



HAL
open science

Development and psychometric properties of the Balance in Daily Life (BDL) scale in a population of frail older people

Claire de Labachellerie, Emilie Viollet, Sandrine Alonso, Célia Dauvergne, Mylène Blot, Fabrice Nouvel, Willy Fagart, Thierry Chevallier, Anthony Gelis, Arnaud Dupeyron

► To cite this version:

Claire de Labachellerie, Emilie Viollet, Sandrine Alonso, Célia Dauvergne, Mylène Blot, et al.. Development and psychometric properties of the Balance in Daily Life (BDL) scale in a population of frail older people. *Maturitas*, 2024, 187, pp.108064. 10.1016/j.maturitas.2024.108064 . hal-04670558

HAL Id: hal-04670558

<https://imt-mines-ales.hal.science/hal-04670558v1>

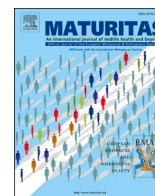
Submitted on 12 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Original article

Development and psychometric properties of the Balance in Daily Life (BDL) scale in a population of frail older people

Claire De Labachellerie^{a,*}, Emilie Viollet^a, Sandrine Alonso^b, Célia Dauvergne^a, Mylène Blot^a, Fabrice Nouvel^a, Willy Fagart^a, Thierry Chevallier^b, Anthony Gelis^{c,d}, Arnaud Dupeyron^{a,e}

^a Centre of Medical Device Evaluation – Handicap (CEDM-H), CHU Nîmes, Univ Montpellier, Nîmes, 4 rue du professeur Robert Debré, 30900 Nîmes, France

^b Department of Biostatistics, Epidemiology, Public Health and Innovation in Methodology (BESPIM), CHU Nîmes, Univ Montpellier, 4 rue du professeur Robert Debré, 30900 Nîmes, France

^c Centre Neurologique Mutualiste Propara, 263 rue du caducée, 34090 Montpellier, France

^d Epsylon Laboratory, 2033 avenue Bouisson Bertrand, 34090 Montpellier, France

^e M2H Laboratory, Euromov Digital Health in Motion, 700 avenue du Pic Saint-Loup, 34090 Montpellier, France



ARTICLE INFO

Keywords:

Balance in Daily Life

Elderly

Psychometric properties

Équilibre et Vie quotidienne

EVQ

ABSTRACT

Background: Balance disorders in older people cause falls, which can have serious functional and economic consequences. No existing scale relates fall risk to daily life situations. This study describes the development, psychometric properties and construct validity of the Balance in Daily Life (BDL) scale, comprising seven routine tasks including answering a phone, carrying a heavy bag, and sitting down and getting up from a chair.

Methods: Frail patients aged 65 years or more were prospectively recruited from the geriatric rehabilitation department of a French university hospital. Inclusion criteria included autonomous walking over 20 m and modified Short Emergency Geriatric Assessment score 8–11. Patients with motor skills disorders and comprehension or major memory difficulties were excluded. Patients were assessed on Day 3 and Day 30 with the Balance in Daily Life scale, Timed Up and Go, one-leg stance time, sternal nudge and walking-while-talking tests. The scale was assessed for acceptability, quality, unidimensionality, internal consistency, reliability, temporal stability, responsiveness and construct validity.

Results: 140 patients (83 ± 6 years) were recruited, of whom 139 were assessed at Day 0 and 133 at Day 30. Acceptability was satisfactory (134/139 patients completed the test), quality assessment showed a slight floor effect (6 % of patients with minimal score) and evaluation of item redundancy found no strong correlation (Spearman < 0.7). Unidimensionality was verified (Loevinger H coefficient > 0.5 for all items except item 6 = 0.4728). Internal consistency was good (Cronbach alpha = 0.86). Reliability and temporal stability were excellent (ICC = 0.97 and ICC = 0.92). Responsiveness was verified by significant score change $p < 0.0001$ between Day 0 and Day 30 (decreased by 1 [0; 2] point), in line with other score changes. Construct validity revealed that the Balance in Daily Life scale was convergent with results of the timed up-and-go and one-leg stance time ($p < 0.0001$ for both) and tended to be higher for participants who had not fallen in the previous 6 months ($p = 0.0528$). The new questionnaire was divergent to sternal nudge tests ($p = 0.0002$) and not related to the walking-while-talking test ($p = 0.5969$).

Conclusion: The Balance in Daily Life scale has good psychometric properties for this population. Its simplicity and innovative nature mean that it can be applied in institutions while being easily modifiable to domestic settings.

Study registration on clinicaltrials.gov: NCT0334382.

* Corresponding author at: CEDM-H (Centre d'Evaluation du Dispositif Médical-Handicap), CHU Nîmes Carémeau, 4 rue du Professeur Robert Debré, 30900 Nîmes, France.

E-mail address: cedm-handicap@chu-nimes.fr (C. De Labachellerie).

<https://doi.org/10.1016/j.maturitas.2024.108064>

Received 22 October 2023; Received in revised form 31 May 2024; Accepted 5 July 2024

Available online 8 July 2024

0378-5122/© 2024 CHU Nîmes. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Approximately one-third of people over the age of 65 fall at least once a year [1] and half fall repeatedly [2]. The short- and long-term consequences include fractures, chronic pain, functional deficit, disability, loss of autonomy, loss of confidence with fear of recurrence, depression, social withdrawal, institutionalisation and death [3–5]. Falls cost around \$64.4 billion annually in the USA [6]. The number of falls and their repercussions tend to increase with advancing age and thus with the aging world population [7]. Risk factors include acute or chronic pathology, frailty, history of falls, use of sedative drugs, muscle weakness, balance deficit, walking disorders, visual deficit and cognitive disorders [2,8–12]. However, preventing falls and their consequences is complex because they are multifactorial events involving multiple intrinsic risk factors plus environmental or contextual factors [13].

