Journal of Healthcare Quality Research 38 (2023) 197–205

Fundación Española de Calidad Asistencial

Journal of Healthcare Quality Research

www.elsevier.es/jhqr

ORIGINAL ARTICLE

Applying machine learning methods to predict the hospital re-admission within 30 days of total hip arthroplasty and hemiarthroplasty

J.-M. Wu[a,b], B.-W. Cheng[b], C.-Y. Ou[b], J.-E. Chiu[b], S.-S. Tsou[a,*]

[a] *Tungs' Taichung MetroHarbor Hospital, Taichung City, Taiwan, ROC*
[b] *Department of Industrial Engineering and Management, National Yunlin University of Science and Technology, Yunlin City, Taiwan, ROC*

Received 12 May 2022; accepted 29 November 2022
Available online 27 December 2022

KEYWORDS
Total hip arthroplasty;
Hemiarthroplasty;
Machine learning;
Logistic regression;
Random forest;
Artificial neural network

Abstract
Background: Total hip arthroplasty (THA) and hemiarthroplasty are common treatments for severe hip joint disease. To predict the probability of re-admission after discharge when patients are hospitalized will support providing appropriate health education and guidance.
Methods: The research aims to use logistic regression (LR), decision trees (DT), random forests (RF), and artificial neural networks (ANN) to establish predictive models and compare their performances on re-admissions within 30 days after THA or hemiarthroplasty. The data of this study includes patient demographics, physiological measurements, disease history, and clinical laboratory test results.
Results: There were 508 and 309 patients in the THA and hemiarthroplasty studies respectively from September 2016 to December 2018. The accuracies of the four models LR, DT, RF, and ANN in the THA experiment are 94.3%, 93.2%, 97.3%, and 93.9%, respectively. In the hemiarthroplasty experiment, the accuracies of the four models are 92.4%, 86.1%, 94.2%, and 94.8%, respectively. Among these, we found that the RF model has the best sensitivity and ANN model has the best area under the receiver operating characteristic (AUROC) score in both experiments.
Conclusions: The THA experiment confirmed that the performance of the RF model is better than the other models. The key factors affecting the prognosis after THA surgery are creatinine, sodium, anesthesia duration, and dialysis. In the hemiarthroplasty experiment, the ANN model showed more accurate results. Poor kidney function increases the risk of hospital re-admission. This research highlights that RF and ANN model perform well on the hip replacement surgery outcome prediction.

* Corresponding author.
E-mail address: tunghospitalbi@gmail.com (S.-S. Tsou).

https://doi.org/10.1016/j.jhqr.2022.11.009
2603-6479/

PALABRAS CLAVE
Artroplastia total de cadera;
Hemiartroplastia;
Regresión logística;
Árboles de decisión;
Bosques aleatorios;
Red neuronal artificial

Aplicación de los métodos de aprendizaje automatizado para predecir el reingreso al hospital durante los 30 días posteriores a la artroplastia total de cadera y la hemiartroplastia

Resumen
Objetivos: La artroplastia total de cadera (*total hip arthroplasty* [THA]) y la hemiartroplastia son tratamientos comunes para tratar problemas graves de la articulación de la cadera. El poder predecir la posibilidad de reingreso de un paciente contribuirá a poder ofrecerle una adecuada educación y orientación sanitaria durante su hospitalización.
Métodos: Esta investigación llevó a cabo Regresiones Logísticas *(Logistic Regression* [LR]), Árboles de Decisión (*Decision Trees* [DT]), Bosques Aleatorios (*Random Forests* [RF]) y Redes Neuronales Artificiales (*Artificial Neural Networks* [ANN]) a fin de establecer modelos predictivos y comparar su eficacia en los reingresos durante los 30 días posteriores a la THA o la hemiartroplastia. El presente estudio engloba los datos demográficos, las mediciones fisiológicas, los antecedentes clínicos y los resultados de los análisis clínicos de los pacientes.
Resultados: Se estudiaron 508 pacientes de THA y 309 de hemiartroplastia desde septiembre de 2016 hasta diciembre de 2018. El índice de precisión mostrado por los cuatro modelos LR, DT, RF y ANN en el experimento de THA alcanzó respectivamente el 94,3%, el 93,2%, el 97,3% y el 93,9%. En el experimento de hemiartroplastia, el índice de precisión de los cuatro modelos fueron del 92,4%, del 86,1%, del 94,2% y del 94,8%, respectivamente. Entre estos, descubrimos que el modelo RF mostró la mejor sensibilidad y el modelo ANN mostró la mejor puntuación acerca del área bajo la característica operativa del receptor (*area under the receiver operating characteristic* [AUROC]) en ambos experimentos.
Conclusiones: Nuestra experiencia de THA confirmó que el rendimiento de RF era mejor que el de los demás modelos. Los factores clave que inciden en el pronóstico tras la cirugía de THA son la creatinina, el sodio, la duración de la anestesia y la diálisis. En los resultados en la hemiartroplastia, el modelo ANN mostró resultados más precisos. La alteración del funcionamiento renal incrementa el riesgo de reingreso en el hospital. La presente investigación subraya que los modelos RF y ANN funcionan bien con respecto a la predicción del resultado de la cirugía de reemplazo de cadera.

