Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Machine Learning Identification of Factors Associated with Fall History Among Community-Dwelling Older Adults

Mohammadtaghi Amiri-Khorasani

amirikhorasani@ssrc.ac.ir

Shahid Bahonar University of Kerman

Francis M. Bui

University of Saskatchewan

Saeid Bahiraei

Shahid Bahonar University of Kerman

Ayyub Sheikhi

Shahid Bahonar University of Kerman

Research Article

Keywords: Aging, Accidental falls, Physical activity, Physical functional performance, Machine learning

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11438391/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Falls in older adults rarely reflect a single deficit, yet screening often relies on isolated measures, and machine-learning models can overstate performance when they use information available only after a fall. This cross-sectional study examined multidomain factors associated with ordered past-year fall-history strata in community-dwelling older adults and evaluated whether machine-learning models that exclude post-fall variables can classify these strata.

Methods

Community-dwelling adults aged 55 years and older completed assessments of demographic and social factors, anthropometry, physical function, health conditions, psychological-cognitive status, and physical activity. The analytic dataset comprised 147 records. Fall location, fall situation, and fear of falling were excluded from the predictor set to reduce circular prediction. Between-stratum profiles were compared statistically, and four algorithms were evaluated using stratified five-fold internal validation, with a secondary lower-versus-higher-burden binary contrast. Reporting follows the STROBE guideline.

Results

Hearing, comorbidity, lower-limb length, foot length, economic status, activity type, physical activity, lower-limb power, balance, and daily function differed across strata. Random forest gave the highest four-class accuracy (0.946; kappa 0.917); a support vector machine provided detailed class-level performance (accuracy 0.918). In the binary analysis random forest showed strong discrimination (AUC 0.968).

Conclusions

Fall-history burden was linked to a multidomain profile of activity, function, sensory status, health burden, and social context rather than any single measure. Findings are internally validated and hypothesis-generating, not a prospective prediction tool.

Background

Falls are a major threat to autonomy and participation in later life. The event may be sudden, but the vulnerability that precedes it often develops gradually through small losses in mobility reserve, lower-limb function, sensory processing, confidence, and environmental adaptability. Falls and fall-related concerns are common and consequential among older adults, and their frequency and correlates have been documented across diverse populations [1, 2]. For those working in falls prevention and

rehabilitation, the practical challenge is not only to identify who has fallen, but to understand the broader physical-activity and functional profile that accompanies fall-history burden.

Community-dwelling older adults differ widely in daily movement exposure, health status, social resources, access to care, and environmental constraints. Where population aging is accelerating and community-based preventive systems vary across regions, screening approaches developed elsewhere may not fully represent local patterns of family support, occupational history, medication use, physical activity, and walking environments [3]. Fall-related vulnerability is therefore better understood as a person-environment-behaviour profile than as a single biomedical label.

Routine fall-risk screening often relies on previous falls, fear of falling, gait speed, balance scales, or brief mobility tests such as the Timed Up and Go [4]. These measures are clinically useful and remain appropriate for frontline assessment. Their limitation is breadth. An older adult may not show a severe deficit on one test, yet may still present a high-burden profile when sensory impairment, cardiovascular disease, inactivity, frailty, medication burden, and social-context variables are considered together.

Machine-learning approaches can help organise heterogeneous screening data, and several studies have applied them to fall risk using gait, wearable, and clinical inputs [5–9]. However, they must be used conservatively. In fall research, high classification accuracy can be misleading if a model uses information that would not be available during a genuine screening encounter [10]. Variables describing the location or situation of a fall, for example, may sharpen classification but are available only after the event has occurred. Similarly, fear of falling can be clinically meaningful but may be too proximal to the outcome if psychological risk contributed to the assigned fall-history stratum. A defensible screening model should therefore separate information used to define the target from information entered as predictors.

The present study examined whether ordered past-year fall-history strata were associated with a multidomain profile of activity, function, health burden, sensory status, and social context, rather than presenting a deployment-ready tool. The aims were to identify between-stratum determinants of fall-history burden, to compare leakage-aware machine-learning models for four-class classification, and to test whether a simpler higher-burden binary analysis improved interpretability for community screening.

