ORIGINAL ARTICLE

Machine-learning-based approach for nonunion prediction following osteoporotic vertebral fractures

Shinji Takahashi[1] · Hidetomi Terai[2] · Masatoshi Hoshino[2] · Tadao Tsujio[3] · Minori Kato[1] · Hiromitsu Toyoda[1] · Akinobu Suzuki[1] · Koji Tamai[1] · Akito Yabu[1] · Hiroaki Nakamura[1]

Abstract

Purpose An osteoporotic vertebral fracture (OVF) is a common disease that causes disabilities in elderly patients. In particular, patients with nonunion following an OVF often experience severe back pain and require surgical intervention. However, nonunion diagnosis generally takes more than six months. Although several studies have advocated the use of magnetic resonance imaging (MRI) observations as predictive factors, they exhibit insufficient accuracy. The purpose of this study was to create a predictive model for OVF nonunion using machine learning (ML).

Methods We used datasets from two prospective cohort studies for OVF nonunion prediction based on conservative treatment. Among 573 patients with acute OVFs exceeding 65 years in age enrolled in this study, 505 were analyzed. The demographic data, fracture type, and MRI observations of both studies were analyzed using ML. The ML architecture utilized in this study included a logistic regression model, decision tree, extreme gradient boosting (XGBoost), and random forest (RF). The datasets were processed using Python.

Results The two ML algorithms, XGBoost and RF, exhibited higher area under the receiver operating characteristic curves (AUCs) than the logistic regression and decision tree models (AUC = 0.860 and 0.845 for RF and XGBoost, respectively). The present study found that MRI findings, anterior height ratio, kyphotic angle, BMI, VAS, age, posterior wall injury, fracture level, and smoking habit ranked as important features in the ML algorithms.

Conclusion ML-based algorithms might be more effective than conventional methods for nonunion prediction following OVFs.

Introduction

Osteoporotic vertebral fractures (OVFs) comprise a major health concern for the elderly population [1]—they degrade the quality of life and reduce life expectancy [2, 3]. Although most OVFs heal within a few months [4], some exhibit nonunion even after conservative treatment, including osteoporotic medication, bed rest, and braces [4]. Nonunion can lead to severe back pain and neurological deficits [5, 6], often requiring surgical intervention [7]. As the efficacy of surgical intervention is higher when it is administered early [8], early nonunion prediction is essential for effective treatment. Several risk factors for nonunion have been reported in the literature [5, 9–13]—in particular, magnetic resonance imaging (MRI) has been extensively utilized for prediction. However, there is still room for improvement in conventional methods for nonunion prediction.

Recently, machine learning (ML) has been widely incorporated in prediction models [14, 15]. The most widely used ML methods are supervised learning methods [14], which search for algorithms that produce general hypotheses based on externally supplied instances and output predictions for future instances [15]. Accurate disease

✉ Shinji Takahashi
stakahashi@omu.ac.jp

1 Department of Orthopaedic Surgery, Osaka Metropolitan University Graduate School of Medicine, 1-4-3 Asahi-Machi, Abeno-Ku, Osaka 545-8585, Japan

2 Department of Orthopaedic Surgery, Osaka City General Hospital, 2-15-16, Miyakojima Hon-Dori, Miyakojima-Ku, Osaka, Japan

3 Department of Orthopaedic Surgery, Shiraniwa Hospital, 6-10-1. Shiraniwadai, Ikoma City, Nara, Japan

outcome prediction is one of the most interesting and challenging tasks for physicians. ML-based techniques can be used to discover and identify patterns and relationships between them by analyzing complex datasets [16]. Consequently, ML methods have become popular tools among medical researchers. There are several ML methods, such as general linear regression models, decision trees, random forests (RFs), and extreme gradient boosting (XGBoost) [17, 18]. In this study, we hypothesized that these models would improve the prediction accuracy of conventional nonunion prediction models. The purpose of this study is to create an ML-based nonunion prediction model.

Materials and methods

Study design

We consider datasets based on two previously conducted multicenter cohort studies[4, 19] (Fig. 1). The first cohort study was conducted to investigate the risk factors for poor prognosis in OVF patients at 25 institutions between 2005 and 2009. The second cohort study was conducted to investigate changes in MRI observations six months after the OVF onset at 11 institutions between 2012 and 2015. The inclusion and exclusion criteria, treatment protocols, and MRI protocols of the two cohort studies were identical, as described here.
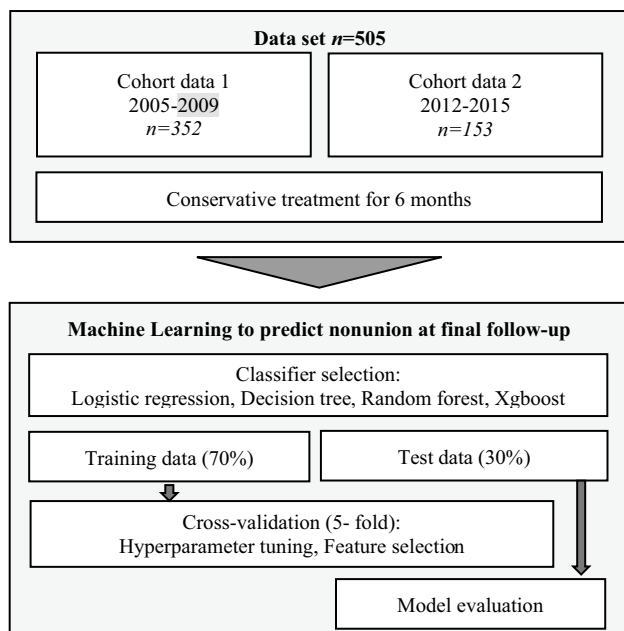

Fig. 1 Study flow chart

Written informed consent was obtained from all study participants. Each study protocol was approved by the Institutional Review Board of the representative institution.