Introduction

Previous research has established that hip replacement surgery accounts for a large proportion of surgical procedures, and the demand for this surgery is increasing.[1] Total hip arthroplasty (THA) and hemiarthroplasty are common treatments for severe joint disease. In THA both the femoral head and acetabulum are replaced, while hemiarthroplasty only replaces the femoral head of patients with joint disease.[2] Under the Diagnosis Related Groups (DRG) policy, hospitals are held responsible for all the treatment costs of the disease. The THA and hemiarthroplasty cases form part of wider DRG cases in Taiwan. Therefore, it is advantageous to establish models for monitoring medical quality.

It is well known that the outcomes vary for THA and hemiarthroplasty given different reasons or indications for the procedures. Hemiarthroplasty is generally reserved for older, sicker patients that suffer from femoral neck fractures, whereas total hip arthroplasty patients are more heterogeneous, being often younger, healthier and more active. Recent evidence reports that hemiarthroplasty has worse outcomes compared to THA in terms of the incidence of re-admission within 90 days of discharge and all causes of complications.[3] Unplanned re-admissions after surgery not only expose potential medical quality issues but also consume healthcare resources. Most solutions to reduce the re-admission rate aim to strengthen pre-discharge health education and post-discharge follow-up.[4,5] Evidence shows that because these interventions were delivered near or after discharge and the care plan was not graded, the improvements in patients' outcomes were limited. If physicians can identify high-risk groups early in patients' hospital stays, this will help to formulate the next medical plan.

Other studies have shown 30-day re-admission rates of 2.2–6.8% in THA patients.[6,7] Regardless of the surgical approach, the re-admission rate within 30 days of the first hip replacement surgery was 8.4%.[8] The Centers for Medicare and Medicaid Services (CMS) Quality Index emphasizes a 30-day re-admission target, which is consistent with Taiwan's health insurance policy. Therefore, the 30-day re-admission rate is a reasonable time interval to evaluate medical results. One study has explored the factors that influence the quality of the prognosis for hip replacement. In geriatric hip fracture patients, advanced age, gender (being male), and decreased body mass index (BMI) are risk factors in 30-day mortality studies.[9] Another study showed that patients with end stage renal disease (ESRD) and chronic kidney disease (CKD) have a poor prognosis following hemiarthroplasty surgery.[10] Kidney function often

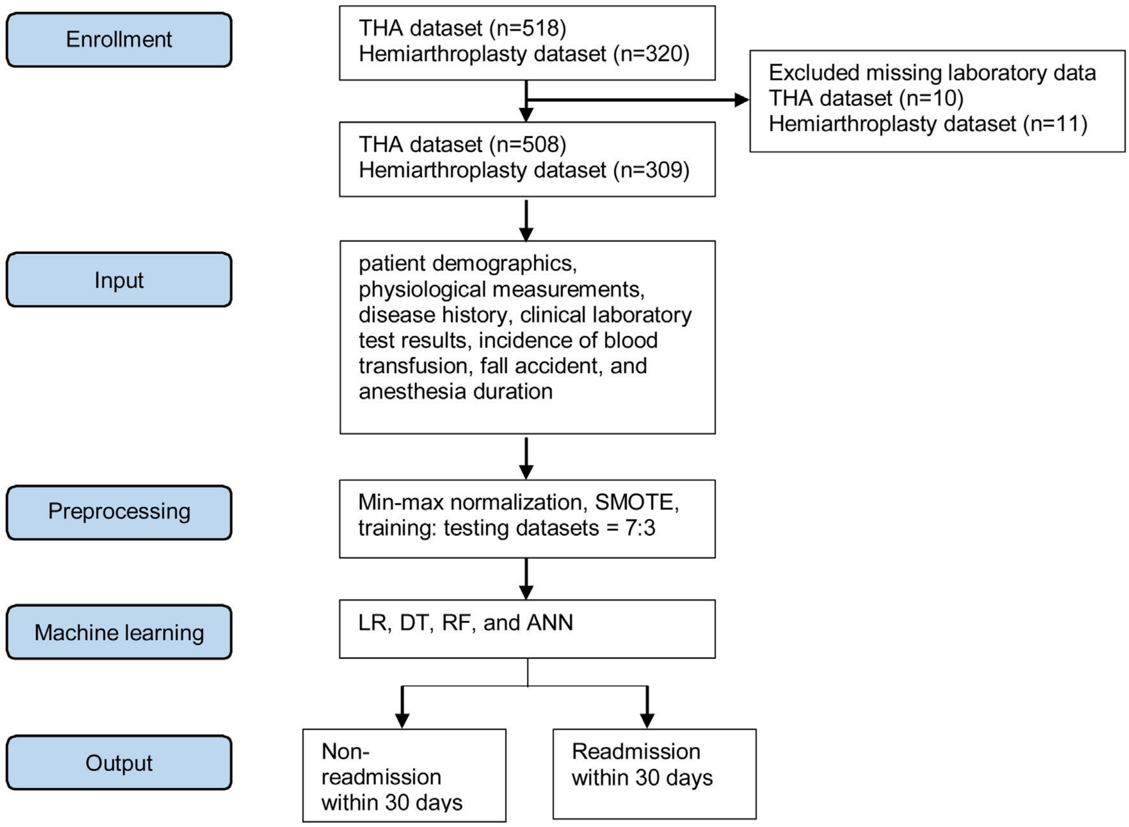

Figure 1 Experiment flow diagram in the study.

begins to change before a patient is diagnosed with ESRD or CKD. Early intervention can be facilitated if postoperative conditions can be predicted before renal disease diagnosis.

