Research on Diagnosis Method of Common Knee Diseases Based on Subjective Symptoms and Random Forest Algorithm

Guangjun Wang[1], Mengxia Hu[2], Linlin Lv[3], Hanyuan Zhang[4], Yining Sun[5], Benyue Su[6], Zuchang Ma[7]*
Anhui Province Key Laboratory of Medical Physics and Technology, Institute of Intelligent Machines, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China[1, 4, 5, 7]
Science Island Branch of Graduate School, University of Science and Technology of China, Hefei 230026, China[1, 4, 5, 7]
The University Key Laboratory of Intelligent Perception and Computing of Anhui Province, Anqing Normal University, Anqing 246013, China[1, 2, 3, 6]
Department of Sports Medicine and Arthroscopic Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei 230022, China[4]

Abstract—Knee diseases are common diseases in the elderly, and timely and effective diagnosis of knee diseases is essential for disease treatment and rehabilitation training. In this study, we construct a diagnostic model of common knee diseases based on subjective symptoms and random forest algorithm to realize patients' self-initial diagnosis. In this paper, we first constructed a questionnaire of subjective symptoms of knee, and set up a questionnaire system to guide users to fill out the questionnaire correctly. Then clinical data collection is carried out to obtain clinical questionnaire data. Finally, the diagnostic analysis of three common diseases of knee joint is carried out by random forest machine learning method. Through leave-one-out cross validation, the accuracy of meniscus injury, anterior cruciate ligament injury and knee osteoarthritis diseases are 0.79, 0.84, 0.81 respectively; the sensitivity is 0.79, 0.84, 0.88 respectively; and the specificity is 0.80, 0.84, 0.79 respectively. The results show that the method can achieve a good effect of self-diagnosis, and can provide a knee joint disease screening a convenient and effective approach.

Keywords—Knee diseases; subjective symptoms; random forest algorithm; self-diagnosis

I. Introduction

Knee disease is a highly prevalent condition that has a significant impact on the quality of life, particularly among older individuals. The knee is a complex structure, and as individuals age, the bones, cartilage, ligaments, and other functional components undergo degenerative changes. This gradual degeneration can lead to various knee conditions, including osteoarthritis, meniscus injuries, and ligament injuries [1-2]. Early detection of knee diseases is crucial for effective control and management. However, there is often a lack of awareness about knee diseases, and minor pain and dysfunction are frequently overlooked or attributed to old age or knee osteoarthritis. This delay in recognizing and treating knee diseases results in missed opportunities for prevention and early intervention [3-4].

Knee diseases encompass a wide range of symptoms and dysfunctions, underscoring the importance of recognizing abnormal subjective knee symptoms for early treatment and management. Researchers have explored multiple models and methods based on knee risk factors, symptoms, and etiology to develop diagnostic models for knee diseases and investigate symptom-based identification [5]. For example, Lim et al. employed deep learning algorithms to predict osteoarthritis using a Korean database [6], while Snoeker et al. designed a questionnaire for diagnosing meniscal injuries [7]. Wang Pei's team conducted ordered logistic regression analysis to identify factors influencing knee osteoarthritis grading, ultimately establishing a diagnostic model for grading knee osteoarthritis [8]. Bisson et al. designed a web-based symptom questionnaire with 26 questions and 126 entries to establish a differential diagnosis of knee disorders, providing patients with potential disorder types. However, the tool's sensitivity (58%) and specificity (48%) were found to be insufficient [9].

Despite the promise shown in clinical applications, these studies primarily focus on diagnosing specific types of knee diseases. The identification and diagnosis of different types of knee diseases and the provision of appropriate treatments present greater challenges. In summary, symptom-based screening and differential diagnosis of knee disorders currently face significant hurdles, resulting in low overall identification effectiveness. The subjective nature and ambiguity of symptom definition and acquisition, coupled with limitations in diagnostic models, pose significant difficulties in applying these models to the general population. Factors such as the subjective and variable nature of symptoms, the need for analyzing the relationship between diseases and symptoms, and the suitability of modeling methods all contribute to the complexity of diagnosing knee diseases based solely on subjective symptoms. Addressing these challenges is crucial for improving the performance of diagnostic models and enhancing the accuracy of knee disease identification.

To address the challenges mentioned above, this study proposes a diagnostic model for knee diseases based on subjective symptoms and random forests. The main research work is outlined as follows:

*Corresponding Author

Definition and screening of the knee subjective symptom questionnaire: The researchers define and screen a questionnaire for knee subjective symptoms through data collection, expert validation, and patient assessment methods. A questionnaire assistance system is designed to guide patients in accurately completing the questionnaire.

Data collection and statistical analysis: Clinical questionnaire experiments are conducted to gather research data on subjective knee symptoms. Statistical analysis is performed to examine the relationship between knee diseases and the distribution of major diseases and symptoms. Univariate logistic regression analysis is utilized to filter out irrelevant symptoms, reducing computational complexity and improving the accuracy of the diagnostic model.

Construction of a diagnostic model: A diagnostic model for knee diseases is built based on subjective symptoms, supplemented with the random forest algorithm. The researchers select optimal model parameters through grid search and explore optimal diagnostic thresholds to achieve the best diagnostic performance. The importance of diagnostically significant symptoms for each disease is further investigated to enhance the interpretability of the model.

The research framework for the diagnosis and screening of knee diseases based on subjective symptoms is illustrated in Fig. 1. By implementing this research framework, the aim is to improve the accuracy and interpretability of knee disease diagnosis, providing a valuable tool for early screening and management of knee diseases based on subjective symptoms.

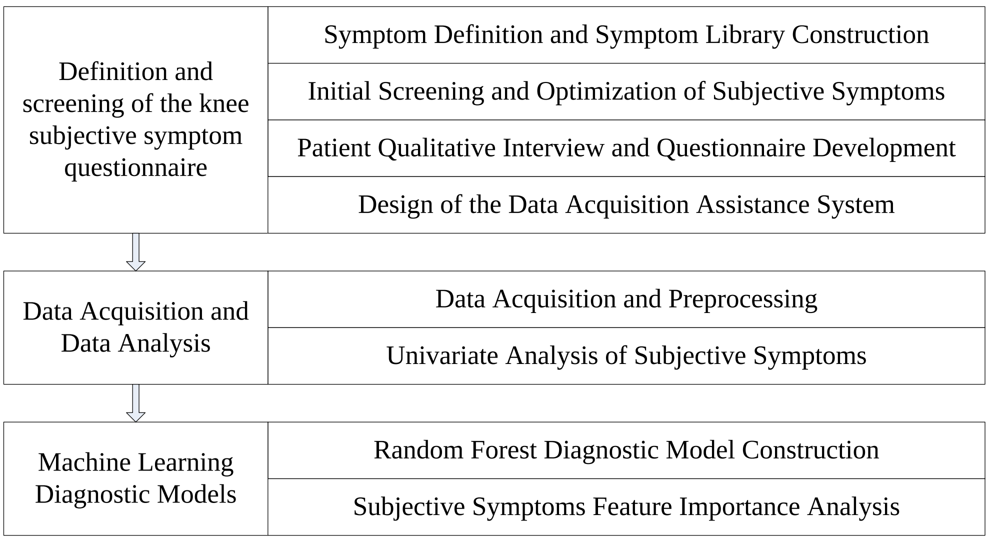

Fig. 1. Research framework for the diagnosis of knee diseases based on subjective symptoms and random forest algorithm.

II. Design of a Subjective Symptom Acquisition System

A. *Subjective Symptom Definition*

When designing the subjective symptom questionnaire, this study adhered to two core principles: firstly, ensuring that the items effectively reflect knee joint diseases and functional status, and secondly, ensuring that the items have clear and easily understandable meanings for the general public. To achieve these goals, the study employed various methods, including data surveys, expert consultations, and qualitative interviews, to construct a comprehensive library of subjective symptom questionnaires.

