Comparing Sensitivity, Specificity, and Accuracy of Fall Risk Assessments in Community-Dwelling Older Adults

Kworweinski Lafontant 1,2, Amber Blount 3,4, Jethro Raphael M Suarez 1,5, David H Fukuda 6, Jeffrey R Stout 1,5, Evette M Trahan 2, Nichole R Lighthall 3, Joon-Hyuk Park 4,5, Rui Xie 2,6, Ladda Thiamwong 2,5

1Institute of Exercise Physiology and Rehabilitation Science, University of Central Florida, Orlando, FL, USA; 2College of Nursing, University of Central Florida, Orlando, FL, USA; 3Department of Psychology, University of Central Florida, Orlando, FL, USA; 4Mechanical and Aerospace Engineering Department, University of Central Florida, Orlando, FL, USA; 5Disability, Aging, and Technology Cluster, University of Central Florida, Orlando, FL, USA; 6Department of Statistics and Data Science, University of Central Florida, Orlando, FL, USA

Correspondence: Ladda Thiamwong, UCF College of Nursing, 12201 Research Parkway, Suite 300, Orlando, FL, 32826, USA, Tel +1 407 823 5091, Email ladda.thiamwong@ucf.edu

Purpose: The US Centers for Disease Control and Prevention (CDC) has implemented the Stopping Elderly Accidents, Deaths, and Injuries (STEADI) initiative. This initiative provides an algorithm for fall risk screening. However, the algorithm has the potential to overcategorize individuals as high risk for falling upon initial screening, which may burden clinicians with the task of recategorizing individuals after follow-up testing. Therefore, this study aimed to compare the accuracy, sensitivity, and specificity of fall risk appraisal between the STEADI, Short Fall-Efficacy Scale International (FES-I), and portable balance system (BTrackS) assessments in community-dwelling older adults.

Patients and Methods: This cross-sectional analysis included 122 community-dwelling older adults, comprising 94 women and 28 men. Center-of-pressure postural sway was assessed using the BTrackS, fear of falling was assessed using the Short FES-I questionnaire, and all participants completed the STEADI checklist. Each assessment categorized participants as either high or low fall risk and fall risk appraisal was compared between groups using McNemar tests.

Results: The STEADI checklist (high risk: n = 62; low risk: n = 60) significantly differed in fall risk appraisal compared to the BTrackS (high risk: n = 44; low risk: n = 78; p = 0.014) and the Short FES-I (high risk: n = 42; low risk: n = 80; p = 0.002). Compared to the BTrackS, the STEADI checklist had a specificity of 62.8%, sensitivity of 70.5%, and accuracy of 65.6%. Compared to the Short FES-I, the STEADI checklist had a specificity of 67.5%, sensitivity of 81.0%, and accuracy of 72.1%.

Conclusion: The STEADI checklist appears to overcategorize individuals as high fall risk more frequently than direct assessments of postural sway and fear of falling. Further research is needed to examine potential improvements in accuracy when combining the STEADI checklist with direct assessments of postural sway and/or fear of falling.

Plain Language Summary: Fall risk assessments are crucial for preventative care in older adults. However, the demands of clinical practice require an accurate and time-efficient method. The U.S Centers for Disease Control and Prevention (CDC) has implemented a fall risk checklist through the Stopping Elderly Accidents, Deaths, and Injuries (STEADI) initiative. However, the STEADI checklist might cost clinicians more time than expected, as some patients initially classified as high risk for falling may not actually be at high risk. This leads to unnecessary follow-up assessments. In this study, we compared the STEADI checklist to direct measures of postural sway (balance) using the BTrackS system and fear of falling using the Short FES-I survey to determine how they differed in classifying community-dwelling older adults as high versus low fall risk. Our results show that the STEADI checklist classifies older adults as high risk more frequently than the BTrackS and Short FES-I. Considering that the follow-up assessments for a high-risk classification by the STEADI checklist include a balance test, we suggest that combining a balance test such as the BTrackS with a questionnaire or checklist may yield better screening outcomes and accurately identify high-risk individuals in a timely manner. Further research is needed to determine the effectiveness of this combination and to establish a true gold standard method for fall risk appraisal.
Introduction
Falls can have serious consequences for older adults. In fact, unintentional falls are a leading cause of injury and injury-related death among older adults in America. As a result of falls, survivors often experience physical function limitations and frailty as they continue to age. Given the significant medical risk that falling poses to older adults, it is crucial to include accurate fall risk screening in clinical practice.