Methods

Design and participants

This cross-sectional analytical study used statistical comparison and supervised machine-learning methods to examine ordered past-year fall-history strata in community-dwelling older adults. Reporting follows the STROBE guideline for cross-sectional studies. Data were collected from geriatric care centres and comprehensive health centres between April and June 2026. Participants were adults aged 55 years or older who lived in the community and could complete the assessment protocol.

The initial target sample was estimated with Cochran's formula using a 95% confidence level and a 5% margin of error. After eligibility screening, data-quality checks, and removal of incomplete or inconsistent records, 147 complete records were available for analysis. Exclusion criteria were severe cognitive impairment, advanced musculoskeletal disease preventing safe walking, lower-limb surgery during the previous month, Parkinson's disease, diagnosed dementia, disabling neurological disease, severe uncorrected visual impairment, systemic disease markedly affecting balance, and incomplete assessment data. Records with more than 20% missingness across core predictors or unresolved logical inconsistencies were excluded before modelling.

The study was approved by the Ethics Committee of Kerman University of Medical Sciences (approval code: IR.KMU.REC.1405.248) and was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants.

Measures

The assessment covered demographic and social factors, anthropometry, body composition, physical function, medical conditions, psychological-cognitive status, and activity-related variables. Demographic and social variables included age category, sex, marital status, education, previous occupation, economic status, living arrangement, smoking status, self-perceived health, caregiver status, and family-physician access.

Anthropometric variables included height, weight, body mass index, calf circumference, waist circumference, hip circumference, lower-limb length, and foot length. Measurements were obtained with a calibrated digital scale, wall-mounted stadiometer, and non-elastic anthropometric tape. Body-composition measures were obtained by bioelectrical impedance analysis.

Physical function was assessed using activities of daily living, instrumental activities of daily living, balance, lower-limb proprioception, mobility-aid use, weekly physical activity, the 2-minute step test, back scratch test, chair sit-and-reach, chair stand, handgrip strength, Timed Up and Go, 6-minute walk, knee extensor strength, Four Square Step Test, arm curl, and 30-second sit-to-stand, drawing on established functional-fitness and mobility assessments for older adults [4, 11–13]. Handgrip strength was measured with a calibrated dynamometer, and the best of three attempts was retained.

Medical variables included visual and hearing impairment, respiratory function, blood-pressure abnormalities, diabetes, cardiovascular condition, sleep-medication use, sleep quality, sedative or antidepressant use, chronic obstructive pulmonary disease, peripheral vascular disease, stroke history, frailty, urinary incontinence, nutritional status, comorbidity count, medication category, pain-medication use, osteoporosis, dyslipidemia, dorsopathy, cancer, and orthostatic hypotension during active standing. Psychological and cognitive variables included depressive symptoms, stress, anxiety, dizziness, health-related quality-of-life items, social-support measures, and cognitive-screening information.

Outcome and leakage-aware feature strategy

The outcome was the ordered four-level fall-history stratum recorded in the cleaned dataset: Class 0, Class 1, Class 2, and Class 3. The distribution was Class 0, n = 44 (29.9%); Class 1, n = 70 (47.6%); Class 2, n = 16 (10.9%); and Class 3, n = 17 (11.6%). Because the dataset did not contain a prospective incident-fall outcome, the outcome is described conservatively as fall-history burden rather than future fall risk.

Before modelling, a leakage-aware predictor strategy was applied, consistent with guidance on transparent reporting of prediction models [10]. Fall location and fall situation were excluded because they are available only after a fall has occurred. Fear of falling was also excluded from the primary predictive feature set because the available study documentation indicated that psychological fall-related risk may have contributed to the fall-history classification. Removing these variables made the model less optimistic but more appropriate for a screening scenario.

