## A. SPECIFIC AIMS

In the United States, where the average American can expect to undergo seven surgical operations during a lifetime, each year at least 150,000 patients die and 1.5 million develop a medical complication within 30 days after surgery[1,2]. Reducing postoperative complications (PC) by 20% could potentially save thousands of lives and significantly reduce healthcare costs[3]. The risk for PC arises from the interactions between a patient's preoperative health and physiologic capacity to withstand surgery-related stress, modulated by the type and quality of surgery and anesthesia that the patient undergoes[4]. Cost-effective strategies implemented in a timely fashion can ameliorate risk for PC or prevent their progression to more severe stages. The ability to implement such preventive strategies depends on the timely and accurate identification of patients at the greatest risk of PC[5]. Assessment of surgical risk requires timely, accurate and dynamic synthesis of the large amount of clinical information throughout the continuum of perioperative care. Today we do not have the ability to accurately predict and quantify, for a given patient, a *personal and real-time* risk for PC that dynamically integrates preoperative risk with the risk incurred by events during surgery. Current surgical risk scores are limited to either a preoperative physician's subjective risk assessment or calculated scores with modest accuracy and limited usability. The interventions that could help prevent PC are applied without consideration of a patient's personal risk profile or not applied at all because the risk is underestimated. Paradoxically the abundant real-time physiologic, laboratory, and other clinical data streams are available in the perioperative electronic health records (EHR) but their magnitude and complexity often overwhelms physicians' ability to comprehend, retain, and organize the information in an optimal and timely way.

We propose that clinical data contains unexplored richness sufficient for the development of *automated and personalized* diagnostic tools that currently do not exist[6,7]. The *objective* of this proposal is to implement technological advances in data science and engineering in innovative steps to develop an intelligent perioperative system (IPS), composed of computers and physicians interacting in real time, which can generate usable medical knowledge with both increased speed and accuracy using complex clinical data obtained in the perioperative period. Our multidisciplinary team of experts in medicine and engineering will address methodological challenges including real-time data integration and processing, data analytics and knowledge exchange between computers and physicians. We will refine and validate *computational algorithms* that reconcile information from streams of clinical data, including multi-dimensional physiologic time series to *accurately predict and quantify* the risk of major complications *at different time points*-prior to surgery and immediately after surgery. The system will utilize a *high-performance computing environment* to enable real-time data-intensive predictive modeling of multi-dimensional clinical data streams. In a prospective clinical study we will validate the accuracy of the system, user acceptance and satisfaction and the impact on clinical decision making. This methodology will provide a significant step towards personalized perioperative medicine utilizing clinical data acquired during routine medical care. We propose three specific aims:

**Specific Aim 1. Refine and validate predictive risk models for major complications using EHR integrated with intraoperative physiologic time series.** Using a retrospective temporal dataset for 10,000 patients we will test the hypothesis that computational algorithms using EHR data preoperatively and EHR enriched with intraoperative physiologic time series data postoperatively are more accurate in dynamically predicting the risk for major complications compared to existing risk scores.

**Specific Aim 2**. **Implement and validate two-way knowledge exchange between predictive risk models and physicians.** Using an interactive knowledge exchange application we will test the hypothesis that active learning that takes higher-level input from physicians improves the performance of predictive models while abductive reasoning that provides a most-likely explanation of the model prediction facilitates their acceptance by physicians.

**Specific Aim 3. Implement and evaluate an intelligent perioperative system for automated risk analysis using real-time EHR data**. We will implement a prototype of an intelligent perioperative system using real-time clinical data streams with data analytics and knowledge exchange layers to quantify the risk for major complications with minimal time delay. In a prospective clinical study of 60 physicians we will validate the diagnostic performance of our predictive risk models, compare them with the physicians' risk assessment and measure change in physicians' risk perception after knowledge exchange with the system.

The *expected outcome* of this proposal is a new framework allowing development, clinical implementation and validation of an **open-source, customizable at scale, plug and play intelligent system that within a typical clinical workflow can integrate and analyze data, extract and exchange information with physicians, and ultimately *both automate and personalize* clinical decision-making**.

## B. SIGNIFICANCE

**B.1. Major postoperative complications impact the health of individuals and populations**. In the United States each year 150,000 patients die and 1.5 million develop a medical complication after surgery[2]. Since postoperative complications (PC) cause a two-fold increase in the 30-day mortality and incur an annual cost of $25 billion[1,2,8], even a small reduction in their occurrence could save lives and reduce healthcare costs[3]. We have demonstrated that **postoperative acute kidney injury (AKI**) affected up to 30% of patients, was associated with increased risk for **sepsis**, severe **respiratory failure (RF), cardiovascular complications, prolonged intensive care unit (ICU) stay** and increased cost of care, and even after resolution was associated with a risk for chronic kidney disease and death years after surgery[9-20]. Together with others we have found that sepsis, prolonged RF and ICU stay are associated with chronic critical illness and increased long-term mortality and disability[21-28]. Thus improvement in risk stratification and prevention of major PC among a diverse group of patients may tremendously improve their future health and quality of life.

**B.2. With timely interventions major complications are preventable.** Although no single therapy can prevent sepsis or AKI, many cost-effective strategies implemented in a timely fashion can ameliorate risk or prevent further progression to more severe stages. Systematic review of randomized controlled trials and meta-analyses reveals several interventions that can reduce PC and mortality in adult surgery when indicated, including hemodynamic optimization[29-33], use of neuraxial anesthesia and volatile agents[34-38], insulin for glycemic control[39-42], non-invasive ventilation[43-45], leukodepleted red blood cell transfusion[46,47], remote ischemic preconditioning[48,49] and preoperative use of statins[50]. Based upon contemporary consensus guidelines for the management of high-risk patients for AKI[51], we have developed standardized clinical protocols combining clinical risk stratification and biomarkers to enhance AKI risk stratification for patients undergoing surgery[52,53]. We have developed computerized sepsis surveillance tools and decision–support systems for early management of patients with sepsis in the University of Florida Health (UFH) surgical ICU[54-57]. The ability to implement such preventive strategies depends on the timely and accurate identification of surgical patients at the greatest risk of PC[5].

**B.3. Risk for complications changes throughout the perioperative period.** The risk for PC arises from the interactions between a patient's preoperative health and physiologic capacity to withstand surgery-related stress[4]. **In the preoperative period,** knowing the extent to which preoperative health predisposes a patient for PC can facilitate a discussion about the risks and benefits of surgery to decrease the uncertainty regarding outcomes. An accurate risk assessment (RA) allows physicians to identify patients who would benefit the most from strategies that can offset the risk. Some of these strategies, like invasive monitoring, are not only costly but carry their own risks. Others, like the avoidance of nephrotoxic medications in patients at risk for AKI, are easy to implement if the risk is identified. And yet both are often applied without consideration of a patient's preoperative risk or are not applied at all because risk is underestimated. **In the postoperative period,** the preoperative risk is modified by the response to the events experienced during surgery. The intraoperative hemodynamic and anesthetic management can either exacerbate or ameliorate surgery-related stress. Without dynamic adjustment of the risk at the end of surgery, the opportunity to initiate preventive therapies for high-risk patients, such as triage to ICU for early hemodynamic optimization, sepsis protocol, non-opioid analgesia and early mobilization, may be missed. Assessment of surgical risk requires timely, accurate and dynamic synthesis of the large amount of clinical information throughout the continuum of perioperative care.

**B.4. Current perioperative risk scores underutilize information in electronic health records.** Current perioperative risk stratification is limited to a physician's subjective RA or risk scores (RS) of modest accuracy and often requiring elaborate data extraction. The American Society of Anesthesiologists (ASA) physical status classification relies on physicians' subjective assessment of a patient's preoperative health and is commonly used by anesthesiologists to predict surgical mortality risks[58]. The Physiologic and Operative Severity Score for the Enumeration of Mortality and Morbidity (POSSUM) hasn't been adopted widely as it requires elaborate manual data collection[59]. A surgical APGAR score uses a crude summary of risk[60] and its adoption in practice has been met with skepticism[5]. The majority of AKI risk scores are limited to cardiac surgery and have modest accuracy[61,62]. No validated risk scores exist for sepsis or ICU admission. Recent risk models for RF have improved accuracy but like other models either underutilize or implement "dimensional reductionism" for complex intraoperative physiologic data[63-67]. To date, none of the existing surgical risk scores have explored the potential for automation and personalization with EHR[58-60,68].

**B.5. Intraoperative physiologic time series data is underutilized for risk assessment.** Prior to surgery physicians process large volumes of clinical data and use their clinical judgment to synthesize the information,

quantify risk and formulate treatment plans[69,70]. During surgery, intraoperative monitors capture complex physiologic time series data reflecting ongoing response to anesthesia and surgery. Several channels of cardio-respiratory waveforms are displayed continuously, derived indices are reported minute-by-minute; medications, fluid balance, and ventilator settings are monitored regularly and critical laboratory results need to be reviewed as they become available. Given time constraints and increasing information load on physicians[71,72], most of the gathered data can only be screened for overt abnormalities and, in their absence, subsequently discarded. Potentially valuable information that could be used to understand the patient's physiologic response to surgery-related stress is easily missed and underutilized for risk stratification, as it often requires computational capacity beyond human ability.

**B.6. Opportunity to move towards precision perioperative medicine using clinical data.** We envision an intelligent system that combines machine and human intelligence to provide automated, dynamic and personalized risk assessment for major complications before they become manifest. This approach would allow management to be tailored to a patients' "personal clinical profile"[6] using clinical data gathered during routine care. On an institutional level, accurate prediction of perioperative risk may help to quantify the complexity of work being undertaken, improve patient triage and treatment planning and provide a method for documenting risk-adjusted outcomes.

**C. INNOVATION**

We propose that clinical data contains unexplored richness sufficient for the development of *automated and personalized* diagnostic tools that currently do not exist[6,7]. This proposal implements technological advances in data science and engineering in innovative steps to pilot and validate a plug and play, scalable open-source intelligent system that in real-time will integrate and analyze data, extract and exchange information with physicians, and *both automate and personalize* clinical decision-making.

**C.1. The architecture for real-time data integration and analytics**. We will design and pilot a real-time, open-source, plug and play software system for perioperative risk calculation. We will translate our experience with data engineering and middleware to adapt the system to other clinical environments with an EHR by developing a set of software solutions that enables fast prototyping and automates the activities and transactions in any given business ecosystem[73].