1) Preliminary definition of symptom item library: Constructing a symptom library that can comprehensively and effectively reflect the characteristics of various knee joint diseases is a challenge due to the subjective and ambiguous nature of symptoms. The research team initially conducted a data survey, gathering relevant literature and books on knee joint diseases to compile a list of possible symptoms. For example, knee osteoarthritis may cause symptoms such as joint pain, morning stiffness, and snapping, while meniscus injuries may result in joint tenderness. Based on this information and by referring to existing joint scales such as WOMAC and KSS, the team constructed a library of subjective symptoms for knee joints. Through discussions with knee-related experts and clinicians, as well as consideration of the clinical complaints of patients, the initial definition of knee joint subjective symptoms was established, and a symptom library was created. This library included medical history, etiology, subjective feelings, functional status, and various specific conditions and degrees, resulting in the design of 104 questionnaire items and 302 subjective symptom options.

2) Optimization of symptom item library Based on expert experience: The initial subjective symptom questionnaire contained a significant amount of redundant information, necessitating screening and optimization. The research team sought expert consultation by inviting 10 experienced doctors to analyze the feasibility, necessity, and comprehensibility of the questionnaire. The experts optimized the item selection and content of the questionnaire by removing redundant and difficult-to-understand symptoms, adding necessary symptoms, and modifying the definitions and explanations of certain symptoms. This process resulted in an optimized screening questionnaire for subjective symptoms, incorporating the suggestions provided by the experts.

3) Further optimization of symptom item library Based on patient feedback: As patients are the ultimate users of the questionnaire, their understanding and applicability are of utmost importance. To address this, the study conducted qualitative interviews with 30 knee joint patients who had been diagnosed with knee joint diseases. Based on the feedback received from the patients, the symptom item library was further optimized. This iterative process led to the finalization of the subjective knee joint symptom questionnaire, which was refined and improved based on patient input. Table I shows the subjective symptoms and corresponding values included in the final questionnaire.

By following this design process, the subjective symptom acquisition system ensures that the knee disease-related symptoms are effectively represented and easily understood by the general population. This system plays a crucial role in accurately capturing subjective symptoms for further analysis and diagnosis of knee diseases.

TABLE I. DEFINITION OF SUBJECTIVE SYMPTOMS OF KNEE DISEASE

Number	Definition of subjective symptoms	value	Number	Definition of subjective symptoms	value
S1	Flexion Limit	False=0 True=1	S16	Pain Activity	False=0 True=1
S2	Extension Limit	False=0 True=1	S17	Pain Rest	False=0 True=1
S3	Snapping	False=0 True=1	S18	Pain Hyperalgesia	False=0 True=1
S4	Locking	False=0 True=1	S19	Pain Hyperflexion	False=0 True=1
S5	Instability	False=0 True=1	S20	Pain Wandering	False=0 True=1
S6	Knee Dislocation	False=0 True=1	S21	Tend Knee Space	False=0 True=1
S7	Patellar Dislocation	False=0 True=1	S22	Tend Above Patella	False=0 True=1
S8	Stiffness	False=0 True=1	S23	Tend Patella	False=0 True=1
S9	Injure	False=0 True=1	S24	Tend Blow Patella	False=0 True=1
S10	Injure Zip	False=0 True=1	S25	Tend Knee Eye	False=0 True=1
S11	Knee Varus	False=0 True=1	S26	Tend Tibial Tubercle	False=0 True=1
S12	Knee Knock	False=0 True=1	S27	Tend LCL	False=0 True=1
S13	Quadriceps Atrophy	False=0 True=1	S28	Tend Iliotibial Band	False=0 True=1
S14	Swelling	False=0 True=1	S29	Tend MCL	False=0 True=1
S15	Pain	False=0 True=1	S30	Tend Popliteal Fossa	False=0 True=1

B. Design of the Data Acquisition Assistance System

To improve the comprehensibility of symptoms and enhance the accuracy of data acquisition, a subjective symptom questionnaire assistance system was designed in this study. The system aims to assist patients in effectively completing the questionnaire. The main components of the system are outlined below:

1) Skip mechanism: The questionnaire incorporates a skip mechanism to reduce redundancy and enhance efficiency. For instance, if a user does not experience pain in the pain section, subsequent pain-related questions will be skipped, eliminating the need to answer irrelevant questions.

2) Tutorial system: The questionnaire incorporates a tutorial system to familiarize users with the questionnaire completion process.

- Overall Tutorial: A brief video, approximately 3 minutes in duration, introduces the questionnaire content to users. Digital media technology, including video animation and interactive media, is employed to help users grasp the questionnaire's content and guide them in providing accurate responses.
- Step-by-Step Tutorial: Complex questions within the questionnaire are accompanied by explanations. Users can click on the question title to access detailed explanations, facilitating their understanding and accurate completion of the questionnaire.

3) 3D Visualization for assisted answering: To address issues of ambiguity and improve the understanding of certain questions, the questionnaire system incorporates 3D visualization. For example, a 3D visualization model is designed for the pressure pain questionnaire. Users can interact with the model to indicate the location of their self-perceived pressure pain. The system automatically fills in the corresponding options based on the user's selection.

The subjective symptom questionnaire assistance system enhances the usability and accuracy of the subjective symptom questionnaire. The skip mechanism reduces redundancy, the tutorial system provides guidance, and the 3D visualization feature facilitates precise responses. These features collectively contribute to more effective and efficient data collection for knee diseases.

III. MATERIALS AND METHODS

A. Random Forest Algorithm

The random forest algorithm is an integrated machine learning algorithm that utilizes decision trees as its base learners. Each decision tree is constructed based on a different subset of the training data and features. As a result, each decision tree is unique and independent, and the collective learning results of multiple decision trees are considered as the final output of the random forest algorithm. This algorithm helps to reduce the variance present in individual decision trees. Random forest is particularly effective in handling classification tasks with complex interactions among attributes. It can also adapt well to datasets with noise or missing values, and the training process is relatively fast. The random forest algorithm is versatile, capable of performing classification, regression, and outlier detection tasks [10-11].

The calculation steps of the random forest algorithm are outlined in Algorithm 1. This algorithm combines the predictions of multiple decision trees to arrive at the final prediction [12].

In the random forest algorithm, each decision tree is trained on a different subset of the data, promoting diversity and reducing overfitting. The final prediction is made by aggregating the predictions of all the decision trees in the ensemble model. This ensemble-based algorithm helps to improve the accuracy and robustness of the model.

Random forest has been widely used in various fields due to its effectiveness and versatility in handling complex datasets. In the context of the study, the random forest algorithm can be employed for tasks such as knee disease classification, regression analysis of symptom severity, or identifying outliers in the dataset.

The calculation steps of the random forest algorithm are shown in Algorithm 1.

Algorithm 1: Random forest algorithm steps

Input: Training set $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, $x_i = (x_i^1, x_i^2, \ldots, x_i^n)$, $x_i^1, x_i^2, \ldots, x_i^n$ is a characteristic of k：

Output: Classification result1

1： $for\ i = 1 : K$

2： A subset is constructed by randomly selecting k samples from the training set $S_i, S_i \subset D$

3： The features contained in each feature are randomly selected from each feature $(m \le n)$

4： $Gini(S_i) = 1 - \sum_J \left(\frac{|y_J|}{|S_i|}\right)^2$ J is the number of features contained in y

5： while(!Generate(T_i))

6： $for\ j = 1 : m$

7： $Gini(S_i, x^j) = \frac{|S_{i1}|}{|S_i|} Gini(S_{i1}) + \frac{|S_{i2}|}{|S_i|} Gini(S_{i2})$

8： $\Delta Gini(x^j) = Gini(S_i) - Gini(S_i, x)$

9： $endfor$

10： The smallest corresponding node is selected as the classification node;

11： end

12： $endfor$

13： The generated individual trees constitute the random forest classifier, which is the final output classification result of the random forest through the voting strategy.

