

# HARVARD Unexpected Deaths in Patients Admitted to Hospital from Emergency Department T.H. CHAN

Mathias Blom<sup>1</sup>, Isabelle Mieling<sup>1</sup>, Linying Zhang<sup>1</sup>, Nancy Zhang<sup>1</sup>

<sup>1</sup>Harvard T.H. Chan School of Public Health, Boston, Massachusetts

## **Background:**

Ideally, inpatient healthcare services would be of high quality and readily available for all. However, several academic studies suggest that quality of services may vary. For example, Swedish researchers have shown that inpatient services are less accessible during hospital overcrowding (i.e. when hospital beds are scarce) [1]. Moreover, patients seeking emergency department care during times of high Emergency Department utilization experience worse outcomes than those who come at less busy times [2]. Emergency Department overcrowding has also been related to an increased incidence of in-hospital deaths [3, 4, 5]. Although not explicitly addressed in these studies, it is not unlikely that part of the increase in in-hospital mortality could be attributed to an admission-bias, where overcrowding in the Emergency Department is a symptom of hospital overcrowding, causing only the sickest patients to be admitted to a hospital bed [1]. Connections have been established between healthcare organization stress and patient outcomes and there may be other potentially time-varying confounders that impact patient outcomes. This is of academic as well as public interest, as knowledge about such confounders may be useful for managerial decisions. Hence, the aim of this study is to explore whether the proportion of in-hospital deaths in patients admitted to the hospital from the Emergency Department is subject to significant variation over time.

# **Methods:**

The study was conducted as an observational study on prospectively collected data from a large academic hospital Emergency Department over the course of three years. This study allowed for predicting in-hospital deaths using a granular set of variables, including disease severity, an array of laboratory results and various patient demographics.

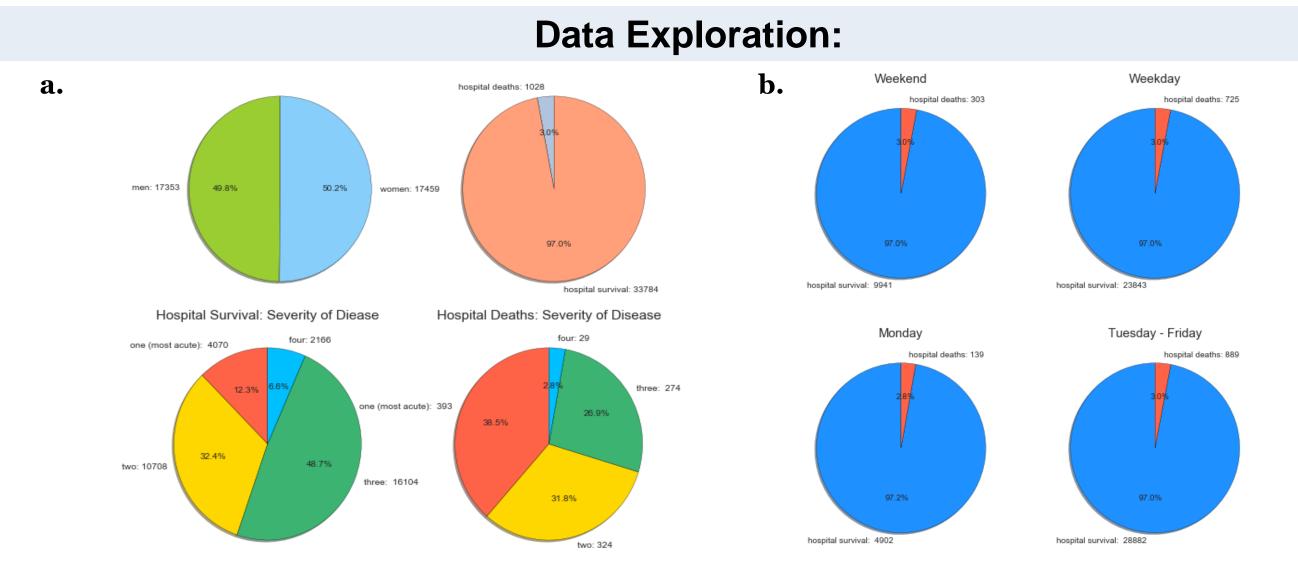


Figure 1. a. General characteristics of ED dataset. b. Proportion of hospital death and survival on day of entry.

