Fall Risk Prediction Model for Chinese Community-Dwelling Older Adults Aged 60+

Zhihao Jin [1,*], Xinghua Ma [2]

[1] Department of Statistics and Finance, University of Science and Technology of China, Hefei, China

[2] Zhujiang College, South China Agricultural University, Guangzhou, China

* Corresponding Author Email: jinzhihao@mail.ustc.edu.cn

Abstract. Background: Falls among community-dwelling Chinese older adults are a significant public health issue. This study aimed to develop a predictive model for fall risk in individuals aged ≥60 years. Methods: We analyzed data from four waves of the China Health and Retirement Longitudinal Study (CHARLS, 2011-2018) to assess fall status and associated factors. Socio-demographic, socio-economic, and emotional variables were summarized using descriptive statistics. Multiple machine-learning methods (Random Forest, Boruta, XGBoost, and ABESS) identified key predictors, which were incorporated into a logistic regression model. Model performance was evaluated via the AUC, and the optimal model was visualized with a nomogram. Results: Sex, history of chronic diseases, hearing status, sleep duration, self-reported pain, and depression score were significantly associated with fall risk. The model demonstrated good predictive accuracy (AUC=0.685,95% CI: 0.650-0.721). Conclusion: Our predictive model provides an effective tool for early identification of high-risk elderly individuals, facilitating personalized interventions to reduce falls and enhance quality of life.

Keywords: Elderly, Falls, Risk Assessment, Prediction Model.

1. Introduction

As global populations age, the proportion of elderly people is expected to increase from 10% to 16%, meaning that by 2050, approximately one in every six people will be aged 65 or older [1]. This trend presents a range of health challenges, with falls becoming a major cause of accidental death and injury among the elderly [2]. It is estimated that approximately one-third of individuals aged 65 and older experience at least one fall each year, with this rate increasing as age advances [3]. Falls not only pose a threat to the physical health of the elderly but can also lead to psychological trauma and significant economic burden [4]. In China, there are approximately 20 million elderly individuals, with around 25 million fall incidents occurring annually. The direct medical costs are approximately 5 billion RMB, and the societal cost is estimated to be between 60 to 80 billion RMB [5].

The risk factors for falls are diverse, including intrinsic factors (such as muscle weakness and neurological defects) and extrinsic factors (such as inadequate lighting and inappropriate footwear) [5]. While clinical assessment methods such as the Berg Balance Scale and Tinetti Test have been shown to be effective, they rely on subjective judgment by healthcare professionals, which may lead to inconsistencies [7, 10-13]. Additionally, the limited availability of professional resources makes continuous fall risk assessment difficult.

In recent years, sensor-based technologies have been widely explored for fall risk assessment, including environmental sensing technologies, optical motion capture systems, and wearable sensors [14-17]. These technologies can provide objective and quantitative metrics, but many require laboratory settings for use, limiting their applicability in clinical practice. To better prevent falls, the Global Steering Committee has proposed the P4 model (Predict, Personalize, Prevent, and Participate), which aims to optimize fall prevention care for the elderly through personalized and predictive approaches. However, implementing this model faces several challenges, including the low execution rate of fall risk screening by physicians and time constraints [19, 20]. Technological advances offer the potential to overcome these barriers, particularly with the development of wearable sensor technology that can capture and analyze quantitative movement data, thereby improving fall risk assessment.

This study selects demographic and fall-related risk factors to provide fall risk assessment and prevention recommendations for elderly patients. By integrating physical activity, hearing assessment and treatment, and environmental evaluations and modifications, a fall prediction model for community-dwelling elderly is developed to reduce fall incidence, improve the quality of life for elderly individuals, and alleviate the economic burden on healthcare systems.

2. Materials and Methods

2.1. Data Source

This study used data from the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative longitudinal survey collecting health and socioeconomic data from individuals aged 45 and above in China. CHARLS applies a multistage probability sampling strategy, providing detailed socioeconomic, health, and demographic information to support aging-related research. The CHARLS dataset includes baseline data collected in 2011 and three follow-up waves in 2013, 2015, and 2018. The baseline survey (2011–2012) covered 10,257 households and 17,708 individuals across 450 communities in 150 counties from 28 provinces. For this study, the cohort comprised individuals aged 60 years or older, including those aged 60+ at baseline and those reaching this age in subsequent waves, ensuring comprehensive coverage of elderly populations.

2.2. Selection and Interpretation of Study Variables

2.2.1 Fall Status and Assessment

CHARLS collected fall data through the baseline survey by asking participants, "Have you fallen in the past two years?" and in follow-up surveys, "Have you fallen since the last visit?" This data was used to assess fall incidence among elderly individuals. The study population consisted of elderly individuals living in communities. The exclusion criteria for this study were: (1) Missing fall data: Studies with incomplete key data on fall events (e.g., details of each fall, hospitalization, or injury severity) were excluded to ensure data completeness; (2) Participants with specific medical diagnoses (e.g., dementia, Parkinson's disease, or stroke): These populations may require specialized approaches for fall risk assessment and intervention, making them non-representative of the general elderly population; (3) Severe physical disabilities (e.g., wheelchair users, leg disabilities, muscle atrophy, paralysis requiring round-the-clock care); (4) Severe sensory impairments (e.g., blindness or severe vision impairment, complete hearing loss); (5) Individuals living in specialized environments (e.g., long-term care facilities): The fall risk factors in these settings differ from those of community-dwelling elderly; (6) Specific types: (i) Participants who have received fall-related health education, (ii) Participants who have undergone orthopedic surgeries (e.g., spinal, knee, or hip surgeries), (iii) Participants receiving nutritional treatment (e.g., calcium or vitamin supplementation), (iv) Participants with urinary incontinence.