Statistical and machine-learning analysis

Descriptive statistics are presented as mean ± standard deviation or frequency with percentage. Because many variables were ordinal or non-normally distributed, between-class comparisons used the Kruskal–Wallis test for numeric and ordinal variables. Categorical associations were tested with chi-square tests. Effect sizes were reported as eta-squared based on the Kruskal–Wallis H statistic for numeric variables and Cramer's V for categorical variables. Exact probability values are reported.

Machine-learning analyses were performed in Python using pandas, NumPy, scikit-learn, and matplotlib [14]. Continuous variables were imputed with the median and standardised to zero mean and unit variance. Categorical variables were imputed with the most frequent category and one-hot encoded. Ordinal predictors were retained as ordered numeric variables when the coding represented a meaningful rank. Four algorithms were compared: multinomial logistic regression, random forest, gradient boosting, and radial-basis-function support vector machine. Models were evaluated with stratified five-fold internal validation. Performance metrics were accuracy, balanced accuracy, weighted F1-score, macro F1-score, Cohen's kappa, mean absolute error, and ROC-AUC when class probabilities were available.

A secondary analysis collapsed the outcome into lower fall-history burden (Classes 0 and 1) and higher burden (Classes 2 and 3). This binary endpoint was added because community screening often requires a simpler triage question: which older adults should be prioritised for more detailed fall-prevention assessment? Synthetic oversampling was not used because the dataset was small and minority-class inflation could produce optimistic estimates [15]. Random forest feature importance was used descriptively to summarise the predictor profile in the binary model.

Results

Sample distribution and between-stratum profile

The final dataset included 147 participants. Most were classified in Class 1 (n = 70), followed by Class 0 (n = 44), Class 3 (n = 17), and Class 2 (n = 16). The smaller higher-burden strata are consistent with community screening, where the most vulnerable subgroups are clinically important but numerically limited.

Several variables differed across fall-history strata (Table 1). The largest numeric and ordinal contrasts were observed for hearing status, comorbidity status, lower-limb length, foot length, economic status, hip circumference, activity type, physical activity level, sleep-medication use, lower-limb power, balance score, and activities of daily living. Among categorical variables, occupation, cardiovascular condition, sex, pacemaker status, family-physician access, and incontinence showed the strongest associations. The pattern supported a multidomain physical-activity and function perspective rather than reliance on one mobility test.

Table 1
Selected numeric and ordinal characteristics by fall-history stratum. Values are mean ± SD; H from Kruskal–Wallis tests; class labels are ordered fall-history strata from the cleaned dataset.

Variable	Class 0	Class 1	Class 2	Class 3	H	p	eta²(H)
Hearing	1.93 ± 0.25	1.39 ± 0.49	1.00 ± 0.00	1.59 ± 0.51	52.29	< .001	.345
Comorbidity	1.45 ± 0.50	1.23 ± 0.42	2.00 ± 0.00	1.00 ± 0.00	44.98	< .001	.294
Lower-limb length	93.64 ± 9.08	97.23 ± 4.13	105.12 ± 3.59	100.29 ± 5.19	39.65	< .001	.256
Foot length	26.61 ± 2.03	27.97 ± 2.41	29.38 ± 0.50	28.06 ± 2.86	33.66	< .001	.214
Economic status	2.09 ± 0.56	1.46 ± 0.53	1.62 ± 0.50	1.47 ± 0.72	29.78	< .001	.187
Hip circumference	50.30 ± 10.64	52.46 ± 20.57	56.75 ± 1.44	46.82 ± 6.57	29.48	< .001	.185
Activity type	1.20 ± 0.41	1.54 ± 0.50	1.62 ± 0.50	1.00 ± 0.00	27.69	< .001	.173
Physical activity	2.36 ± 1.12	2.46 ± 1.02	2.25 ± 1.00	1.06 ± 0.24	26.23	< .001	.162
Sleep-medication use	1.84 ± 0.37	1.89 ± 0.32	2.00 ± 0.00	1.41 ± 0.51	25.74	< .001	.159
Lower-limb power	24.63 ± 11.88	34.59 ± 6.71	32.06 ± 4.21	27.59 ± 7.95	21.97	< .001	.133
Balance	15.29 ± 7.73	10.27 ± 3.96	10.53 ± 1.04	14.14 ± 8.05	18.61	< .001	.109
Activities of daily living	2.98 ± 0.15	2.99 ± 0.12	3.00 ± 0.00	2.76 ± 0.44	18.60	< .001	.109

