads to a reduction in the incidence of adverse postoperative outcomes. The incorporation of machine-learning forecasting algorithms into perioperative care will complement the expertise of clinicians, and has the potential to increase both safety and efficiency.

Strengths and Limitations

One of the greatest strengths of this project is the novel use of machine learning techniques to harness the abundant data in the perioperative electronic medical record. Unlike traditional risk prediction models, which utilize data from a single time point and therefore incorporate only a small fraction of the available information about the patient, our algorithms will take advantage of the rich time-series data generated in the operating rooms and, more broadly, in perioperative settings (e.g., preoperative assessment clinic, postoperative recovery area). Another strength is the efficiency of the proposed modeling techniques, which will need to quickly process large amounts of data. The use of group-based learning will increase the accuracy of the derived models by separating groups of patients who likely have different relationships between underlying features and the predicted outcomes.

This project does have limitations that should be noted. Because the forecasting algorithms will utilize large quantities of data, generalizability of the results and implementation of the algorithms at other health care facilities will depend upon the availability of high-quality input data. In particular, the preoperative evaluation and medical history may not be documented in an electronic format with discrete analyzable fields at some other institutions. Even when such data are available, differences in formatting will require caution during implementation at other hospitals.

Ethics and Dissemination

This study has been approved by the Human Research Protection Office at Washington University in St. Louis. As noted earlier in this document, a waiver of informed consent has been granted for all participants. This work will be funded largely by a grant from the National Science Foundation (award number 1622678) and from a grant from the Agency for Healthcare Research and Quality (R21 HS24581-01).

Once this investigation has been completed, we intend to publish the results in a peer-reviewed publication. We also intend to present the results of this work at professional conferences for both the anesthesiology and computer science communities. In accordance with the recent proposal from the International Committee of Medical Journal Editors, patient-level data will be made available within six months after publication of the primary manuscript.(51) Data will be provided to researchers who submit a methodologically sound research proposal including a protocol and statistical analysis plan. No patient-identifying fields (including dates) will be included in the shared dataset. Age will be provided in years, unless the patient is older than 89 years. In this case, age will be reported as “>89 years.” Any dates will be presented as “number of days since index surgery.”

REFERENCES

1. Kertai MD, Pal N, Palanca BJ, Lin N, Searleman SA, Zhang L, et al. Association of perioperative risk factors and cumulative duration of low bispectral index with intermediate-term mortality after cardiac surgery in the B-Unaware Trial. Anesthesiology. 2010;112(5):1116-27.
2. Kertai MD, Palanca BJ, Pal N, Burnside BA, Zhang L, Sadiq F, et al. Bispectral index monitoring, duration of bispectral index below 45, patient risk factors, and intermediate-term mortality after noncardiac surgery in the B-Unaware Trial. Anesthesiology. 2011;114(3):545-56.
3. Walsh M, Devereaux PJ, Garg AX, Kurz A, Turan A, Rodseth RN, et al. Relationship between intraoperative mean arterial pressure and clinical outcomes after noncardiac surgery: toward an empirical definition of hypotension. Anesthesiology. 2013;119(3):507-15.
4. Walsh M, Garg AX, Devereaux PJ, Argalious M, Honar H, Sessler DI. The association between perioperative hemoglobin and acute kidney injury in patients having noncardiac surgery. Anesthesia and analgesia. 2013;117(4):924-31.
5. Devereaux PJ, Yang H, Yusuf S, Guyatt G, Leslie K, Villar JC, 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.
6. Kheterpal S, Tremper KK, Englesbe MJ, O'Reilly M, Shanks AM, Fetterman DM, et al. Predictors of postoperative acute renal failure after noncardiac surgery in patients with previously normal renal function. Anesthesiology. 2007;107(6):892-902.
7. Sharifpour M, Moore LE, Shanks AM, Didier TJ, Kheterpal S, Mashour GA. Incidence, predictors, and outcomes of perioperative stroke in noncarotid major vascular surgery. Anesthesia and analgesia. 2013;116(2):424-34.
8. Bhave PD, Goldman LE, Vittinghoff E, Maselli J, Auerbach A. Incidence, predictors, and outcomes associated with postoperative atrial fibrillation after major noncardiac surgery. American heart journal. 2012;164(6):918-24.
9. Willingham MD, Karren E, Shanks AM, O'Connor MF, Jacobsohn E, Kheterpal S, et al. Concurrence of Intraoperative Hypotension, Low Minimum Alveolar Concentration, and Low Bispectral Index Is Associated with Postoperative Death. Anesthesiology. 2015;123(4):775-85.
10. Willingham M, Ben Abdallah A, Gradwohl S, Helsten D, Lin N, Villafranca A, et al. Association between intraoperative electroencephalographic suppression and postoperative mortality. British journal of anaesthesia. 2014;113(6):1001-8.
11. Aranake A, Gradwohl S, Ben-Abdallah A, Lin N, Shanks A, Helsten DL, et al. Increased risk of intraoperative awareness in patients with a history of awareness. Anesthesiology. 2013;119(6):1275-83.
12. Sessler DI, Sigl JC, Kelley SD, Chamoun NG, Manberg PJ, Saager L, et al. Hospital stay and mortality are increased in patients having a "triple low" of low blood pressure, low bispectral index, and low minimum alveolar concentration of volatile anesthesia. Anesthesiology. 2012;116(6):1195-203.
13. Wang Y, Chen W, Heard K, Kollef MH, Bailey TC, Cui Z, et al. Mortality Prediction in ICUs Using A Novel Time-Slicing Cox Regression Method. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2015;2015:1289-95.
14. Kellum JA, Lameire N. Diagnosis, evaluation, and management of acute kidney injury: a KDIGO summary (Part 1). Critical care. 2013;17(1):1.
15. Hackmann G, Chen M, Chipara O, Lu C, Chen Y, Kollef M, et al. Toward a two-tier clinical warning system for hospitalized patients. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2011;2011:511-9.

16. Bailey TC, Chen Y, Mao Y, Lu C, Hackmann G, Micek ST, et al. A trial of a real-time alert for clinical deterioration in patients hospitalized on general medical wards. Journal of hospital medicine : an official publication of the Society of Hospital Medicine. 2013;8(5):236-42.
17. Chen W, Chen Y, Mao Y, Guo B, editors. Density-based logistic regression. Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining; 2013: ACM.
18. Dowdy S, Weardon S. Statistics for Research. Hoboken, NJ: Wiley; 1983.
19. George A, Wild C. Nonlinear Regression. Hoboken, NJ: Wiley; 2003.
20. Haralick RM, Shanmugam K. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics. 1973(6):610-21.
21. Arthanari T, Dodge Y. Mathematical Programming in Statistics. Hoboken, NJ: Wiley; 1993.
22. Penzel T, Kantelhardt JW, Grote L, Peter JH, Bunde A. Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. IEEE transactions on bio-medical engineering. 2003;50(10):1143-51.
23. Mao Y, Chen W, Chen Y, Lu C, Kollef M, Bailey T, editors. An integrated data mining approach to real-time clinical monitoring and deterioration warning. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining; 2012: ACM.
24. Pincus SM. Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(6):2297-301.
25. Loforte R, Carrault G, Mainardi L, Beuche A, editors. Heart rate and respiration relationships as a diagnostic tool for late onset sepsis in sick preterm infants. 2006 Computers in Cardiology; 2006: IEEE.
26. Nollo G, Faes L, Pellegrini B, Porta A, Antolini R, editors. Synchronization index for quantifying nonlinear causal coupling between RR interval and systolic arterial pressure after myocardial infarction. Computers in Cardiology 2000; 2000: IEEE.
27. Lin J, Khade R, Li Y. Rotation-invariant similarity in time series using bag-of-patterns representation. Journal of Intelligent Information Systems. 2012;39(2):287-315.
28. Baydogan MG, Runger G, Tuv E. A bag-of-features framework to classify time series. IEEE transactions on pattern analysis and machine intelligence. 2013;35(11):2796-802.
29. Deng H, Runger G, Tuv E, Vladimir M. A time series forest for classification and feature extraction. Information Sciences. 2013;239:142-53.
30. Bagnall A, Davis LM, Hills J, Lines J, editors. Transformation Based Ensembles for Time Series Classification. SDM; 2012: SIAM.
31. Caiado J, Crato N, Peña D. A periodogram-based metric for time series classification. Computational Statistics & Data Analysis. 2006;50(10):2668-84.
32. Bagnall A, Janacek G. A run length transformation for discriminating between auto regressive time series. Journal of classification. 2014;31(2):154-78.
33. Ye L, Keogh E, editors. Time series shapelets: a new primitive for data mining. Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining; 2009: ACM.
34. Lines J, Davis LM, Hills J, Bagnall A, editors. A shapelet transform for time series classification. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining; 2012: ACM.
35. Grabocka J, Schilling N, Wistuba M, Schmidt-Thieme L, editors. Learning time-series shapelets. Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining; 2014: ACM.
36. Rakthanmanon T, Keogh E, editors. Fast shapelets: A scalable algorithm for discovering time series shapelets. Proceedings of the 13th SIAM international conference on data mining; 2013: SIAM.
37. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological cybernetics. 1980;36(4):193-202.