The United States Centers for Disease Control and Prevention (CDC) has implemented the Stopping Elderly Accidents, Deaths, and Injuries (STEADI) initiative. This initiative provides a fall risk screening algorithm for clinicians to use. The algorithm begins with a checklist designed to categorize patients as either low or high risk for falling. The checklist consists of 12 questions that assess psychological fears/concerns, as well as physical and emotional risk factors related to falls. In the STEADI algorithm, a checklist score of at least 4, or answering “yes” to at least one of three key questions, classifies the patient as at risk. This classification leads to further assessments and interventions, such as the Timed-Up-and-Go (TUG), 4-stage balance assessment, and recommendations for fall risk reduction strategies. The STEADI checklist has been shown to effectively reduce the risk of future falls by correctly identifying those at risk and providing a fall prevention plan. It has also been found to be valid and reliable among Thai older adults, even beyond US populations. The goal of any screening tool is to accurately assess risk in a timely manner. Accuracy includes both sensitivity (correctly identifying those at high risk) and specificity (correctly identifying those at low risk). However, the STEADI algorithm has the potential for reclassifying those initially identified as high risk to low risk after follow-up assessments. This suggests that the STEADI checklist may have high sensitivity but low specificity. However, there is limited literature comparing the STEADI checklist to other fall risk appraisal tools.

The Short Fall-Efficacy Scale International (Short FES-I) is a validated questionnaire used to assess the fear of falling, also known as fall efficacy. This fear can lead to physical activity avoidance and a decline in physical function even before a fall occurs. The Short FES-I is a shorter version of the Falls Efficacy Scale questionnaire, consisting of seven questions that focus on psychological concerns related to falling. Scores of 10 or higher on this questionnaire are considered to indicate an increased risk of falling. In addition to the Short FES-I, postural sway assessments can provide a direct physical evaluation of fall risk. One such assessment tool is the BTrackS Balance System (Balance Tracking Systems, San Diego, CA, USA), which is a lightweight and portable device available for commercial use. This device measures postural sway by tracking the center of pressure on a force platform. The assessment consists of four 20-second trials, and scores above 30cm are indicative of a higher risk of falling. Previous research has shown that the BTrackS has good predictive validity in assessing fall risk.

Although the STEADI checklist is favored by the CDC for its simplicity and brevity as a fall risk assessment, it is unclear whether this ease of use compromises its accuracy. This study aimed to compare the accuracy, sensitivity, and specificity of fall risk appraisal between the STEADI checklist, Short FES-I, and BTrackS assessments in community-dwelling older adults. Our hypothesis was that there would be significant variations in fall risk appraisal among the three assessments.

Materials and Methods
Participants
This study was conducted as part of a federally funded research study (NIH Grant #R03AG06799), and all methods have been previously reported. A total of 122 community-dwelling older individuals, consisting of 94 women and 28 men, participated in this cross-sectional study. The study was carried out in accordance with the Declaration of Helsinki and was pre-registered on ClinicalTrials.gov (NCT06063187). All study procedures were approved by the University of Central Florida Institutional Review Board (IRB# STUDY00002473) and all participants provided written informed consent prior to participation. The study took place in community centers located in low-income areas within the Orlando, Florida, region. Various strategies were used to recruit participants, including the distribution of fliers, inclusion...
in local newsletters, word-of-mouth, engagement with clinical sites, and collaboration with community partners. Participants received a $30 gift card as compensation upon completing the study. Participants were included in this study if they were at least 60 years old, met the low-income criteria based on the 2019 United States Census guidelines, and were living independently. Individuals who were currently receiving treatment from a rehabilitation facility were not eligible to participate.