One of the first steps in assessing the risk of falling is to evaluate gait and balance [14], but routinely used balance assessment tests have limited ability to predict falls [15–18]. Current physical performance-based tools such as alternate step test, adjusted maximum step length, Berg Balance Scale (BBS), Dynamic Gait Index, Functional Gait Assessment, five-times sit-to-stand test, Zur balance scale, getting up from lying on the floor, one-leg stance-time (OLST), Tinetti Performance-Oriented Mobility Assessment (POMA), risk assessment, stair ascent, test battery, timed gait, Timed Up and Go (TUG) test, and walking while talking (WWT) test underestimate the impact of the environment and attentional load in daily life situations where dual-task solicitations are common and unpredictable [19]. Motor capacities are altered during other tasks (cognitive, motor or visuo-spatial) [20,21] and the characteristics of walking in real-life situations are mimicked by dual-tasks in the laboratory [22]. The value of dual-task assessments as a predictor of falls is debated [23,24], but many authors consider them useful [25–30]. To our knowledge, there are no scales for evaluating balance in real-life situations. We believe that the current lack of a scale for assessing balance in real-life may explain the difficulty of current tests in predicting the risk of falling.

The French EVQ (Equilibre dans la Vie Quotidienne, in English: the Balance and Daily Life = BDL) scale was developed to reflect real life. The BDL is scored by observing and rating patients' compensation strategies to maintain balance in seven situations of daily life. It is used as a screening tool for falls, but also helps to orient rehabilitation to prevent fall recurrence. We have used the BDL in our centre for many years. The activities assessed involve voluntary gestures that monopolise the attention (bi-manual praxis, memory, auditory and visual attention, gnosis) and more automatic tasks (postural adaptation, motor patterns, walking, turning round, sitting down). These activities also demonstrate how the patient manages a walking aid.

Such tools need stringent validation before application in clinical practice. Psychometric validation of assessment tools consists of successively evaluating: acceptability, quality, saturation phenomena (floor or ceiling effect) and item redundancy (correlation between items), unidimensionality, internal consistency, reliability, temporal stability, responsiveness, and construct validity (convergence and/or divergence with other validated tools) [31,32]. The aim of this study was to describe development and psychometric properties of the BDL scale in frail older patients.

2. Methods

2.1. Scale development

The scale was developed in two steps: 1. item generation; and 2. theoretical analysis [33]. Items were generated by the geriatric rehabilitation service of the University Hospital Center of Nîmes. The first version was gradually refined from 2005 to 2007, with several professionals, mainly occupational therapists and doctors in physical rehabilitation medicine, providing feedback to establish the items and

scoring system. The scale was disseminated during presentations to geriatric professionals at conferences and published in the journal of the French Association of Occupational Therapists in Geriatrics [34]. The content of the items and their scoring further evolved with informal feedback from professionals using the scale in their current practice (in France, Quebec, Switzerland, Belgium), feedback from patients who had taken the scale, and professionals in the sector. Items such as making the bed were removed (difficult to put in place, not necessarily done by the target population), the item carrying a light load (a pillow) was replaced with a heavier load to be more relevant (requiring greater postural adaptations and object masking vision of the floor) and because patients described fatigue when carrying out tasks. Feedback from professionals led to the production of a test booklet to standardise the test conditions. This first stage was not standardised, but refinement was continued until feedback was consistently positive.

In 2007, we used a standardised questionnaire to evaluate content validity: relevance and usefulness, methodology, opinion of the test booklet and general opinion of the scale. This questionnaire was sent to professionals in the geriatric sector (occupational therapists, physiotherapists, physical rehabilitation and geriatric doctors). We analysed the answers for a final revision, in particular with regard to the scoring and verbatim used in the questions. The assessment has not changed since 2007 and we use it daily in the hospital. In this study, we evaluated the scale psychometric properties in order to disseminate it more widely.

2.2. Trial design

This study was designed to describe psychometric properties of the BDL scale in frail older patients, following the COSMIN checklist [31]. The study was approved by the local ethical committee (CPP Sud Méditerranée III, #2021.06.05) and registered on clinicaltrials.gov (#NCT03343821).

An occupational therapist (OT1) performed the initial evaluation on Day 0, including a standard balance assessment: Timed Up and Go (TUG) test, one-leg stance time (OLST), sternal nudge test, walking-while-talking test (WWT), and the BDL assessment. A 5-minute resting period was observed between the balance tests and the BDL. A second occupational therapist (OT2) independently and simultaneously scored the BDL to test inter-rater reliability. OT1 and OT2 were blinded to each other's ratings. Intra-rater reliability was assessed through a second evaluation on Day 3 (± 2 days) by OT1 alone (test-retest method). Upon discharge on Day 30 (± 7 days), OT1 repeated the assessments from inclusion, including the third BDL assessment.

2.3. Participants

Patients were consecutively included upon arrival at Nîmes University Hospital's department of Physical Medicine and Geriatric Functional Rehabilitation. Inclusion criteria were age ≥ 65 years and frailty (Short Emergency Geriatric Assessment (SEGA)) score 8–11 [35]. Patients had to be able to walk 20 m independently (with or without a walking aid) with full weight on both feet. Exclusion criteria were: major praxis disorders preventing task completion, failure to understand simple commands and major memory disorders. All patients provided written consent prior to inclusion. During hospitalisation, patients received standard physiotherapy to restore their independence and mobility to enable return home and limit fall risk.