38. Schroff F, Kalenichenko D, Philbin J, editors. Facenet: A unified embedding for face recognition and clustering. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015.
39. He K, Zhang X, Ren S, Sun J, editors. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Proceedings of the IEEE International Conference on Computer Vision; 2015.
40. Khosla A, Cao Y, Lin CC-Y, Chiu H-K, Hu J, Lee H, editors. An integrated machine learning approach to stroke prediction. Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining; 2010: ACM.
41. Shi J, Yin W, Osher S, Sajda P. A fast hybrid algorithm for large-scale l1-regularized logistic regression. Journal of Machine Learning Research. 2010;11(Feb):713-41.
42. Moore RC, DeNero J. L1 and L2 regularization for multiclass hinge loss models. 2011.
43. Schwarz G. Estimating the dimension of a model. The annals of statistics. 1978;6(2):461-4.
44. Grünwald PD. The minimum description length principle: MIT press; 2007.
45. Breiman L. Bagging predictors. Machine learning. 1996;24(2):123-40.
46. Liu X-Y, Wu J, Zhou Z-H. Exploratory undersampling for class-imbalance learning. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 2009;39(2):539-50.
47. Elgammal A, Duraiswami R, Davis LS. Efficient kernel density estimation using the fast gauss transform with applications to color modeling and tracking. IEEE transactions on pattern analysis and machine intelligence. 2003;25(11):1499-504.
48. Ferguson TS. An inconsistent maximum likelihood estimate. Journal of the American Statistical Association. 1982;77(380):831-4.
49. Zinkevich M, Weimer M, Li L, Smola AJ, editors. Parallelized stochastic gradient descent. Advances in neural information processing systems; 2010.
50. Abu-Hanna A, de Keizer N. Integrating classification trees with local logistic regression in Intensive Care prognosis. Artificial Intelligence in Medicine. 2003;29(1):5-23.
51. Taichman DB, Backus J, Baethge C, Bauchner H, de Leeuw PW, Drazen JM, et al. Sharing Clinical Trial Data--A Proposal from the International Committee of Medical Journal Editors. The New England journal of medicine. 2016;374(4):384-6.

AUTHORS' CONTRIBUTIONS

Bradley A Fritz, MD, contributed to overall study design, initial draft of protocol, and critical revision of protocol.

Yixin Chen, PhD, contributed to development of methods for creation of forecasting algorithms

Teresa M Murray-Torres, MD, contributed to study design and critical revision of protocol

Stephen Gregory, MD, contributed to study design and critical revision of protocol

Arbi Ben Abdallah, PhD, contributed to statistical methods for validation of forecasting algorithms and to critical revision of protocol

Alex Kronzer contributed to study design and critical revision of protocol

Sherry McKinnon contributed to study design and critical revision of protocol

Thaddeus Budelier contributed to critical revision of protocol

Daniel L Helsten, MD, contributed to critical revision of protocol

Troy S Wildes, MD, contributed to study design and critical revision of protocol

Anshuman Sharma, MD, contributed to study design and critical revision of protocol

Michael S Avidan, MBBCh, contributed to overall study design and critical revision of protocol

FUNDING STATEMENT

This work will be funded by a grant from the National Science Foundation (award number 1622678) and from a grant from the Agency for Healthcare Research and Quality (R21 HS24581-01).

COMPETING INTERESTS STATEMENT

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: all authors had financial support from the National Science Foundation and the Agency for Healthcare Research and Quality for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work."

BMJ Open

Protocol for a Retrospective Study Using Machine Learning Techniques to Develop Forecasting Algorithms for Postoperative Complications: The ACTFAST-2 Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020124.R2
Article Type:	Protocol
Date Submitted by the Author:	22-Jan-2018

Complete List of Authors:	Fritz, Bradley; Washington University School of Medicine in St. Louis, Anesthesiology Chen, Yixin; Washington University in Saint Louis, Computer Science and Engineering Murray-Torres, Teresa; Washington University School of Medicine in St. Louis, Anesthesiology Gregory, SH; Washington University School of Medicine in St. Louis, Anesthesiology Ben Abdallah, Arbi; Washington University School of Medicine in St. Louis, Anesthesiology Kronzer, Alex; Washington University School of Medicine in St. Louis, Anesthesiology McKinnon, Sherry; Washington University School of Medicine, Anesthesiology Budelier, Thaddeus; Washington University School of Medicine in St. Louis, Anesthesiology Helsten, Daniel; Washington University School of Medicine in St. Louis, Anesthesiology Wildes, Troy; Washington University in Saint Louis School of Medicine, Anesthesiology Sharma, Anshuman; Washington University School of Medicine in St. Louis, Anesthesiology Avidan, Michael; Washington University School of Medicine, Anesthesiology

Primary Subject Heading:	Anaesthesia
Secondary Subject Heading:	Health informatics
Keywords:	Adult anaesthesia < ANAESTHETICS, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

Protocol for a Retrospective Study Using Machine Learning Techniques to Develop Forecasting Algorithms for Postoperative Complications: The ACTFAST-2 Study

Bradley A Fritz, MD, Resident Physician[1]

Yixin Chen, PhD, Professor[2]

Teresa M Murray-Torres, MD, Resident Physician[1]

Stephen Gregory, MD, Instructor[1]

Arbi Ben Abdallah, PhD, Assistant Professor[1]

Alex Kronzer, Database Analyst[1]

Sherry McKinnon, Project Manager[1]

Thaddeus Budelier, Research Coordinator[1]

Daniel L Helsten, MD, Associate Professor[1]

Troy S Wildes, MD, Associate Professor[1]

Anshuman Sharma, MD, Professor[1]

Michael S Avidan, MBBCh, Professor[1]

1 – Department of Anesthesiology, Washington University in St. Louis, St. Louis, MO, USA

2 – Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO, USA

Corresponding Author: Bradley Fritz

1 Barnes Jewish Hospital Plaza

St. Louis, MO, 63108

(314) 747-4155

bafritz@wustl.edu

ABSTRACT

Introduction – Mortality and morbidity following surgery are pressing public health concerns in the United States. Traditional prediction models for postoperative adverse outcomes demonstrate good discrimination at the population level, but the ability to forecast an individual patient's trajectory in real time remains poor. We propose to apply machine-learning techniques to perioperative time-series data to develop algorithms for predicting adverse perioperative outcomes.

Methods and Analysis – This study will include all adult patients who had surgery at our tertiary care hospital over a four-year period. Patient history, laboratory values, minute-by-minute intraoperative vital signs, and medications administered will be extracted from the electronic medical record. Outcomes will include in-hospital mortality, postoperative acute kidney injury, and postoperative respiratory failure. Forecasting algorithms for each of these outcomes will be constructed using density-based logistic regression after employing a Nadaraya-Watson kernel density estimator. Time-series variables will be analyzed using first- and second-order feature extraction, shapelet methods, and convolutional neural networks. The algorithms will be validated through measurement of precision and recall.

Ethics and Dissemination – This study has been approved by the Human Research Protection Office at Washington University in St. Louis. The successful development of these forecasting algorithms will allow perioperative health care clinicians to predict more accurately an individual patient's risk for specific adverse perioperative outcomes in real time. Knowledge of a patient's dynamic risk profile may allow clinicians to make targeted changes in the care plan that will alter the patient's outcome trajectory. This hypothesis will be tested in a future randomized controlled trial.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- Will utilize modeling techniques that take advantage of the rich time-series data that are available, rather than data from a single time point
- Will utilize efficient modeling techniques that can process large amounts of data quickly
- Will utilize group-based learning to increase model accuracy by separating groups of patients who likely have different relationship between underlying features and predicted outcomes
- Dissemination to other health care facilities may be limited by the availability of high-quality preoperative and intraoperative input data in a usable format

INTRODUCTION

An estimated 40 million people undergo surgery every year in the United States. Postoperative mortality rate at one year for surgical inpatients is between 5 and 10 percent,(1, 2) and an estimated 10 percent of surgical patients suffer major in-hospital morbidity.(3-8) Perioperative morbidity and mortality are therefore pressing public health concerns. Many patient characteristics, including comorbid medical conditions, associate strongly and independently with perioperative mortality and major morbidity.(1, 2, 9-11) While many of these characteristics are not modifiable, some perioperative risk factors, such as intraoperative blood pressures and anesthetic concentrations,(1, 2, 9, 10) can be modified in real time. Although the association between perioperative variables and postoperative outcomes has been well established at the population level using approaches such as standard logistic regression, (1, 2, 9, 10, 12) the ability to utilize deviations in physiological parameters in real time to dynamically forecast the trajectory of each individual patient remains poor.

There is a gap in the field with an opportunity to assess the potential utility of machine learning-based forecasting algorithms to anticipate adverse perioperative outcomes, guide interventions, and improve overall quality of care. Standard forecasting models, such as logistic regression, linear regression, and other statistical modeling procedures, have long been used to identify and prioritize risk factors for adverse outcomes. Although most of these statistical techniques have been shown to have moderate predictive values, they are limited in their prognostic ability and practical use.(1, 2, 6, 9, 10) In contrast to standard forecasting models, we have demonstrated machine learning and data mining approaches for patients on intensive care units that generate markedly superior prediction for outcomes such as mortality.(13) Our methods differ from standard statistical techniques in their ability to effectively incorporate time-series data. Most standard modeling techniques for surgical patients are based on a snapshot scheme, which only considers the data values at a given moment. They are not competent in extracting features from time-series data, especially in real-time fashion, such as temporal trends and shapes. Therefore, the objective of this study is to utilize machine-learning techniques to build forecasting algorithms that use patient characteristics and high-fidelity intraoperative time-series data to predict adverse perioperative outcomes.