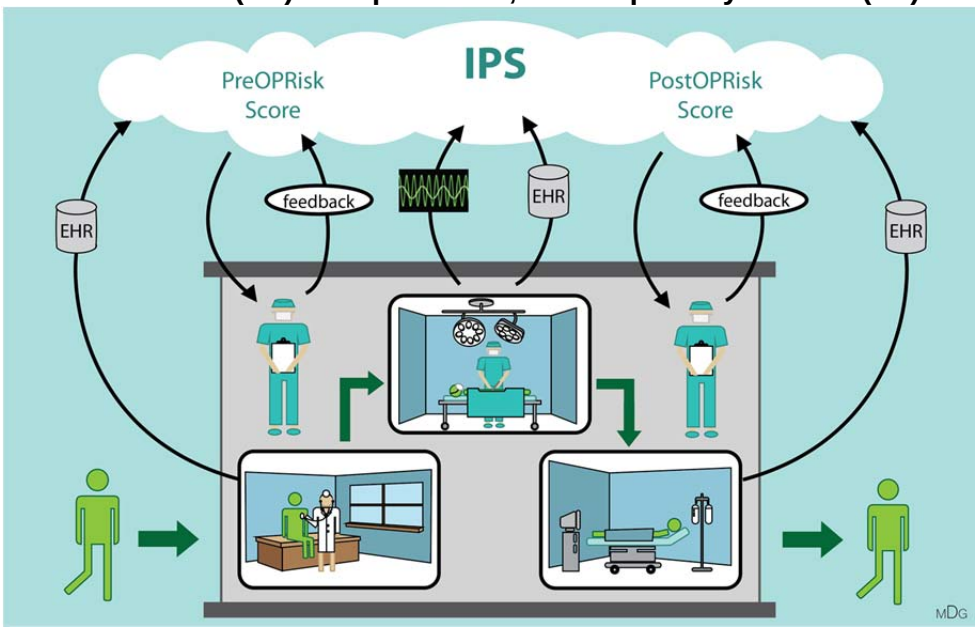
**C.2. Use of intraoperative physiological time series for predictive algorithms.** We will address the issue of processing and integrating multidimensional clinical data from EHR with high-volume intraoperative physiologic time series in a way that enables real-time data-intensive predictive modeling. Although recent studies in the ICU[74,75] and our preliminary work have shown that physiologic data trajectories of patients with PC might be used for risk prediction, the wealth of dynamic information in this type of data is currently not used for real-time risk prediction. Our time series processing techniques will enable us to fully utilize the information captured by physiologic trend data in a dynamic manner.

**C.3. Knowledge exchange between computer and human intelligence.** While most active learning literature is based on the use of low-level annotated data points to improve predictive models, we will implement recent advances in active learning that takes higher-level knowledge input from physicians to improve the performance of predictive models. Use of abductive reasoning and lineage tracking and representation to explain models' output and facilitate their acceptance by physicians is a novel application that expands on our research in information retrieval and databases.

**C.4. Multidisciplinary team and technology to deliver predictive analytics within the clinical workflow.** The team of experts in medicine, engineering and informatics will build upon UF resources and research in data science, engineering and surgical outcomes. Recently funded by NIGMS, UF P50 Sepsis and Critical Illness Research Center (SCIRC) (AB and TOB) will longitudinally follow 400 sepsis patients to identify mechanism of chronic critical illness. We will link perioperative data to SCIRC follow-up data for a subset of patients with sepsis and AKI. UF investments in high-performance computing (HPC) (XL) give us a unique opportunity to use these resources with reduced cost. UF Informatics Institute (AB, PP, PR, DW) provides a collaborative network of experts in data science. UF has built the infrastructure for the OneFlorida Clinical Research Consortium (9 clinical systems providing care for ~9.7M Floridians) to conduct large-scale clinical trials. As a result our Integrated Data Repository (IDR) (GL and WH) has developed routines to extract, transform, and load EHR data from their native format and coding systems to standard formats (including standard terminologies required by Meaningful Use) providing data harmonization necessary for our system.

## D. APPROACH

**D.0. Overview of intelligent perioperative system.** An "intelligent perioperative system (IPS)" is built on the conceptual framework that integration of data, computational algorithms and clinical judgment allow "precision, automation and personalization" in perioperative risk stratification using real-time analytics of clinical data obtained during routine care (Fig 1). The IPS is an open-source, scalable, plug and play software system deployed on high-performance computers for perioperative risk calculation. It allows the integration and data analytics of EHR and physiological data with minimal time delay. *Computational algorithms predict and quantify the risk for PC using all available data during preoperative evaluation and at the end of surgery when decisions about postoperative care are made based on patient's anticipated hospital course.* The IPS augments clinical decision-making through the process of two-way knowledge exchange with physicians by allowing them to provide feedback to the system that is subsequently used to re-train and improve the predictive models. **The predictive primary outcomes are major PC: 1. Sepsis (SS), 2. Acute kidney injury (AKI), 3. Cardiovascular (CV) complications, 4. Respiratory failure (RF) and 5. ICU admission > 48 hours (**occurring in the first 30 postoperative days and defined using the American College of Surgeons–National Surgical Quality Improvement Program's definitions[76] except for AKI that is defined using consensus KDIGO criteria[77]). Our previous work has demonstrated the high prevalence of major PC (between 6% and 39%) associated with increases in short and long-term mortality and cost[9-20]. We will also use secondary outcomes a*. Hospital death*; b. *Discharge other than home*; c. *Readmission within 30-days of discharge; d. One-year survival after surgery (all available in IDR).*




**D.1. Specific Aim 1. Refine and validate predictive risk models for major complications using EHR integrated with intraoperative physiologic time series.**

The *objective* of this aim is to refine the first-generation of computational algorithms developed by us to predict the risk for postoperative sepsis, acute kidney injury, cardiovascular complications, prolonged respiratory failure and ICU admission. For each complication we calculate the risk at two time points: preoperatively-days to hours before the surgery and postoperatively-at the end of surgery, using her data enriched with intraoperative physiologic time series and trend data. Using a retrospective temporal dataset for 10,000 patients we will refine and validate algorithms and test the *hypothesis* that computational algorithms are more accurate in predicting the risk for major complications compared to existing risk scores. The *rationale* is that our models will act as a real-time indicator of patient's response to surgical stress and provide dynamic information about changes in her health state after surgery. Upon completing this aim we *expect* to have validated two *probabilistic risk scores (PreOpRisk and PostOpRisk*) for each complication.

**D.1.1. Justification and feasibility.** The majority of current perioperative risk scores, including ASA, POSSUM, NSQIP, surgical APGAR, RF and AKI scores, rely on physicians' subjective RA or implement "dimensional reductionism" when dealing with physiologic time series and trend data[58-68]. Many of them suffer the limitation of a simplistic time-invariant design using manual data collection. The recent use of open-source ICU databases to develop a clinical alert system for hypotension using machine learning algorithms applied on high-resolution vital signs data[74,75] emphasized the importance of medical datasets for advancement of computational methods[78,79]. The largest existing intraoperative Multicenter Perioperative Outcomes Group' database has membership-only access, lacks high-dimensional physiologic data and includes limited postoperative outcomes compared to our plan of study[80]. Using the experience of UFH IDR[81,82] we have developed a perioperative dataset DECLARE linked with long-term survival demonstrating the relationship between the adverse effect of AKI and other major PC[11].

### D.1.2. Preliminary studies.

**D.1.2.1. Mortality and cost associated with PC.** Our group has published important papers on the epidemiology and outcomes of postoperative AKI and sepsis[9-20])[83,84]. Using the perioperative Declare dataset (Vision Grant for PI AB) we studied the outcomes of PC among ~ 70,000 surgical patients at UFH between 2000 and 2010 demonstrating the high prevalence and increase in mortality and cost associated with major PC: AKI, SS, RF, ICU and CV complications (Fig 2).

| PC | Prevalence | Hospital mortality | One year mortality |
|---|---|---|---|
| SS | 6% | 29% | 43% |
| CV | 8% | 20% | 32% |
| RF | 14% | 22% | 36% |
| ICU | 32% | 10% | 32% |
| AKI | 39% | 8% | 21% |

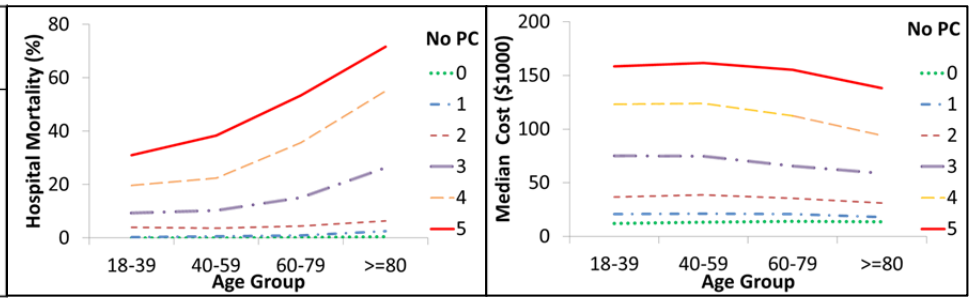



Figure 2. Prevalence, mortality and cost associated with major complications.

**D.1.2.2. Validation of preoperative probabilistic risk models.** Among 51,457 surgical patients we have validated algorithms quantifying the risk for postoperative AKI, SS, RF, CV and ICU stay using all available preoperative variables in EHR. We optimized use of nominal variables with multiple levels (eg. procedure and ZIP codes) using a classification tree while for numerical variables we fitted a nonlinear function for each PC. We compared several machine learning and dimensionality reduction approaches to obtain strong models with area under curve (AUC) > 0.8 (Fig 3). The best accuracy was obtained using generalized additive models (GAM) and principal component analysis (PCA). For an individual patient the algorithm produces a probabilistic risk score for each PC after surgery. Optimizing two cut-offs for each score allowed us to stratify patients into low, medium and high risk categories.

| | AKI AUC (95% CI) | |
|---|---|---|
| Model Types | Full Model | Reduced Model |
| Logistic Regression | 0.82 (0.82, 0.83) | 0.82 (0.82, 0.83) |
| GAM | 0.83 (0.82, 0.83) | 0.83 (0.82, 0.83) |
| Naïve Bayes | 0.80 (0.79, 0.80) | 0.80 (0.79, 0.80) |
| Support Vector Machine | 0.82 (0.81, 0.83) | 0.86 (0.85, 0.86) |

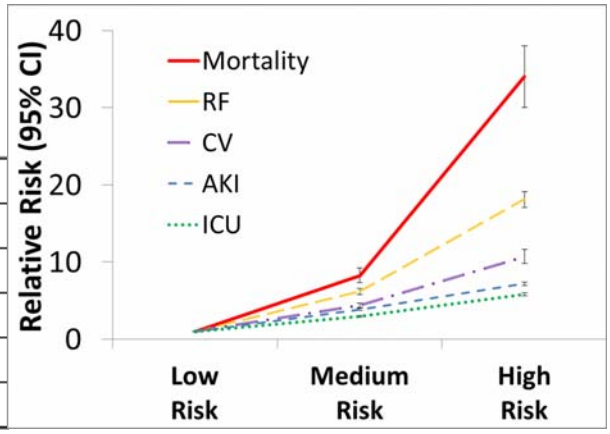



Figure 3. Performance of preoperative risk models.

