This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2859581, IEEE Journal of Biomedical and Health Informatics

An Analytical Framework for TJR Readmission Prediction and Cost-Effective Intervention

Hyo Kyung Lee, *Student Member, IEEE*, Rebecca Jin, Yuan Feng, Philip A. Bain, Jo Goffinet, Christine Baker and Jingshan Li, *Fellow, IEEE*

***Abstract*—This paper introduces an analytical framework for assessing the cost-effectiveness of intervention strategies to reduce total joint replacement (TJR) readmissions. In such a framework, a machine learning-based readmission risk prediction model is developed to predict an individual TJR patient's risk of hospital readmission within 90 days post-discharge. Specifically, through data sampling and boosting techniques, we overcome the class imbalance problem by iteratively building an ensemble of models. Then, utilizing the results of the predictive model, and by taking into account the imbalanced misclassification costs between readmitted and non-readmitted patients, a cost analysis framework is introduced to support decision making in selecting cost-effective intervention policies. Finally, using this framework, a case study at a community hospital is presented to demonstrate the applicability of the analysis.**

***Index Terms*—Total joint replacement (TJR), readmission, risk prediction, intervention, machine learning, cost-effectiveness, class imbalance, cost imbalance.**

I. INTRODUCTION

Total joint replacement (TJR) is a surgical procedure to remove an arthritic or damaged joint and replace with a prosthesis (a metal, plastic or ceramic device) designed to replicate the movement of a normal and healthy joint [1], where hip and knee are the most common joints for replacement. TJR procedures are performed for more than a million cases each year [2], and the numbers are projected to increase significantly with growing demand for improved mobility, greater expectation in quality of life, and increased rates of diagnosis and treatment of advanced arthritis [3].

Although most replacement procedures are successful, unplanned hospital readmissions (episode where a patient who had been discharged from a hospital is admitted again within a specified time interval) may occur due to infection, joint-specific problems (such as dislocation, misalignment), sequelae (e.g., postoperative pain), etc. [4]. With the recent addition of total joint replacement to the hospital readmissions reduction program in 2012, the Centers for Medicare and Medicaid Services (CMS) started reducing payments to hospitals with excessive TJR readmissions. Moreover, hospital leaders recognize that restrictions and financial penalties on readmissions are expected to increase as healthcare systems attempt to transit from volume-based care to value-based care. As such, the economic burden of readmission cost and penalty has been driving many investigators to devote substantial efforts to reducing TJR readmissions.

Most models created to date focus primarily on identifying the risk factors that are associated with readmissions after TJR operations [5]–[7]. Based on the identified risk factors, a few studies go one step further to develop predictive/classification models that can predict whether a patient will be readmitted or not, where logistic regression is generally used as the classification algorithm. Although logistic regression may be a standard approach in many classification problems, it may lead to misleading results due to cost imbalance and class imbalance that commonly exist in readmission problems. Here, the imbalanced misclassification cost implies misclassifying a readmitted patient as a non-readmitted case has a much higher misclassification cost than misclassifying a non-readmitted patient as a readmitted one. In addition, the readmitted portion is typically small among all TJR patients, which leads to class imbalance in the dataset. Therefore, a complete picture of the classification performance needs to be considered to accurately evaluate the performance of predictive models.

Another limiting aspect of existing studies on readmissions is that most prevailing literature only focuses on finding the predictors and risk factors without linking the findings to actionable guidelines of intervention plans. However, if we consider the fundamental objective of predictive modeling, what is the purpose of identifying high-risk patients? The answer to the question is quite obvious: to provide patient-specific followup interventions to the identified high-risk patients in order to prevent them from being readmitted. Ideally, readmission risk prediction models should provide clinically relevant stratification of risk and present information to trigger a transitional care intervention [8]. However, while there exist a handful of perceived to be effective intervention strategies to choose from, there are no widely accepted readmission risk based models for estimating the cost effectiveness of intervention options. As intervention options can be costly, the cost-effectiveness of each intervention method should be analyzed carefully in order to most effectively use the limited resources available for preventing readmissions. Thus, methods that can analyze the cost-effectiveness of intervention procedures are in great need. Eventually, such methods can potentially aid in devising personalized intervention plans.

This work was supported in part by NSF Grant CMMI-1536987.

H.K. Lee, R. Jin, Y. Feng and J. Li are with Department of Industrial and Systems Engineering, University of Wisconsin, Madison, WI 53706, USA. Y. Feng is also with Department of Automation, Tsinghua University, Beijing, China. hlee555@wisc.edu, rcjin@wisc.edu, fengyuan1216@gmail.com, jingshan.li@wisc.edu.

P.A. Bain is with SSM Health Dean Medical Group, 1821 S. Stoughton Rd, Madison, WI 53716, USA. philip.bain@ssmhealth.com

J. Goffinet and C. Baker are with SSM Health St. Mary's Hospital, 700 South Park Street, Madison, WI 53715, USA. Jo.Goffinet@ssmhealth.com, christine.baker@ssmhealth.com

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2859581, IEEE Journal of Biomedical and Health Informatics

In this paper, an analytical framework for machine learning-based TJR readmission prediction and cost-effectiveness intervention is introduced. By extending a preliminary work [9] where a logistic regression predictive model was developed for 105 patient records, the main contribution of this paper is presented in two forms. First, machine learning algorithms and techniques to alleviate the class imbalance problem are utilized to overcome the limitation in traditional readmission predictive modeling approaches and thus improve predictability. Secondly, rather than just improving the predictive model, we utilize the prediction results to evaluate the cost of providing interventions. Guidelines in choosing the appropriate intervention strategy are presented using the total cost of providing interventions and the resulting readmission penalty. That being said, we provide an intervention cost-effectiveness framework to overcome cost imbalance and utilize the predictive model results to compare the total costs of various intervention strategies. The unequal misclassification costs and class imbalance are handled in a unified framework, and guidelines to choose the most suitable intervention are provided.

The remainder of the paper is structured as follows: Section II introduces the classifier algorithm used to develop the readmission predictive model. Based on the results from the predictive model, intervention cost-effectiveness analysis is carried out in Section III and sensitivity study on the analysis results is presented in Section IV. A detailed illustration of the proposed framework is provided via a case study in Section V and final conclusions are formulated in Section VI. All proofs are given in the Appendix.

II. READMISSION PREDICTION

Many studies have been devoted to identifying risk factors for TJR readmissions. For example, it is summarized in [6] that the independent predictors include black race, male sex, discharge to inpatient rehabilitation, increased duration of hospital stay, unilateral replacement, decreased age, and decreased distance between home and hospital. In [10], age, male gender, cancer history, elevated BUN (blood urea nitrogen), a bleeding disorder, and high ASA (American Society of Anesthesiologists) class are predictors for knee replacement, while obesity, steroid use, a bleeding disorder, dependent functional status, and high ASA class are for hip arthroplasty. It is found in [7] that surgical causes, mostly infection, arthrofibrosis, and cellulitis are the primary causes for readmission, and procedure type, long hospital stay, discharge destination, and a fluid/electrolyte abnormality are all associated with readmissions. Moreover, age and Charlson comorbidity index are also identified to have more effects on the probability of adverse outcomes following the total knee arthoplasty (TKA) [5]. As such, there is a demonstrable heterogeneity on the identified risk factors among studies which makes it difficult to generalize the findings.

Using the identified risk factors, logistic regression has been a prevailing tool to predict a patient's readmission probability where the accuracy or area under the receiver operating curve (AUROC) is used to evaluate model performance [9], [11]. Logistic regression seeks to minimize the error rate (i.e., the percentage of the incorrect prediction of class labels) by implicitly assuming all misclassification errors have equal cost. Thus, such a technique may result in misleading results due to cost imbalance prevalent in readmission predictive modeling. Indeed, it is concluded in many existing studies that most prediction models created to date have poor predictive ability [12]–[14]. It is suggested in [15] that although AUROC is a commonly used general metric of model performance, it should be supplemented by task-specific measures that depend on different misclassification costs for false positives and false negatives. Among various classification algorithms, it is identified in [16] and [17] that the models with high AUROC turn out to have low positive predictive value, which implies the use of a single measure such as AUROC that is insensitive to misclassification costs can erroneously eliminate important clinical risk predictors for consideration.