B. Performance Evaluation Indicators

The performance of a disease diagnostic model can be assessed using various indicators, including Accuracy, Sensitivity, Specificity, Receiver Operating Characteristic (ROC) curve, and the Area Under the Curve (AUC) [13]. These indicators provide valuable insights into the model's ability to correctly classify individuals with and without a specific knee disease.

To calculate these indicators, the following four parameters are defined:

True Positives (TP): The number of individuals correctly diagnosed with a particular knee disease.

True Negatives (TN): The number of individuals correctly diagnosed as not having the knee disease.

False Positives (FP): The number of individuals incorrectly diagnosed with the knee disease (false alarms).

False Negatives (FN): The number of individuals incorrectly diagnosed as not having the knee disease when they actually have it (missed diagnoses).

Taking meniscus injury as an example, the classification confusion matrix is shown in Table II:

TABLE II. CLASSIFICATION CONFUSION MATRIX

Diagnostic	Predicted Meniscus Injury	Predicted No Meniscus Injury
Actual Meniscus Injury	TP	FN
Actual No Meniscus Injury	FP	TN

The indicator parameters are defined as follows.

1) The accuracy rate is mainly used to measure the accuracy of the diagnosis of the total sample, i.e., the number of samples correctly diagnosed as a proportion of the total number of samples. The formula for calculating the accuracy rate is as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

2) Sensitivity is the proportion of samples diagnosed as positive out of all positive samples, also known as the true positive rate (TPR), the recall rate. The formula for calculating sensitivity is as follows:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{2}$$

3) Specificity is the proportion of samples with a negative diagnosis to all negative samples, and the formula for calculating specificity is as follows:

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{3}$$

4) False-positive rate is the proportion of incorrectly diagnosed positive samples to all negative samples, and the formula for calculating the false-positive rate is as follows:

$$FPR = \frac{FP}{TN + FP} \tag{4}$$

5) AUC (Area Under Curve, AUC) indicates the area under the ROC (receiver operating characteristic curve, ROC) curve. The formula for calculating AUC is as follows:

$$\text{AUC} = \int_{x=0}^{1} TPR(FPR^{-1}(x))dx \tag{5}$$

where, TPR represents the true positive rate and FPR represents the false positive rate.

6) Youden index, also known as the correct index, is the sum of sensitivity and specificity minus 1. The larger the index, the better the screening test is, and the more truthful it is. The Youden index can be applied when it is assumed that the false-negative (missed diagnosis) and false-positive (misdiagnosis) rates are of equal significance. The method of evaluating the veracity of a screening test indicates the total ability of the screening method to detect true patients versus non-patients.

$$\text{Youden index} = \text{Sensitivity} + \text{Specificity} - 1 \tag{6}$$

C. *Feature Importance*

In order to improve the interpretability of the machine learning model, we calculated and ranked the importance of each symptom in the classification model. Random forest samples datasets from the sample set by bootstrap method with putback, and each dataset   constitutes a decision tree.   is the out-of-bag dataset, i.e., the dataset that was not sampled in when the   dataset was generated. It is used as a test set to compute the importance of each feature using the Random Forest model, and the importance ranking of each feature is obtained by ranking the importance probability of each feature. For example, the steps of feature F importance calculation are as follows. The process of feature importance algorithm is shown in Algorithm 2:

Algorithm 2: Feature Importance Algorithm Steps

1: Using the random forest model based $\overline{D_i}$ calculating the error of each decision tree $\varepsilon = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]$;

2: Randomly upset the order of the features in the sample set and again use the random forest model to calculate to get the error $\bar{\varepsilon} = [\overline{\varepsilon_1}, \overline{\varepsilon_2}, \cdots, \overline{\varepsilon_n}]$, The difference between the two calculation errors is $d = \varepsilon - \bar{\varepsilon}$;

3: Calculate the average of the differences $mean(d)$ and (statistics) standard deviation $\sigma(d)$;

4: The significance of a feature is the ratio of the mean to the standard deviation, i.e., $importance(F) = mean(d)/\sigma(d)$.

IV. Results

A. *Data Collection*

This study utilized the knee subjective symptom questionnaire assistance system for data collection and analysis of knee diseases and symptoms. The collection of clinical data was conducted in adherence to ethical guidelines and approved by the Ethics Committee Board of the Hefei Institute of Materials Research, Chinese Academy of Sciences, following the principles outlined in the Declaration of Helsinki. Informed consent forms were obtained from all participants after providing them with information about the study.

The gold standard for diagnosing knee diseases in the patients was based on assessments using MRI scans and comprehensive evaluations by physicians. The inclusion criteria for participants in the experiment were a definite diagnosis of a specific knee disease through MRI and other necessary examinations, the ability to provide complete clinical examination data and case information, and voluntary participation with an understanding of the nature of the experiment. Exclusion criteria included patients with myocardial infarction, serious infectious diseases, malignant diseases, or cognitive impairment.

The process of data collection and pre-processing involved the following steps: Firstly, patients independently completed the knee subjective symptom questionnaire through the questionnaire system. Subsequently, physicians reviewed each item of the questionnaire with the patients, verifying the diagnostic information related to the type of knee disease, which could include multiple disease types. This ensured the accuracy and validity of the questionnaire data. Finally, the data was exported from the questionnaire system for further analysis.

Data collection primarily took place from January 2021 to April 2021 at the orthopedic outpatient clinic of a tertiary hospital in Hefei, Anhui Province, China. Three orthopedic specialists were involved in acquiring and verifying the questionnaire data. A total of 157 valid data cases were obtained, comprising 76 male and 81 female participants, with no significant difference observed. The average age of the participants was 56±7, predominantly middle-aged and elderly patients.

B. *Univariate Analysis*

In this study, the subjective symptom questionnaire designed for knee disease symptoms is comprehensive, but not all symptoms are directly or strongly correlated with a specific disease. Therefore, it is crucial to filter and remove redundant symptom information to improve the efficiency of disease diagnosis. Symptom screening is an essential step in data preprocessing for data mining and machine learning analysis. Effective symptom selection can significantly enhance the classification performance and overall robustness of the model. The primary objective of symptom selection is to eliminate features with weak or no relevance to the classification target and utilize only the most effective features to achieve accurate classification or prediction results.