**D.1.2.3. Intraoperative physiologic time-series enrich risk stratification for PC.** In the subset of 7215 patients, to enrich AKI prediction with intraoperative events we used 3 physiologic time series with ~ 600 observations per patient (mean arterial blood pressure (MAP), minimum alveolar concentration and heart rate.) We also included labs, medications and estimated blood loss. We extracted time series features (e.g. mean, long and short term variability) and added them to the PreOP model. Fig 4A shows distribution for AKI risk score based on time spent at different intraoperative MAP values. The combined model AUC was 0.86 (95% CI 0.84-

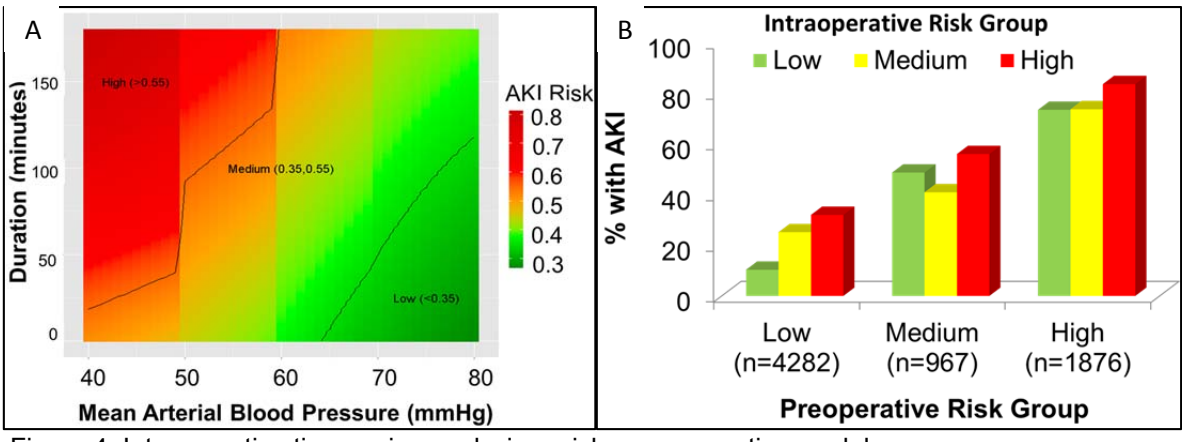



Figure 4. Intraoperative time series analysis enriches preoperative models.

0.87). After addition of intraoperative data we re-classified patients with low preoperative AKI risk to medium (n=556, 13%) and high PostOP risk (n=200, 5%) groups (Fig 4B).

**D.1.3. Research design.** Our objective is to a) Refine computational risk algorithms described in D1.2.2-3, b) Perform their internal validation in a retrospective cohort against the outcomes, and c) compare them against other surgical risk scores. *The computational risk algorithms calculate dynamic probabilistic risk scores for each PC at two time points: before surgery using preoperative EHR (PreOPRisk) and after surgery using integrated EHR and intraoperative physiologic time series to reflect pathophysiologic response to surgical stress (PostOPRisk).*

**D.1.3.1. Study design and participants.** Retrospective cohort study of 10,000 patients ≥18 years who underwent any type of inpatient surgery at UFH between April 2013 and July 2015. We will use a temporal deidentified data repository of EHR integrated with medium resolution intraoperative physiologic time series with annotated primary and secondary outcomes. Using 70/30 cross-validation we will split the retrospective cohort into a training cohort for model learning and a validation cohort for testing the diagnostic accuracy. We will improve the accuracy of the models that will be included in the real-time system by both feature enrichment and use of new analytical approaches for time series analyses.

**D.1.3.2. Sample size and feasibility.** A sample size of 10,000 patients will offer 80% power to detect a difference as low as 0.02 to 0.05 between a new models' sensitivity (Se) and the null Se (Se0) of 0.80 presuming a complication prevalence between 39% and 6%[85]. This sample size allows estimation of a 95% confidence interval (CI) width for specificity (Sp) of 0.02 and 95% CI width for Se between 0.02 and 0.1 (for all possible values of Sp, Se and PC prevalence). It will result in the maximum 95% CI width for AUC of 0.06 when complication prevalence is 5%, 0.04 when prevalence is 10% and 0.02 when prevalence is >20%, higher AUC values giving narrower CI[86]. UFH performs annually > 10,000 surgical cases so we anticipate a sufficient number of records.

**D.1.3.3. Data integration and harmonization.** Our objective is to develop a temporal deidentified dataset with complete perioperative records for patients fulfilling inclusion criteria. The IDR team (GL) currently maintains 5 years of integrated Epic® EHR data and claims data from UFH and has substantial experience in harmonizing data with key standards, such as i2b2 common data model, the RxNorm medication terminology and the Logical Observation identifiers Names and Codes (LOINC) standards. Co-I Dr. Hogan is an international expert in data standards, having developed, researched, and implemented numerous standards[87-89]. The policies and procedures of IDR are designed to secure privacy and confidentiality of the individuals represented in the data as described in detail in App_1A-C. We will develop extract, transform, and load (ETL) routines to take data from their native format and coding systems in EHR and convert them to the standard format and coding systems (including standard terminologies required by Meaningful Use) to create the *de-identified data repository (DDR) on our high-performance computing platform*. New database schemas will be optimized for both data loading and retrieval performances. Data integrity, regression, and performance testing will be carried out against data and processes that already exist in the IDR. Structured data deidentification will be automated using multiple step pseudonyms under the UF Honest Broker Protocol to remove all 18 HIPAA identifiers and loaded into the DDR after gaining IRB approval.

**D.1.3.4. Data elements and feature enrichment.** Each patient's complete perioperative record contains an array of heterogeneous variables extracted from the EHR and besides features in current models (demographics, encounters, diagnoses, procedures, medications, laboratory results) will include themicrobiology results, weight/height and vital signs, fluids intake and output, organ support, functional status, family history, plan(s) of care, and tobacco history. We will apply a set of diagnostic codes to identify patients at high risk for frailty and introduce it as a new predictive feature (App_1D). The multiple intraoperative physiologic time series (eg. blood pressure, HR, MAC, oxygen saturation, bispectral index, etc.) are drawn from intraoperative monitors, where data are down-sampled from 1Hz capture of monitor data to automatic readings every one-minute and exported to Epic in HL7 format. We will use all available time series to expand on work already done using IDR data for the DECLARE dataset (D.1.2.1.)..

**D.1.3.5. Optimizing intraoperative physiologic time series.** We will expand normalization steps used in the preliminary analyses to correct time-series artifacts including expert feedback to eliminate extremes and imputation based on local interpolation[90,91,92]. Since the time series may contain redundant information with significant dependency between time series values, we will develop a more compact way to represent time series by fitting data to mathematical models and exploiting properties (e.g. entropy) of those models. We will evaluate a range of different time series compression approaches such as piecewise aggregate approximation

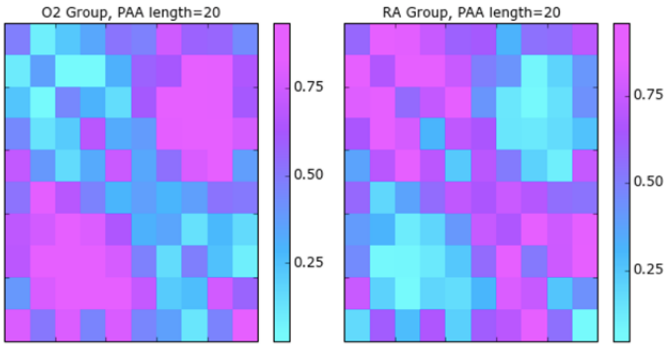


Figure 5. PAA compression of pulse-oximetry time series for patients on oxygen and room air.

(PAA)[93] or symbolic aggregate approximation (SAX)[94] which allow for faster computation of high-volume data, while preserving a faithful representation. We have (Co-I PR) applied SAX and PAA to compress pulse oximetry time series and pain scores and in a novel way visualize simultaneous change in pulse oximetry and pain over time (Fig 5). Our preliminary analysis also indicates that the time series exhibits non-stationary behavior. We will address this issue by considering a range of time scales over which signal statistics are computed. This will allow us to decompose the original signals into long-term trends and short-term behavior that potentially carry complementary information.

**D.1.3.6. Novel machine learning approaches for time series analyses.** For the preoperative models we have used a set of unimodular risk factors in different probabilistic models (e.g., GAM, Bayesian, SVM)[95]. This approach allows a straightforward medical interpretation and most of the factors are robust with respect to artifacts or missing data. As we and others have demonstrated, while suitable for preoperative models, this approach ignores the information contained in the dynamic nature of intraoperative physiologic data needed for postoperative risk models[74,75]. In addition to adding new features to current algorithms we will expand our computational approach. We will extract different features that characterize physiologic data on different time scales, therefore transforming time series into a feature space. We will extract statistical features (e.g. mean, variance), global temporal and shape features (e.g. the overall trend and slope), as well as other time/frequency domain features (e.g. Fourier Transform features). As a novel approach we will include shapelets as another set of features. Shapelets are a small subsection of the time series that are particularly discriminating. We will rely on fast shapelet finding methods to extract discriminant shapelets in our high volume data[96,97]. Shapelets provide a fast measure of identifying signature patterns among physiologic data of patients with and without PC (Fig 6). The complete set of extracted physiologic features along with clinical features from EHR will be provided to a probabilistic classifier to quantify the risk of PC. We will test the ability to improve the performance of our classification algorithms by applying feature selection and dimensionality-reduction methods such as PCA[98] used in preliminary analyses or locally linear embedding[99].

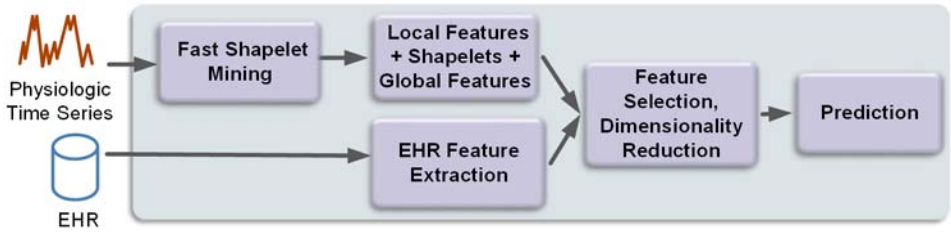


Figure 6. Analysis of time series.

**D.1.3.7. Using time discretization for prediction of physicians' actions.** Our initial algorithms have not accounted for the difference in physician's' action in response to change in physiologic parameters, such as hypotension. Management options including type and volume of fluid or the use of vasopressors may vary among physicians facing the same type of changes in physiologic parameters. These differences may either ameliorate or aggravate the physiologic stress of surgery given a patient's preoperative health. In order to account for this effect we will utilize a patient-management outliers approach for the risk-stratification. We will modify the fixed time segmentation approach that segments data in each medical record into segments of equal time duration[100,101]. The time series data in each segment is converted into a feature-space representation, representing a patient state at a given time. The patient state at this time is then linked to the actions taken in the previous time segment. This approach allows us to generate multiple examples of patient-state/action pairs for each patient, which is then used to train our statistical outlier models. We plan to use 1-hour time-segmentation and a uniform 1-hour action-response window. Thus patient state examples are generated from the record every 1 hour, and future actions associated with each of these states are determined using the 1 hour time-window following that state.

