

Short physical performance battery: Pilot study of a human motion capture app (MobiSPPB)

DIGITAL HEALTH

Volume 11: 1–12

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/20552076251346575

journals.sagepub.com/home/dhj



Lucas Küppers¹ , Richard Pfannenstiel², Arezoo Bozorgmehr¹ , Stephan Jonas² ,
Birgitta Weltermann¹ and Lara Marie Reimer²

Abstract

Background: A standardized fall risk assessment can guide targeted interventions. The widely used short physical performance battery (SPPB) for mobility assessment covers balance, gait speed, and lower limb strength, but is time-consuming and requires trained raters. The newly developed video-based smartphone application called MobiSPPB provides a rater-independent SPPB assessment. This study evaluated the technical validity and reliability of the MobiSPPB app compared to the standard rater-based SPPB. In addition, the ability to detect disease-related movement patterns was investigated.

Methods: Using a standardized experimental setting, 10 healthy participants performed the SPPB with and without movement impairments simulated by an instant aging suit. Two experienced raters rated the SPPB performance, and a smartphone recorded at the same time. The MobiSPPB app analyzed videos via vision-based human motion capture techniques. Spearman's correlations, the intraclass correlation coefficient (ICC), and receiver operating characteristic curves were calculated.

Results: There was a strong correlation between the app and standard SPPB (Spearman's Correlation of 0.869, 95% confidence interval (CI) of 0.79–0.92, $p < 0.001$). Compared with the standard assessment, the app presented a more significant ICC in the test–retest reliability analysis (0.936, 95% CI of 0.87–0.97, $p < 0.001$). Detecting disease-related movement patterns achieved high accuracy in capturing severe impairments such as hemiplegia (area under the curve (AUC) 93%). Inconsistencies between the raters indicated that the app provides more objective assessments.

Conclusions: The technical validation of the MobiSPPB app was successful in a standardized experimental setting and requires further testing in clinical practice.

Keywords

Geriatric assessment, mobility assessment, falls, frailty, sarcopenia, short physical performance battery, vision-based human motion capture technology, mHealth

Received: 11 January 2025; accepted: 15 May 2025

Background

Decreased muscle strength in community-dwelling older adults reduces their functional ability and balance control, resulting in increased falls.¹ In a cross-sectional US study with 11,344 participants above age 65, 30.4% suffered from frailty, and 42.9% from sarcopenia.² A recent meta-analysis of 104 international studies reported a 26.5% prevalence of falls in older adults.³ Neurological conditions such as Parkinson's disease (PD) or stroke can also significantly increase the risk of falling.⁴ Falls are responsible for more than 80% of nonvertebral fractures, with respective short- and even long-term care needs,

hospitalizations, and healthcare costs.^{5,6} Standardized assessments of mobility and fall risk in older adults are needed to guide targeted interventions such as balance

¹Institute of General Practice and Family Medicine, University Hospital Bonn, Bonn, Germany

²Institute for Digital Medicine, University Hospital Bonn, Bonn, Germany

Corresponding author:

Lucas Küppers, Institute of General Practice and Family Medicine, University Hospital Bonn, University of Bonn, Venusberg-Campus 1, 53127, Bonn, Germany.

Email: lucas.kueppers@ukbonn.de



and muscle training.⁷ The timed up-and-go (TUG) test and short physical performance battery (SPPB) are widely used mobility assessments.⁸ First introduced in 1994, the SPPB evaluates balance in three different foot poses, gait speed in a 4-meter walk, and lower limb strength in rising from a chair five times in a row.⁹ A total score ranging from 0 to 12 is calculated on the basis of the time required to complete each item.¹⁰ Lower SPPB scores are linked to important outcomes such as disability, frailty, sarcopenia, all-cause mortality and heart failure hospitalization.^{9,11–13} However, physician time constraints and staff training requirements are some of the key barriers to routine implementation.^{14–16} Thus, easily applicable and inter-rater-independent assessment tools are needed for an objective, time-saving assessment and longitudinal follow-up of patients.

Mobile medical apps and automation processes have gained prominence in recent years for geriatric mobility assessments.¹⁷ In a systematic review of 28 studies, fall risk assessment with wearable sensors was feasible in older adults regardless of their cognitive status.¹⁸ In a recent systematic review of 40 studies, an instrumented timed up-and-go test (iTUG) using cameras, sensors and optoelectronic systems was able to increase the predictive value of the TUG for several outcomes, such as frailty.¹⁹ Also, a smartphone-sensor supported TUG correctly identified different neurological conditions, such as PD.²⁰ Jung et al. introduced a multisensor-based SPPB that includes load cell detectors and LiDAR sensors. A study involving 40 elderly individuals equipped with these sensors revealed that the electronically obtained SPPB scores strongly correlated with rater-based scores.²¹ Duncan et al. developed a camera-based tool with human motion capture (HMC) techniques that achieved high accuracy rates for all SPPB items in a study with eight participants.²² While the described systems provide promising digital solutions for automated mobility assessments, they are linked to several shortcomings. First, most of them use the TUG. While this is a widely used mobility assessment, the SPPB provides more detailed insights into different domains such as balance and strength, which are of high importance. The German guideline for geriatric assessment therefore highlights its increasing importance of use.²³ Furthermore, the described SPPB systems require extensive use of additional hardware such as body-worn sensors or camera systems. The implementation of such systems into daily medical practice is thus linked to barriers, supporting the need for affordable systems using consumer-grade hardware.

The field of HMC encompasses the comprehensive process of digitally capturing, processing, and analyzing human motion and has facilitated mobility assessments.²⁴ It can simultaneously assess entire body movements for one or multiple individuals and complex movements of specific body parts such as the face or hands.²⁵ Image processing systems (IPSs), as a part of HMC, analyze image data

via various machine learning techniques, eliminating the need for subjects to wear transponder devices or have markers placed on their bodies.²⁶ With the integration of specialized neural engines into modern processors, IPS can now employ sophisticated machine learning models for image recognition on consumer-grade devices.^{27,28}

The progress of IPS on mobile devices and the shortcomings of the standard SPPB suggest that a newly developed mobile HMC-based smartphone application (MobiSPPB) might offer potential in geriatric mobility assessment. Using a standardized experimental setting, this study evaluated the technical validity and reliability of the MobiSPPB app compared to the standard rater-based SPPB.