Patients

Details of the patients are described in our previous studies [4, 19]. The inclusion criteria were age greater than 65 years and diagnosis of fresh OVF. The exclusion criteria were pathological fractures, two or more new fractures, malignant cancer, dementia, and high-energy injuries. In the first cohort study, 485 patients were enrolled. Among them, 15 died, six underwent surgery, 11 were excluded because of other diseases, and 33 were lost during the follow-up procedure. Consequently, 420 patients completed the six-month follow-up. Finally, 352 patients (72.6% follow-up rate) with the required data, including examinations and completed questionnaires at both the time of enrollment and the six-month follow-up, were analyzed in the present study. In the second prospective multicenter cohort study, 218 patients with symptomatic OVFs were eligible for participation. Among them, three died, eight were excluded because of other diseases, and 54 were lost during the follow-up period. As a result, 153 patients (125 females and 28 males) completed the six-month follow-up (70.2% follow-up rate) and were analyzed in the present study.

At the time of enrollment, the severity of pain was subjectively assessed using the Visual Analogue Scale (VAS), which identified the average level of back pain experienced by patients over the week preceding the study.

Imaging assessment

At the time of enrollment and the six-month follow-up, patients were examined using plain radiography. The kyphotic angle of the fractured vertebrae was determined using the segmental angle in weight-bearing positions. The relative height of the anterior wall (%) of the affected vertebra was calculated as follows: $[2 \times$ affected vertebral height/(lower vertebral height + upper vertebral height)$] \times 100$[4]. Bone mineral density was measured at the toe and hip in the first and second cohorts, respectively.

MRI

The patients were examined using MRI at the time of enrollment and the six-month follow-up. The signal change patterns within the fractured vertebral bodies identified via MRI were classified based on midsagittal and bilateral parasagittal (medial aspect of pedicles) T1WI and T2WI. Signal change patterns on T1WI were classified into three categories—diffuse low, confined low, and no signal change—whereas those on T2WI were classified into four categories—high

(Fig. 2A), confined low, diffuse low (Fig. 2B), and no signal change. The intensity of the high signal changes on T2WI was defined to be similar to that of the cerebrospinal fluid. If at least two of the three slices exhibited a low signal change over more than 50% of the area of the vertebral body, the observation was considered to be a diffuse low signal change. With respect to intra- and inter-reliability, the kappa values of the MRI observations were in the range of 0.712–0.942 [4].

Nonunion definition

Plain radiographs were assessed based on comparisons of lateral profiles between supine and weight-bearing positions because intravertebral cleft can be easily identified in the supine position. Nonunion was defined as a recognizable intravertebral cleft and apparent segmental motion on dynamic plain radiographs (difference greater than 5° between supine and weight-bearing positions) observed at the six-month follow-up.

Treatment strategy

Treatment with braces was continued for 2–3 months, with soft and hard braces prescribed to 60 and 40% of the patients, respectively. The patients were allowed to maneuver into an erect posture when the brace was applied. Additionally, patients were prescribed anti-osteoporotic and pain-relief medications according to their individual status.

Prediction by the surgeon

Nonunion was predicted based on MRI observations. T2 high or T2 diffuse low-signal change was previously reported as a predictive factor for nonunion [5, 19, 20].
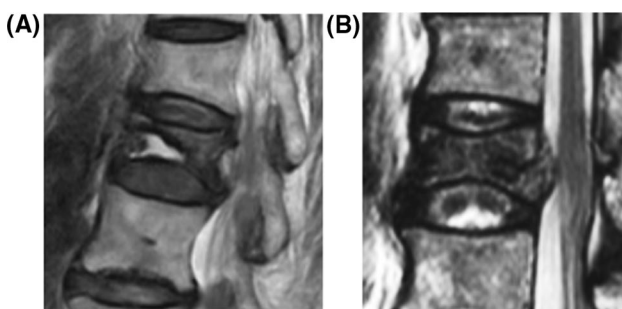

Fig. 2 MRI observations in the acute phase (**A**) High-signal change on T2 weighted image was observed at the fractured vertebral body. (**B**) Diffuse low-signal change on T2 weighted image was observed at the fractured vertebral body

Machine learning

Four ML models—logistic regression, decision tree, XGBoost, and RF—were utilized for nonunion prediction following OVF (Fig. 1). A Python (v. 3.7.6) library called scikit-learn (v. 0.22) was used for this purpose. The specific parameters of the ML models are depicted in Fig. 1. The scikit-learn package was used for logistic regression, decision trees, and RF. XGBoost was used as the gradient boosting machine, which features split finding algorithms that handle sparse data with node default directions, address weighted data using merge and prune operations, and efficiently enumerate over all possible splits for splitting threshold optimization [17, 21].