**D.1.4. Data analyses and interpretation.** *For each PC the algorithms will calculate* two *risk scores, PreOpRisk and PostOpRisk, corresponding to the probability of developing the complication that ranges from 0 to 1. A.* To test accuracy of preoperative algorithms for predicting probability for each PC we will measure the diagnostic performance of each score for discrimination (AUC, Se, Sp, predicted values, likelihood ratios,

accuracy, discrimination slope, Brier's score and misclassification rate) and calibration (Hosmer-Lemeshow (HL) test)[102]. B. These statistics will be calculated using optimal cut-off values for each risk score to stratify patients into risk categories. We will expand on our previous approach using the Youden Index to determine optimal cut-offs. C. To test the hypothesis that the addition of intraoperative physiologic time series improves the accuracy of risk prediction at the end of surgery we will compare the diagnostic accuracy of PostOPRisk scores to PreOPRisk using AUC comparison, net reclassification improvement (NRI) and integrated discrimination improvement (IDI)[103]. D. For each surgical case ASA, NSQIP, Surgical Apgar, RF and AKI risk scores[58-68] will be compared with both PreOPRisk and PostOPRisk scores for diagnostic performance. Statistical computing will be performed with SAS, version 9.3 and R software.

**D.1.5. Expected outcomes**. Upon completing this aim we expect to have a curated temporal deidentified database hosted on HPC platform (Y1) and validated algorithms for PreOPRisk (Y1) and PostOPRisk scores for each complication (Y2). Once completed SA1 will provide the computational framework for the analytical layer of IPS.

**D.1.6. Potential problems and alternative approaches**. We will address the issue of missing data in time series by developing a set of algorithms that utilizes a limited number of physiologic time series. A review of related literature for feature extraction in similar applications will also ensure that we will identify and extract relevant and informative features. The same will be done with feature selection and classification techniques to make sure we are using models that can guarantee consistent and accurate results. In the case of low accuracy results we will employ available diagnostic techniques to identify the source of problem.

**D.2. Specific Aim 2. Implement and validate two-way knowledge exchange between predictive risk models and physicians.**

While machine intelligence, realized by machine learning algorithms over high-performance computing infrastructure, is superior for high-volume data processing, human intelligence has the advantage in comprehensive medical knowledge, experience, and direct interaction with patients. *The objective of this aim is to develop and implement the knowledge exchange layer within IPS.* We will test the hypothesis that two-way knowledge transfer between algorithms and physicians' clinical reasoning improves the performance of models and facilitates their acceptance by physicians. The *rationale* is that through fusion of human and machine intelligence IPS will augment physician's decision-making. After completing SA2 we *expect* to have interactive software for knowledge exchange and a feedback loop that inputs physicians' opinion for IPS learning.

**D.2.1. Justification and feasibility.** While machine learning methods use past data to perform classification over new data, human input is essential for supervised learning algorithms by providing tagged training data[104]. Active learning algorithms also need to interactively query humans to obtain the desired outputs at new data points[104]. Input from domain experts are also used to tune the model through feature selection, feature weight setting and result interpretation[105]. The mismatch between the volume of information and humans' abilities to comprehend, retain, and organize that information can impair decision-making process[106]. Machine input through integration and interpretation of a large volume of complex data can be essential to eliminate certain errors or misjudgment and to reduce the labor-intensity of care [107]. Drs. Wang and Rashidi have on-going work in using expert input to improve machine learning results[108-111].

**D.2.2.Preliminary studies.**

**D.2.2.1. Determining weights for features in the predictive risk model.** We have developed an approach to identify top *k* features of the model that contribute to a patient's personal risk profile. For each feature in the model we calculated the additive part of the risk function and determined how it differs from an average value of a given risk factor across the data set. The difference reflects the degree to which this risk factor is responsible for changes in risk of individual patients compared to an average risk across the population. Using a validation cohort we have developed weighted features for algorithms in D.1.2.2 for each patient's record.

**D.2.2.2. Algorithms and interactive software for knowledge exchange with physicians.** Using algorithms described in D.1.2.2. we have developed web-based software (App_2) that exchanges risk prediction between models and physicians in five steps: 1. Present patient's data as a case synopsis; 2. Solicit physicians prediction for each PC using a risk analog score (RAS) from 1-100; 3. Present a model-derived risk score for each PC listing the top features contributing to the risk of individual patient; 4. Re-solicit physicians' RAS; 5. Present the expert with recommendations for interventions based on model-generated risk categories. The application's frontend is designed using HTML, JQuery and interfaces with the backend with PHP. The physician's feedback was pushed back to the relational database in the backend.

**D.2.3.Research design.** Our objective is to *A)* refine the system for transfer of the *knowledge about the individual patient's risk* from predictive models to the physician described in D.2.2.1, *B)* retrain the predictive models using expert feedback and *C)* develop an interactive interface for two-way knowledge exchange.

**D.2.3.1. Knowledge exchange between predictive models and physicians using active learning and feature selection.** While extracting important features for a population[112] is a well-studied problem, important feature extraction for *a particular patient* is not. We will expand our approach in D.2.2.1. to use features ranking according to their deviation from the general case and a lineage tracing system that can be built into the machine learning algorithms to trace important features in different components of a classifier for a particular patient. We will develop algorithms to enable abductive reasoning that provides top-K most-likely explanation given a predictive model output. Important features with high probability of PC can be presented via user interface to the medical experts to explain the reason behind the risk for given patient. The doctor's response to such knowledge would be recorded in the database for future active learning. While most active learning literature is based on the use of low-level annotated data points to improve predictive models, we will use higher-level knowledge input from physicians to improve the performance of predictive models. We envision using expert feedback as a way to transfer knowledge from physicians to algorithms to (1) modify weights and features in classification models, and (2) obtain new training data for retraining and online/active learning. Medical experts can delete/add certain features and change the weight of certain features while reviewing the explanation of PreOPRisk and PostOPRisk scores. Such expert knowledge will be incorporated into an updated classifier model. We expect to see improvement in the Se, Sp, and F-measure of the modified classifier incorporating expert feedback. We plan to implement an online-learning/re-training algorithm to train new versions of the classification model as new data comes in during prospective evaluation of the system. In order to keep the model update process stable, we expect to learn a new model after a 6-month interval during which a significant amount of new data points are collected. The new model replaces the old model only if the evaluation on the retrospective data shows statistically significant improvement in the model accuracy.

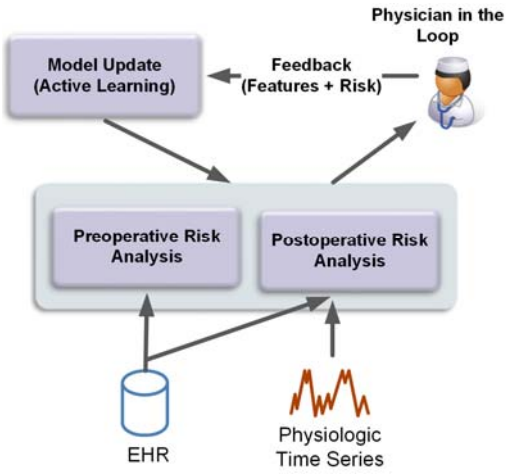


Figure 7. Process for knowledge exchange.

**D.2.3.2 Develop interface for knowledge exchange between models and physicians.** We will develop interactive software to exchange information between a prediction model and an expert in the knowledge exchange layer of IPS system. The interface aims at (1) displaying extracted knowledge and predicted likelihood of PCs to inform the physician about information that is important but buried in EMR; and (2) collecting feedback to fine tune the predictive model to achieve higher accuracy. The initial prototype of the knowledge exchange is conducted in four phases as described in D.2.2.2. (App_2). Objective measures of user's attention and information processing (time spent and links clicked) will be tracked automatically with software. The physician's feedback will be pushed back to the database. Using an iterative developmental process the prototype will be first evaluated using a focus group of 8 physicians to further improve the interface using users' feedback. Beta testing will be performed using 10 individual sessions with physicians. Finally we will conduct usability testing with two groups of 8 physicians. We will use a usability testing protocol developed following the standards outlined at www.usability.gov and each group leader will follow same script (App_3). All group sessions will be video-taped. Participants will be asked to test the software and describe their experience simultaneously using the "think aloud" technique. First, we will ask participants to explore the software on their own; second, we will give participants specific tasks to accomplish, simulating real-life situations users might face. Once participants are finished using the software, we will ask them to summarize their impressions and their overall experience (e.g., their level of satisfaction and the difficulties they encountered while using the system). Findings suggesting changes to the software will be incorporated into the final version.

**D.2.4. Data analyses and interpretation.** During re-training phase all modified models incorporating expert feedback will be tested for improvement in performance by comparing Se, Sp, AUC and F-measures.

**D.2.5. Expected results**. After completing this aim we will have fully deployed a web-based application of the software (Y2). We will implement an active learning algorithm during prospective evaluation of the IPS (Y3-5).

**D.2.6. Potential problems and alternative solutions**. One potential problem is the uncertainty in the knowledge transferred between machine and human intelligence. While the classifier may provide wrong explanations for the risk scores, the physician may also give erroneous feedback on the modification of the features and weights. This will be mitigated in our system using feedback from a surgeon, an anesthesiologist, and a resident. If the disagreement and uncertainty introduces too much noise we plan to take a holistic approach and model all the information with uncertainty and a probabilistic model. We have current work on probabilistic data integration from different machine learning algorithms and crowd-sourcing. We will explore different ways of presenting risk frequencies during the software development phase as these techniques have shown to improve understanding of the risk[113,114].

### D.3. Specific Aim 3: Implement and evaluate an intelligent perioperative system for real-time automated risk analysis.

The *objective* of this aim is to implement an ***intelligent perioperative system (IPS),*** an open-source, scalable, plug and play software system deployed on high-performance computers for perioperative risk calculation. The IPS uses real-time clinical data streams for data analytics to quantify the risk for major complications and exchange the knowledge of that risk with physicians with minimal time delay. We will test the working hypothesis that IPS is more accurate than physicians in forecasting and quantifying the risk for major PC using same clinical data in real time. The successful completion of the proposed research will establish a methodological framework for the implementation of the IPS for health risk determination and will validate the prognostic performance of the predictive models for PC.

**D.3.1. Justification and feasibility.** The IPS will be designed as an open-source, scalable, plug and play software ecosystem that will enable and automate the activities and transactions in any given complex work environment[73,115]. We have implemented similar platforms for synergistic collaboration among networking, cloud, HPC, and domain scientists that offer cyber-infrastructure for data-intensive and computation-intensive applications[116,117]. Several principles are critical for the implementation of such a system in health care and we will use our previous experience to address key methodological issues. To protect patient's privacy and enhance functionality isolation without sacrificing efficiency and manageability we will apply deregulation middleware and subsystem modularization as in our previous work[118,119] We will use our experience with an efficient data management layer of working sets[120] to support effective and real-time data analytics for highly dimensional data such as physiologic time series[121]. To boost multiple complex data analytic and query and knowledge fusion tasks we will use runtime management options applied in our previous work[122]. We will provide developer-friendly programming models for data analytics solutions to facilitate their implementation in different environments[123].

**D.3.2. Preliminary studies.**

**D.3.2.1. Previous work on software ecosystems.** Our recent assimilation projects for CognitiveEngine, GatorCloud, HiCloud and SMART[124] demonstrate the capability of our team to undertake a project of this magnitude. *CognitiveEngine*, an intelligent engine with real-time deep learning capabilities allows computational models composed of multiple processing layers to learn representations of data with multiple levels of abstraction. *GatorCloud,* a software-defined networking-enabled campus cloud provided a 10-fold boost for the UF on- and off-campus gateway networks from 10G to 100G bandwidth, and offered novel cloud services, including full-fledged computing, data, network and software as services. In the *HiCloud* project we built a suite of algorithms and programming models for easy-to-use weather forecasting applications. In the *SMART* project we developed a platform for highly dynamic scientific applications requiring intensive computational domains such as global high-resolution weather forecasting.