Methods

This research presents a cross-sectional pilot study assessing the technical validity and reliability of a HMC app for the short physical performance battery (MobiSPPB) in healthy participants with simulated mobility impairments. Standardized data were recorded of 10 healthy subjects who wore the instant aging suit GERT[®] to simulate different geriatric conditions.²⁹ We aimed to evaluate the app in an experimental setting to validate its functionality before testing it with patients. Each SPPB test was rated by two human raters and the MobiSPPB app. The following three hypotheses were evaluated. The MobiSPPB app:

- has good validity according to its correlation with the standard rater-based SPPB (H1),
- has high test–retest reliability (H2), and
- can detect specific disease-related movement patterns (H3).

Participant recruitment and ethical approval

Ten healthy, young volunteers were recruited via email and notice boards among students at the University of Bonn. Participants were eligible if they were older than 18 years, could consent, and had no motor impairments. The study complied with the Declaration of Helsinki. Each participant provided written informed consent before the SPPB runs, which addressed potential risks associated with using instant aging suits (e.g. falling). Volunteer liability insurance covered these risks. The Ethics Committee of the Medical Faculty of the University of Bonn did not raise ethical or professional objections to the study (reference number 191/23-EP, date of approval: 10 August 2023).

Study design and data collection

Research involving the GERT[®] suit has demonstrated its effectiveness in simulating age-related physical limitations. Vieweg et al. detailed the suit's ability to replicate

impairments such as joint stiffness and reduced coordination, highlighting its potential in generating reference data for understanding the physical performance of older populations.²⁹ The University of Bonn uses GERT[®] to provide medical students firsthand experience of the physical challenges associated with aging. The decision was made to use GERT[®] in this research, as the study's goal is to pilot the technical setup of MobiSPPB before assessing it in a clinical environment with older adults. Each participant performed the SPPB in four different successive GERT[®] configurations (Figure 1):

- Healthy (without an instant aging suit)
- Poststroke with hemiplegia (stiffened knee joint and bandaged right arm)
- PD (weighted vest, light foot weights, and elastic strap around the legs to simulate a small-step gait)
- Frailty (stiff shoes, a weighted vest, light hand weights, heavy foot weights, one-sided knee stiffness, and vision-impairing glasses).

Before the performance of the SPPB, participants were shown the exact movements and positions by one of the study investigators. To measure inter-rater reliability, participants were asked to perform the SPPB twice for each condition, resulting in 80 performed SPPB runs (10 participants $\times$ 4 simulated conditions $\times$ 2 runs). The SPPB item for gait speed was performed with deviations from the instructions described in the standard SPPB protocol: instead of only recording the 4-meter walk, participants were asked to sit on a chair, stand up, and then walk the 4 meters. The raters and the MobiSPPB app only started to record the time when the participants started to walk, so that the measured time was consistent with the standard SPPB protocol. All other SPPB items were performed according to the standard protocol. The recording setup consisted of a tripod holding a smartphone, a chair for the chair-rise test, and a taped line on the floor to mark the 4-meter distance. The tripod position was also marked on the floor to ensure a reproducible setup. The tripod height was set to each participant's hip height to ensure good visibility of the legs (Figure 2). It took participants approximately one hour to perform the above-described test succession. The two measurement methods were subsequently applied: standard SPPB assessment by two experienced raters was performed via paper and pencil, while videos were recorded and later analyzed in the MobiSPPB app. The two raters were placed at least 1 m apart from each other and did not have any knowledge of the app's ratings to minimize bias. Both raters received training for SPPB rating, including explanation of the procedure and test runs to ensure inter-rater reliability. An iPhone 14 Pro was used for the video recordings and analysis.

Smartphone application

The self-developed MobiSPPB is a native iOS application and uses recent vision-based HMC technology to perform an automated SPPB assessment (Figure 3). The app calculated the duration and the SPPB score (for the total battery and per individual item) and predicted movement patterns via the different rule-based algorithms described in the following sections. The standard paper and pencil assessments were transferred into an Excel file for further analysis. The final dataset used for the analysis consisted of all standard and automatic ratings of the SPPB runs (including durations and scores) and the predicted movement patterns.

Requirements and system design: The decision to opt for a purely optical capture approach was rooted in the envisioned system's user-friendliness and its intended application domain in hospitals, practices, or home settings by eliminating the need for additional sensors. The application was designed to enable inexperienced users to conduct the SPPB correctly on the app by guiding them through the battery protocol. During the performance of individual SPPBs, the system should be further able to identify movement patterns, providing more comprehensive insight into the patient's overall health status.

Video integration: To automatically evaluate the individual SPPB runs, the app required videos depicting the participants performing the SPPB. The system was designed to decouple the evaluation from the video recording. This allowed immediate input with real-time recordings from the smartphone's integrated camera or sequential input with prerecorded videos stored in the device's media gallery. The latter option was used in this study to evaluate the evaluation algorithms' proficiency.

Pose detection: BlazePose, the underlying algorithm of the Google MLKit's Vision API, was integrated into the MobiSPPB app.³⁰ It was specifically designed for real-time motion analysis on mobile devices and can be integrated into Android and iOS applications. BlazePose provides a skeletal topology capturing 33 distinct key points of the patient's body.

SPPB evaluation algorithms: Different rule-based algorithms were developed for the MobiSPPB app based on calculated joint angles and distances between body landmarks. The app continuously monitored the participants' elapsed duration and potential movement patterns. Scores were assigned according to the original SPPB protocol, depending on how long the participant took to complete each SPPB item.

The balance item consisted of three poses: the side-by-side, semitandem, and tandem. For recognizing all three poses, the underlying algorithm mainly focused on the position of the subject's ankle joints and the distance between them (Figure 4). Upon the participant taking the demanded pose, the app recognized the pose and enabled the examiner to manually initiate the item by pushing the



Figure 1. Simulated geriatric conditions with the instant aging suit GERT®.



Figure 2. The recording setup.

start button. After the participant held the position successfully for 10 s, the app automatically concluded the item.

The duration required for a participant to walk a 4-meter distance was measured to assess gait speed. Once the participant started walking, the app detected the participant's movement and timing started automatically. The system continuously monitored the participants' progress as they walked toward the camera. When the vertical coordinate of either foot heel has wholly crossed the virtual course finish line, the item concluded, and the timer stopped (Figure 5).