TABLE III. THE UNIVARIATE LOGISTIC REGRESSION ANALYSES RESULTS OF THE THREE DISEASES

Diseases	MT			ACL injurie			KOA		
symptom	**B**	**OR (95%CL)**	**P**	**B**	**OR (95%CL)**	**P**	**B**	**OR (95%CL)**	**P**
Flexion Limit	-0.171	0.843(0.447-1.590)	0.598	0.269	1.308(0.646-2.648)	0.455	0.353	1.424(0.723-2.805)	0.307
Extension Limit	0.499	1.648(0.795-3.415)	0.179	0.698	2.009(0.928-4.348)	0.077	-1.148	0.317(0.123-0.818)	**0.017**
Snapping	-0.484	0.616(0.325-1.168)	0.138	-1.305	0.271(0.129-0.570)	**<0.001**	0.990	2.692(1.303-5.563)	**0.007**
Locking	-0.055	0.946(0.454-1.973)	0.883	-0.119	0.888(0.390-2.024)	0.078	0.731	2.077(0.979-4.405)	0.057
Instability	0.622	1.862(0.981-3.536)	0.057	2.522	12.833(4.976-33.101)	**<0.001**	-0.372	0.690(0.348-1.366)	0.286
Knee Dislocation	1.661	5.265(2.315-11.976)	**<0.001**	2.767	15.906(6.522-38.792)	**<0.001**	-1.265	0.282(0.102-0.779)	**0.015**
Patellar Dislocation	1.814	6.138(1.259-29.931)	**0.025**	-1.281	0.278(0.034-2.261)	0.231	-20.510	0.00(0.000-)	0.999
Stiffness	0.321	1.349(0.626-3.035)	0.425	-0.549	0.577(0.219-1.523)	0.267	1.272	3.567(1.582-8.044)	**0.002**
Injure	0.436	1.546(0.783-3.051)	0.209	3.597	36.492(4.855-274.269)	**<0.001**	-1.710	1.181(0.087-0.376)	**<0.001**
Injure Zip	-0.138	0.871(0.398-1.907)	0.729	2.283	9.810(4.140-23.242)	**<0.001**	-1.125	0.325(0.117-0.901)	**0.031**
Knee Varus	-0.379	0.685(0.061-7.712)	0.759	-20.255	0.000(0.000-)	0.999	1.516	4.553(0.403-51.453)	0.220
Knee Knock	-20.950	0.000(0.000-)	0.999	-20.282	0.000(0.000-)	0.999	0.101	1.106(0.196-6.253)	0.909
Quadriceps Atrophy	2.197	9.000(1.057-76.652)	**0.044**	2..908	18.324(2.136-157.203)	**0.008**	-20.480	0.00(0.000-)	0.999
Swelling	0.306	1.359(0.709-2.604)	0.356	1.007	2.737(1.332-5.622)	**0.006**	-0.903	0.405(0.191-0.861)	**0.019**
Pain	-21.671	0.000(0.000-)	0.999	-2.388	0.092(0.018-0.462)	**0.004**	-0.103	0.902(0.216-3.765)	0.887
Pain Activity	-0.993	0.370(0.192-0.715)	**0.003**	-0.194	0.824(0.405-1.675)	0.592	0.120	1.128(0.565-2.249)	0.733
Pain Rest	21.603	>100(0.000-)	0.999	-0.423	0.655(0.071-6.028)	0.708	1.240	3.457(0.559-21.383)	0.182
Pain Hyperalgesia	-0.501	0.606(0.289-1.269)	0.184	0.546	1.726(0.806-3.696)	0.160	-0.671	0.511(0.223-1.174)	0.114
Pain Hyperflexion	-1.145	0.318(0.162-0.626)	**0.001**	-0.593	0.533(0.270-1.131)	0.105	0.102	1.107(0.547-2.243)	0.777
Pain Wandering	0.578	1.783(0.460-6.911)	0.403	-20.310	0.000(0.000-)	0.999	0.605	1.831(0.470-7.139)	0.384
Tend Knee Space	-2.712	0.066(0.025-0.173)	**<0.001**	-0.334	0.716(0.328-1.564)	0.402	0.405	1.500(0.666-3.380)	0.328
Tend Above Patella	-0.796	0.451(0.046-4.438)	0.495	-0.127	0.081(0.089-8.707)	0.914	1.943	6.978(0.707-68.869)	0.096
Tend Patella	0.904	2.469(1.094-5.570)	**0.030**	-1.058	0.347(0.113-1.062)	0.064	0.305	1.356(0.589-3.122)	0.474
Tend Blow Patella	0.087	0.917(0.149-5.646)	0.925	0.590	1.805(0.291-11.192)	0.526	-20.460	0.00(0.000-)	0.999
Tend Knee Eye	0.035	1.036(0.224-4.791)	0.964	-0.847	0.429(0.050-3.667)	0.439	-0.132	0.877(0.164-4.683)	0.878
Tend Tibial Tubercle	21.539	>100(0.000-)	1.000	22.201	>100(0.000-)	1.000	-20.422	0.00(0.000-)	1.000
Tend LCL	-20.893	<0.001(0.000-)	1.000	22.201	>100(0.000-)	1.000	-20.422	0.00(0.000-)	1.000
Tend Iliotibial Band	-20.904	<0.001(0.000-)	0.999	-20.246	0.000(0.000-)	0.999	-20.431	0.00(0.000-)	1.000
Tend MCL	1.055	2.871(0.510-16.163)	0.232	1.748	5.744(1.012-32.594)	**0.048**	-0.846	0.429(0.049-3.775)	0.446
Tend PoplitealFossa	1.034	2.812(0.250-31.685)	0.403	20.255	>100(0.000-)	0.999	-20.441	0.00(0.000-)	0.999

For feature selection, we employed the logistic regression algorithm as the Univariate Analysis method. Univariate logistic regression analysis explores the potential correlation between the dependent variable and each independent variable by establishing a functional relationship between the value of the independent variable and the probability of the occurrence of the event defined by the dependent variable. In this paper, univariate regression analysis was utilized to identify multiple symptoms that can effectively characterize knee disorders. Table III presents the results of univariate logistic regression analyses for three specific diseases: meniscus injury, anterior cruciate ligament injury, and knee osteoarthritis.

Symptoms with a p-value of less than 0.05 were included as significant symptoms of the disease based on the variable inclusion criteria for statistical significance. As can be illustrated in Table III, the significant symptoms of meniscus injury disease include seven symptoms: Knee Dislocation, Patellar Dislocation, Quadriceps Atrophy, Pain Activity, Pain Hyperflexion, Tend Knee Space, and Tend Patella, and the significant symptoms of ACL injury disease include nine symptoms such as Snapping, Instability, Knee Dislocation, Injure, Injury Zip, Quadriceps Atrophy, Swelling, Pain, and Tend MCL. The notable symptoms of knee osteoarthritis include seven symptoms of Extension Limit, Snapping, Knee Dislocation, Stiffness, Injure, Injure Zip, and Swelling. These symptoms differed significantly in the incidence of knee disease, so these factors needed to be screened for inclusion in the subsequent diagnostic modeling index system.

C. *Disease Diagnosis*

The experiment aimed to build a diagnostic screening model for knee diseases based on the significant symptoms identified through univariate analysis of each illness. The performance of the classification model was evaluated using leave-one-out cross-validation, a technique where each data point is used as a test sample once while the remaining samples were used for training the model.

To assess the effectiveness of our proposed method, we compared the diagnostic performance of the optimized random forest-based model with other models, including support vector machine (SVM), logistic regression (LR), AdaBoost (AB), and MLP neural network (Multi-Layer Perceptron, MLP). We evaluated the diagnostic performance regarding various indicators such as AUC, accuracy, sensitivity, specificity, and

Yoden's index. To address data imbalance, we used the SMOTE (Synthetic Minority Over-sampling Technique) method for data augmentation, which helped achieve analyzable results.

Furthermore, to improve the interpretability of the model, we employed a variable importance calculation method specifically designed for the random forest-based model. This method allowed us to calculate and rank the importance of each variable in the model.

The specific results obtained from the experiment are as follows:

1) Model performance analysis: We evaluated the performance of the subjective symptom-based diagnostic model for knee diseases proposed in this study. The model, supplemented with the random forest algorithm, was used for diagnosing meniscus injury, anterior cruciate ligament (ACL) injury, and knee osteoarthritis. Table IV show the diagnostic results of the model for three diseases. The experimental results showed that the AUC values of the diagnostic model for all three diseases were more significant than 0.8. This indicates that the model constructed in this study is suitable for the diagnostic task of common knee diseases. ACL injury showed the highest diagnostic performance among the three diseases, with an AUC value of 0.92. Meniscus injury had an AUC of 0.87, and knee osteoarthritis, after excluding a large amount of concomitant disease data and using SMOTE data augmentation, achieved an AUC of 0.85. Additionally, this study evaluated the comprehensive performance and found that the AUC value showed the best performance among the comprehensive performance of the three disease diagnosis.