Moreover, readmission datasets are typically unbalanced by nature, that is, examples of one class greatly outnumber examples of the other class [18]. Since TJR readmission prevalence is typically low (usually around 3-6% [6], [10]), unless controlling the class ratios when sampling, it is highly likely that majority of the collected dataset will come from the non-readmitted cases. It is argued in [19] that for such imbalanced datasets, traditional data mining algorithms tend to favor classifying examples as belonging to the majority class, so that they would be ineffective in identifying the minority class. In fact, it has been observed that class imbalance and cost imbalance are closely related [20]. Paper [21] compares various class imbalance techniques when dealing with data that have both an imbalanced class distribution and unequal error costs and observes that methods designed for class imbalance problem can help in cost-sensitive learning methods. Thus it is feasible to handle unequal misclassification costs and class imbalance in a unified framework.

Therefore, in this study, techniques to handle class imbalance problem are used to develop the readmission predictive model. As class imbalance is a common problem in data mining, a large number of techniques have been devoted to addressing the issue [22]–[24] where there are largely two types of approaches; data level approach and algorithm level approach. Data level approach resamples the data distribution in order to decrease the effect of the skewed class distribution while algorithm level approach creates or modifies the learning algorithms that exist. For TJR readmissions, representative methods for each approach (random under sampling (RUS) and synthetic minority oversampling technique (SMOTE) for data level and cost-sensitive learning for algorithm level) on various machine learning algorithms have been tested, such as logistic regression (LG), support vector machine (SVM), C 4.5 decision tree (DT) and random forest (RT). However, both approaches cannot achieve an acceptable recall rate (all below 70%, see Appendix A for a detailed performance comparison). Thus, another machine learning technique, ensemble methods which train several different weak learners and combine their decisions to output a single class label [22], is sought. Although directly applying ensemble learning algorithms does not solve the class imbalance problem, their combination with other techniques (such as sampling or cost-sensitive methods)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2859581, IEEE Journal of Biomedical and Health Informatics

have led to positive results in the literature. Among various ensemble methods, random undersampling boost (RUSBoost) [19] is selected as it has been recognized as one of the top performing methods for imbalanced datasets [22], [25]. Specifically, we modify the data distribution through data sampling and use boosting technique to alleviate class imbalance by iteratively building an ensemble of models.

Regarding the base learner algorithm for RUSBoost, decision tree is selected due to two reasons. First, RUSBoost utilizes boosting methods which consist of iteratively learning weak classifiers and adding them to a final strong classifier [26]. While there exist various learning mechanisms, a common feature of boosting algorithms is that the base classifiers should be weak learners; this is why the most commonly used base classifiers are tree induction algorithms [22]. Although more complex/stronger learners can be considered as the base classifier, using a more robust learner may hinder the diversity of the ensembles, thus possibly reducing the performance [27]. Second reason for selecting decision tree is due to the tradeoff between accuracy and interpretability [28]. Complex machine learning methods may increase overall prediction accuracy, but these models may be overly complicated, thus limiting the interpretability or validity of the results [15]. Compared to other machine learning methods, decision tree has the advantage of interpretability by humans, and has similar performance compared with more complex algorithms in dealing with relatively small and medium datasets. Such a feature is critical in healthcare domain since understanding and validating the model are crucial.

III. COST-EFFECTIVE INTERVENTION

A. Cost Evaluation

While there is a large body of research on readmission predictive modeling, studies addressing the post-discharge intervention options are rather scarce. Currently there are no widely accepted models of TJR for estimating the cost effectiveness of intervention options. Paper [29] asserts the needs for a model that can identify appropriate interventions for specific groups of patients instead of merely predicting the likelihood of readmissions. However, to the best of our knowledge, there is no prevailing TJR literature which can utilize predictive modeling when planning intervention procedures. Therefore, to link predictive modeling and interventions, we directly utilize the predictive model results to provide actionable guidelines on interventions for high-risk patients.

Specifically, confusion matrix (or error matrix) generated from the readmission predictive model is used, which contains information about actual and predicted classifications (see an example for a two-class classifier in Table I, where α_L, β_L, α_H, β_H are the number of patients, and subscripts H and L refer to high and low predicted readmission risks, respectively). Such complete information allows more detailed analysis than calculating the mere proportion of correct classifications, i.e., accuracy. In addition, corresponding to a confusion matrix, a cost matrix can be generated to provide the costs associated with the outcomes in the confusion matrix. Combining the two matrices, the total cost can be evaluated [30], [31].

TABLE I
CONFUSION MATRIX

		True	
		Non-readmit	Readmit
Predicted	Non-readmit	α_L	β_L
	Readmit	α_H	β_H

In this study, to link the predictive model results to actionable guidelines, we specify the total cost in terms of a particular intervention strategy. To do this, we define total cost as the sum of all the costs of providing interventions to the predicted patients and the penalties incurred by the readmitted patients. For each intervention strategy, introduce the following notation to define the associated costs.

R = Readmission penalty per person,
I = Cost of providing followup intervention per person,
p = Success probability, i.e., probability of preventing readmission, in followup interventions.

Remark 1: Note that the actual way CMS calculates readmission penalty is not according to per readmitted case. Rather, it is based on the hospital's excess readmission ratio, which is a measure of the hospital's readmission performance compared to the national average for the hospital's set of patients. Such an excess readmission ratio links directly to the hospital's readmission penalty, the greater the rate of excess readmissions, the higher the penalty. Thus, defining readmission penalty per person is not accurate in terms of how CMS calculates the penalty. Here we consider such a per person definition only for cost analysis purpose. It can be obtained by dividing the total readmission penalty for a given period by the total number of readmissions occurred in the period. Thus, it represents how much penalty a readmitted case is responsible of. Throughout this paper, we assume a hospital has an estimated readmission penalty of $\$909,135$ for 55 readmissions in a given year. Thus, the readmission penalty per person (R) is estimated to be $\$909,135/55 = \$16,530$.

The total cost for hospital readmission is largely generated from two sources: 1) cost of providing followup interventions to the patients, and 2) cost of readmission penalty. Here, interventions are provided to the high-risk patients (i.e., the patients predicted to be readmitted), of whom, some patients are successfully prevented from getting readmitted (assuming with probability p) while the remaining ones are readmitted (with probability $1-p$). Now, if a hospital plans to intervene according to the classifier results, interventions would be provided to the predicted readmitted patients, $\alpha_H + \beta_H$. Of those, α_H patients are truly non-risk patients without readmissions, thus interventions are regardless to them. On the other hand, β_H patients are truly high-risk patients and depending on intervention success probability, some of them will be readmitted while others will not. Also, there exist β_L patients that are actually high-risk patients but the predicting classifier fails to detect. These patients will not receive intervention, thus all of them will be readmitted.

Let C_{interv}. $C_{readmit}$, and C_{total} denote the intervention cost, readmission cost, and total cost, respectively. Then, the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2859581, IEEE Journal of Biomedical and Health Informatics

intervention cost is applied to $\alpha_H + \beta_H$ patients, i.e.,

$$C_{interv} = (\alpha_H + \beta_H) \cdot I. \tag{1}$$

The readmission cost covers β_L truly high-risk readmitted patients who are not identified by the classifier and a proportion $(1-p)$ of β_H patients whose interventions are not effective.

$$C_{readmit} = [\beta_L + \beta_H \cdot (1-p)] \cdot R. \tag{2}$$

Finally, the total cost includes both C_{interv} and $C_{readmit}$.

$$C_{total} = (\alpha_H + \beta_H) \cdot I + [\beta_L + \beta_H \cdot (1-p)] \cdot R. \tag{3}$$

B. Cost-Effectiveness Analysis

As previously mentioned, there are a handful of perceived to be effective intervention options to choose from. Intervention procedures for reducing readmissions have been thoroughly studied and implemented at many hospitals. Such procedures may include scheduling additional followup visits, conducting additional phone calls, etc. Then among the various options, which one should the hospital implement? That is, if a hospital plans to choose among various intervention approaches, how to evaluate whether an intervention is worth implementing or not? To help hospital managers in making such decisions, we provide an analytical framework to determine the cost-effectiveness of an intervention policy.