METHODS AND ANALYSIS

Study Design

Our central hypothesis is that with sufficient knowledge of patient characteristics coupled with repeated, high-fidelity time series data from the perioperative electronic medical record, advanced models can be constructed for individual patients that will forecast adverse perioperative outcomes. To test this hypothesis, we will conduct an observational cohort study of adult patients who undergo surgery at Barnes-Jewish Hospital in St. Louis, Missouri. First, we plan to develop forecasting algorithms for specific adverse perioperative outcomes using historical data. Next, we plan to validate these algorithms by determining whether they can be used to reliably forecast individual adverse perioperative outcomes.

Patient Population and Sample Size

This study will include all adult patients who had surgery in the 48 operating rooms at Barnes-Jewish Hospital in St. Louis, Missouri between June 1, 2012 and August 31, 2016. Patients who receive anesthesia care in areas outside the main operating rooms, such as the obstetric suite or the outpatient surgery suite, will not be included. Barnes-Jewish Hospital is a 1,252-bed academic university-affiliated adult tertiary care hospital, performing approximately 19,000 surgeries a year. We therefore anticipate that gathering data from a 4.25-year period will lead to a total sample size of approximately 80,000-81,000 surgeries for algorithm development and validation.

The Human Research Protection Office at Washington University in St. Louis has granted a waiver of informed consent for all subjects enrolled in this study. This study has been determined to involve no more than minimal risk to participants, as no additional data will be collected beyond that already contained in the electronic record. For the same reason, the waiver of consent will not adversely affect the participants' rights and welfare. It is impracticable to conduct this research without a waiver of consent because 100% participation from the patients is imperative to obtain scientifically sound data.

Data Acquisition

For this project, we will use a variety of electronic medical record sources to cover the entire perioperative period. Much of the relevant information will be imported from MetaVision®(iMDsoft, Wakefield, MA), an anesthesiology information management software system that is the perioperative electronic clinical documentation system currently utilized by the Department of Anesthesiology. MetaVision® captures comprehensive clinical data beginning with the preoperative assessment and continuing throughout the duration of the perioperative period. Information captured preoperatively includes patients' past medical and surgical histories, chronic medical issues, medications used, and functional capacity. Intraoperatively, minute-by-minute vital signs are captured, in addition to fluid balances, ventilator parameters, and anesthetic medications administered. Blood pressure measurements are available at intervals ranging from once per minute to once every five minutes, while other vital signs are captured once per minute. Thus, a three-hour procedure would have about 180 measurements for each vital sign. All data fields are alphanumeric and are captured in a uniform and granular manner allowing for easy coding and data analysis. Reports from MetaVision® are commonly used to support many patient safety and quality improvement initiatives in addition to numerous research studies.

Postoperative outcome data will be obtained from Sunrise Clinical Manager (Allscripts, Chicago, IL), the electronic medical record currently used for inpatient care at Barnes-Jewish Hospital. Data will also be obtained from several registries, including the Systematic Assessment and Targeted Improvement of Services Following Yearlong Surgical Outcomes Surveys (SATISFY-SOS) patient-reported outcomes registry (NCT02032030), the National Surgical Quality Improvement Program (NSQIP) database, the Society of Thoracic Surgeons (STS) database. Preoperative and postoperative laboratory values will be obtained from the Center for Biomedical Informatics at Washington University, which hosts the data repository where these data are stored once they are processed by the laboratory. In general, a preoperative complete blood count is available if the patient is undergoing major surgery with potential significant blood loss or if other clinical reasons are present. Electrolytes and renal function are available if there is clinical reason to suspect an abnormality (including, but not limited to, patients with hypertension, diabetes mellitus, or chronic kidney disease). Additional tests, such as hepatic function and coagulation studies, are available on smaller sets of patients in whom the tests are clinically indicated. A data dictionary has been included as supplementary files (Data Dictionary v4 Tab1, Tab2, Tab3, and Tab4) detailing all the data elements that will be captured for this study.

The specific outcomes that will be predicted by the forecasting algorithms will include in-hospital mortality, postoperative acute kidney injury, and postoperative respiratory failure. In-hospital mortality will be ascertained from Sunrise Clinical Manager. Postoperative acute renal failure will be defined according to the KDIGO criteria(14): an increase in serum creatinine of 0.3 mg/dL, increase in serum creatinine to 1.5 times the baseline value, or initiation of renal replacement therapy within 48 hours of surgery end time. Patients receiving renal replacement therapy prior to surgery, patients with no baseline creatinine available within 30 days prior to surgery, and patients undergoing kidney transplant or dialysis access procedures will be excluded from analysis of this outcome. Postoperative respiratory failure will be defined as mechanical ventilation for greater than 48 hours or unplanned postoperative intubation within 48 hours. These events will be extracted from clinical documentation recorded by respiratory therapists in Sunrise Clinical Manager. Patients receiving mechanical ventilation prior to surgery will be excluded from analysis of this outcome.

Data Analysis, Part 1 – Forecasting Algorithm Development

We will develop hybrid learning techniques to combine the strength of generative models such as histogram and kernel density estimation and discriminative models such as support vector machines, logistic regressions, and kernel machines to improve predictions of adverse perioperative outcomes (in-hospital mortality, postoperative acute renal failure, postoperative respiratory failure). The goal is to deliver superior prediction quality with good interpretability and high computational efficiency that supports fast processing of big data. Based on our preliminary work using density-based logistic regression (DLR) to develop an early clinical deterioration warning system for patients in the general wards of Barnes-Jewish Hospital,(15, 16) we propose to develop novel hybrid data mining/machine learning algorithms that exploit both non-parametric and parametric techniques. For each target outcome, we plan to develop a model that will predict the likelihood of the postoperative outcome in real time using preoperative features and time-series data from the preceding 60 minutes.

DLR first applies a Nadaraya-Watson kernel density estimator, a non-parametric transformation, on the input data to extract features that conform best to the true distribution of data, and then applies the parametric logistic regression model on the transformed features. The resulting model exhibits five desirable properties: nonlinear separation ability, high efficiency, good interpretability, ability to handle mixed data types including numerical and categorical ones, and support for multi-way classification. Our previous results using Barnes-Jewish Hospital clinical data showed that DLR achieves better classification accuracy than state-of-the-art nonlinear classifiers such as support vector machines and kernel logistic regression but is also much more efficient than nonlinear models.(17) In fact, DLR has the same asymptotic complexity as linear classifiers and can scale up to very large datasets in practice.(17)

To analyze the collected time-series data, we need to extract features that capture temporal patterns, such as a rapid temperature increases or abnormal heart rate fluctuations. To make predictions at a given point in time, time-series values from the preceding 60 minutes will be used. Missing values will be handled using linear interpolation. We will first extract a large pool of time-series features including: first-order features such as variance, skewness, and kurtosis, and second-order features such as energy, entropy, correlation, inertia, and local homogeneity.(18, 19) The second-order features are known to be robust under noises.(20, 21) Self-similarity is widely observed in human physiological signs. Detrended fluctuation analysis (22) measures the degree of self-similarity in time series and has been applied to analyze heartbeat and oxygen levels.(23) Approximate entropy measures the degree of unpredictability in a time series.(24) Spectral analysis has also been used to analyze clinical time-series.(22) We will also consider cross-sign features including correlation,(25) coherence,(25) lagged regression, nonlinear regression,(19) and the synchronization index.(26) We will also extract features based on the bag-of-patterns approach(27-29) and autocorrelation.(30-32) In addition, we will also generate features based on shapelets.(33) A shapelet is a subseries that is used to compare against each time-series. For a shapelet with length *l* and a time series *T*, the shapelet gives a feature value which is the minimum Euclidean distance between the shapelet and any subseries of *T* with length *l*. Efficient methods have been developed to find good shapelets, based on length estimation and optimized search.(34-36)

We will also develop a novel deep learning method to extract more robust features from time-series. A leading method for feature selection from time series has been the shapelet method. However, we have shown that deep learning methods can significantly improve over shapelet. Deep learning methods, especially those using convolutional neural networks (CNNs),(37) have achieved great success in learning useful representations (features) from images.(38, 39) However, its uses in time-series classification are very limited. We plan to apply CNNs to time-series data to generate good representations. We note that the convolutional layers in CNNs can be viewed as a collection of local filters over the input space; the filters' weights are learnt through back propagation. The filters in CNNs regulate the time series in different frequency bands, and the dot product operations in the CNNs measure distances between two subseries. Thus, CNNs can be viewed as a more general framework than shapelet learning which can adaptively find the suitable down-sampling rates and scales of the shapelets.

Our preliminary work has shown that it is beneficial to use a large feature set: the modeling accuracy increases as more features are used and the top features in the final model include features from different categories.(23) With the above features, we will address overfitting. An overfit model will generally have poor predictive performance and interpretability. We will investigate three schemes to avoid over-fitting including: 1) using feature selection methods, such as forward feature selection based on F-score or area-under-curve score,(40) to find the most discriminative features; 2) adding regularization terms (such as L1,(41) L2,(42) Akaike information criterion, Bayesian information criterion,(43) minimum description length,(44) or a probabilistic prior) to the optimization objective; and 3) using meta-techniques such as bootstrap aggregation (45) and exploratory undersampling (46) to further address overfitting and class imbalance.