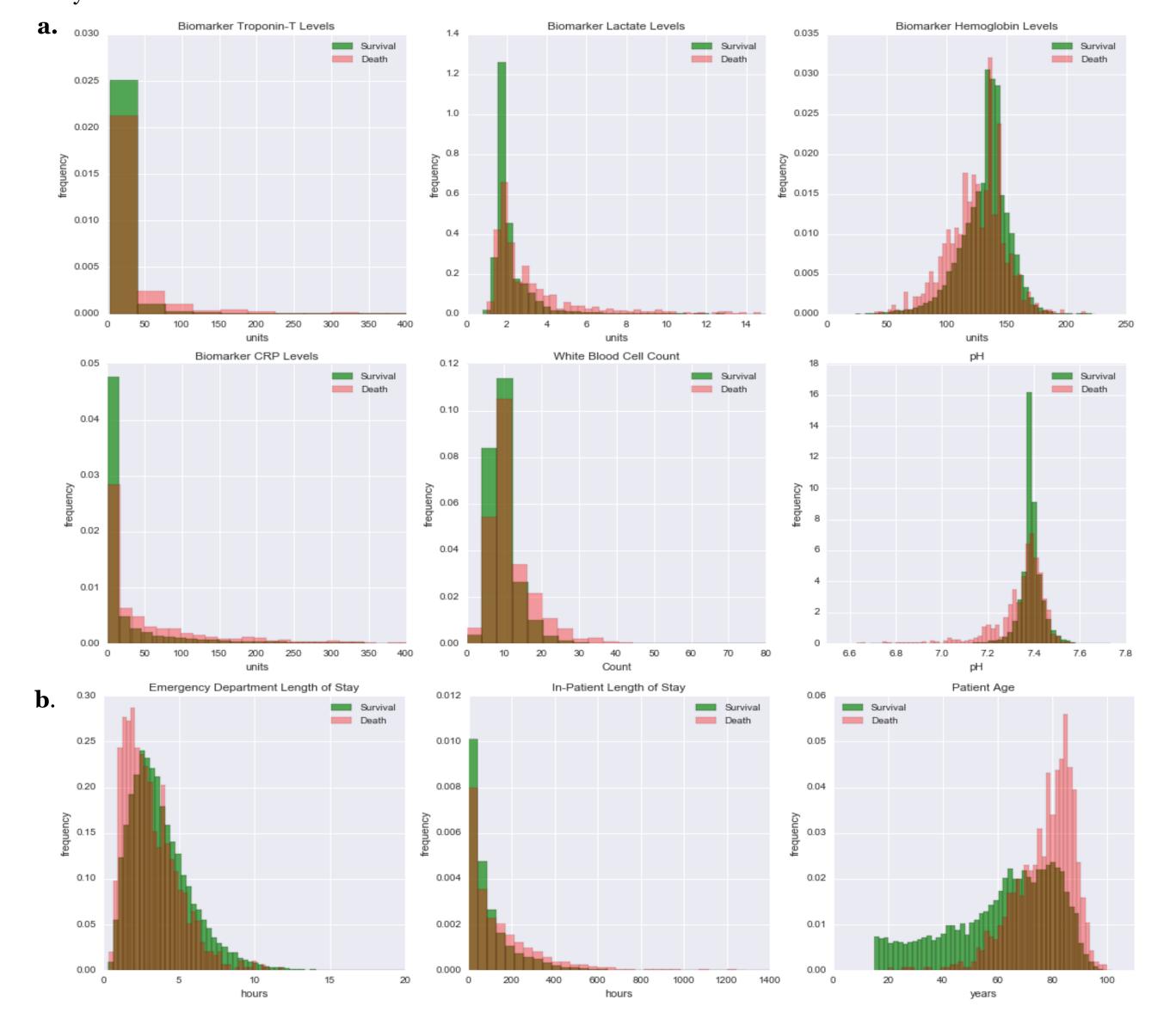


Figure 2. a. Distribution of biomarker levels, white blood cell count, pH among patients. b. Distribution of ED length of stay, in-patient length of stay, and patient age. Survival and death refer to patients who died during hospital stay and patients who did not.