The chair stand item algorithm divided the item sequence into four states: the participant sat, rose from the chair, stood up straight, and lowered back down. To assess whether the participant was fully seated, the position of the hips was captured. While the participant was sitting on the chair in the start position, the vertical hip coordinate was stored when the examiner started the timer. When the hip joints approached the stored position after standing upright, the algorithm recognized the participant sitting on the chair again. The quotient of the hip-to-knee and knee-to-ankle distances was used to determine whether the participant was standing upright (Figure 6). The distance between the participant's hands and thighs was tracked to detect the position of the participant's crossed arms during the exercise. The app recognized when the participant completed the five chair rises and stopped the timer.

Movement pattern detection: A rule-based algorithm was developed using all detected abnormalities as inputs to predict geriatric conditions (hemiplegia, PD, frailty). The abnormality detection algorithm was built on specific predefined movement abnormalities that were identified

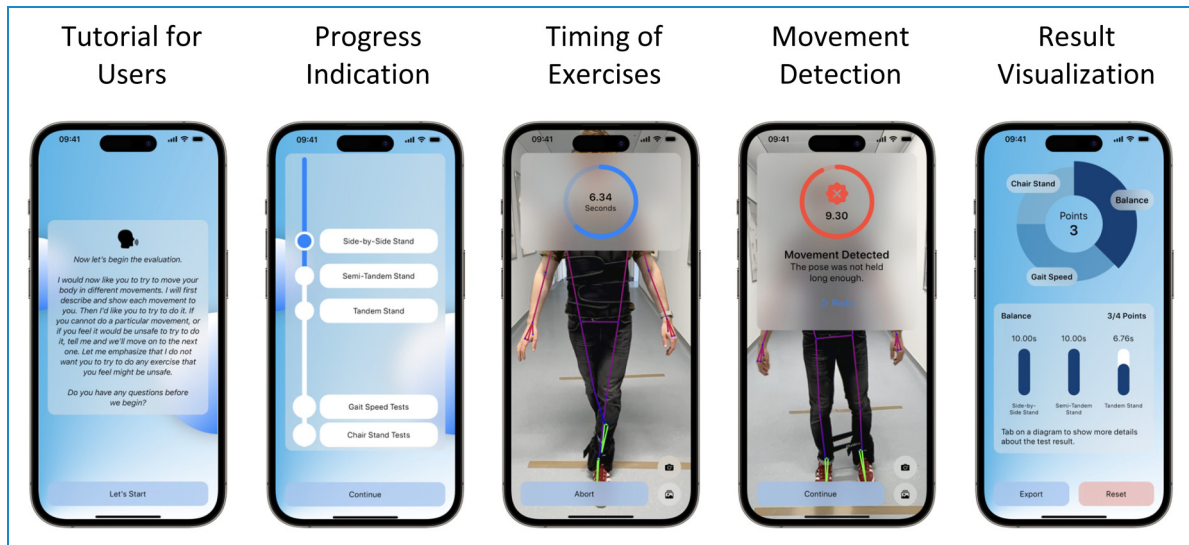


Figure 3. User interface of the MobiSPPB application.



Figure 4. Distances between ankle joints and foot positions for the side-by-side, semitandem and tandem pose.

with the help of an experienced physiotherapist before the study was conducted. The movement abnormalities were specific for each of the SPPB items and were further differentiated by whether they were strongly indicative or might hint at a specific condition. For the balance item, constant bending of one arm while the other was straight was considered a strong indication of hemiplegia. Spreading the arms to maintain balance or significant swaying while holding the pose was considered a low frailty indicator. Abnormally short step sizes during the gait speed items were a strong signal for PD. Limited joint mobility in the legs indicated frailty, whereas general asymmetry due to paralysis in one leg was typical of hemiplegia. Walking in serpentine had weak implications for PD and frailty. If the system detected an abnormal spatial offset between the feet during the chair stand item, suggesting an asymmetrical load on the legs, it was interpreted as a strong indication of hemiplegia and frailty. Based on these rules, a scoring system was created that assigned scores to each possible impairment, indicating the likelihood of the simulated

condition. The score for each impairment was initially set to zero. It was incremented each time a movement abnormality typical for the respective condition was detected during SPPB performance, also considering the degree to which a specific abnormality was indicative of the respective condition. The algorithm considered the condition with the highest score the most likely one.

Statistical analysis

Statistical disclosure control was applied, and only pseudonymized datasets were used for statistical analyses to ensure participant confidentiality. To test H1, correlation analysis was used to compare the results of the MobiSPPB app's total battery score with those of the standard SPPB. Similarly, the recorded durations and the calculated scores for the balance, gait speed, and chair stand items were compared. The relationships between automatically and standard generated results were also visually inspected via scatter plots and heatmaps. A correlation coefficient was subsequently calculated.

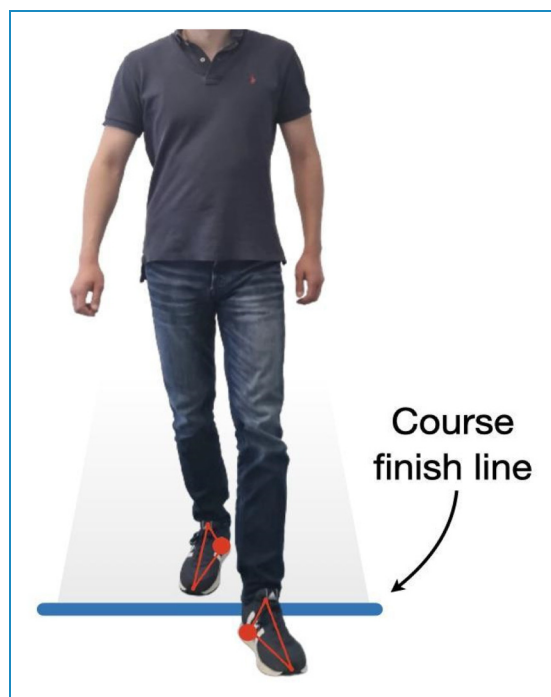


Figure 5. Foot positions while crossing the course finish line for the gait speed item.

The durations and scores were not assumed to be normally distributed across the study population, as the simulation of different conditions influenced the individual performance. The Shapiro–Wilk test was used to check whether the variables followed a normal distribution. Pearson’s correlation was applied if the data were normally distributed, and alternatively, Spearman’s correlation was utilized.