We plan to use bin-based kernel density estimation, another non-parametric technique, to process the input features in each dimension. In previously described DLR, we use the Nadaraya-Watson kernel density estimator for each data point in each dimension, which has time complexity of $O(mN^2)$ where m is the number of dimensions and N is the number of data points. Therefore, it is still slow for big datasets with a large N. Bin-based kernel density estimation differs from the Nadaraya-Watson kernel density estimator in that we divide each dimension into equal-sized bins and estimate the density for each bin instead of each data point. This will reduce the time into $O(mB^2)$ where $B \ll N$ is the number of bins. Note that instead of using a simple histogram count for each bin, we will use a Gaussian kernel function to smooth the density estimation across bins. The time complexity can be further reduced to O(mB) using techniques such as Gauss transformation.(47) Such dramatic reduction of computing time will enable us to process large datasets and perform quick model-building. We will also combine the kernel density estimator-based features with other parametric models such as Cox regression.

We will leverage a hierarchical optimization algorithm for training DLR,(17) which automatically learns free parameters in the model under a maximum likelihood framework. This optimization formulation not only learns the coefficients in the model, but also provides a way to automatically select the kernel bandwidth in the Nadaraya-Watson estimator or the bin size in the bin-based kernel density estimation, which is absent in previous work. We will also employ techniques including stochastic gradient descent (48) and its parallelized implementation (49) to further enhance the scalability of the training algorithm.

Our algorithm will utilize group-based modeling. The idea is to first use a few key features to divide the patients into some major categories, and then train a separate classifier for each category. The intuition is that from clinical knowledge, we know that some different groups of patients have drastically different behaviors and should correspond to different statistical models. Mixing such vastly different groups together to train a single model may not give the best result. Therefore, it is instrumental to identify important sub-populations of patients, before we use sophisticated hybrid algorithms to accurately model the patients in each group. For a simple example, we can group the patients into a few age ranges, e.g., <45, 45-55, 56-65, etc. Although age can be used as a feature in a single classifier for all patients, such explicit division leads to multiple, more specific classifiers. It can be viewed as a hybrid algorithm combining a decision tree with other classifiers. We may also use metrics defined on multiple attributes to group the patients. Features that will be used as classifiers will include age, sex, and surgery type (cardiac versus non-cardiac). To systematically integrate such clinical knowledge into modeling, we plan to study hybrid models that are mixture of two or more classifiers. For example, we can construct a global decision tree whose nodes denote patient groups, where each group is modeled by a local classifier such as DLR. Different nodes may use different types of classifiers. Previous work on a similar idea has demonstrated improved performance (50) in an intensive care prognosis application.

Data Analysis, Part 2 – Forecasting Algorithm Validation

After algorithm development, the forecasting algorithms will be tested for accuracy of their predictive performances in two ways. First, algorithm validity will be tested within the historical database by dividing the database into training, validation, and testing datasets. Second, the performance of the developed algorithms will be additionally validated prospectively (out-of-sample performance), using precision and recall.

For initial model training and validation, the historical database will be divided into a training dataset (60% of the database), a validation dataset (20% of the database), and a testing dataset (20% of the database). Each training, validation, or testing example will be a 60-minute epoch randomly selected from a single surgery. More than one epoch from the same surgery may be included if the surgery lasted long enough to generate more than one distinct 60-minute epoch. However, all epochs from the same surgery will be included either all in the training dataset, all in the validation dataset, or all in the testing dataset. Because we expect that our target outcomes will be relatively rare events, overall classification accuracy is not likely to be a useful measure of model performance. Instead, we will use precision (true positives/[true positives + false positives]) and recall (true positives/[true positives + false negatives]). We will optimize model parameters using the training dataset. Then we will pre-specify our desired recall and use the validation dataset to select the decision threshold that leads to the highest precision without sacrificing our desired recall. Then we will apply our model to the testing dataset and report the observed precision and recall. The overall flow of algorithm training and validation is outlined in Figure 1.

Additionally, we propose to perform a validation test of the predictive performance of the developed algorithms prospectively, using patient records that did not belong to the learning database. For this evaluation, we will apply our model to the prospectively-collected data. We will report the observed precision and recall as measures of model performance.

Prespecified Secondary Analyses

In addition to the primary algorithms described above (in-hospital mortality, postoperative acute kidney injury, and postoperative respiratory failure), we anticipate using the acquired data to develop prediction algorithms for additional outcomes. These outcomes are outlined in Table 1.

Table 1. Prespecified Secondary Outcomes

Data Source	**Outcome**
Sunrise Clinical Manager	- Thirty-day hospital readmission - Intensive care unit admission - Postoperative delirium
National Surgical Quality Improvement Program (NSQIP) database	- Thirty-day mortality - Thirty-day hospital readmission - Unplanned intubation - Postoperative sepsis - Postoperative myocardial infarction - Postoperative cerebrovascular accident

	- Postoperative pulmonary embolism - Postoperative deep vein thrombosis - Postoperative cardiac arrest requiring cardiopulmonary resuscitation
Society of Thoracic Surgeons database	- Thirty-day mortality - Thirty-day hospital readmission - Postoperative atrial fibrillation - Postoperative venous thromboembolism - Postoperative acute respiratory distress syndrome
SATISFY-SOS registry	- Patient-reported thirty-day readmission - Patient-reported postoperative myocardial infarction - Patient-reported postoperative cardiac arrest - Patient-reported postoperative heart failure - Patient-reported postoperative cerebrovascular accident - Patient-reported postoperative venous thromboembolism - Patient-reported postoperative respiratory arrest - Patient-reported postoperative pneumonia - Patient-reported severe postoperative pain lasting greater than one day - Patient-reported severe postoperative nausea and vomiting lasting greater than one day - Return to work 30 days after surgery - Quality of life 30 days after surgery - Ability to perform activities of daily living 30 days after surgery

DISCUSSION

Implications and Future Directions

We anticipate that the successful development of machine learning-based algorithms for predicting adverse postoperative outcomes will impact the perioperative care of surgical patients in important ways. Because our algorithms will utilize time-series data, we expect to be able to use them in real time to provide perioperative health care clinicians with dynamic predictions of their patients' risks for specific adverse outcomes. Because the features in our models will include modifiable risk factors such as blood pressure and concentrations of anesthetic agents, we believe clinicians will be able to make changes that may alter their patients' risk trajectories. The models may also help clinicians make decisions regarding their patients' postoperative disposition (intensive care unit versus hospital ward; inpatient admission versus discharge). To be feasible and efficient, we suggest that the forecasting algorithms could be incorporated into a telemedicine paradigm, such as an anesthesiology control tower for a perioperative suite. Once the forecasting algorithms are developed, we intend to conduct a randomized controlled trial to investigate whether implementation of the algorithms in the operating rooms leads to a reduction in the incidence of adverse postoperative outcomes. The incorporation of machine-learning forecasting algorithms into perioperative care will complement the expertise of clinicians, and has the potential to increase both safety and efficiency.

Strengths and Limitations

One of the greatest strengths of this project is the novel use of machine learning techniques to harness the abundant data in the perioperative electronic medical record. Unlike traditional risk prediction models, which utilize data from a single time point and therefore incorporate only a small fraction of the available information about the patient, our algorithms will take advantage of the rich time-series data generated in the operating rooms and, more broadly, in perioperative settings (e.g., preoperative assessment clinic, postoperative recovery area). Another strength is the efficiency of the proposed modeling techniques, which will need to quickly process large amounts of data. The use of group-based learning will increase the accuracy of the derived models by separating groups of patients who likely have different relationships between underlying features and the predicted outcomes.

This project does have limitations that should be noted. Because the forecasting algorithms will utilize large quantities of data, generalizability of the results and implementation of the algorithms at other health care facilities will depend upon the availability of high-quality input data. In particular, the preoperative evaluation and medical history may not be documented in an electronic format with discrete analyzable fields at some other institutions. Even when such data are available, differences in formatting will require caution during implementation at other hospitals.

Ethics and Dissemination

This study has been approved by the Human Research Protection Office at Washington University in St. Louis. As noted earlier in this document, a waiver of informed consent has been granted for all participants. This work will be funded largely by a grant from the National Science Foundation (award number 1622678) and from a grant from the Agency for Healthcare Research and Quality (R21 HS24581-01).

Once this investigation has been completed, we intend to publish the results in a peer-reviewed publication. We also intend to present the results of this work at professional conferences for both the anesthesiology and computer science communities. In accordance with the recent proposal from the International Committee of Medical Journal Editors, patient-level data will be made available within six months after publication of the primary manuscript.(51) Data will be provided to researchers who submit a methodologically sound research proposal including a protocol and statistical analysis plan. No patient-identifying fields (including dates) will be included in the shared dataset. Age will be provided in years, unless the patient is older than 89 years. In this case, age will be reported as “>89 years.” Any dates will be presented as “number of days since index surgery.”

