Original Manuscript

Development and Validation of the ReTos Scale as a Support Tool for Fall Risk Assessment in a Large Sample of Older Italian Inpatients

Evaluation & the Health Professions
2026, Vol. 0(0) 1–9

DOI: 10.1177/01632787261476234
journals.sagepub.com/home/ehp
Mary Ann Liebert
A Part of Sage

Laura Rasero[1], Paolo Iovino[1], Angela Brandi[2], Monica Simonetti[3], Riccardo Tartaglia[4], and Tommaso Bellandi[5]

Abstract

Aim of this study was to develop and evaluate the validity and reliability of the ReTos (Regione Toscana) scale for identifying older inpatients at risk of falls. 15,159 inpatients aged ≥65 years from 41 Italian hospitals were enrolled with six-month follow-up. Items were generated through literature review and expert consensus. Content validity was assessed by a multidisciplinary panel. Exploratory factor analysis (EFA) examined dimensionality. Internal consistency was estimated using McDonald's Omega. Criterion-related validity was tested through logistic regression, and discriminative performance assessed using the area under the ROC curve. Mean age was 79 years (52.7% female). During follow-up, 278 patients experienced at least one fall (incidence 0.59 per 1,000 patient-days). EFA reduced the original 22 items to 11 items across five factors: motor disorders, mental state, diseases/medications, anamnesis and sensory impairment. Several items were significantly associated with fall occurrence. Internal consistency was acceptable (Ω = 0.79). Discriminative performance was modest (AUC = 0.62; 95% CI 0.58–0.66), with sensitivity 0.70 and specificity 0.52 at a cut-off ≥7; positive predictive value was low (0.03), reflecting low event incidence. The ReTos scale demonstrated satisfactory psychometric properties but modest discriminative performance, supporting its use as an adjunct to clinical judgment rather than a stand-alone tool.

Keywords: accidental falls, ReTos scale, aged, validity, reliability, sensitivity and specificity, psychometrics

Background

Falls are the second leading cause of unintentional injury deaths worldwide, with nearly 700,000 fatalities each year (World Health Organization, 2021). Older adults are particularly vulnerable due to age-related declines in strength, balance, and cognition, as well as multimorbidity and polypharmacy (Appeadu & Bordoni, 2023). In the United States, one in four adults over 65 experiences a fall annually (Kakara, 2023). Hospitalization further increases the risk, with recent European data reporting inpatient fall rates between 3.4% and 3.9% (Bernet et al., 2024). In Italy, falls remain a major safety concern, ranking fourth among adverse events after surgical, diagnostic, and therapeutic errors (Amato et al., 2024).

Falls in hospitals lead to serious physical and psychological harm, prolonged hospitalization, and higher mortality (Cui et al., 2023; James et al., 2020). Prevention therefore represents a key patient safety priority. Multifactorial strategies are recommended, combining individual risk evaluation and targeted interventions. Screening tools play a crucial role by synthesizing known risk factors to guide prevention. However, existing instruments show limited predictive validity; a review of 26 scales concluded that two tools should be used concurrently to improve accuracy (Park, 2018).

According to guidelines issued by the National Institute for Health and Care Excellence (NICE, 2013, 2025), risk assessment should complement rather than replace clinical judgment, focusing on factors such as medication review, mobility, cognition,

[1]Department of Health Sciences, University of Florence, Florence, Italy
[2]Nursing and Midwifery Department, Careggi University Hospital, Florence, Italy
[3]Independent Healthcare Consultant, Florence, Italy
[4]Department of Engineering Science, Marconi University, Rome, Italy
[5]Patient Safety Unit, Tuscany North West Trust, Italy

Corresponding Author:
Paolo Iovino, Department of Health Sciences, University of Florence, Via Giovanni Battista Morgagni 48, 50134 Florence, Italy.
Email: paolo.iovino@unifi.it

and sensory impairment. Recent evidence also highlights the need for user-friendly tools that combine psychometric soundness with ergonomic usability to ensure adoption in clinical practice (Czuber et al., 2024; Ploegmakers et al., 2024). Developing such integrated instruments remains a priority for improving fall prevention among hospitalized older adults. Therefore, further studies are needed to develop and evaluate fall-risk screening tools for hospitalized older people. The ReTos scale was developed to address these gaps by providing a concise, clinically oriented tool that integrates multiple domains of fall risk into a structured set of indicators. Unlike many existing instruments, ReTos combines a multidomain approach with item-level weighting based on expert judgment, while maintaining feasibility for routine use in hospital settings. In addition, its development and validation were conducted on a large multicenter inpatient sample, enhancing the robustness of the findings. Therefore, the aim of this study was to develop and test the validity and reliability of a new supportive tool for evaluating fall risk factors in a large cohort of older patients in a variety of hospital settings.