To address H2, the total scores of the two subsequent SPPB runs per MobiSPPB app and standard SPPB were compared via the intraclass correlation coefficient (ICC).³¹ The two-way random effects model (2,1) was selected. Since each participant was assessed twice by the same raters, with each rater evaluating multiple participants, this ICC model was chosen in line with the statistical methods used in similar studies.^{32,33}

H3 was tested by calculating the movement detection algorithm’s accuracy score. Furthermore, a heatmap was plotted to compare the app’s predictions against the actual conditions to gain deeper insights into misclassifications. The classifier employed multiclass classification, which was designed for datasets with all classes being mutually exclusive. For this type of classifier, the evaluation metrics for individual classes were averaged to evaluate the algorithm’s overall performance across the dataset. A macro-average approach was applied for this purpose. A receiver operating characteristic (ROC) curve was generated to evaluate the classifier model’s performance. The area under the curve (AUC) derived from the ROC curve signified the classifier’s ability to distinguish between classes. In this

study, the ROC curve was generated for each class individually by disaggregating the multiclass problem into a series of binary problems via the one-vs-rest approach. The macroaverage was computed by summing the individual values for true positives, true negatives, false positives, and false negatives across all classes.³⁴ All significance levels were set at $p < 0.05$.

Results

Participants

Half of the participants were female ($n = 5$). All participants were over the age of 18, had no pre-existing medical conditions and didn’t take any medications regularly. The participants exhibited a baseline performance score of 10.0 ± 1.376 (SEM 0.308, automatic measurement) and 8.95 ± 1.190 (SEM 0.266, standard measurement) in the “healthy” condition. All participants completed the recordings, resulting in a dataset without missing data.

Comparison of the MobiSPPB app and standard SPPB

The automatic scoring of the app and the standard rater-based scoring yielded average total scores across all runs of 8.225 ± 1.746 (SEM 0.142) and 7.588 ± 1.262 (SEM 0.192), respectively (Figure 7). The Shapiro–Wilk test performed on both variables revealed a non-normal distribution for each. Therefore, Spearman’s correlation was used to statistically evaluate the relationship between the total scores obtained by automatic and standard scoring, resulting in a correlation coefficient of 0.869 (confidence interval (CI) 95%: 0.79–0.92, $p < 0.001$).

The comparison of each SPPB item included recorded durations and calculated scores. The comparison of total balance scores revealed that the automatic evaluation assigned full scores for 70 out of 80 SPPB runs (Figure 7) with automatic scores of 3.862 ± 0.379 (SEM 0.043) and standard scores of 3.950 ± 0.218 (SEM 0.025). The remaining runs received a score of only two or three because of faulty detection of movements even if participants were able to maintain the balance positions. Similarly, the human raters inaccurately evaluated some balance items, resulting in full scores for 76 out of 80 runs. The correlation analysis between the automatic and standard assessment scores for the balance items yielded a Spearman’s correlation coefficient of 0.106 (CI 95%: -0.12 – 0.32 , $p = 0.35$). The analysis of the gait speed item revealed a strong correlation between the measured durations and derived scores (Figure 8), with Spearman’s correlation coefficients of 0.965 (CI 95%: 0.94–0.98, $p < 0.001$) for durations) and 0.896 (CI 95%: 0.83–0.94, $p < 0.001$) for scores. The gait speed duration took 6.974 ± 2.150 s (SEM 0.242) in the automatic and 7.522 ± 2.200 s (SEM 0.248) in the standard

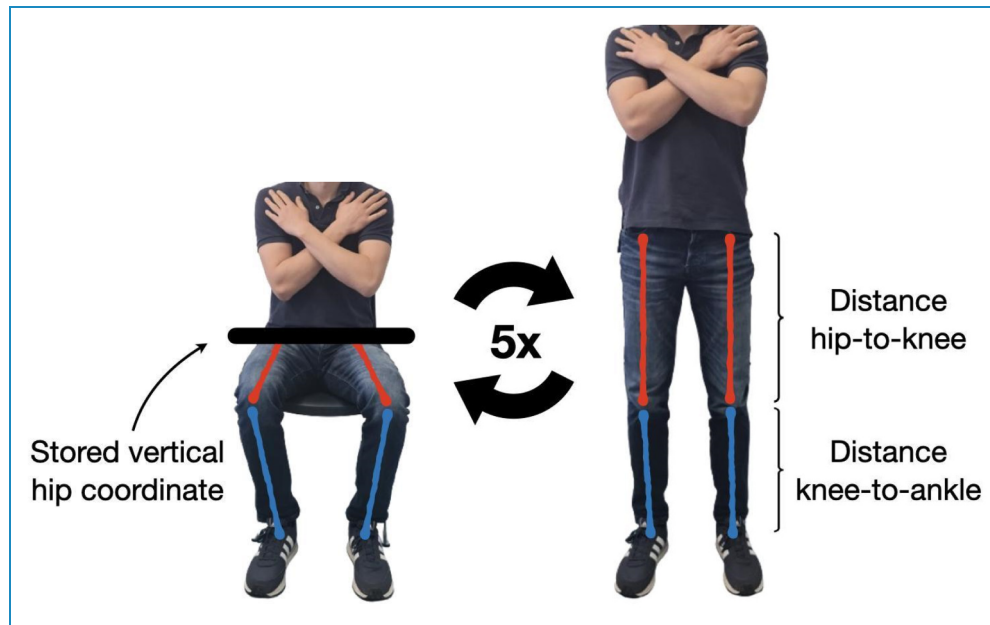


Figure 6. Hip coordinate and lower extremity ankle distances for the chair stand item.