First, let B be the total readmission budget. Depending on the objective of analysis, B can be defined in various ways. If the objective is to compare with current expenditures, for hospitals that are currently implementing some form of interventions, B can represent the sum of current readmission penalty and the cost of providing the intervention. If a hospital has no intervention in progress, B can merely be the current readmission penalty. Furthermore, if a hospital is interested in decreasing current expenditure, B can be viewed as an upper limit on the budget the hospital is willing to spend on readmissions. In either case, the total cost incurred by implementing an intervention should be less than the readmission budget B. Thus, for a given intervention with cost I and success rate p, the following constraint holds:

$$(\alpha_H + \beta_H) \cdot I + [\beta_L + \beta_H \cdot (1-p)] \cdot R \leq B. \tag{4}$$

Since B is a budget constraint imposed by the hospital, it should be regarded as a fixed constant. In addition, as I represents the intervention cost and p is the probability of success, we need to have

$$I \geq 0, \qquad 0 \leq p \leq 1. \tag{5}$$

Then, a linear programming method can be conducted to find the feasible regions of I and p, that is, finding the pair of I and p values under which the given intervention strategy is cost-effective. Formally, the problem can be analytically formulated as finding the feasible region determined by the three constraints in (4) and (5). Using the feasible regions determined by these constraints, we can arrive the following:

Proposition 1: The following arguments hold:

(1) Given intervention cost I and budget B, the minimum success probability p_{min} for cost-effective intervention is:

- When $B \geq R(\beta_L + \beta_H)$

$$p_{min} = \begin{cases} 0, & \text{if } I < \frac{B-(\beta_L+\beta_H)R}{\alpha_H+\beta_H}, \\ \frac{(\alpha_H+\beta_H)I+(\beta_L+\beta_H)R-B}{\beta_H R}, & \text{if } I \geq \frac{B-(\beta_L+\beta_H)R}{\alpha_H+\beta_H}. \end{cases} \tag{6}$$

- When $\beta_L R < B < R(\beta_L + \beta_H)$

$$p_{min} = \frac{(\alpha_H + \beta_H)I + (\beta_L + \beta_H)R - B}{\beta_H R}. \tag{7}$$

(2) Given success probability p and budget B, the maximal allowed intervention cost I_{max} is:

$$I_{max} = \frac{B - R(\beta_L + \beta_H) + \beta_H R p}{\alpha_H + \beta_H}. \tag{8}$$

Proof: See Appendix B. ■

Remark 2: Note that it is impossible to achieve a total cost lower than B if the given budget is less than $\beta_L R$, i.e., no feasible solution exists. This implies that even with a highly effective intervention that incurs minimal cost, it is impossible to satisfy the given budget constraint.

A graphical illustration of the feasible regions is shown in Figure 1. Here we consider the intervention budget relative to the readmission penalty ($\frac{B}{R}$), that is, compared to the readmission penalty, how much budget is allowed?

Given that the minimal budget condition is guaranteed ($B > \beta_L R$), depending on whether the budget is restricted or not, the feasible regions vary. The shaded areas in Figure 1 correspond to the feasible regions, i.e., the combinations of horizontal axis (I) and vertical axis (p) of an intervention satisfying constraint $I(\alpha_H + \beta_H) + R[\beta_L + \beta_H(1-p)] \leq B$ are cost-effective.

In the left-side figure, when the given budget is large ($\frac{B}{R} \geq \beta_L + \beta_H$), for every low cost intervention (i.e., $I < \frac{B-(\beta_L+\beta_H)R}{\alpha_H+\beta_H}$), it is always cost-effective to implement such an intervention regardless of its success probability. Whereas in the right-side figure for a smaller budget ($\beta_L < \frac{B}{R} < \beta_L + \beta_H$), such a "free" region does not exist. In both cases, as an intervention becomes more costly, the minimum success probability needs to be increased (i.e., intervention needs to be more effective) in order to be considered cost-effective, which coincides with intuition. If an intervention costs more than $\frac{B-\beta_L R}{\alpha_H+\beta_H}$, regardless of how effective the intervention is, not implementing this intervention is always more cost-effective. Such a graphical representation could directly provide answers to the hospital about minimal success probability needed to achieve and maximal allowed intervention cost, where closed form solutions of both are provided in Proposition 1.

C. Stratification of Risk Group

While providing interventions to all high-risk patients can guide hospitals in identifying target patient groups, in some cases, it may be too costly to give interventions to all the identified patients. Thus, to further narrow down the target groups, we can segment patients into different risk groups. Cutoff scores for each risk group can either be determined by the healthcare professionals or be generated from the distribution of the risk scores. Now, with the introduction of risk groups, we need to solve a multi-label classification problem

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2859581, IEEE Journal of Biomedical and Health Informatics

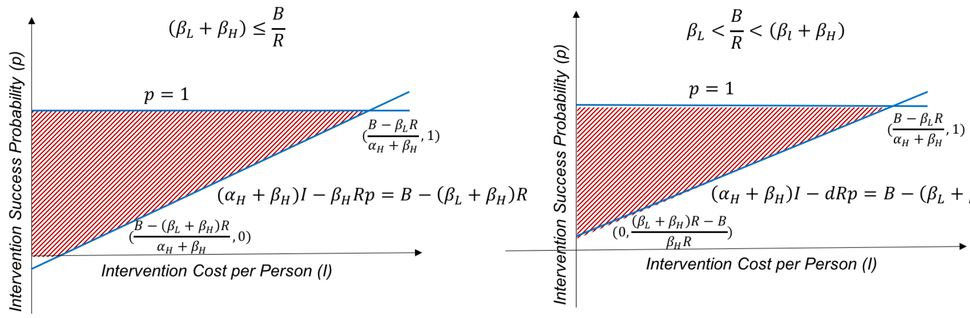

Fig. 1. Feasible regions

rather than the previous binary classification problem. That is, instead of predicting whether a patient will be readmitted or not, we need to classify among multiple class labels. For the TJR risk group definition, we are looking at a 3-class classification problem with the confusion matrix generated as Table II, where the patients are segmented into high-, medium- and low-risk (represented by 'H', 'M', and 'L', respectively) groups. Clearly this can also be generalized to any n-class classification problem.

TABLE II
CONFUSION MATRIX FOR 3-CLASS RISK GROUP CLASSIFICATION

		True	
		Non-readmit	Readmit
Predicted	Low Risk	α_L	β_L
	Medium Risk	α_M	β_M
	High Risk	α_H	β_H

As the purpose of risk stratification is in identifying the population group the hospital should focus intervention resources on, varying degree of interventions should be carried out depending on the patient's risk group. In accordance with intuition, the most intense level of interventions should be carried out to the high-risk group, and the least or possibly no interventions to the low-risk group. Accordingly, cost and success probability of interventions should vary among different risk groups. Thus, we introduce I_r and p_r, $r \in \{H, M, L\}$, to denote the cost and success probability of patients in risk group r, respectively. Note that the readmission penalty cost per person (R) is the same across all risk groups since such a penalty does not take into account a patient's readmission risk score. Then, using the information provided in Table II, the total cost can be calculated as follows:

$$C_{total} = \sum_{r \in \{H,M,L\}} [(\alpha_r + \beta_r)I_r + \beta_r(1 - p_r)R]. \quad (9)$$

Analogous to (4) and (5), now the constraints become

$$\sum_{r \in \{H,M,L\}} [(\alpha_r + \beta_r)I_r + \beta_r(1 - p_r)R] \leq B, \quad (10)$$

$$I_r \geq 0, \qquad 0 \leq p_r \leq 1. \quad (11)$$

For a given $r \in \{H, M, L\}$, let X_r represent the cost for patients in other risk groups, i.e.,

$$X_r = \sum_{s \in \{H,M,L\}, s \neq r} \{(\alpha_s + \beta_s)I_s + \beta_s(1 - p_s)R\}. \quad (12)$$

Then the minimal success probability and maximal allowed budget can be calculated as follows:

Proposition 2: The following arguments hold:

(1) Given intervention costs I_r and budget B, the minimum success probability $p_{r_{min}}$ to make an intervention cost-effective is:
- When $B \geq X_r + R\beta_r$,

$$p_{r_{min}} = \begin{cases} 0, & \text{if } I < \frac{B - X_r - R\beta_r}{\alpha_r + \beta_r}, \\ \frac{(\alpha_r+\beta_r)I_r + X_r + R\beta_r - B}{R\beta_r}, & \text{if } I \geq \frac{B - X_r - R\beta_r}{\alpha_r + \beta_r}. \end{cases} \quad (13)$$

- When $X_r < B < X_r + R\beta_r$,

$$p_{r_{min}} = \frac{(\alpha_r + \beta_r)I_r + X_r + R\beta_r - B}{R\beta_r}. \quad (14)$$

(2) Given success probabilities p_r and budget B, the maximal allowed intervention cost $I_{r_{max}}$ is:

$$I_{r_{max}} = \frac{B - X_r - R(1 - p_r)\beta_r}{\alpha_r + \beta_r}. \quad (15)$$

Proof: See Appendix B. ■

IV. SENSITIVITY ANALYSIS

The proposed analytical framework provides a quantitative supporting tool for hospital management to evaluate the impact of intervention policies and make necessary decisions. However, since the cost is evaluated by combining confusion matrix and cost matrix, the cost-effectiveness analysis results heavily depend on the confusion matrix values. Therefore, a sensitivity analysis on the confusion matrix needs to be carried out to check the robustness of the analysis results.

