Video game-based application for fall risk assessment: a proof-of-concept cohort study

Antao Ming,[a,c] Tanja Schubert,[a,c] Vanessa Marr,[a] Jaqueline Hötzsch,[a] Sebastian Stober,[b] and Peter R. Mertens[a,*]

[a]University Clinic for Nephrology and Hypertension, Diabetes and Endocrinology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany
[b]Artificial Intelligence Lab, Otto-von-Guericke University Magdeburg, Magdeburg, Germany

eClinicalMedicine
2024;78: 102947
Published Online xxx
https://doi.org/10.1016/j.eclinm.2024.102947

Summary

Background Fall(s) are a significant cause of morbidity and mortality especially amongst elderly with polyneuropathy and cognitive decline. Conventional fall risk assessment tools are prone to low predictive values and do not address specific vulnerabilities. This study seeks to advance the development of an innovative, engaging fall prediction tool for a high-risk cohort diagnosed with diabetes.

Methods In this proof-of-concept cohort study, between July 01, 2020, and May 31, 2022, 152 participants with diabetes performed clinical examinations to estimate individual risks of fall (timed “up and go” (TUG) test, dynamic gait index (DGI), Berg-Balance-Scale (BBS)) and participated in a video game-based fall risk assessment with sensor-equipped insoles as steering units. The participants engaged in four distinct video games, each designed to address capabilities pertinent to prevent fall(s): skillfulness, reaction time, sensation, endurance, balance, and muscle strength. Data were collected during both, seated and standing gaming sessions. By data analyses using binary machine learning models a classification of participants was achieved and compared with actual fall events reported for the past 24 months.

Findings Overall 22 out of 152 participants (14.5%) underwent at least one episode of fall during the past 24 months. Adjusted risk classification accuracies of TUG, DGI, and BBS reached 58.7%, 58.3%, and 47.5%, respectively. Data analyses from gaming sessions in seated and standing positions yielded two models with six predictors from the four games with accuracies of 82.8% and 88.6% (area under the receiver-operating-characteristic curve 0.84 (95% confidence interval (CI): 0.77–0.91) and 0.91 (95% CI: 0.85–0.97), respectively). Key capabilities that were distinctly different between the groups related to endurance (0.6 ± 0.1 vs. 0.5 ± 0.2; $p = 0.03$) and balance (0.7 ± 0.2 vs. 0.6 ± 0.2; $p = 0.05$). The AI-driven analysis allowed to extract a list of game features that showed highly significant predictive values, e.g., reaction times in specific task, deviation from ideal steering routes in parcours and pressure-related parameters.

Interpretation Thus, video game-based assessment of fall risk surpasses traditional clinical assessment tools and scores (e.g., TUG, DGI, and BBS) and may open a novel resource for patient evaluation in the future. Further research with larger, heterogeneous cohorts is needed to validate these findings and especially predict future fall risk probabilities in clinical as well as outpatient settings.

Funding This project was funded by the Ministry of Science, Economics, and Digitalization of the State of Saxony-Anhalt and the European Fund for Regional Development under the Autonomy in Old Age Program (Funding No: ZS/2016/05/78615, ZS/2018/12/95325) and Healthy Cognition and Nerve function (HeyCoNer, ZS/2023/12/183088).

Copyright © 2024 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Timed “up and go” test; Dynamic gait index; Berg-Balance-Scale; Fall risk assessment; Diabetes mellitus; Video games; Machine learning algorithms

*Corresponding author. University Clinic for Nephrology and Hypertension, Diabetes and Endocrinology, Otto-von-Guericke University Magdeburg, Leipziger Str. 44, 39120, Magdeburg, Germany.
E-mail address: peter.mertens@med.ovgu.de (P.R. Mertens).
[c]Contributed equally to the study.

Research in context

Evidence before this study

We searched PubMed database, from January 1, 2000, to March 31, 2024, for papers published in English using the following title search terms: "fall risk assessment" and "diabetes", "fall risk prediction" and "diabetes", "fall risk assessment" and "diabetic", and "fall risk prediction" and "diabetic". Our search yielded three results, highlighting that falls are a leading cause of morbidity and mortality in the elderly adults with diabetes, who are at heightened risk due to factors such as hypoglycemia, neuropathy, and cognitive decline. Traditional assessment tools for fall risk like the Timed "Up and Go" (TUG) test and Dynamic Gait Index (DGI) are frequently used but show limited predictive accuracy in high-risk populations, particularly those with diabetes. Emerging sensor-based technologies have shown promise in fall detection, but few studies have combined these tools with machine learning in diabetic cohorts.

Added value of this study

Our study introduces a novel fall risk assessment method that combines video game-based tasks with sensor-equipped insoles and machine learning, which is completed within ~15 min. This approach demonstrated high predictive accuracies when performed in seated (82.8%) and standing positions (88.6%) and outperformed traditional tools such as TUG, DGI, and the Berg Balance Scale. Our method provides detailed, AI-driven insights into fall risk factors such as balance, endurance, and reaction time. Thus, a novel comprehensive assessment tool for fall risk is introduced, that may be especially useful for elderly high-risk diabetes patients.

Implications of all the available evidence

This innovative video game-based approach has the potential to improve fall risk assessment in older populations with diabetes. The fast and standardized performance may transform screening for at risk individuals. Future research should validate these findings in large, diverse cohorts and pinpoint underlying conditions that determine fall risk (medication, cognition, polyneuropathy).

Introduction

Falls represent the second leading global cause of mortality due to accidental and unintentional injuries.[1,2] The incidence of falls amongst the elderly aged ≥65 years sums up to 25% experiencing at least one fall annually.[3] This risk is further elevated in individuals with chronic diseases that impact joint integrity, muscle strength, balance, cognitive function, resulting in frailty.[4] Specifically, patients with diabetes are subject to additional risk factors, such as hypoglycemia,[5] loss of protective sensation from peripheral neuropathy,[6] visual impairment,[7] altered glucose levels due to insulin and antidiabetic therapy.[8,9] As a result, this cohort of patients exhibits an even higher incidence compared to their non-diabetic peers.[10,11]

Over the past decades, various assessment tools have been devised and described to predict fall risk, including the Hendrich II Fall Risk Model,[12] Comprehensive Geriatric Assessment,[13] timed "up and go" (TUG) test,[14] Dynamic Gait Index (DGI),[15] and Berg-Balance-Scale (BBS).[16] These methods often lack the sensitivity needed to detect subtle changes in fall risk.[17] Moreover, these approaches may not adequately address the unique vulnerabilities inherent in patients with diabetes with the combination of confounding factors (e.g., hypoglycemia,[5] peripheral neuropathy,[18] cognitive decline[19]), highlighting the need for innovative solutions.[8]

Recent advancements in sensor-based technologies have garnered significant attention for the assessment of fall risk in older adults.[20] Wearable devices, such as accelerometers and gyroscopes, are now widely employed for real-time monitoring of gait and balance. These devices enable continuous data acquisition, providing detailed insights into mobility patterns outside traditional clinical environments.[21] Additionally, in–home monitoring systems and smartphone-based applications further extend the ability to detect irregularities in movement patterns associated with fall risk.[22,23] Such sensor-driven tools allow for interpretation of movement and balance, offering a more comprehensive view on patient health than what is typically captured during periodic clinical evaluations.[24]

Building on these innovations, interactive video game applications have been increasingly integrated with sensor technologies to assess and enhance balance and coordination.[25] Such applications utilize sensors embedded in footwear or motion-tracking devices to capture detailed information on motor function, which may not be evident during standard clinical assessments.[22,26] In this project our strategy is not to focus on movements or gait changes and instead design a setup where the participants are in a seated or standing position with the instruction to make adaptations to the plantar pressures recorded by sensors to steer a video game.

In this proof-of-concept study, we apply established clinical risk assessment scores for fall(s) in elderly patients and compare these results with a video game-based approach. In the latter, distinct features are extracted out of a 15 min lasting parcours and analyzed by means of machine learning algorithms. The overall findings were correlated with actual events reported for the past 24 months.

Methods

Ethics

The study was conducted following approval of the study protocol by the local ethical committee of the Otto-von-Guericke University Magdeburg, Germany (approval 28/17; March 13, 2017). All study participants provided written informed consent upon a detailed explanation of the study protocol, test procedure, and data policy.

Study design and participants

In the proof-of-concept study, a game-based assessment platform was conceptualized, developed, and integrated into a clinical investigation at the Clinic of Nephrology and Hypertension, Diabetes and Endocrinology at the Otto-von-Guericke University Magdeburg, Germany (Supplementary Figure S1). All participants were also enrolled in the Smart Prevent Diabetic Feet study.[27,28]

During the period extending from July 01, 2020 to May 31, 2022, the patient cohort underwent a rigorous screening process. Participants were furnished with a comprehensive briefing regarding the study methodology, evaluative procedures, and the data governance policies underpinning the research, subsequent to which written informed consent was procured. Inclusion criteria stipulated that only individuals diagnosed with diabetes mellitus aged 18 years or more were considered eligible for the study. Exclusion criteria were meticulously delineated to preclude the participation of individuals presenting with any of the following conditions: significant macroangiopathy localized to the lower extremities; any form of physical deformity (inclusive of amputations and limb malformations necessitating the use of orthopedically modified footwear); active neuropathic ulceration of the foot; visual impairments characterized by a visual acuity metric less than 0.8, with the caveat that corrective measures for myopia and hyperopia were permissible; any form of myopathy or neuromotor pathology; a history of myocardial infarction within the preceding 12 weeks; heart failure categorized as New York Heart Association Functional Classification III or IV; recent transient ischemic attacks or cerebrovascular accidents; clinically significant tremor; or any impediment to granting consent or utilizing a mobile phone, for whatever cause.