Height and weight were assessed for each participant using a digital physician scale (Health-O-Meter, Model 402KL, McCook, IL, USA) without shoes. Body fat percentage was assessed using bioelectrical impedance analysis (BIA) with an InBody s10 device (Biospace, Seoul, South Korea) in the seated position. Participants were all assessed between 9:00 AM and 12:00 PM and were asked to maintain normal dietary habits and arrive adequately hydrated. Prior to BIA assessment, participants were instructed to fast for 3–4 hours, abstain from exercise for a minimum of 6 hours, and avoid caffeine and alcohol for 24 hours. Skin was prepared using an InBody Tissue (Biospace, Seoul, South Korea), and all jewelry was removed. Participants were also given a demographic survey in which they self-reported age and race/ethnicity. The participants in this study exhibited a diverse range of ethnic backgrounds, with 9.8% (n = 12) identifying as African American, 3.3% (n = 4) as Asian, 17.2% (n = 21) as Hispanic, and 69.7% (n = 85) as non-Hispanic White. Height, weight, body fat percentage, and demographic information are provided to characterize the sample (Table 1), as none of those variables were primary outcomes for this study.

### Fall Risk Assessment and Data Collection

All participants completed the STEADI checklist, Short FES-I questionnaire, and BTrackS balance assessment. Participants who only spoke Spanish were given the validated Spanish versions of the STEADI checklist and Short FES-I questionnaire. Instructions and prompts for the BTrackS balance assessment were provided in either English or Spanish, depending on the participant’s primary language.

The BTrackS Balance System consists of a balance force plate and the BTrackS Assess Balance software (version 5.5.9). It measures center-of-pressure sway by calculating the distance traveled while maintaining a static posture, following the inverse pendulum model. Participants first had a 20-second practice trial that did not count towards their final score. This was followed by 3 additional 20-second trials, which were averaged to determine their final score. During each trial, participants were instructed to stand on the pre-marked balance plate with their feet approximately 30cm apart, hands on hips, and eyes closed. To minimize the risk of falling, a walker or sturdy piece of furniture was placed within their reach during each trial.

For the BTrackS assessment, a postural sway score ≤ 30cm indicated a low risk for falling, while a score > 30cm indicated a high risk for falling. For the Short FES-I questionnaire, a score ≤ 10 indicated a low risk for falling, while a score > 10 indicated a high risk for falling. As for the STEADI checklist, a score ≤ 4 indicated a low risk for falling, while a score > 4 indicated a high risk for falling. Additionally, participants who reported feeling unsteady when standing or walking, fears about falling, or a history of falls within the past year on the STEADI checklist (ie, answering yes to questions 1, 3, and/or 5) were automatically classified as high risk for falling, regardless of their total score.

<table>
<thead>
<tr>
<th>Table 1 Participant Characteristics</th>
<th>All Participants (N = 122)</th>
<th>Female Participants (n = 94)</th>
<th>Male Participants (n = 28)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>74.8 ± 7.3</td>
<td>75.1 ± 7.3</td>
<td>74.1 ± 7.6</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>166.0 ± 8.6</td>
<td>163.0 ± 6.1</td>
<td>175.0 ± 9.0</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>74.4 ± 16.1</td>
<td>71.3 ± 15.7</td>
<td>84.8 ± 12.9</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.9 ± 5.0</td>
<td>26.8 ± 5.4</td>
<td>27.1 ± 3.4</td>
</tr>
<tr>
<td>Body Fat %</td>
<td>27.9 ± 13.7</td>
<td>31.1 ± 12.4</td>
<td>17.7 ± 12.7</td>
</tr>
</tbody>
</table>