2.4. BDL assessment (Appendix 1)

Participants performed seven routine domestic tasks such as answering a ringing phone, carrying a 2 kg shopping bag for 10 m, opening and closing a door to enter and exit a room, setting the table for four people, moving a chair 2 m to the table, sitting down and getting up from a chair without armrest and carrying a small pan of water to the table. These tests require postural adaptations and multiple task

management (praxis, memory, attention), thus highlighting the strategies used. Each task is scored as 1 = Normal (no instability), 2 = Appropriate (the patient uses safety strategies to perform the task, adaptation by anticipation, no instability), 3 = Abnormal (the patient uses reflex adaptation strategies, adaptation by correction, fall-avoidance strategies, minor instability) and 4 = Very abnormal (ineffective postural adaptation reflex, major instability). In case of inability to perform the task for a reason other than balance disorders, no score was assigned to the task.

2.5. Scoring system of BDL

The calculation is made if at least half of the BDL items have been completed according to the formula: score = sum of test responses \times [7 / (number of tests completed and scored)]. Because of the frequent health problems (hearing impairment, osteoarthritis, etc.) among frail older people, the proposed scoring system allows for the calculation of an overall score if at least four of the seven proposed tests have been performed.

2.6. Other test assessments

ADL (activities in Daily Life) and IADL (Instrumental Activities In Daily Life) questionnaires were administered to assess patients' functional abilities [36]. For construct validity assessment, we used tests recommended by French health authority to evaluate balance and fall risk in autonomous older people: Timed Up and Go test (TUG) (considered normal if <20 s), One Leg Stance Time (OLST), Sternal nudge test and Walking While Talking test (WWT) [37]. In the TUG test, participants performed three successive tests preceded by a non-scored learning test. Tasks were: getting up from an ordinary chair, walking about 3 m, turning around and coming back to sit down without the help of another person (with or without a mobility device). This test is reliable (ICC = 0.99) [38]. A mobility deficit is considered to exist from 20 s upwards, with a significant deficit above 29 s [38]. In OLST, patients balanced on one leg. The test was considered abnormal when the participant held the position for <5 s. This test is reliable and has been described as a predictor of injurious falls [39,40]. The threshold of 5 s was chosen in accordance with the Tinetti Balance Subscale [41]. The sternal nudge test assesses the response of the participant in a bipodal standing position to a sternal nudge. A normal response to this test consists of a slight sway of the body only; an abnormal response included stepping back, staggering, falling, or a 'startle' response of the upper limbs [42]. The WWT consisted of asking a question "what do you think of..." while the patient was walking. The test is impaired when patients stop walking to answer a question. This test is associated with the risk of falling within 6 months (specificity = 95 %, sensitivity = 48 %, positive predictive value = 83 % and negative predictive value = 76 %) [20].

2.7. BDL validation process

The validation process consisted of the assessment of: (1) global and individual acceptability of tests; (2) global and individual quality (absence of saturation, ceiling or floor effects and redundancy) of tests; (3) unidimensionality; (4) internal consistency; (5) reproducibility of scoring and temporal stability (repeatability); (6) responsiveness; (7) construct validity to assess convergence and divergence [31,32].

2.8. Statistical analysis

A sample size of 140 participants was calculated with a priori participant-to-item ratio method [43]. According to the 28 rating possibilities (four possible answers for seven items) using a ratio of five participants for each variable.

Psychometric properties were assessed following current recommendations [31,32]. Acceptability was assessed by number and

proportion of completed tests. Acceptable tests must have a non-completion rate < 5 %; a rate > 10 % means the test is disputable [44]. Test quality was assessed by the absence of saturation, ceiling or floor effects and redundancy between items using the non-parametric Spearman rank correlation coefficient (negligible: < 0.3 , poor: 0.3–0.49, moderate: 0.5–0.69, good: 0.7–0.89, excellent: >0.9) [45]. Scale unidimensionality was assessed using factor analysis and the Mokken scale procedure (MSP) [46]. This procedure automatically partitions the tests into one or more sets by defining the dimensions and possibly a set of un-scalable tests using Loevinger's H coefficients. The properties were considered strong when $H \geq 0.5$, medium when $H \geq 0.4$ and weak otherwise. Internal consistency was estimated using Cronbach's α coefficient 95 % CI [47]. A score > 0.7 was considered reliable [48].

Reproducibility of measurements made by two observers at Day 0 and temporal stability or reliability between Day 0 and Day 3 were assessed using the intraclass correlation coefficient [49] (ICC; 95 % CI) for global score (poor: <0.5 , moderate: 0.5–0.75, good: 0.75–0.9, excellent: > 0.9) [49] and weighted kappa coefficient (poor: <0 , slight: 0–0.2, fair: 0.21–0.4, moderate: 0.41–0.6, substantial: 0.61–0.8, almost perfect: 0.81–1) [50] (K; 95 % CI) for each test. The Bland-Altman graph method [51] was used to evaluate the existence of bias. The symmetry McNemar test was used to assess discordance between two tests.

Responsiveness was assessed between Day 0 and 30 using the Wilcoxon signed rank test, McNemar test and non-parametric Spearman rank correlation. Finally, we assessed convergence and divergence compared to other tools using the non-parametric Spearman rank correlation coefficient and Wilcoxon test as a measure of construct validity. All analyses were performed using SAS software, version 9.4 (SAS Institute Inc., Cary, NC).

3. Results

140 patients (83 ± 6 years old, BMI: 25 ± 5 kg/m², women = 94 (67 %)) were recruited from 26/01/2018 to 18/06/2019 (Table 1). At "Day 3", 138 patients were evaluated within 3 [2; 4] days post-inclusion and at "Day 30", 133 patients were evaluated within 20 [14; 28] days. The Day 30 test was carried out earlier in all cases because the patients were discharged prior to day 30, and it would have been onerous to bring them back. The flow chart is presented in Fig. 1.