Machine-learning classification

Internal validation performance for all models is shown in Table 2. In the four-class analysis, random forest achieved the highest accuracy and agreement (accuracy = 0.946; weighted F1 = 0.945; kappa = 0.917), followed by logistic regression, the support vector machine, and gradient boosting. Because the support vector machine produced a stable, interpretable class-by-class solution across the ordered strata, its class-level performance is reported in detail in Table 3 and Fig. 1 as a representative four-class model; the confusion matrix in Fig. 1 corresponds to this model.

Table 2
Internal validation performance of leakage-aware machine-learning models (stratified five-fold internal validation). ROC-AUC shown as — when class probabilities were not used. Binary = higher-burden (Classes 2–3) versus lower-burden (Classes 0–1).

Analysis	Model	Acc.	Bal. acc.	Weighted F1	Macro F1	Kappa	MAE	ROC-AUC
Four-class	Logistic regression	.932	.913	.932	.914	.896	.143	.972
Four-class	Random forest	.946	.923	.945	.933	.917	.116	.987
Four-class	Gradient boosting	.884	.858	.882	.854	.824	.238	.976
Four-class	SVM (RBF)	.918	.893	.917	.893	.876	.156	—
Binary	Logistic regression	.918	.883	.918	.883	.766	.082	.880
Binary	Random forest	.932	.881	.930	.898	.796	.068	.968
Binary	Gradient boosting	.939	.917	.939	.913	.826	.061	.961
Binary	SVM (RBF)	.898	.870	.899	.858	.716	.102	.866

Table 3
Class-specific performance of the representative four-class support vector machine (pooled out-of-fold predictions from stratified five-fold internal validation).

Class	Precision	Recall	F1	Support
0 (lower history)	.952	.909	.930	44
1 (moderate history)	.931	.957	.944	70
2 (high history)	.941	1.000	.970	16
3 (very high history)	.750	.706	.727	17

The support vector machine classified Classes 0, 1, and 2 with strong performance, but recall for the highest-burden stratum was lower (Table 3). This result is clinically important. Even when overall accuracy is high, a screening model that misses a subset of higher-burden participants should not be used as a stand-alone triage tool.

Discussion

This study examined ordered past-year fall-history strata in community-dwelling older adults using a multidomain physical-activity, function, and health profile. Three findings are most relevant. First, fall-

history burden was associated with more than one domain: activity-related variables, sensory status, lower-limb characteristics, health burden, and social context all contributed. Second, leakage-aware machine-learning models classified the ordered strata with strong internal performance, but the highest-burden class remained the main source of clinically important error. Third, a binary higher-burden analysis provided a more interpretable community-screening signal than the four-class model alone.

The results support a multidomain view of mobility vulnerability in aging. Older adults seldom become vulnerable because of one isolated impairment. Risk-relevant profiles often accumulate across sensory processing, activity exposure, chronic disease, lower-limb function, and confidence in movement. In this sample, hearing status, physical activity level, lower-limb power, balance, and activities of daily living were not isolated findings; they formed a coherent profile of functional reserve and daily movement capacity.

The physical-activity signal is important. Participants with higher fall-history burden tended to show less favourable activity-related patterns and functional performance. This finding aligns with the idea that reduced movement exposure may both reflect and reinforce mobility vulnerability. Lower activity can accelerate deconditioning, reduce opportunities to practise balance recovery, and increase dependence on others. Conversely, poor balance or sensory impairment can make daily activity feel less safe, creating a cycle of avoidance and declining capacity.