2.2.2 Factors Associated with Fall Risk

We selected 21 relevant fall-related factors from the CHARLS dataset based on extensive literature reviews of studies conducted across Europe, the United States, and India. These factors encompass various dimensions, including demographics, sensory function, health behaviors, health status, functional status, psychological status, and social and economic status. Detailed information about these factors can be found in Table S1 in the appendix.

2.3. Variable Selection

In model development, researchers typically select various independent variables based on theoretical considerations and expert experience, often leading to the inclusion of variables with minimal or no predictive value. Such redundant variables can increase computational complexity, reduce model accuracy, and elevate data collection costs, especially if they are resource-intensive to measure. To address these challenges, this study employed four advanced variable selection methods—Random Forest (RF), Boruta, XGBoost, and ABESS—to rigorously screen the 21 candidate variables, identifying the most relevant factors for inclusion in the predictive model [21-25].

2.4. Model Construction

After identifying significant fall-related factors using Random Forest (RF), Boruta, XGBoost, and ABESS methods, the dataset with imputed missing values was randomly divided into a training set (70%) and a validation set (30%). Six machine learning algorithms—Naive Bayes (Bayes), Multi-layer Perceptron (MLP), K-Nearest Neighbor (KNN), Logistic Regression (LR), Random Forest (RF), and XGBoost—were applied to construct predictive models for fall risk among elderly individuals [24-29].

2.5. Model Evaluation

The predictive performance of the logistic regression-based fall risk model, developed from variables selected through Random Forest (RF), Boruta, XGBoost, and ABESS methods, was evaluated using statistical metrics, including goodness-of-fit and mean squared error (MSE). The area under the receiver operating characteristic (ROC) curve (AUC) was calculated to assess the model's discrimination ability. Additionally, to facilitate clinical interpretation and personalized risk assessment, a nomogram was constructed based on the identified significant predictors using the rms package in R software. The nomogram was further validated through multi-center data, ensuring its reliability and practical applicability.

3. Results

3.1. Data Inclusion

Following the inclusion and exclusion criteria, data from four waves of the CHARLS survey (baseline in 2011, and follow-up surveys in 2013, 2015, and 2018) were analyzed. After these four rounds, a total of 5,985 participants aged 60 years and above were ultimately included in the final dataset. Detailed participant characteristics and the data inclusion process are illustrated in Figure 1.

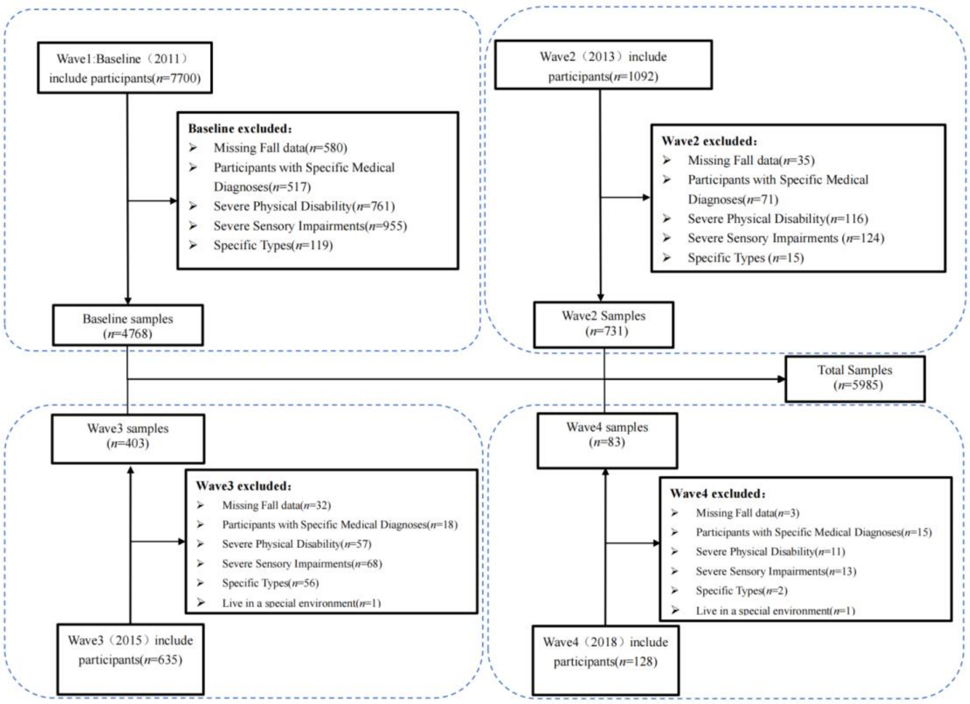

Figure 1. Detailed Participant Characteristics and Data Inclusion Process

3.2. Demographic Descriptive Statistics

A total of 5,985 community-dwelling elderly individuals aged 60 and above were included in the study. Demographic and health characteristics are summarized in Table 1. Among the participants, the median age was 66 years (IQR: 62–71), with 49.42% males and 50.58% females. Over half (50.83%) had an education level below primary school. Geographically, participants were mainly from the eastern (32.40%) and western (31.45%) regions, with rural residents accounting for 52.97%. Regarding health status, 72.93% reported having chronic diseases, 27.80% experienced pain, and 55.92% did not engage in regular physical exercise. Most participants rated their health status as "average" (52.65%). In terms of daily living ability, 14.55% showed some degree of dependence, while 85.45% were completely independent. Sleep durations between 6.5 to 8.5 hours were reported by 37.23% of participants. Depression (≥11 points) was present in 29.32% of the elderly population. Fall history indicated that 16.06% of participants had experienced at least one fall. Additionally, 93.15% had health insurance, 79.13% lived with their spouses, and 66.67% reported receiving minimal intergenerational financial support.