In aggregate, 17 explanatory variables, including the variables listed in Table 1, were used in the ML models. Before developing the prediction models, the collected data were divided into two categories—70% were allocated to the training dataset and 30% to the test dataset. The cases in the training dataset were used to develop ML and logistic regression models. The cases in the testing dataset were used to validate and compare the performance of the developed models. Each ML method had its own set of hyperparameters—such as the number of layers in XGBoost or the number of trees in RF. Five-fold cross-validation was used to determine the optimal hyperparameters. This cross-validation process was used only for model development; the performance of the final models was evaluated using the testing dataset.

Recursive feature elimination (RFE), known as wrapper feature selection, employs ML models for the computation of relevance scores of features [22–24]. It first trains a model with an entire feature set and computes a relevance score for each feature. In the next step, the feature with the least relevance score is neglected, and the model is retrained to compute the relevance scores of new features. This process is continued until the desired number of features remains in the feature set.

Performance evaluation

As mentioned previously, the data were randomly divided into two groups—a training group (used to train the model) and a testing group (used to evaluate the model's generalization ability)—with relative proportions of 70 and 30%, respectively. The cases in the testing dataset were used to validate and compare the performances of the developed models. Optimal hyperparameter values were identified using five-fold cross-validation to avoid overfitting during model development. The cross-validation process was used only during model development, and the performance of the final models was evaluated using the testing dataset. To identify the most important

Table 1 Comparison of baseline data between nonunion and union groups

	Nonunion		Union		P-value
	n = 85		n = 420		
	Mean (SD) or n(%)		Mean (SD) or n(%)		
Age (years)	78.4	(6.7)	76.4	(6.8)	0.012
Height (cm)	152.1	(7.4)	151.6	(7.5)	0.596
Weight (kg)	52.7	(9.8)	50.6	(8.6)	0.046
BMI (kg/m²)	22.7	(3.7)	22.0	(3.5)	0.088
Sex (female)	53	(62.4)	274	(65.2)	0.612
Smoker					
Non	69	(81.2)	306	(72.9)	
Former	9	(10.6)	67	(16.0)	
Current	7	(8.2)	447	(11.2)	0.274
History of oral steroid use	7	(8.2)	32	(7.6)	0.846
Old OVF					
0	44	(51.8)	266	(63.3)	0.130
1	39	(45.9)	148	(35.2)	
≥ 2	2	(2.4)	6	(1.4)	
VAS at enrollment (cm)	8.3	(2.1)	7.9	(2.1)	0.113
Level					
Thoracic	6	(7.1)	40	(9.5)	0.002
Thoracolumbar	75	(88.2)	299	(71.2)	
Lumbar	4	(4.7)	81	(19.3)	
MRI T1					
Diffuse low	64	(76.3)	265	(63.1)	
Confined low	21	(24.7)	155	(36.9)	0.031
MRI T2					
Diffuse low	27	(31.8)	61	(14.5)	
Confined low	26	(30.6)	284	(67.6)	
Iso	0	(0)	47	(11.2)	
High	32	(37.7)	28	(6.7)	< 0.001
Posterior wall injury	48	(56.5)	123	(29.3)	< 0.001
Kyphotic angle (degrees)	13.0	(7.1)	9.2	(6.2)	< 0.001
Anterior height ratio (%)	75.8	(16.6)	84.5	(13.5)	< 0.001
T-score of BMD	− 3.01	(0.79)	− 2.82	(0.87)	0.092
Cohort					
First	55	(15.6)	297	(84.4)	0.272
Second	30	(19.6)	123	(80.4)	

BMI: body mass index, *OVF*: osteoporotic vertebral fracture, *VAS*: visual analog scale, *BMD*: bone mineral density

features for nonunion prediction, XGBoost and RF were used to rank the variables in terms of importance. SHAP (SHapley Additive exPlanation) interpretation was used (model-agnostic) to compute feature importance using RF [25]. We trained XGBoost using all generator features and then obtained the feature importance scores (F-score) [26]. The top 10 variables present in all four ML algorithms were considered during the subsequent stages of

model development. The area under the receiver operating characteristic curve (AUC) was used as the performance metric for each model—higher AUC values correspond to better predictive performances. The AUC results were considered excellent for AUC values between 0.9 and 1, good for AUC values between 0.8 and 0.9, fair for AUC values between 0.7 and 0.8, poor for AUC values between 0.6 and 0.7, and failed for AUC values between 0.5 and 0.6 [27]. In addition, accuracy was used as a performance quality metric for the ML classifiers.

Data comparison

The chi-square test or Fisher's exact test was used to compare categorical variables. The t-test was used to compare continuous variables. Statistical significance was taken to correspond to $p < 0.05$. All p values were two-sided.