3.1. Acceptability

All patients completed all items of the BDL except for four patients who did not react to the phone ringing (due to hearing and/or attentional problems) and one individual who had difficulty carrying the shopping bag due to pain from thumb basal joint arthritis.

3.2. Quality

The distribution of responses to each item is presented in Table 2. A slight floor effect was observed in the global score with 6 % of individuals having a minimum score, although this did not correspond to instability of any single item. Individual items demonstrated variable floor effects ranging from 10 % for Q6 (sitting down and getting up from a chair) to 21 % for Q1 (answering a ringing phone). All items showed moderate redundancy except for Q6 (sitting down and getting up from a chair) which was weakly or not correlated to the other items. No correlations were strong or very strong.

3.3. Unidimensionality

One unique dimension was defined by factor analysis and confirmed by the MSP. The Loevinger's H coefficient was good (0.55). All tests showed good properties except Q6 (sitting and standing), which was of average quality ($H = 0.47$). Q6 had the lowest floor and highest ceiling

Table 1

Characteristics of the study population at inclusion (Day 0). Data shown as number (%) or median [IQR] and mean (SD). * missing data for 35 patients, † missing data for 73 patients.

	n (%)
SEGA score	
9	42 (40 %)
10	26 (25 %)
11	37 (35 %)
ADL score	4.5 (4; 5)
	4.5 (1)
IADL score [†]	3 (2; 4)
	3 (1)
Technical assistance with walking	
None	22 (16 %)
Walking stick	28 (20 %)
Rollator	84 (60 %)
Walking stick and rollator	6 (4 %)
Fall within the previous 6 months	85 (61 %)
Number of falls in the previous 6 months	1 (0; 2)
	1 (2)
Living place	
Home without caregivers	48 (34 %)
Home with caregivers	90 (64 %)
Institution	2 (2 %)
BDL score	14 (14; 15)
	14 (3)
Timed Up and Go time (seconds)	28 (17,5; 42,5)
	34 (25)
Normal (<20 s)	44 (31 %)
Significantly impaired (>29 s)	63 (45 %)
One-leg stance time (seconds)	1 (0; 3)
	2 (3)
Impaired (<5 s)	121 (86 %)
Walking-while-talking test: impaired	26 (19 %)
Sternal nudge test: impaired	60 (43 %)

SEGA: Short Emergency Geriatric Assessment, ADL: Activities of Daily Living; IADL: Instrumental Activities of Daily Living, BDL: Balance in Daily Life.

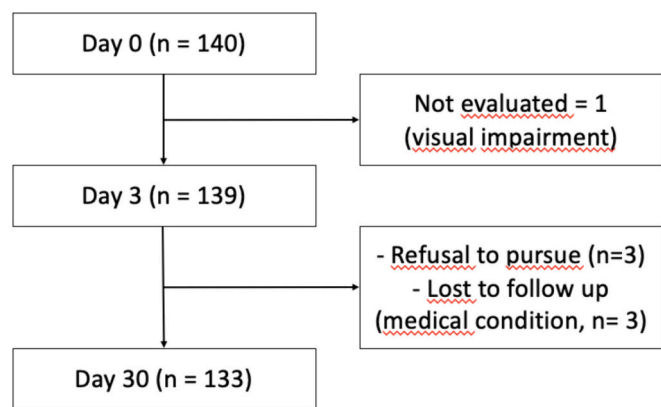


Fig. 1. Flowchart.

effect.

3.4. Internal consistency

The BDL seemed reliable with a Cronbach's alpha coefficient estimated at 0.86 [0.81; 0.89]. The step-by-step Cronbach α backward procedure confirmed the general agreement between tests measuring the same concept (latent trait) with a stable Cronbach's alpha coefficient estimate always exceeding the 0.7 threshold (from 0.83 to 0.86). The monotone increasing variation of Cronbach α when tests were successively removed confirmed the unidimensionality of BDL.

3.5. Reproducibility and temporal stability

Inter-rater reproducibility score on Day 0 was excellent (ICC = 0.97 [0.96; 0.98]) with a bias of 0.04 (0.68). The Bland and Altman graph did not show any systematic bias (Appendix 2). Individual item inter-rater reproducibility was good for items 1, 3, 4, 5, 6 and 7 and excellent for items 2. The McNemar symmetry test was not significant for any item.

Temporal stability, assessed 3 days [2; 4] apart, was good (ICC = 0.92 [0.89; 0.94]) with a bias of 0.28 (1.14). The Bland and Altman graph did not show systematic bias (Appendix 3). The item-by-item assessment was moderate for all items. McNemar's symmetry test was not significant for any item.

3.6. Responsiveness

Sensitivity to change (responsiveness) showed a statistically significant change in BDL score between Day 0 and Day 30 ($p < 0.0001$). The BDL score decreased by 1 [0; 2] point between the two assessments, whereas the TUG test decreased by 6 [2; 12] seconds, and the OLST increased by 1 [0; 4] second ($p < 0.0001$ for all tests). The sternal nudge and walking-while-talking tests did not change significantly over time (respectively $p = 0.0679$ and $p = 0.5485$).

No correlation between the change of the BDL score, the TUG and the OLST test was observed. As we found no correlation, we conducted a complementary analysis whereby the change of the different scores was classed as improvement, stability or worsening according to the threshold values in the literature (e.g. TUG: normal <20s and significant deficit >29 s or OLST: deficit if <5 s) (Table 3). The McNemar tests showed a dissymmetry in the allocation of patients to the different levels of change between the BDL and all other tests ($p < 0.0001$). The BDL test more frequently led to a more favourable outcome than the other tests.