Table 1. Demographic and Clinic-pathologic Variables of the datasets

Variable Name		60+ Elderly (Total N = 5985)	
		Frequency (N)	Percentage (%)
Age		66 (62, 71)	
Sex	Male	2958	49.42
	Female	3027	50.58
Education Level	Below Primary School	3042	50.83
	Primary School	1590	26.57
	Junior High School	822	13.73
	High School or Above	531	8.87
Region	Northeast	421	7.03
	Western	1882	31.45
	Central	1743	29.12
	Eastern	1939	32.40
Residence	Urban	2815	47.03
	Rural	3170	52.97
Myopia Status	Poor	1224	20.45
	Fair	2725	45.53
	Good	1451	24.24
	Very Good	524	8.76
	Excellent	61	1.02
Hyperopia Status	Poor	1342	22.42
	Fair	2510	41.94
	Good	1372	22.92
	Very Good	670	11.19
	Excellent	91	1.53
Hearing Status	Poor	643	10.74
	Fair	2716	45.38
	Good	1807	30.19
	Very Good	739	12.35
	Excellent	80	1.34
Sleep Duration	≤4.5h	1016	16.97
	≤6.5h	2280	38.10
	≤8.5h	2228	37.23
	>8.5h	461	7.70

Exercise	No	3347	55.92
	Yes	2638	44.08
Chronic Disease	No	1620	27.07
	Yes	4365	72.93
Chronic	0	1620	27.07
	1	2118	35.39
	2	1193	19.93
	≥3	1054	17.61
Self-reported Pain	No	4321	72.20
	Yes	1664	27.80
Self-rated Health	Very Poor	189	3.16
	Poor	1252	20.92
	Average	3151	52.65
	Good	995	16.62
	Very Good	398	6.65
Life Satisfaction	Not Satisfied at All	105	1.75
	Not Very Satisfied	508	8.49
	Somewhat Satisfied	3975	66.42
	Very Satisfied	1235	20.63
	Extremely Satisfied	162	2.71
Activities of Daily Living (ADL) Score	0(Complete Dependence)	9	0.15
	1(Dependence in 5 activity)	25	0.42
	2(Dependence in 4 activity)	38	0.63
	3(Dependence in 3 activity)	76	1.27
	4(Dependence in 2 activity)	178	2.97
	5(Dependence in 1 activity)	545	9.11
	6(Completely Independent)	5114	85.45
ADL Categorization	Dependent (score 0-5)	871	14.55
	Completely Independent (score 6)	5114	85.45
Marital Status	Married with Spouse	4736	79.13
	Other	1249	20.87
Health Insurance	No	410	6.85
	Yes	5575	93.15
Intergenerational Economic Support	Very Little	3990	66.67
	Moderate	1289	21.53
	Quite a Lot	706	11.80
Depression Score		7 (3, 12)	
Falls	No	5024	83.94
	Yes	961	16.06

3.3. Variable Selection and Modeling Results

3.3.1 Variable Selection

Based on additional screening of the 21 candidate variables using Random Forest (RF), Boruta, XGBoost, and ABESS methods, six optimal predictors were ultimately identified: sex, hearing status, sleep duration, self-reported pain, chronic diseases, and depression score. Figure 2 presents a Venn diagram illustrating the overlap and consensus of selected variables among the four methods.

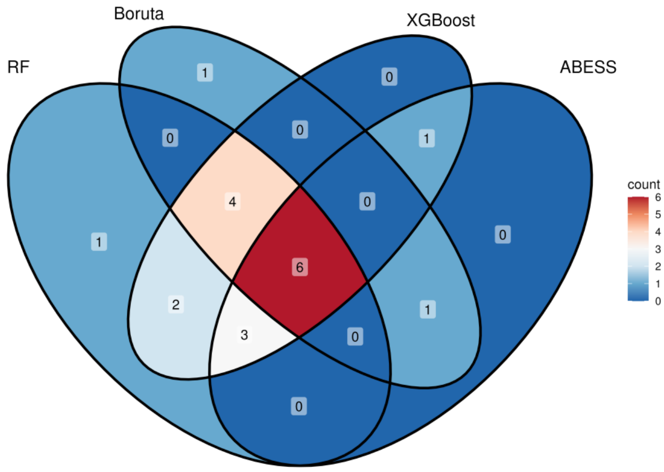

Figure 2. Venn diagram illustrating variable selection results using RF, Boruta, XGBoost, and ABESS methods

3.3.2 Modeling Results

The results of the logistic regression analysis are summarized in Table 2. Multiple factors significantly influenced fall risk among elderly individuals. Poorer hearing was strongly associated with increased fall risk, with significantly lower odds observed among individuals reporting "fair" (OR=0.377, 95% CI: 0.254–0.560, *P*<0.001), "good" (OR=0.370, 95% CI: 0.263–0.521, *P*<0.001), "very good" (OR=0.252, 95% CI: 0.175–0.362, *P*<0.001), and "excellent" (OR=0.227, 95% CI: 0.150–0.345, *P*<0.001) hearing compared to those with "poor" hearing status.