Existing research recognizes the key role of machine learning methods in predicting patient outcomes and to establishing clinical decision-making tools. Data from several studies suggest that the random forest and least absolute shrinkage and selection operator (LASSO) regression achieved good performance.[11,12] The specific objective of this present study was to establish a predictive model of re-admission within 30 days after THA or hemiarthroplasty. This project utilized logistic regression (LR), decision trees (DT), random forests (RF), and artificial neural networks (ANN) to construct prediction models and explored the key factors related to patient outcomes. The study was designed to assist medical staff in the early prediction of re-admissions in hospitalized patients undergoing hip replacement surgery and to remind physicians to make plans accordingly during the hospitalization period. In order to reduce the re-admission rate, medical staff could follow this clinical reference to strengthen the implementation of graded care for groups at high-risk of re-admission.

Methods and materials

Data collection

The data collected was based on the Health Information System database provided by a teaching hospital in Taiwan from September 2016 to December 2018.

Inclusion and exclusion criteria

Inclusion criteria were patients who underwent THA or hemiarthroplasty as primary surgery. The exclusion criteria were patients without clinical laboratory test results, such as renal function, coagulation function, and electrolytes, before surgery but during hospitalization. Fig. 1 presents 518 and 320 subjects in the THA and hemiarthroplasty datasets, respectively. Of these 10 and 11 were excluded from each respective study based on missing clinical laboratory test results.

Data analysis and ethical considerations

All data were pre-processed to eliminate all identifiable personal information. This retrospective research was approved by the Institutional Research Board (Ethical Committee of Tungs' Taichung MetroHarbor Hospital, No. 107078).

This article aimed to explore the application of machine learning models to predict re-admissions that occur within 30 days of discharge after THA or hemiarthroplasty. Since many researchers have utilized gender, physiological measures, and disease history to measure the likelihood of re-admission after surgery,[13–16] we collected the above-mentioned variables. In addition, we collected other important variables such as incidence of blood transfusion, accidental fall (leading up to the surgery), and duration of anesthesia.

Experimental analysis

MATLAB R2019a was used for model training. The research process used by the authors is shown in Fig. 1. The data were normalized using min–max normalization to avoid unit differences (such as age versus glucose level) or numerals of different sizes, which can confound the results of statistical analysis. We applied the oversampling method based on the synthetic minority oversampling technique (SMOTE) to standardize the imbalanced datasets to a ratio of 1:1. SMOTE is a preprocessing method used to adjust the distribution of groups in the dataset to avoid under-representation issues due to imbalanced data. The data were divided into training and testing datasets in a ratio of 7:3 (shown in Appendix A).

The authors used LR, DT, RF, and ANN to compare their predictive capabilities. LR is a statistical model for basic binary classification using the logistic function.[17] The DT approach is a tree-like structure used to aid decision making, consisting of a decision diagram and its possible consequences, including resource costs and utility.[18] The RF approach is an ensemble learning method based on multiple decision Trees.[19] The ANN approach is a kind of perceptron with a learning function formed by imitating biological neural networks. The ANN method is more flexible than LR because the decision boundary can be nonlinear and can be used to solve problems that are more difficult to solve with rule-based programming.[17]

Measures of model performance

A confusion matrix was conducted to examine the model classification results with true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Accuracy was determined by the proportion of correct results among all predicted results (Eq. (1)). Precision was determined from the positive predictive value which refers to the proportion of positive results in tests that are true positives (Eq. (2)). Sensitivity was determined from the proportion of actual positive samples that received a positive result (Eq. (3)). When the sensitivity is low, it means that many patients who are re-admitted within 30 days are misclassified as non-readmission cases. The F1-score (Eq. (4)) is the harmonic mean of the precision and sensitivity. The area under the receiver operating characteristic (AUROC) score is a statistical performance measure used in classification. The various measures are summarized below:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad (2)$$

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \quad (3)$$

$$\text{F1-score} = \frac{2 * TP}{(2 * TP + FP + FN)} \quad (4)$$

Results

Participant characteristics

There were 508 patients in the THA dataset, of which 49.6% were women. The average age in the ''non-readmission group'' was 59.58 years old. In the ''re-admission within 30 days'' group it was very nearly the same at 59.21 years old. In the hemiarthroplasty dataset, there were 309 patients (68.3% women) with average ages of 77.32 years and 77.91 years in the non-readmission and re-admission groups, respectively. The rate of re-admission within 30 days was 3% (14/508) in the THA dataset and 7% (21/309) in the hemiarthroplasty dataset.

T-tests, Fisher's exact tests, and chi-square tests were used to compare the differences between re-admission within 30 days and non-readmission. Table 1 reports a general overview of the characteristics of the datasets, including social demographics, physiological measures, disease history, and clinical laboratory test results. As can be seen in Table 1, there are significant differences in anesthesia duration ($p = 0.045$), creatinine levels ($p < 0.001$), and sodium levels ($p = 0.010$) between the two groups in the THA dataset. There is no difference between the two groups in the hemiarthroplasty dataset.