Clinical examinations and fall risk assessment

At time of enrolment a comprehensive questionnaire was conducted to elicit detailed information on the patient's medical history, diabetes mellitus (type, duration, treatment regimen, sensory abnormalities, and daily movement limitations), manifestations of autonomic diabetic neuropathy (such as dizziness, cardiac arrhythmias, urinary anomalies, and altered sweating patterns), diabetes-related comorbidities, and physical activities, including sports involvement, handedness, and lower limb preference (Supplementary Figure S1).

The presence of peripheral neuropathy was assessed by a study physician using the German iterations of scores for Neuropathy Disability (NDS) and Neuropathy Symptoms (NSS).[29–31] These validated instruments represent a standard in clinical evaluation with grading of neuropathic impairments and symptoms, respectively. Peripheral neuropathy diagnosis was contingent upon an NDS score of 6 or higher, or a NDS score of 3 or more concurrent with an NSS score of 5 or above, in alignment with the German Diabetes Association's clinical guidelines.[29,32,33] Cognitive function was gauged using the Montreal Cognitive Assessment (MoCA), a 30-point scale where higher scores denote superior cognitive abilities.[34] An additional point was conferred to individuals with education levels of 12 years or less. Scores below 26 were indicative of cognitive decline.[35]

A questionnaires about fall incidents was set up. The participants were asked about events within the last 24 months and the likely causes thereof. Data were also collected on the use of walking aids, the maximum walking distance unaided and without rest, usage of handrails, balance difficulties in the dark, and dizziness with head movements or positional changes. A translated version of this questionnaire is provided in the Supplementary appendix.

Study device and game-based application

The gaming setup for each study participant consisted of a pair of size-matched slippers embedded with sensor-equipped insoles (ActiSense System®, IEE S.A. Luxembourg), an Android tablet (Samsung Galaxy Tab A7 SM-T500) with a video game application, and a set of headphones (Fig. 1a). A detailed explanation of the study setup has been published.[36,37] Briefly, the insoles serve as control units for steering of video games with eight pressure sensors positioned at defined positions within the heel, lateral arch, metatarsals 1, 3, and 5, hallux, and toes. These sensors quantify plantar pressures with a sensitivity of 3.4 mbar, ranging from 250 mbar to 7 bar, and a maximum sampling rate of 500 Hz. The electronic control unit facilitates data synchronization with smart devices, automatic foot side detection, 16 GB internal storage, and a 10-h energy supply. The insole is covered with a protective sponge material and integrated into the gaming slipper with a supportive ethylene-vinyl acetate (EVA-50) layer below (Fig. 1b).

During the gaming session, the insole is initialized to measure plantar foot pressures at 200 Hz and transmit the data to the game application installed on the Android tablet via Bluetooth 5.0 in real-time. The setup allows the participant to play video games both in a seated and standing position by modulating plantar foot pressure values (Fig. 1c and d). Each gaming session included four games.

Apple-Catch: Participants operate a carriage in an autumnal scenario using plantar pressure to catch apples falling from a tree (Fig. 1e). Precision in

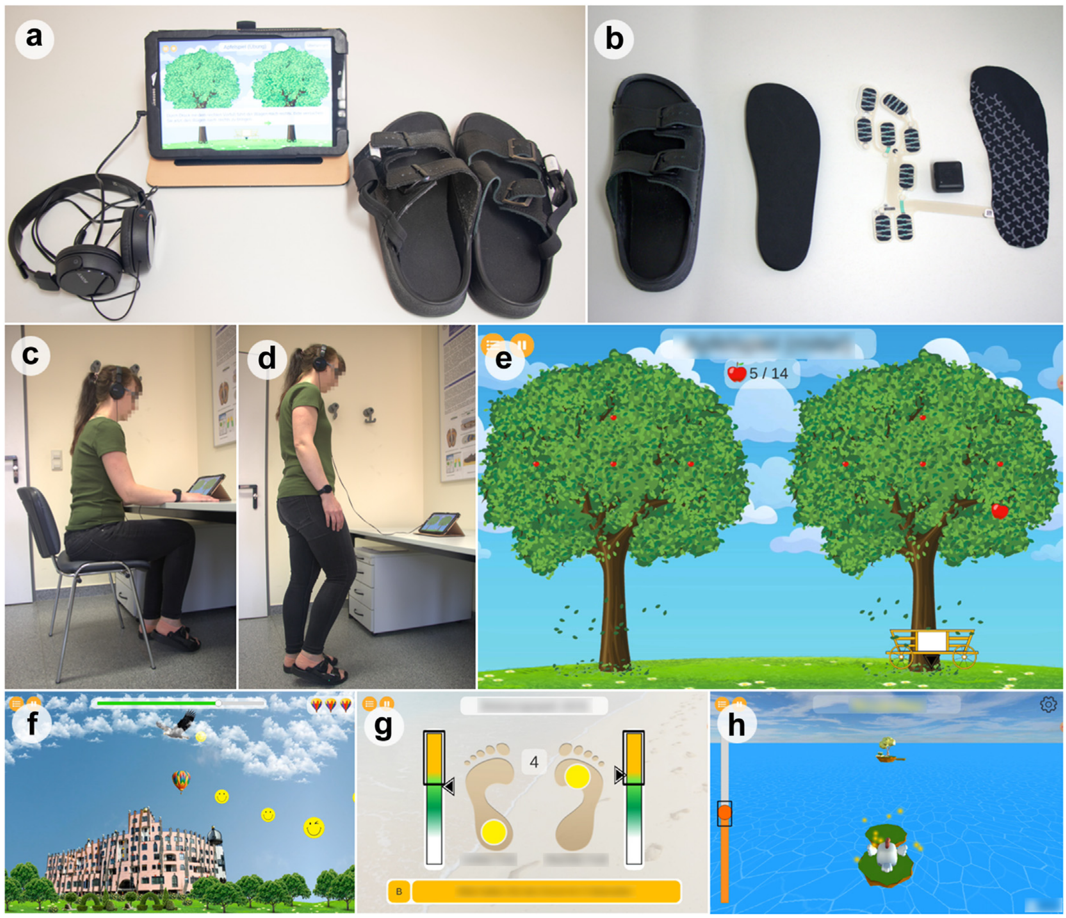


Fig. 1: **Video game-based assessment of fall risk using sensor-equipped insoles in seated and standing positions.** (a) Overview of the materials provided to each study participant, including a pair of gaming slippers, an Android tablet with a video game application connected via Bluetooth and a set of headphones. The setup enables participants to control video games through the modulation of plantar pressure values. (b) Illustration of the sensor-equipped insole integrated into the footwear (from left to right): gaming slipper, supportive ethylene-vinylacetat-50 layer, sensor layer, electronic control unit, and protective sponge layer. (c and d) Participants were instructed to perform the initial gaming session in a seated position, followed by the same set of games in a standing position without contacting chairs or tables. Each session comprised four distinct video games: (e) Apple-Catch, (f) Balloon-Flying, (g) Cross-Pressure, and (h) Island-Jump.

maintaining target pressure is required for optimal carriage positioning. After each attempt, the carriage automatically returns to a central baseline.

Balloon-Flying: Participants maneuver a balloon across a skyline, with altitude controlled by varying forefoot pressure (Fig. 1f). Players avoid obstacles and collect points represented by smileys, requiring quick adjustments to avoid collisions and restarts.

Cross-Pressure: Participants apply pressure as indicated by green (low) and yellow (high) markers on specific foot areas for set durations (Fig. 1g). Achievements are confirmed through graphical indicators, and the game advances if no valid response is recorded within 25 s.

Island-Jump: Participants guide a virtual bird across islands using plantar pressure to control jump distances and directions (Fig. 1h). Failure to adjust pressure within the indicated range results in the bird falling and the need to repeat the attempt. Optimal scores require adherence to predefined pressure thresholds, with deviations prompting restarts.

Collectively, these games engage participants in complex tasks requiring nuanced control of foot pressure, facilitating the assessment of balance, postural stability, and fall risk in an interactive manner. Gaming data are transferred to a central study server for analyses using predefined hypotheses[36,37] and for feature extraction using machine learning algorithms (detailed below). An introductory video on the game-based application is provided in the Supplementary Materials.

Video gaming session for fall risk evaluation

Each participant played four consecutive games, initially in a seated position (Fig. 1c), followed by an optional repetition in standing position (Fig. 1d). Headphones were provided to enhance patient attention to the game, precluded interactions with the study personel and to ensure standardization of the test procedure. Prior to the seated gaming session, participants underwent a process of subject-level data normalization that involved 12 standardized steps aimed at adjusting pressure measurements within a range of 0–1 (Supplementary Figure S2). As first step the participants were seated in a balanced position and apply minimal and maximal pressures to the insoles for 5 s each. Furthermore, participants were instructed to stand up for 5 s and then maintain balance on both or single feet for an additional 5 s. Areas of pressure application were again indicated by yellow/orange/red color. The entire pressure normalization process lasted about 2 min.

Thereafter, participants listened to standardized tutorials and engaged in interactive practice of pressure application to familiarize themselves with the game controls and setup. Each gaming session lasted about 15 min (Fig. 1e–h). Further details regarding pressure normalization and game setup have been published in our previous work[36,37] and are also available in the Supplementary Materials.

Safety was a priority during the game-based assessments, and all sessions were conducted under the supervision of trained personnel. Participants were closely monitored to ensure that they maintained their balance throughout the standing session. No adverse events were recorded during the study.