Table 2. Logistic Regression Analysis Results

Variable		Regression Coefficient	Wald Statistic	*P-value*	OR (Odds Ratio)	OR 95% CI	
						Lower	Upper
Sex		-0.344	24.952	<0.001	0.709	0.612	0.821
Hearing Status	Poor		67.375	<0.001			
	Fair	-0.976	23.397	<0.001	0.377	0.254	0.560
	Good	-0.994	32.462	<0.001	0.370	0.263	0.521
	Very Good	-1.378	55.417	<0.001	0.252	0.175	0.362
	Excellent	-1.482	48.460	<0.001	0.227	0.150	0.345
Sleep Duration	≤4.5h		21.632	<0.001			
	≤6.5h	-0.143	0.733	0.392	0.866	0.624	1.203
	≤8.5h	-0.232	2.410	0.121	0.793	0.592	1.063
	>8.5h	-0.587	14.442	<0.001	0.556	0.411	0.753
Chronic Disease	0		12.286	0.006			
	1	-0.042	3.502	0.061	0.959	0.791	1.165
	2	0.200	7.851	0.005	1.221	0.988	1.508
	≥3	0.275	0.001	0.012	1.316	1.061	1.633
Pain Status		0.489	29.697	<0.001	1.630	1.388	1.912
Depression Score		0.030	6.149	0.013	1.030	1.017	1.043

Sleep duration also impacted fall risk significantly, with elderly individuals sleeping more than 8.5 hours showing a lower risk of falling (OR=0.556, 95% CI: 0.411–0.753, *P*<0.001) compared to those sleeping ≤4.5 hours.

The presence and number of chronic diseases were significant predictors. Individuals with two chronic diseases had a significantly higher risk of falling compared to those without chronic diseases (OR=1.221, 95% CI: 0.988–1.508, *P*=0.005), whereas having three or more chronic diseases did not show a statistically significant difference.

Additionally, self-reported pain significantly increased fall risk (OR=1.630, 95% CI: 1.388–1.912, *P*<0.001). Higher depression scores were associated with slightly elevated fall risk (OR=1.030, 95% CI: 1.017–1.043, *P*=0.013).

Overall, these findings suggest that hearing status, sleep duration, chronic disease status, pain, and depression significantly contribute to fall risk among elderly individuals, highlighting essential indicators for targeted fall prevention strategies.

3.4. Nomogram and Calibration Curve

Based on the logistic regression analysis results, the nomogram in Figure 3 visualizes the relationship between multiple predictor factors (such as sex, hearing status, sleep duration, chronic diseases, self-reported pain and depression score) and fall risk among elderly individuals. Factors such as history of chronic diseases and depression scores have a significant positive impact on fall risk. As the presence of chronic diseases and depression scores rise, the risk of falling significantly increases. Individuals with poorer hearing condition and those with shorter sleep duration (e.g., ≤4.5 hours) have a higher risk of falling, while those with longer sleep durations (>8.5 hours) exhibit a lower fall risk. Elderly individuals who report pain have a higher fall risk. By accumulating the scores of each variable, the nomogram can be used to calculate an individual's fall risk, providing a scientific basis for the development of personalized preventive interventions.