Results

Comparison between the nonunion and union groups

Table 1 presents a comparison between the nonunion and union groups in terms of demographic, clinical, and radiological data. Age was observed to have higher values in the nonunion group than in the union group (78.4 vs. 76.4 years, $p = 0.012$). The nonunion group exhibited higher weight than the union group (52.7 vs. 50.6 kg, $p = 0.046$). No significant differences in sex, smoking habits, history of oral steroid use, old OVF, and VAS score for LBP at enrollment were noticed between the two groups. In the nonunion group, OVFs were more frequent at the thoracolumbar level (88.2% vs. 71.2%, $p = 0.002$). In terms of MRI observations, diffuse low-signal change in T1 weighted image was more frequently observed in the nonunion group than in the union group (76.3% vs. 63.1%, $p = 0.031$). On T2 weighted images, diffuse low- and high-signal changes were more frequently observed in the nonunion group than in the union group (31.8% vs. 14.5% and 37.7% vs. 6.7%, respectively, $p < 0.001$). Posterior wall injury was more frequent in the nonunion group than in the union group (56.5% vs. 29.3%, $p < 0.001$). The kyphotic angle and anterior height ratio were more severe in the nonunion group than in the union group (13.0° vs. 9.2°, $p < 0.001$ and 75.8% vs. 84.5%, $p < 0.001$, respectively). The T-scores of bone mineral density exhibited no apparent differences. Moreover, there was no difference in nonunion incidence between the first and second cohorts.

Table 2 Result of multivariate logistic regression model for nonunion

	Odds ratio	95% CI lower	Upper	p-value
Age	1.056	1.009	1.106	0.018
Sex	0.652	0.349	1.22	0.181
BMI	1.084	0.993	1.184	0.072
T-score	0.983	0.973	0.995	0.006
Anterior height ratio (%)	0.979	0.960	0.999	0.042
Posterior wall injury	1.786	0.946	3.372	0.074
MRI T2				
Confined low or iso	ref			
Diffuse low or high	7.638	4.145	14.073	<.0001
Level				
Lumbar	ref			
Thoracic	1.637	0.377	7.11	0.511
TL	2.65	0.862	8.15	0.089

BMI: Body mass index, *CI*: Confidence interval, *TL*: Thoracolumbar

Performance evaluation of machine learning algorithms

Table 2 presents the results obtained using the logistic regression model for nonunion incidence. Greater age, lower anterior height ratio, and T2 WI observations, including confined high and diffuse low-signal changes, exhibited a significant increase in adjusted ORs. Figure 3 depicts the results obtained using a decision tree with a depth of 3. Among the 13 patients with T2 high-signal change, VAS ≥ 9 cm, and age ≥ 72 years, 12 exhibited nonunion at the final follow-up consultation.

ML-based systems, including RF and XGBoost, used the following variables: age, sex, VAS, smoking habit, MRI T2 weighted image, posterior wall injury, old OVF, kyphotic angle, steroid use, BMI, anterior wall height, and fracture level. Table 3 presents the optimal hyperparameter values and the performances of the ML classifiers. The ranking of features in terms of importance for the RF and XGBoost models is depicted in Figs. 4A and B, respectively. Table 3 and Fig. 5 present a comparison of the test AUC for nonunion prediction using the proposed models. RF and XGBoost
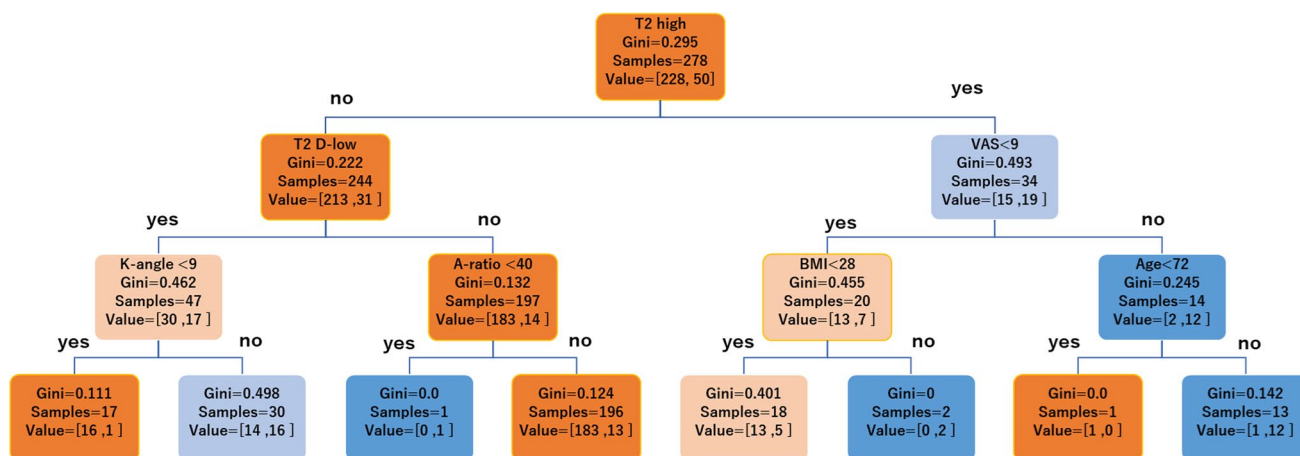

Fig. 3 Decision tree for this study, which aids the analysis to yield the best prediction. Classification results for nonunion and union are highlighted in blue and orange, respectively. Smaller values of the Gini index correspond to darker colors. BMI: Body mass index; VAS: Visual analog scale