**Note:** Data are presented as mean ± standard deviation.

**Abbreviations:** cm, centimeters; kg, kilograms; BMI, Body Mass Index.
Statistical Analyses
All statistical analyses were performed using jamovi version 2.4.1.\textsuperscript{15–17} A Kolmogorov–Smirnov test revealed that the data did not follow a normal distribution. Consequently, nonparametric tests were employed. Kendall’s Tau correlation coefficients were utilized to evaluate the relationship between STEADI Checklist, Short FES-I, and BTrackS scores as continuous variables. McNemar tests were used for comparing how each measure categorized participants as high or low fall risk. Sensitivity was calculated as [True Positives/(True Positives + True Negatives)], specificity as [True Negatives/(True Negatives + False Positives)], and accuracy as [(True Positives + True Negatives)/N].\textsuperscript{5,17} Positive likelihood ratios were calculated as [Sensitivity / (1 – Specificity)] and negative likelihood ratios were calculated as [(1 – Sensitivity) / Specificity].\textsuperscript{17} The threshold for statistical significance was set at $p < 0.05$.

Results
Table 2 provides an overview of the scores obtained by participants on each fall risk assessment. The results indicate a significant positive correlation between Short FES-I scores and STEADI scores ($\tau_b = 0.51$, $p < 0.001$), as well as between Short FES-I scores and BTrackS scores ($\tau_b = 0.14$, $p = 0.04$). However, no significant correlation was found between BTrackS and STEADI scores ($\tau_b = 0.11$, $p = 0.10$).

Figure 1 shows the fall risk appraisal for each assessment. The significance of differences in fall risk classification between the STEADI checklist and BTrackS ($X^2 = 6.1$, $p = 0.014$), as well as between the Short FES-I and STEADI checklist ($X^2 = 9.53$, $p = 0.002$), were determined using the McNemar test. However, no significant difference was found in the classification of fall risk between the BTrackS and Short FES-I ($X^2 = 0.087$, $p = 0.77$).

When compared to the BTrackS as the criterion measure, the STEADI checklist demonstrated a specificity of 62.8%, sensitivity of 70.5%, accuracy of 65.6%, positive likelihood ratio of 2.13, and negative likelihood ratio of 0.5. Similarly, when compared to the Short FES-I as the criterion measure, the STEADI checklist exhibited a specificity of 67.5%, sensitivity of 81.0%, accuracy of 72.1%, positive likelihood ratio of 3.54, and negative likelihood ratio of 0.4. Conversely, the BTrackS, when compared to the Short FES-I as the criterion measure, displayed a specificity of 70.0%, sensitivity of 47.6%, accuracy of 62.3%, positive likelihood ratio of 1.34, and negative likelihood ratio of 0.4.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>STEADI Score</th>
<th>Short FES-I Score</th>
<th>BTrackS Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.7</td>
<td>10.2</td>
<td>31.2</td>
</tr>
<tr>
<td>Median</td>
<td>2.0</td>
<td>9.0</td>
<td>27.0</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.91</td>
<td>3.97</td>
<td>17.50</td>
</tr>
<tr>
<td>Interquartile Range</td>
<td>4.0</td>
<td>5.0</td>
<td>15.5</td>
</tr>
</tbody>
</table>

Note: BTrackS scores are presented in centimeters.

Abbreviation: STEADI, Stopping Elderly Accidents Deaths & Injuries.

![Figure 1](https://doi.org/10.2147/CIA.S453966)
0.63. Lastly, in comparison to the BTrackS as the criterion measure, the Short FES-I yielded a specificity of 71.79%, sensitivity of 45.5%, accuracy of 62.3%, positive likelihood ratio of 1.32, and negative likelihood ratio of 0.62.