## **Model Building:**

To investigate the importance of time variables in predicting hospital death, weighted logistic regression without using time variables, random forest model with time variables, and random forest with only time variables were optimized through hyper-parameter tuning and crossvalidation.

### Weighted Logistic Regression:

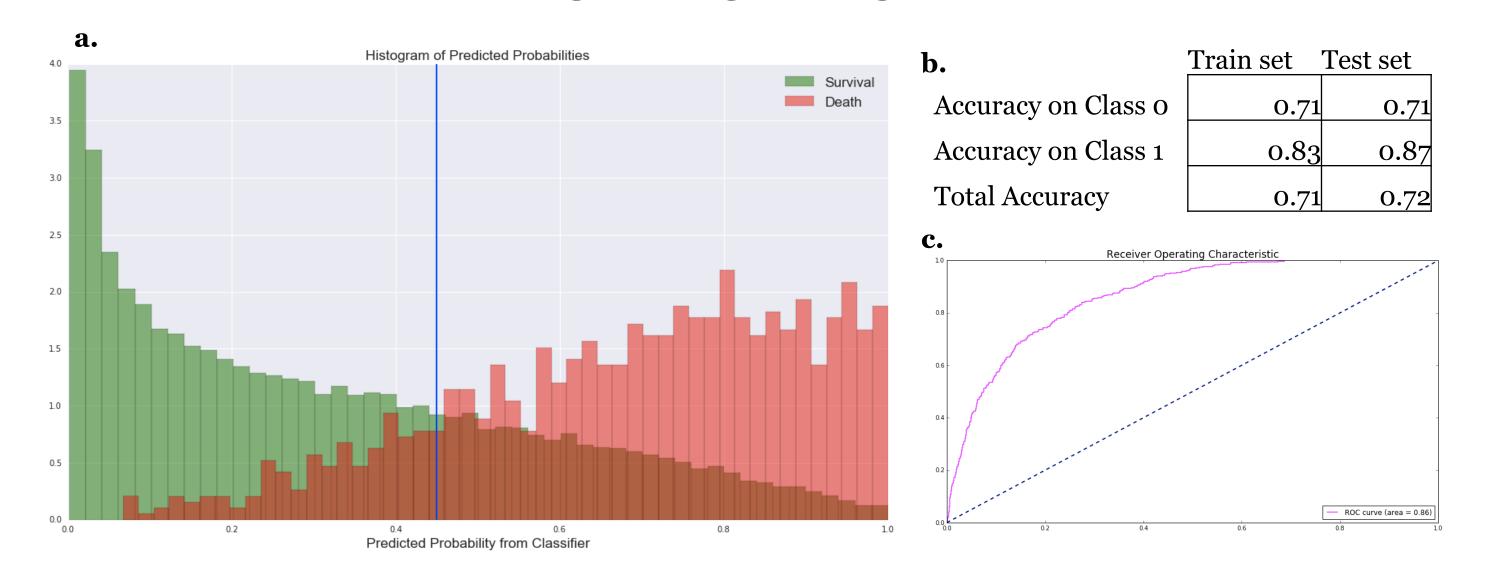
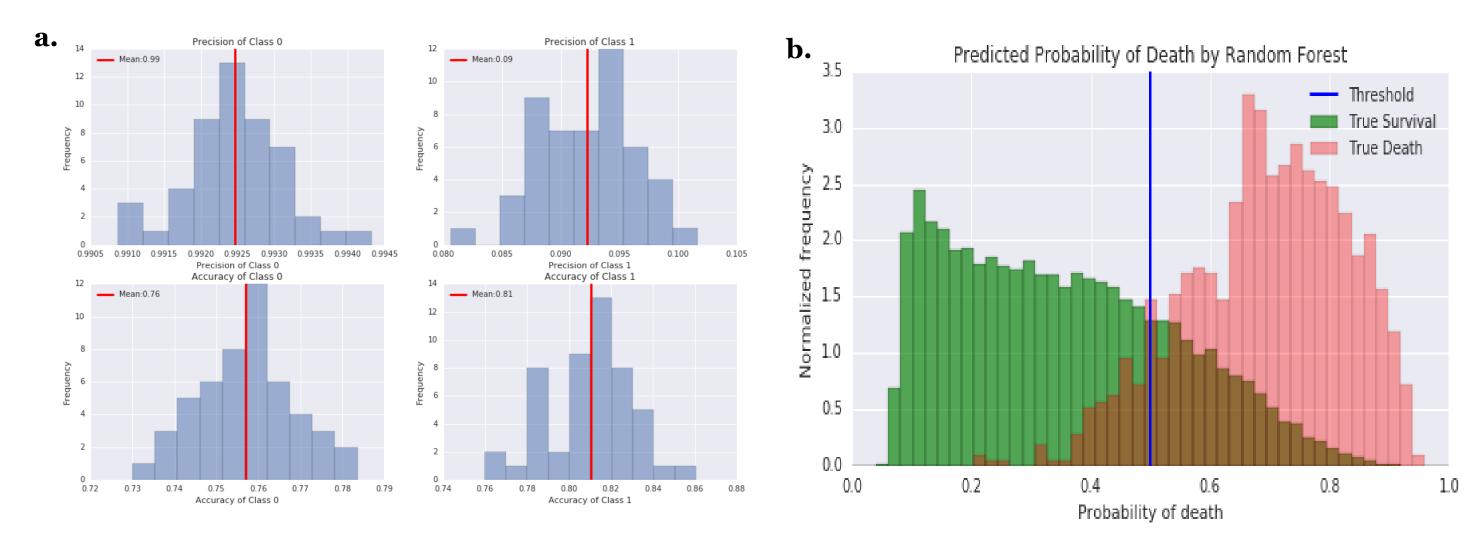