Table 3 Optimal values of hyperparameters and comparison of area under the receiver operating characteristic curve among the different models during the prediction of nonunion

Model	Optimal values of hyperparameters	AUC (95% CI)	Accuracy
Logistic regression		0.745 (0.656–0.832)	0.784
Decision tree	Maximum depth = 3 Criterion = Gini index	0.726 (0.637–0.804)	0.757
Random forest	Maximum depth = 10 Number of estimators = 100	0.860 (0.784–0.916)	0.840
Gradient boosting machine	Maximum depth = 8 Number of estimators = 200 gamma = 0.4	0.845 (0.768–0.905)	0.824

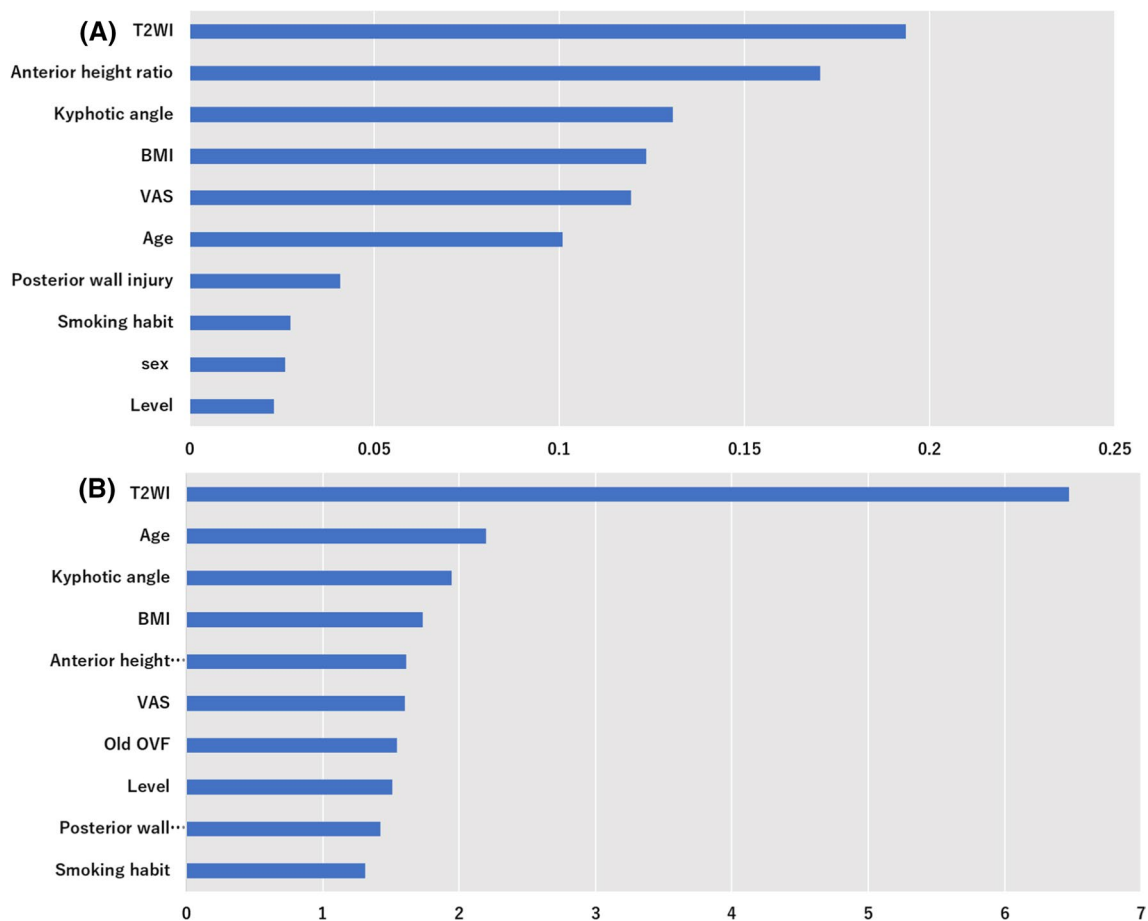*AUC*: Area under curve, *CI*: Confidence interval

Fig. 4 Variance importance plot. (**A**) Variance importance plot for the RF model. (**B**) Variance importance plot for the XGBoost model. T2WI: T2 weighted image; BMI: Body mass index; VAS: Visual analog scale

exhibited good prediction accuracy, while logistic regression and decision tree exhibited fair results. RF exhibited the highest AUC (0.860, 95% CI: 0.784–0.916), followed by XGBoost (0.845, 95% CI 0.768–0.905). In both models, T2 WI observation was the most important feature (Fig. 4A and 4B,). Additionally, anterior height ratio, kyphotic angle, BMI, VAS, age, posterior wall injury, fracture level, and smoking habit were ranked as the top risk factors.

Discussion

This study presents the first attempt to predict nonunion in OVFs based on ML. Evaluation of risk factors of nonunion following OVF onset is essential for efficient treatment, including early surgical intervention [8]. This study compared the predictive abilities of three ML models and a logistic regression model for nonunion following OVF onset. The results demonstrated that RF exhibited the largest AUC and the highest accuracy in nonunion prediction. The XGBoost model also exhibited good AUC and accuracy.