Prediction of outcome

Total hip arthroplasty

In the THA experiment, 692 cases were used for training, and 296 were used for testing. The confusion matrix of the THA testing dataset in the four algorithm models is shown in Appendix B. We compared the accuracy of the four models. The LR, DT, RF, and ANN models classified individuals correctly in 94.3%, 93.2%, 97.3%, and 93.9% of the cases, respectively. The comparison results of this experiment show that RF had the greatest accuracy (0.973), sensitivity (0.973), precision (0.973), and F1-score (0.973). The AUROC score of the testing dataset in the ANN model with 15 hidden layers had the highest score of 0.989 with a sensitivity score of 0.912, precision score of 0.964, and F1-score in 0.938. These results are shown in Fig. 2 and Table 2.

The simple LR analysis is shown in Appendix C. Dialysis (OR = 37.923, 95% CI = 2.247–640.051, *p*-value = 0.012), creatinine levels (OR = 1.933, 95% CI = 1.177–3.175, *p*-value = 0.009), and sodium levels (OR = 0.816, 95% CI = 0.698–0.954, *p*-value = 0.011) were significant independent variables. The multiple LR is shown in Appendices D and E. Using the conditional forward method, anesthesia duration (OR = 1.010, 95% CI = 1.001–1.020, *p*-value = 0.038), sodium levels (OR = 0.794, 95% CI = 0.676–0.932, *p*-value = 0.005), and dialysis (OR = 7.352, 95% CI = 3.015–953.953, *p*-value = 0.007) were the significant predictors of 30-day re-admission. When using the conditional backward method, anesthesia duration (OR = 1.010, 95% CI = 1.001–1.020, *p*-value = 0.030), sodium levels (OR = 0.816, 95% CI = 0.703–0.948, *p*-value = 0.008), and creatinine levels (OR = 2.022, 95% CI = 1.208–3.384, *p*-value = 0.007) were the significant predictors of 30-day re-admission.

Table 1 Characteristics of patients who received THA or hemiarthroplasty operation.

Variable	THA		p-Value	Hemiarthroplasty		p-Value
	Non-readmission (n = 494)	Readmission within 30 days (n = 14)		Non-readmission (n = 288)	Readmission within 30 days (n = 21)	
Age	59.58 (13.64)	59.21 (13.37)	0.921	77.32 (10.46)	77.91 (11.53)	0.807
BMI	25.88 (4.40)	25.47 (5.42)	0.732	23.20 (3.68)	23.46 (2.48)	0.750
Sex						
Female	245 (49.6%)	7 (50.0%)	0.976	196 (68.1%)	15 (71.4%)	0.748
Male	249 (50.4%)	7 (50.0%)		92 (31.9%)	6 (28.6%)	
Dialysis						
No	493 (99.8%)	13 (92.9%)	0.054†	280 (97.2%)	20 (95.2%)	0.474†
Yes	1 (0.2%)	1 (7.1%)		8 (2.8%)	1 (4.8%)	
CKD						
No	485 (98.2%)	13 (92.9%)	0.246	252 (87.5%)	18 (85.7%)	0.737
Yes	9 (1.8%)	1 (7.1%)		36 (12.5%)	3 (14.3%)	
CAD						
No	488 (98.8%)	14 (100%)	1.000†	281 (97.6%)	20 (95.2%)	0.434†
Yes	6 (1.2%)	0 (0%)		7 (2.4%)	1 (4.8%)	
DM						
No	419 (84.8%)	11 (78.6%)	0.460†	196 (68.1%)	14 (66.7%)	0.895
Yes	75 (15.2%)	3 (21.4%)		92 (31.9%)	7 (33.3%)	
HTN						
No	310 (62.8%)	9 (64.3%)	0.907	115 (39.9%)	6 (28.6%)	0.303
Yes	184 (37.2%)	5 (35.7%)		173 (60.1%)	15 (71.4%)	
Blood transfusion						
No	278 (56.3%)	8 (57.1%)	0.949	148 (51.4%)	9 (42.9%)	0.450
Yes	216 (43.7%)	6 (42.9%)		140 (48.6%)	12 (57.1%)	
Fall accident						
No	478 (96.8%)	14 (100%)	1.000†	103 (35.8%)	10 (47.6%)	0.276
Yes	16 (3.2%)	0 (0%)		185 (64.2%)	11 (52.4%)	
Glucose	119.86 (48.43)	114.57 (51.18)	0.687	157.42 (79.04)	165.86 (120.05)	0.651
Anesthesia duration	165.71 (38.73)	186.79 (40.13)	0.045*	147.59 (35.41)	154.52 (52.87)	0.405
Creatinine	0.70 (0.38)	1.15 (1.70)	<0.001*	1.16 (1.39)	1.11 (1.56)	0.867
Potassium	3.89 (0.38)	4.06 (0.57)	0.101	3.93 (0.53)	3.81 (0.46)	0.289
Sodium	139.14 (2.72)	137.21 (3.51)	0.010*	136.82 (3.58)	136.00 (6.43)	0.342
INR	0.97 (0.07)	0.96 (0.04)	0.385	1.01 (0.14)	0.99 (0.05)	0.662
PT	10.14 (0.64)	9.99 (0.35)	0.398	10.46 (1.40)	10.33 (0.49)	0.678
APTT	27.38 (2.27)	26.89 (2.30)	0.427	26.76 (3.28)	26.45 (2.86)	0.674

Continuous variables were expressed as means and standard deviations. THA, total hip arthroplasty; BMI, body mass index; CKD, chronic kidney disease; CAD, cardiac artery disease; DM, diabetes; HTN, hypertension; INR, international normalized ratio; PT, prothrombin time; APTT, activated partial thromboplastin time.