Materials and Methods

Study Design and Setting

This prospective, multicenter observational study was conducted over six months in 41 hospitals (150 wards) of the Tuscany Region (Italy) within the National Health System. It was part of a regional fall-prevention program coordinated by the Regional Center for Clinical Risk Management, the Regional Health Service, and the University of Florence. Wards joined voluntarily following a public call. Ethical approval was granted by the institutional review board of Azienda USL Toscana Centro, and patient-level data were securely stored with restricted access.

Study Sample and Data Collection

Older inpatients aged ≥65 years with an admission beyond 24 hours were enrolled in the study. We excluded those with an unconscious status or who were terminally ill. Data were collected by ward nurses from 41 hospitals of the Tuscany Region, covering Internal Medicine, Pneumology, Neurology, Geriatrics, Cardiology, Oncohematology, Orthopedics, and Surgery units. All participating nurses received standardized training on screening and prevention procedures. At admission, they completed the ReTos scale along with three established fall-risk tools (Morse Fall Scale, Conley Scale, and STRATIFY). Fall events were prospectively recorded during hospitalization. The six-month period refers to the overall data collection timeframe across participating wards.

Instruments

The ReTos was used to evaluate the risk of falls. The final version of this tool is composed of 11 items distributed into the domains of motor disorders (3 items), mental state (2 items), current diseases and medications (2 items), anamnesis (2 items), and sensory impairments (2 items). Scores assigned to each item are weighted from a score of 2, assigned to item 5, to a score of 6, assigned to item 11 (Supplemental Table S1). The ReTos yields a total score ranging from 0 to 49, with higher scores indicating a higher risk of falls. A fall reporting form was designed to describe the characteristics of the event and its consequences, the place and circumstances of the fall, and possible contributory factors. A sociodemographic and clinical form was also used to collect data on hospital stay (length of stay, clinical setting), sociodemographic (age, sex) and clinical characteristics (diagnosis at discharge). These data were retrieved from the deidentified regional health dataset through a nosologic code, specific to each hospital admission. Alongside the ReTos, three widely used fall-risk assessment tools (Morse Fall Scale, Conley Scale, and STRATIFY) were also administered in order to allow direct comparison of discriminative performance within the same patient population.

Instrument Development

The ReTos scale was developed and validated following the COSMIN checklist (Mokkink et al., 2010). Item generation was informed by a targeted literature review aimed at identifying risk factors for falls in hospitalized older adults. The search was conducted in PubMed, CINAHL, and Scopus using combinations of the keywords "falls," "risk assessment," "older adults," and "hospital." Two researchers independently screened the literature to identify relevant studies, clinical guidelines, and existing fall-risk assessment tools. Key risk domains were extracted and synthesized to inform the development of the initial item pool.

Content validity was established through a multistep process combining evidence from the literature and expert consensus. A multidisciplinary panel of ten experts (geriatric nurses, physicians, physiotherapists, and clinical risk managers), each with at least five years of experience in fall prevention or research, was purposively selected. Candidate items were initially derived from a targeted literature review and from commonly used validated fall-risk assessment tools (e.g., Morse, STRATIFY, Hendrich II, Conley). These items were independently reviewed by the expert panel using a structured approach, with each item evaluated in terms of relevance, comprehensiveness, and clarity (Polit & Beck, 2006). The consensus process was conducted through iterative rounds of structured discussion, during which items were progressively refined, merged, or excluded to reduce redundancy and improve interpretability. Additional items were proposed by the expert panel when considered clinically relevant but insufficiently represented in the literature. Consensus was reached when no further modifications were proposed by

the panel. The resulting preliminary version of the ReTos scale included 22 items covering key domains of fall risk in hospitalized older adults.

Statistical Analysis

Continuous variables were summarized as means and standard deviations, and categorical variables as frequencies and percentages. Prevalence was defined as the proportion of patients experiencing at least one fall during hospitalization. Incidence was calculated as the number of fall events per 1,000 patient-days, using total length of stay as the denominator. Statistical significance was set at $p < 0.05$ (two-tailed).