In comparison, we previously reported that the prediction accuracy of spine surgeons based on MRI observations was 75.5%–79.1% using this dataset [5]. Therefore, ML-based models remarkably improved the accuracy compared to human prediction.

ML is not limited by parametric or model-based assumptions that may be difficult to determine a priori. Therefore, ML-based approaches have been used to detect patterns in data to automate complex tasks or make predictions and offer different advantages [28]. Especially, ensemble learning methods comprise popular and powerful ML tools for multivariate regression and classification problems that are used to predict nonlinear relationships [29, 30]. RF constituents of the ensemble are tree-structured predictors, each of which is constructed using an injection of randomness [28]. The XGBoost algorithm balances the influence of each tree by adding weights to reduce overfitting using only a random subset of descriptors during the construction of a tree [31]. The main idea behind the ensemble methodology is to aggregate multiple weighted models to obtain a combined model that outperforms every
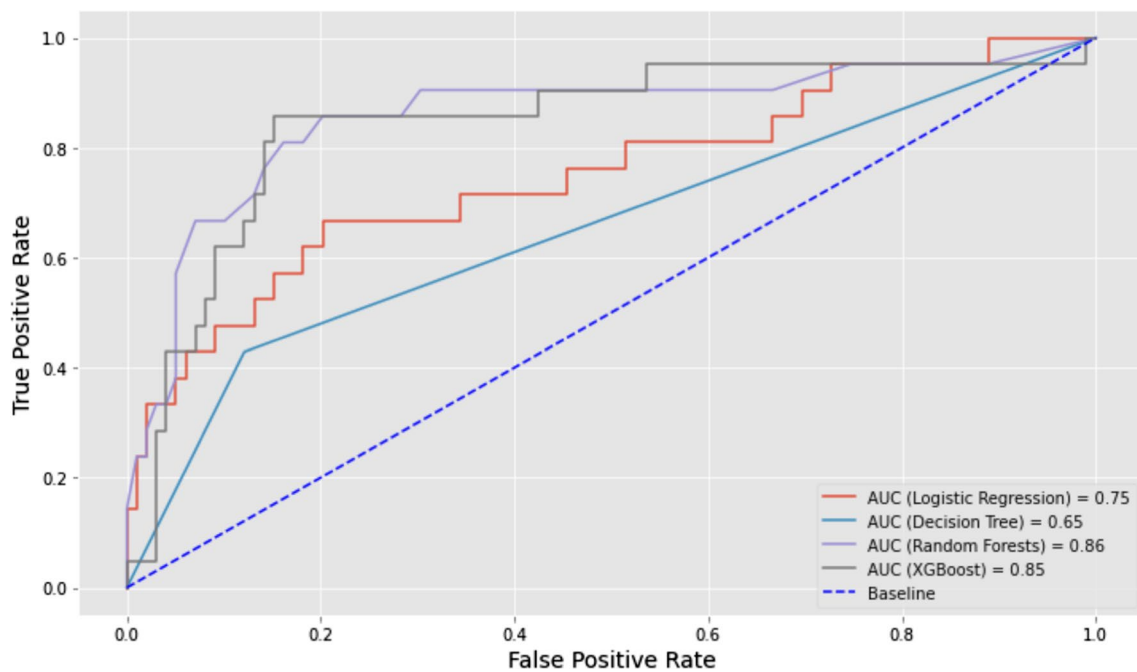
Fig. 5 Areas under the receiver operating characteristic curve for nonunion prediction. AUC: area under the curve

single model in it [32]. Our results demonstrate that these algorithms are effective and efficient ML methods.

The incidence of nonunion in OVFs was reported to range between 13.5 and 19.6% [5, 12, 19]. Patients with nonunion in OVFs generally present functional deterioration, severe pain, and limited mobility. A systematic review [33] reported that the current evidence for using orthotic devices in patients with OVF was inconsistent and of limited quality; therefore, clinicians should undertake careful consideration before prescribing them in practice. Two randomized controlled trials [34, 35] demonstrated no apparent effect of braces on the treatment of OVFs. Some patients required surgical intervention. Early intervention was reported to lead to better clinical and radiological outcomes of vertebral augmentation [8, 36] because the vertebral height of the anterior wall progressed steadily from incidence during the six months of follow-up in patients with non-union [5]. Therefore, early nonunion prediction is essential to improve treatment strategies for OVFs.

The present study established MRI findings, anterior height ratio, kyphotic angle, BMI, VAS, age, posterior wall injury, fracture level, and smoking habit as risk factors in the RF and XGBoost models with high accuracy. Several risk factors of nonunion, e.g., posterior wall injury, instability of fractured vertebrae, and MRI observations depicting high-signal changes on T2WI, diffuse low-signal changes on T2WI, and diffuse low-signal changes on T1WI, have been reported. Our previous studies [5, 19] showed that diffuse low- and high-signal changes on T2 MRI were correlated