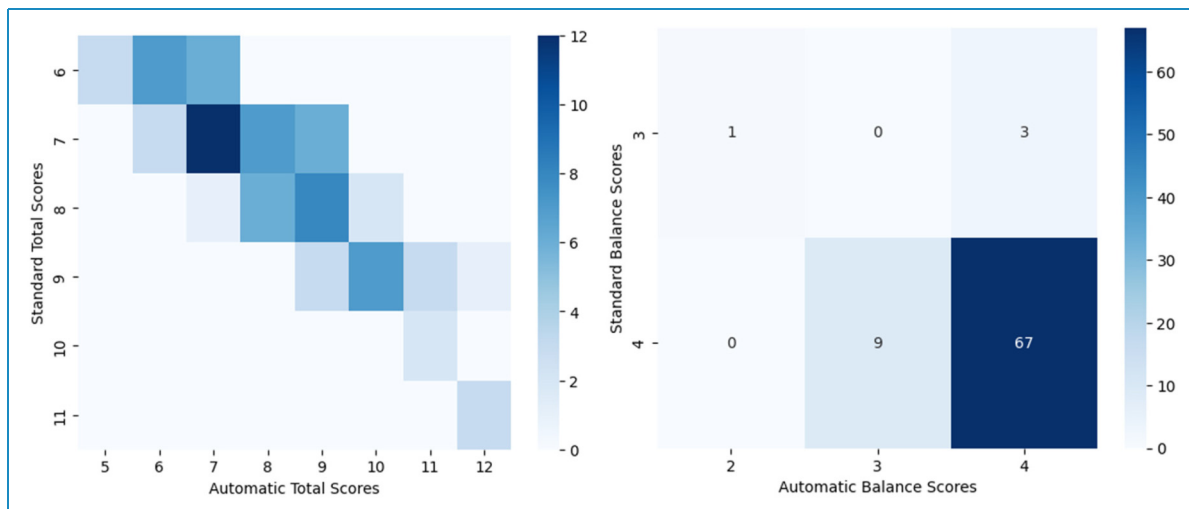


Figure 7. Heatmaps comparing automatic and standard assessments of total battery scores (left) and balance items (right).

measurement. The gait speed scores resulted in an automatic score of 2.388 ± 1.019 (SEM 0.115) and a score of 2.388 ± 1.019 (SEM 0.115) in the standard measurement. The chair stand item results also showed a strong correlation between the automatic and standard assessments (Figure 9), with Spearman's correlation coefficients of 0.945 (CI 95%: 0.91–0.97, $p < 0.001$) for durations and 0.827 (CI 95%: 0.73–0.89, $p < 0.001$) for scores. Durations for the chair stand item were 15.572 ± 3.174 s (SEM 0.357) for the automatic measurements and 18.337 ± 3.748 (SEM 0.422) for the standard measurements. The respective scores yielded 1.975 ± 0.935 (SEM 0.105) for the automatic measurement and 1.500 ± 0.671 (SEM 0.075) for the standard measurement.

Test–retest reliability

A comparison of the total scores for the two subsequent SPPB runs per automatic and standard SPPB evaluation is depicted in Figure 10. Both automatic runs yielded mean scores of 8.1 ± 1.823 (SEM 0.288) and 8.35 ± 1.703 (SEM 0.269). The ICC derived from the app's automated evaluation was 0.936 ($p < 0.001$), with a 95% CI of 0.87–0.97. In contrast, the results from standard assessments yielded an ICC of 0.870 ($p < 0.001$), with a 95% CI of 0.76–0.93. The consecutive standard assessment runs were scored at 7.475 ± 1.219 (SEM 0.193) and 7.7 ± 1.324 (SEM 0.209).

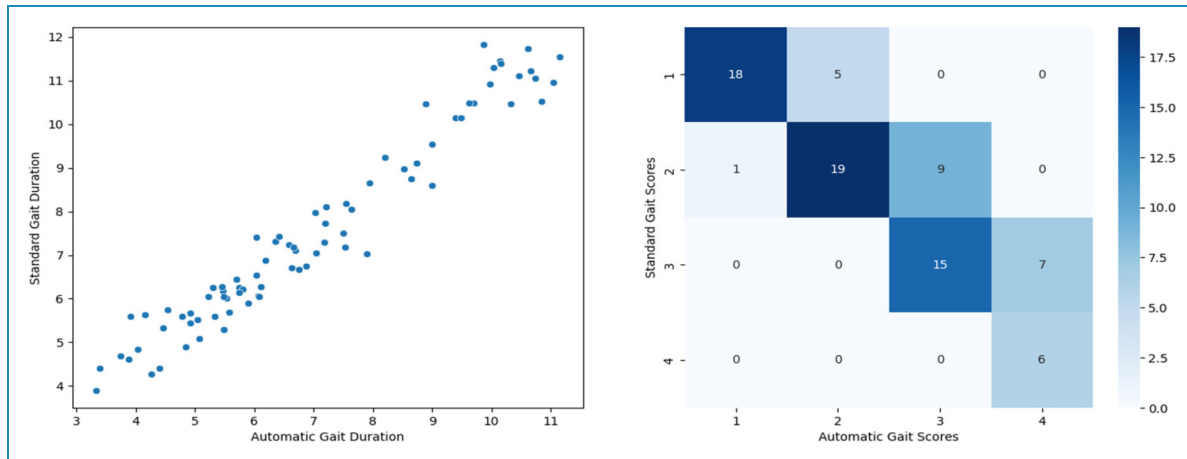


Figure 8. Scatter plot and heatmap comparing standard and automatic assessments for the gait speed item.

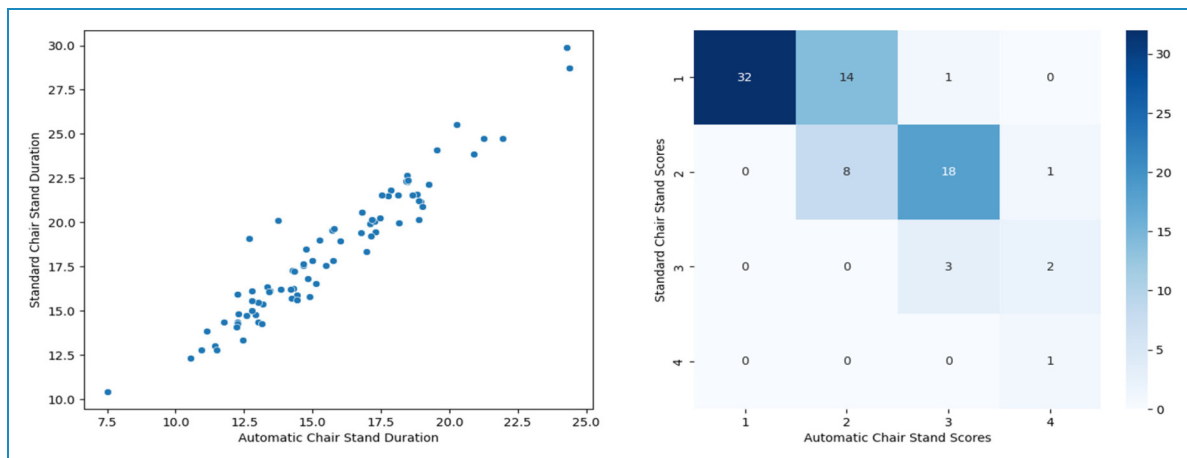


Figure 9. Scatter plot and heatmap comparing standard and automatic assessments for the chair stand item.