Table 1. Change in physicians' risk assessment after interaction with algorithms

| PC | Events | | Non-events | |
|---|---|---|---|---|
| | Risk underestimated | Score increased | Risk overestimated | Score decreased |
| AKI | 37/57 (65%) | 26/37 (70%) | 44/93 (47%) | 28/44 (64%) |
| CV | 17/43 (40%) | 12/17 (71%) | 78/107 (73%) | 56/78 (72%) |
| SS | 21/39 (54%) | 16/21 (76%) | 76/111 (69%) | 52/76 (68%) |

**D.3.2.2. Comparing predictive risk models with physicians risk assessment.** Using a software prototype in D.2.2.2 we have compared risk assessments of 20 physicians with computer algorithm risk scores for 150 patients for each of the five PC. AUC of physicians' RA ranged between 0.54 and 0.69. Algorithms' risk scores performed significantly better than physicians' RA for four of the PC with AUC ranging from 0.78 to 0.85 ($p<0.002$). After interaction with the software

most physicians who initially underestimated or overestimated risk changed their RA towards computer score resulting in improvement of prediction.

**D.3.3. Research design.** Our objective is to *A) Implement a prototype of IPS in clinical work-flow, B) Prospectively evaluate and refine the performance of the system*.

**D.3.3.1. Architecture of intelligent perioperative system.** The IPS is a software ecosystem that consists of three conceptual layers: an infrastructure layer, a data engineering layer, and a data analytics layer. The IPS is deployed in a HPC environment for real-time data integration of EHR and intraoperative physiologic time series to execute computational algorithms that can dynamically predict and quantify the risk for PC. The IPS augments clinical decision-making through the process of two-way knowledge exchange with physicians (Fig 8). The data engineering layer pulls data input from the real-time data repository (RDR) where multiple clinical data streams are integrated. It includes a stream and big data processing engine, data management, runtime management, and programming models and application program interfaces. The big data processing engine is based on the open-source Apache SparkTM (The Apache Software Foundation)[125], featuring fast in-memory stream processing (Spark Streaming), online query (Spark SQL), and comprehensive machine learning library (MLlib). IPS manages a suite of stores for apps (data analytics algorithms), models (training models, risk models), knowledge (trained or fused parameters), and data (working sets). IPS also manages multiple jobs and tasks, allocates computing, networking, and storage resources to execute data analytic models and supports two-way knowledge exchange. Following the developer-friendly simplistic programming model and APIs, our data analytics algorithms can be executed to handle large volume of data in real-time streams for many objectives concurrently at scale. The terminal devices include mobile devices and web applications. The big data capability is enabled by the in-memory distributed file system Tachyon and Hadoop distributed file system HDFS. The resources are managed by our software-defined ecosystem mechanisms and policies based on our experiences in building the GatorCloud.

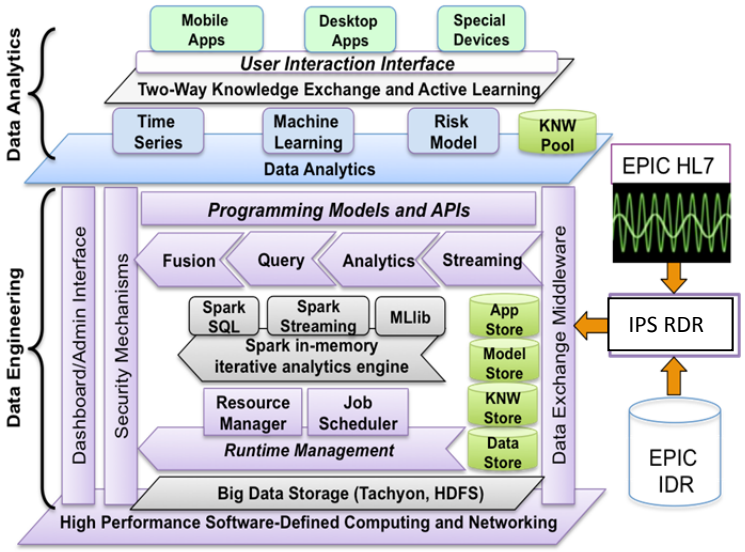


Figure 8. The structure of IPS software ecosystem.

**D.3.3.2. Implementation of IPS in clinical work-flow.** We will implement the IPS in the clinical workflow in the following steps: 1. Creation of IPS Read-time Data Repository (RDR) within the clinical applications environment by replicating DDR and by running the ECTL scripts from SA1 to perform system validation and regression tests against DDR. 2. Migration of the HPC Storage & Analytics platform into the clinical application environment. This new platform is validated by running automated testing scripts from SA1 while RDR is loaded with the same data from DDR. Any subsequent addition to this platform will follow a strict promotion process first developed and quality assured in a different system, followed by a review and approval from the Steering Committee. 3. Real-time data integration. Since direct, real-time data feed from a patient care systems such as an EHR (Epic) is not practical due to regulations and vendor product limitations, we will leverage the clinical messaging system to obtain the necessary clinical data in real-time. HL7 messages going through the interface engine will be duplicated and send to IPS databases. These messages will be parsed and stored in a landing zone and further processed as they arrive. 4. Development of new ECTL process using real-time data. Scripts will be developed to extract, cleanse, transform, and load necessary data from the landing zone into IPS database tables. Due to the asynchronous nature of these messages generated various systems and mechanisms are built to ensure only complete messages are processed and saved in the RDR.

**D.3.3.3. Study design.** We will conduct an observational study where all physicians working within adult inpatient operative practices at UFH will be recruited. We plan for three 6-month study intervals and two 4-month retraining intervals. For the entire study duration the IPS will calculate PreOPRisk and PostOPRisk scores for all inpatient surgical cases cared by enrolled physicians while operating in a silent mode with no interaction with physicians. After the first 6 months of operating only in silent mode we will perform two 6-month study intervals where IPS will additionally operate in interactive mode. During interactive mode the IPS will

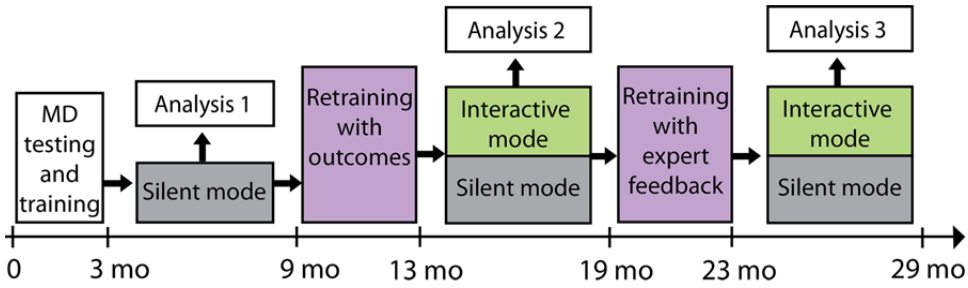


Figure 7. Timeline of the clinical validation study.

deliver PreOPRisk and PostOPRisk scores for all *first weekday surgical cases* to the enrolled physicians to solicit their risk assessment (RA) and feedback on calculated Risk Scores (RS). Study intervals will be separated by re-training periods where models will be retrained using both expert feedback and completed outcomes collected for all cases performed by enrolled operative practices during study intervals (Fig 7).

**D.3.3.4. Study participants.** We will enroll all attending surgeons and anesthesiologists in the adult inpatient operative practices. We will exclude obstetric and pediatric practices. We anticipate enrolling 60 physicians (~ 80% of faculty based on preliminary survey). Informed consent will be obtained from each participant prior to study initiation. Because numeracy and general cognitive abilities (e.g., working memory capacity[126]) affect risk judgment[127], we will test each physician at the time of enrollment for: a. *Objective and specialized numeracy skills*[128,129] and b. *Cognitive reflection task (CRT)* to measure a cognitive style linked to more careful and thorough information processing during decision making[130]. We will also obtain demographic information, training and work experience. All information will be deidentified at the end of study.

**D.3.3.5. Predictive outcomes.** Primary outcomes are major PC- sepsis, AKI, CV, RF and ICU admission > 48 hours, occurring in the first 30 postoperative days and defined as described in D0[76 77]. Secondary outcomes are hospital death, discharge other than home, readmission within 30-days of discharge and one-year survival after surgery (using IDR link to Social Security Death Index).

**D.3.3.6. Study intervention.** For all surgical cases performed during three 6-month intervals the IPS will automatically calculate PreOPRisk (at T1-prior to surgery) and PostOPRisk (at T2-end of surgery) scores for each PC without interfering with the clinical work-flow. For all first week-day cases the research coordinator will solicit physicians' RA for PC using knowledge exchange software on a dedicated iPad. We will quantify physicians' RA in absolute terms[131] using RAS (0 to 100 slider scale from "no chance" of the PC to "certainty). At T1 and T2 the corresponding IPS Risk scores will be presented afterwards with the explanations of the most important features contributing to the risk. Physicians will be asked to provide interactive feedback for each score and explanations. Once this is completed physicians' RA will be retested on the same scale. Depending on the risk score the software will present physicians with guidelines recommendations for high-risk patients for each major PC with links to scientific evidence behind the recommendation. Physicians will be allowed to explore all recommendations and objective measures of *attention* and *risk information processing* (*time spent and links clicked*) will be tracked automatically with software. The final set of questions posed to physicians will be related to the usefulness of the information provided and whether it will impact their decision for therapy.

**D.3.3.7. System learning and model retraining.** We have planned two 4-months re-training periods for system learning and model refinement through an iterative process of re-running algorithms with final primary and secondary outcomes and physician feedback to improve prediction as outlined in D2.3.1.

**D.3.3.8. Sample size and feasibility.** Between July 2012 and April 2013 16,206 weekday adult surgical cases were performed at UFH and 72% of them involved at least overnight stay. Per each 6-month study interval we expect to have at least 5,000 surgical cases performed for all practices and at least 1,000 physician RAS compared to IPS risk score. A sample size of 5,000 will offer 80% power to detect a difference between 0.03 and 0.06 when comparing models' Se with Se0 of 0.80 when PC prevalence ranges between 39% and 5%, respectively[85]. This sample size allows estimations of Sp with 95% CI width of 0.02-0.04 and of Se 95% CI width of 0.04-0.18 for all possible values of Se, Sp and the PC prevalence. It will result in the maximum 95% CI width for AUC of 0.08 when PC prevalence is 5%, 0.06 when prevalence is 10% and 0.04 when prevalence is >20%[86]. Same sample size will detect a difference between 0.03 and 0.11 in Se between preoperative and postoperative predictive models with a power of 80% assuming different values for Se0 between 0.6 and 0.9 and for PC prevalence between 5% and 30%. On average 3-12 physicians are involved in first morning surgical cases during weekdays and for most of the cases we will have separate scores from anesthesiologists and surgeons. With the potential drop in recruitment goal to 3 physicians per day we anticipate total of 1200 - 2400 physicians' RA per each 6-month interval. The predictive accuracy for physicians' RA (Se 0.63, Sp 0.75

and AUC 0.69) was set using preliminary data and previous reports[132]. A sample size of 1,200 physicians' RA will detect a difference between 0.1 and 0.2 between Se of IPS predictive models and physician's RA assuming prevalence of PC between 39% and 5%, respectively[85].