Scale development proceeded in five steps. (1) Exploratory factor analysis (EFA) was conducted on the initial item pool using polychoric correlation matrices, given the ordinal nature of the variables, with Varimax rotation applied to aid interpretability of the latent structure of the initial 22-item pool. Factors with eigenvalues >1 were retained, and item retention was guided by factor loadings ≥0.32 together with considerations of interpretability, redundancy, and clinical feasibility, with the aim of developing a parsimonious tool for routine inpatient screening. Sample adequacy was confirmed with KMO and Bartlett tests. (2) Criterion-related validity was examined via univariate logistic regression models including each item separately as a predictor of fall occurrence. No covariates were included. (3) Internal consistency was assessed with Cronbach's α (for domains) and McDonald's Ω (for the total scale), considering ≥0.70 acceptable. (4) Inter-rater reliability was evaluated in a subsample of 150 patients independently assessed by two trained nurses. Each nurse completed the ReTos scale separately at patient admission, following standardized assessment procedures. Agreement at the item level was estimated using weighted Cohen's κ coefficients, while intraclass correlation coefficients (ICCs) were calculated using a two-way random-effects model to assess absolute agreement for scale scores. The sample size refers to independent patient observations, with agreement computed across paired ratings. (5) Item weighting was derived through the Thurstone paired-comparison method (Krabbe, 2008), involving seven geriatric nurses who assigned relative importance to each indicator. Discriminative performance was evaluated using receiver operating characteristic (ROC) curve analysis, with the area under the curve (AUC) used to assess discrimination. Sensitivity, specificity, and predictive values were calculated at the selected cut-off. Differences between AUCs of ReTos and comparator scales (Morse, Conley, STRATIFY) were tested using the DeLong test. Missing data were minimal and handled using a complete-case approach. Given the low proportion of missing values, no imputation procedures were considered necessary. Analyses were conducted with STATA 18 (StataCorp, College Station, TX).

Results

Descriptives of the Sample

We collected data on a total sample of 15,159 inpatients (males: n = 7,150, 47.2%; females: n = 8,009, 52.8%). The most frequent diseases were cardiovascular (n = 1,399, 9.2%), and respiratory (n = 956, 4.2%) while the average length of stay was 13.19 days (range: 7.3 in coronary units and cardiological settings and 28.4 in long hospitalization settings). 18.5% of the sample underwent surgical interventions. The sample was mostly admitted to medicine wards (58.9%), followed by orthopedic (10.7%), surgery (7.6%), and cardiology (5.6%) units.

At the end of the study data collection (6 months) 278 patients (1.8%) experienced at least one fall. The incidence rate was 0.59 falls per 1,000 patient-days. Among fallers, 154 (55.4%) were males (mean age = 78.9, SD = 6.4) and 124 were females (mean age = 80.1, SD = 7.3). Sixteen patients experienced recurrent falls during hospitalization. Falls were most frequent in medical wards (n = 155), followed by neurology (n = 18) and oncology units (n = 18).

Exploratory Factor Analysis

Factorability of the data was considered adequate to proceed with the main analysis (Kaiser-Meyer-Olkin index = 0.95; Bartlett test of sphericity: $p < 0.001$). The initial EFA of the 22-item pool identified 7 factors with eigenvalues >1. Six items ("*urine and fecal incontinence*", "*use of a medical walker*", "*use of a wheelchair*", "*cling to furnishings*", "*use of psychotropic drugs*" and "*intravenous infusion*") showed factor loadings below 0.32 and were removed. The EFA was then repeated without these items. During this second analysis, five additional items ("*frequent fecal or urine elimination*", "*use of walking aids*", "*difficulty moving in the bed*", "*difficulty maintaining a stable upright position*", and "*disoriented, confused*") were excluded because they did not contribute to a stable and interpretable factor solution. The final EFA retained 11 items loading on five factors. Based on the conceptual meaning of the retained items, these factors were labelled motor disorders, mental state, diseases and medications, anamnesis, and sensory impairment. The five-factor solution explained 60.62% of the total cumulative variance (Table 1).

Criterion-Related Validity

In the univariate logistic models (Table 2), most ReTos items were significantly associated with fall occurrence ($p < 0.01$) except for the items of the "diseases and medications" (i.e., "presence of diseases", $p = 0.07$, and "medications", $p = 0.28$) and "sensory impairment" factors (i.e., "vision impairment", $p = 0.82$, and "hearing impairment", $p = 0.31$) and one item in the "anamnesis" factor ("vertigo episodes in the previous six months", $p = 0.05$).