In comparing the assessments of the two raters, box plots visually represent the absolute discrepancies in recorded exercise durations as assessed by the two raters for each SPPB item (Figure 11). The median time differences for the balance and chair stand items were less than 0.5 s, whereas for the gait speed item, the median difference surpassed the 1.5 s threshold.

Objectivity

Outliers extending beyond the box plot whisker boundaries revealed inconsistencies between raters reaching up to a maximum of 3.4 s, specifically in the context of the gait speed item (Figure 11).

Prediction of movement patterns

A heatmap illustrates the algorithm's predictions compared with actual conditions simulated by participants wearing the instant aging suits (Figure 12). The system accurately

identified the impairing condition in 72 out of 80 SPPB runs, achieving a detection accuracy of 90%. The obtained AUC score was 0.93, indicating the robust discriminative performance of the classifier across multiple classes. The system correctly predicted the condition simulating hemiplegia for all conducted runs. Only eight misclassifications were observed. Five runs simulating frailty were erroneously predicted as PD, and in three other runs, confusion arose between healthy and PD conditions. Among the incorrect predictions observed, the algorithm showed uncertainty between the correctly and incorrectly selected conditions in six instances, assigning equal scores to both categories (Table 1). In all the cases, the prediction with the second-highest score aligned with the correct condition.

Discussion

This pilot study of the HMC app MobiSPPB showed high validity and reliability compared to the standard rater-based SPPB. It provided objectivity by eliminating rater

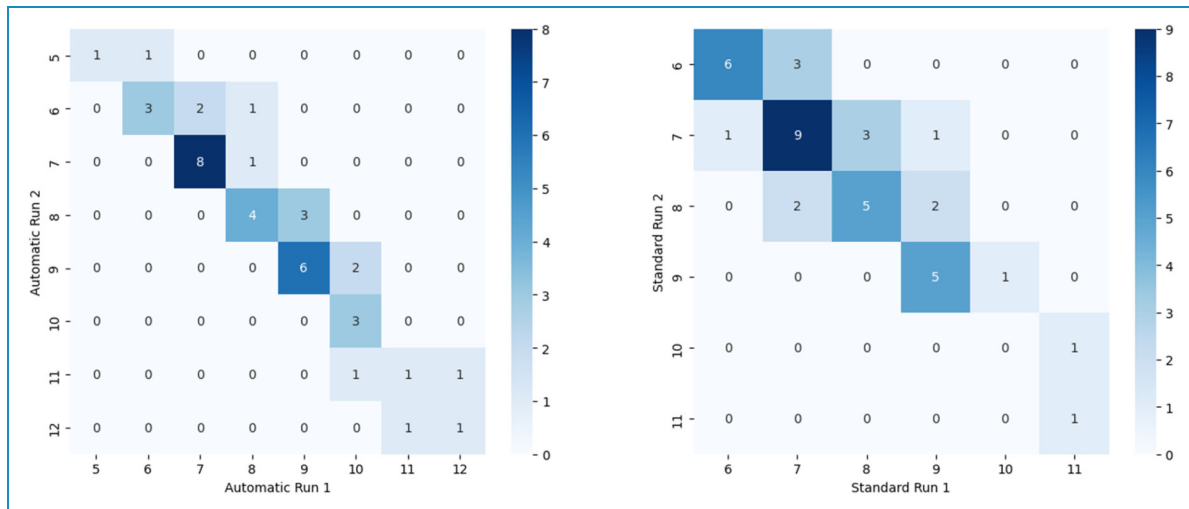


Figure 10. Heatmaps comparing total scores for two consecutive SPPB runs. Left: automatically obtained results. Right: standard assessment outcomes. SPPB: short physical performance battery.

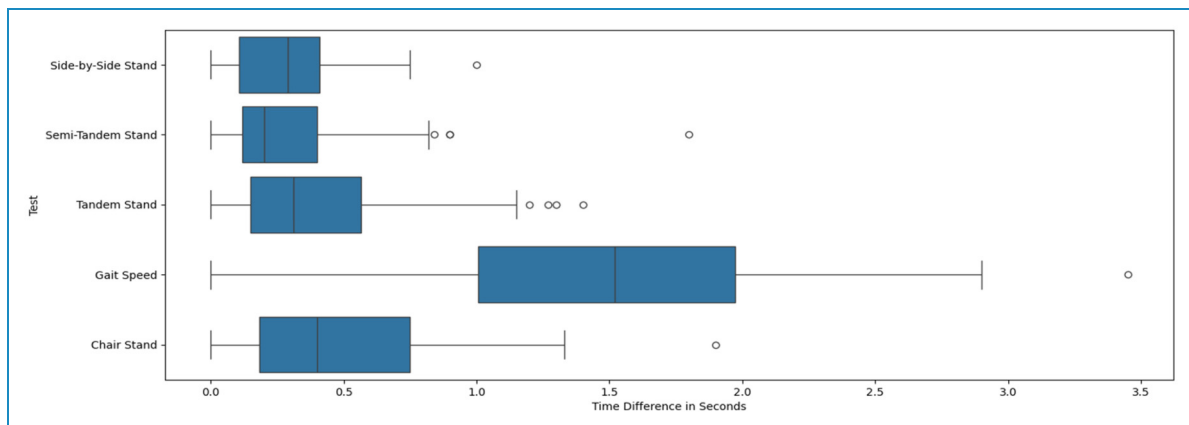


Figure 11. Box plot of the inter-rater differences in the standard assessment.

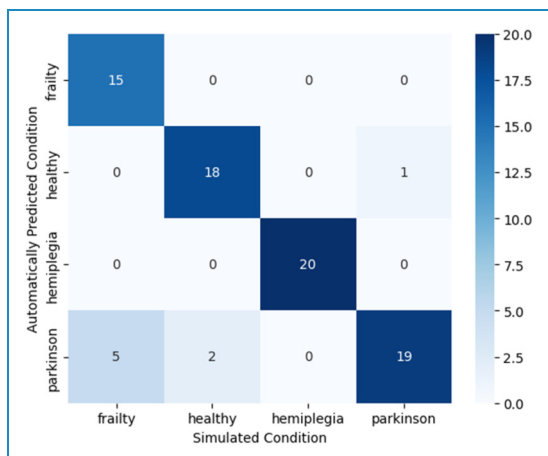


Figure 12. Heatmap of system predictions for simulated conditions.

discrepancies and showed a 90% accuracy in detecting the simulated impairments.