In addition, the sensitivity analysis results can be used as a guideline for hospitals in selecting appropriate predictive model. There exist numerous studies on readmission risk predictive models with various classifier algorithms. Different models have varying degree of accuracies, and even when the overall accuracy may be the same, the exact values in the confusion matrix may vary. For a binary classification problem, accuracy of the model depends on how many patients can be correctly classified into their corresponding classes, either classifying a truly non-readmitted patient into a non-readmitted class or classifying a truly readmitted patient into a readmitted class. Generally there is a tradeoff between the two objectives, thus we need to decide which class to focus

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2859581, IEEE Journal of Biomedical and Health Informatics

on based on their relative importance. Therefore, we seek to compare the impact of two improvement directions under different budget constraints.

The two objectives are formally characterized in Table III. An increase in the correctly classified case is represented by $\Delta\alpha$ for non-readmitted case and $\Delta\beta$ for readmitted case. Since increasing $\Delta\alpha$ is focusing on correctly identifying non-readmitted patients, this objective can be considered as giving interventions to less people to avoid unnecessary intervention costs. On the other hand, increasing $\Delta\beta$ aims to correctly identify readmitted patients, which is equivalent to providing interventions to more people in order to avoid readmission penalty. Note that the actual numbers of non-readmitted patients and readmitted patients need to remain the same regardless of the classifier algorithm being used. Thus, the increased number of correctly classified cases needs to be subtracted from the misclassified cases.

TABLE III
CONFUSION MATRIX FOR SENSITIVITY ANALYSIS

		True	
		Non-readmit	Readmit
Predicted	Non-readmit	$\alpha_L + \Delta\alpha$	$\beta_L - \Delta\beta$
	Readmit	$\alpha_H - \Delta\alpha$	$\beta_H + \Delta\beta$

First, we seek to find monotonicity properties of $\Delta\alpha$ and $\Delta\beta$ with respect to the total cost. In other words, will a more accurate predictive model reduce the total cost, where the total cost for an improved accuracy model can be formulated as:

$$\widetilde{C}_{total} = (\alpha_H - \Delta\alpha + \beta_H + \Delta\beta)I + [\beta_L - \Delta\beta + (\beta_H + \Delta\beta)(1-p)]R. \quad (16)$$

By evaluating the partial derivative of total cost with respect to $\Delta\alpha$ and $\Delta\beta$, we obtain

Proposition 3: The improved total cost $\widetilde{C}_{total}$ is monotonically decreasing with respect to the number of reduced mis-predicted high-risk patients $\Delta\alpha$. In addition, $\widetilde{C}_{total}$ is monotonically decreasing with respect to the number of increased correctly classified high-risk patients $\Delta\beta$ if $I < pR$.

Proof: See Appendix B. ■

As one can see, $\Delta\alpha$ always brings reduction in total cost, while $\Delta\beta$ is beneficial only when $I < pR$. If the intervention cost is too high relative to its success probability, increasing $\Delta\beta$ results in giving costly interventions to too many patients with minimal reduction in readmissions.

Now, for interventions that satisfy $I < pR$, since both $\Delta\alpha$ and $\Delta\beta$ reduce total cost, we compare their impacts through cost-effectiveness analysis. Depending on the given budget ($\frac{B}{R}$) and the increased number of correctly classified cases, the varying impacts of $\Delta\alpha$ and $\Delta\beta$ are shown in Figure 2. For comparison purpose, the original curve is also included, shown as the solid line. In addition, $\Delta\alpha$ and $\Delta\beta$ need to be increased by the same amount, which is denoted as n, where $n < \min(\beta_L, \alpha_H)$.

As illustrated in Figure 2, when the given budget is large (Cases 1 and 2), the readmission penalty plays a less significant role in decision making. Since providing interventions is more of a prevention to reduce readmission penalty, preventing readmission has less significant cost benefit under a less restricted budget. Instead, it is more cost-effective to provide interventions to less people, which corresponds to increasing $\Delta\alpha$. But there exists an exception in Case 1 when the model has an extremely low positive predictive value (of the patients predicted as readmitted cases, how many are truly readmitted). Here the given conditions can be rearranged to $(\beta_L + \beta_H) \leq \frac{B}{R} < \frac{2\beta_L + \alpha_H + \beta_H - n}{2}$, thus $(\beta_L + \beta_H) < \frac{2\beta_L + \alpha_H + \beta_H - n}{2}$ must be true in order for the condition to be satisfied. This is equivalent to $n < \alpha_H - \beta_H$. Since n is an increased number of cases, $\alpha_H > \beta_H$ is an implicit condition that should hold as well. Thus, when the model has a low positive predictive value (too many false-positives), efforts should be focused on increasing $\Delta\beta$ to increase the positive predictive value. Therefore, under this condition, increasing $\Delta\beta$ is a more cost-effective way compared to increasing $\Delta\alpha$.

On the other hand, if the budget is very limited (Cases 3 and 4), readmission penalty becomes a significant burden, thus it is cost-effective to prevent as many readmitted patients as possible. However, as intervention cost increases, the cost of providing intervention grows significantly as well. Thus for Case 4 where interventions are given to too many people (due to increased number of identifications of readmitted cases n), intervention cost outweighs the benefit brought from the reduction of readmission penalty. Hence, increasing $\Delta\alpha$ becomes a better option for interventions with large cost. Therefore, depending on the given budget and how much improvement is considered, efforts should be focused on correctly identifying either the non-readmitted or the readmitted patients.

V. CASE STUDY

In this section, we demonstrate the applicability of the proposed cost-effectiveness analysis framework through a case study at St. Mary's Hospital (SMH), which is a 440-bed not-for-profit hospital in Madison, WI, offering a full range of inpatient and outpatient treatment and diagnostic services in primary care and nearly all specialties. Among the endeavors for quality improvement, reducing TJR patients' readmission rate has received substantial attention.

A. Setting and Variable Selection

A total of 525 patient records are studied. Among them, 477 are not readmitted within 90 days, while 48 are readmitted. The dataset consists of 11% of readmitted cases and 89% of non-readmitted cases, thus being highly imbalanced.

Based on extensive literature review, consultation with physicians, surgeons, nurses, and other healthcare professionals, and by analyzing patient medical records, a total of 33 risk factors are identified in a preliminary study [9]:

(1) Demographic: 8 variables.
- Age.
- Gender (male/female).
- With patient support at home (Y/N).
- BMI (body mass index, which is a measure of body fat) > 40 (Y/N).
- Insurance (none/public/private).
- Employment status.

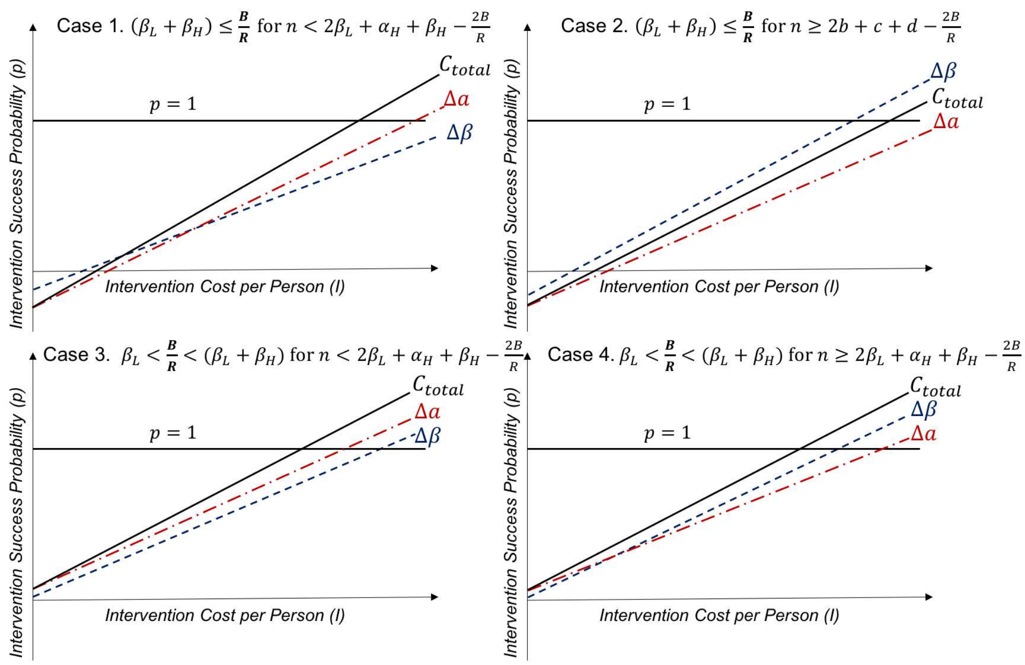

Fig. 2. Sensitivity analysis

- Median income (zip code).
- Active use of MyChart (a secure patient portal).