3.7. Construct validity

The construct validity was estimated at Day 0. Spearman correlation coefficients were weak or non-existent between the BDL score and the number of falls in the previous 6 months (0.2072), TUG (0.3774) and OLST (-0.3754). BDL scores of patients who had fallen in the previous 6 months (14 [14; 16]) tended to be higher than for non-fallers (14 [13; 15]), although without reaching significance ($p = 0.0528$).

There was a difference in BDL score according to the deficit in the TUG test with median BDL scores estimated at 12.5 [9; 14] for individuals with a normal test (<20 s); 14 [14; 16] for those with a deficit (20–29 s); and 15 [14; 16] for those with a significant deficit (>29 s) ($p < 0.0001$). There was a difference in the BDL score according to the deficit detected by the OLST with median BDL scores estimated at 14 [14; 16] for individuals with a deficit (<5 s) and 10 [8; 14] for the others ($p < 0.0001$). For the sternal nudge test, there was a difference in score with a median estimated BDL score of 14.5 [14; 16] for individuals without imbalance and 14 [10; 15] for the others. There was no statistically significant difference between the BDL score and the results of the WWT test ($p = 0.5969$).

4. Discussion

This validation study demonstrated that the BDL scale is reliable with good psychometric properties in an older population at risk of falls. It is innovative, complements the existing assessments and is easily performed. Risk factors for falls are split into three domains: sensory-motor, cognitive and environmental. The sensory-motor domain is evaluated by motor strength tests such as the sit-to-stand test [52], tests evaluating balance such as strength platform measurement, or more functional tests such as the TUG test and the OLST. The cognitive domain is influenced by cognitive and behavioral disorders such as fear of falling, taking risks or drug interactions. This component is assessed by dual-task tests like the WWT test [25,28,53]. Finally, the environmental domain is difficult

Table 2
Results of the Balance in Daily Life (BDL) score on Day 0.

Item	N (/140)	Mean \pm standard deviation	Distribution of responses			
			1 Normal	2 Appropriate	3 Abnormal	4 Very abnormal
Q1: Go and answer the phone when it rings	136	1.88 \pm 0.55	21 %	71 %	7 %	1 %
Q2: Carry a shopping bag	139	2.04 \pm 0.62	14 %	71 %	12 %	3 %
Q3: Enter and exit a room	140	1.93 \pm 0.58	19 %	72 %	7 %	2 %
Q4: Set the table for four	140	2.02 \pm 0.56	12 %	76 %	10 %	2 %
Q5: Move a chair	140	1.86 \pm 0.51	19 %	72 %	3 %	1 %
Q6: Sit down and get up from a chair	140	2.24 \pm 0.70	10 %	61 %	24 %	5 %
Q7: Bring a pan of water to the table	140	1.97 \pm 0.56	13 %	79 %	4 %	4 %
Item	N (/140)	Median [q1; q3]	Min	Max	% floor	% ceiling
BDL score (/28)	140	14 [14; 15]	7	21	6 %	0 %

Table 3
Change of each test from Day 0 to Day 30.

N = 133	Improvement	Stability	Worsening
Timed Up and Go test	45 (34 %)	83 (62 %)	5 (4 %)
One-leg stance time	14 (11 %)	117 (88 %)	2 (2 %)
Sternal nudge test	10 (8 %)	103 (77 %)	20 (15 %)
Walking-while-talking test	14 (11 %)	108 (81 %)	11 (8 %)
BDL	71 (53 %)	42 (32 %)	20 (15 %)

to evaluate, lacking validated evaluation scales. The BDL incorporates all three domains, evaluating balance (sensory-motor domain) in a dual-task context (cognitive domain), within daily life scenarios (environmental domain). We believe that getting patients to perform standardised daily living activities with a low cognitive load, independent of social-cultural factors, is an important asset of this evaluation. The postural reactions induced by the environmental context would likely be more marked if evaluated at home.

While the procedure for developing the scale did not comply with recommended standards [33], we nevertheless followed the two stages classically described: generation of the items from 2005 to 2007 and collecting feedback as soon as possible. Circulation of the questionnaire through the French Association of Occupational Therapists in Geriatrics enabled us to obtain feedback from professionals in the sector concerned. The theoretical analysis was done through a standardised questionnaire sent to professionals in the sector and a qualitative analysis of their feedback.

Psychometric properties were promising; acceptability was good, with only 5/140 (3.5 %) of patients unable to perform 1/7 of the test items. In comparison, the inability to perform the OLST and the TUG test reached 46.9 % and 4.7 % respectively in patients using a walking aid [40], while 84 % of our population used a walking aid. The possibility of carrying out this test with a walking aid is beneficial in the older fragile population. With regard to quality, the slight floor effect observed can also be ascribed to the population characteristics (60 % used a rollator, 64 % lived with caregivers, average IADL score of 3, age 83 \pm 6 years, etc.). It would be interesting to see whether the same ceiling effects present in numerous tests for assessing balance in older people would be present in community-dwelling older adults [54]. Inter-rater reliability and temporal stability were comparable to those described in other balance assessment tests in older people [40]. Regarding responsiveness, we found a significant sensitivity to change, with a gain of one point on the BDL after one month of rehabilitation (median score of 14 in the assessment at D1 and D30, but significant non-parametric Wilcoxon test). The BDL score evolved in the same direction as the other balance assessment scales regularly used in geriatric medicine, with no correlation found. Given the lack of correlation, we conducted a complementary analysis by classifying the data into categories rather than continuous values. The asymmetric distribution of BDL scores, more frequently showing a favourable outcome, would be interesting to explore further. This may be due to the greater discriminatory effect of

the BDL score, which shows more subtle improvements in balance than the other tests evaluated. The aim of our scale is to assess balance in daily life, which no other scale currently does, making it difficult to compare scales. The lack of correlation between the BDL and the other tests or between the change of the BDL and the change of the other scales confirms the different objective of the BDL. The limitations of available assessments in predicting falls have been demonstrated [55]. The current recommendations are to use at least two different tools [18]. The cross-domain element potentially makes the BDL workable as a single evaluation, but this requires validation.