* p-Value < 0.05.

† Fisher's exact test.

Both duration of anesthesia and sodium levels are significant predictors in the forward and backward regression selection methods. The different factors between the forward and backward regression are dialysis and creatinine levels. These two factors indicate poor renal function and both p-values are less than 0.05 in the simple LR analysis. Therefore, the risk factors used are anesthesia duration, sodium levels, and poor renal function.

The AUROC score of the decision tree testing dataset is 0.933. The final model has 6 independent variables as predictors: creatinine levels, potassium levels, glucose levels, dialysis, CKD, and age.

Hemiarthroplasty

In the hemiarthroplasty experiment, 404 cases were used for training, and 172 were used for testing. The confusion matrix of the hemiarthroplasty testing dataset in the four algorithm models are shown in Appendix F. Table 3 shows that the ANN model had the highest accuracy at 0.948, while the RF model

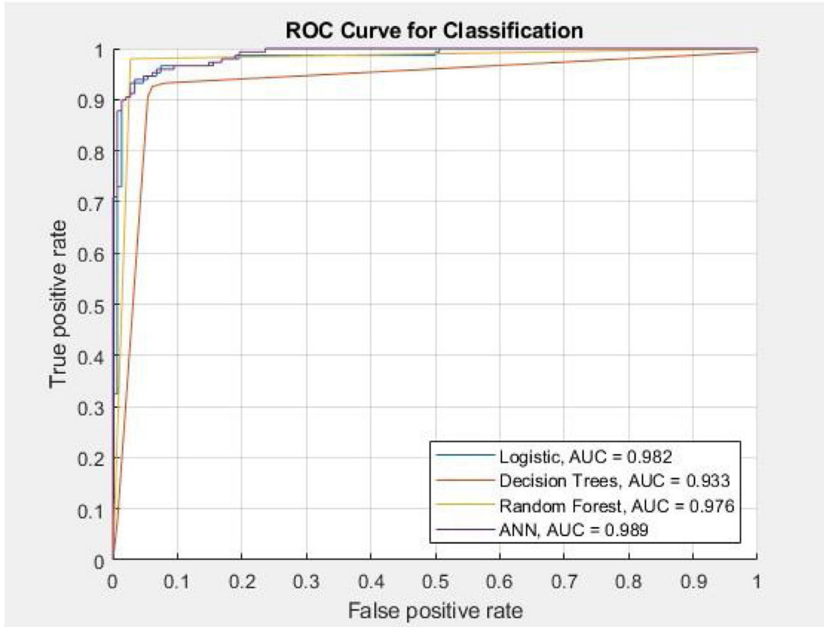

Figure 2 The AUROC score in the THA testing dataset was 0.982, 0.933, 0.976, 0.989 in the four algorithm models.

Table 2 The results of the four algorithm models in the THA testing dataset.

Algorithm	Accuracy	Sensitivity	Precision	F1-score
LR	0.943	0.939	0.946	0.942
DT	0.932	0.926	0.938	0.932
RF	0.973	0.973	0.973	0.973
ANN	0.939	0.912	0.964	0.938

Table 3 The results of the four algorithm models in the hemiarthroplasty testing dataset.

Algorithm	Accuracy	Sensitivity	Precision	F1-score
LR	0.924	0.919	0.929	0.924
DT	0.861	0.849	0.869	0.859
RF	0.942	0.919	0.963	0.941
ANN	0.948	0.907	0.987	0.946

had an accuracy of 0.942. The RF and LR models showed the best sensitivity (0.919), while the ANN model had the best precision (0.987) and F1-score (0.946). The ANN model had the highest AUROC score across the testing dataset at 0.955, and the DT the lowest (0.860) (Fig. 3). The simple LR analysis is shown in Appendix G without significant independent variables.

Discussion

As mentioned in the literature review, the DRG policy aims to reduce the financial burden on medical insurance and make the utility of medical resources more efficient. In prognosis indicators, the incidence of surgery-related complications after the hip replacement was 6.94%, and it was affected by the surgical method.[20] Complications are numerous and do not necessarily lead to hospitalization, so complications were not included in the study results.

Meanwhile, the 30-day and 90-day mortality rates after hip replacement were 0.30% and 0.65%, respectively.[21] In Taiwan, elderly patients over 80 years old have a 7-day mortality rate of 0.6% after that surgery.[22] It can be seen that the mortality rate may be lower regardless of age, so this indicator has not been included to in this study.