(2) Preoperative (Preop): 14 variables.
- Preoperative hemoglobin.
- ASA class, which refers to the physical status classification system for assessing the fitness of patients before surgery, adopted by ASA.
- Depression.
- CCI (Charlson comorbidity index), predicting one-year mortality for a patient who may have a range of comorbid conditions.
- Hypertension.
- COPD (chronic obstructive pulmonary disease).
- OSA (obstructive sleep apnea)/asthma.
- CAD (coronary artery disease).
- Arthritis.
- Diabetes.
- Smoking (never/former/current).
- Preoperative patient reported pain score.
- Alcohol usage (never /normal: ≤ 5 drinks/day /heavy: > 5 drinks/day).
- Preoperative fall risk score (No: 0–2/Yes: ≥ 3).

(3) Operative: 4 variables
- Type of joint (hip/knee).
- Type of anesthesia (regional/general).
- Total operative time.
- Number of joints (1: unilateral; 2: bilateral under one anesthetic).

(4) Postoperative: 5 variables.
- Postoperative complications.
- Number of positive CAM/Delirium assessment score, where CAM stands for confusion assessment method, a standardized tool to identify and recognize delirium.
- Maximum postoperative patient reported pain score.
- Average postoperative patient reported pain score.
- Last postoperative patient reported pain score.

(5) Discharge: 2 variables: Length of stay; and Discharge location (home /home with services /skilled nursing facility (SNF)).

B. Predictive Model

To deal with a highly imbalanced dataset, RUSBoost which is considered as one of the top performing algorithms for imbalanced data, is selected. Patient data for the 33 variables defined in Subsection V-A are used as inputs to the model. Given these predictors, the record of patient i can be presented by a tuple (x_i, y_i), where x_i's correspond to the set of risk factors $(x_i^1, \ldots, x_i^{33})$ and y_i indicates the class label, i.e., whether patient i belongs to a readmitted case or a non-readmitted case. Thus, the development of the predictive model becomes a supervised learning task in which the objective is to infer a function that associates the relationship between the risk factors and the class label by learning from the set of labeled patient records. Specifically, we assume the relationship can be represented by a decision tree classifier and RUSBoost algorithm is adopted to infer the structure of such a tree. The main procedure of RUSBoost is outlined below.

Algorithm RUSBoost

Given: $S = (x_1, y_1), \ldots, (x_n, y_n)$: Set S of n patient records, N: Desired percentage of total instances to be represented by the minority class (readmitted).

Step 0. Initialization: Initialize the weights of each patient example $W_1(i) = \frac{1}{n}$ for all i. Set iteration number $t = 1$.

Step 1.1. Random Undersampling: Remove randomly from the majority class (class 0: non-readmitted) examples to con-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2859581, IEEE Journal of Biomedical and Health Informatics

struct a new training dataset S'_t where $N\%$ belongs to the minority class.

Step 1.2. Training: Using dataset S'_t, train the decision tree h_t which predicts the class label given the risk predictors.

Step 1.3. Update: Calculate the weight update parameter α_t using the pseudo-loss ϵ_t, which is calculated based on the original training dataset S and current weights W_t. Using α_t, update W_{t+1} for the next iteration.

Step 2. Repeat: Set $t = t+1$ and continue Step 1 until $t \leq T$, i.e., T iterations have been conducted.

Step 3. Weighted Vote: Take the weighted vote of the decision trees h_t for $t = 1, \ldots, T$, to obtain the final decision tree ensemble H.

Remark 3: Note here we only provide a simplified version of RUSBoost algorithm to illustrate how the algorithm is formulated. The detailed formulas for calculating the pseudo-loss ϵ's, weight update parameter α's and weighted vote for the final decision tree ensemble H are omitted as they are not the focus of this paper. The complete description and detailed steps of RUSBoost algorithm can be found in [19].

An ensemble of decision trees is developed as the risk prediction model using the illustrated RUSBoost algorithm. Then, to validate the developed model, a 10-fold cross-validation is used to estimate the prediction accuracy. Specifically, we train each method 10 times, each time on 90% of the dataset, withholding a different 10% for the final evaluation.

Using the developed predictive model, the confusion matrix in this study is generated and presented in Table IV.

TABLE IV
CONFUSION MATRIX: RUSBOOST

		True	
		Non-readmit	Readmit
Predicted	Non-readmit	420	11
	Readmit	57	37

From Table IV, it is shown that the overall prediction accuracy δ of the model is

$$\delta = \frac{\text{Correctly predicted number}}{\text{Total number of patients}} = \frac{420+37}{(420+37)+(57+11)} = 87.05\%.$$

Moreover, as there exists significant cost imbalance where the readmission cost is much higher, correctly identifying the patient who will be readmitted is more critical. Thus, we introduce recall rate $\widetilde{\delta}$ as such a measurement,

$$\widetilde{\delta} = \frac{\text{Correctly predicted number of readmissions}}{\text{Total number of readmissions}} = \frac{37}{37+11} = 77.08\%.$$

To illustrate the advantage of the model compared to the traditional predictive modeling approach, a logistic regression model is also developed. The resulting confusion matrix is shown in Table V. As one can see, logistic regression can obtain a higher overall prediction accuracy ($\frac{470+16}{525} = 92.57\%$) but the recall rate is significantly lower ($\frac{16}{32+16} = 33.33\%$). This implies the model can only identify one third of truly high-risk patients and the remaining two thirds of high-risk patients will be misclassified as non-readmission patients. Thus the resulting readmission penalty would be significantly higher than that from the proposed RUSBoost prediction model. Therefore, by considering accuracies in both overall prediction and recall rate, the proposed model returns a more reliable result. In fact, for all possible values of intervention success probability and intervention cost, RUSBoost always gives a lower total cost than that of logistic regression (see Figure 3), which shows the importance of considering imbalance.

TABLE V
CONFUSION MATRIX: LOGISTIC REGRESSION

		True	
		Non-readmit	Readmit
Predicted	Non-readmit	470	32
	Readmit	7	16

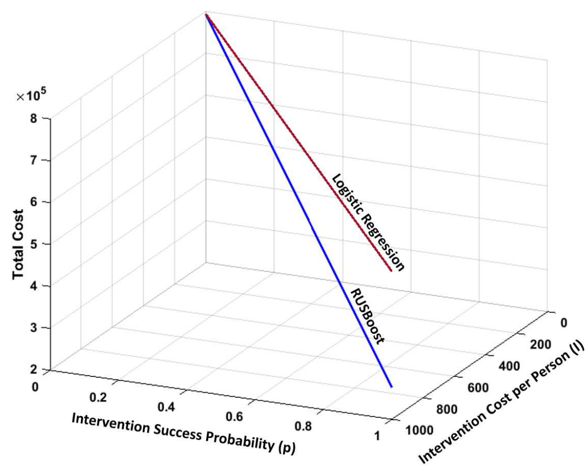

Fig. 3. Cost comparison between logistic regression and RUSBoost methods

C. Cost-Effectiveness Analysis

Using the confusion matrix in Table IV and setting $R = 16,530$, equation (3) can be evaluated as

$$C_{total} = (57+37)\cdot I - [11 + (37(1-p)]\cdot 16,530 = 94I - 611,610p + 793,440.$$

Then constraints (4) and (5) form a feasible region. From Proposition 1, the minimal success probability p_{min} for a given budget B and intervention cost I can be evaluated as

- When $B \geq 793,440$,

$$p_{min} = \begin{cases} 0, & \text{if } I < \frac{B-793,440}{94}, \\ \frac{(94I+793,440-B}{611,610}, & \text{if } I \geq \frac{B-793,440}{94}. \end{cases}$$