These findings demonstrate the effectiveness of the diagnostic screening model for knee diseases based on subjective symptoms. The optimized random forest-based model, along with the use of appropriate machine learning algorithms, showed promising performance in diagnosing common knee diseases. The evaluation metrics such as AUC, accuracy, sensitivity, specificity, and Yoden's index provided comprehensive insights into the model's diagnostic capabilities.

2) Comparison of model performance based on different models: We used support vector machines, logistic regression, AdaBoost algorithm, and MLP neural networks to construct prediction models for comparative analysis. The hyperparameter settings for these four models were carried out in the same manner as the Random Forest model.

A random forest algorithm was used to construct a diagnostic screening model for meniscus injury, anterior cruciate ligament (ACL) injury, and knee osteoarthritis. Four common machine-learning algorithms were used for comparative analysis. Table V presents the accuracy, sensitivity, specificity, and Youden's index for diagnosing the three common knee disorders. Fig. 2 to Fig. 4 display the AUC curves for diagnosing the three disorders.

TABLE IV. THE DIAGNOSTIC RESULTS OF THE MODEL FOR THREE DISEASES

Disease Types	AUC	Accuracy	Sensitivity	Specificity	Threshold	Jordon Index
Meniscus injury	0.87	0.79	0.79	0.80	0.53	0.61
Anterior Cruciate Ligament Injury	0.92	0.84	0.84	0.84	0.30	0.73
Knee osteoarthritis	0.85	0.81	0.88	0.79	0.19	0.68
comprehensive performance	0.88	0.81	0.84	0.81	0.34	0.67

TABLE V. PERFORMANCE OF A DIAGNOSTIC MODEL FOR COMMON KNEE DISEASES IN SELECTING OPTIMAL THRESHOLDS

Disease Types	AUC	Accuracy	Sensitivity	Specificity	Threshold	Jordon Index
Meniscus injury	0.87	0.79	0.79	0.80	0.53	0.61
Anterior Cruciate Ligament Injury	0.92	0.84	0.84	0.84	0.30	0.73
Knee osteoarthritis	0.85	0.81	0.88	0.79	0.19	0.68
comprehensive performance	0.88	0.81	0.84	0.81	0.34	0.67

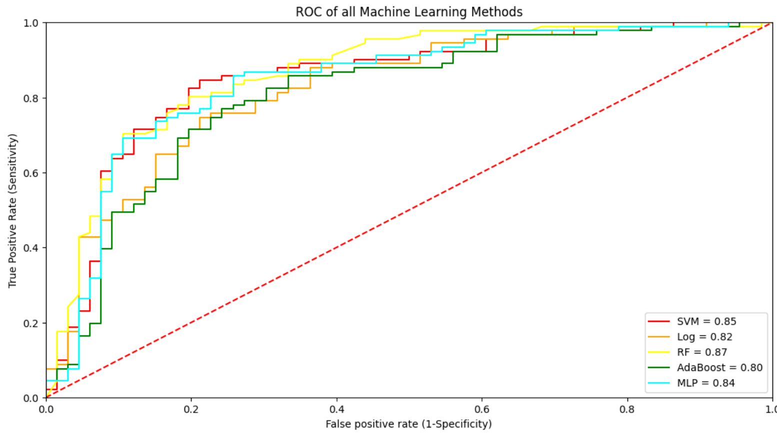

Fig. 2. Machine learning algorithm to meniscus injury disease AUC curve.

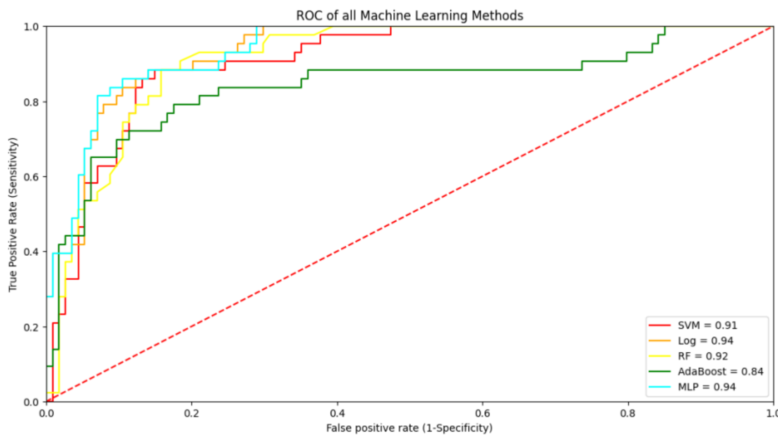

Fig. 3. Anterior cruciate ligament injury disease machine learning algorithm AUC curve.

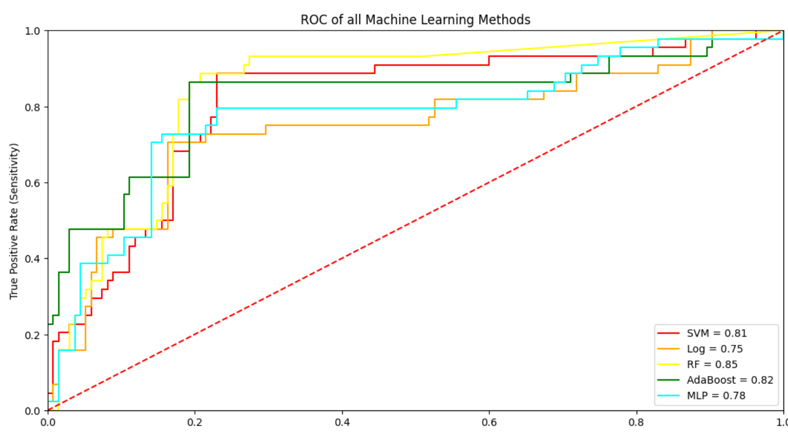

Fig. 4. Knee osteoarthritis disease machine learning methods AUC curve.

Specifically, for the diagnosis of meniscus injury disease, the Random Forest (RF) model achieved the highest AUC of 0.87, followed by the Support Vector Machine (SVM) model at 0.85, the MLP neural network model at 0.84, the Logistic Regression (Log) model at 0.82, and the AdaBoost model at 0.80.

For diagnosing ACL injury disease, the MLP neural network model and Log model performed the best with an AUC of 0.94, followed by the RF model at 0.92, the SVM model at 0.91, and the AdaBoost model at 0.84.

In the knee osteoarthritis disease diagnostic model, the RF model achieved the highest AUC of 0.74 on the test set, followed by the Log model at 0.73, the MLP neural network model and the AdaBoost model, both with an AUC of 0.72, and the SVM model performed slightly worse with an AUC of 0.63. The overall performance of the knee osteoarthritis disease diagnostic model was significantly improved by enhancing the model with SMOTE data, with an AUC of 0.80, accuracy of 0.80, sensitivity of 0.78, and specificity of 0.81.

Compared to the other four classifiers random forest model has higher AUC value and Yoden index, and better comprehensive performance for disease diagnosis. The random forest model has good ability to diagnose and screen common knee joint diseases, and therefore has some practical value, which further indicates that the disease diagnosis and screening based on subjective symptoms proposed in this study is feasible.