The machine-learning results should be interpreted with restraint. The strong internal performance shows that the multidomain variables contained structured information about fall-history burden. It does not show that the model can prospectively predict future falls, a distinction emphasised in current prediction-model reporting guidance [10]. The sample was modest, the higher-burden groups were small, and validation was internal. The decision to remove fall-location, fall-situation, and fear-of-falling variables was therefore essential. Models that use post-fall context may appear accurate but are not appropriate for pre-event screening.

The binary analysis may be the most useful applied finding. Community and rehabilitation services often need to decide who should receive a more detailed fall-prevention assessment rather than assign older adults to one of four exact categories. A binary higher-burden model is easier to explain, easier to calibrate, and better aligned with referral decisions. Even so, sensitivity for higher-burden cases should be prioritised in future work, potentially through larger samples, external validation, threshold calibration, and prospective follow-up.

Several limitations should be acknowledged. The design was cross-sectional and used fall-history strata rather than prospectively observed incident falls. The sample was drawn from a single region and was unevenly distributed across strata. Some variables were self-reported. Reliability statistics for the present assessment sessions were not available in the supplied dataset. The sex distribution was imbalanced, which may affect generalisability. Finally, feature-importance values are descriptive and should not be interpreted as causal effects.

Conclusions

Past-year fall-history burden in community-dwelling older adults was associated with a multidomain profile involving physical activity, function, sensory status, lower-limb characteristics, health burden, and social context. Leakage-aware machine-learning models summarised this profile with strong internal performance, but the findings should not be treated as a prospective fall-prediction tool. For practice, the results support broader screening strategies that combine functional testing with activity, sensory, medical, and contextual information. A compact multidomain screen may better identify adults who should receive more detailed fall-prevention assessment than any single mobility test or fall-history question, although prospective multicentre validation is required before clinical use.

Abbreviations

ADL
Activities of daily living
AUC
Area under the receiver operating characteristic curve
BMI
Body mass index
IADL
Instrumental activities of daily living
MAE
Mean absolute error
RBF
Radial basis function
ROC
Receiver operating characteristic
SD
Standard deviation
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology
SVM
Support vector machine
TUG
Timed Up and Go

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of Kerman University of Medical Sciences (approval code: IR.KMU.REC.1405.248) and was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Authors' information

Not applicable.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution

MA-K conceived and designed the study, collected and analysed the data, and drafted the manuscript. FMB contributed to study design, methodological oversight, and critical revision of the manuscript. SB contributed to data collection and interpretation. AS contributed to statistical analysis and interpretation. All authors read and approved the final manuscript.

Acknowledgement

The authors thank the participating older adults and the staff of the collaborating health centres.

Data Availability

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. 1. Tavan H, Azadi A. The frequency of fall, fear of fall and its related factors among older adults in Iran: a systematic review and meta-analysis. Int J Afr Nurs Sci. 2024;20:100660.
2. 2. World Health Organization. WHO global report on falls prevention in older age. Geneva: World Health Organization; 2007.
3. 3. World Health Organization. Ageing and health. Geneva: World Health Organization; 2024.
4. 4. Podsiadlo D, Richardson S. The Timed Up and Go: a test of basic functional mobility for frail older persons. J Am Geriatr Soc. 1991;39(2):142–8.
5. 5. Bardus M, El Rassi R, Chahrour M, Al-Hajj S, Nehme J, El Asmar K. Application of machine learning algorithms for predicting falls in older adults: a systematic review. Geriatrics. 2021;6(4):88.
6. 6. Dormosh N, Medlock S, Abu-Hanna A, Zonneveld HI, de Rooij SE, Moll van Charante EP. A systematic review of fall prediction models for community-dwelling older adults. Age Ageing. 2024;53(7):afae131.
7. 7. Gonzalez-Castro A, Leiros-Rodriguez R, Prada-Garcia C, Benitez-Andrades JA. The applications of artificial intelligence for assessing fall risk and detection in older adults: systematic review. J Med Internet Res. 2024;26:e54934.
8. 8. Lockhart TE, Soangra R, Yoon H, Wu T, Frames CW, Weaver R, et al. Prediction of fall risk among community-dwelling older adults using a wearable system. Sci Rep. 2021;11:20976.
9. 9. Noh B, Youm C, Goh E, Lee M, Park H, Jeon H. XGBoost based machine learning approach to predict the risk of fall in older adults using gait outcomes. Sci Rep. 2021;11:12183.
10. 10. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378.
11. 11. Berg K, Wood-Dauphinee S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. Physiother Can. 1989;41(6):304–11.
12. 12. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport. 1999;70(2):113–9.
13. 13. Rikli RE, Jones CJ. Development and validation of a functional fitness test for community-residing older adults. J Aging Phys Act. 1999;7(2):129–61.
14. 14. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825–30.
15. 15. He H, Garcia EA. Learning from imbalanced data. IEEE Trans Knowl Data Eng. 2009;21(9):1263–84.