Table 1

Exploratory Factor Analysis of the ReTos Scale (n = 15,159)

Item	Motor disorders	Mental state	Diseases & medications	Anamnesis	Sensory impairment	h^2
1.Needs of assistance	0.55					0.69
2. Impaired mobility	0.64					0.56
3. Difficulty in performing ADL	0.61					0.52
4. Impaired judgement/Lack of safety awareness		0.60				0.58
5. Agitation		0.52				0.71
6. Presence of diseases			0.43			0.76
7. Medications			0.41			0.82
8. History of falls (last 6 months)				0.39		0.79
9. Vertigo episodes (last 6 months)				0.41		0.79
10. Vision impairment					0.40	0.78
11. Hearing impairment					0.40	0.81

[a]*Note.* h^2 = communality coefficient. Extraction method: Principal axis factoring. Rotation method: Varimax. The five-factor solution explained 60.62% of the total cumulative variance. ADL, activities of daily living.

Internal Consistency and Inter-rater Agreement

Internal consistency was acceptable, with McDonald's Omega values of 0.70 for the motor disorder factor, 0.60 for the mental state factor, 0.71 for the disease and medication factor, 0.68 for the anamnesis factor, and 0.68 for the sensory impairment factor. The Omega coefficient for the entire scale was 0.79. Inter-rater reliability analysis showed satisfactory agreement across items, with weighted Cohen's kappa values ranging from 0.66 to 0.82 (mean $\kappa = 0.74$). The overall intraclass correlation coefficient (ICC, two-way random effects model) for the total ReTos score was 0.79 (95% CI: 0.72–0.84), indicating good reproducibility between raters (Table 3).

Scoring Weights Assignment

Supplemental Table S1 reports the matrix of the normal deviates of the paired comparisons according to the Thurstone method. The items with the lowest weights were "vertigo

Table 2

Univariate Logistic Regressions of ReTos Indicators Predicting Fall Events

Factors and items	n	%	OR	p	95% CI
Motor disorders					
Needs of assistance	3,141	20.7	1.75	<0.001	1.35–2.26
Impaired mobility	5,895	38.9	2.26	<0.001	1.78–2.88
Difficulty in performing ADL	7,132	47.0	1.88	<0.001	1.47–2.40
Mental state					
Impaired judgement/Lack of safety awareness	2,484	16.4	1.64	<0.001	1.24–2.17
Agitation	1,365	9.0	1.62	0.010	1.14–2.29
Diseases and medications					
Presence of diseases	10,821	71.4	1.29	0.070	0.98–1.71
Medications	11,902	78.5	0.86	0.280	0.65–1.13
Anamnesis					
History of falls in the six previous months	3,545	23.4	1.61	<0.001	1.25–2.07
Vertigo episodes in the six previous months	4,301	28.4	1.29	0.050	1.00–1.67
Sensory impairment					
Vision impairment	2,285	15.1	1.04	0.82	0.75–1.44
Hearing impairment	2,867	18.9	0.85	0.310	0.62–1.17

Note. OR, odds ratio; CI, confidence interval; *p*, *p* value; ADL, activities of daily living.

Table 3

Inter-rater Agreement of the ReTos Scale by Domain (n = 150)

Domain	Number of items	Weighted Cohen's κ (range)	Mean κ	ICC (95% CI) for domain score
Motor disorders	3	0.68–0.82	0.76	0.80 (0.72–0.86)
Mental state	2	0.66–0.74	0.70	0.77 (0.69–0.84)
Diseases & medications	2	0.67–0.75	0.71	0.78 (0.70–0.85)
Anamnesis	2	0.66–0.73	0.69	0.76 (0.68–0.83)
Sensory impairment	2	0.67–0.72	0.69	0.75 (0.67–0.82)
Total scale	11	—	—	0.79 (0.72–0.84)

Note. Agreement statistics are based on a subsample of 150 inpatients independently assessed by two trained nurses across wards. Weighted Cohen's κ was calculated at the item level; values ≥0.60 indicate acceptable agreement, ≥0.75 good agreement. ICC = intraclass correlation coefficient (two-way random effects model).

episodes in the six previous months" and "agitation". The highest weights were assigned to items "impaired mobility" and "difficulty in performing ADL".

Sensitivity and Specificity

Based on the Youden index, the optimal cut-off of the ReTos was ≥7, with a corresponding sensitivity and specificity of 0.70 and 0.52, respectively, indicating modest discriminative performance. The positive and negative predictive values were 0.03 and 0.99, respectively. The area under the ROC curve was 0.62 (Figure 1).

The ROC analysis confirmed that the AUC of the ReTos (0.62) was comparable to that of the Morse (0.61), Conley (0.59), and STRATIFY (0.53), with no statistically significant differences (DeLong test, $p > 0.05$). The optimal cut-off identified by the Youden index for the ReTos was ≥7, while the other scales showed variable thresholds, none of which improved discriminative performance (Supplemental Table S2).