3) Symptom feature importance for random forest algorithms: The highest importance scores constructed for different knee diseases varied considerably. The feature importance scores of the random forest algorithm are shown in Fig. 5 to Fig. 7. The first four subjective symptoms were taken as follows: meniscus injury disease with the following order of significance: Tend Knee Space, Knee Dislocation, Pain Hyperflexion, Pain Activity, ACL injury disease with the following order of significance: Knee Dislocation, Instability, Injure Zip, Injure. Significant symptoms of osteoarthritis of the knee and their ranking are as follows: Injure, Stiffness, Snapping and Extension Limit. These are important references for disease analysis.

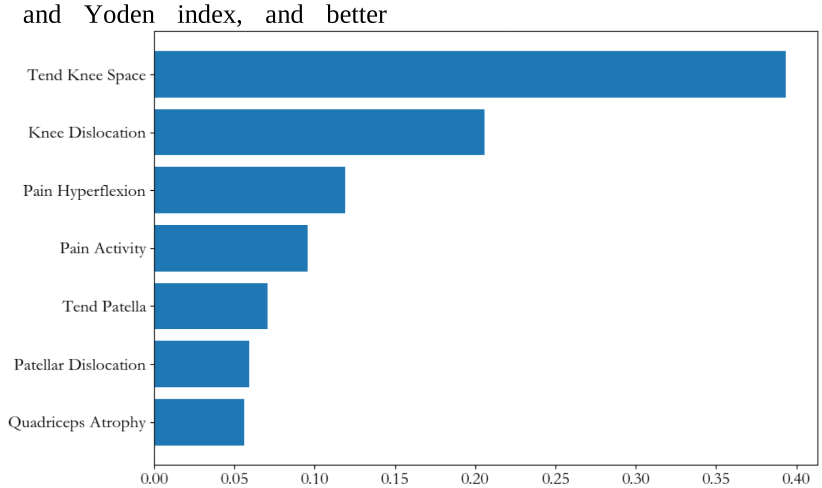

Fig. 5. Ranking the symptom feature importance of diagnostic model for MT injuries.

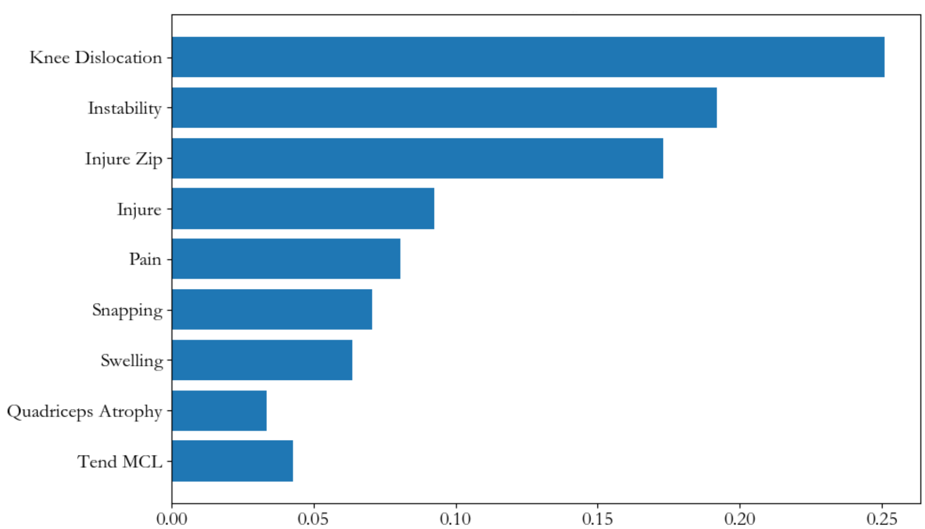

Fig. 6. Ranking the symptom feature importance of diagnostic model for ACL injuries.

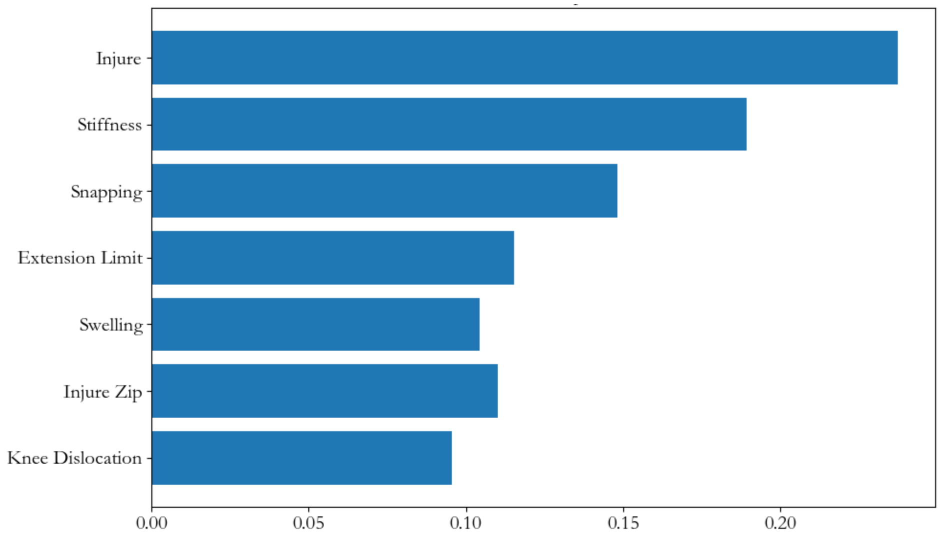

Fig. 7. Ranking the symptom feature importance of diagnostic model for KOA.

V. Discussion

A. Disease Diagnosis Effectiveness Analysis

In this study, a diagnostic screening model for common knee diseases was constructed using a random forest algorithm. The results of the study revealed variations in the diagnostic accuracy for different diseases. ACL injuries were generally better identified, meniscal injuries had slightly lower accuracy, and knee osteoarthritis showed the lowest diagnostic effectiveness. One of the reasons behind the lower diagnostic effectiveness for meniscal injuries and knee osteoarthritis is their strong concurrency.

Pathologically, ACL injuries involving ligamentous structures tend to exhibit more distinct symptoms and a certain degree of specificity. On the other hand, meniscal injuries and knee osteoarthritis are both cartilage diseases and share similar symptoms, making it more challenging to differentiate between them.

When comparing different machine learning methods, the random forest model demonstrated advantages in this study. Random forest is an ensemble learning algorithm that combines multiple decision trees into a single predictive model. It mitigates overfitting issues and enables parallel operation since there are no dependencies between weak learners. Random forests have shown excellent performance in various classical problems, including disease diagnosis. Another integrated learning algorithm, AdaBoost, also performed well in this study. AdaBoost, based on boosting, is particularly effective in handling categorical variables. However, the MLP neural network algorithm, commonly used for unstructured and complex data, was not the most suitable choice for the structured data and categorical features of this study. The random forest andAdaBoost models were more appropriate and demonstrated good performance. Among the compared models, the random forest model generally outperformed the AdaBoost model in this study, supporting previous findings that highlighted the superior classification performance of random forests. Logistic regression, although widely used in the biomedical field due to its simplicity and interpretability, may not have been as effective in addressing the classification problem of this study.

This study achieved better results than the current symptom-based diagnosis of knee disorders. The diagnostic rate for meniscal injuries based on detailed history and clinical examination by doctors is typically around 80% to 85%. The developed diagnostic system in this study showed similar diagnostic accuracy to that of doctors' initial diagnoses, indicating its high clinical value. Table VI provides a comparison of the diagnosis of knee diseases based on symptoms and risk factors.

Overall, the random forest-based diagnostic screening model demonstrated superior performance in identifying knee disorders compared to other machine learning methods and the current symptom-based diagnosis. Removing the interference of concurrent diseases improved the diagnostic accuracy for knee osteoarthritis. These findings highlight the potential of machine learning algorithms in improving disease diagnosis and could have implications for clinical practice.