Figure 3: Logistic regression without time variables was created to predict hospital death. a. Predicted probability of death for both true survival and death with a 0.45 classification threshold. **b.** Prediction accuracy on each class and overall accuracy for both train and test set. c. Receiver Operating Characteristic (ROC) metric to evaluate logistic regression output quality.

### **Random Forest:**



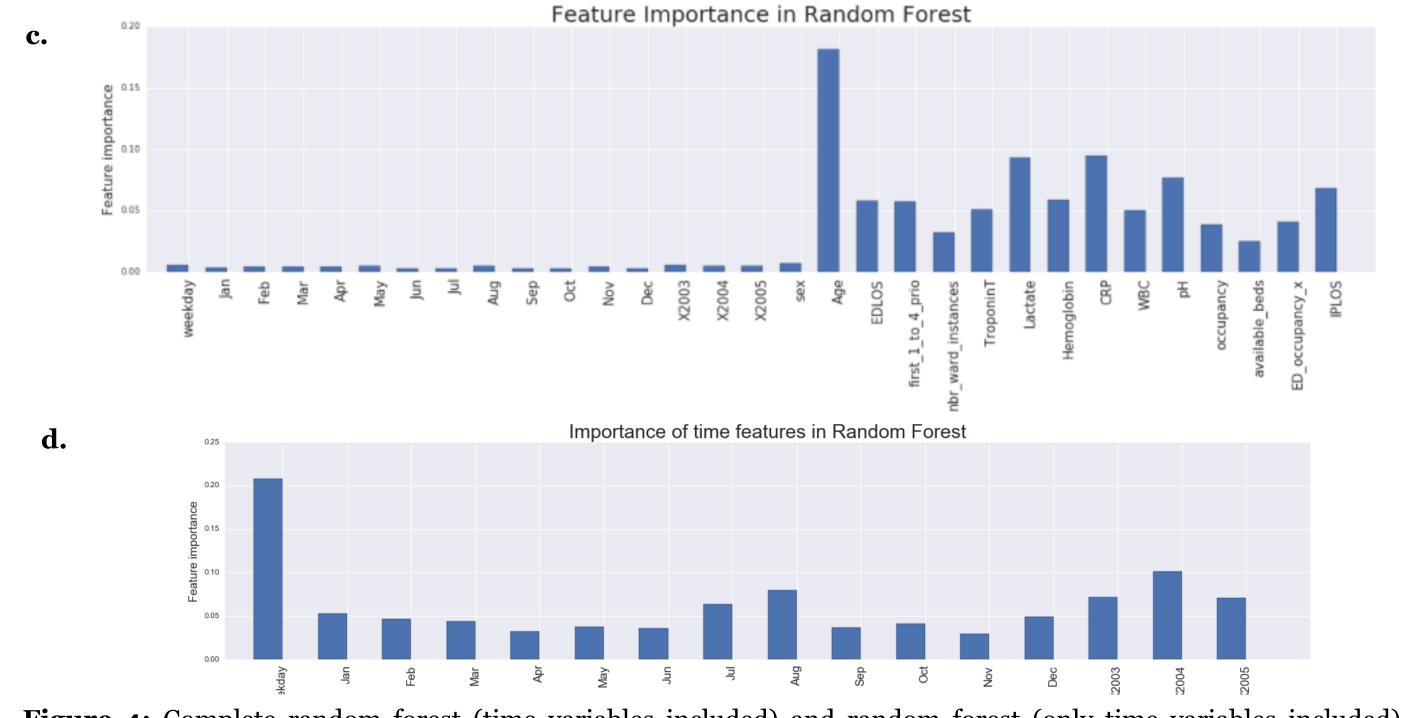
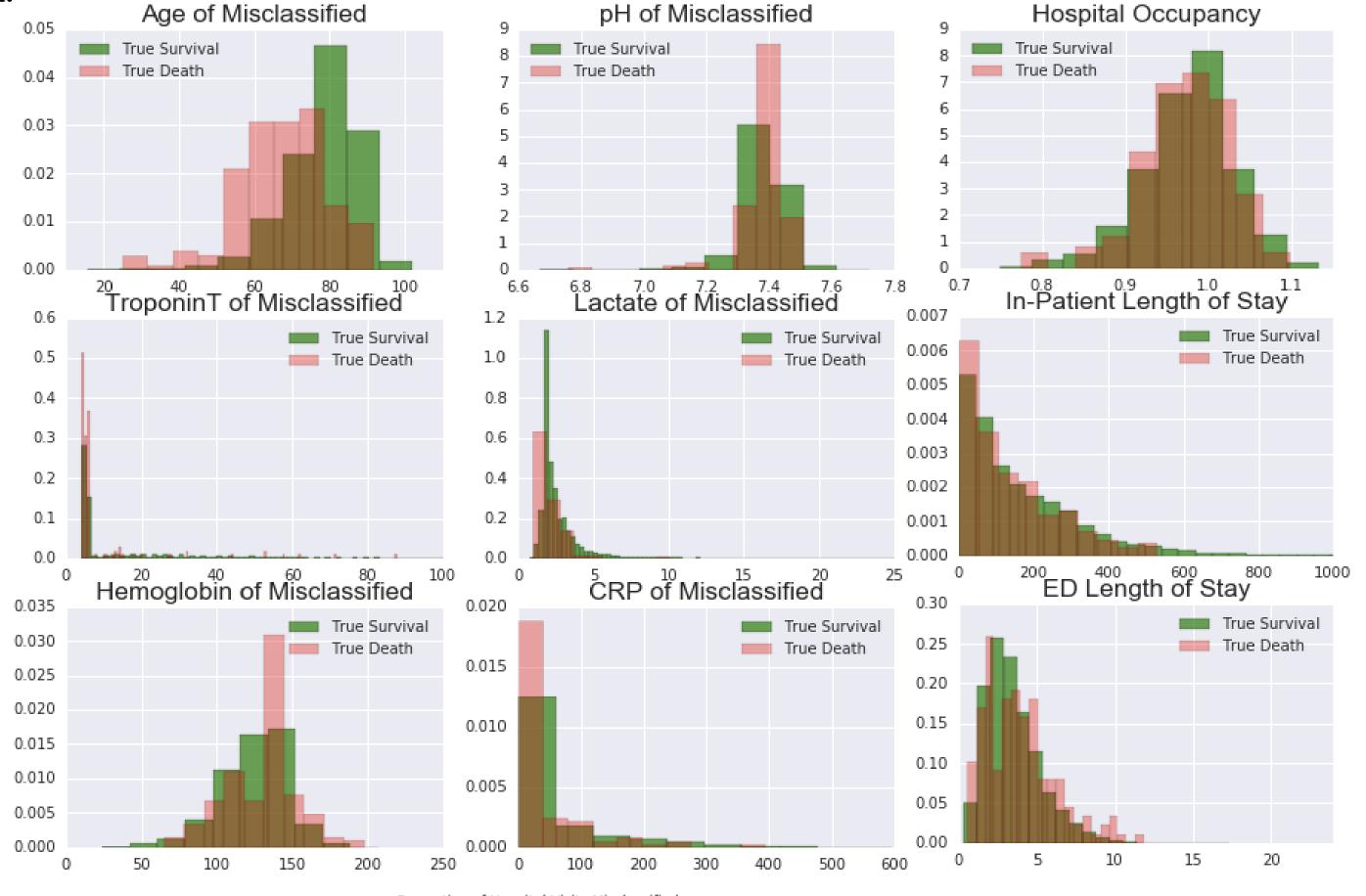
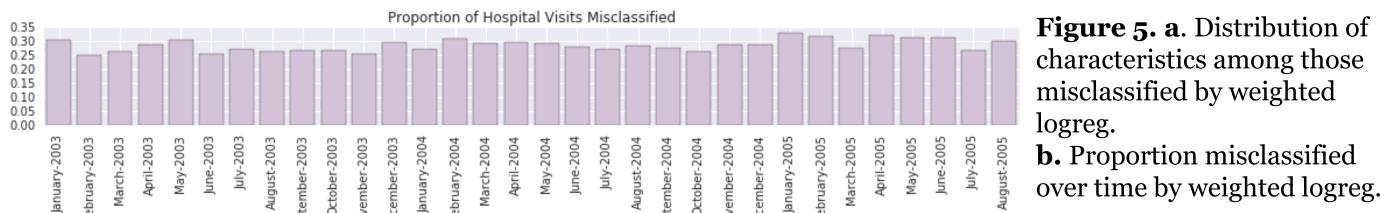
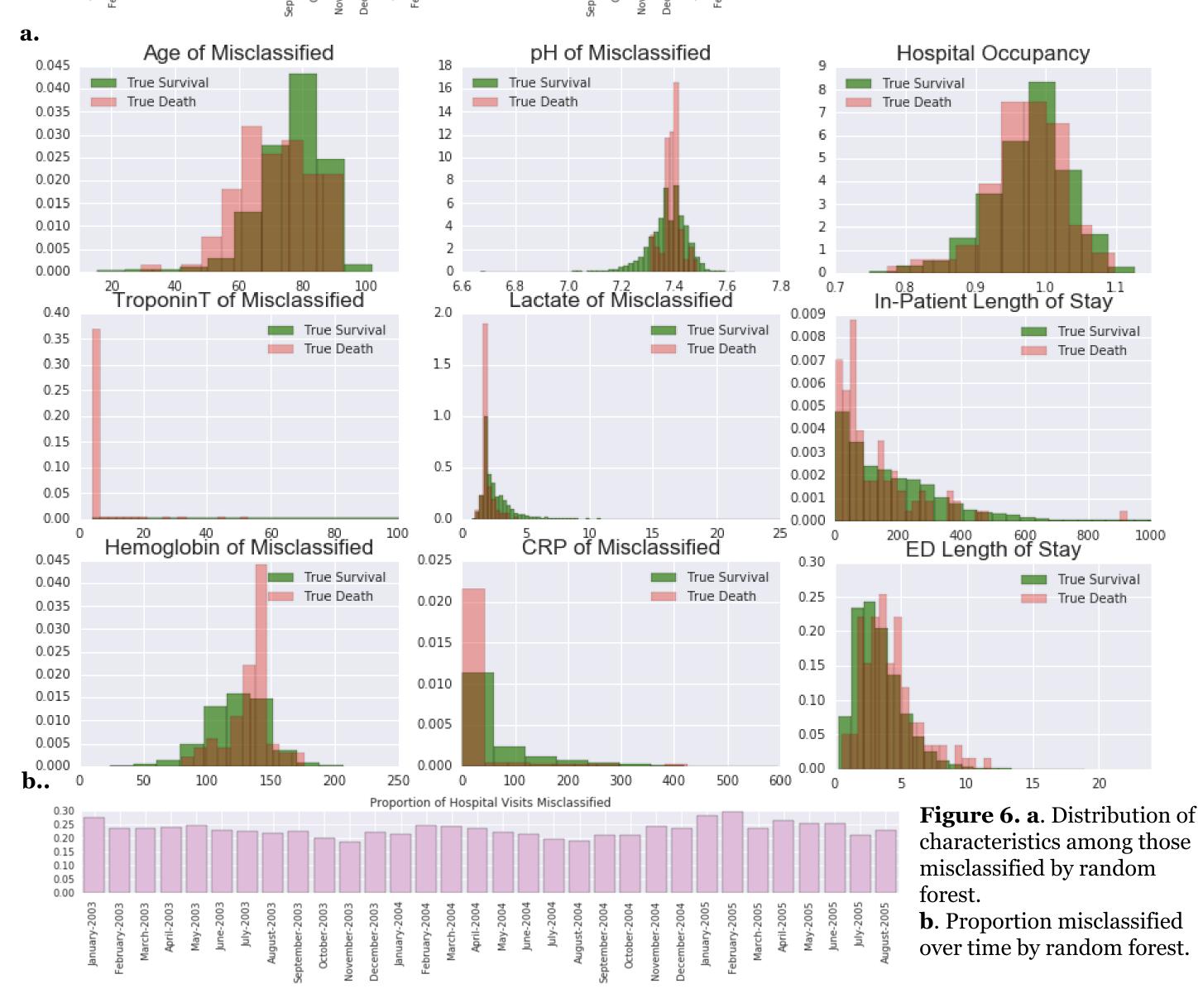