Discussion

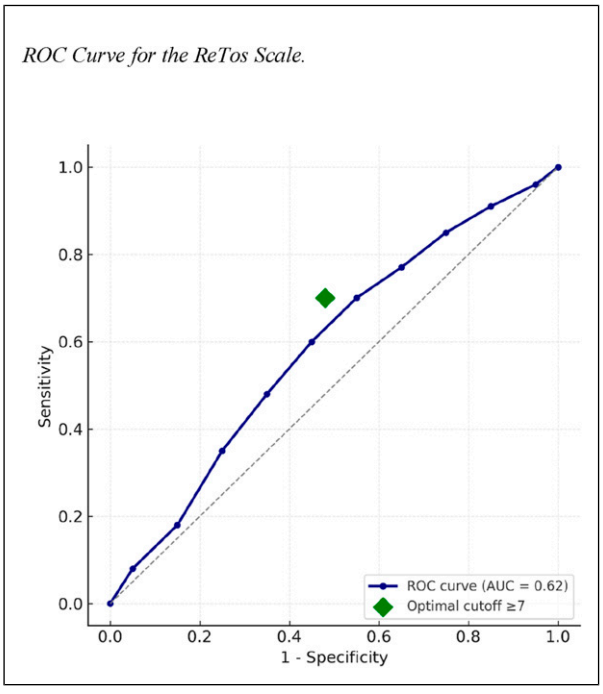
In this study we developed a simple tool for assessing the fall risk of elderly inpatients admitted to a variety of Italian clinical settings. The findings suggest that the ReTos has satisfactory psychometric properties and may be considered as a supportive tool for routine screening assessments. Several instruments are present in the literature, but only a few of them are available to measure the risk of falls in older inpatients.

Validity and Reliability

The reliability of the ReTos was found to be satisfactory for the overall scale but it was at the threshold of acceptability at the factor level, and specifically for the factors "mental state", "anamnesis", and "sensory impairment", which is partly explained by the limited number of items per factor. Indeed, reliability coefficients tend to increase with the number of indicators included. In our case, while some two-item factors (e.g., motor disorder, Ω = 0.70) achieved minimally acceptable values, the overall pattern suggests that subscales composed of only two indicators are inherently less stable. Only the total scale, which aggregates a larger set of items, reached a more satisfactory level of internal consistency (Ω = 0.79). However, the ReTos was not conceived as a purely reflective latent construct, but rather as a structured set of clinically relevant risk indicators. In line with current

Figure 1

ROC Curve for the ReTos Scale



Note. The area under the curve (AUC) was 0.62 (95% CI 0.58 to 0.66), indicating modest discriminative ability. Based on the Youden index, the optimal cut-off point was ≥7, which yielded a sensitivity of 0.70 and a specificity of 0.52

guideline recommendations, the identification and management of individual risk factors is prioritized over reliance on aggregated total scores. In this perspective, individual items represent the primary unit of interpretation, supporting targeted clinical decision-making. The total score, supported by acceptable internal consistency, may provide a useful summary measure, but should not replace item-level assessment.

We also found that most of our hypotheses were satisfied, since most ReTos items were significantly associated with fall occurrence. The item investigating medications was not significant. It is possible that this result was due to the characteristics of our sample, because most patients (78.5%) took medications, thus reducing the variability of this indicator.

Comparison With Literature

The literature shows that only a few instruments effectively assess fall risk among hospitalized older adults. A systematic review concluded that existing scales should be used in combination, as none alone demonstrate adequate sensitivity or specificity (Park, 2018). ReTos may complement established tools such as STRATIFY or the Hendrich II Fall Risk Model, which include fewer items and may overlook relevant predictors. By summarizing key patient risk factors, ReTos supports—rather than replaces—clinical judgment, facilitating systematic assessment and individualized prevention planning.

Items related to sensory impairment were not significantly associated with fall risk, diverging from prior evidence (Jin et al., 2024). However, a recent study found that only dual sensory impairment predicted post-discharge falls, not those occurring during hospitalization (Yamada et al., 2024). The ReTos showed satisfactory reliability overall, though some two-item factors were near the acceptability threshold—an expected finding given the limited number of indicators per factor. Thus, only the total score is recommended for clinical screening. The Thurstone paired-comparison method assigned the highest weights to *impaired mobility* and *difficulty in performing ADL*, consistent with literature identifying gait and balance impairment as major determinants of falls (Wang et al., 2025). It is worth noticing that the use of the Thurstone method was intended to reflect the perceived clinical relevance of different risk indicators rather than to maximize statistical prediction. In this sense, weighting represents a pragmatic attempt to align the scoring system with expert knowledge, although its added value in terms of discriminative performance remains to be demonstrated.