**Discussion**

The purpose of this study was to compare the specificity, sensitivity, and accuracy in fall risk appraisal between the STEADI, Short FES-I, and BTrackS assessments among community-dwelling older adults. Our hypothesis that the three assessments would significantly differ in fall risk appraisal was partially supported. Fall risk appraisal significantly differed between the STEADI checklist and BTrackS, as well as between the STEADI checklist and Short FES-I questionnaire. However, there was no significant difference between the appraisals of the Short FES-I and the BTrackS. The STEADI checklist categorized the most participants as high fall risk compared to the Short FES-I and BTrackS assessments, suggesting a potentially diminished ability to distinguish between low and high fall risk.

The STEADI checklist is the first step in the CDC’s fall risk screening algorithm. If a person is categorized as high fall risk, additional assessments and interventions are conducted. The purpose of the STEADI checklist is to accurately identify older adults who are at a higher risk of falling in a timely manner. This is important because time constraints and concerns often prevent the implementation of other fall risk appraisal strategies. However, categorizing patients as high fall risk when they actually are not leads to unnecessary time spent on additional assessments and interventions. In the STEADI algorithm, patients may initially be flagged as high fall risk based on the checklist, but then reclassified as low fall risk after further assessments like the TUG and 4-stage balance test. This can result in repetitive cycles of classification and reclassification, which can be costly in terms of time and money in clinical practice.

The Short FES-I and BTrackS assessments have been validated as assessments of fall risk in older adults, but with consideration from different perspectives. The Short FES-I directly assesses fear of falling, while the BTrackS directly assesses center-of-pressure postural sway during a static stance. Compared to a classification of fall risk by direct measurement of postural sway, the STEADI checklist demonstrated low accuracy (65.57%) likely due to a low amount of specificity (62.82%), representing a diminished ability to correctly identify those with low fall risk. Postural sway assessments such as the BTrackS have also been shown to correlate well with the fear of falling, which is assessed with the Short FES-I. Compared to a classification of fall risk by a direct measurement of fear of falling, the STEADI checklist demonstrated low-to-moderate accuracy (72.13%) like due to a low amount of specificity (67.5%). The STEADI checklist includes questions aimed at gauging fear of falling and postural sway; however, it may be more beneficial and time-efficient to measure these factors directly. The Short FES-I contains 7 multiple choice questions and the BTrackS assessment is approximately 2-minutes long when accounting for participant prompting. Furthermore, previous research from our group has demonstrated the potential for additional insights into factors of fall risk when using the combination of Short FES-I and BTrackS assessments. This may allow for more targeted interventions to reduce fall risk by identifying specific factors that require resolution, many of which may not be typically assessed by clinicians during fall risk screenings. Including direct assessments of postural sway and fear of falling appears to provide a more accurate fall risk appraisal than the STEADI checklist alone.

Although the STEADI checklist provides valuable insight into different fall risk factors, such as medication use and fall history, it may not be sufficient to rely solely on a single instrument to cover all variables. Likewise, previous research has suggested that it is more effective to use both objective and subjective measures when assessing older adults, as subjective reports may not always be entirely accurate. The BTrackS offers an objective assessment of postural sway and can be used in conjunction with the subjective self-reporting included in either the Short FES-I or STEADI assessments. Combining subjective self-reporting with an objective assessment of postural sway would provide clinicians with information on various fall risk factors and aid them in making more accurate decisions. The present results demonstrate weak correlations between the BTrackS and STEADI assessments, as well as between the BTrackS and Short FES-I, which support the idea that the BTrackS offers a unique assessment not provided by STEADI and Short FES-I. The BTrackS and Short FES-I had moderate specificity compared to each other, and previous research has shown that their combination can be used beyond the binary high/low fall risk appraisal.