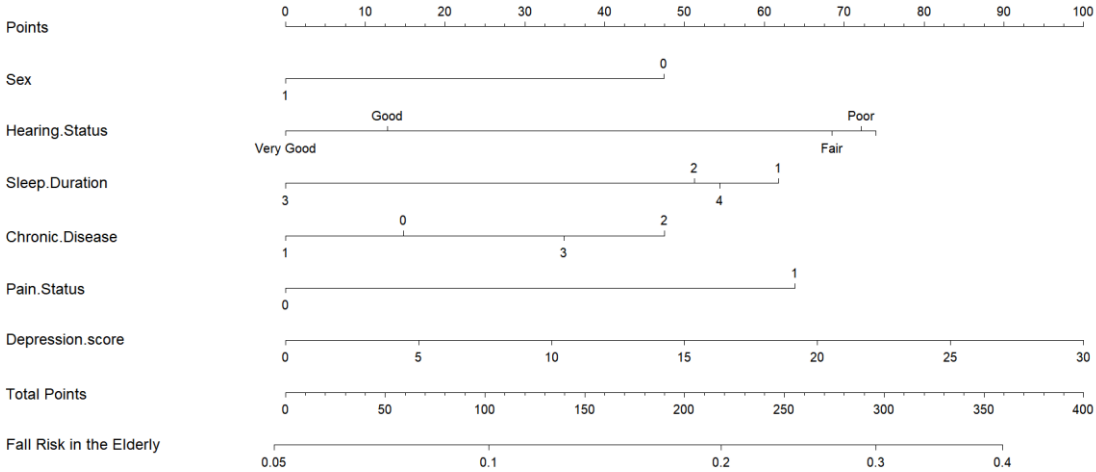

Figure 3. Logistic Regression Nomogram

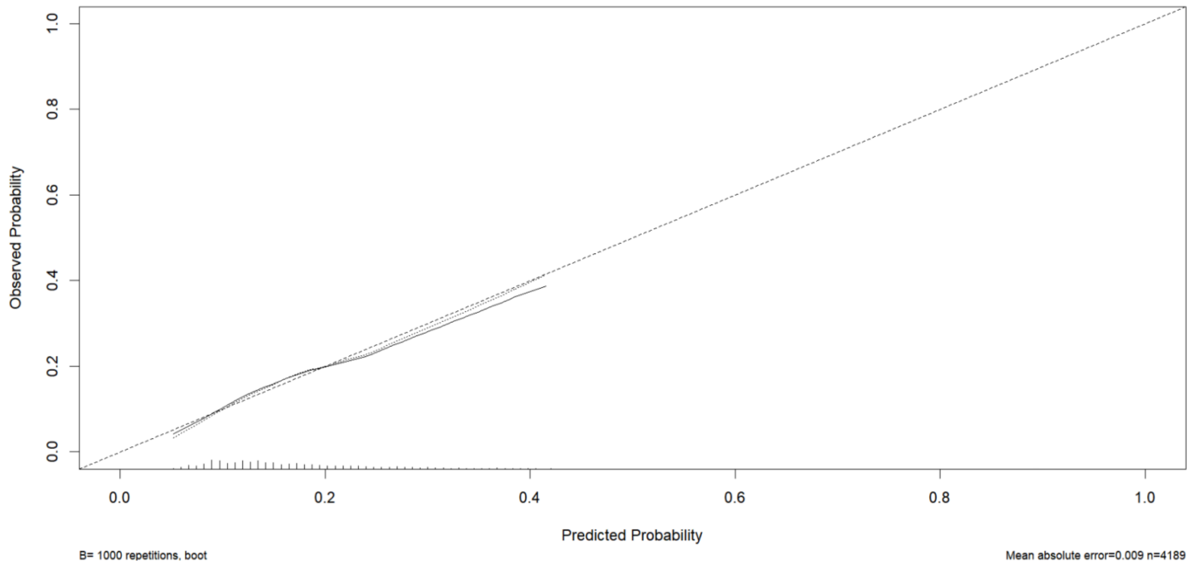

Figure 4. Calibration Curve of the Model

Figure 4 illustrates the calibration curve of the model. The predicted results of the model generally align closely with the 45-degree ideal line, indicating good overall predictive performance. However, in the lower 40% of the predicted probability range (low-risk zone), there is a noticeable deviation. Specifically, the model tends to underestimate the fall risk, with predicted fall risk being lower than the actual incidence rate. This suggests that, within the low-risk group, the model underestimates the fall risk. While the overall prediction performance is good, the deviation in the calibration curve indicates that the prediction accuracy in the low-risk group needs further optimization to improve the model’s precision in this particular subgroup.

3.5. ROC Curve

Figure 5 illustrates the ROC curves for six machine learning models on the training set and the optimal model on the validation set. Among the evaluated models (Naive Bayes, MLP, KNN, Logistic Regression, Random Forest, and XGBoost), Logistic Regression demonstrated the best predictive performance on the training set. When validated using the test set, this optimal model achieved an AUC of 0.685 (95% CI: 0.650–0.721), indicating moderate discriminatory ability in predicting fall risk among elderly individuals.

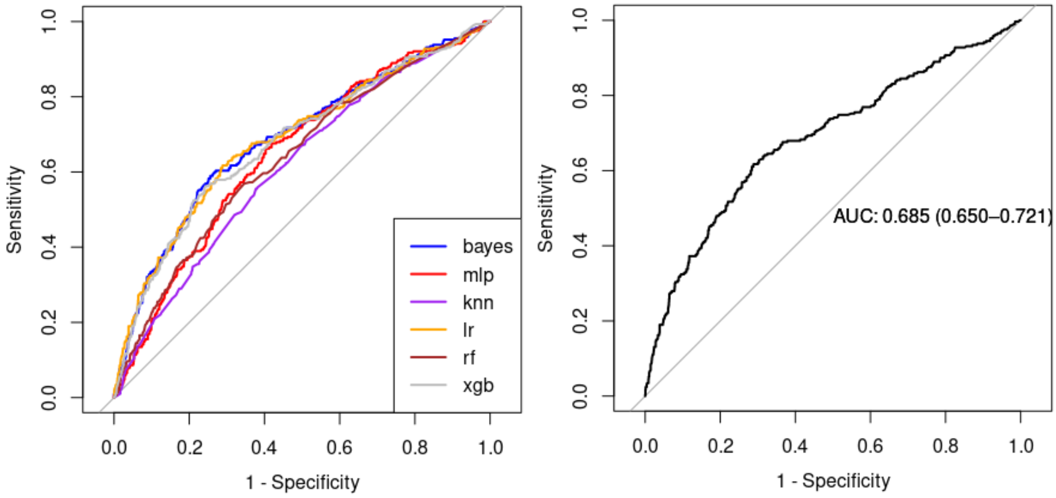

Figure 5. ROC curves of six machine learning models in the training set (left) and the optimal logistic regression model in the validation set (right)

4. Discussion

Falls pose substantial threats to the physical and psychological health of elderly populations, increasing the economic burden on families and healthcare systems, particularly in developing countries like China. Given the rapidly growing elderly demographic and the limited healthcare resources available, developing an accurate, practical, and cost-effective fall risk prediction tool is crucial for timely intervention and prevention strategies.

In this study, we utilized data from the China Health and Retirement Longitudinal Study (CHARLS) and employed multiple machine-learning-based variable screening methods (Random Forest, Boruta, XGBoost, and ABESS) to systematically identify significant predictors of fall risk among elderly individuals aged 60 years and above in Chinese communities. The final multivariate logistic regression model included factors such as hearing status, sleep duration, chronic diseases, self-reported pain, depression score, and sex, all of which significantly influenced fall risk. Specifically, elderly individuals with poor hearing, shorter sleep durations, presence of chronic diseases, pain, and higher depression scores demonstrated an increased likelihood of falls.

These findings align with previous international studies emphasizing the importance of sensory impairments, chronic illnesses, psychological status, and sleep quality in fall prevention [30-36]. Notably, hearing impairment significantly increased fall risk, consistent with evidence suggesting that diminished auditory function negatively impacts balance and spatial awareness, contributing to higher susceptibility to falls among the elderly [37-39].

Despite robust predictive performance, certain limitations exist in our study. First, individuals with severe disabilities, profound sensory impairments (e.g., blindness or complete deafness), or those residing in long-term care facilities were excluded. Consequently, our results may not generalize to these high-risk subgroups. Future research should incorporate these populations to better understand their unique fall-risk profiles and improve model inclusivity.