- When $181,830 < B < 793,440$

$$p_{min} = \frac{94I + 793,440 - B}{94}.$$

In addition, the maximal allowance on intervention cost I_{max} can be calculated as a function of budget B and success probability p,

$$I_{max} = \frac{B - 793,440 + 611,610p}{94}.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2859581, IEEE Journal of Biomedical and Health Informatics

To provide further guidance on interventions, risk stratification is conducted as well. After consulting with the healthcare professionals at SMH, dividing patients into 3 groups (high/medium/low) is considered practical, and the proportions of each group are determined upon discussion. The top 10 percentile are grouped as high-risk group, the next top 25 percentile as medium-risk group and the remaining 65 percentile as low-risk group. The performance of the predictive model under such risk group stratification is presented in Table VI. From (12), we obtain

TABLE VI
CONFUSION MATRIX FOR RISK GROUP CLASSIFICATION

		True	
		Non-readmit	Readmit
Predicted	Low-risk	340	1
	Medium-risk	114	17
	High-risk	23	30

$$X_r = \begin{cases} 131I_M + 53I_H - 281,010p_M \\ \qquad -495,900p_H + 776,910, & r = L, \\ 341I_L + 53I_H - 16,530p_L \\ \qquad -495,900p_H + 512,430, & r = M, \\ 341I_L + 131I_M - 16,530p_L \\ \qquad -281,010p_M + 297,540, & r = H. \end{cases}$$

Using Proposition 2, for given budget B and intervention cost I_r, the minimum success probability $p_{r_{min}}$ can be evaluated. When success probabilities p_r and budget B are known, the maximal allowed intervention cost $I_{r_{max}}$ can be calculated.

D. Intervention Process

Although the minimum success probability $p_{r_{min}}$ and maximal allowed intervention cost $I_{r_{max}}$ can provide guidance in selecting cost-effective interventions, no specific intervention strategies are available yet at SMH to compare the cost-effectiveness. Thus, to seek opportunities for possible intervention strategies, a detailed analysis of the current post-surgery care process at SMH is carried out.

When a TJR patient is discharged, disposition option is determined based on the patient's progress during hospitalization, available support at home, and medical complications. Currently, there are three options for disposition: stay at a SNF; stay at home but receive home care visits provided by professional services (referred to as home service); and stay at home without any additional services (self care). At the time of discharge, although it may not be a quantifiable number, most healthcare professionals have their own estimate on each patient's readmission risk level. Based on this estimate, and considering other environmental factors as well, they suggest an appropriate disposition option to the patient.

Each discharge option has varying degrees of care support where SNF is at the most intense level of care facility and self care is the least. In general, the types of care services provided to the patients are similar, but the additional support provided by the care facility makes the difference. Patients mainly receive three types of care processes. First, the orthopedic care coordinator nurse calls the patient within 24 hours after discharge to check the recovery status. If the nurse feels that the patient's condition may not be stable, successive phone calls are conducted during the following days. In addition to the phone calls, patients also need to visit the orthopedic clinic on a regular basis. All TJR patients are scheduled to visit the orthopedic physician in 2 weeks, 1 month, 6 weeks, 3 months, and 1 year post discharge. Some patients may also be required to receive physical therapy (PT) exercises on a weekly basis, and PT service is received in different forms depending on where the patient is discharged to. SNF patients typically receive rehab therapy at the built-in PT facility within the nursing care facility, home service patients receive PT by home care professionals (therapist, nurse) during their visits to patient homes, and self care patients generally visit an external PT center for exercises. In addition to the three main care services, home service patients receive additional nurse visits on a regular basis and SNF patients receive additional care from the healthcare professionals at the facility. Thus, even though the general care process is similar, the degree of care provided to the patient varies depending on the discharged location.

According to SMH's current discharge policy, the ideal discharge options for high-, medium- and low-risk patients are SNF, home service and self care, respectively. However, except for self care, not all patients are able to be discharged to their ideal option due to eligibility issues (e.g., insurance). Thus, only a certain proportion of patients are able to be discharged to their ideal locations, and the remaining non-eligible patients are discharged to a lower intensity care facility, which may increase their risks in readmission. Now, to avoid unnecessary increase in readmission risks, hospitals can incentivize patients not by directly paying the cost of staying at care facilities (SNF or home service), but through agreements with their affiliated care facilities to make them not turn patients away. In this sense, hospitals can advise high-risk patients who may be non-eligible to be admitted to external nursing facilities to stay at affiliated SNFs, and high- and medium-risk patients who are not eligible for home service to receive services from the affiliated home service providers. Thus, increasing eligibility proportions can be viewed as an intervention strategy.

In this sense, there can be three patient groups the hospital can focus their incentives on. The high-risk patients who are not eligible to be admitted to SNF, the high-risk patients who are even not eligible to receive home services, and the medium-risk patients who are not eligible for home service. Note that, patients will not be discharged to a higher level of care facility than their risk group level. That is, medium-risk patients will not be discharged to SNF and low-risk patients will not be discharged to SNF or home service. Upon discussion with healthcare professions, they agree that providing unnecessary intensive care to the patients not only is a waste of healthcare resource but also may lead to unexpected adverse effects. The discharge options for patient risk groups and the associated intervention options are illustrated in Figure 4, where h_{SNF}, h_{HS}, and m_{HS} refer to the proportion of high-risk patients who are eligible for SNF, the proportion of high-risk patients who are not eligible for SNF but eligible for home services, and the proportion of medium-risk patients who are eligible for home services, respectively. Increasing h_{SNF},

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2859581, IEEE Journal of Biomedical and Health Informatics

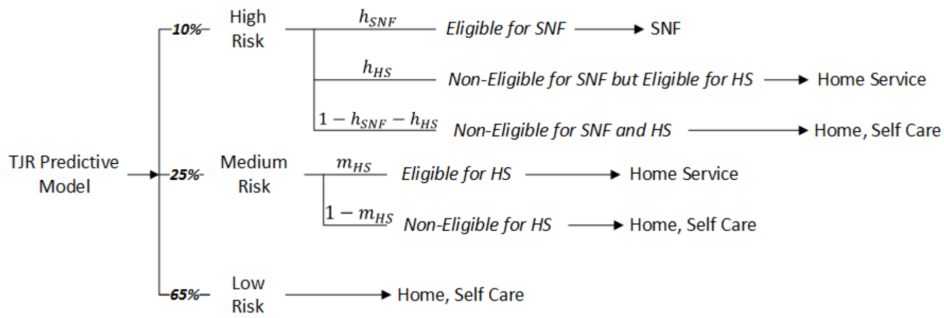

Fig. 4. Decision of discharge disposition option

h_{HS}, and m_{HS} can each be considered as an intervention strategy, denoted as H_{SNF}, H_{HS}, and M_{HS}, respectively.

E. Intervention Cost-Effectiveness Analysis

To quantify the relationship between discharge location and patient risk group, the readmission probability of each disposition option depending on patient risk level is identified and shown in Table VII. These numbers are determined upon consultation with clinical managers and nurses at SMH. Analogous to TKA, THA refers to total hip arthoplasty. Typically TKA has a higher readmission rate than THA.

TABLE VII
DISCHARGE DISPOSITION IMPACT ON READMISSION PROBABILITY

Location	SNF	Home service	Self care
High-risk	TKA: 4.32% THA: 4%	Increase to 10%	Increase to 20%
Medium-risk	-	TKA: 3.7% THA: 2.8%	Increase to 5.5%
Low-risk	-	-	TKA: 2.4% THA: 1.8%

Using Table VII, the intervention success probability can be derived for each intervention option, H_{SNF}, H_{HS}, and M_{HS}. Such success probability (p) can be viewed as, of current total number of readmitted cases, how many cases can be prevented if certain intervention strategy were implemented.

$$
\begin{aligned}
p &= \frac{\text{Reduced number of readmissions}}{\text{Current number of readmissions}} \\
&= \frac{\text{Reduced readmission probability}}{\text{Current readmission probability}}
\end{aligned}
$$

Note that since H_{SNF} and H_{HS} only provide interventions to high-risk patients, intervention probabilities for medium- and low-risk patients remain unaffected. Similarly, M_{HS} does not affect high- and low-risk patients' success probabilities. For illustration purpose, only the success probabilities for TKA patients are presented. The same approach can be applied to THA patients as well. Specifically, denote p_i^k as the success probability for risk i group patients under intervention option k. Then, for each intervention option H_{SNF}, H_{HS} and M_{HS},

$$
\begin{aligned}
&p_L^{H_{SNF}} = 0, \quad && p_M^{H_{SNF}} = 0, \quad && p_H^{H_{SNF}} = 0.784, \\
&p_L^{H_{HS}} = 0, && p_M^{H_{HS}} = 0, && p_H^{H_{HS}} = 0.5, \\
&p_L^{M_{HS}} = 0, && p_M^{M_{HS}} = 0.327, && p_H^{M_{HS}} = 0.
\end{aligned}
$$