Construct validity showed the same trend, with a link between the BDL score and the TUG test and the OLST, but no link was found with the WWT test and the sternal nudge test. After 1 month of rehabilitation, the sternal nudge and the WWT test did not improve. This could be explained by the greater cognitive impact of these tests [56], but we did not carry out cognitive tests on our population to test this hypothesis. The psychometric qualities of our scale appear to be similar to those of the TUG and OLST [40], both in terms of construct quality and responsiveness, but the lack of correlation found suggests that it does not measure the same aspects. The BDL scale is unique in measuring balance in daily life, possibly explaining these small differences. However, the BDL scale does not seem to measure the same dimensions as the WWT tests and the sternal nudge. Looking at falls over the last 6 months, the BDL score was slightly higher for patients experiencing a fall, but the *p*-value of 0.0528 did not cross the significance threshold, although there does seem to be a trend. The difficulty of judging this type of correlation has already been described, as the assessments are made retrospectively, often some time after the fall. [40]. The patient's condition may have changed considerably, particularly if a medical event leading to hospitalisation precipitated the fall. Moreover, widely used tests such as the TUG test have been criticised for failing to differentiate between fallers and non-fallers [17].

While balance assessments in older patients mainly try to predict falls, we argue that the more pertinent aim ought to be preventing falls. Medical management combines optimisation of drug treatment, nutrition, balance work and rehabilitation. The BDL assessment performed by occupational therapists could target patients requiring complementary work. Indeed, it is essential to differentiate between an "appropriate" rating i.e. patients using effective compensatory strategies by anticipation without modifying their balance, and an "abnormal" rating i.e. those using compensatory strategies by correction and whose balance is beginning to deteriorate. Patients apply different compensatory strategies, for example, decomposition of the half-turn, support from one/both hands when lowering and/or standing up, and increasing the support polygon by spreading the feet. The aim of this rehabilitation approach is to develop compensatory techniques pro-actively rather than reactively, once it has become too late. Patients rated "abnormal" become priority targets for this rehabilitation. Ineffective balance compensation strategies can be optimised by reinforcement until they become reflexive and/or anticipatory. One month of rehabilitation oriented in this way improves the BDL score as well as other balance scores

(TUG and OLST). However, we still need to study whether this approach and this assessment reflect the risk of falling in a prospective study.

This questionnaire could also be used at home. The original French version is already used by occupational therapists and is listed by the French health authorities (HAS) as a complementary screening tool for falls, even prior to validation [57].

5. Limitations

This study has limitations. Firstly, it was conducted in the department where the scale was developed and has been used for many years. Thus, the therapists have training in administering the scale. However, we have already disseminated it outside our department and informal feedback is that, in teams experienced in balance disorders in geriatric settings, the scoring is relatively intuitive. Nevertheless, an external assessment would be beneficial.

It should also be noted that the target population (frail older people) is relatively limited and this scale is unlikely to be useful in other populations, particularly more active older people. In addition, we did not carry out a quantified assessment of cognitive impairment. Only a general assessment was recorded in the modified SEGA score, which did not allow further analysis of cognitive functions and the BDL score.

The degree of impairment is indicated by a number from 1 to 4, but there is currently no global score or threshold value. It has not yet been studied quantitatively.

6. Future research

The BDL change between Day 0 and 30 was mostly classed as favourable. It is not clear whether these changes reflect the functional rehabilitation care given or if they are the result of a learning curve. The effect of learning from BDL was not evaluated and could be the subject of further research. The next step to validating the BDL is via a prospective evaluation linking BDL score and falls. Furthermore, we need to test the value of this tool in guiding rehabilitation intervention or home adaptation.

7. Conclusion

The BDL scale demonstrated good temporal stability, acceptability, quality, unidimensionality, internal consistency, responsiveness and construct validity in a population of frail older patients. This scale is unique in its integrated assessment of balance in daily life and is easy to use. Following validation, it can be added to the scales already available for evaluating balance of older patients at risk of falling.

Ethical statement

The study was approved by the local ethical committee (CPP Sud Méditerranée III, #2021.06.05) and registered on clinicaltrials.gov (#NCT0334382).

All patients provided written consent prior to inclusion.

Contributors

Claire de Labachellerie contributed to interpreting the data, and wrote and revised the manuscript.

Emilie Viollet contributed to study concept and design, and interpreting the data.

Sandrine Alonso contributed to study concept and design, composed the statistical dataset, and performed the analyses.

Célia Dauvergne contributed to study concept and design, and interpreting the data.

Myène Blot contributed to interpreting the data and critical revision of the manuscript.

Fabrice Nouvel contributed to study concept and design, and

interpreting the data.

Willy Fagart contributed to interpreting the data and critical revision of the manuscript.

Thierry Chevallier contributed to study concept and design.

Anthony Gelis contributed to study concept and design, interpreting the data and critical revision of the manuscript.

Arnaud Dupeyron contributed to study concept and design, interpreting the data and critical revision of the manuscript.