Figure 4: Complete random forest (time variables included) and random forest (only time variables included) were optimized to predict hospital death. a. The distribution of precision and accuracy from complete random forest on both survival and death classes on the test set from simulations. b. Predicted probability of death for both survival and death classes with a 0.5 classification threshold based on the complete model. c. Importance of features in complete random forest with 31 predictors. d. Importance of features in random forest with only time-associated predictors.

# **Misclassification Error Analysis:**







### **Conclusions and Future Directions:**

In the process of building a classifier to predict hospital death based on hospital lab results, time of entry, or both lab results and time of entry, we deal with the tradeoffs in making type I or type II errors when our outcome is rare. Further investigation could improve on accurately predicting such a rare outcome and quantifying our error tradeoff, perhaps by considering realistic cost and resource constraints of a hospital.

Our best classifier trained with the inclusion of time variables, the random forest classifier, did not perform dramatically better than the best classifier built excluding the time variables. Analysis of the misclassified points by each classifier shows similarities and differences in the distribution of characteristics segregated by true outcome. An interesting further investigation would be to attempt matching each hospital death to a similar patient who survived, and investigating whether patterns in time variables exist in such cases.

Despite the fact that people in the US as well as overseas increasingly rely on unscheduled healthcare services [6-8], the availability of such has been reported to decrease, with discontinuing of Emergency Department services cited as an important driver [9-10]. The rate by which this happens is not likely to be constant across geographical areas, potentially aggravating inequalities that are already prevalent in a multitude of countries. The increasing availability of methods for predicting binary outcomes in high dimensional data, along with abundant data substrates, calls for future initiatives aiming at monitoring healthcare quality in real-time, to provide patients with up to date benchmarks of performance across providers and facilities.

o – Pines JM, McCarthy ML. Executive summary: interventions to improve quality in the crowded emergency department. Academic Emergency Medicine. 2011;18(12):1229-33.