Next, the cost of providing each intervention is determined using the average total cost incurred during a patient's stay at the care location. Specifically, SNF cost is estimated to be $\$30,000$, home service cost is $\$2,000$, and self care at home incurs no cost. As H_{SNF}, H_{HS} and M_{HS} increase eligibility for patient admissions, the average cost of staying at SNF or using home services are used as the intervention cost. Particularly, let I_i^k be the intervention cost for risk i group patients under intervention option k. Then for each intervention option H_{SNF}, H_{HS} and M_{HS}, such costs are:

$$
\begin{aligned}
&I_L^{H_{SNF}} = 0, \quad && I_M^{H_{SNF}} = 0, \quad && I_H^{H_{SNF}} = \$30,000, \\
&I_L^{H_{HS}} = 0, && I_M^{H_{HS}} = 0, && I_H^{H_{HS}} = \$2,000, \\
&I_L^{M_{HS}} = 0, && I_M^{M_{HS}} = \$2,000, && I_H^{M_{HS}} = 0.
\end{aligned}
$$

By associating intervention parameters with confusion matrix, the intervention and readmission cost can be evaluated.

$$
\begin{aligned}
I^s &= (\alpha_L + \beta_L)I_L^s + (\alpha_M + \beta_M)I_M^s + (\alpha_H + \beta_H)I_H^s, \\
R^s &= [\beta_L(1 - p_L^s) + \beta_M(1 - p_M^s) + \beta_H(1 - p_H^s)]R, \\
&\quad s = H_{SNF}, H_{HS}, M_{HS}.
\end{aligned}
$$

Summing up the intervention cost and readmission cost, we can calculate the total cost associated with an intervention strategy. By comparing total cost with intervention budget, the cost-effectiveness of implementing any of the intervention strategies can be evaluated. If multiple strategies are cost-effective, the total costs can be compared, and along with other considerations, the optimal intervention policy can be determined. The results of cost-effectiveness analysis are shown in Table VIII. With the lowest intervention cost and readmission cost, incentivizing the high-risk patients to be eligible for home service is identified as the most cost-effective strategy.

TABLE VIII
TOTAL COSTS OF INTERVENTIONS

	H_{SNF}	H_{HS}	M_{HS}
Intervention Cost	$1,590,000	$106,000	$262,000
Readmission Cost	$404,654	$545,490	$701,550
Total Cost	$1,994,654	$651,490	$963,550

However, this result largely depends on the readmission probabilities based on patient's discharge location (Table VII). The readmission probabilities when patients are discharged to their ideal location are based on the collected data, but the increased probabilities when patients are discharged to a lower level of care facility are based on healthcare professional's estimates. Thus, the robustness of the results needs to be assessed where there are three variables that have uncertainties associated with them; readmission probability of 1) high-risk patients discharged to home service, 2) high-risk patients discharged home, self care, and 3) medium-risk patients discharged home, self care. Therefore, to validate the robustness of cost evaluation, a 3-way sensitivity analysis on the readmission probabilities needs to be carried out. Figure 5 plots the regions of the variables where each intervention strategy is preferred. Incentivizing the high-risk patients to be eligible for SNF (H_{SNF}) costs the most for all cases. Thus increasing the proportion of home service is more favored than increasing that of SNF. Specifically, among the two options H_{HS} and M_{HS}, the cost-effectiveness regions where incentivizing the high-risk patients is the most favored (H_{HS}) lie on the left part of each graph, and the regions where incentivizing the medium-risk patients is the most favored (M_{HS}) are on the right. As the readmission probability of high-risk patients discharged to home increases, incentivizing high-risk patients becomes more valuable, thus the cost-effectiveness region of H_{HS} increases, which is represented by the graph's shifting to the right.

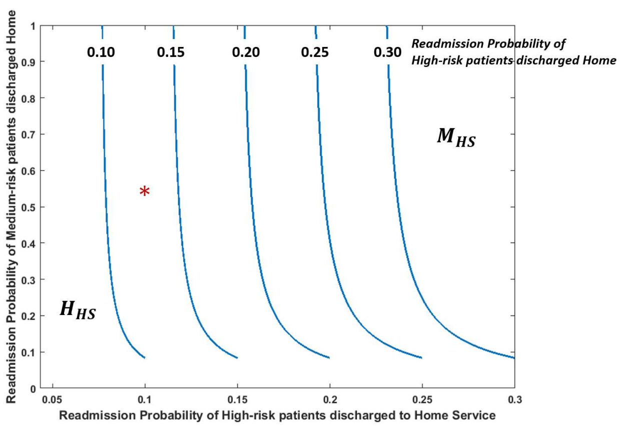

Fig. 5. Three-way sensitivity analysis

When the readmission probability of high-risk patients discharged to home service is low, intervention H_{HS} is always favored regardless of medium-risk patients' readmission probability. However, as the readmission probability of high-risk patients increases, H_{HS} becomes less effective, thus incentivizing medium-risk patients M_{HS} becomes more cost-effective. When the readmission probability of medium-risk patients discharged to home increases, incentivizing medium-risk patients becomes more critical and M_{HS} is favored.

The current readmission probability setting is correspondent to the asterisk (*) point in Figure 5. With the readmission probability of high-risk patients discharged to home being set as 0.20, the asterisk point lies on the left part of the graph, so that the favored intervention is H_{HS}. However, as the readmission probability of high-risk patients discharged home decreases, H_{HS} region shrinks as well, and when the probability is 0.10, H_{HS} is no longer considered cost-effective and M_{HS} is favored instead. Thus, depending on the robustness of readmission probability estimates, the favored intervention can be determined between H_{HS} and M_{HS}.

VI. Conclusions

This paper introduces a new framework to overcome the limitations in traditional predictive modeling approaches and assess the cost-effectiveness of interventions. The main contribution is in linking the predictive model to actionable intervention guidelines. Currently in prevailing TJR literature, predictive modeling is not utilized when planning post-discharge interventions. Therefore, there are many potential applications for such a model as it provides a means of establishing the relative cost effectiveness of various intervention options in TJR post-discharge care, and allows healthcare interventions to be valued and compared. The proposed framework can be a useful guide for hospital managers in planning patient specific intervention strategies, and can be extended to other diseases to evaluate various treatment or intervention procedures, which are of significant importance to support decision making for personalized care [32], [33].

Appendix A: Performance Comparison

TABLE A.1
Performance comparison

(a) Baseline performance of each Classifier

	LG	SVM	DT	RF
Accuracy	92.57%	90.47%	91.04%	91.81%
Recall	33.33%	31.25%	31.25%	10.41%

(b) Data level approach applied to each classifier

	Data+LG	Data+SVM	Data+DT	Data+RF
Accuracy	90.05%	90.57%	88.48%	91.79%
Recall	66.67%	66.67%	69.79%	61.45%

(c) Algorithm level approach applied to each classifier

	Alg+LG	Alg+SVM	Alg+DT	Alg+RF
Accuracy	82.28%	80.19%	86.85%	91.62%
Recall	64.58%	68.75%	52.08%	59.37%

Note that RUSBoost gives accuracy of 87.05% and recall rate of 77.08%. Compared to data level approach and algorithm level approach, RUSBoost significantly improves the recall rate with a slight sacrifice in overall accuracy.

Appendix B: Proofs

Proof of Proposition 1: Constraint (4) can be rewritten as

$$I(\alpha_H + \beta_H) - [B - (\beta_L + \beta_H)R] \le p\beta_H R.$$

Then expressions (6) and (7) follow. Moving the bracket to the right side of the inequality, expression (8) holds. ■

Proof of Proposition 2: When $r = L$, expand (10) as

$$\begin{aligned}&(\alpha_L + \beta_L)I_L + b_L(1-p_L)R + (\alpha_M + \beta_M)I_M \\ &+ \beta_M(1-p_M)R + (\alpha_H + \beta_H)I_H + \beta_H(1-p_H)R \\ &= (\alpha_L + \beta_L)I_L + \beta_L(1-p_L)R + X_L \le B,\end{aligned}$$

which can be rewritten as

$$(\alpha_L + \beta_L)I_L - [B - X_L - R\beta_L] \le p_L\beta_L R.$$

Then analogously to the proof of Proposition 1, all arguments follow. Similar steps can be applied to the cases of $r = M$ and $r = H$. ∎

Proof of Proposition 3: From (16), we have

$$\frac{\partial \widetilde{C}_{total}}{\partial \Delta\alpha} = -I < 0.$$

When $I < pR$, we obtain

$$\frac{\partial \widetilde{C}_{total}}{\partial \Delta\beta} = I - pR < 0.$$

∎

Acknowledgement: The authors thank the help of K. Blanchar, J. Hettenbach, S. Longley, K. Nelson at SMH.