Game feature extraction

Following the gaming sessions, distinct parameters were extracted from the acquired data sets to evaluate the performance of study participants.[36,37] Specific parameters were delineated for each game task and utilized to evaluate game performance.

In the Apple-Catch game, reaction time (s) measures the time elapsed from the moment the apple starts falling until the participant initiates movement to adjust the carriage. A shorter reaction time indicates faster sensorimotor coordination, crucial for timely adjustments in response to environmental stimuli. Anticipation time (s) captures the time the participant spends preparing for the apple's descent before making any carriage movement. Prolonged anticipation time may reflect hesitancy or over-cautiousness, possibly indicative of reduced confidence in motor control or cognitive slowing. Time inside the catching area (s) measures the duration for which the apple remains within the designated target area, indicating the participant's ability to align the carriage with the apple's path. Longer times suggest improved balance and precise motor control. Time outside the catching area (s) represents the time the apple spends outside the optimal catching zone. Increased time outside the area may indicate difficulty in maneuvering the carriage, potentially signaling impaired coordination. Frequency outside the catching area (n) records the number of times the apple exits the target area as it falls. Higher frequencies reflect greater instability or challenges in maintaining carriage alignment, which may be linked to motor impairment. Final virtual distance measures the distance between the final apple position and the center of the target area. Greater distances indicate poorer precision in positioning the carriage, highlighting deficiencies in motor accuracy or control. Apple caught (yes/no) is a binary indicator of task success, assessing whether the apple was successfully caught by the carriage. This feature provides a direct measure of task performance and motor coordination. Normalized pressure represents the pressure applied by the participant's feet on the sensors during task performance, normalized over time. Maintaining appropriate pressure is essential for accurate carriage movement, with deviations potentially reflecting imbalances or incorrect weight distribution. Pressure difference between successive frames captures the change in pressure between consecutive time frames. Large fluctuations could indicate instability in the participant's balance or uneven weight shifts during the task. Pressure gradient between successive frames examines the rate of pressure change across successive frames. A smoother gradient reflects more controlled, stable movements, whereas abrupt changes may suggest poor balance or motor control. Pressure time integral represents the cumulative pressure applied over the course of the task. It reflects the overall force exerted, which can provide insights into how participants distribute weight over time to control the carriage.

For the Balloon-Flying game, the feature smiley count (n) captures the number of smiley faces collected by the participant during the task. Smiley collection reflects the player's ability to navigate accurately and effectively, with higher counts indicating better coordination and task performance. Collision frequency (n) records the number of times the balloon collides with obstacles during the flight. A higher collision frequency suggests difficulty in controlling the balloon's altitude, potentially indicating reduced motor control or slower reaction times. Minimal virtual distance to smiley 1–4 measure the closest distance between the balloon and each smiley during the task. A smaller distance indicates better precision in maneuvering the balloon to collect smileys without excessive deviations. Virtual deviation from ideal flying route quantifies how much the balloon deviates from the optimal flight path. A larger deviation indicates greater difficulty in maintaining the desired trajectory, potentially signaling balance or coordination issues. Normalized pressure represents the level of pressure applied by the participant's forefoot to control the balloon's altitude. Maintaining appropriate pressure is crucial for smooth altitude adjustments. Pressure difference between successive frames measures how the pressure changes between consecutive time frames. Greater pressure variability may indicate instability in control or difficulty maintaining steady pressure. Pressure gradient between successive frames captures the rate of change in pressure across time. A smoother gradient indicates better control, while abrupt changes suggest a lack of coordination. Pressure time integral records the total pressure applied over the duration of the task, reflecting the participant's effort and consistency in pressure application.

In the Cross-Pressure game, anticipation time (s) captures the time elapsed between the initiation of the task and the participant's first pressure application. Shorter anticipation times indicate faster response and reaction abilities, while longer times may suggest delays in motor coordination or hesitation. Time outside optimal pressure zone (s) measures the amount of time the participant spends applying pressure that deviates from the target zone (either too high or too low). Longer durations outside the optimal zone may indicate difficulty in maintaining the required pressure, reflecting poor motor control. Relaxation time (s) captures the time required for the participant to release pressure after applying the target force. A shorter relaxation time indicates better control over the release phase, whereas longer times may suggest slow motor function. Normalized pressure (left or right foot) represents the actual pressure applied by each foot, normalized over time. This parameter evaluates how accurately the participant can maintain the required pressure for a given task. The normalized pressure provides insights into overall motor performance and precision. Pressure difference between successive frames (left or right foot) captures the variation in pressure between consecutive time frames during the task. Large pressure differences may indicate inconsistent control or an inability to smoothly transition between pressure levels. Pressure gradient between successive frames (left or right foot) presents the rate of change in pressure over time, indicating the smoothness or abruptness of the pressure application. A smooth gradient suggests controlled and steady pressure application, while sharp gradients may reveal instability or difficulty in maintaining target pressure. Pressure time integral (left or right foot) records the cumulative pressure applied over the entire duration of the task. This metric reflects the participant's effort in applying and maintaining pressure, and it helps assess their ability to consistently meet the required target over time.

For the Island-Jump game, attempt count (n) records the number of attempts made to successfully guide the bird from one island to another. Multiple attempts may indicate difficulty in controlling pressure or reacting to the task demands. Deviation from optimal pressure (%) captures how far the applied pressure deviates from the predefined optimal range required to successfully complete the jump. A greater deviation suggests that the participant is struggling to maintain the required force, leading to potential restarts. Anticipation time (s) captures the time between task initialization and the participant's first pressure application. Shorter anticipation times reflect faster cognitive-motor responses, while longer times may indicate hesitation or slower reaction times. Execution time (s) is the time taken by the participant to apply the required pressure to guide the bird across the islands. Efficient execution is essential to avoid deviations that cause the bird to fall. Mean pressure during execution phase captures the average pressure applied during the execution phase of the jump. Maintaining pressure close to the predefined target is critical for successfully completing the task. Pressure difference between successive frames represents the variation in pressure applied between consecutive frames. A higher difference indicates instability or inconsistency in maintaining the target pressure during the task. Pressure gradient between successive frames captures the rate of pressure change over time. A smooth gradient suggests controlled and gradual pressure adjustment, while abrupt changes may reflect difficulty in maintaining steady control. Pressure time integral measures the cumulative pressure applied over the entire task duration. This metric provides insights into how consistently the participant applies pressure throughout the task, which is essential for understanding their ability to maintain motor control.

In summary, the aforementioned parameters were deemed as the principal metrics for game performance. Additionally, the concept of task combinations was introduced, wherein sets of game tasks with similar specifications were grouped. Taking Apple-Catch game as an example, task combination for the left foot (TCL1) encompassed all tasks requiring the left foot to control carriage movement for apple collection, while task combination for the right foot comprised tasks that relied on the right foot. The sum, mean, and standard deviation of primary features across game tasks within task combinations were classified as secondary features. For example, in the Apple-Catch game the reaction time for TCL1 was considered a secondary feature, calculated as the average reaction time across tasks involving the left foot. Overall, a total of 5244 distinct features reflecting the players' performance across the four games were extracted for each participant's gaming data set.

Statistics

Descriptive statistics for categorical variables are reported as proportions and frequencies. For continuous variables, data are summarized using the mean and standard deviation (SD). To compare groups, Chi-square tests are applied to categorical variables. Under finite sample conditions additional correction methods (e.g., Yates' continuity correction) and Fisher's exact tests were employed to improve the reliability of the test results. The Shapiro–Wilk test assesses the normality of continuous variables. For normally distributed variables, t-tests determine group differences; for non-normally distributed variables, Mann–Whitney U tests are employed. Comparisons across multiple groups utilize either the Kruskal–Wallis H test or one-way analysis of variance, as appropriate. Statistical significance is defined as two-sided P values below 0.05, with pairwise comparisons among multiple groups adjusted using the Holm-Bonferroni method.[38] Correlation and linear regression analyses are conducted to explore relationships between variables, using Pearson or Spearman correlation coefficients based on data distribution. Data analysis is performed using R programming language (version 4.2.1; R Foundation for Statistical Computing) and related open-source libraries as well as Zstats (version 1.0; Hangzhou Yunxiang Statistical Technology Co, Ltd). A comprehensive list of R packages is included in the Supplementary Materials.

Machine learning algorithms

Multiple algorithms were employed to deal with limited and imbalanced data sets where predictors far outnumber observations, such as gradient boosting machine (GBM), lasso and elastic-net regularized generalized linear models (GLMNET), k-nearest neighbor (KNN), penalized logistic regression (PLR), random forest (RF), and support vector machines (SVM). The algorithms were selected for its unique capability to analyze and interpret intricate data structures efficiently. Notably, the elastic net algorithm, as implemented in the GLMNET package,[39] is particularly advantageous in scenarios where the number of predictors vastly exceeds the number of observations. This method employs a regularization technique controlled by the elastic net penalty parameter alpha, which can be adjusted to emphasize either L1 regularization (lasso regression for variable selection) or L2 regularization (ridge regression for parameter shrinking).[40] GBM stood out for its sequential learning method, systematically correcting mistakes from previous models. This characteristic is invaluable for imbalanced datasets, as it focuses on refining predictions for harder-to-identify instances, thereby improving overall model accuracy. Its flexibility and iterative refinement process effectively tackle the dual challenges of data imbalance and the abundance of predictors, revealing essential patterns within the data.[41]

Model development for fall risk prediction

The primary aim of this study was to classify patients into the categories “without” (w/o falls) or “with” a history of fall(s) using game features extracted from the acquired data sets. The seated and standing game sessions challenge different physiological and biomechanical aspects, with seated tasks focusing on muscle control and cognitive-motor coordination, and standing tasks emphasizing balance-related issues. Therefore, the data from seated and standing positions were fitted separately to better evaluate posture-specific risks associated with falls.