Few studies have investigated the automation of the SPPB. While the multisensor-based SPPB by Jung et al. provides favorable evaluation results, a technical setup with multiple sensors and specialized equipment for individual tests is too complicated for geriatric assessment with patients. This complexity limits its use to scientific investigations. In the motivation of their work, Jung et al. recommend applying their sensor-based SPPB as a criterion for selecting study populations in clinical research.²¹ In contrast, the system by Duncan et al. utilizes HMC techniques with a setup comprising a Raspberry Pi computer, three cameras, two motor-driven cameras, and a smartphone to control the system. The authors acknowledge their system's complex interface, which requires the configuration of technical details.

Table 1. Presumption scores of the algorithm for instances of incorrect predictions.

Simulated condition	Prediction	Healthy	Hemiplegia	PD	Frailty
Frailty	PD	0.26	0.07	0.33	0.33
Healthy	PD	0.33	0.00	0.41	0.26
PD	Healthy	0.34	0.03	0.34	0.28
Frailty	PD	0.26	0.07	0.33	0.33
Healthy	PD	0.33	0.00	0.41	0.26
Frailty	PD	0.24	0.10	0.33	0.33
Frailty	PD	0.22	0.09	0.34	0.34
Frailty	PD	0.24	0.10	0.33	0.33

PD: Parkinson's disease.

Note: Bold values mark the six instances where equal scores were assigned to both categories.

Moreover, concerns arise about the robustness of the algorithms, particularly in the context of analyzing patients deviating from the ideal movement sequence of the SPPB.²²

The MobiSPPB app differs markedly from these two approaches. The app offers not only the time to complete the tests but also an algorithm-based detection of movement abnormalities. Its use is independent of the examiner's experience and skills by providing step-by-step instructions. It is easily applicable and requires only a hand-held or tripod-mounted smartphone. The app has potential even for an assessment at home by a next-of-kin.

According to a systematic review with 31 studies, no single tool can reliably predict falls.³⁵ Only gait speed offers potential as a helpful fall predictor.³⁵ The literature disagrees regarding associations between the SPPB and falls.^{36–39} The calculation of gait speed by the MobiSPPB app may potentially predict falls in geriatric and primary care routine assessments. In Germany, the cost of treating hip fractures, one of the most common fall-related complications, is expected to increase by 128% between 2002 and 2050.⁴⁰ Therefore, reliable and early detection of a patient's risk of falling would be highly beneficial for preventing the consequences of falls and subsequent healthcare costs. To prevent falls in community-dwelling older adults with abnormal MobiSPPB performance, specific physical training regimens, physiotherapy, and dietary plans could be helpful. Research has shown consistently that the intervention with the highest evidence for falls prevention is physical exercise, ideally encompassing functional, strength, and balance training.^{41,42} The use of apps for physical fitness, showed a significant improvement in balance and walking abilities in a meta-analysis of 14 studies by Ambrens et al.⁴³ In the future, the MobiSPPB app might be able to track treatment effects over time.

Limitations

In this experimental setting, the app did not detect any imbalances, because all participants were healthy and able to keep the different balance positions. In 4.6% of the balance tests, the app was confused by the participant's clothing or suit configuration. The rule-based algorithm had difficulties in detecting low-level motoric abnormalities such as small steps typical for PD, which may be overcome by machine learning models in future.

This study was designed as a technical validation of the standard SPPB measures using a motion capture app. The exploratory analysis of movement patterns across the simulated conditions was intended to illustrate the potential of such technology for more detailed biomechanical assessment. However, internal validation and comparisons with alternative methods were not conducted, as the study population consisted of a small sample of healthy young participants using an instant aging suit to simulate mobility impairments. Given that the movement characteristics of older adults are likely to differ substantially (e.g. due to neurocognitive disorders), future research should focus on clinical populations to validate these findings and benchmark the approach against established movement analysis techniques.

Conclusions

The technical validation of the MobiSPPB app was successful. The app proved to be feasible in an experimental setting. It provided objectivity across users and was able to distinguish specific disease-related movement patterns. A follow-up study will assess the validity of the MobiSPPB app in the target population with community-dwelling older adults.

Acknowledgments

The authors owe thanks to all participants for their study support.

ORCID iDs

Lucas Küppers  <https://orcid.org/0009-0000-9253-3020>
 Arezoo Bozorgmehr  <https://orcid.org/0000-0003-3020-1332>
 Stephan Jonas  <https://orcid.org/0000-0002-3687-6165>
 Birgitta Weltermann  <https://orcid.org/0000-0002-1285-1545>
 Lara Marie Reimer  <https://orcid.org/0000-0002-6546-8531>