Second, the study primarily relied on self-reported health assessments, potentially introducing reporting biases due to subjective interpretations. Future studies should integrate more objective measures, such as physiological data captured via wearable sensors or smart-home devices, including gait analysis, muscle strength tests, or mobility assessments [40]. This approach could substantially improve model accuracy and practical applicability.

The model demonstrated moderate predictive accuracy (AUC = 0.685, 95% CI: 0.650–0.721) when validated internally, indicating good clinical utility. Its straightforward design allows easy integration into community healthcare practices, assisting healthcare providers in identifying elderly individuals at elevated fall risk. Regular assessments using this prediction model can facilitate timely, personalized interventions—such as hearing aids, sleep management strategies, chronic disease management, pain relief interventions, and psychological support—which could significantly reduce fall incidence.

From a policy perspective, our findings provide evidence-based guidance for public health interventions targeting elderly fall prevention. Government agencies and healthcare organizations can leverage this model to formulate tailored strategies and resource allocation plans, thereby alleviating the healthcare burden associated with elderly falls.

Finally, while the logistic regression model developed herein has practical value, future research should further optimize and enhance model performance. Incorporating additional factors such as environmental conditions, social support networks, and emotional states could enrich the predictive capability of the model. Additionally, advanced predictive modeling approaches, including deep learning algorithms and ensemble modeling techniques, should be explored to capture complex interactions among variables and improve overall predictive accuracy.

In conclusion, the developed fall risk prediction model effectively identifies elderly individuals at risk of falling, providing practical support for clinical decision-making and public health interventions. With ongoing advancements in technology and data collection methods, future studies will further refine fall prediction strategies, promoting healthier aging populations globally.

5. Conclusion

This study developed a community-based fall risk prediction tool for elderly individuals using a logistic regression model. The results highlighted significant associations between fall risk and factors such as sex, chronic disease history, hearing status, sleep duration, self-reported pain and depression score. The study shows that elderly individuals who are female, have a history of chronic disease, poor hearing status, and body pain are at a higher risk of falling. The model provides a useful tool for early fall risk assessment in community-dwelling elderly individuals and offers scientific evidence for developing personalized interventions to reduce fall incidents, improve quality of life, and alleviate the healthcare burden. Despite some limitations, such as reliance on self-reported data and the exclusion of certain high-risk groups, future research could improve the model by including more objective data, expanding the sample, and refining the algorithm to enhance predictive accuracy. In summary, this study presents a practical and operational tool for fall prevention in elderly populations, with broad application potential and clinical value.

References

[1] Chamie J. Population Ageing: An Inescapable Future. https://www.globalissues.org/news/2022/01/05/29746.

[2] Colon-Emeric C S, McDermott C L, Lee D S, et al. Risk Assessment and Prevention of Falls in Older Community-Dwelling Adults: A Review. JAMA, 2024, 331 (16): 1397-1406.

[3] Phelan E A, Ritchey K. Fall Prevention in Community-Dwelling Older Adults. Ann Intern Med, 2018, 169 (11): C81-C96.

[4] Rikkonen T, Sund R, Koivumaa-Honkanen H, et al. Effectiveness of exercise on fall prevention in community-dwelling older adults: a 2-year randomized controlled study of 914 women. Age Ageing, 2023, 52 (4).

[5] Wang J, Chen Z, Song Y. Falls in aged people of the Chinese mainland: epidemiology, risk factors and clinical strategies. Ageing Res Rev, 2010, 9 Suppl 1: S13-S17.

[6] Hsieh K L, Chen L, Sosnoff J J. Mobile Technology for Falls Prevention in Older Adults. J Gerontol A Biol Sci Med Sci, 2023, 78 (5): 861-868.

[7] Howcroft J, Kofman J, Lemaire E D. Review of fall risk assessment in geriatric populations using inertial sensors. J Neuroeng Rehabil, 2013, 10 (1): 91.

[8] Kang C M. Risk Factors for Falls Among Older Adults with Diabetes in the Community: A Zero-Inflated Negative Binomial Regression Model. J Community Health Nurs, 2025, 42 (1): 1-10.

[9] Bourke R, Doody P, Perez S, et al. Cardiovascular Disorders and Falls Among Older Adults: A Systematic Review and Meta-Analysis. J Gerontol A Biol Sci Med Sci, 2024, 79 (2).

[10] Lima C A, Ricci N A, Nogueira E C, et al. The Berg Balance Scale as a clinical screening tool to predict fall risk in older adults: a systematic review. Physiotherapy, 2018, 104 (4): 383-394.

[11] Omana H, Bezaire K, Brady K, et al. Functional Reach Test, Single-Leg Stance Test, and Tinetti Performance-Oriented Mobility Assessment for the Prediction of Falls in Older Adults: A Systematic Review. Phys Ther, 2021, 101 (10).

[12] Yoshimoto Y, Tanaka M, Sakamoto A. Effective Combined Assessments of Weight Bearing Ratio and Four Square Step Test in Predicting Falls in Discharged Stroke Patients. J Stroke Cerebrovasc Dis, 2021, 30 (3): 105582.

[13] Park S H, Lee Y S. The Diagnostic Accuracy of the Berg Balance Scale in Predicting Falls. West J Nurs Res, 2017, 39 (11): 1502-1525.

[14] Chen M, Wang H, Yu L, et al. A Systematic Review of Wearable Sensor-Based Technologies for Fall Risk Assessment in Older Adults. Sensors (Basel), 2022, 22 (18).