In the selection process of “variables” we retained only one feature from each set of highly correlated variables with correlation coefficients >0.75 to ensure that the retained features contribute uniquely and independently to the model.[42] The selected features were then ranked using the “varImp” function from the R Caret package.[43] This function estimates variable importance by measuring each feature’s contribution to the model’s predictive performance, which varies in dependence of the model type. For example, in random forests, importance is calculated by measuring the accuracy drop when a variable is permuted in the out-of-bag data. In linear models, it is based on the absolute value of the t-statistics of the coefficients, while in boosted trees, the importance is aggregated over all boosting iterations.

Thereafter, in line with widely recommended feature selection techniques[44,45] and our previous experience from related work,[36,37] we performed model tuning to determine the optimal number of features that maximize cross-validated performance while minimizing model complexity. Including more features would introduce noise or multicollinearity, while focusing on the top features the model ensures an interpretable and efficient algorithm without compromising predictive power. As a result, we retained the top six ranked variables as the final predictors (“most important”) for the classification model (Supplementary Figure S3).

Due to the proof-of-concept nature of the study and the limited data available, the test data set was not separately partitioned from the entire data set. Instead, a three-fold cross-validation repeated ten times was employed during the training phase to mitigate overfitting and provide a more accurate estimate of model performance. This approach involved repeatedly partitioning the training data set into three subsets of approximately equal size, ensuring each subset contained the same proportion of labels as the complete dataset. In each iteration, two subsets were used for model training, and the remaining subset was used for validation.

Different feature combinations and multiple aforementioned classifiers were evaluated during the modeling process using the Cohens Kappa as the performance metric due to the unbalanced nature of the acquired data. Youden’s Index was utilized to select the optimal predicted probability cut-off for the calculation of adjusted accuracy.[46]

Role of funding source

The Ministry of Science, Economics, and Digitalization of the State of Saxony-Anhalt and the European Fund for Regional Development provided financial support for this study under the Autonomy in Old Age Program

(Funding Nos.: ZS/2016/05/78615, ZS/2018/12/95325) and the Healthy Cognition and Nerve Function Program (HeyCoNer, ZS/2023/12/183088). The Medical Faculty of the Otto-von-Guericke University Magdeburg served as the primary sponsor. The funders had no role in the study design, data collection, data analysis, data interpretation, or the writing of the report.

Results

Characteristics of study participants

A total of 152 patients diagnosed with diabetes were included in the study. The demographic characteristics and clinical profiles of the study cohort, along with the results of clinical assessments for fall risks (Romberg's test, Unterberger's stepping test, TUG, DGI, and BBS), are summarized in Table 1. The mean age was 65.0 years (SD 10.5) with a preponderance of male participants (95 out of 152, 62.5%), the average diabetes duration was 17.6 years (SD 13.5). 114 out of 152 patients (75.0%) were diagnosed with diabetes type 2. The majority of participants were aged 50 years or older, with nine out of 152 participants (5.9%) under the age of 50 years. All participants engaged in the gaming session while seated (n = 152, Fig. 1c), in addition 102 (67.1%) participated the same gaming session in a standing position immediately following the seated session (Fig. 1d, Supplementary Table S1). The remainder 50 participants (32.9%) opted out of the optional standing session due to time constraints. Analysis showed no significant difference in the number of falls between participants who completed only the seated session (n = 50) and those who completed both sessions (n = 102, Supplementary Table S2). No clear association was found between skipping the standing session and fall risk. For the participants performing both, seated and standing gaming sessions, 29 out of 102 (28.4%) were female at a mean age of 63.9 years (SD 11.9) with an average diabetes duration of 17.7 years (SD 14.1). 72 out of 102 patients (70.6%) were diagnosed with type 2 diabetes.

At the time of the gaming session, all participants were asked about fall(s), circumstances related to the events, as well as consequences thereof. Overall 22 participants (14.5%) reported 38 events, in four participants more than one event occurred (Table 2). The types of fall(s) were categorized as stumbling (n = 16), slipping (n = 3, due to black ice), and fainting (n = 2), according to the criteria introduced by Schu et al.[47] The events were reported for the periods –6, –12, and –24 months to the date of questionnaire. Three events necessitated hospitalization and five resulted in injuries, e.g., bone fractures. Amongst co-morbidities psychologic disorders were reported by four participants (e.g., depression, post-traumatic stress disorders).

Fall(s) were more prevalent in women, with 14 of 57 (24.5%) women experiencing falls compared to 8 of 95 (8.4%) men (difference in proportions = 0.31, 95% CI: 0.06–0.55; p = 0.01). Participants with a limited walking distance (<1000 m) had a higher fall rate (26.9%, 7 out of 26) compared to those able to walk more than 1000 m (11.9%, 15 of 126; p = 0.02). Unexpectedly, in patients with diabetes fall(s) were more common in those with a recent diagnosis (11.3 ± 7.3 years) compared to patients without falls (18.6 ± 14.0 years; p = 0.04). However the time of diabetes onset is often obscure in patients diagnosed with diabetes type 2 given the slow onset of disease. Overall, 21 out of 22 falls occurred in participants aged 50 and above, and there was no significant difference in the age distribution between those who experienced falls and those who did not (67.5 ± 9.3 vs. 64.6 ± 10.7; p = 0.23, Table 1).

The applied clinical tests performed in 113 participants were correlated with actual fall numbers (overall 14 events). In the TUG test, three out of 14 (21.4%) participants who experienced falls had a pathological result compared to four out of 99 (4.0%) patients without falls. This showed a trend toward significance for patients at risk with a timed "up and go" of 11 s (difference in proportions = 0.17, 95% CI: –0.09–0.43; p = 0.05). The other clinical tests (DGI with cut-off of 10 points; BBS with cut-off of below 45 points) failed to correctly classify patients at high risk of fall(s) (Table 1).

In addition, the device functioned reasonably well throughout the study, and all participants were able to complete the game tasks at the set difficulty level. There were short term interruptions of Bluetooth connectivity in eight out of 254 game sessions (3.1%, five seated and three standing). In all case the game session resumed after reconstitution of connection. All data were recorded and analyzed.

Hypothesis-driven analyses of gaming data

In our study, we hypothesized that a specific set of motor and cognitive skills are crucial for effective locomotion and fall prevention, and that these skills can be systematically evaluated using gaming data (Fig. 2). These hypotheses build upon previous studies[36,37] that identified essential components of postural control, cognitive processing, and peripheral nerve function. Based on these foundational hypotheses, we classified six key capabilities that capture essential clinical aspects of movement control: balance, and strength, factors that are known to play a vital role in fall prevention, particularly in at-risk populations. Our hypothesis-driven approach integrates insights from previous research and expert knowledge. These six key capabilities have been categorized and defined as follows. Reaction time: speed and accuracy of task comprehension and response, which is particularly important for avoiding falls in sudden or unexpected situations. Sensation: precision in pressure application during gaming tasks, mirroring an individual's ability to fine-tune motor control, critical for both stability and mobility. Skillfulness: overall performance and achievements across different gaming tasks, serving as an aggregate indicator of coordination, precision, and adaptability in motor tasks. Muscle strength: ability to adjust to varying levels of plantar pressure, particularly relevant for generating and maintaining sufficient force to stabilize the body. Balance: coordination of plantar pressure distribution related to postural control, which is critical for maintaining stability and preventing falls. Endurance: capacity to sustain motor output over time, representing the ability to maintain consistent performance throughout extended physical tasks, which is the basis for overall locomotion safety.

The analysis of data sets from the seated gaming session (n = 152) identified significant between-group disparities in several key functional variables, e.g., endurance (0.6 ± 0.1 vs. 0.5 ± 0.2; p = 0.03), sensation (0.6 ± 0.1 vs. 0.5 ± 0.1; p = 0.06) and muscle strength (0.8 ± 0.1 vs. 0.7 ± 0.1; p = 0.06).

In the standing position the performance quality of the participants (n = 102) decreased markedly for all key capabilities, especially endurance, sensation and muscle strength. Overall, "balance" remained as the main discriminator of participants "with" vs. "without" fall(s) (0.7 ± 0.1 vs. 0.6 ± 0.2; p = 0.05). Fig. 2 illustrates the detailed distribution of these game scores between groups "with" and "without" fall history.