with an increased risk of nonunion in the acute phase. In addition, other studies [12, 20] demonstrated that MRI revealed the association of delayed union with T1 and T2 diffuse low-signal changes and T2 high-signal changes in the acute phase. Posterior wall injury and anterior vertebral height are also risk factors for nonunion [5, 12, 19]. Inose et al. [12] demonstrated that the degree of pain is also an important predictor of nonunion. Although instability evaluation may be an important step in instability prediction, approximately 30% of patients have been reported to be incapable of undergoing dynamic radiography because of severe pain [5]. However, bone mineral density was not reported as a risk factor for nonunion. Therefore, dynamic radiography results in the acute phase and bone mineral density were not included in this model. The performance of the proposed RF and XGBoost models demonstrated that the T2WI observation was the most important feature, confirming the reports of previous studies [5, 12, 19, 20]. RF demonstrated that the degree of compression and pain VAS score were important features. Additionally, obesity and smoking were identified as important features in the proposed RF and XGBoost models. Smoking and obesity have been reported as significant risk factors for nonunion in long bone patients [37]. Smoking can lead to local vasoconstriction, and the high affinity of carbon monoxide to hemoglobin can decrease the total amount of oxygen transported to the healing site [37, 38]. The literature on obesity and fracture healing is similarly inconclusive, with some authors reporting no correlation and others reporting

an increased nonunion risk [32]. Obesity can include several conditions, including diabetes, vascular disease, and fracture load, which suggest that obesity is currently not a reliable model of impaired fracture healing [32]. In the current study, information regarding diabetes and vascular disease was not included in the ML models.

This study suffers from certain limitations. Osteoporosis treatment was not included in the ML-based models. However, there have been no reports revealing the effect of osteoporosis agents on the bone healing process in OVF. Further, owing to the small sample size, the data included in the training set for validation were insufficient. Because of this limitation, we did not search for a better ML model; rather, we trained the considered model using various hyperparameters in a predetermined manner. Moreover, the models are only valid corresponding to data collected using the measurement equipment used in this study. Therefore, further research is necessary to generalize the model to higher-dimensional data.

In conclusion, the ML-based system developed in this study based on MRI observations is an effective tool for nonunion prediction following OVFs. The proposed RF and XGBoost models exhibited good performance. Further research is warranted to increase the sample size and apply the proposed ML system to other populations.

Declarations

Conflict of interest The authors declare no conflicts of interest.

References

1. Lyles KW, Gold DT, Shipp KM et al (1993) Association of osteoporotic vertebral compression fractures with impaired functional status. Am J Med 94:595–601. https://doi.org/10.1016/0002-9343(93)90210-G
2. Kado DM, Duong T, Nevitt MC et al (2003) Incident vertebral fractures and mortality in older women: a prospective study. Osteoporos Int 14:589–594. https://doi.org/10.1007/s00198-003-1412-5
3. Lips P, van Schoor NM (2005) Quality of life in patients with osteoporosis. Osteoporos Int 16:447–455. https://doi.org/10.1007/S00198-004-1762-7/TABLES/4
4. Takahashi S, Hoshino M, Takayama K et al (2017) Time course of osteoporotic vertebral fractures by magnetic resonance imaging using a simple classification: a multicenter prospective cohort study. Osteoporos Int 28:473–482. https://doi.org/10.1007/s00198-016-3737-x
5. Takahashi S, Hoshino M, Takayama K et al (2016) Predicting delayed union in osteoporotic vertebral fractures with consecutive magnetic resonance imaging in the acute phase: a multicenter cohort study. Osteoporos Int 27:3567–3575. https://doi.org/10.1007/s00198-016-3687-3
6. Ahmadi SA, Takahashi S, Hoshino M et al (2019) Association between MRI findings and back pain after osteoporotic vertebral fractures: a multicenter prospective cohort study. Spine J 19:1186–1193. https://doi.org/10.1016/j.spinee.2019.02.007
7. Kim DH, Vaccaro AR (2006) Osteoporotic compression fractures of the spine; current options and considerations for treatment. Spine J 6:479–487. https://doi.org/10.1016/J.SPINEE.2006.04.013
8. Takahashi S, Hoshino M, Terai H et al (2018) Differences in short-term clinical and radiological outcomes depending on timing of balloon kyphoplasty for painful osteoporotic vertebral fracture. J Orthop Sci 23:51–56. https://doi.org/10.1016/j.jos.2017.09.019
9. Ha K-Y, Lee J-S, Kim K-W, Chon J-S (2006) Percutaneous vertebroplasty for vertebral compression fractures with and without intravertebral clefts. J Bone Jt Surg Br 88:629–633. https://doi.org/10.1302/0301-620X.88B5.17345
10. Min H-K, Ahn J-H, Ha K-Y et al (2019) Effects of anti-osteoporosis medications on radiological and clinical results after acute osteoporotic spinal fractures: a retrospective analysis of prospectively designed study. Osteoporos Int 30:2249–2256. https://doi.org/10.1007/s00198-019-05125-0
11. Iwata A, Kanayama M, Oha F et al (2017) Does spinopelvic alignment affect the union status in thoracolumbar osteoporotic vertebral compression fracture? Eur J Orthop Surg Traumatol 27:87–92. https://doi.org/10.1007/s00590-016-1844-1
12. Inose H, Kato T, Ichimura S et al (2020) Risk factors of nonunion after acute osteoporotic vertebral fractures: a prospective multi-center cohort study. Spine (Phila Pa 1976) 45:895–902. https://doi.org/10.1097/BRS.0000000000003413
13. Ha KY, Kim YH (2013) Risk factors affecting progressive collapse of acute osteoporotic spinal fractures. Osteoporos Int 24:1207–1213. https://doi.org/10.1007/s00198-012-2065-z
14. Jordan MI, Mitchell TM (2015) Machine learning: trends, perspectives, and prospects. Science 349:255–260
15. Saraswat P (2022) Supervised machine learning algorithm: a review of classification techniques. In: Smart Innovation, Systems and Technologies. pp 477–482
16. Kourou K, Exarchos TP, Exarchos KP et al (2015) Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 13:8–17
17. Chen T, Guestrin C (2016) XGBoost: a scalable tree boosting system. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, New York, NY, USA, pp 785–794
18. Segal MR (2004) Machine learning benchmarks and random forest regression. Biostatistics 1–14
19. Tsujio T, Nakamura H, Terai H et al (2011) Characteristic radiographic or magnetic resonance images of fresh osteoporotic vertebral fractures predicting potential risk for nonunion. Spine (Phila Pa 1976) 36:1229–1235. https://doi.org/10.1097/BRS.0b013e3181f29e8d
20. Kanchiku T, Imajo Y, Suzuki H et al (2013) Usefulness of an early MRI-based classification system for predicting vertebral collapse and pseudoarthrosis after osteoporotic vertebral fractures. J Spinal Disord Tech 27:61–65
21. Sagi O, Rokach L (2018) Ensemble learning a survey Wiley Interdiscip. Rev Data Min Knowl Discov 8(4):1249
22. Misra P, Yadav AS (2020) Improving the classification accuracy using recursive feature elimination with cross-validation. Int J Emerg Technol 11:659–665
23. Gunduz H (2021) An efficient stock market prediction model using hybrid feature reduction method based on variational autoencoders and recursive feature elimination. Financ Innov 7:1–24. https://doi.org/10.1186/s40854-021-00243-3