**D.3.4. Data analyses and interpretation.** All analyses will be performed after all predictive outcomes are collected for all calculated IPS risk scores (Fig 7). *For Analyses 1-3 we* will prospectively validate the accuracy of all risk models for model discrimination (Se, Sp, predicted values, likelihood ratios, accuracy, AUC, F-measure) and calibration[102]. *For Analysis 3* we will assess the effect of retraining with expert feedback by comparing the performance of predictive models before and after re-training period (Se, Sp, AUC comparison). *For Analysis 3* we will use the final set of models to test improvement in performance of PostOPRisk score compared to PreOPRisk score for each PC. Optimal cut-offs will be determined for each score to risk stratify patients, and NRI and IDI will be used to assess improvement in the classification of patients into low and high-risk categories[103]. *For Analyses 2 - 3* we will test the hypothesis that IPS models are superior to physicians' RA by comparing performance of PreOPRisk and PostOPRisk scores to paired physician's scores using AUC comparisons, NRI and IDI. To assess the change in physicians' RA after knowledge exchange with the IPS system at each time point of intervention (T1-before surgery and T2-at the end of surgery) we will measure the proportion of users changing their absolute risk perception, proportion of users moving their perceived risk closer to the system prediction and absolute amount of change and whether repetitive exposure to software improves risk prediction after providers were tested repetitively. The absolute change of physicians risk assessment after knowledge exchange will be evaluated by the Wilcoxon signed rank sum test appropriately, especially in cases which are underestimated or overestimated initially by physicians. Attention, information processing and usefulness of information in the software (% times physicians choose to look into explanations, agreement to use consensus recommendations when the algorithm predicts high risk) will be assessed in descriptive analyses. To test the clinical usefulness of IPS risk scores we will utilize decision curve analysis to plot the net benefit achieved by making decisions based on the model predictions[133]. Using generalized estimating equations and mixed models we will assess a. the effect of numeracy and CRT on change in risk perceptions and b. response to explanations of risk by the knowledge transfer tool. Statistical computing will be performed with SAS software v9.3, R and TreeAge Pro 2012.

**D.3.5. Expected results**. The IPS hosting in the clinical environment will be completed by the end of Y2. Refinement of each score will be done through Y3-5 during implementation phase. Retraining of the final models will be done in the Y3-5 using the database from the completed prospective study with labeled prediction end-points. In Y5 all data analyses will be completed and the prospectively collected database will be de-identified.

**D.3.6. Potential problems and alternative solutions**. Risk stratification and comparison of Se, Sp, and accuracy based on PreOPRisk and PostOPRisk Scores will be compared at optimal cut-off values. Optimal cut-off values for risk score to stratify patients in risk categories will be calculated expanding on our previous approach using the Youden Index and prevalence for lower cut-off and the point where the accuracy curve starts to plateau for higher cut-off[11,134]. These cut-off values obtained from various risk score sets might vary. In that case, separate cut-offs will be calculated to risk stratify patients for IPS Risk Scores, and physicians' RA. To evaluate model reclassification before and after retraining, patients will be stratified using cut-offs derived from the data in each phase.

**E. TIMETABLE.**

| | Year 1 | | | Year 2 | | | Year 3 | | | Year 4 | | | Year 5 | | | Teams |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Specific Aim 1.** | | | | | | | | | | | | | | | | |
| De-identified Data Repository | | | | | | | | | | | | | | | | GL, XL, TOB, AB |
| Refine and evaluate models | | | | | | | | | | | | | | | | PP, AB, PR, PM |
| **Specific Aim 2.** | | | | | | | | | | | | | | | | |
| Active learning and feature selection | | | | | | | | | | | | | | | | AB, PR, DW |
| Interactive software for knowledge exchange | | | | | | | | | | | | | | | | AB, PR, DW |
| **Specific Aim 3.** | | | | | | | | | | | | | | | | |
| Real-time system implementation | | | | | | | | | | | | | | | | XL, GL, AB |
| Prospective evaluation of the system | | | | | | | | | | | | | | | | AB, XL, PR, TOB |

**REFERENCES:**


1. Lee PH, Gawande AA. The number of surgical procedures in an American lifetime in 3 states. J Am Coll Surg. 2008;207(3):S75-S.
2. Weiser TG, Regenbogen SE, Thompson KD, Haynes AB, Lipsitz SR, Berry WR, Gawande AA. An estimation of the global volume of surgery: a modelling strategy based on available data. The Lancet. 2008;372(9633):139-44.
3. Hall BL, Hamilton BH, Richards K, Bilimoria KY, Cohen ME, Ko CY. Does Surgical Quality Improve in the American College of Surgeons National Surgical Quality Improvement Program: An Evaluation of All Participating Hospitals. Ann Surg. 2009;250(3):363-76
4. Grocott MPW, Pearse RM. Perioperative medicine: the future of anaesthesia? Br J Anaesth. 2012;108(5):723-6.
5. Gawande AA, Regenbogen SE. Critical need for objective assessment of postsurgical patients. Anesthesiology. 2011;114(6):1269-70.
6. Nair P. QnAs with Leroy E. Hood. Proc Natl Acad Sci U S A. 2011;108(23):9325.
7. Galas DJ, Leroy HE. Systems Biology and Emerging Technologies Will Catalyze the Transition from Reactive Medicine to Predictive, Personalized, Preventive and Participatory (P4) Medicine. IBC. 2009;1(2):1-4.
8. Khuri SF, Henderson WG, DePalma RG, Mosca C, Healey NA, Kumbhani DJ, Program tPitVNSQI. Determinants of Long-Term Survival After Major Surgery and the Adverse Effect of Postoperative Complications. Ann Surg. 2005;242(3):326-43.
9. Korenkevych D, Ozrazgat-Baslanti T, Thottakkara P, Hobson CE, Pardalos P, Momcilovic P, Bihorac A. The Pattern of Longitudinal Change in Serum Creatinine and 90-Day Mortality After Major Surgery. Ann Surg. 2015.
10. Hobson CE, Ozrazgat-Baslanti T, Kuxhausen A, Thottakkara P, Efron PA, Moore FA, Moldawer LL, Segal MS, Bihorac A. Cost and Mortality Associated With Postoperative Acute Kidney Injury. Ann Surg. 2014. PMCID: PMC4247993.
11. Bihorac A, Brennan M, Ozrazgat-Baslanti T, Bozorgmehri S, Efron PA, Moore FA, Segal MS, Hobson CE. National surgical quality improvement program underestimates the risk associated with mild and moderate postoperative acute kidney injury. Crit Care Med. 2013;41(11):2570-83. PMCID: PMC3812338.
12. Bihorac A, Delano MJ, Schold JD, Lopez MC, Nathens AB, Maier RV, Layon AJ, Baker HV, Moldawer LL. Incidence, clinical predictors, genomics, and outcome of acute kidney injury among trauma patients. Ann Surg. 2010;252(1):158-65. PMCID: PMC3357629.
13. Huber M, Ozrazgat-Baslanti T, Thottakkara P, Efron PA, Feezor R, Hobson CE, Bihorac A. Mortality and cost of acute and chronic kidney disease after vascular surgery. Ann Vasc Surg. 2015;29(4).
14. Vaught A, Ozrazgat-Baslanti T, Javed A, Morgan L, Hobson C, Bihorac A. Acute kidney injury in major gynaecological surgery: an observational study. BJOG. 2014. PMCID: PMC4334755.
15. White LE, Hassoun HT, Bihorac A, Moore LJ, Sailors RM, McKinley BA, Valdivia A, Moore FA. Acute kidney injury is surprisingly common and a powerful predictor of mortality in surgical sepsis. J Trauma Acute Care Surg. 2013;75(3):432-8. PMCID: PMC3823059.
16. Arnaoutakis GJ, Bihorac A, Martin TD, Hess PJ, Jr., Klodell CT, Ejaz AA, Garvan C, Tribble CG, Beaver TM. RIFLE criteria for acute kidney injury in aortic arch surgery. J Thorac Cardiovasc Surg. 2007;134(6):1554-60; discussion 60-1.
17. Bihorac A, Yavas S, Subbiah S, Hobson CE, Schold JD, Gabrielli A, Layon AJ, Segal MS. Long-term risk of mortality and acute kidney injury during hospitalization after major surgery. Ann Surg. 2009;249(5):851-8.
18. Hobson CE, Yavas S, Segal MS, Schold JD, Tribble CG, Layon AJ, Bihorac A. Acute kidney injury is associated with increased long-term mortality after cardiothoracic surgery. Circulation. 2009;119(18):2444-53.
19. Bihorac A, Schold JD, Hobson CE. Acute kidney injury and long-term outcomes: more to learn. Kidney Int. 2010;77(8):745.
20. Bihorac A, Schold JD, Hobson CE. Long-term mortality associated with acute kidney injury requiring dialysis. JAMA. 2010;303(3):229.
21. Gentile LF, Cuenca AG, Efron PA, Ang D, Bihorac A, McKinley B, Moldawer LL, Moore FA. Persistent inflammation and immunosuppression: A common syndrome and new horizon for surgical intensive care. J Trauma Acute Care Surg. 2012;72(6):1491-501.
22. Vanzant EL, Lopez CM, Ozrazgat-Baslanti T, Ungaro R, Davis R, Cuenca AG, Gentile LF, Nacionales DC, Cuenca AL, Bihorac A, Leeuwenburgh C, Lanz J, Baker HV, McKinley B, Moldawer LL, Moore FA, Efron