Group	overall (n = 152)	w/o falls (n = 130)	with falls (n = 22)	p
Sex (female)	57 (37.5%)	43 (33.1%)	14 (63.6%)	**0.01**
Age (years)	65.0 (10.5)	64.6 (10.7)	67.5 (9.3)	0.23
>65	85 (55.9%)	70 (53.8%)	15 (68.2%)	
50–65	58 (38.2%)	52 (40.0%)	6 (27.3%)	0.48
<50	9 (5.9%)	8 (6.2%)	1 (4.5%)	
BMI (kg/m^2)	30.2 (6.1)	29.8 (5.8)	32.3 (7.3)	0.13
Diabetes type				0.59
Type 1	38 (25.0%)	34 (26.2%)	4 (18.2%)	
Type 2	114 (75.0%)	96 (73.8%)	18 (81.8%)	
Diabetes duration (years)	17.6 (13.5)	18.6 (14.0)	11.3 (7.3)	0.04
Walking assistance				0.17
Rollator	4 (2.6%)	3 (2.3%)	1 (4.5%)	
Cane	5 (3.3%)	3 (2.3%)	2 (9.1%)	
Walking distance (m)				**0.02**
<500	13 (8.6%)	9 (6.9%)	4 (18.2%)	
Until 1000	13 (8.6%)	10 (7.7%)	3 (13.6%)	
Usage of handrail while climbing stairs	75 (49.3%)	61 (46.9%)	14 (63.6%)	0.22
Balance problems in darkness	18 (11.8%)	15 (11.5%)	3 (13.6%)	0.99
Dizziness during head nodding or turning	21 (13.8%)	16 (12.3%)	5 (22.7%)	0.33
Dizziness during body posture changes	39 (25.7%)	33 (25.4%)	6 (27.3%)	0.99
NSS				0.29
Moderate (5–6)	33 (21.7%)	28 (21.5%)	5 (22.7%)	
Severe (7–10)	60 (39.5%)	48 (36.9%)	12 (54.5%)	
NDS				0.51
Moderate (6–8)	55 (36.2%)	47 (36.2%)	8 (36.4%)	
Severe (9–10)	15 (9.9%)	14 (10.8%)	1 (4.5%)	
MoCA (≤25 points)	58 (38.2%)	48 (36.9%)	10 (45.5%)	0.60
Clinical assessments	n = 113	n = 99	n = 14	
Romberg's test (falling)	4 (3.5%)	3 (3.0%)	1 (7.1%)	0.42
Unterberger's stepping test (left/right)	35 (30.9%)	31 (31.4%)	4 (28.6%)	0.98
TUG (≥11 s)	7 (6.2%)	4 (4.0%)	3 (21.4%)	0.05
DGI (<10 points)	16 (14.2%)	12 (12.1%)	4 (28.6%)	0.21
BBS (≤45 points)	5 (4.4%)	5 (5.1%)	0	0.99

Categorical variables are presented as n (%), and continuous variables are described as mean [SD]. Group comparisons were performed using Chi-square tests or Fisher's exact test for categorical variables, and t-tests or Mann–Whitney U tests for continuous variables according to their distribution. Bold values are marked as significant. BBS, Berg-Balance-Scale; BMI, body-mass index; DGI, Dynamic Gait Index; MoCA, Montreal Cognitive Assessment; NDS, Neuropathy Disability Score; NSS, Neuropathy Symptom Score; TUG, timed "up and go" test.

Table 1: **Demographic and clinical profiles of study participants.**

Game-based risk factors for fall(s)

Following hypothesis-driven analyses data entry into features extraction programs was performed to yield game-based features (overall 5244 distinct features per data set) and to correlate these with a history of fall(s).[36,37] After preprocessing and inter-correlation analyses to eliminate features with high multicollinearity, 1181 and 1232 independent game features were identified from the seated and standing sessions, respectively. Logistic regression analyses were performed on these features to evaluate their predictive power with adjustment for age, sex, BMI, diabetes type, and duration of diabetes. Overall, 65 (seated) and 97 (standing) game features were identified as independent predictors for history of fall(s). Supplementary Figure S4 illustrates representative game features along with their odds ratios for history of fall(s). Apple-Catch and Cross-Pressure games contributed more risk factors compared to BF and IJ games.

Machine learning models for fall risk prediction

To identify the most informative predictors for history of fall(s), feature ranking methods were applied to both, the hypothesis-driven key capabilities as well as extracted game features. Subsequently, the top six ranked variables were selected from both approaches and a multiple binary classification model was established to distinguish participants with a history of fall(s) from those without. Due to the different number of participants in the seated and standing gaming session, two series of models were developed using their respective data sets (for sessions in seated vs. standing position). The global model explanation and ROC curves for both models are illustrated in Fig. 3 (left and right).

Regarding the seated gaming session with 152 data sets the GLMNET model using six hypothesis-driven key scoring parameters (Fig. 3, model 1) yielded an area under the receiver-operating-characteristic curve (AUC-ROC) of 0.78 (95% confidence interval (CI): 0.69–0.87; adjusted accuracy of 74.4%) in classifying participants "with" and "without" a history of fall(s). In comparison, subsequent models that employed features

ID	Events (n)	Fall type	Event within (months)			Hospitalization\1	Bone fracture, injuries	Psychologic disorder
			−6 to 0	−12 to 0	−24 to 0			
P477	10	Stumbling			X			Depression
P493	5	Stumbling			X			
P270	3	Stumbling			X			
P544	2	Stumbling			X			
P283	1	Stumbling			X			
P304	1	Stumbling			X			
P313	1	Stumbling			X			
P102	1	Stumbling			X			PTSD
P449	1	Stumbling		X		X	X	
P451	1	Stumbling		X				Depression
P492	1	Stumbling			X		X	
P534	1	Stumbling	X			X	X	
P106	1	Stumbling		X				
P231	1	Stumbling		X				
P238	1	Stumbling	X					
P257	1	Stumbling	X					
P606	1	Slipping	X					Depression
P317	1	Slipping	X			X	X	
P551	1	Slipping	X					
P407	1	Fainting	X				X	
P522	1	Fainting	X					
P505	1	Unknown		X				

The fall type is determined according to the criteria introduced by Shu et al. (2021): stumbling is loss of balance caused by stepping onto an unperceived object, resulting in a shift of the center of gravity and potential fall. Slipping occurs when the frictional force opposing foot movement is insufficient to counteract the horizontal shear force, causing the legs to slide and resulting in a fall, commonly in environments with smooth or wet surfaces. Fainting or syncope is a sudden and temporary loss of consciousness due to impaired cerebral perfusion and transient brain hypoxia, leading to a collapse. PTSD, post-traumatic stress disorders.

***Table 2:* Summary on fall events.**

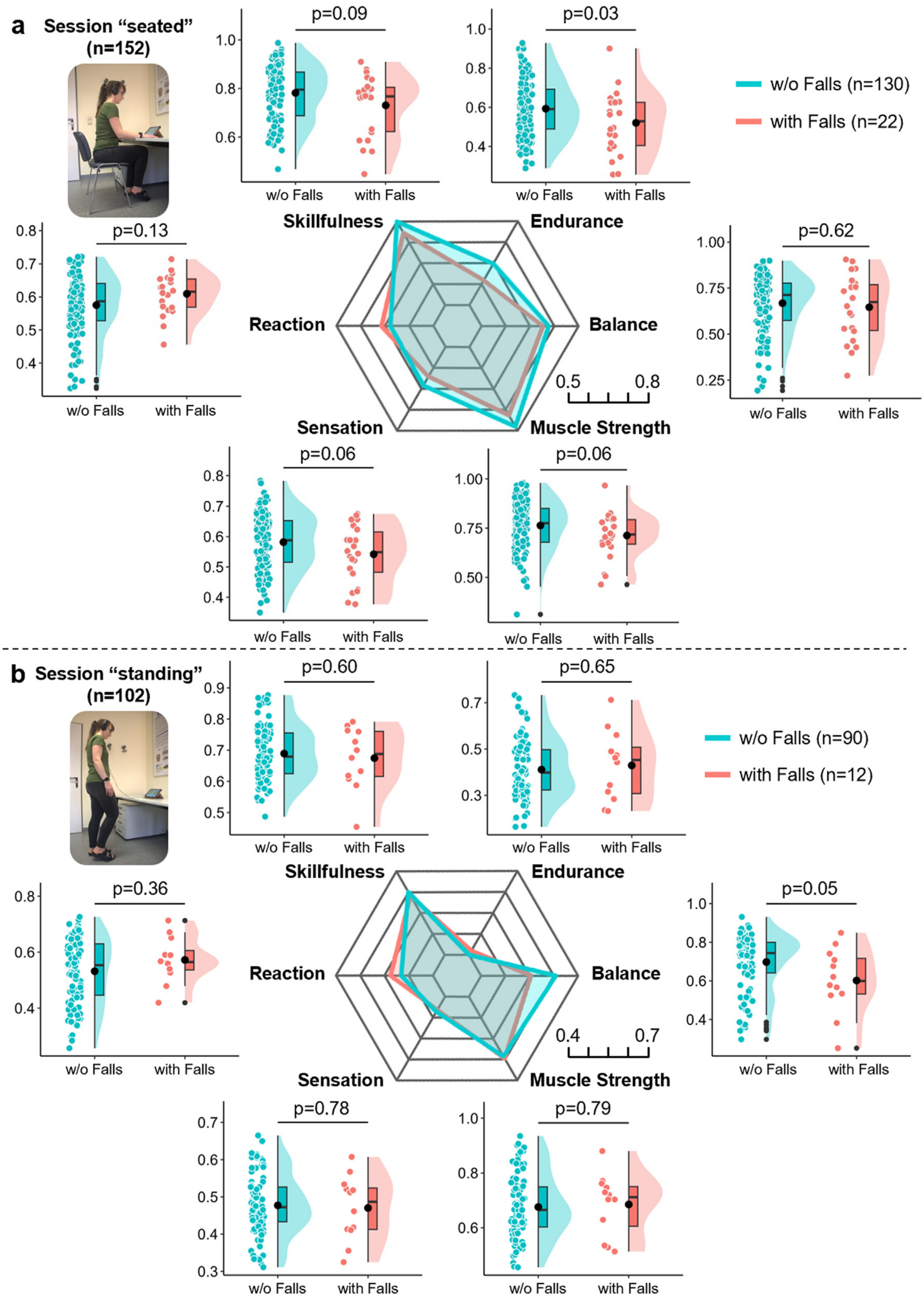
extracted from various games (Fig. 3, models 2–6) demonstrated superior predictive performance, with AUC-ROC values of 0.84 (95% CI: 0.77–0.91) and an adjusted accuracy of 82.8% (Fig. 3, model 6).