B. Analysis of Significant Symptoms of Disease

The subjective knee symptoms constructed in this paper can characterize knee disease states with good accuracy for diagnostic prediction through machine learning. Symptoms that characterize the functional state of knee diseases need to meet the needs of diagnosing the effectiveness of the disease, public comprehensibility, and other needs, and there are greater challenges. In this paper, through multiple rounds of research, collation and expert selection and optimization, 30 subjective knee symptoms are finally identified for characterizing knee diseases. Further diagnostic prediction through machine learning algorithms achieves a high correct rate and verifies the validity of symptom definition and selection.

TABLE VI. COMPARISON ON DIAGNOSIS OF KNEE DISEASES BASED ON SYMPTOMS AND RISK FACTORS

	Related Research and Methods	**Results**
Bisson et al.	Constructing a web-based symptom checker for multiple knee disorders to establish a differential diagnosis of knee injuries, allowing patients to determine the correct diagnosis from a checklist.	Overall diagnosis of the disease: 91% sensitivity and 23% specificity; 58% sensitivity and 48% specificity when patients used the tool.
Elkin et al. [14]	To construct a questionnaire-based diagnostic expert system for multiple knee diseases, a Bayesian method was used in model 1 and a heuristic method was used in model 2, and disease importance and term importance weights were added to combine models 1 and 2 to form models 3 and 4.	Accuracy of correct diagnosis in the 1st order in the expert model: model 1: 43.3%, model 2: 43.3%, model 3: 47.8%, model 4: 40.7%.
Lim et al.	Deep learning algorithm was used to predict the diagnosis of osteoarthritis of the knee in the Korean Health and Nutritional Status Database.	Sensitivity 67%, specificity 73%, accuracy 71.97%, AUC 76%.
Ratzlaff et al. [15]	Web-based questionnaire survey to predict diagnosis of knee osteoarthritis and hip osteoarthritis.	For knee osteoarthritis diagnosis: sensitivity 73%, specificity 96%.
Roux et al.	Phone-based questionnaire to predict diagnosis of knee osteoarthritis and hip osteoarthritis.	For the diagnosis of osteoarthritis of the knee: sensitivity 87%, specificity 93%.
Snoeker et al.	A digital-based questionnaire to predict diagnosis of meniscal injury disorders.	For meniscus diagnosis: sensitivity 86.1%, specificity 45.5%, AUC 0.76.
Wang Pei et al.	Establishment of a diagnostic grading model for osteoarthritis of the knee on the basis of data from pathogenic factors, symptoms, signs, physical examination and various scales using logistic regression methods.	Overall accuracy for knee osteoarthritis was 67%, sensitivity 50%, specificity 75%, and AUC of 0.88.
This paper	Differential diagnosis of meniscus injury, anterior cruciate ligament injury and osteoarthritis of the knee by self-developed subjective symptom questionnaire combined with random forest machine learning method for common knee diseases.	Diagnostic performance for MT: an AUC of 0.87, accuracy of 0.79, sensitivity of 0.79 and specificity of 0.80; ACL injury disease: 0.92, 0.84, 0.84 and 0.84; KOA: 0.85, 0.81, 0.88 and 0.79.

Through the clinical questionnaire data collection process, it can be found that patients are able to fill in the questionnaire by themselves through the questionnaire and its auxiliary system, and the detection error rate of the doctor's verification is low, which fully demonstrates that ordinary patients can understand the questionnaire content, and it can be applied to further clinical promotion.

Machine learning algorithms build models with better performance compared to statistical methods but have the disadvantage of poor model interpretability. Machine learning involves learning to train, construct a model and predict new input data. To increase the transparency of the model and provide health education for residents in practical applications, we calculated the impact of each symptom on the performance of the diagnostic model. Through the one-way analysis and diagnostic model importance analysis, significant symptoms can be filtered out for differential diagnosis of knee diseases, and further through the machine learning method and the importance of the calculation of the ranking, to obtain the more important symptoms for each disease, the specific analysis is as follows:

1) The notable symptoms of meniscus injury disease and their symptoms of high importance are compression Tend Knee Space, Knee Dislocation, Pain Hyperflexion and Pain Activity, which are important references for the analysis of the disease. Since the meniscus is present in the knee space, pressure pain in the knee space is a prominent symptom in the diagnosis of meniscus injury [16]. The meniscus is the role of the spacer for the knee activity to form a cushioning effect; once the injury lesion, the knee activity will be exacerbated by pain, so Pain Hyperflexion is increased, Pain Activity is increased, and other symptoms [17]. Generally, meniscal disease does not cause symptoms of dislocation sensation. In this study, there existed a large number of symptoms of meniscus in combination with ACL, and dislocation sensation was a significant symptom of ACL injury and thus was included among the significant symptoms of meniscus, which is a reflection of the complexity of knee disease.

2) The notable symptoms of ACL injury disease and its symptoms of higher importance are Knee Dislocation, Instability, Injure Zip, and Injure, respectively. ACL injury disease indicates the presence of damage and rupture of the ACL. The anterior cruciate ligament (ACL) is located in the knee, connecting the femur and tibia, and its main role is to limit the tibia's forward shift. It is an important static and kinetic anterior stabilizing structure of the knee, which prevents the tibia's anterior shift in flexion, prevents the knee from over-extending in extension, controls knee rotation, and controls knee internal and external rotation at different flexion angles, and has a proprioceptive function [18]. When an injury rupture of the ACL occurs, there is a noticeable sense of dislocation, which later develops into a sense of instability. The main cause of ACL rupture is injury, which accounts for more than 70% of the cases. Therefore, the presence of a history of trauma in the patient is the main causative factor in the development of the disease [19]. A tearing sound accompanies ligament tear injuries. Further intra-articular hemorrhage leads to swelling, pain, and, in most cases, inability to continue with the original sport or even limited extension and hyperflexion activities.

3) The prominent symptoms of osteoarthritis disease of the knee and their higher importance are Stiffness, Snapping, Injure, and Swelling, respectively. This is almost identical to the clinical diagnostic criteria for the knee. Clinical diagnostic criteria for knee osteoarthritis usually consider ① knee pain, ② snapping, ③ morning stiffness time$\leq$30min, ④ age $\geq$38 years old, ⑤ bony enlargement, Osteoarthritis of the knee is diagnosed if①②③④, or ①②⑤ or ①④⑤ are fulfilled [20]. From the perspective of subjective symptoms, pain was the basic symptom, while joint friction sound (sensation), morning stiffness and bony enlargement were important for clinical diagnosis, which verified the reliability of the subjective and significant symptom analysis in this study.

From the distribution of disease symptoms, it can be found that pain is the first complaint of knee diseases, and in this questionnaire case, different diseases caused pain symptoms almost 100%. Therefore, pain symptoms are the basic symptoms but are not significant for the differential diagnosis of the disease.

Stiffness is an important symptomatic feature of knee osteoarthritis. Patients may experience stiffness in the joints in the morning or after rest, which can be relieved by activity due to verification of cartilage or joint adhesions.

Snapping due to cartilage destruction and rough joint surface, bone friction sound (sense) occurs during joint movement.

Injure is an important symptom in the history of knee disease and an important factor in the development of Osteoarthritis of the knee. Everyday knee injuries are prone to cause cartilage damage or lesions, which gradually form Osteoarthritis of the knee.

Swelling symptoms in this study were generally compared to the difference in the knee compared to the healthy side or the previous one, so bony enlargement symptoms were included. As Osteoarthritis of the knee progresses, patients experience swelling and bony enlargement of the knee. Swelling is caused by fluid and cells collecting around the joint due to inflammation, resulting in swelling, pain, and warmth in the joint area. The swelling may get worse as the inflammation increases. On the other hand, bone enlargement is caused by damage to the cartilage in the knee. The cartilage loses its ability to protect the articular bones, causing them to be exposed to friction and wear and tear. Over time, the articular bones regrow, forming bone spurs and osteophytes. These bony protrusions cause pain and stiffness when the joint moves.