Based on the literature and international guidelines, the assessment results of this study focus on 30-day re-admissions. Our study found that the 30-day re-admission rate for hemiarthroplasty (7%) is higher than THA (3%). A report has shown that among the treatment results of patients with hip fractures, the re-admission rate within 90 days of hemiarthroplasty was significantly higher than that of THA ($p < 0.001$).[3]

The present study was designed to compare the performance of four types of machine learning (LR, DT, RF, and ANN). DT is a non-parametric method, and the path can be transformed into a hierarchical model of if-then-rules. RF reforms the issue of overfitting the training set in DT, and is generally better than DT.[19] Backpropagation is used for supervised learning and recursively updates weights to reduce errors.[23] The application of machine learning in orthopedics continues to increase, and most papers focus on the fields of osteoarthritis detection and prediction, bone and cartilage image segmentation, and spine pathology detection.[24] In addition, there is a Bootstrap aggregation model based on a DT and an extreme gradient enhancement model to predict the probability of hip fracture.[25]

The results of this study indicate that LR, DT, RF, and ANN models are suitable for predicting the likelihood of re-admission within 30 days. In the THA dataset, the accuracy of the RF model (0.973) is greater than that of ANN (0.939), LR (0.943), and DT (0.932), while ANN is 0.948 and RF is 0.942 in the hemiarthroplasty dataset. The authors are inclined to

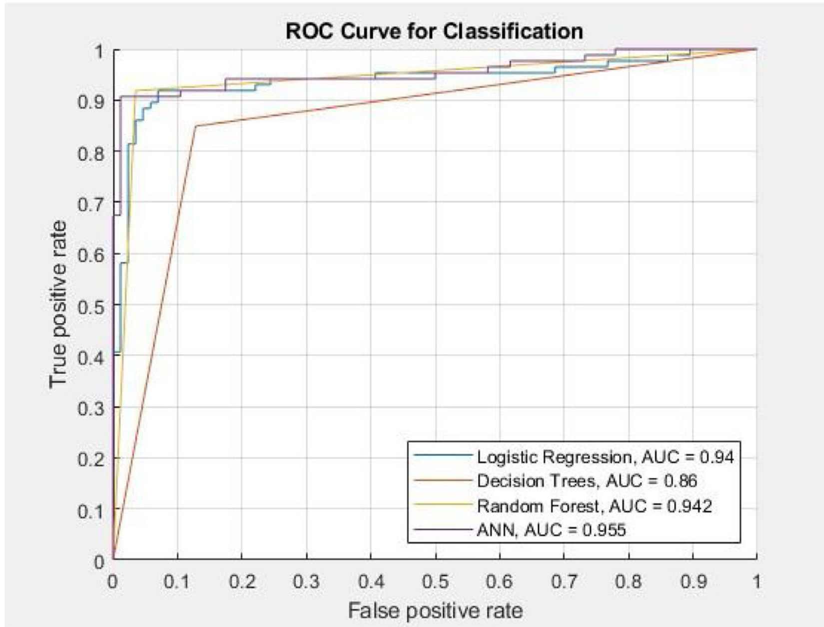

Figure 3 The AUROC score in the hemiarthroplasty testing dataset was 0.940, 0.860, 0.942, 0.955 in LR, DT, RF, and ANN, respectively.

use accuracy to compare the performance of the four models rather than ROC analysis because of the discreteness of the data.[26] At present, these four models are complementary to the medical field.

A comparison of the findings with those of previous studies confirms that LR is generally applicable to disease prediction. The marked difference in our study is the inclusion of clinical laboratory test results in addition to the machine learning methods. Previous studies looked at postoperative complications using e.g. LASSO and enhanced regression to predict 30-day mortality and cardiac complications after THA,[12,27] or used a self-developed questionnaire to predict complications using binary LR analysis.[28] One study used the patient's cognitive assessment process to predict postoperative hip disability and osteoarthritis by building a machine learning model with LASSO.[29] The major difference between this current work and other studies is the addition of clinical laboratory test results on the day of admission. As a result, risk assessment and prediction of prognosis can be carried out on the day of admission, providing useful insights into clinical care.

In the THA group, it was found that creatinine levels, potassium levels, glucose levels, dialysis, CKD, and age are related factors that affect the depth of decision tree branching. This finding is similar to previous studies that showed that advanced age,[13,14] diabetes,[15] renal insufficiency,[13] CKD,[16] prolonged hospitalization time and greater than two emergency re-admissions are important predictors for total hip replacement.[14]

THA and hemiarthroplasty surgery are sufficiently different that their prognoses cannot be compared with each other. In fact, it is more important to ascertain the key factors of hemiarthroplasty because its prognosis is worse than that of THA. Previous studies were limited to patients with hip fractures. In a study on femoral neck fractures, a multivariate LR model adjusted for age, gender, and treatment methods found that ESRD (OR = 3.09) and CKD (OR = 1.43) had an increased risk of re-admission within 90 days.[10] It can be seen from the DT of our hemiarthroplasty experiment that creatinine levels, glucose levels, sodium levels, potassium levels, APTT, INR, dialysis, and age are all related factors affecting the outcome.