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the German Federal Ministry of Education and Research, (grant number 01ZZ2022). This publication was supported by the Open Access Publication Fund of the University of Bonn.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Dhillion RJS and Hasni S. Pathogenesis and management of sarcopenia. *Clin Geriatr Med* 2017; 33: 17–26.
2. Sanford AM, Morley JE, Berg-Weger M, et al. High prevalence of geriatric syndromes in older adults. *PLoS One* 2020; 15: e0233857.
3. Salari N, Darvishi N, Ahmadianpanah M, et al. Global prevalence of falls in the older adults: a comprehensive systematic review and meta-analysis. *J Orthop Surg Res* 2022; 17: 334.
4. Homann B, Plaschig A, Grundner M, et al. The impact of neurological disorders on the risk for falls in the community dwelling elderly: a case-controlled study. *BMJ Open* 2013; 3: e003367.
5. Blain H, Masud T, Dargent-Molina P, et al. A comprehensive fracture prevention strategy in older adults: the European Union Geriatric Medicine Society (EUGMS) statement. *Aging Clin Exp Res* 2016; 28: 797–803.
6. Florence CS, Bergen G, Atherly A, et al. Medical costs of fatal and nonfatal falls in older adults. *J Am Geriatr Soc* 2018; 66: 693–698.
7. Sadaqa M, Németh Z, Makai A, et al. Effectiveness of exercise interventions on fall prevention in ambulatory community-dwelling older adults: a systematic review with narrative synthesis. *Front Public Health* 2023; 11: 1209319.
8. Soubra R, Chkeir A and Novella JL. A systematic review of thirty-one assessment tests to evaluate mobility in older adults. *Biomed Res Int* 2019; 2019: 1354362.
9. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol* 1994; 49: M85–M94.
10. Guralnik JM, Ferrucci L, Pieper CF, et al. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *J Gerontol A Biol Sci Med Sci* 2000; 55: M221–M231.
11. Beaudart C, McCloskey E, Bruyère O, et al. Sarcopenia in daily practice: assessment and management. *BMC Geriatr* 2016; 16: 170.
12. Pavasini R, Guralnik J, Brown JC, et al. Short physical performance battery and all-cause mortality: systematic review and meta-analysis. *BMC Med* 2016; 14: 215.
13. Fuentes-Abolafio IJ, Stubbs B, Pérez-Belmonte LM, et al. Physical functional performance and prognosis in patients with heart failure: a systematic review and meta-analysis. *BMC Cardiovasc Disord* 2020; 20: 512.
14. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing* 2019; 48: 16–31.
15. O'Shaughnessy Í, Fitzgerald C, Hayes C, et al. Stakeholders' experiences of comprehensive geriatric assessment in an inpatient hospital setting: a qualitative systematic review and meta-ethnography. *BMC Geriatr* 2023; 23: 821.
16. Seematter-Bagnoud L and Büla C. Brief assessments and screening for geriatric conditions in older primary care patients: a pragmatic approach. *Public Health Rev* 2018; 39: 8.
17. Merchant RA, Hui RJY, Kwek SC, et al. Rapid geriatric assessment using mobile app in primary care: prevalence of geriatric syndromes and review of its feasibility. *Front Med (Lausanne)* 2020; 7: 261.
18. 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: 15.
19. Ortega-Bastidas P, Gómez B, Aqueveque P, et al. Instrumented timed up and go test (iTUG)-more than assessing time to predict falls: a systematic review. *Sensors (Basel)* 2023; 23: 3426.
20. Ponciano V, Pires IM, Ribeiro FR, et al. Is the timed-up and go test feasible in mobile devices? A systematic review. *Electronics* 2020; 9: 528.
21. Jung H-W, Roh H, Cho Y, et al. Validation of a multi-sensor-based kiosk for short physical performance battery. *J Am Geriatr Soc* 2019; 67: 2605–2609.
22. Duncan L, Zhu S, Pergolotti M, et al. Camera-based short physical performance battery and timed up and go assessment for older adults with cancer. *IEEE Trans Biomed Eng* 2023; 70: 2529–2539.
23. Krupp S. für die AG assessment der Deutschen Gesellschaft für Geriatrie e. V., S1-Leitlinie Geriatriisches Assessment der Stufe 2, Living Guideline, Version 11.01.2024, AWMF-Register-Nr. 084-002LG.
24. Menolotto M, Komaris D-S, Tedesco S, et al. Motion capture technology in industrial applications: a systematic review. *Sensors (Basel)* 2020; 20: 5687.

25. Moeslund TB and Granum E. A survey of computer vision-based human motion capture. *Comput Vis Image Underst* 2001; 81: 231–268.
26. Zhong D and Chang S-F. Real-time view recognition and event detection for sports video. *J Vis Commun Image Represent* 2004; 15: 330–347.
27. Apple Machine Learning Research. Deploying transformers on the apple neural engine, <https://machinelearning.apple.com/research/neural-engine-transformers> (accessed 17 May 2024).
28. Bazarevsky V, Grishchenko I, Raveendran K, et al. BlazePose: On-device real-time body pose tracking: arXiv, 2020.
29. Vieweg J and Schaefer S. How an age simulation suit affects motor and cognitive performance and self-perception in younger adults. *Exp Aging Res* 2020; 46: 273–290.
30. Grishchenko I, Bazarevsky V, Zangir A, et al. BlazePose GHUM holistic: Real-time 3D human landmarks and pose estimation: arXiv, 2022.
31. Koo TK and Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15: 155–163.
32. Freire AN, Guerra RO, Alvarado B, et al. Validity and reliability of the short physical performance battery in two diverse older adult populations in Quebec and Brazil. *J Aging Health* 2012; 24: 863–878.
33. Ortega-Pérez de Villar L, Martínez-Olmos FJ, Junqué-Jiménez A, et al. Test-retest reliability and minimal detectable change scores for the short physical performance battery, one-legged standing test and timed up and go test in patients undergoing hemodialysis. *PLoS One* 2018; 13: e0201035.
34. Grandini M, Bagli E and Visani G. Metrics for multi-class classification: an overview: arXiv, 2020.
35. Beck Jepsen D, Robinson K, Ogliari G, et al. Predicting falls in older adults: an umbrella review of instruments assessing gait, balance, and functional mobility. *BMC Geriatr* 2022; 22: 615.
36. Lauretani F, Ticinesi A, Gionti L, et al. Short-physical performance battery (SPPB) score is associated with falls in older outpatients. *Aging Clin Exp Res* 2019; 31: 1435–1442.
37. Welch SA, Ward RE, Beauchamp MK, et al. The short physical performance battery (SPPB): a quick and useful tool for fall risk stratification among older primary care patients. *J Am Med Dir Assoc* 2021; 22: 1646–1651.
38. Li W, Rao Z, Fu Y, et al. Value of the short physical performance battery (SPPB) in predicting fall and fall-induced injury among old Chinese adults. *BMC Geriatr* 2023; 23: 574.
39. Freiburger E, Fabbietti P, Corsonello A, et al. Short physical performance battery is not associated with falls and injurious falls in older persons: longitudinal data of the SCOPE project. *Eur Geriatr Med* 2024; 15: 831–842.
40. Konnopka A, Jerusel N and König H-H. The health and economic consequences of osteopenia- and osteoporosis-attributable hip fractures in Germany: estimation for 2002 and projection until 2050. *Osteoporos Int* 2009; 20: 1117–1129.
41. Guirguis-Blake JM, Perdue LA, Coppola EL, et al. Interventions to prevent falls in older adults: updated evidence report and systematic review for the US preventive services task force. *JAMA* 2024; 332: 58–69.
42. Sherrington C, Fairhall N, Wallbank G, et al. Exercise for preventing falls in older people living in the community: an abridged Cochrane systematic review. *Br J Sports Med* 2020; 54: 885–891.
43. Ambrens M, Alley S, Oliveira JS, et al. Effect of eHealth-delivered exercise programmes on balance in people aged 65 years and over living in the community: a systematic review and meta-analysis of randomised controlled trials. *BMJ Open* 2022; 12: e051377.