All authors saw and approved the final version and no other person made a substantial contribution to the paper.

Funding

This work was supported by an internal grant by the University Hospital of Nîmes.

Ethical approval

The study was approved by the local ethical committee (CPP Sud Méditerranée III, #2021.06.05). All patients provided written consent prior to inclusion.

Provenance and peer review

This article was not commissioned and was externally peer reviewed.

Data sharing and collaboration

There are no linked research data sets for this paper. Data will be made available on request.

Declaration of competing interest

The authors declare that they have no competing interest.

Acknowledgements

We wish to thank Pierre Rataboul for help in drafting the protocol, Teresa Sawyers for English translation of this article, Sarah Kabani for editing assistance, Brigitte Lafont for project management and Sarah el Sherif and Noemie Coneau for collecting data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.maturitas.2024.108064>.

References

- [1] J. Craig, et al., The high cost to health and social care of managing falls in older adults living in the community in Scotland, *Scott Med J* 58 (4) (2013) 198–203.
- [2] D.A. Jehu, et al., Risk factors for recurrent falls in older adults: A systematic review with meta-analysis, *Maturitas* 144 (2021) 23–28.
- [3] P. Kannus, et al., Fall-induced injuries and deaths among older adults, *JAMA* 281 (20) (1999) 1895–1899.
- [4] V.S. Stel, et al., Consequences of falling in older men and women and risk factors for health service use and functional decline, *Age Ageing* 33 (1) (2004) 58–65.
- [5] P. Kannus, et al., Prevention of falls and consequent injuries in elderly people, *Lancet* 366 (9500) (2005) 1885–1893.
- [6] J.C. Davis, et al., International comparison of cost of falls in older adults living in the community: a systematic review, *Osteoporos Int* 21 (8) (2010) 1295–1306.
- [7] P. Kannus, et al., Fall-induced deaths among elderly people, *Am J Public Health* 95 (3) (2005) 422–424.
- [8] M.E. Tinetti, M. Speechley, S.F. Ginter, Risk factors for falls among elderly persons living in the community, *N Engl J Med* 319 (26) (1988) 1701–1707.
- [9] Rubenstein, L.Z., Falls in older people: epidemiology, risk factors and strategies for prevention. *Age Ageing*, 2006. 35 Suppl 2: p. ii37-ii41.
- [10] S.W. Muir, K. Gopaul, M.M. Montero Odasso, The role of cognitive impairment in fall risk among older adults: a systematic review and meta-analysis, *Age Ageing* 41 (3) (2012) 299–308.