Our findings, while preliminary, suggest that the key factors for THA and hemiarthroplasty are different. The re-admission rate of hemiarthroplasty patients with coagulation dysfunction (APTT < 23.42 sec, INR < 1.12) is higher. It is necessary to identify high-risk groups early and provide corresponding intervention measures. The present results are significant in at least two major respects. First, the utility of the clinical laboratory test results, gender, and disease history of the patient at the time of admission is effective for prediction. Secondly, whether machine learning algorithms are non-parametric methods, linear models, or neural networks, they all have good predictive capabilities when using representative features. The AUROC score of the testing dataset all reached 0.93 or more in the THA experiment, while all reached 0.86 in the hemiarthroplasty.

In clinical practice, we apply machine learning to find more accurate predictive models, screen out important risk factors, and conduct preliminary research on preventive measures. Future research could be directed toward improving patient health status and focusing on renal insufficiency, as well as paying attention to preoperative assessment of patients' electrolyte status. Furthermore, due to the small sample size of our study, caution must be applied, as the findings might not be extrapolated to all patients. Further research should be undertaken to investigate the big data across hospitals to establish a generalized model.

Conclusion

This study set out to explore the application of machine learning models to predict re-admissions within 30 days of discharge after two types of hip replacements. The THA experiments confirmed that the AUROC score was 0.982, 0.933, 0.976, 0.989, and the hemiarthroplasty experiments showed that it was 0.940, 0.860, 0.942, 0.955 in LR, DT, RF, and ANN, respectively. The current data generally highlights the applicability of the four kinds of machine learning.

In the THA dataset, the highest accuracy among the four models is 97.3% in the RF model. It was found that the key factors associated with re-admissions within 30 days are creatinine levels, sodium levels, anesthesia duration, and dialysis from the multivariate LR. In the hemiarthroplasty dataset, the accuracies of the ANN model and the RF model were similar to each other being 94.8% and 94.2%, respectively. Although no significant key factors were found in simple LR, creatinine level is the most important factor in the RF model.

The analysis of independent variables undertaken here, such as demographics, physiological measurements, anesthesia duration, and disease history, especially the results of clinical laboratory tests, has extended our knowledge of model features. However, the scope of this study was limited in terms of the subject number of re-admissions within 30 days. Although the current study is based on a small sample of participants, the findings suggest that RF and ANN have good predictive capabilities.

Authors' contributions

Jia-Min Wu drafted this article and contributed to the final version of the manuscript. Chin-Yu Ou designed the research framework and developed the AI models. Jing-Er Chiu wrote, reviewed, and edited the work. Bor-Wen Cheng approved the final submitted version. Shung-Sheng Tsou is a surgeon who supervised the work and revised the work critically for important intellectual content.

Funding

This research was supported by the Ministry of Science and Technology, Taiwan, R.O.C. under Grant number MOST 108-2221-E-224-014-MY3.

Conflict of interest

The author(s) declared no potential conflicts of interest concerning the research, authorship, or publication of this article.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhqr.2022.11.009.

References

1. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg. 2007;89:780–5, http://dx.doi.org/10.2106/jbjs.F.00222.
2. Lewis DP, Wæver D, Thorninger R, Donnelly WJ. Hemiarthroplasty vs total hip arthroplasty for the management of displaced neck of femur fractures: a systematic review and meta-analysis. J Arthroplasty. 2019;34:1837–43, http://dx.doi.org/10.1016/j.arth.2019.03.070, e1832.
3. Nichols CI, Vose JG, Nunley RM. Clinical outcomes and 90-day costs following hemiarthroplasty or total hip arthroplasty for hip fracture. J Arthroplasty. 2017;32:S128–34, http://dx.doi.org/10.1016/j.arth.2017.01.023.
4. Zheng Q-Y, Geng L, Ni M, Sun J-Y, Ren P, Ji Q-B, et al. Modern instant messaging platform for postoperative follow-up of patients after total joint arthroplasty may reduce re-admission rate. J Orthop Surg Res. 2019;14:464, http://dx.doi.org/10.1186/s13018-019-1407-3.
5. Pelt CE, Gililland JM, Erickson JA, Trimble DE, Anderson MB, Peters CL. Improving value in total joint arthroplasty: a comprehensive patient education and management program decreases discharge to post-acute care facilities and post-operative complications. J Arthroplasty. 2018;33:14–8, http://dx.doi.org/10.1016/j.arth.2017.08.003.
6. Avram V, Petruccelli D, Winemaker M, de Beer J. Total joint arthroplasty readmission rates and reasons for 30-day hospital readmission. J Arthroplasty. 2014;29:465–8, http://dx.doi.org/10.1016/j.arth.2013.07.039.
7. Vorhies JS, Wang Y, Herndon J, Maloney WJ, Huddleston JI. Readmission and length of stay after total hip arthroplasty in a national medicare sample. J Arthroplasty. 2011;26:119–23, http://dx.doi.org/10.1016/j.arth.2011.04.036.
8. Martin CT, Gao Y, Pugely AJ. Incidence and risk factors for 30-day readmissions after hip fracture surgery. Iowa Orthop J. 2016;36:155–60.
9. Schuijt HJ, Bos J, Smeeing DPJ, Geraghty O, van der Velde D. Predictors of 30-day mortality in orthogeriatric fracture patients aged 85 years or above admitted from the emergency department. Eur J Trauma Emerg Surg. 2019, http://dx.doi.org/10.1007/s00068-019-01278-z.
10. Hsiue PP, Seo LJ, Sanaiha Y, Chen CJ, Khoshbin A, Stavrakis AI. Effect of kidney disease on hemiarthroplasty outcomes after femoral neck fractures. J Orthop Trauma. 2019;33:583–9, http://dx.doi.org/10.1097/bot.0000000000001576.
11. Kunze KN, Karhade AV, Sadauskas AJ, Schwab JH, Levine BR. Development of machine learning algorithms to predict clinically meaningful improvement for the patient-reported health state after total hip arthroplasty. J Arthroplasty. 2020;35:2119–23, http://dx.doi.org/10.1016/j.arth.2020.03.019.
12. Harris AHS, Kuo AC, Bowe T, Gupta S, Nordin D, Giori NJ. Prediction models for 30-day mortality and complications after total knee and hip arthroplasties for veteran health administration patients with osteoarthritis. J Arthroplasty. 2018;33:1539–45, http://dx.doi.org/10.1016/j.arth.2017.12.003.
13. Belmont PJ, Goodman GP, Hamilton W, Waterman BR, Bader JO, Schoenfeld AJ. Morbidity and mortality in the thirty-day period following total hip arthroplasty: risk factors and incidence. J Arthroplasty. 2014;29:2025–30, http://dx.doi.org/10.1016/j.arth.2014.05.015.
14. Ali AM, Loeffler MD, Aylin P, Bottle A. Factors associated with 30-day readmission after primary total hip arthroplasty: analysis of 514 455 procedures in the UK National Health Service. JAMA Surg. 2017;152, http://dx.doi.org/10.1001/jamasurg.2017.3949, e173949.