[15] Bezold J, Krell-Roesch J, Eckert T, et al. Sensor-based fall risk assessment in older adults with or without cognitive impairment: a systematic review. Eur Rev Aging Phys Act, 2021, 18 (1): 15.

[16] Antcliff S R, Witchalls J B, Wallwork S B, et al. Developing a multivariate prediction model of falls among older community-dwelling adults using measures of neuromuscular control and proprioceptive acuity: A pilot study. Australas J Ageing, 2023, 42 (3): 463-471.

[17] Andreasson K, Hesselstrand R. Fall Prevention in Community-Dwelling Older Adults. N Engl J Med, 2020, 382 (26): 2580.

[18] Montero-Odasso M, Van der Velde N, Alexander N B, et al. New horizons in falls prevention and management for older adults: a global initiative. Age Ageing, 2021, 50 (5): 1499-1507.

[19] Smith M L, Stevens J A, Ehrenreich H, et al. Healthcare providers' perceptions and self-reported fall prevention practices: findings from a large new york health system. Front Public Health, 2015, 3: 17.

[20] Casey C M, Parker E M, Winkler G, et al. Lessons Learned From Implementing CDC's STEADI Falls Prevention Algorithm in Primary Care. Gerontologist, 2017, 57 (4): 787-796.

[21] Hu J, Szymczak S. A review on longitudinal data analysis with random forest. Brief Bioinform, 2023, 24 (2).

[22] Huang Y, Huang W, Ma X, et al. Nomogram for predicting mild cognitive impairment in Chinese elder CSVD patients based on Boruta algorithm. Front Aging Neurosci, 2025, 17: 1431421.

[23] Zhou H, Xin Y, Li S. A diabetes prediction model based on Boruta feature selection and ensemble learning. BMC Bioinformatics, 2023, 24 (1): 224.

[24] Shin H. XGBoost Regression of the Most Significant Photoplethysmogram Features for Assessing Vascular Aging. IEEE J Biomed Health Inform, 2022, 26 (7): 3354-3361.

[25] Ogunleye A, Wang Q G. XGBoost Model for Chronic Kidney Disease Diagnosis. IEEE/ACM Trans Comput Biol Bioinform, 2020, 17 (6): 2131-2140.

[26] Ilic M, Srdjevic Z, Srdjevic B. Water quality prediction based on Naive Bayes algorithm. Water Sci Technol, 2022, 85 (4): 1027-1039.

[27] Xiong Y, Ye M, Wu C. Cancer Classification with a Cost-Sensitive Naive Bayes Stacking Ensemble. Comput Math Methods Med, 2021, 2021: 5556992.

[28] Virgolici O, Virgolici H. Predicting Prediabetes Using Simple a Multi-Layer Perceptron Neural Network Model. Stud Health Technol Inform, 2023, 305: 168-171.

[29] Lorencin I, Andelic N, Spanjol J, et al. Using multi-layer perceptron with Laplacian edge detector for bladder cancer diagnosis. Artif Intell Med, 2020, 102: 1017-1046.

[30] Gopinath B, McMahon C M, Burlutsky G, et al. Hearing and vision impairment and the 5-year incidence of falls in older adults. Age Ageing, 2016, 45 (3): 409-414.

[31] Yesantharao L V, Vohra V, Cheng M, et al. Olfactory Dysfunction and Balance Dysfunction are Associated with Increased Falls in Older Adults. Laryngoscope, 2023, 133 (8): 1964-1969.

[32] Coulter J S, Randazzo J, Kary E E, et al. Falls in Older Adults: Approach and Prevention. Am Fam Physician, 2024, 109 (5): 447-456.

[33] Mesrine S, Boutron-Ruault M C, Clavel-Chapelon F. Chronic pain and risk of falls in older adults. JAMA, 2010, 303 (12): 1147-1148, 1148-1149.

[34] Moncayo-Hernandez B A, Duenas-Suarez E P, Reyes-Ortiz C A. Relationship between Social Participation, Children's Support, and Social Frailty with Falls among Older Adults in Colombia. Ann Geriatr Med Res, 2024, 28 (3): 342-351.

[35] Zhang L, Liu S, Li Y, et al. Associations of Sleep Quality with Gait Speed and Falls in Older Adults: The Mediating Effect of Muscle Strength and the Gender Difference. Gerontology, 2022, 68 (1): 1-7.

[36] Mesas A E, Lopez-Garcia E, Rodriguez-Artalejo F. Self-reported sleep duration and falls in older adults. J Sleep Res, 2011, 20 (1): 21-27.

[37] Lin F R, Ferrucci L. Hearing loss and falls among older adults in the United States. Arch Intern Med, 2012, 72 (4): 369-371.

[38] Yeo B, Tan V, Ng J H, et al. Hearing Loss and Falls: A Systematic Review and Meta-Analysis. JAMA Otolaryngol Head Neck Surg, 2025.

[39] Riska K M, Peskoe S B, Kuchibhatla M, et al. Impact of Hearing Aid Use on Falls and Falls-Related Injury: Results From the Health and Retirement Study. Ear Hear, 2022, 43 (2): 487-494.

[40] Feng Y, Liu Y, Fang Y, et al. Advances in the application of wearable sensors for gait analysis after total knee arthroplasty: a systematic review. Arthroplasty, 2023, 5 (1): 49.

Supplemental Information

Based on extensive literature review (see Table S1), we selected 21 potential fall-related risk factors a priori. A multivariate logistic regression model was then developed using these factors to examine their association with falls in the elderly.