ReTos demonstrated acceptable sensitivity but modest specificity, consistent with findings from other inpatient fall-risk tools. Similar results were reported for the STRATIFY scale (sensitivity = 0.67, specificity = 0.61) in hospitalized adults (Cortés et al., 2023). This pattern may reflect the characteristics of our sample, including the relatively low incidence of falls and the heterogeneity of diagnoses and care settings, which can reduce discrimination power. Moreover, the very low PPV (0.03) reflects the low prevalence of falls in this study sample and should not be interpreted in isolation. PPV is strongly dependent on outcome prevalence, and similar values have been reported in comparable inpatient settings. However, this also implies that a substantial proportion of patients identified as high risk may not experience a fall, with potential implications for resource allocation and clinical workload. More specifically, this performance profile should imply a trade-off between sensitivity and specificity. While higher sensitivity may reduce missed cases, the relatively low specificity and very low positive predictive value indicate that a substantial proportion of patients may be incorrectly classified as high risk. This may result in increased clinical workload and resource utilization, particularly if the tool is applied indiscriminately.

When considering all performance indicators—sensitivity, specificity, predictive values, and AUC (0.62), ReTos showed a performance profile comparable to that of the tested tools. These results confirm that ReTos should not be interpreted as a stand-alone predictive tool, but rather as a structured aid to support clinical judgment and identify relevant risk factors (Park, 2018). In this perspective, its relatively higher sensitivity and very high negative predictive value may help reduce missed cases when integrated into multifactorial assessment strategies.

Strengths and Limitations

This study has several strengths. The large multicenter sample enhances generalizability and provides a robust basis for psychometric testing. The progressive item reduction and weighting process allowed the development of a concise and clinically oriented tool—an approach that differs from many existing instruments assigning equal weights to items. In addition, the comparison with established scales (Morse, Conley, STRATIFY) supports the consistency of the observed performance profile with previous research.

However, several limitations should be acknowledged. First, the study design was not intended to evaluate the impact of the ReTos scale on fall incidence or patient outcomes, and the relatively short follow-up period further limits conclusions on its effectiveness in clinical practice. Moreover, criterion-related validity was assessed using univariate logistic regressions only; therefore, the independent predictive contribution of each item and the risk of model overfitting could not be evaluated. Second, some clinically relevant indicators identified in the literature, such as incontinence or use of walking aids, were not retained in the final structure; although this was consistent with the empirical reduction strategy adopted to derive a brief and feasible tool, it may have reduced the breadth of content covered by the instrument.

Structural validity was not further assessed through confirmatory factor analysis, which limits the robustness of the conclusions regarding dimensionality. In addition, internal consistency at the domain level was limited, supporting the interpretation of the factor structure as exploratory rather than as a stable multidimensional construct. Third, the weighting procedure was based on a relatively small panel of experts and no comparison with unweighted scoring was performed, preventing conclusions on the incremental value of this approach.

Fourth, variability in length of stay implies differences in time at risk that were not explicitly modeled in the regression analyses. Although incidence rates per patient-days were calculated, the use of logistic regression does not account for time-to-event and may therefore provide a simplified representation of fall risk. External validity is also limited by the regional scope of the study and the voluntary participation of hospitals, which may introduce selection bias at the institutional level. Finally, the low number of fall events resulted in substantial class imbalance, which likely contributed to the low positive predictive value and limited the discriminative performance of the ReTos. Therefore, the instrument should be regarded as a screening support tool rather than a stand-alone predictive model. External validation in independent and more diverse populations is warranted.

Practice and Research Implications

This study offers meaningful implications for clinical practice and future research. The latest NICE guideline (NICE, 2025) discourages the use of total risk scores with arbitrary cut-offs, emphasizing instead the identification and management of individual risk factors. In this context, the ReTos scale represents a useful complementary tool to support clinicians in systematically identifying fall-related factors and prioritizing tailored preventive strategies, without oversimplifying patients into high- or low-risk categories. ReTos provides a concise and clinically feasible approach to the systematic assessment of fall-risk factors, particularly for older inpatients. Its concise structure and focus on clinically relevant domains address the ongoing need for user-friendly, evidence-based instruments that fit routine nursing workflows (Jin et al., 2021). However, implementation of the ReTos scale should be carefully considered within clinical workflows, balancing the potential benefit of identifying at-risk patients against the burden associated with false-positive classifications.