The GLMNET model using six hypothesis-driven key scoring parameters from the standing gaming session with 102 data sets achieved an AUC-ROC of 0.84 (95% CI: 0.75–0.93; adjusted accuracy of 81.4%) in classifying participants "with" and "without" a history of fall(s) (Fig. 3, model 7). Five models incorporating specific features extracted from different games (Fig. 3, models 8–12) further improved the discrimination, the maximal AUC-ROC was calculated at 0.91 (95% CI: 0.85–0.97; adjusted accuracy of 88.6%; Fig. 3, model 12).

Overall, models trained with specific game features by machine learning outperformed those using hypothesis-driven key capability parameters. The gaming sessions in "standing" position delivered superior predictive values on history of fall(s) compared to the "seated" sessions (Fig. 3, models 7–12 vs. 1–6). The final game features of model 6 and 12 are visualized using rain cloud plots in Supplementary Figure S5. Principal component analysis of the final feature space for these two models are presented in Supplementary Figure S6. Besides, the machine learning models significantly outperformed clinical assessments of fall(s) (i.e., TUG test, DGI, BBS, see Supplementary Table S3).

Discussion

The primary objective of this study was to evaluate the feasibility and efficacy of a video game-based platform for classifying participants with (n = 22) and without (n = 130) a history of fall(s). The occurrence rate of fall events in the 24 months prior to the gaming session was 14.5%. In comparison, Cheng et al. reported a fall incidence of 9.79% in a six-month follow-up study involving 950 elderly patients with diabetes.[17] Cross-sectional studies conducted in Singapore and the United States have documented higher fall rates of 17.53%[7] and 37.5%,[19] respectively, which may be attributed to differences age of study populations and comorbidities, such as diabetes, retinopathy, vasculopathy and co-medications. Regarding the clinical outcomes of fall(s) in our cohort, five out of 22 events (22.7%) were injurious, with three cases necessitating hospitalization. Consistent with findings reported by Cheng et al.[17] and Rashida et al.,[48] our data indicate correlations between fall(s) and factors such as sex and maximum walking distance without assistance. However, age, balance impairment (assessed by TUG, DGI, or BBS), cognitive status (evaluated by MoCA test),


a Session "seated" (n=152)
w/o Falls (n=130)
with Falls (n=22)
p=0.09
p=0.03
p=0.13
p=0.62
p=0.06
p=0.06
Skillfulness
Endurance
Reaction
Balance
Sensation
Muscle Strength
0.5 0.8
w/o Falls
with Falls
b Session "standing" (n=102)
w/o Falls (n=90)
with Falls (n=12)
p=0.60
p=0.65
p=0.36
p=0.05
p=0.78
p=0.79
0.4 0.7

and peripheral neuropathy (assessed by NSS and NDS) do not correlate with history of fall(s) in our study.

In addition, both hypothesis-driven key capabilities and game features extracted from the gaming data set significantly correlate with history of fall(s). These features provide a nuanced perspective on fall risk in contrast to the demographic and clinical risk factors reported by Cheng et al.[17] and Rashida et al.[48] They mirror the real-world physical capabilities of the game player, such as visual observation, decision-making, reaction time, anticipation, task execution, balance, endurance, and motor coordination during game engagement.[49] In addition, the game session executed in a "standing position" has even higher predictive power on the history of fall(s) in comparison to the "seated" session.

The obtained classification models using game features achieve AUC-ROC scores of 0.84 and 0.91 for "seated" and "standing" gaming sessions, respectively. These models outperform those reported by Cheng et al. (AUC-ROC 0.69)[17] and Rashida et al. (AUC-ROC 0.80),[48] underscoring the effectiveness of using comprehensive game-based data and machine learning algorithms for fall risk assessment in an elderly cohort. Our approach offers enhanced sensitivity and specificity, surpassing traditional clinical assessments like TUG, DGI, and BBS.

Integrating machine learning algorithms with data from gamification sessions or other gadgets that enable sensor recordings is revolutionizing health assessments and will likely enhance personalized care.[50] This approach has the potential to instigate behavioral changes, ensuring that individuals are not only informed about their health status but are also incentivized to maintain or improve it. Examples of these advancements include algorithms that decipher complex patterns within large data sets, such as assessments for imbalance and vestibular dysfunction, treatments for phantom limb pain, and evaluations of cognitive function.[51–55] Leveraging these machine learning techniques will transform data into actionable health strategies, significantly improving monitoring accuracy and leading to tailored treatment plans.[56,57]

The innovative character of our approach has several important basic concepts. First, the games are designed in an easy to understand manner. Second, the tutorial on the performance of the games is standardized and allows for short sequences where the patients may test responsiveness of the sensors and steering units. Third, the setup of patient positioning in front of the tablet and on a chair without armrest is kept constant. Fourth, the insoles are covered with protective sponge material and integrated into standardized slippers. The core material of the insole consists of supportive ethylene-vinyl acetate (EVA, 50° shore). The latter provides a sufficient counter-resistance for pressure application. Fifth, external support during the testing is reduced to a minimum by supplying the patients with headphones that minimize communication. By strictly adhering to this concept comparability of data sets is meaningful possible. Most technical challenges were overcome by the established system, except that at times the sensor positioning was suboptimal. This may be solved by integration of more than 8 sensors per insole. Although numerous external advisors objected and pointed out that patients with advanced age are not capable of playing video games our experience in the study cohort proves the opposite. None of the enrolled patients failed in performing the games at the current difficulty level.

One may speculate why the clinical assessment scores do to meet the precision of the gaming data sets. The latter obtain more information on the execution and coordination of plantar pressure changes. It is our belief that other video games that are designed to collect sensor data with movements (of the lower and upper limbs) will have to solve the issue of standardization, given the complex nature of movement.[58–60]

In comparison with prospective studies that utilize future event data as labels, as suggested by Shany et al.,[61] the retrospective design employed in this study introduces several potential limitations. Retrospective data, often based on self-reported falls, is subject to recall bias, and post-fall interventions, such as physical therapy, may alter participants' risk profiles, potentially confounding the analysis. Additionally, the retrospective nature limits our ability to account for changes in fall risk over time, such as deteriorations in balance or strength.

We wish to emphasize that the achieved algorithms are clearly distinct from ones for future events. These may be even more informative and helpful to prevent events (e.g., carpet removal from floors, barriers and steps in apartments, adjustments of medications that have a sedative effect, such as pain killers, antidepressants). While models predicting future falls may offer preventative opportunities, our analysis on past falls allows for the identification of enduring risk factors, such as neuromuscular deficits, impaired postural stability, and motor coordination dysfunctions. These persistent risk factors may provide critical insights to guide targeted interventions in the prevention of future falls. Furthermore, our approach enables an immediate

***Fig. 2:* Hypothesis-driven key capabilities and scoring of video games.** Six game performance parameters are visualized with two spider diagrams to indicate the player's skillfulness, endurance, balance, sensation, reaction, and muscle strength for both "seated" (a) and "standing" sessions (b). A rain cloud plot presents the distribution difference of each parameter between groups "with" and "without" falls. All results are provided as arbituary units.