REFERENCES

1. Kertai MD, Pal N, Palanca BJ, Lin N, Searleman SA, Zhang L, et al. Association of perioperative risk factors and cumulative duration of low bispectral index with intermediate-term mortality after cardiac surgery in the B-Unaware Trial. Anesthesiology. 2010;112(5):1116-27.
2. Kertai MD, Palanca BJ, Pal N, Burnside BA, Zhang L, Sadiq F, et al. Bispectral index monitoring, duration of bispectral index below 45, patient risk factors, and intermediate-term mortality after noncardiac surgery in the B-Unaware Trial. Anesthesiology. 2011;114(3):545-56.
3. Walsh M, Devereaux PJ, Garg AX, Kurz A, Turan A, Rodseth RN, et al. Relationship between intraoperative mean arterial pressure and clinical outcomes after noncardiac surgery: toward an empirical definition of hypotension. Anesthesiology. 2013;119(3):507-15.
4. Walsh M, Garg AX, Devereaux PJ, Argalious M, Honar H, Sessler DI. The association between perioperative hemoglobin and acute kidney injury in patients having noncardiac surgery. Anesthesia and analgesia. 2013;117(4):924-31.
5. Devereaux PJ, Yang H, Yusuf S, Guyatt G, Leslie K, Villar JC, 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.
6. Kheterpal S, Tremper KK, Englesbe MJ, O'Reilly M, Shanks AM, Fetterman DM, et al. Predictors of postoperative acute renal failure after noncardiac surgery in patients with previously normal renal function. Anesthesiology. 2007;107(6):892-902.
7. Sharifpour M, Moore LE, Shanks AM, Didier TJ, Kheterpal S, Mashour GA. Incidence, predictors, and outcomes of perioperative stroke in noncarotid major vascular surgery. Anesthesia and analgesia. 2013;116(2):424-34.
8. Bhave PD, Goldman LE, Vittinghoff E, Maselli J, Auerbach A. Incidence, predictors, and outcomes associated with postoperative atrial fibrillation after major noncardiac surgery. American heart journal. 2012;164(6):918-24.
9. Willingham MD, Karren E, Shanks AM, O'Connor MF, Jacobsohn E, Kheterpal S, et al. Concurrence of Intraoperative Hypotension, Low Minimum Alveolar Concentration, and Low Bispectral Index Is Associated with Postoperative Death. Anesthesiology. 2015;123(4):775-85.
10. Willingham M, Ben Abdallah A, Gradwohl S, Helsten D, Lin N, Villafranca A, et al. Association between intraoperative electroencephalographic suppression and postoperative mortality. British journal of anaesthesia. 2014;113(6):1001-8.
11. Aranake A, Gradwohl S, Ben-Abdallah A, Lin N, Shanks A, Helsten DL, et al. Increased risk of intraoperative awareness in patients with a history of awareness. Anesthesiology. 2013;119(6):1275-83.
12. Sessler DI, Sigl JC, Kelley SD, Chamoun NG, Manberg PJ, Saager L, et al. Hospital stay and mortality are increased in patients having a "triple low" of low blood pressure, low bispectral index, and low minimum alveolar concentration of volatile anesthesia. Anesthesiology. 2012;116(6):1195-203.
13. Wang Y, Chen W, Heard K, Kollef MH, Bailey TC, Cui Z, et al. Mortality Prediction in ICUs Using A Novel Time-Slicing Cox Regression Method. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2015;2015:1289-95.
14. Kellum JA, Lameire N. Diagnosis, evaluation, and management of acute kidney injury: a KDIGO summary (Part 1). Critical care. 2013;17(1):1.
15. Hackmann G, Chen M, Chipara O, Lu C, Chen Y, Kollef M, et al. Toward a two-tier clinical warning system for hospitalized patients. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2011;2011:511-9.

16. Bailey TC, Chen Y, Mao Y, Lu C, Hackmann G, Micek ST, et al. A trial of a real-time alert for clinical deterioration in patients hospitalized on general medical wards. Journal of hospital medicine : an official publication of the Society of Hospital Medicine. 2013;8(5):236-42.
17. Chen W, Chen Y, Mao Y, Guo B, editors. Density-based logistic regression. Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining; 2013: ACM.
18. Dowdy S, Weardon S. Statistics for Research. Hoboken, NJ: Wiley; 1983.
19. George A, Wild C. Nonlinear Regression. Hoboken, NJ: Wiley; 2003.
20. Haralick RM, Shanmugam K. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics. 1973(6):610-21.
21. Arthanari T, Dodge Y. Mathematical Programming in Statistics. Hoboken, NJ: Wiley; 1993.
22. Penzel T, Kantelhardt JW, Grote L, Peter JH, Bunde A. Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. IEEE transactions on bio-medical engineering. 2003;50(10):1143-51.
23. Mao Y, Chen W, Chen Y, Lu C, Kollef M, Bailey T, editors. An integrated data mining approach to real-time clinical monitoring and deterioration warning. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining; 2012: ACM.
24. Pincus SM. Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(6):2297-301.
25. Loforte R, Carrault G, Mainardi L, Beuche A, editors. Heart rate and respiration relationships as a diagnostic tool for late onset sepsis in sick preterm infants. 2006 Computers in Cardiology; 2006: IEEE.
26. Nollo G, Faes L, Pellegrini B, Porta A, Antolini R, editors. Synchronization index for quantifying nonlinear causal coupling between RR interval and systolic arterial pressure after myocardial infarction. Computers in Cardiology 2000; 2000: IEEE.
27. Lin J, Khade R, Li Y. Rotation-invariant similarity in time series using bag-of-patterns representation. Journal of Intelligent Information Systems. 2012;39(2):287-315.
28. Baydogan MG, Runger G, Tuv E. A bag-of-features framework to classify time series. IEEE transactions on pattern analysis and machine intelligence. 2013;35(11):2796-802.
29. Deng H, Runger G, Tuv E, Vladimir M. A time series forest for classification and feature extraction. Information Sciences. 2013;239:142-53.
30. Bagnall A, Davis LM, Hills J, Lines J, editors. Transformation Based Ensembles for Time Series Classification. SDM; 2012: SIAM.
31. Caiado J, Crato N, Peña D. A periodogram-based metric for time series classification. Computational Statistics & Data Analysis. 2006;50(10):2668-84.
32. Bagnall A, Janacek G. A run length transformation for discriminating between auto regressive time series. Journal of classification. 2014;31(2):154-78.
33. Ye L, Keogh E, editors. Time series shapelets: a new primitive for data mining. Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining; 2009: ACM.
34. Lines J, Davis LM, Hills J, Bagnall A, editors. A shapelet transform for time series classification. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining; 2012: ACM.
35. Grabocka J, Schilling N, Wistuba M, Schmidt-Thieme L, editors. Learning time-series shapelets. Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining; 2014: ACM.
36. Rakthanmanon T, Keogh E, editors. Fast shapelets: A scalable algorithm for discovering time series shapelets. Proceedings of the 13th SIAM international conference on data mining; 2013: SIAM.
37. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological cybernetics. 1980;36(4):193-202.

38. Schroff F, Kalenichenko D, Philbin J, editors. Facenet: A unified embedding for face recognition and clustering. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015.
39. He K, Zhang X, Ren S, Sun J, editors. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Proceedings of the IEEE International Conference on Computer Vision; 2015.
40. Khosla A, Cao Y, Lin CC-Y, Chiu H-K, Hu J, Lee H, editors. An integrated machine learning approach to stroke prediction. Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining; 2010: ACM.
41. Shi J, Yin W, Osher S, Sajda P. A fast hybrid algorithm for large-scale l1-regularized logistic regression. Journal of Machine Learning Research. 2010;11(Feb):713-41.
42. Moore RC, DeNero J. L1 and L2 regularization for multiclass hinge loss models. 2011.
43. Schwarz G. Estimating the dimension of a model. The annals of statistics. 1978;6(2):461-4.
44. Grünwald PD. The minimum description length principle: MIT press; 2007.
45. Breiman L. Bagging predictors. Machine learning. 1996;24(2):123-40.
46. Liu X-Y, Wu J, Zhou Z-H. Exploratory undersampling for class-imbalance learning. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 2009;39(2):539-50.
47. Elgammal A, Duraiswami R, Davis LS. Efficient kernel density estimation using the fast gauss transform with applications to color modeling and tracking. IEEE transactions on pattern analysis and machine intelligence. 2003;25(11):1499-504.
48. Ferguson TS. An inconsistent maximum likelihood estimate. Journal of the American Statistical Association. 1982;77(380):831-4.
49. Zinkevich M, Weimer M, Li L, Smola AJ, editors. Parallelized stochastic gradient descent. Advances in neural information processing systems; 2010.
50. Abu-Hanna A, de Keizer N. Integrating classification trees with local logistic regression in Intensive Care prognosis. Artificial Intelligence in Medicine. 2003;29(1):5-23.
51. Taichman DB, Backus J, Baethge C, Bauchner H, de Leeuw PW, Drazen JM, et al. Sharing Clinical Trial Data--A Proposal from the International Committee of Medical Journal Editors. The New England journal of medicine. 2016;374(4):384-6.

FIGURE LEGEND

Figure 1. Data flow for algorithm training and validation using the historical database.

AUTHORS' CONTRIBUTIONS

Bradley A Fritz, MD, contributed to overall study design, initial draft of protocol, and critical revision of protocol.

Yixin Chen, PhD, contributed to development of methods for creation of forecasting algorithms

Teresa M Murray-Torres, MD, contributed to study design and critical revision of protocol

Stephen Gregory, MD, contributed to study design and critical revision of protocol

Arbi Ben Abdallah, PhD, contributed to statistical methods for validation of forecasting algorithms and to critical revision of protocol

Alex Kronzer contributed to study design and critical revision of protocol

Sherry McKinnon contributed to study design and critical revision of protocol

Thaddeus Budelier contributed to critical revision of protocol

Daniel L Helsten, MD, contributed to critical revision of protocol

Troy S Wildes, MD, contributed to study design and critical revision of protocol

Anshuman Sharma, MD, contributed to study design and critical revision of protocol

Michael S Avidan, MBBCh, contributed to overall study design and critical revision of protocol

FUNDING STATEMENT

This work will be funded by a grant from the National Science Foundation (award number 1622678) and from a grant from the Agency for Healthcare Research and Quality (R21 HS24581-01).

COMPETING INTERESTS STATEMENT

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: all authors had financial support from the National Science Foundation and the Agency for Healthcare Research and Quality for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work."