Self-administered tools have shown promise in empowering older adults but remain of limited clinical applicability (Sun et al., 2019). More recent work has focused on ergonomically designed fall-risk instruments that integrate usability into their development (Garvin et al., 2025). Future research should evaluate the performance of the ReTos in more homogeneous populations and across diverse healthcare settings, as well as examine its responsiveness in comparison with other validated tools. Prospective studies are needed to determine whether its integration into multifactorial assessment strategies translates into improvements in fall incidence, patient safety, and quality of care in hospitalized older adults. In addition, the real-world impact and cost-effectiveness of the ReTos scale remain to be established, and further research is warranted to assess its implications for both patient outcomes and healthcare resource utilization. Furthermore, no internal cross-validation or split-sample validation was performed, which may limit the robustness of the observed performance estimates.

Conclusion

We developed the ReTos scale as a structured screening support tool to assist the assessment of fall risk among older inpatients across diverse clinical settings. The instrument demonstrated satisfactory psychometric properties, although its discriminative performance was modest. Rather than serving as a stand-alone predictive instrument, ReTos is intended to support the systematic identification of clinically relevant fall-risk factors and complement clinical judgment within a comprehensive multifactorial assessment. Further validation studies are needed to confirm its structural validity, evaluate its performance in different healthcare settings, and determine its impact on clinical outcomes and routine practice.

Acknowledgments

The authors wish to sincerely thank Donatella Rizzini and Stefania Rodella for their pivotal contributions to the feasibility and execution of the study, and for their valuable expertise in the statistical analyses.

ORCID iD

Paolo Iovino https://orcid.org/0000-0001-5952-881X

Ethical Considerations

Approval was obtained from the institutional review board of Azienda USL Toscana Centro (Italy)

Author Contributions

• LR (Laura Rasero): Conceptualization; study design; supervision; critical revision of the manuscript for important intellectual content.
• TB (Tommaso Bellandi): Conceptualization, methodology; project administration; data curation coordination; review of manuscript.
• MS (Monica Simonetti): Data collection supervision; investigation; resources; manuscript editing.
• RT (Riccardo Tartaglia): Conceptualization, methodology; interpretation of findings; critical revision; validation.
• PI (Paolo Iovino): Conceptualization; formal analysis; data interpretation; drafting of the manuscript; corresponding author responsibilities.

• AB (Angela Brandi): Clinical coordination; facilitation of field implementation; manuscript review and editing.
All authors read and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

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

Data Availability Statement

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

Patient and Public Involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Supplemental Material

Supplemental material for this article is available online.

References

Amato, S., Aniuskevich, A., Lombardi, A. M., & Amato, F. (2024). Healthcare risk management and monitoring of patient falls in Italian national service. The experience of a local health authority. *Igiene e Sanità Pubblica*, *89*(2), 30–40.

Appeadu, M. K., & Bordoni, B. (2023). Falls and fall prevention in older adults. In *Statpearls [internet]*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK560761/

Bernet, N. S., Everink, I. H., Hahn, S., Bauer, S., & Schols, J. M. (2024). Comparing risk-adjusted inpatient fall rates internationally: Validation of a risk-adjustment model using multicentre cross-sectional data from hospitals in Switzerland and Austria. *BMC Health Services Research*, *24*(1), 331. Article 331. https://doi.org/10.1186/s12913-024-10839-x

Cortés, O. L., Vásquez, S. M., & Mendoza, A. C. (2023). Validation of the stratify scale for the prediction of falls among hospitalized adults in a tertiary hospital in Colombia: A retrospective cohort study. *Scientific Reports*, *13*(1), 21640. Article 21640. https://doi.org/10.1038/s41598-023-48330-y

Cui, Y., Liu, B., Qin, M.-Z., Liu, Q., Ye, H., & Zhou, J. (2023). Effects of early mental state changes on physical functions in elderly patients with a history of falls. *BMC Geriatrics*, *23*(1), 564. Article 564. https://doi.org/10.1186/s12877-023-04274-6

Czuber, N. K., Garabedian, P. M., Rice, H., Tejeda, C. J., Dykes, P. C., & Latham, N. K. (2024). Human-centered design and development of a fall prevention exercise app for older adults in primary care settings. *Applied Clinical Informatics*, *15*(3), 544–555. https://doi.org/10.1055/a-2267-1727