24. Darst BF, Malecki KC, Engelman CD (2018) Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. BMC Genet 19:65. https://doi.org/10.1186/s12863-018-0633-8

25. Lundberg SM, Erion GG, Lee S-I (2018) consistent individualized feature attribution for tree ensembles

26. Li N, Li B, Gao L (2020) Transient stability assessment of power system based on XGBoost and factorization machine. IEEE Access 8:28403–28414. https://doi.org/10.1109/ACCESS.2020.2969446

27. Šimundić AM (2011) Measures of diagnostic accuracy. pp 13–55

28. ElShawi R, Sherif Y, Al-Mallah M, Sakr S (2021) Interpretability in healthcare: a comparative study of local machine learning interpretability techniques. Comput Intell 37:1633–1650. https://doi.org/10.1111/coin.12410

29. Lee H-C, Yoon S, Yang S-M et al (2018) Prediction of acute kidney injury after liver transplantation: machine learning approaches vs. Logist Regres Model J Clin Med 7:428. https://doi.org/10.3390/jcm7110428

30. Corradi JP, Thompson S, Mather JF et al (2018) Prediction of incident delirium using a random forest classifier. J Med Syst 42:1–10. https://doi.org/10.1007/S10916-018-1109-0/FIGURES/4

31. Sheridan RP, Wang WM, Liaw A et al (2016) Extreme gradient boosting as a method for quantitative structure-activity relationships. J Chem Inf Model 56:2353–2360. https://doi.org/10.1021/acs.jcim.6b00591

32. Ryan G, Magony R, Gortler H et al (2021) Systemically impaired fracture healing in small animal research: a review of fracture repair models. J Orthop Res 39:1359–1367

33. Goodwin VA, Hall AJ, Rogers E, Bethel A (2016) Orthotics and taping in the management of vertebral fractures in people with osteoporosis: a systematic review. BMJ Open 6:e010657. https://doi.org/10.1136/bmjopen-2015-010657

34. Kim HJ, Yi JM, Cho HG et al (2014) Comparative study of the treatment outcomes of osteoporotic compression fractures without neurologic injury using a rigid brace, a soft brace, and no brace: a prospective randomized controlled non-inferiority trial. J Bone Jt Surg Am 96:1959–1966. https://doi.org/10.2106/JBJS.N.00187

35. Kato T, Inose H, Ichimura S et al (2019) Comparison of rigid and soft-brace treatments for acute osteoporotic vertebral compression fracture: a prospective, randomized, multicenter study. J Clin Med 8:198. https://doi.org/10.3390/jcm8020198

36. Minamide A, Maeda T, Yamada H et al (2018) Early versus delayed kyphoplasty for thoracolumbar osteoporotic vertebral fractures: the effect of timing on clinical and radiographic outcomes and subsequent compression fractures. Clin Neurol Neurosurg 173:176–181. https://doi.org/10.1016/j.clineuro.2018.07.019

37. Liska F, Haller B, Voss A et al (2018) Smoking and obesity influence the risk of nonunion in lateral opening wedge, closing wedge and torsional distal femoral osteotomies. Knee Surg, Sport Traumatol Arthrosc 26:2551–2557. https://doi.org/10.1007/s00167-017-4754-9

38. Sloan A, Hussain I, Maqsood M et al (2010) The effects of smoking on fracture healing. Surgeon 8:111–116