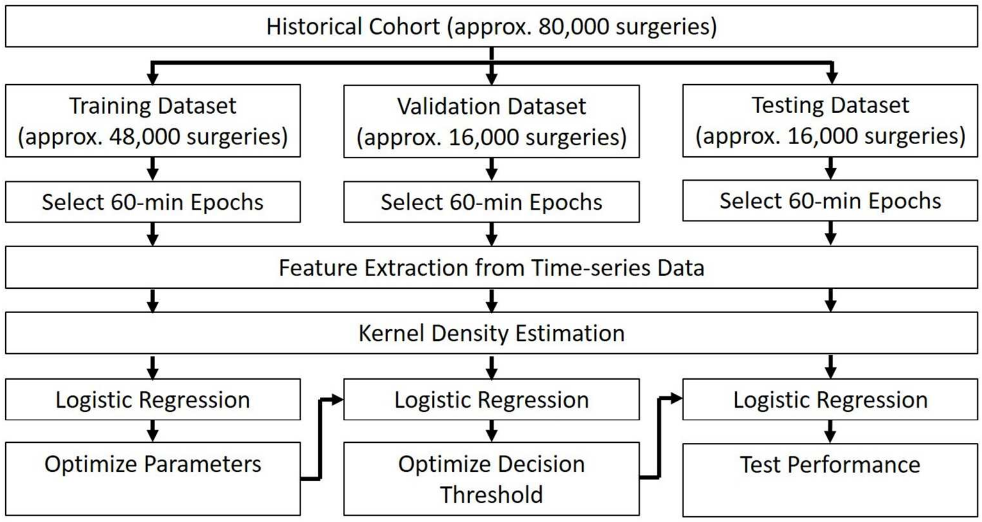

Figure 1. Data flow for algorithm training and validation using the historical database.

123x66mm (300 x 300 DPI)

Field	Output		Description
PatientID			Indentifier
Surg_Type			Text field
Anesthesia_Type		1	General
		2	Block
		3	MAC
		4	Spinal
		5	Epidural
		6	CSE
		8	Cancelled Procedure
		9	Cancelled Procedure after selected anesthesia type(s)
		10	converted to General
SEX		1	Male
		2	Female
		3	Unknown
RACE		5	Unknown
		6	Hispanic
		7	Black American
		8	Other
		9	White
		10	American Indian
		11	Asian
		12	American Indian or Alaska Native
		13	Black or African American
		14	Native Hawaiian or other Pacific Islander
		15	Some other Race
HEIGHT	Continuous		CM
WEIGHT	Continuous		KG
Ideal_Body_Weight	Continuous		Ideal weight at designated height, sex, etc.
BMI	Continuous		Body Mass Index
CCI	Integer sore 0-41		Charlson Comorbidity Index
FUNCTIONAL_CAPACITY		5	-
		6	>10 METs
		7	6-10 METS

	8	4-6 METS
	9	<4 METS
	10	Ambulates with assistance only
	11	Cannot assess
ASA	1	1
	2	2
	3	3
	4	4
	5	5
	6	1E
	7	2E
	8	3E
	9	4E
	10	5E
	11	6
HTN	1	Patient has hypertension
CAD	1	Patient has coronary artery disease
CAD_PRIORMI	1	Patient has previous Myocardial Infarction
CHF	1	Patient has congestive heart failure
CHF_Diastolic_Function	20	normal
	21	stage I - Impaired relaxation
	22	stage II - Pseudonormal
	23	stage III - Restrictive
	24	unknown/unspecified
	25	--
	26	unspecified dysfunction
LVEF	30	unknown
	31	>70%
	32	60-70%
	33	50-60%
	34	40-50%
	35	30-40%
	36	20-30%
	37	10-20%

	38	<10%
	39	unspecified normal
	40	unspecified mildly reduced
	41	unspecified moderately reduced
	42	unspecified severely reduced
	43	--
VALVULAR_DISEASE	61	Mild
	62	Mild-moderate
	63	Moderate
	64	Moderate-Severe
	65	Severe
	66	Unknown/unspecified
	67	--
AFIB	4	permanent (AF episode greater than 1 year)
	5	first/one detected episode (less than 7 day duration)
	6	paroxysmal (multiple episodes <7 days)
	7	persistent (one or more episodes >7 days)
	8	unknown
	9	--
PPM_ICD	1	Patient has pacemaker
CV_TIA_STROKE	1	Patient has had a stroke
PAD	1	Patient has peripheral artery disease
DVT	1	Patient has had deep vein thrombosis
PE	1	Patient has had a pulmonary embolism
DM	1	Patient has Diabetes
Outpatient_Insulin	103	none
	104	previous
	105	current
	106	insulin pump
	107	--
CKD	1	Patient has chronic kidney disease
Dialysis_History	90	ongoing peritoneal dialysis
	95	never
	96	ongoing hemodialysis

		97	past dialysis
		98	--
PHTN		1	Patient has pulmonary hypertension
COPD		1	Patient has chronic obstructive pulmonary disease
ASTHMA		1	Patient has asthma
OSA		1	Patient has sleep apnea
StopBang_Total	continuous		Higher score is worse
StopBang_Observed		1	Yes
		2	No
		3	Don't know
StopBang_Pressure		1	Yes
		2	No
		3	Don't know
StopBang_Snore		1	Yes
		2	No
		3	Don't know
StopBang_Tired		1	Yes
		2	No
		3	Don't know
CIRRHOSIS		1	Patient has cirrhosis
CANCER_HX		1	patient has history of cancer
GERD		1	Patient has Gastroesophageal Reflux Disease
ANEMIA		1	Patinet has history of anemia
COOMBS_POS		1	Patient has has a positive Coombs test
DEMENTIA		1	Patient has history of dementia
SMOKING_EVER		1	Patient reports having smoked
ULCER		1	Patient has history of Ulcer
CREATININE	continuous		Most recent serum creatinine value, manually entered during preoperative examination
PLATELET	continuous		Most recent platelet count, manually entered during preoperative examination
PreOp_Diastolic	continuous		Diastolic blood pressure during preoperative examination
PreOp_Systolic	Continuous		Systolic blood pressure during preoperative examination

Field	Output	Description
PatientID	Identifier	
ANESTHESIA START	Time	
TIME	Time	
DESIN	Continuous	Inhaled desflurane concentration (volume percent)
DESOUT	Continuous	Exhaled desflurane concentration (volume percent)
ISOIN	Continuous	Inhaled isoflurane concentration (volume percent)
ISOOUT	Continuous	Exhaled isoflurane concentration (volume percent)
N2OIN	Continuous	Inhaled nitrous oxide concentration (volume percent)
N2OOUT	Continuous	Exhaled nitrous oxide concentration (volume percent)
SEVOIN	Continuous	Inhaled sevoflurane concentration (volume percent)
SEVOOUT	Continuous	Exhaled sevoflurane concentration (volume percent)
TOTALMAC	Continuous	End-tidal anesthetic concentration (MAC [minimum alveolar concentration] units)
TOTALMACAGEADJ	Continuous	End-tidal anesthetic concentration (Age-adjusted MAC units)
BIS_INDEX	Continuous	Bispectral index
BIS_SR	Continuous	Suppression ratio (output from bispectral index monitor)
DIASTOLIC	Continuous	Diastolic blood pressure
SYSTOLIC	Continuous	Systolic blood pressure
BP_MEAN	Continuous	Mean arterial blood pressure
HR	Continuous	Heart rate
PULSE	Continuous	Pulse
TEMP	Continuous	Temperature
TEMP_CORE	Continuous	Temperature
SPO2	Continuous	Pulse Oximeter
CO2	Continuous	End-tidal carbon dioxide
RR	Continuous	Respiratory Rate
PEEP	Continuous	Positive End-Expiratory Pressure
PIP	Continuous	Positive Inspiratory Pressure
MV	Continuous	Minvute Ventilation
CVP	Continuous	Central Venous Pressure
URINE_OUTPUT	Continuous	Urine Output
Est_Blood_Loss	Continuous	Estimated Blood Loss
TIDAL_VOLUME	Continuous	Tidal Volume
GLU_ART	Continuous	Glucose value

GLU_VEN	Continuous	Glucose value
HCT	Continuous	Hematocrit
INR	Continuous	International normalized ratio
Plateau pressure	Continuous	Plateau Pressure (airway)
PLATELET	Continuous	Platelet Count
POTASSIUM	Continuous	Potassium level
SECSURPPRESSED	Continuous	Seconds of electroencephalogram suppression
DRUG_TYPE	DRUGS	Types of medication
	DRUGS IMPORT	
	DRUGS IMPORT Allergies	
	DRUGS REMOVED FROM SERVICE	
	DRUGS-ANTIBIOTICS	
	DRUGS-ANTIEMETICS	
	DRUGS-Beta blockers	
	DRUGS-CARDIACS	
	DRUGS-HYPNOTICS	
	DRUGS-LOCAL ANESTHESIA	
	DRUGS-NARCOTICS	
	DRUGS-OTHER	
	DRUGS-OTHER A-E	
	DRUGS-OTHER F-J	
	DRUGS-OTHER K-O	
	DRUGS-OTHER P-T	
	DRUGS-OTHER U-Z	
	DRUGS-OXYTOCICS	
	DRUGS-RELAXANTS	
	DRUGS-REVERSALS	
	DRUGS-TIMERS	
	FLUIDS	
DRUGS	See below	
DRUG_AMT	Continuous	
DRUG_MINUTES	Continuous	