15. Marchant MH Jr, Viens NA, Cook C, Vail TP, Bolognesi MP. The impact of glycemic control and diabetes mellitus on perioperative outcomes after total joint arthroplasty. J Bone Joint Surg. 2009;91:1621–9, http://dx.doi.org/10.2106/jbjs.H.00116.
16. Miric A, Inacio MCS, Namba RS. The effect of chronic kidney disease on total hip arthroplasty. J Arthroplasty. 2014;29:1225–30, http://dx.doi.org/10.1016/j.arth.2013.12.031.
17. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform. 2002;35:352–9, http://dx.doi.org/10.1016/S1532-0464(03)00034-0.
18. Quinlan JR. Induction of decision trees. Mach Learn. 1986;1:81–106, http://dx.doi.org/10.1007/BF00116251.
19. Breiman L. Random forests. Mach Learn. 2001;45:5–32, http://dx.doi.org/10.1023/A:1010933404324.
20. Aggarwal VK, Elbuluk A, Dundon J, et al. Surgical approach significantly affects the complication rates associated with total hip arthroplasty. Bone Joint J. 2019;101-B:646–51, http://dx.doi.org/10.1302/0301-620x.101b6.Bjj-2018-1474.R1.
21. Berstock JR, Beswick AD, Lenguerrand E, Whitehouse MR, Blom AW. Mortality after total hip replacement surgery. Bone Joint Res. 2014;3:175–82, http://dx.doi.org/10.1302/2046-3758.36.2000239.
22. Tsai M-C, Ng Y-Y, Chen W-M, Tsai S-W, Wu S-C. The effects of cement fixation on survival in elderly patients with hip hemiarthroplasty: a nationwide cohort study. BMC Musculoskelet Disord. 2019;20:628, http://dx.doi.org/10.1186/s12891-019-3013-2Sunday, December 25, 2022 at 2:46 am.
23. DeGregory KW, Kuiper P, DeSilvio T, et al. A review of machine learning in obesity. Obes Rev. 2018;19:668–85, http://dx.doi.org/10.1111/obr.12667.
24. Cabitza F, Locoro A, Banfi G. Machine learning in orthopedics: a literature review. Front Bioeng Biotechnol. 2018:6, http://dx.doi.org/10.3389/fbioe.2018.00075.
25. Kruse C, Eiken P, Vestergaard P. Machine learning principles can improve hip fracture prediction. Calcif Tissue Int. 2017;100:348–60, http://dx.doi.org/10.1007/s00223-017-0238-7.
26. Muschelli J. ROC and AUC with a binary predictor: a potentially misleading metric. J Classif. 2020;37:696–708, http://dx.doi.org/10.1007/s00357-019-09345-1.
27. Harris AHS, Kuo AC, Weng Y, Trickey AW, Bowe T, Giori NJ. Can machine learning methods produce accurate and easy-to-use prediction models of 30-day complications and mortality after knee or hip arthroplasty? Clin Orthop Relat Res. 2019;477:452–60, http://dx.doi.org/10.1097/CORR.0000000000000601.
28. Kunze KN, Li J, Movassaghi K, Wiggins AB, Sporer SM, Levine BR. Internal validation of a predictive model for complications after total hip arthroplasty. J Arthroplasty. 2018;33:3759–67, http://dx.doi.org/10.1016/j.arth.2018.08.011.
29. Sniderman J, Stark RB, Schwartz CE, Imam H, Finkelstein JA, Nousiainen MT. Patient factors that matter in predicting hip arthroplasty outcomes: a machine-learning approach. J Arthroplasty. 2021;36:2024–32, http://dx.doi.org/10.1016/j.arth.2020.12.038.