Despite the usefulness of center-of-pressure assessments such as the BTrackS there is often a financial barrier to implementing them, unlike questionnaires like the STEADI checklist and Short-FES-I. While it may be worth conducting
a cost-benefit analysis on a clinician level, the present results indicate the need for further investigation into the STEADI algorithm’s placement of balance assessments. Instead of performing balance assessments as the second step in the algorithm, including a balance assessment in the initial screening would provide a more specific evaluation compared to relying solely on subjective assessments. This is especially important considering that falls are a direct result of balance loss. Although concerns about the time required for balance testing are understandable, previous evidence has shown that balance tests can be conducted within a reasonable timeframe, even in emergency departments. The algorithm includes a 4-stage balance test that requires no equipment and takes about 1 minute to complete. The 4-stage balance test’s 1-minute duration, as well as the 2-minute duration of the BTrackS assessment, falls within the 3-minute timeframe considered feasible for clinicians in time-sensitive settings. While we have limited knowledge about the correlation between the 4-stage balance test and direct center-of-pressure measures, it may be worthwhile to explore it as a potentially cost-effective alternative to more expensive balance assessments. This could encourage clinicians to incorporate balance testing alongside subjective questionnaires when screening for fall risk, potentially leading to more accurate fall risk classification.

**Study Limitations**
The small sample size of male participants did not allow for a high-quality comparison of fall risk classification between female and male participants, and future research should attempt to determine whether sex differences are present. This study was also cross-sectional in design; therefore, we were unable to provide follow-up information on falls sustained after fall risk assessment. It should be noted that there is no gold standard fall risk assessment. The present study used the BTrackS and Short FES-I as criterion measures because the STEADI checklist was the primary fall risk assessment of interest, given its CDC recommendation. However, CDC recommendation does not equate gold standard status, particularly within a global scope. A strong case could be made for other questionnaires or direct assessments of postural sway to be considered the best fall risk assessment method, but the decision of which assessment to use often relies on the clinician’s access to technology and preferences. The lack of a gold standard to compare to is a limitation of any study investigating assessment methods, and future research should aim to validate a gold standard fall risk assessment for older adults.

**Conclusion**
Although the CDC’s STEADI checklist is highly sensitive, it may not be as specific and accurate in assessing fall risk compared to the Short FES-I and BTrackS. As a result, additional follow-up assessments are necessary to confirm fall risk appraisal according to the STEADI algorithm, as the STEADI checklist appears to overcategorize individuals as high fall risk. However, by combining a balance assessment like the BTrackS with the STEADI checklist, or a different questionnaire such as the Short FES-I, a more efficient and accurate fall risk appraisal may be achieved, which requires less follow-up testing and provides both objective and subjective insights into fall risk. Previous research has supported the combined use of the BTrackS and Short FES-I as a fall risk screening tool, but further research is needed to validate these findings. Additionally, future research should explore the combined use of the BTrackS or similar postural sway assessments with the STEADI checklist as a fall risk screening tool.

**Data Sharing Statement**
The deidentified data that support the findings of this study are available from the corresponding author, LT, upon reasonable request.

**Funding**
This work was supported by the National Institute on Aging under Grant R03AG06799; the National Institute on Minority Health and Health Disparities under Grant R01MD018025; and the Office of the Director, Chief Officer for Scientific Workforce Diversity, Office the National Institutes of Health under supplemental Grant number 3R01MD018025-02S1.
Disclosure

KL, AB, and JRS received financial support from the National Institutes of Health under supplemental Grant number R01MD018025-02S1. JRS, ET, NL, JP, RX and LT received financial support from the National Institute on Aging (R03AG06799), the National Institute on Minority Health and Health Disparities (R01MD018025) and the Office of the Director, Chief Officer for Scientific Workforce Diversity (COSWD) Office (3R01MD018025-02S1) of the National Institutes of Health. LT also received support from the National Science Foundation (NSF2222662). DHP has no conflicts of interest to report. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

11. Goble DJ, Baweja HS. Normative data for the btracks balance test of postural sway: results from 16,357 community-dwelling individuals who were 5 to 100 years old. Phys Ther. 2018;98(9):779–785. doi:10.1093/ptj/pzy062
20. Qin H, Xiong S. Center-of-pressure based postural sway measures: reliability and ability to distinguish between age, fear of falling and fall history. Int J Ind Ergonom. 2015;47:37–44. doi:10.1080/07488208.2015.102.004