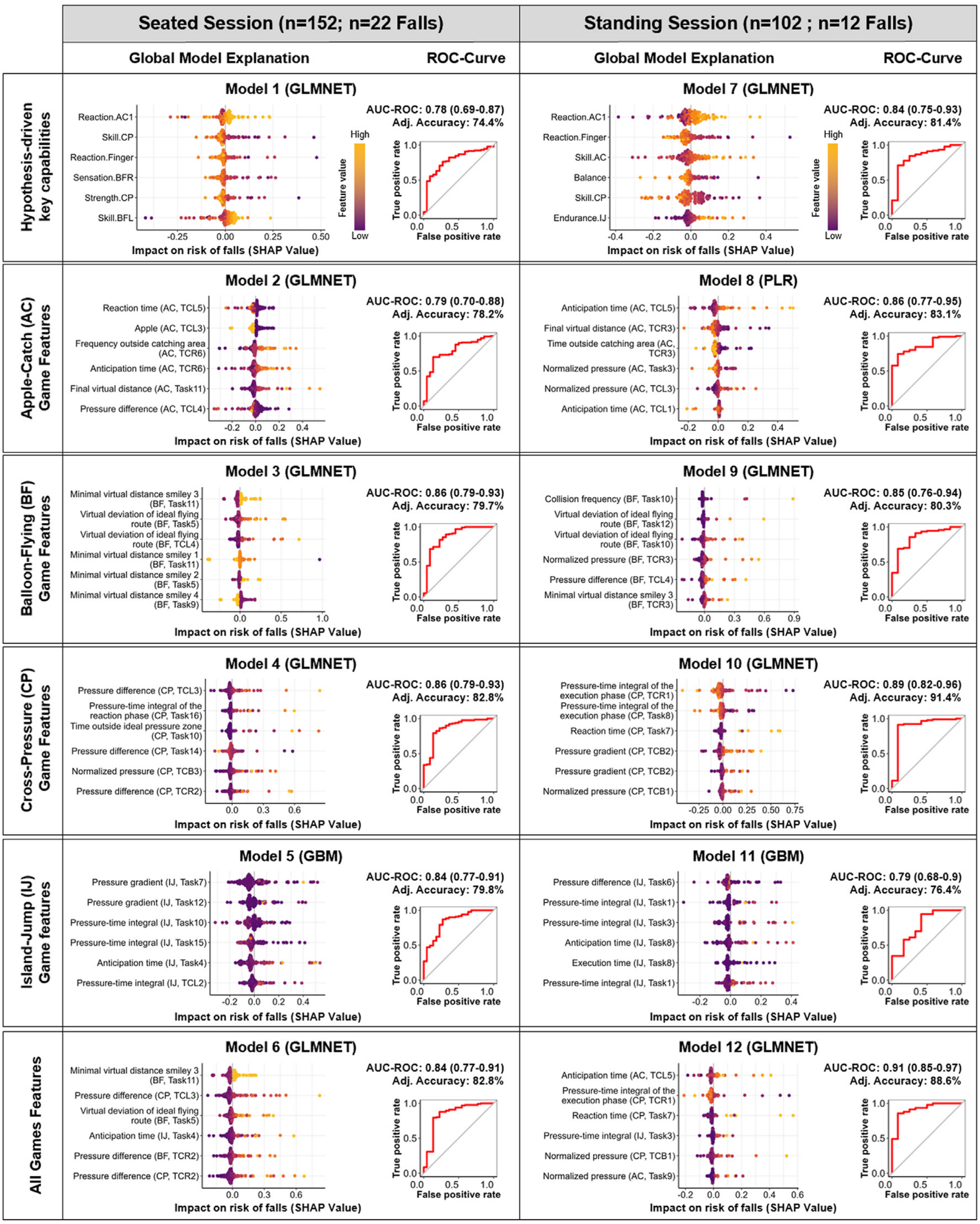


Fig. 3: **Comprehensive explanation and comparative analysis of machine learning models for fall risk assessment.** Models 1–6 were trained using 152 data sets obtained from the "seated" gaming session, while 102 data sets from the "standing" gaming session were utilized for models 7–12. The sources of model features are listed vertically on the left side of the figure. Final features for each model (n = 6) are sorted by their ranked importance, as determined by the "varImp" function of the R Caret library. A SHAP summary dot plot ranks the features for each model. In these plots, the probability of falls increases with the SHAP value of a feature. Each dot represents a SHAP value for an individual participant, with one dot per feature per participant. The dot colors indicate the actual feature values for each participant, where bright yellow signifies higher feature values and dark purple indicates lower feature values. The dots are stacked vertically to illustrate density. SHAP values above zero for specific features influence the decision towards the "with Falls" class. An ROC curve illustrates the predictive performance of each

assessment of the video game-based features' effectiveness in predicting fall risk, without waiting for future events. This makes the method particularly useful in clinical settings, where prompt feedback is essential. In addition, the use of past event data lays a solid foundation for continuous model refinement. When novel data on future falls will become available, models will be iteratively validated and optimized for accuracy and robustness.

However, we acknowledge that our algorithm is more suited to identifying individuals with a history of falls, rather than predicting future falls directly. A combination of retrospective fall data with balance and functional assessments, such as TUG, DGI, and BBS, may provide a more holistic view of current fall risk and future predictions.

Further limitation of this study is the lower-than-expected fall rate, with only 22 out of 152 participants (14.5%) reporting falls over the past two years, compared to the 25–30% fall rate in elderly populations reported by other groups.[19] This discrepancy may be due to underreporting given the retrospective self-reporting of the patients which is prone to recall bias. On the other hand, real event differences may also exist due to better health management amongst the population cohort with diabetes. Additionally, comparing fall rates over 24 months with studies using different time frames (e.g., 9.79% over 6 months) complicates interpretation. Thus, the observed fall rate should be contextualized with caution, recognizing the potential for recall bias and differences in follow-up periods. Moreover, the low number of falls presents a challenge for the machine learning models, as accurate predictions rely on well-labeled outcome data. Underreporting of falls may have reduced the predictive power of the models and restrict the generalizability of the results or applicability in clinical assessment centers. To address this, future studies will focus on larger and more diverse cohorts to validate the findings and enhance their applicability. Furthermore, two falls due to fainting were included as they represent an important subtype of falls in older adults. While differences from accidental falls exist, we opted for the inclusion of both cases as in the patients recalled the events of falls and fainted thereafter. In this regard the fainting was not related to an Adam Stokes attack but rather constitute a dysregulation of blood pressure levels.

Regarding the hypothesis-driven analyses of the game data from the standing position, participants demonstrated a marked decline in performance across all key capabilities, with "balance" remaining the primary distinguishing factor between those with and without a history of falls (0.7 ± 0.1 vs. 0.6 ± 0.2; p = 0.05). A possible explanation is that the standing game session placed an unusually high demand on postural balance, as participants were required to maintain their posture while completing various game challenges. This may have heightened the differences in balance performance, overshadowing or diminishing the sensitivity of other key capabilities.

Our findings indicate that patients who experienced falls tend to have a shorter duration of diabetes. This contradicts the clinical intuition. However, one has to keep in mind that patients with longer diabetes duration may use mobility aids as a protective factor or have adopted compensatory strategies. Nonetheless a validation in larger cohorts is needed. Moreover, future studies should incorporate multiple sessions across both seated and standing conditions to better track learning and performance progression over time, offering a more comprehensive evaluation of game data.

The model's performance, derived from cross-validation, needs confirmation through an independent testing data set. Variability in neuropathic manifestations and patient effort impact results. Recurrent game interactions might induce a training effect, potentially biasing conclusions. Technological adeptness variations could influence game performance. The fixed sequence of gaming tasks warrants the exploration of randomized sequences in future research. As the system remains a prototype and the sensor-equipped insoles are not commercially available, reproducibility by other researchers is currently limited. It is our next goal to deploy the game platform in an outpatient context to assess its broader applicability. In addition, future work will explore data integration and transformation methods that could enable more effective application of data augmentation techniques (e.g., Synthetic Minority Over-sampling Technique) to better address class imbalance.

This study demonstrates the potential of a video game-based platform and machine learning to assess performance in relation to fall history in a cohort of adults diagnosed with diabetes. The AI-derived algorithms showed promise in identifying patients with a history of falls. However, there are limitations, such as retrospective data collection that are prone to underreporting and may explain relative low fall rates. Future studies should validate the findings in larger cohorts, putatively integrating this approach into clinical practice.

model. The adjusted accuracy is calculated based on the optimal predicted probability cutoff selected by Youden's Index. SHAP, SHapley additive explanation; AUC-ROC, area under the curve of receiver operating characteristic; GLMNET, lasso and elastic-net regularized generalized linear models; GBM, gradient boosting machine; PLR, penalized logistic regression; AC, Apple-Catch game; BF, Balloon-Flying game; CP, Cross-Pressure game; IJ, Island-Jump game; TCL, task combination for the left foot; TCR, task combination for the right foot.

Contributors

AM and PRM conceptualized and designed the study. AM managed app development. AM, TS, JH, SS, and PRM contributed to the acquisition, analysis, or interpretation of data. AM and SS did the statistical analysis. AM and PRM verified the underlying data and drafted the manuscript. AM, VM, SS, and PRM critically revised the manuscript for important intellectual content. PRM wrote the grant application, was the initiator of the game platform, supervised the study and data analyses. All authors read and approved the final version of the manuscript.

Data sharing statement

We are open to any reasonable requests for the original data directed through the corresponding author, providing that the requested data relates to that published in the article and does not compromise any future publication or any other related issues.

Declaration of interests

AM and PRM filed patents related to the assessment of fall risk in elderly patients (202415EP) and the evaluation of peripheral neuropathy (202205PCT). All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

We would like to thank C. Moeller-Wohl, A. Petrow, and I. Siddiquee for their support in the performance of the study, such as software development, recruitment process, and test organization. We thank T. Meyer (IEE S.A., Luxembourg) for fruitful interactions in the establishment of sensor-equipped insoles as well as V. Herbst for advice on insole design.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2024.102947.