Figures

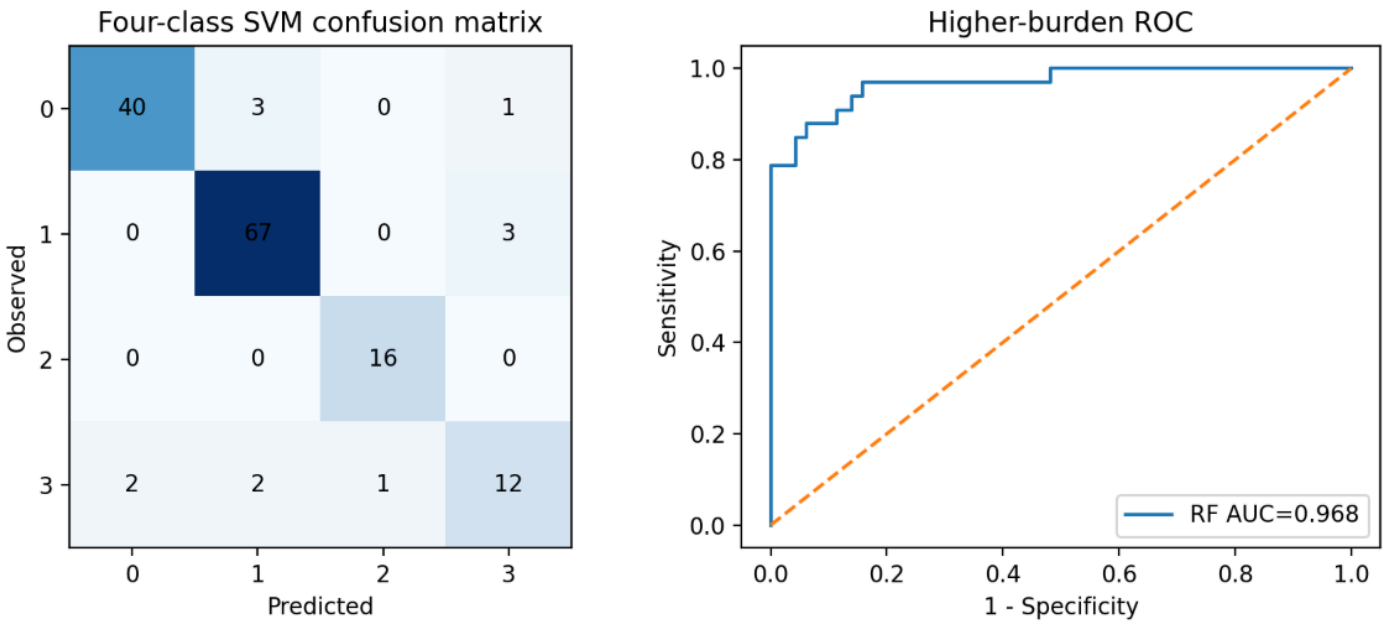


Figure 1

Model performance and error structure. Left: confusion matrix for the representative four-class support vector machine. Right: receiver operating characteristic curve for the binary higher-burden random forest model.