PA. Persistent inflammation, immunosuppression, and catabolism syndrome after severe blunt trauma. J Trauma Acute Care Surg. 2014;76(1):21-9; discussion 9-30.
23. Yende S, Linde-Zwirble W, Mayr F, Weissfeld LA, Reis S, Angus DC. Risk of cardiovascular events in survivors of severe sepsis. Am J Respir Crit Care Med. 2014;189(9):1065-74. PMCID: PMC4098105.
24. Iwashyna TJ, Cooke CR, Wunsch H, Kahn JM. Population burden of long-term survivorship after severe sepsis in older americans. J Am Geriatr Soc. 2012;60(6):1070-7.
25. Unroe M, Kahn JM, Carson SS, Govert JA, Martinu T, Sathy SJ, Clay AS, Chia J, Gray A, Tulsky JA, Cox CE. One-year trajectories of care and resource utilization for recipients of prolonged mechanical ventilation: a cohort study. Ann Intern Med. 2010;153(3):167-75. PMCID: 2941154.
26. Iwashyna TJ. Survivorship will be the defining challenge of critical care in the 21st century. Ann Intern Med. 2010;153(3):204.
27. Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. JAMA: The Journal of the American Medical Association. 2010;304(16):1787-94.
28. Hopkins RO, Jackson JC. Short- and long-term cognitive outcomes in intensive care unit survivors. Clin Chest Med. 2009;30(1):143-53, ix.
29. Gurgel ST, do Nascimento P, Jr. Maintaining tissue perfusion in high-risk surgical patients: a systematic review of randomized clinical trials. Anesth Analg. 2011;112(6):1384-91.
30. Hamilton MA, Cecconi M, Rhodes A. A systematic review and meta-analysis on the use of preemptive hemodynamic intervention to improve postoperative outcomes in moderate and high-risk surgical patients. Anesth Analg. 2011;112(6):1392-402.
31. Poeze M, Greve JW, Ramsay G. Meta-analysis of hemodynamic optimization: relationship to methodological quality. Crit Care. 2005;9(6):R771-9. PMCID: 1414050.
32. Brienza N, Giglio MT, Marucci M, Fiore T. Does perioperative hemodynamic optimization protect renal function in surgical patients? A meta-analytic study. Crit Care Med. 2009;37(6):2079-90.
33. Cecconi M, Corredor C, Arulkumaran N, Abuella G, Ball J, Grounds RM, Hamilton M, Rhodes A. Clinical review: Goal-directed therapy-what is the evidence in surgical patients? The effect on different risk groups. Crit Care. 2013;17(2):209. PMCID: 3679445.
34. Guay J, Choi P, Suresh S, Albert N, Kopp S, Pace NL. Neuraxial blockade for the prevention of postoperative mortality and major morbidity: an overview of Cochrane systematic reviews. Cochrane Database Syst Rev. 2014;1:CD010108.
35. Popping DM, Elia N, Van Aken HK, Marret E, Schug SA, Kranke P, Wenk M, Tramer MR. Impact of epidural analgesia on mortality and morbidity after surgery: systematic review and meta-analysis of randomized controlled trials. Ann Surg. 2014;259(6):1056-67.
36. Rodgers A, Walker N, Schug S, McKee A, Kehlet H, van Zundert A, Sage D, Futter M, Saville G, Clark T, MacMahon S. Reduction of postoperative mortality and morbidity with epidural or spinal anaesthesia: results from overview of randomised trials. BMJ. 2000;321(7275):1493. PMCID: PMC27550.
37. Landoni G, Greco T, Biondi-Zoccai G, Nigro Neto C, Febres D, Pintaudi M, Pasin L, Cabrini L, Finco G, Zangrillo A. Anaesthetic drugs and survival: a Bayesian network meta-analysis of randomized trials in cardiac surgery. Br J Anaesth. 2013;111(6):886-96.
38. Landoni G, Biondi-Zoccai GG, Zangrillo A, Bignami E, D'Avolio S, Marchetti C, Calabro MG, Fochi O, Guarracino F, Tritapepe L, De Hert S, Torri G. Desflurane and sevoflurane in cardiac surgery: a meta-analysis of randomized clinical trials. J Cardiothorac Vasc Anesth. 2007;21(4):502-11.
39. Giakoumidakis K, Eltheni R, Patelarou E, Theologou S, Patris V, Michopanou N, Mikropoulos T, Brokalaki H. Effects of intensive glycemic control on outcomes of cardiac surgery. Heart Lung. 2013;42(2):146-51.
40. Haga KK, McClymont KL, Clarke S, Grounds RS, Ng KY, Glyde DW, Loveless RJ, Carter GH, Alston RP. The effect of tight glycaemic control, during and after cardiac surgery, on patient mortality and morbidity: A systematic review and meta-analysis. J Cardiothorac Surg. 2011;6:3. PMCID: PMC3023693.
41. van den Berghe G, Wouters P, Weekers F, Verwaest C, Bruyninckx F, Schetz M, Vlasselaers D, Ferdinande P, Lauwers P, Bouillon R. Intensive insulin therapy in critically ill patients. N Engl J Med. 2001;345(19):1359-67.
42. Maerz LL, Akhtar S. Perioperative glycemic management in 2011: paradigm shifts. Curr Opin Crit Care. 2011;17(4):370-5.
43. Stephan F, Barrucand B, Petit P, Rezaiguia-Delclaux S, Medard A, Delannoy B, Cosserant B, Flicoteaux G, Imbert A, Pilorge C, Berard L, Bi POPSG. High-Flow Nasal Oxygen vs Noninvasive Positive

Airway Pressure in Hypoxemic Patients After Cardiothoracic Surgery: A Randomized Clinical Trial. JAMA. 2015;313(23):2331-9.
44. Del Sorbo L, Ferguson ND. High-Flow Nasal Cannulae or Noninvasive Ventilation for Management of Postoperative Respiratory Failure. JAMA. 2015;313(23):2325-6.
45. Antonelli M, Conti G, Bufi M, Costa MG, Lappa A, Rocco M, Gasparetto A, Meduri GU. Noninvasive ventilation for treatment of acute respiratory failure in patients undergoing solid organ transplantation: a randomized trial. JAMA. 2000;283(2):235-41.
46. van de Watering LM, Hermans J, Houbiers JG, van den Broek PJ, Bouter H, Boer F, Harvey MS, Huysmans HA, Brand A. Beneficial effects of leukocyte depletion of transfused blood on postoperative complications in patients undergoing cardiac surgery: a randomized clinical trial. Circulation. 1998;97(6):562-8.
47. Bilgin YM, van de Watering LM, Eijsman L, Versteegh MI, Brand R, van Oers MH, Brand A. Double-blind, randomized controlled trial on the effect of leukocyte-depleted erythrocyte transfusions in cardiac valve surgery. Circulation. 2004;109(22):2755-60.
48. Thielmann M, Kottenberg E, Kleinbongard P, Wendt D, Gedik N, Pasa S, Price V, Tsagakis K, Neuhauser M, Peters J, Jakob H, Heusch G. Cardioprotective and prognostic effects of remote ischaemic preconditioning in patients undergoing coronary artery bypass surgery: a single-centre randomised, double-blind, controlled trial. Lancet. 2013;382(9892):597-604.
49. Zarbock A, Schmidt C, Van Aken H, Wempe C, Martens S, Zahn PK, Wolf B, Goebel U, Schwer CI, Rosenberger P, Haeberle H, Gorlich D, Kellum JA, Meersch M, Renal RI. Effect of remote ischemic preconditioning on kidney injury among high-risk patients undergoing cardiac surgery: a randomized clinical trial. JAMA. 2015;313(21):2133-41.
50. de Waal BA, Buise MP, van Zundert AA. Perioperative statin therapy in patients at high risk for cardiovascular morbidity undergoing surgery: a review. Br J Anaesth. 2015;114(1):44-52.
51. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. Clinical Practice Guideline for Acute Kidney Injury. Kidney inter, Suppl. 2012;2(1):1-138.
52. Bihorac A, Chawla LS, Shaw AD, Al-Khafaji A, Davison DL, Demuth GE, Fitzgerald R, Ng Gong M, Graham DD, Gunnerson K, Heung M, Jortani S, Kleerup E, Koyner JL, Krell K, Letourneau J, Lissauer M, Miner J, Nguyen HB, Ortega LM, Self WH, Sellman R, Shi J, Straseski J, Szalados JE, Wilber ST, Walker MG, Wilson J, Wunderink R, Zimmerman J, Kellum JA. Validation of Cell-Cycle Arrest Biomarkers for Acute Kidney Injury Using Clinical Adjudication. Am J Respir Crit Care Med. 2014.
53. Bihorac A. Guiding AKI Prevention Using Biomarkers. Critical Connections. 2015;14(2):1.
54. Croft CA, Moore FA, Efron PA, Marker PS, Gabrielli A, Westhoff LS, Lottenberg L, Jordan J, Klink V, Sailors RM, McKinley BA. Computer versus paper system for recognition and management of sepsis in surgical intensive care. J Trauma Acute Care Surg. 2014;76(2):311-7; discussion 8-9.
55. McKinley BA, Moore LJ, Sucher JF, Todd SR, Turner KL, Valdivia A, Sailors RM, Moore FA. Computer protocol facilitates evidence-based care of sepsis in the surgical intensive care unit. J Trauma. 2011;70(5):1153-66; discussion 66-7.
56. Moore LJ, Turner KL, Todd SR, McKinley B, Moore FA. Computerized clinical decision support improves mortality in intra abdominal surgical sepsis. Am J Surg. 2010;200(6):839-43; discussion 43-4.
57. Sucher JF, Todd SR, Jones SL, Throckmorton T, Turner KL, Moore FA. Robotic telepresence: a helpful adjunct that is viewed favorably by critically ill surgical patients. Am J Surg. 2011;202(6):843-7.
58. Lake AP, Williams EG. ASA classification and perioperative variables: graded anaesthesia score? Br J Anaesth. 1997;78(2):228-9.
59. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. Br J Surg. 1991;78(3):355-60.
60. Gawande AA, Kwaan MR, Regenbogen SE, Lipsitz SA, Zinner MJ. An Apgar score for surgery. J Am Coll Surg. 2007;204(2):201-8.
61. Hobson CE, Singhania G, Bihorac A. Acute kidney injury in the surgical patient. Crit Care Clin. 2015.
62. Huen SC, Parikh CR. Predicting acute kidney injury after cardiac surgery: a systematic review. Ann Thorac Surg. 2012;93(1):337-47. PMCID: PMC3286599.
63. Gupta H, Gupta PK, Fang X, Miller WJ, Cemaj S, Forse RA, Morrow LE. Development and validation of a risk calculator predicting postoperative respiratory failure. Chest. 2011;140(5):1207-15.
64. Ramachandran SK, Nafiu OO, Ghaferi A, Tremper KK, Shanks A, Kheterpal S. Independent predictors and outcomes of unanticipated early postoperative tracheal intubation after nonemergent, noncardiac surgery. Anesthesiology. 2011;115(1):44-53.