References

1 Falls. https://www.who.int/news-room/fact-sheets/detail/falls. Accessed June 2, 2024.
2 CDC. *About older adult fall prevention. Older adult fall prevention*; 2024. https://www.cdc.gov/falls/about/index.html. Accessed June 3, 2024.
3 Bergen G, Stevens MR, Burns ER. Falls and fall injuries among adults aged ≥65 years - United States, 2014. *MMWR Morb Mortal Wkly Rep*. 2016;65:993–998.
4 Immonen M, Haapea M, Similä H, et al. Association between chronic diseases and falls among a sample of older people in Finland. *BMC Geriatr*. 2020;20:225.
5 Lee AK, Juraschek SP, Windham BG, et al. Severe hypoglycemia and risk of falls in type 2 diabetes: the Atherosclerosis Risk in Communities (ARIC) Study. *Diabetes Care*. 2020;43:2060–2065.
6 Timar B, Timar R, Gaiţă L, Oancea C, Levai C, Lungeanu D. The impact of diabetic neuropathy on balance and on the risk of falls in patients with type 2 diabetes mellitus: a cross-sectional study. *PLoS One*. 2016;11:e0154654.
7 Gupta P, Aravindhan A, Gand ATL, et al. Association between the severity of diabetic retinopathy and falls in an Asian population with diabetes: the Singapore epidemiology of eye diseases study. *JAMA Ophthalmol*. 2017;135:1410–1416.
8 Yau RK, Strotmeyer ES, Resnick HE, et al. Diabetes and risk of hospitalized fall injury among older adults. *Diabetes Care*. 2013;36:3985–3991.
9 Lapane KL, Jesdale BM, Dubé CE, Pimentel CB, Rajpathak SN. Sulfonylureas and risk of falls and fractures among nursing home residents with type 2 diabetes mellitus. *Diabetes Res Clin Pract*. 2015;109:411–419.
10 Riandini T, Khoo EYH, Tai BC, et al. Fall risk and balance confidence in patients with diabetic peripheral neuropathy: an observational study. *Front Endocrinol*. 2020;11:573804.
11 Rashedi V, Iranpour A, Mohseni M, Borhaninejad V. Risk factors for fall in elderly with diabetes mellitus type 2. *Diabetes Metab Syndr*. 2019;13:2347–2351.
12 Hendrich AL, Bender PS, Nyhuis A. Validation of the Hendrich II fall risk model: a large concurrent case/control study of hospitalized patients. *Appl Nurs Res*. 2003;16:9–21.
13 Parker SG, McCue P, Phelps K, et al. What is Comprehensive Geriatric Assessment (CGA)? An umbrella review. *Age Ageing*. 2018;47:149–155.
14 Podsiadlo D, Richardson S. The timed 'Up & Go': a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc*. 1991;39:142–148.
15 Mañago MM, Cameron M, Schenkman M. Association of the Dynamic Gait Index to fall history and muscle function in people with multiple sclerosis. *Disabil Rehabil*. 2020;42:3707–3712.
16 Berg K, Wood-Dauphine S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Can*. 1989;41:304–311.
17 Cheng Z, Li X, Xu H, Bao D, Mu C, Xing Q. Incidence of accidental falls and development of a fall risk prediction model among elderly patients with diabetes mellitus: a prospective cohort study. *J Clin Nurs*. 2023;32:1398–1409.
18 Brown SJ, Handsaker JC, Bowling FL, Boulton AJM, Reeves ND. Diabetic peripheral neuropathy compromises balance during daily activities. *Diabetes Care*. 2015;38:1116–1122.
19 Blackwood J. Cognitive function and falls in older adults with type 2 diabetes mellitus. *J Geriatr Phys Ther*. 2019;42:E91–E96.
20 Tunca C, Salur G, Ersoy C. Deep learning for fall risk assessment with inertial sensors: utilizing domain knowledge in Spatio-temporal gait parameters. *IEEE J Biomed Health Inform*. 2020;24:1994–2005.
21 Chen M, Wang H, Yu L, et al. A systematic review of wearable sensor-based technologies for fall risk assessment in older adults. *Sensors*. 2022;22:6752.
22 van der Kolk NM, de Vries NM, Kessels RPC, et al. Effectiveness of home-based and remotely supervised aerobic exercise in Parkinson's disease: a double-blind, randomised controlled trial. *Lancet Neurol*. 2019;18:998–1008.
23 Wang X, Cao J, Zhao Q, et al. Identifying sensors-based parameters associated with fall risk in community-dwelling older adults: an investigation and interpretation of discriminatory parameters. *BMC Geriatr*. 2024;24:125.
24 Bezold J, Krell-Roesch J, Eckert T, Jekauc D, Woll A. Sensor-based fall risk assessment in older adults with or without cognitive impairment: a systematic review. *Eur Rev Aging Phys Act*. 2021;18:15.
25 Johnson D, Deterding S, Kuhn K-A, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. *Internet Interv*. 2016;6:89–106.
26 Burgon T, Casebeer L, Aasen H, et al. Measuring and improving evidence-based patient care using a web-based gamified approach in primary care (QualityIQ): randomized controlled trial. *J Med Internet Res*. 2021;23:e31042.
27 Ming A, Walter I, Alhajjar A, Leuckert M, Mertens PR. Study protocol for a randomized controlled trial to test for preventive effects of diabetic foot ulceration by telemedicine that includes sensor-equipped insoles combined with photo documentation. *Trials*. 2019;20:521.
28 Ming A, Alhajjar A, Walter I, et al. Telemedical monitoring of plantar temperature in diabetic patients at risk of foot ulcers. *Dtsch Arztebl Int*. 2024;121:9–16.
29 Ziegler D, Keller J, Maier C, Pannek J, German Diabetes Association. Diabetic neuropathy. *Exp Clin Endocrinol Diabetes*. 2014;122:406–415.
30 Weintrob N, Amitay I, Lilos P, Shalitin S, Lazar L, Josefsberg Z. Bedside neuropathy disability score compared to quantitative sensory testing for measurement of diabetic neuropathy in children, adolescents, and young adults with type 1 diabetes. *J Diabetes Complications*. 2007;21:13–19.
31 Dyck PJ. Detection, characterization, and staging of polyneuropathy: assessed in diabetics. *Muscle Nerve*. 1988;11:21–32.
32 Tsilingiris D, Schimpfle L, von Rauchhaupt E, et al. Dysmetabolism-related early sensory deficits and their relationship with peripheral neuropathy development. *J Clin Endocrinol Metab*. 2023;108:e979–e988.
33 Pfannkuche A, Alhajjar A, Ming A, Walter I, Piehler C, Mertens PR. Prevalence and risk factors of diabetic peripheral neuropathy in a diabetics cohort: register initiative "diabetes and nerves". *Endocr Metab Sci*. 2020;1:100053.
34 Trzepacz PT, Hochstetler H, Wang S, Walker B, Saykin AJ. Relationship between the Montreal cognitive assessment and mini-mental state examination for assessment of mild cognitive impairment in older adults. *BMC Geriatr*. 2015;15:107.

35 Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53:695–699.
36 Ming A, Lorek E, Wall J, et al. Unveiling peripheral neuropathy and cognitive dysfunction in diabetes: an observational and proof-of-concept study with video games and sensor-equipped insoles. *Front Endocrinol*. 2024;15. https://www.frontiersin.org/journals/endocrinology/articles/10.3389/fendo.2024.1310152. Accessed March 5, 2024.
37 Ming A, Clemens V, Lorek E, et al. Game-based assessment of peripheral neuropathy combining sensor-equipped insoles, video games, and AI: proof-of-concept study. *J Med Internet Res*. 2024;26:e52323.
38 Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6:65–70.
39 Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33:1–22.
40 Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin Epigenetics*. 2019;11:123.
41 Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurorobot*. 2013;7:21.
42 Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155–163.
43 Kuhn M. Building predictive models in R using the caret package. *J Stat Softw*. 2008;28:1–26.
44 Pudjihartono N, Fadason T, Kempa-Liehr AW, O'Sullivan JM. A review of feature selection methods for machine learning-based disease risk prediction. *Front Bioinform*. 2022;2. https://doi.org/10.3389/fbinf.2022.927312.
45 Bacanin N, Zivkovic M, Antonijevic M, et al. Addressing feature selection and extreme learning machine tuning by diversity-oriented social network search: an application for phishing websites detection. *Complex Intell Syst*. 2023;9:7269–7304.
46 Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3:32–35.
47 Shu F, Shu J. An eight-camera fall detection system using human fall pattern recognition via machine learning by a low-cost android box. *Sci Rep*. 2021;11:2471.
48 Bibi R, Yan Z, Ilyas M, Shaheen M, Singh SN, Zeb A. Assessment of fall-associated risk factors in the Muslim community-dwelling older adults of Peshawar, Khyber Pakhtunkhwa, Pakistan. *BMC Geriatr*. 2023;23:623.
49 Schoene D, Lord SR, Verhoef P, Smith ST. A novel video game–based device for measuring stepping performance and fall risk in older people. *Arch Phys Med Rehabil*. 2011;92:947–953.
50 Seshadri DR, Thom ML, Harlow ER, et al. Wearable Technology and analytics as a complementary toolkit to optimize workload and to reduce injury burden. *Front Sports Act Living*. 2021;2. https://www.frontiersin.org/articles/10.3389/fspor.2020.630576. Accessed February 27, 2024.
51 Mekruksavanich S, Jitpattanakul A. Hybrid convolution neural network with channel attention mechanism for sensor-based human activity recognition. *Sci Rep*. 2023;13:12067.
52 Yeh S-C, Huang M-C, Wang P-C, et al. Machine learning-based assessment tool for imbalance and vestibular dysfunction with virtual reality rehabilitation system. *Comput Methods Programs Biomed*. 2014;116:311–318.
53 Ortiz-Catalan M, Guðmundsdóttir RA, Kristoffersen MB, et al. Phantom motor execution facilitated by machine learning and augmented reality as treatment for phantom limb pain: a single group, clinical trial in patients with chronic intractable phantom limb pain. *Lancet*. 2016;388:2885–2894.
54 Jung H-T, Daneault J-F, Lee H, et al. Remote assessment of cognitive impairment level based on serious mobile game performance: an initial proof of concept. *IEEE J Biomed Health Inform*. 2019;23:1269–1277.
55 Nicolaidou I, Antoniades A, Constantinou R, et al. A virtual emergency telemedicine serious game in medical training: a quantitative, professional feedback-informed evaluation study. *J Med Internet Res*. 2015;17:e150.
56 Sarker IH. Machine learning: algorithms, real-world applications and research directions. *SN Comput Sci*. 2021;2:160.
57 Sabry F, Eltaras T, Labda W, Alzoubi K, Malluhi Q. Machine learning for healthcare wearable devices: the big picture. *J Healthc Eng*. 2022;2022:4653923.
58 Clark RA, Bryant AL, Pua Y, McCrory P, Bennell K, Hunt M. Validity and reliability of the nintendo wii balance board for assessment of standing balance. *Gait Posture*. 2010;31:307–310.
59 Yong Joo L, Soon Yin T, Xu D, et al. A feasibility study using interactive commercial off-the-shelf computer gaming in upper limb rehabilitation in patients after stroke. *J Rehabil Med*. 2010;42:437–441.
60 Barry G, Galna B, Rochester L. The role of exergaming in Parkinson's disease rehabilitation: a systematic review of the evidence. *J Neuroeng Rehabil*. 2014;11:33.
61 Shany T, Wang K, Liu Y, Lovell NH, Redmond SJ. Review: are we stumbling in our quest to find the best predictor? Over-optimism in sensor-based models for predicting falls in older adults. *Healthc Technol Lett*. 2015;2:79–88.