Parameter	Drug		Drug Type
	68	Morphine	DRUGS-NARCOTICS
	69	FentaNYL	DRUGS-NARCOTICS
	74	Midazolam	DRUGS-HYPNOTICS
	76	Propofol	DRUGS-HYPNOTICS
	81	Succinylcholine	DRUGS-RELAXANTS
	83	Pancuronium	DRUGS-RELAXANTS
	84	Vecuronium	DRUGS-RELAXANTS
	85	Rocuronium	DRUGS-RELAXANTS
	86	Atropine	DRUGS-CARDIACS
	87	Neostigmine	DRUGS-REVERSALS
	88	Naloxone	DRUGS-REVERSALS
	89	Flumazenil	DRUGS-REVERSALS
	100	DOPamine	DRUGS-CARDIACS
	101	EPHEDrine	DRUGS-CARDIACS
	102	Esmolol	DRUGS-Beta blockers
	103	HydrALAZINE	DRUGS-CARDIACS
	105	Labetalol	DRUGS-Beta blockers
	108	Nitroglycerin	DRUGS-CARDIACS
	110	Norepinephrine	DRUGS-CARDIACS
	112	Phenylephrine	DRUGS-CARDIACS
	195	Vasopressin	DRUGS-CARDIACS
	200	Dexamethasone	DRUGS-OTHER
	203	Hydrocortisone	DRUGS-OTHER
	207	MethylPREDNISolone	DRUGS-OTHER
	218	Lactated Ringers	FLUIDS
	220	Normal Saline	FLUIDS
	226	Mannitol 20%	FLUIDS
	232	Packed RBC	BLOOD PRODUCTS
	234	Platelets	BLOOD PRODUCTS
	235	Fresh frozen plasma	BLOOD PRODUCTS
	236	Cryoprecipitate	BLOOD PRODUCTS
	237	Albumin 5%	BLOOD PRODUCTS
	238	Albumin 25%	BLOOD PRODUCTS

573	Other fluid	FLUIDS
766	Cisatracurium	DRUGS-RELAXANTS
780	Ondansetron	DRUGS-ANTIEMETICS
968	Glycopyrrolate	DRUGS-CARDIACS
1023	Chloroprocaine 3%	DRUGS-LOCAL ANESTHESIA
1121	Ketamine	DRUGS-HYPNOTICS
1198	Metoprolol	DRUGS-Beta blockers
1204	HYDROmorphone	DRUGS-NARCOTICS
1205	Meperidine hydrochloride	DRUGS-NARCOTICS
1444	1/2 Normal saline	FLUIDS
1446	D10-W	FLUIDS
1450	D5 Ringers Lactate	FLUIDS
1451	D5-1/2 Normal Saline	FLUIDS
1452	D5-1/4 Normal Saline	FLUIDS
1455	D5-Normal Saline	FLUIDS
1456	NaHCO3 150mEq in D5W	FLUIDS
1457	D5-W	FLUIDS
1458	D50	FLUIDS
1461	Hextend	FLUIDS
1670	Albuterol MDI	DRUGS-OTHER
1671	Albuterol Neb 0.5%	DRUGS-OTHER
1794	DiphenhydrAMINE	DRUGS-OTHER
1846	Glucagon	DRUGS-OTHER
1950	EPINEPHrine	DRUGS-CARDIACS
1951	Mannitol 25%	FLUIDS
1961	Dexmedetomidine	DRUGS-OTHER
1968	Hyperal	FLUIDS
2071	Packed RBC (ml)	BLOOD PRODUCTS
2073	Platelets (ml)	BLOOD PRODUCTS
2074	FFP (ml)	BLOOD PRODUCTS
2989	Atracurium	DRUGS-RELAXANTS
3057	Hespan	FLUIDS
3387	NaHCO3 150mEq in Water	FLUIDS
3460	Phenylephrine SA	DRUGS-LOCAL ANESTHESIA

3489	FentaNYL SA	DRUGS-LOCAL ANESTHESIA
9097	D5-1/2 Normal Saline 20K	FLUIDS
10291	OxyCODONE/Acetaminophen 5/325 (1	DRUGS-NARCOTICS
10297	OxyCODONE CR (for OxyCONTIN)	DRUGS-NARCOTICS
10298	EPINEPHrine inh 2.25%	DRUGS-CARDIACS
10585	Albuterol Neb 2.5mg	DRUGS-OTHER
10647	Racepinephrine Neb 2.25%	DRUGS-OTHER P-T
10648	D5-Normal Saline 20 KCL	FLUIDS
10652	Ketorolac Tromethamine (for Toradol	DRUGS-OTHER K-O
10665	OxyCODONE IR	DRUGS-NARCOTICS
10667	Morphine Sulfate 1mg/ml PCA	DRUGS-NARCOTICS
10672	HYDROmorphone 0.5mg/ml PCA	DRUGS-NARCOTICS
10689	Propofol (ml)	DRUGS-HYPNOTICS
10821	1/2 Normal Saline 20 KCL	FLUIDS
10910	Voluven	FLUIDS
11292	Normal Saline 20 KCL	FLUIDS
11951	Racepinephrine Neb 2.25% (ml)	DRUGS-OTHER P-T
14162	Ropivacaine 0.2% w/HYDROmorphon	DRUGS-LOCAL ANESTHESIA
14163	Ropivacaine 0.2% w/HYDROmorphon	DRUGS-LOCAL ANESTHESIA
14166	Ropivacaine 0.125% w/HYDROmorph	DRUGS-LOCAL ANESTHESIA
14325	Recomb Factor VII (ml)	BLOOD PRODUCTS
14796	Albuterol Neb	DRUGS-OTHER

SurveyHeaı	0	Reported No Heart Attack on Survey
	1	Reported Heart Attack on Survey
	3	Prefer not to answer
SurveyCarc	0	Reported No Cardiac Arrest on Survey
	1	Reported Cardiac Arrest on Survey
	3	Prefer not to answer
SurveyHeaı	0	Reported No Heart Failure on Survey
	1	Reported Heart Failure on Survey
	3	Prefer not to answer
SurveyStro	0	Reported No Stroke on Survey
	1	Reported Stroke on Survey
	3	Prefer not to answer
SurveyBloo	0	Reported No Leg Blood Clot on Survey
	1	Reported Leg Blood Clot on Survey
	3	Prefer not to answer
SurveyBloo	0	Reported No Lung Blood Clot on Survey
	1	Reported Lung Blood Clot on Survey
	3	Prefer not to answer
SurveyWoι	0	Reported No Wound Infection on Survey
	1	Reported Wound Infection on Survey
	3	Prefer not to answer
SurveyResp	0	Reported No Respiratory Failure on Survey
	1	Reported Respiratory Failure on Survey
	3	Prefer not to answer
SurveyPneι	0	Reported No Pneumonia on Survey
	1	Reported Pneumonia on Survey
	3	Prefer not to answer
SurveyNerv	0	Reported No Nerve Injury on Survey
	1	Reported Nerve Injury on Survey
	3	Prefer not to answer
SurveyGIBl	0	Reported No GI Bleed on Survey
	1	Reported GI Bleed on Survey
	3	Prefer not to answer
SurveyUlce	0	Reported No Ulcer on Survey
	1	Reported Ulcer on Survey
	3	Prefer not to answer
SurveyDeliı	0	Reported No Delerium on Survey
	1	Reported Delerium on Survey
	3	Prefer not to answer
30d_Survey	0	Reported No Readdmission within 30 days on Survey
	1	Reported Readdmission within 30 days on Survey
	3	Prefer not to answer
1y_Survey_	0	Reported No Readdmission within 1 year on Survey
	1	Reported Readdmission within 1 year on Survey
	3	Prefer not to answer

Info_Type	SubField	Output	Description
ADT	AdmitDtm	[DateTime]	Hospital Admit DateTime
	DischargeDtm	[DateTime]	Hospital Discharge DateTime
	ICU_Unit	10400 ICU	ICU Name - Neurosurgery ICU
		4400 ICU	ICU Name - Surgical/Burn/Trauma ICU
		5600 ICU	ICU Name - Cardiothoracic ICU
		8200 ICU	ICU Name - Cardiac ICU
		83 CTICU	ICU Name - Cardiothoracic ICU
	ICU_InDtm	[DateTime]	ICU Admit DateTime
	ICU_OutDtm	[DateTime]	ICU Discharge DateTime
Diagnoses	Value	[Text]	Diagnoses related to delirium/encephalopathy
Readmissions	DischargeDX	[Text]	Discharge Diagnosis
	Re_Adm_IDCode	[Continuous]	Medical record number for readmission
	Re_adm_VisitIDCode	[Continuous]	Patient account number (registration) for readmission
	Re_Adm_Location	[Text]	Bed assignment for readmission
	Re_Adm_AdmitDtm	[DateTime]	DateTime of Readmission
	Re_Adm_AdmitDx	[Text]	Diagnosis for readmission
	DurationBetweenVisits_Days	[Continuous]	Duration (days) between initial discharge and readmission
Labs	SignificantDtm	[DateTime]	DateTime of Lab Test
	Label	A-a Gradient	Alveolar-Arterial Oxygen Gradient (mmHg)
		Albumin Level	Plasma albumin level (g/dL)
		Alkaline Phosphatase Total	Plasma alkaline phosphatase level (units/L)
		ALT	Plasma alanine transaminase level (units/L)
		Anion Gap	Plasma anion gap (mmol/L)
		AST	Plasma aspartate transaminase level (units/L)