Appendix Table S1. Overview of risk factors associated with falls in the elderly

ID	Fall-Related Factors	Categories/Values	Literature
1	**Sex**	Male, Female	Stanaway F F, Cumming R G, Naganathan V, et al. Ethnicity and falls in older men: low rate of falls in Italian-born men in Australia. *Age and Ageing*, **2011**, *40* (5): 595-601.
2	**Age**	Continuous, ≥60 years	Segev-Jacubovski O, Herman T, Yogev-Seligmann G, et al. The interplay between gait, falls and cognition: can cognitive therapy reduce fall risk? *Expert Review of Neurotherapeutics*, **2011**, *11* (7): 1057-1075.
3	**Education Level**	Below primary, Primary school, Middle school, High school and above	Ong M F, Soh K L, Saimon R, et al. Fall prevention education to reduce fall risk among community-dwelling

			older persons: A systematic review. *Journal of nursing management*, **2021**, *29* (8): 2674-2688.
4	**Region**	Northeast, West, Central, East	Liang H, Zhang Z, Lai H, et al. Prevalence and risk factors for falls among older Chinese adults in the community: findings from the CLHLS study. *Brazilian journal of medical and biological research*, **2024**, *57*: e13469.
5	**Residence**	Urban, Rural	Jeon M Y, Jeong H, Petrofsky J, et al. Effects of a randomized controlled recurrent fall prevention program on risk factors for falls in frail elderly living at home in rural communities. *Medical science monitor: international medical journal of experimental and clinical research*, **2014**, *20*: 2283-2291.
6	**Near Vision**	Poor, Fair, Good, Very good, Excellent	Jin H, Zhou Y, Stagg B C, et al. Association between vision impairment and increased prevalence of falls in older US adults. *Journal of the American Geriatrics Society*, **2024**, *72* (5): 1373-1383.
7	**Far Vision**	Poor, Fair, Good, Very good, Excellent	
8	**Hearing Status**	Poor, Fair, Good, Very good, Excellent	Jiang C Y, Han K, Yang F, et al. Global, regional, and national prevalence of hearing loss from 1990 to 2019: A trend and health inequality analyses based on the Global Burden of Disease Study 2019. *Ageing research reviews*, **2023**, *92*: 102124.
9	**Sleep Duration**	≤4.5h, ≤6.5h, ≤8.5h, >8.5h	Zhu C, Sun J, Huang Y, et al. Sleep and risk of hip fracture and falls among middle-aged and older Chinese. *Scientific reports*, **2024**, *14* (1): 23273.
10	**Physical Exercise**	No, Yes	Colón-Emeric C S, McDermott C L, Lee D S, et al. Risk Assessment and Prevention of Falls in Older Community-Dwelling Adults: A Review. *The Journal of the American Medical Association*, **2024**, *331* (16): 1397- 1406.
11	**Chronic Disease Presence**	No, Yes	Ayus J C, Negri A L, Kalantar-Zadeh K, et al. Is chronic hyponatremia a novel risk factor for hip fracture in the elderly? *Nephrol*
12	**Number of Chronic Diseases**	0, 1, 2, ≥3	

			Dial Transplant, **2012**, *27* (10): 3725-3731.
13	**Self-reported Pain**	No, Yes	Alenazi A M, Alanazi M F, Elnaggar R K, et al. Prevalence and risk factors for falls among community-dwelling adults in Riyadh area. *PeerJ*, **2023**, *11*: e16478.
14	**Self-rated Health**	Very poor, Poor, Fair, Good, Very good	Zhang M, Rong J, Liu S, et al. Factors related to self-rated health of older adults in rural China: A study based on decision tree and logistic regression model. *Front Public Health*, **2022**, 10: 952714.
15	**Life Satisfaction**	Extremely satisfied, Very satisfied, Somewhat satisfied, Not very satisfied, Not at all satisfied	Roh S, Lee Y S, Lee K H, et al. Friends, Depressive Symptoms, and Life Satisfaction Among Older Korean Americans. ***Americans. Journal of immigrant and minority health*, 2015**, *17* (4): 1091-1097.
16	**Activities of Daily Living (ADL) Score**	0: complete dependence; 1: dependence in 5 activities; 2: dependence in 4 activities; 3: dependence in 3 activities; 4: dependence in 2 activities; 5: dependence in 1 activity; 6: full independence	Mattle M, Chocano-Bedoya P O, Fischbacher M, et al. Association of Dance-Based Mind-Motor Activities With Falls and Physical Function Among Healthy Older Adults: A Systematic Review and Meta-analysis. *JAMA Netw Open*, **2020**, *3* (9): e2017688.
17	**ADL Categorization**	Severe dependence: 0–1; Moderate dependence: 2–3; Mild dependence: 4–5; Independent: 6	
18	**Depression Score**	Continuous	Shao L, Shi Y, Xie X Y, et al. Incidence and Risk Factors of Falls Among Older People in Nursing Homes: Systematic Review and Meta-Analysis. ***J Am Med Dir Assoc***, **2023**, *24* (11): 1708-1717.
19	**Marital Status and Cohabitation**	Married and living with spouse, Others	Qin Y, Hao X, Lv M, et al. A global perspective on risk factors for frailty in community-dwelling older adults: A systematic review and meta-analysis. ***Arch Gerontol Geriatr***, **2023**, *105*: 104844.
20	**Health Insurance Status**	No, Yes	Yoshikawa A, Smith M L, Ory M G. Differential risk of falls associated with pain medication among community-dwelling older adults by cognitive status. ***Age Ageing*, 2021**, *50* (5): 1578-1585.

21	**Intergenerational Financial Support from Children**	Little, Moderate, Substantial	Chen Y, Wang K, Zhao J, et al. Overage labor, intergenerational financial support, and depression among older rural residents: evidence from China. ***Front Public Health***, **2023**, *11*: 1219703.