Garvin, J. H., Yazzie, V. M., Katsuyama, N. A., Rudloff, T., Worthen-Chaudhari, L. C., & Chaudhari, A. M. (2025). Design of a clinical balance tool for fall risk assessments: A development and usability study. *PLoS One*, *20*(2), e0302080. Article e0302080. https://doi.org/10.1371/journal.pone.0302080

James, S. L., Lucchesi, L. R., Bisignano, C., Castle, C. D., Dingels, Z. V., Fox, J. T., Hamilton, E. B., Henry, N. J., Krohn, K. J., McCracken, D., Nixon, M. R., Roberts, N. L. S., Sylte, D. O., Adsuar, J. C., Arora, A., Briggs, A. M., Collado-Mateo, D., Cooper, C., & Liu, Z. (2020). The global burden of falls: Global, regional and national estimates of morbidity and mortality from the global burden of disease study 2017. *Injury Prevention*, *26*(Suppl 2), i3–i11. https://doi.org/10.1136/injuryprev-2019-043286

Jin, H., Zhou, Y., Stagg, B. C., & Ehrlich, J. R. (2024). Association between vision impairment and increased prevalence of falls in older US adults. *Journal of the American Geriatrics Society*, *72*(5), 1373–1383. https://doi.org/10.1111/jgs.18879

Jin, Y., Kim, H., Jin, T., & Lee, S.-M. (2021). Automated fall and pressure injury risk assessment systems: Nurses' experiences, perspectives, and lessons learned. *CIN: Computers, Informatics, Nursing*, *39*(6), 321–328. https://doi.org/10.1097/CIN.0000000000000696

Kakara, R., Bergen, G., Burns, E., & Stevens, M. (2023). Nonfatal and fatal falls among adults aged≥ 65 years—United States, 2020–2021. *Morbidity and Mortality Weekly Report*, *72*(35), 938–943. https://doi.org/10.15585/mmwr.mm7235a1

Krabbe, P. F. (2008). Thurstone scaling as a measurement method to quantify subjective health outcomes. *Medical Care*, *46*(4), 357–365. https://doi.org/10.1097/MLR.0b013e31815ceca9

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, *19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

National Institute for Health and Care Excellence. (2013). Falls in older people: Assessing risk and prevention (clinical guideline CG161). https://www.nice.org.uk/guidance/cg161

National Institute for Health and Care Excellence. (2025). Falls: Assessment and prevention in older people and in people 50 and over at higher risk. NICE guideline (NG249). https://www.nice.org.uk/guidance/ng249

Park, S.-H. (2018). Tools for assessing fall risk in the elderly: A systematic review and meta-analysis. *Aging-Clinical & Experimental Research*, *30*(1), 1–16. https://doi.org/10.1007/s40520-017-0749-0

Ploegmakers, K. J., Linn, A. J., Medlock, S., Seppälä, L. J., Bahat, G., Caballero-Mora, M. A., Ilhan, B., Landi, F., Masud, T., Morrissey, Y., Ryg, J., Topinkova, E., van der Velde, N., van Weert, J. C. M., & European Geriatric Medicine Society EuGMS Task and Finish group on Fall-Risk Increasing Drugs FRIDs. (2024). A European survey of older peoples'

preferences, and perceived barriers and facilitators to inform development of a medication-related fall-prevention patient portal. *European Geriatric Medicine*, *15*(3), 817–829. https://doi.org/10.1007/s41999-024-00951-w

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Sun, R., Aldunate, R. G., Paramathayalan, V. R., Ratnam, R., Jain, S., Morrow, D. G., & Sosnoff, J. J. (2019). Preliminary evaluation of a self-guided fall risk assessment tool for older adults. *Archives of Gerontology and Geriatrics*, *82*, 94–99. https://doi.org/10.1016/j.archger.2019.01.022

Wang, J., Li, Y., Yang, G.-Y., & Jin, K. (2025). Age-related dysfunction in balance: A comprehensive review of causes, consequences, and interventions. *Aging and Disease*, *2024*(2), 0124–2021. https://doi.org/10.14336/AD.2024.0124-1

World Health Organization. (2021). *Falls*. https://www.who.int/news-room/fact-sheets/detail/falls

Yamada, Y., Nakashima, H., Nagae, M., Watanabe, K., Fujisawa, C., Komiya, H., Tajima, T., Sakai, T., Satake, S., Takeya, Y., Umeda-Kameyama, Y., & Umegaki, H. (2024). Dual sensory impairment predicts an increased risk of postdischarge falls in older patients. *Journal of the American Medical Directors Association*, *25*(9), 105123. Article 105123. https://doi.org/10.1016/j.jamda.2024.105123