65. Brueckmann B, Villa-Uribe JL, Bateman BT, Grosse-Sundrup M, Hess DR, Schlett CL, Eikermann M. Development and validation of a score for prediction of postoperative respiratory complications. Anesthesiology. 2013;118(6):1276-85.
66. Kor DJ, Warner DO, Alsara A, Fernandez-Perez ER, Malinchoc M, Kashyap R, Li G, Gajic O. Derivation and diagnostic accuracy of the surgical lung injury prediction model. Anesthesiology. 2011;115(1):117-28. PMCID: PMC3986041.
67. Surgical Risk Calculator. [cited 06/28/2015]; Available from: http://www.surgicalriskcalculator.com.
68. Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning : data mining, inference, and prediction. 2nd ed. New York: Springer; 2009.
69. Rutkowski AF, Saunders CS. Understanding overload: An emotional cognitive model. Proceedings of the 71st Academy of Management Conference 2011: West meets East: Enlightening, Balancing, Transcending, San Antonio: Academy of Management. 2011.
70. Pluyter JR. Designing immersive surgical training against information technology-related overload in the operating room: Tilburg University; 2012 Contract No.: Document Number|.
71. Jakimowicz JJ, Cuschieri A. Time for evidence-based minimal access surgery training--simulate or sink. Surg Endosc. 2005;19(12):1521-2.
72. Bitterman N. Technologies and solutions for data display in the operating room. J Clin Monit Comput. 2006;20(3):165-73.
73. Mistrík I. Collaborative software engineering. Heildelberg Germany ; London ; New York: Springer; 2010.
74. Pimentel MA, Clifton DA, Clifton L, Watkinson PJ, Tarassenko L. Modelling physiological deterioration in post-operative patient vital-sign data. Med Biol Eng Comput. 2013;51(8):869-77. PMCID: 3709086.
75. Johnson AE, Burgess J, Pimentel MA, Clifton DA, Young JD, Watkinson PJ, Tarassenko L. Physiological trajectory of patients pre and post ICU discharge. Conf Proc IEEE Eng Med Biol Soc. 2014;2014:3160-3.
76. American College of Surgeons National Surgical Quality Improvement Program. User Guide for the 2010 Participant Use Data File2010: Available from: http://site.acsnsqip.org/wp-content/uploads/2012/03/2010-User-Guide_FINAL.pdf.
77. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney inter, Suppl. 2013;3(1):1-150.
78. Hug CW, Szolovits P. ICU acuity: real-time models versus daily models. AMIA Annu Symp Proc. 2009;2009:260-4. PMCID: PMC2815497.
79. Mayaud L, Lai PS, Clifford GD, Tarassenko L, Celi LA, Annane D. Dynamic data during hypotensive episode improves mortality predictions among patients with sepsis and hypotension. Crit Care Med. 2013;41(4):954-62. PMCID: 3609896.
80. Multicenter Perioperative Outcomes Group. The Regents of the University of Michigan; [cited 2013 May 26]; Available from: http://mpog.med.umich.edu/.
81. Erdal BS, Liu J, Ding J, Chen J, Marsh CB, Kamal J, Clymer BD. A database de-identification framework to enable direct queries on medical data for secondary use. Methods Inf Med. 2012;51(3):229-41.
82. Johnson JA, Elsey AR, Clare-Salzler MJ, Nessl D, Conlon M, Nelson DR. University of Florida & Shands Personalized Medicine Program: Clinical implementation of pharmacogenetics. Pharmacogenomics. 2013;In Press.
83. Tighe PJ, Lucas SD, Edwards DA, Boezaart AP, Aytug H, Bihorac A. Use of machine-learning classifiers to predict requests for preoperative acute pain service consultation. Pain Med. 2012;13(10):1347-57.
84. Bihorac A, Korenkevych D, Ozrazgat Baslanti T, Momcilovic P, Pardalos P, Segal MS, Moore FA. Database communication enables machine learning classifiers to predict postoperative acute kidney injury. Crit Care Med. 2013;41(12):714.
85. Li J, Fine J. On sample size for sensitivity and specificity in prospective diagnostic accuracy studies. Stat Med. 2004;23(16):2537-50.
86. Hanley JA, Mcneil BJ. The Meaning and Use of the Area under a Receiver Operating Characteristic (Roc) Curve. Radiology. 1982;143(1):29-36.
87. Hanna J, Joseph E, Brochhausen M, Hogan WR. Building a drug ontology based on RxNorm and other sources. J Biomed Semantics. 2013;4(1):44. PMCID: PMC3931349.
88. Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. J Am Med Inform Assoc. 1997;4(5):342-55. PMCID: PMC61252.
89. Hogan WR, Wagner MM. Free-text fields change the meaning of coded data. Proc AMIA Annu Fall Symp. 1996:517-21. PMCID: PMC2233180.

90. KESTIN I, MILLER B, LOCKHART C. AUDITORY ALARMS DURING ANESTHESIA MONITORING. Anesthesiology. 1988;69(1):106-9.
91. Ebrahim M, Feldman J, Bar-Kana I. A robust sensor fusion method for heart rate estimation. J Clin Monit. 1997;13(6):385-93.
92. Lin C, Liu S, Wang J, Wen Z. Reduction of interference in oscillometric arterial blood pressure measurement using fuzzy logic. IEEE Trans Biomed Eng. 2003;50(4):432-41.
93. Guo C, Li H, Pan D. An Improved Piecewise Aggregate Approximation Based on Statistical Features for Time Series Mining. In: Bi Y, Williams M-A, editors. Knowledge Science, Engineering and Management: Springer Berlin Heidelberg; 2010. p. 234-44.
94. Lin J, Keogh E, Wei L, Lonardi S. Experiencing SAX: a novel symbolic representation of time series. Data Min Knowl Disc. 2007;15(2):107-44.
95. Saria S, Rajani AK, Gould J, Koller D, Penn AA. Integration of early physiological responses predicts later illness severity in preterm infants. Sci Transl Med. 2010;2(48):48ra65.
96. Mueen A, Keogh E, Young N. Logical-shapelets: an expressive primitive for time series classification. Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining; San Diego, California, USA. 2020587: ACM; 2011. p. 1154-62.
97. Lines J, Davis LM, Hills J, Bagnall A. A shapelet transform for time series classification. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining; Beijing, China. 2339579: ACM; 2012. p. 289-97.
98. Pearson K. On lines and planes of closest fit to systems of points in space. Philosophical Magazine. 1901;2(7-12):559-72.
99. Bishop CM. Pattern recognition and machine learning. New York: Springer; 2006.
100. Batal I, Fradkin D, Harrison J, Moerchen F, Hauskrecht M. Mining Recent Temporal Patterns for Event Detection in Multivariate Time Series Data. KDD. 2012;2012:280-8. PMCID: PMC4414327.
101. Batal I, Valizadegan H, Cooper GF, Hauskrecht M. A Temporal Pattern Mining Approach for Classifying Electronic Health Record Data. ACM Trans Intell Syst Technol. 2013;4(4). PMCID: PMC4192602.
102. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. 2010;21(1):128-38.
103. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008;27(2):157-72; discussion 207-12.
104. Murphy KP. Machine learning: a probabilistic perspective. Cambridge, MA MIT Press; 2012.
105. Anderson M, Antenucci D, Bittorf V, Burgess M, Cafarella M, Kumar A, Niu F, Park Y, Ré C, Zhang C. Brainwash: A Data System for Feature Engineering. 6th Biennial Conference on Innovative Data Systems Research (CIDR '13). 2013.
106. Morris AH. Developing and implementing computerized protocols for standardization of clinical decisions. Ann Intern Med. 2000;132(5):373-83.
107. Macdonnell M, Darzi A. A Key To Slower Health Spending Growth Worldwide Will Be Unlocking Innovation To Reduce The Labor-Intensity Of Care. Health Aff (Millwood). 2013;32(4):653-60.
108. Purdy D, Wang DZ. Bonsai: Exploration and Cultivation of Machine Learning Models. Proceedings of Joint Statistical Meetings. 2008.
109. Wang DZ, Chen Y, Goldberg S, Grant C, Li K. Automatic Knowledge Base Construction using Probabilistic Extraction, Deductive Reasoning, and Human Feedback. Proceedings of NAACL-HLT. 2012.
110. Rashidi P, Cook DJ. Ask me better questions: active learning queries based on rule induction. Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining; San Diego, California, USA. 2020559: ACM; 2011. p. 904-12.
111. Goldberg S, Wang DZ, Kraska T. CASTLE: Crowd-Assisted System for Textual Labeling and Extraction. Proceedings of AAAI HCOMP. 2013:51-9.
112. Liu H, Motoda H. Computational methods of feature selection. Boca Raton: Chapman & Hall/CRC; 2008.
113. Reyna VF, Nelson WL, Han PK, Dieckmann NF. How numeracy influences risk comprehension and medical decision making. Psychol Bull. 2009;135(6):943-73. PMCID: 2844786.
114. Peters E, Hart PS, Fraenkel L. Informing Patients. Med Decis Making. 2011;31(3):432-6.
115. Zhao H, Yu Z, Tiwari S, Mao X, Lee K, Wolinsky D, Li X, Figueiredo R. CloudBay: Enabling an Online Resource Market Place for Open Clouds. The 5th IEEE/ACM International Conference on Utility and Cloud Computing. 2012.

116. Li X, Parashar M. Hybrid Runtime Management of Space-Time Heterogeneity for Parallel Structured Adaptive Applications. IEEE Transactions on Parallel and Distributed Systems 2007;18(9).
117. Yang X, Yi Z, Li M, Li X. Mammoth: Autonomic Data Processing Framework for Scientific State-Transition Applications. The ACM Cloud and Autonomic Computing Conference 2013.
118. Yan C, Yang X, Yu Z, Li M, Li X. IncMR: Incremental Data Processing based on MapReduce. The 5th IEEE International Conference on Cloud Computing. 2012.
119. Zhao H, Pan M, Liu X, Li X, Fang Y. Optimal Resource Rental Planning for Elastic Applications in Cloud Market. The 26th IEEE International Parallel and Distributed Processing Symposium. 2012.
120. Zhao H, Liu X, Li X. Hypergraph-based Task-Bundle Scheduling Towards Efficiency and Fairness in Heterogeneous Distributed Systems. The 24th IEEE International Parallel and Distributed Processing Symposium. 2010.
121. Zhao H, Li X. The 9th IEEE International Symposium on Cluster Computing and the Grid. Efficient Grid Task-Bundle Allocation Using Bargaining Based Self-Adaptive Auction. 2009.
122. Li X, Parashar M. The 8th IEEE International Symposium on Cluster Computing and the Grid GridMate: A Portable Simulation Environment for Large-Scale Adaptive Scientific Applications. 2008.
123. Liu X, Zhao H, Yang X, Li X. SinkTrail: A Proactive Data Reporting Protocol for Wireless Sensor Networks. IEEE Transactions on Computers. 2013;62(1):151-62.
124. Scalable Software Systems Laboratory ($S^3$Lab). Scalable Software Systems Laboratory ($S^3$Lab): Current Projects. [cited 2013 06/01]; Available from: http://www.s3lab.ece.ufl.edu/projects/.
125. Apache Spark. [cited 06/24/2015]; Available from: https://spark.apache.org/.
126. Cokely ET, Kelley CM, Gilchrist AL. Sources of individual differences in working memory: contributions of strategy to capacity. Psychon Bull Rev. 2006;13(6):991-7.
127. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. How making a risk estimate can change the feel of that risk: shifting attitudes toward breast cancer risk in a general public survey. Patient Educ Couns. 2005;57(3):294-9.
128. Lipkus IM, Samsa G, Rimer BK. General Performance on a Numeracy Scale among Highly Educated Samples. Med Decis Making. 2001;21(1):37-44.
129. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med. 1997;127(11):966-72.
130. Frederick S. Cognitive Reflection and Decision Making. J Econ Perspect. 2005;19(4):25–42.
131. Harle C, Padman R, Downs J. The impact of web-based diabetes risk calculators on information processing and risk perceptions. AMIA Annu Symp Proc. 2008:283-7. PMCID: 2656037.
132. Hightower CE, Riedel BJ, Feig BW, Morris GS, Ensor JE, Woodruff VD, Daley-Norman MD, Sun XG. A pilot study evaluating predictors of postoperative outcomes after major abdominal surgery: physiological capacity compared with the ASA physical status classification system. Br J Anaesth. 2010.
133. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565-74. PMCID: PMC2577036.
134. Pickering JW, Endre ZH. New metrics for assessing diagnostic potential of candidate biomarkers. Clin J Am Soc Nephrol. 2012;7(8):1355-64.