Basophil Absolute Automated	Concentration of basophils in blood (K/mm^3)
Basophil Percent Automated	Percent of basophils in blood (%)
Bilirubin Direct	Plasma direct bilirubin level (mg/dL)
Bilirubin Total	Plasma total bilirubin level (mg/dL)
Calcium Total	Plasma calcium level (mg/dL)
Chloride Level	Plasma chloride level (mmol/L)
CO2 Total	Plasma bicarbonate level, measured (mmol/L)
CO2, Total Calculated, Arterial	Plasma bicarbonate level, calculated from arterial blood gas (mmol/L)
Creatinine Level	Plasma creatinine level (mg/dL)
Eosinophil Absolute Automated	Concentration of eosinophils in blood (K/mm^3)
Eosinophil Percent Automated	Percent of eosinophils in blood (%)
Glucose Level Fasting	Plasma glucose level, fasting (mg/dL)
Glucose Level Random	Plasma glucose level (mg/dL)
Glucose Ur Qual	Urine glucose, qualitative (negative, 1+, 2+, 3+)
Glucose WB f POC	Plasma glucose level, point of care (mg/dL)
HCT	Hematocrit (%)
Hemoglobin A1c	Hemoglobin A1c (%)
HGB	Hemoglobin level (g/dL)
INR ePOC	International normalized ratio (no units)
Lymphocyte Absolute Automated	Concentration of lymphocytes in blood (K/mm^3)
Lymphocyte Percent Automated	Percent of lymphocytes in blood (%)
MCH	Mean corpuscular hemoglobin (pg)
MCHC	Mean corpuscular hemoglobin concentration (g/dL)
MCV	Mean corpuscular volume (fL)

Monocyte Absolute Automated	Concentration of monocytes in blood (K/mm^3)
Monocyte Percent Automated	Percent of monocytes in blood (%)
MPV	Mean platelet volume (fL)
Neutrophil Absolute Automated	Concentration of neutrophils in blood (K/mm^3)
Neutrophil Percent Automated	Percent of neutrophils in blood (%)
NRBC Abs Auto	Concentration of nucleated red blood cells in blood (K/mm^3)
NRBC Pct Auto	Percent of nucleated red blood cells in blood (%)
pCO2 Art gPOC	Arterial Partial Pressure of Carbon Dioxide (mmHg)
pCO2 Ven gPOC	Venous Partial Pressure of Carbon Dioxide (mmHg)
pCO2, Arterial	Arterial Partial Pressure of Carbon Dioxide (mmHg)
Percent Inspired Oxygen, Arterial	Percent Inspired Oxygen (%)
pH Art gPOC	Arterial pH (no units)
pH Arterial	Arterial pH (no units)
platelet cPOC	Concentration of platelets in blood (K/mm3)
Plt	Concentration of platelets in blood (K/mm^3)
pO2 Art gPOC	Arterial Partial Pressure of Oxygen (mmHg)
pO2 Ven gPOC	Venous Partial Pressure of Oxygen (mmHg)
pO2, Arterial	Arterial Partial Pressure of Oxygen (mmHg)
Potassium Level Plasma	Plasma potassium level (mmol/L)
Protein Level Plasma	Plasma protein level (g/dL)
PT (Seconds)	Prothrombin time (sec)
PT ePOC	Prothrombin time (sec)

		PTT Activated	Partial thomboplastin time, activated (sec)
		PTT Activated.	Partial thomboplastin time, activated (sec)
		RBC	Concentration of red blood cells in blood (M/mm^3)
		RDW CV	Red cell distribution width (%)
		RDW SD	Red cell distribution width (fL)
		Sodium gPOC	Plasma sodium level (mmol/L)
		Sodium iPOC	Plasma sodium level (mmol/L)
		Sodium Level	Plasma sodium level (mmol/L)
		Urea Nitrogen	Plasma urea nitrogen (BUN) level (mg/dL)
		Volume Inspired Oxygen Arterial	--
		WBC	Concentration of white blood cells in blood (K/mm^3)
	Value	[Continuous]	Value of lab test
Medications	SignificantDtm	[DateTime]	DateTime medication order is placed in electronic medical record
	Label	ALPRAZolam Tablet	
		Ampicillin/Sulbactam - CRITICAL SHORTAGE	
		CeFAZolin IVPB	
		Cefepime IVPB - CRITICAL SHORTAGE	
		CefOXitin IVPB	
		CefTRIAXone IVPB	
		Clotrimazole Troche	
		Darunavir	
		Dexmedetomidine Infusion	
		DiazePAM Injection	
		DiphenhydrAMINE Oral	
		Eszopiclone	
		FentaNYL Bolus from CADD	
		FentaNYL Infusion	

FentaNYL Injection
Fluconazole Tablet
Haloperidol (immediate-acting) Lactate Injection
HYDROmorphone Bolus from CADD
HYDROmorphone Injection
HYDROmorphone PCA
Insulin Glargine for LANTUS
Insulin Lispro for HumaLOG

Insulin Lispro-Sliding Scale for HumaLOG
Insulin NPH for HumuLIN-N
Insulin Regular for HumuLIN-R
Insulin Regular Infusion
LORazepam Injection
LORazepam Tablet
Maraviroc
Metoclopramide Injection
Micafungin IVPB
Midazolam Infusion
Midazolam Injection
Morphine Injection
Nystatin Liquid
Ondansetron Injection
Oxacillin IVPB
OxyCODONE Liquid

OxyCODONE Tablet Immediate-Release

OxyCODONE/Acetaminophen 5/325 mg
Piperacillin/Tazobactam
QUEtiapine Tablet
Raltegravir
Ramelteon

		Ritonavir Tablet	
		Trimethobenzamide Injection	
	DosageLow	[Continuous]	Medication dose (low end of range, if applicable)
	DosageHigh	[Continuous]	Medication dose (high end of range)
	UOM	[Categorical]	Unit of measure for dosage
	FrequencyCode	[Categorical]	Frequency of medication order
	Duration_Days	[Continuous]	Number of days during which medication order was active
Clinical Documentation - 01. Vital Signs	SignificantDtm	[DateTime]	DateTime of vital sign record
	Intensity	[Continuous]	Numeric rating scale (0-10) pain score
	Behaviors	Calm	
		Agitated	
	Tool Used	0-10 Scale	
		Faces	
		Non Communicative	
	Patient Goal	[Continuous]	Patient's target pain score
	Location	[Categorical]	Location of pain
	Pain Management	[Categorical]	Nurse response to current pain score
	Effectiveness of Pain Interventions	Obtaining relief	
		Partial relief	
		No relief	
Clinical Documentation - 03. Assessment/IPOC	SignificantDtm	[DateTime]	DateTime of CAM-ICU Assessment
	CAM-ICU Overall Score	[Negative/Positive]	CAM-ICU Result (delirium assessment)
	Feature 1: Acute Onset or Fluctuating Course	[Negative/Positive]	Feature 1 of CAM-ICU
	Feature 2: Inattention	[Negative/Positive]	Feature 2 of CAM-ICU
	Feature 3: Altered Level of Consciousness	[Negative/Positive]	Feature 3 of CAM-ICU

	Feature 4: Disorganized Thinking	[Negative/Positive]	Feature 4 of CAM-ICU
	RASS: If -4 or -5, STOP Reassess later	[Categorical]	Indicates whether CAM-ICU is contraindicated (e.g., because patient is too sedated)
Clinical Documentation - Fall Event Note	SignificantDtm	[DateTime]	
	Fall Date/Time	[DateTime]	DateTime of Fall
	Injury	[Yes/No]	Indicates whether fall resulted in patient injury
	Injury Details	[Text]	Description of injury
Clinical Documentation - Neuro Flowsheet	SignificantDtm	[DateTime]	DateTime of RASS Assessment
	Behaviors	Calm Agitated Restless	
	RASS Score Numeric	[Continuous]	Richmond Agitation Sedation Scale score (-4 to 4)
	RASS Sedation Scale	[Categorical]	Verbal description of RASS numeric score
Clinical Documentation - Intubation Procedure Note	SignificantDtm	[DateTime]	DateTime of Intubation Note
	Intubation Type	[Emergent/Elective]	
	Paralytics Given	[Yes/No]	
	Procedure Date/Time	[DateTime]	DateTime of Intubation
	Sedation Given	[Yes/No]	
Clinical Documentation - Patient Profile	SignificantDtm	[DateTime]	DateTime of assessment
	Does patient currently have a tracheostomy tube?	[Yes/No]	

	Does patient have a history of laryngectomy?	[Yes/No]	
Clinical Documentation - Ventilator Flowsheet	SignificantDtm	[DateTime]	DateTime of ventilator event
	Modes/Bagged	A/C - VC	Assist control - volume control (AC-VC)
		A/C - VC+	Assist control - volume control + (AC-VC+)
		A/C - PC	Assist control - pressure control (AC-PC)
		BiPAP	Bilevel positive airway pressure (BiPAP)
		Continuous Positive Airway Pressure	Continuous positive airway pressure (CPAP)
		SIMV - VC	Synchronized intermittent mandatory ventilation (SIMV)
		SIMV-PC	Synchronized intermittent mandatory ventilation (SIMV)
		SIMV-VC	Synchronized intermittent mandatory ventilation (SIMV)
		SPONT-VC	Pressure support (PSV)
		Synchronized intermittent mandatory ventilation	Synchronized intermittent mandatory ventilation (SIMV)
	Ventilator Pulmonary Event	Extubation	
		Re-intubation	
		Ventilator start	Start of mechanical ventilation
		Ventilator stop	End of mechanical ventilation
		Weaning end	
		Weaning start	