References

[1] American Academy of Orthopaedic Surgeons, *Total joint replacement*, https://orthoinfo.aaos.org/en/treatment/total-joint-replacement/, 2018.

[2] C. Steiner, R. Andrews, M. Barrett, and A. Weiss, "HCUP Projections: Mobility/Orthopedic Procedures 2003 to 2012," *HCUP Projections Report 2012-03*, U.S. Agency for Healthcare Research and Quality, 2012.

[3] S.M. Kurtz, E. Lau, K. Ong, K. Zhao, M. Kelly and K.J. Bozic, "Future young patient demand for primary and revision joint replacement: National projections from 2010 to 2030," *Clinical Orthopedic Related Research*, vol. 467, no. 10, pp. 2606-2612, 2009.

[4] P.N. Ramkumar, C.T. Chu, J.D. Harris, A. Athiviraham, M.A. Harrington, D.L. White, D.H. Berger, A.D. Naik and L.T. Li, "Causes and rates of unplanned readmissions after elective primary total joint arthroplasty: A systematic review and meta-analysis," *American Journal of Orthopedics*, vol. 44, no. 9, pp. 397-405, 2015.

[5] N.F. SooHoo, J.R. Lieberman, C.Y. Ko, and D.S. Zingmond, "Factors predicting complication rates following total knee replacement," *JBJS*, vol. 88, no. 3, pp. 480-485, 2006.

[6] B. Zmistowski, C. Restrepo, J. Hess, D. Adibi, S. Cangoz and J. Parvizi, "Unplanned readmission after total joint arthroplasty: Rates, reasons, and risk factors," *JBJS*, vol. 95, no. 20, pp. 1869-1876, 2013.

[7] W.W. Schairer, T.P. Vail and K.J. Bozic, "What are the rates and causes of hospital readmission after total knee arthroplasty?" *Clinical Orthopaedics and Related Research*, vol. 472, no. 1, pp. 181-187, 2014.

[8] D. Kansagara, H. Englander, A. Salanitro, D. Kagen, C. Theobald, M. Freeman and S. Kripalani, "Risk prediction models for hospital readmission: A systematic review,"*JAMA*,vol. 306, no. 15, pp. 1688-1698, 2011.

[9] H.K. Lee, R. Jin, Y. Feng, P.A. Bain, J. Goffinet, C. Baker and J. Li, "Modeling and analysis of postoperative intervention process for total joint replacement patients using simulations," *Proceedings of IEEE International Conference on Automation Science and Engineering*, pp. 568-573, 2017.

[10] A.J. Pugely, J.J. Callaghan, C.T. Martin, P. Cram and Y. Gao, "Incidence of and risk factors for 30-day readmission following elective primary total joint arthroplasty: Analysis from the ACS-NSQIP," *The Journal of Arthroplasty*, vol. 28, no. 9, pp. 1499-1504, 2013.

[11] A. Hosseinzadeh, M.T. Izadi, A. Verma, D. Precup and D.L. Buckeridge, "Assessing the predictability of hospital readmission using machine learning," *Proceedings of the 25th Innovative Applications of Artificial Intelligence Conference*, pp. 1532-1538, 2013.

[12] J.L. Wolff, B. Starfield and G. Anderson, "Prevalence, expenditures, and complications of multiple chronic conditions in the elderly," *Archives of Internal Medicine*,vol. 162, no. 20, pp. 2269-2276, 2002.

[13] H.H. Pham, D. Schrag, A.S. O'Malley, B. Wu and P.B. Bach, "Care patterns in Medicare and their implications for pay for performance,"*New England Journal of Medicine*,vol. 356, no. 11, pp. 1130-1139, 2007.

[14] A.J.H. Kind, M.A. Smith, J.R. Frytak and M.D. Finch, "Bouncing back: Patterns and predictors of complicated transitions 30 days after hospitalization for acute ischemic stroke,"*Journal of the American Geriatrics Society*, vol. 55, no. 3, pp. 365-373, 2007.

[15] J. Futoma, J. Morris and J. Lucas, "A comparison of models for predicting early hospital readmissions," *Journal of Biomedical Informatics*, vol. 56, pp. 229-238, 2015.

[16] N.R. Cook, "Use and misuse of the receiver operating characteristic curve in risk prediction,"*Circulation*,vol. 115, no. 7, pp. 928-935, 2007.

[17] A.K. Waljee, P.D.R. Higgins and A.G. Singal, "A primer on predictive models," *Clinical and Translational Gastroenterology*, vol. 5, no. 1, pp. e44, 2014.

[18] G.M. Weiss and F. Provost, "Learning when training data are costly: The effect of class distribution on tree induction," *Journal of Artificial Intelligence Research*, vol. 19, pp. 315-354, 2003.

[19] C. Seiffert, T.M. Khoshgoftaar, J. van Hulse and A. Napolitano, "RUSBoost: A hybrid approach to alleviating class imbalance," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 40, no. 1, pp. 185-197.

[20] X.Y. Liu and Z.H. Zhou, "The influence of class imbalance on cost-sensitive learning: An empirical study," *Proceedings of the 6th IEEE International Conference on Data Mining*, pp. 970-974, 2006.

[21] G.M. Weiss, K. McCarthy and B. Zabar, "Cost-sensitive learning vs. sampling: Which is best for handling unbalanced classes with unequal error costs?" *DMIN*, vol. 7, pp. 35-41, 2007.

[22] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince and F. Herrera, "A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches,"*IEEE Transactions on Systems, Man, and Cybernetics, Part C*,vol. 42, no. 4, pp. 463-484, 2012.

[23] M. Ohsaki., P. Wang, K. Matsuda, S. Katagiri, H. Watanabe and A. Ralescu, "Confusion-matrix-based Kernel logistic regression for imbalanced data classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 9, pp. 1806-1819, 2017.

[24] W. Lee, C.H. Jun and J.S. Lee, "Instance categorization by support vector machines to adjust weights in AdaBoost for imbalanced data classification," *Information Sciences*, vol. 381, pp. 92-103, 2017.

[25] T.M. Khoshgoftaar, J. van Hulse and A. Napolitano, "Comparing boosting and bagging techniques with noisy and imbalanced data," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 41, no. 3, pp. 552-568, 2011.

[26] L. Rokach, "Ensemble-based classifiers,"*Artificial Intelligence Review*,vol. 33(1-2), pp. 1-39, 2010.

[27] D.A. Cieslak and N.V. Chawla, "Learning decision trees for unbalanced data," *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 241-256, 2008.

[28] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm and N. Elhadad, "Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission," *Proceedings of the 21th ACM International Conference on Knowledge Discovery and Data Mining*, pp. 1721-1730, 2015.

[29] S. Yu, F. Farooq, A. van Esbroeck, G. Fung, V. Anand and B. Krishnapuram, "Predicting readmission risk with institution-specific prediction models,"*Artificial Intelligence in Medicine*,vol. 65, no. 2, pp. 89-96, 2015.

[30] C. Drummond and R.C. Holte, "Cost curves: An improved method for visualizing classifier performance,"*Machine Learning*,vol. 65, no.1, pp. 95-130, 2006.

[31] C. Drummond and R.C. Holte, "Explicitly representing expected cost: An alternative to ROC representation," *Proceedings of the sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*,pp. 198-207, 2000.

[32] J. Yoon, C. Davtyan and M. van der Schaar, "Discovery and clinical decision support for personalized healthcare," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 4, pp. 1133-1145, 2017.

[33] R. Chen, J. Sun, R.S. Dittus, D. Fabbri, J. Kirby, C.L. Laffer, C.D. McNaughton and B. Malin, "Patient stratification using electronic health records from a chronic disease management program," *IEEE Journal of Biomedical and Health Informatics*, DOI: 10.1109/JBHI.2016.2514264, 2016.