C. *Research Limitations*

This research also has some limitations.

Firstly, the sample size used for constructing the diagnostic model was relatively small, and the uneven distribution of positive and negative cases may have affected the model's training and analysis of the disease. Therefore, the model may not directly apply to clinical diagnosis, but it can still be valuable for self-screening common knee diseases.

Secondly, the data used in the study were collected from a specific region and population in Anhui province, which may limit the generalizability of the findings to other populations. Regional differences and variations in disease duration should be considered when applying the model to different populations.

Lastly, the subjective questionnaire employed in the study mainly utilized dichotomous data and lacked more detailed information regarding symptom typing and severity. Incorporating more comprehensive and detailed symptomatic details in future studies could enhance the diagnosis of a wider range of diseases and provide a better understanding of disease severity.

These limitations should be addressed in future research to improve the diagnostic accuracy and applicability of the model.

VI. Conclusion

In this study, we designed a subjective symptom questionnaire for knee diseases based on easily collected subjective symptoms. We collected clinical data to analyze the relationship between diseases and symptoms. By combining univariate logistic regression analysis and random forest, we developed a diagnostic screening method for common knee diseases using these subjective symptoms. The study demonstrated promising diagnostic performance for the examined common knee diseases.

The area under the curve value was 0.87 for meniscal disease. For ACL injury, the AUC value was 0.92; for knee osteoarthritis, the AUC value was 0.84. Additionally, accuracy, sensitivity, and specificity values were reported for each disease, indicating favorable performance compared to similar studies and proximity to the clinical diagnosis results of physicians. These findings suggest that the proposed method, which utilizes subjective symptoms, holds advantages in screening common knee diseases and may be applicable for self-diagnosing these conditions.

Furthermore, the study presents a general framework for utilizing machine learning methods to predict the risk of developing other chronic diseases. However, it is important to note that further research and validation are necessary to ensure the robustness and generalizability of the proposed diagnostic method.

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability Statement: All data underlying the findings are fully available without restriction. All relevant data are within the paper and appendices.

Fund Project: This research was funded by National Key R&D Program Project, Demonstration of Integrated Application of Community Science and Fitness in Central Region （2022YFC2010200）, the Key Project on Anhui Provincial Natural Science Study by Colleges and Universities under Grant “Key technical research of knee function evaluation and rehabilitation training”(No. KJ2019A0555), Key project of Science and Technology Service Network Program of Chinese Academy of Sciences (No. KFJ-STS-ZDTP-079), Hefei Natural Science Foundation Project (No. 202302), Medical Artificial Intelligence Joint Fund Project (MAI2023 C008). Furthermore, it was supported by key projects of natural science research in Anhui Province's higher education institutions (No. KJ2021A1436), (No. 2022AH052604).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

References

[1] Okamoto S, Ishii Y, Ishikawa M, et al. The effect of gait modification on the response of medial meniscus extrusion during gait in patients with knee osteoarthritis[J]. Gait & posture, 2023, 102: 180-185.

[2] Hutchison L, Grayson J, Hiller C, et al. Relationship Between Knee Biomechanics and Pain in People with Knee Osteoarthritis: A Systematic Review and Meta-Analysis[J]. Arthritis care & research, 2023,75(6):1351-1361.

[3] Hanna M, Jeffrey N K. The influence of meniscal pathology in the incidence of knee osteoarthritis: a review[J]. Skeletal radiology, 2023,52(11).:2045-2055.

[4] Adams B G, Houston M N, Cameron K L. The Epidemiology of Meniscus Injury[J]. Sports Medicine and Arthroscopy Review, 2021, 29(3):24-33.

[5] Qin S, Chi Z, Xiao Y, et al. Effectiveness and safety of massage for knee osteoarthritis: A protocol for systematic review and meta-analysis[J]. Medicine, 2020, 99(44): e22853.

[6] Lim J, Kim J, Cheon S. A deep neural network-based method for early detection of osteoarthritis using statistical data[J]. International journal of environmental research and public health, 2019, 16(7): 1281.

[7] Snoeker B A M, Zwinderman A H, Lucas C, et al. A clinical prediction rule for meniscal tears in primary care: development and internal validation using a multicentre study[J]. British Journal of General Practice, 2015, 65(637): e523-e529.

[8] Wang P, Zhang X, Gao Y, et al. Establishment of grading model of knee osteoarthritis based on clinical research system[J]. China J Orthop Trauma, 2018, 31(6):528-533.

[9] Bisson L J, Komm J T, Bernas G A, et al. How accurate are patients at diagnosing the cause of their knee pain with the help of a web-based symptom checker? [J]. Orthopaedic journal of sports medicine, 2016, 4(2): 2325967116630286.

[10] Cui M, Gang X, Gao F, et al. Risk assessment of sarcopenia in patients with type 2 diabetes mellitus using data mining methods[J]. Frontiers in endocrinology, 2020, 11:123.

[11] Chiew C J, Liu N, Wong T H, et al. Utilizing machine learning methods for preoperative prediction of postsurgical mortality and intensive care unit admission[J]. Annals of surgery, 2020, 272(6):1133-1139.

[12] Zhao H, Zhang X, Xu Y, et al. Predicting the Risk of Hypertension Based on Several Easy-to-Collect Risk Factors: A Machine Learning Method[J]. Frontiers in Public Health, 2021, 9:619429. DOI: 10.3389/fpubh.2021.619429.

[13] Moitra D, Mandal R K. Prediction of Non-small Cell Lung Cancer Histology by a Deep Ensemble of Convolutional and Bidirectional Recurrent Neural Network[J]. Journal of Digital Imaging, 2020, 33(9): 895-902.

[14] Elkin P L, Schlegel D R, Anderson M, et al. Artificial intelligence: bayesian versus heuristic method for diagnostic decision support[J]. Applied clinical informatics, 2018, 9(02): 432-439.

[15] Ratzlaff C, Guermazi A, Collins J, et al. A rapid, novel method of volumetric assessment of MRI-detected subchondral bone marrow lesions in knee osteoarthritis[J]. Osteoarthritis and cartilage, 2013, 21(6): 806-814.

[16] Smith J D, Johnson A B, Wilson C D. Accuracy of McMurray's test, the modified version, and joint-line tenderness in diagnosing chronic meniscal tear in the knee joint: A cross-sectional study. Journal of Orthopedic Research, 2023, 41(2), 123-135.

[17] Garvick S J, Reich S. Meniscal tears[J]. JAAPA: official journal of the American Academy of Physician Assistants, 2020, 33(1):45-46.

[18] Philipp F, Constantin D, Lucca L, et al. Repair of the lateral posterior meniscal root improves stability in an ACL-deficient knee[J]. Knee surgery, sports traumatology, arthroscopy: official journal of the ESSKA, 2018, 26(8):2302-2309.

[19] Rigg J, Perera N, Toohey L, et al. ACL injury occurrence in the Australian High Performance Sports System: a 5-year retrospective analysis[J]. Journal of Science and Medicine in Sport, 2022, 25(Supplement 2):45-46.

[20] Guo J, Chen P, Cai T, et al. Dysfunction Characteristics of Knee Osteoarthritis with the Conception of"Muscles and Bones,Arthralgia and Flaccidity,Asthenia and Sthenia,Dynamic and Static,Hardness and Softness"[J]. Rehabilitation Medicine, 2021, 31(1):69-72.