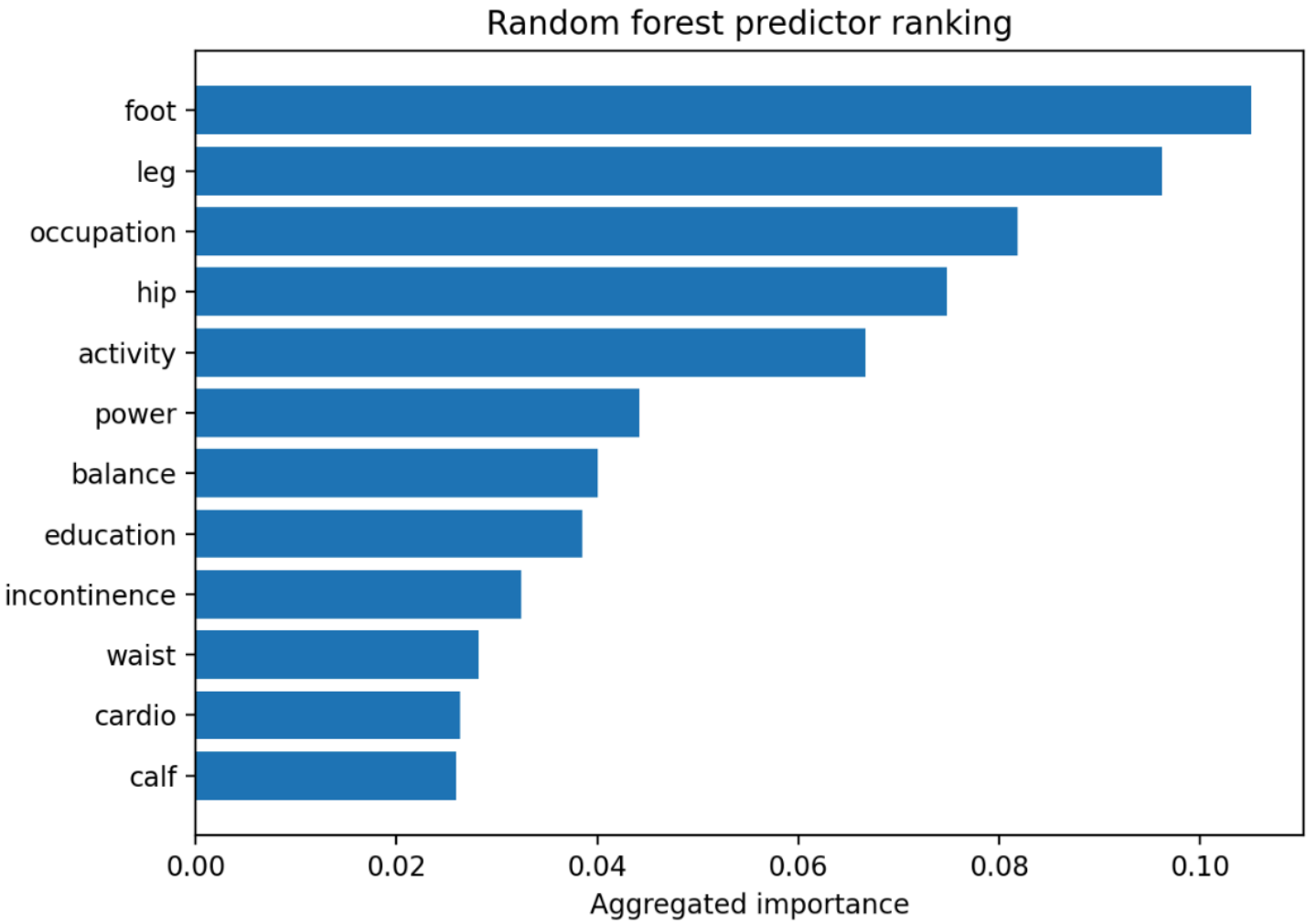


Figure 2

Aggregated random forest predictor importance for the binary higher-burden model.