- [11] G. Kojima, Frailty as a Predictor of Future Falls Among Community-Dwelling Older People: A Systematic Review and Meta-Analysis, *J Am Med Dir Assoc* 16 (12) (2015) 1027–1033.
- [12] M.M. Lusardi, et al., Determining Risk of Falls in Community Dwelling Older Adults: A Systematic Review and Meta-analysis Using Posttest Probability, *J Geriatr Phys Ther* 40 (1) (2017) 1–36.
- [13] R.T. Morello, et al., Multifactorial falls prevention programmes for older adults presenting to the emergency department with a fall: systematic review and meta-analysis, *Inj Prev* 25 (6) (2019) 557–564.
- [14] M. Montero-Odasso, et al., World guidelines for falls prevention and management for older adults: a global initiative, *Age Ageing* 51 (9) (2022).
- [15] O. Beauchet, et al., Timed Up and Go test and risk of falls in older adults: a systematic review, *J. Nutr. Health Aging* 15 (10) (2011) 933–938.
- [16] A.F. Ambrose, G. Paul, J.M. Hausdorff, Risk factors for falls among older adults: a review of the literature, *Maturitas* 75 (1) (2013) 51–61.
- [17] D. Schoene, et al., Discriminative ability and predictive validity of the timed up and go test in identifying older people who fall: systematic review and meta-analysis, *J Am Geriatr Soc* 61 (2) (2013) 202–208.
- [18] S.H. Park, Tools for assessing fall risk in the elderly: a systematic review and meta-analysis, *Aging Clin Exp Res* 30 (1) (2018) 1–16.
- [19] Waterval, N.F.J., et al., Predictability of Fall Risk Assessments in Community-Dwelling Older Adults: A Scoping Review. *Sensors* (Basel), 2023. 23(18).
- [20] Lundin-Olsson, L., L. Nyberg, and Y. Gustafson, “Stops walking when talking” as a predictor of falls in elderly people. *Lancet*, 1997. 349(9052): p. 617.
- [21] B. Wollesen, et al., A taxonomy of cognitive tasks to evaluate cognitive-motor interference on spatiotemporal gait parameters in older people: a systematic review and meta-analysis, *Eur Rev. Aging Phys Act* 16 (2019) 12.
- [22] I. Hillel, et al., Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating the gaps between gait performance in the lab and during 24/7 monitoring, *Eur Rev. Aging Phys Act* 16 (2019) 6.
- [23] A. Zijlstra, et al., Do dual tasks have an added value over single tasks for balance assessment in fall prevention programs? A mini-review. *Gerontology* 54 (1) (2008) 40–49.
- [24] J.C. Menant, et al., Single and dual task tests of gait speed are equivalent in the prediction of falls in older people: a systematic review and meta-analysis, *Ageing Res Rev.* 16 (2014) 83–104.
- [25] J. Verghese, et al., Validity of divided attention tasks in predicting falls in older individuals: a preliminary study, *J Am Geriatr Soc* 50 (9) (2002) 1572–1576.
- [26] A. Bergland, T.B. Wyller, Risk factors for serious fall related injury in elderly women living at home, *Inj Prev* 10 (5) (2004) 308–313.
- [27] O. Beauchet, et al., Stops walking when talking: a predictor of falls in older adults? *Eur J Neurol* 16 (7) (2009) 786–795.
- [28] E.I. Ayers, et al., Walking while talking and falls in aging, *Gerontology* 60 (2) (2014) 108–113.
- [29] S.W. Muir-Hunter, J.E. Wittwer, Dual-task testing to predict falls in community-dwelling older adults: a systematic review, *Physiotherapy* 102 (1) (2016) 29–40.
- [30] E. Smith, T. Cusack, C. Blake, The effect of a dual task on gait speed in community dwelling older adults: A systematic review and meta-analysis, *Gait Posture* 44 (2016) 250–258.
- [31] L.B. Mokkink, et al., The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content, *BMC Med Res Methodol* 10 (2010) 22.
- [32] K. Swan, et al., Measuring what matters in healthcare: a practical guide to psychometric principles and instrument development, *Front Psychol* 14 (2023) 1225850.
- [33] F.F.R. Morgado, et al., Scale development: ten main limitations and recommendations to improve future research practices, *Psicol Reflex Crit* 30 (1) (2017) 3.
- [34] M. Blot, A. Dupeyron, Evaluation de l'équilibre en situation d'activités de la vie journalière: le bilan EVQ, *Revue de l'Association Française des Ergothérapeutes en Gériatrie* 9 (2007).
- [35] N. Oubaya, et al., Screening for frailty in elderly subjects living at home: validation of the Modified Short Emergency Geriatric Assessment (SEGAm) instrument, *J Nutr Health Aging* 18 (8) (2014) 757–764.
- [36] C. Benaim, et al., The assessment of autonomy in elderly people, *Ann Readapt Med Phys* 48 (6) (2005) 336–340.
- [37] H.A.S., Réponse à la saisine du 3 juillet 2012 en application de l'article L.161–39 du code de la sécurité sociale 2012, Haute Autorité de la Santé.
- [38] D. Podsiadlo, S. Richardson, The timed “Up & Go”: a test of basic functional mobility for frail elderly persons, *J Am Geriatr Soc* 39 (2) (1991) 142–148.
- [39] B.J. Vellas, et al., One-leg balance is an important predictor of injurious falls in older persons, *J Am Geriatr Soc* 45 (6) (1997) 735–738.
- [40] M.R. Lin, et al., Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people, *J Am Geriatr Soc* 52 (8) (2004) 1343–1348.
- [41] M.E. Tinetti, Performance-oriented assessment of mobility problems in elderly patients, *J Am Geriatr Soc* 34 (2) (1986) 119–126.
- [42] D. Wild, U.S. Nayak, B. Isaacs, Prognosis of falls in old people at home, *J Epidemiol Community Health* 35 (3) (1981) 200–204.
- [43] E. Anthoine, et al., Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures, *Health Qual Life Outcomes* 12 (2014) 176.
- [44] M. Pritchard, et al., Psychometric properties of discourse measures in aphasia: acceptability, reliability, and validity, *Int J Lang Commun Disord* 53 (6) (2018) 1078–1093.
- [45] Hinkle, D.E., W. Wiersma, and S.G. Jurs, *Applied statistics for the behavioral sciences*. 2003, Houghton Mifflin; [Hi Marketing] (distributor) Boston, Mass., [London]: Boston, Mass., [London].
- [46] B.T. Hemker, K. Sijtsma, I.W. Molenaar, Selection of Unidimensional Scales From a Multidimensional Item Bank in the Polytomous Mokken I RT Model, *Applied Psychological Measurement* 19 (4) (1995) 337–352.
- [47] L.J. Cronbach, Coefficient alpha and the internal structure of tests, *Psychometrika* 16 (3) (1951) 297–334.
- [48] M. Tavakol, R. Dennick, Making sense of Cronbach's alpha, *Int J Med Educ* 2 (2011) 53–55, <https://doi.org/10.5116/ijme.4dfb.8dfd>.
- [49] T.K. Koo, M.Y. Li, A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research, *J Chiropr Med* 15 (2) (2016) 155–163.
- [50] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, *Biometrics* 33 (1) (1977) 159–174.
- [51] J.M. Bland, D.G. Altman, Comparing methods of measurement: why plotting difference against standard method is misleading, *Lancet* 346 (8982) (1995) 1085–1087.
- [52] R.W. Bohannon, Test-retest reliability of the five-repetition sit-to-stand test: a systematic review of the literature involving adults, *J Strength Cond Res* 25 (11) (2011) 3205–3207.
- [53] C.J. Nightingale, S.N. Mitchell, S.A. Butterfield, Validation of the Timed Up and Go Test for Assessing Balance Variables in Adults Aged 65 and Older, *J Aging Phys Act* 27 (2) (2019) 230–233.
- [54] C.K. Balasubramanian, The community balance and mobility scale alleviates the ceiling effects observed in the currently used gait and balance assessments for the community-dwelling older adults, *J Geriatr Phys Ther* 38 (2) (2015) 78–89.
- [55] D. Beck Jepsen, et al., Predicting falls in older adults: an umbrella review of instruments assessing gait, balance, and functional mobility, *BMC Geriatr* 22 (1) (2022) 615.
- [56] J. Verghese, et al., Walking while talking: effect of task prioritization in the elderly, *Arch Phys Med Rehabil* 88 (1) (2007) 50–53.
- [57] H.A.S., Actes d'ergothérapie et de psychomotricité susceptibles d'être réalisés pour la réadaptation à domicile des personnes souffrant de la maladie d'Alzheimer ou d'une maladie apparentée. 2010, Haute Autorité de Santé: Online.