Archives of Gerontology and Geriatrics 89 (2020) 104079

Contents lists available at ScienceDirect

Archives of Gerontology and Geriatrics

journal homepage: www.elsevier.com/locate/archger

Test-retest reliability of upper-limb proprioception and balance tests in older nursing home residents

Luis Galhardas[a,b], Armando Raimundo[a,b], José Marmeleira[a,b,*]

[a] Departamento de Desporto e Saúde, Escola de Ciências e Tecnologia, Universidade de Évora, Évora, Portugal
[b] Comprehensive Health Research Center (CHRC), Portugal

ARTICLE INFO

Keywords:
Reliability
Nursing homes
Proprioception
Balance
Field tests

ABSTRACT

Aim: To examine the test-retest reliability of two upper-limb proprioception tests (Weight Detection Test, or WDT, and Arm Ruler Positioning Test, or ARPT) and two balance tests (Functional Reach Test, or FRT, and Timed Up and Go test, or TUG) in older nursing home residents.

Methods: Fifty-three nursing home residents (85.9 ± 3.9 years) participated in this study. Outcome measures were assessed on two occasions, 10–14 days apart. The same rater administered all tests. The relative reliability was estimated using the intraclass correlation coefficients (ICCs) with a two-way mixed-effects model. The absolute reliability was analyzed using the standard error of the mean (SEM) to estimate the minimal detectable change (MDC) at the 95 % confidence level. Systematic bias was studied using the paired-samples *t*-test or the Wilcoxon signed-rank test.

Results: The WDT (ICC = 0.84), ARPT (ICC = 0.87) and FRT (ICC = 0.85) had good relative reliability, and the TUG (ICC = 0.99) had excellent reliability. Our results suggest acceptable measurement precision: the SEMs were equal to 1.0 points, and 0.3 cm, 1.5 cm, and 0.5 s for the WDT, ARPT, FRT, and TUG, respectively. The mean difference between sessions was 0.3 points (1.4 %; w = -1.37, p = 0.17) in the WDT, 0.1 cm (-0.74 %; t = 0.41, p = 0.68) in the ARPT, 0.1 cm (0.45 %; w = -0.33, p = 0.74) in the FRT, and 0.2 s (1.37 %; w = -2.28, p = 0.02) in the TUG.

Conclusions: This study showed that the four field-usable motor tests had good to excellent test-retest reliability and had acceptable measurement precision in older nursing home residents. These tests could be valuable clinical tools for assessing proprioception and balance in nursing home residents.

1. Introduction

As a large and growing number of older adults live in institutional settings (World Health Organization, 2015), several factors should be considered in promoting their independence, autonomy, and quality of life. The assessment of nursing home residents' abilities is of extreme importance for health personnel to plan appropriate (and individualized) intervention activities, monitor patient progress, and evaluate the effects of the interventions (Lee, Yu, Hsueh, Chen, & Hsieh, 2017; Ries, Echternach, Nof, & Gagnon Blodgett, 2009). Nursing home residents tend to be considerably frail (Cadore et al., 2015; González-Vaca et al., 2014; Tabue-Teguo et al., 2015), and, therefore, it is important to design interventions that have a positive impact on their functional capacity. Unfortunately, few studies have focused on the reliability of assessment methods in institutionalized older adults. Thus, although some studies analyzed the psychometric properties of several motor tests in older adults (Botolfsen, Helbostad, Moe-nilssen, & Wall, 2008; Olson & Zareh, 2019; Ries et al., 2009; Rosa, Perracini, & Ricci, 2019), they were conducted mainly in community-dwelling subjects.

Proprioception plays a vital role in controlling human movement, which is the basis of all activities of daily living (Henry & Baudry, 2019; Hillier, Immink, & Thewlis, 2015; Zou et al., 2019). Proprioceptive information includes body segment static position, displacement, acceleration, velocity, and muscle sense of tension/effort (Han, Waddington, Adams, Anson, & Liu, 2016; Hillier et al., 2015; Ogard, 2011). Several studies have highlighted the importance of proprioception in older people, showing, for instance, its relevance to body stability and the prevention of falls (Martínez-Amat et al., 2013; Sohn & Kim, 2015). Research has revealed that proprioception declines with aging, which may affect postural control and balance and increase the

* Corresponding author at: Departamento de Desporto e Saúde, Universidade de Évora, Pavilhão Gimnodesportivo da Universidade de Évora, Prolongamento da Rua de Reguengos de Monsaraz, 14, 7000-727, Évora, Portugal.
E-mail addresses: galhardas.luis@hotmail.com (L. Galhardas), jmarmel@uevora.pt (J. Marmeleira).

https://doi.org/10.1016/j.archger.2020.104079
Received 10 January 2020; Received in revised form 9 April 2020; Accepted 10 April 2020
Available online 22 April 2020
0167-4943/

risk of falls (Henry & Baudry, 2019; Zou et al., 2019). Despite the relevance of proprioception, few investigations have studied the reliability of this ability in older adults (Ko, Simonsick, Deshpande, & Ferrucci, 2015; Rinderknecht et al., 2018).

Most studies of proprioception in older adults have focused on the lower limb (specifically, the knee or ankle) (Henry & Baudry, 2019; Ko et al., 2015; Zou et al., 2019), probably due to the high frequency of balance problems in the older adults (González-Vaca et al., 2014; Henry & Baudry, 2019). However, it is desirable to assess upper-limb proprioception as well, given its direct potential effect on the ability to perform activities of daily living. Furthermore, most studies of proprioception focus on joint position sense, and there is a dearth of research on other components of proprioception (including muscle tension sense) that also play major roles in movement control (Hillier et al., 2015; Proske & Gandevia, 2012). In this context, the present study analyzes the reliability of two field proprioception tests for the upper limb, namely, the Weight Detection Test (WDT) and the Arm Ruler Positioning Test (ARPT), which assess muscle tension sense and joint position sense, respectively.

We also chose to examine the reliability of the Functional Reach Test (FRT) and the Timed Up and Go test (TUG), two important tests of functional physical fitness that focus balance in older adults (Duncan, Weiner, Chandler, & Studenski, 1992; Podsiadlo & Richardson, 1991). In the context of nursing home residences, balance should be carefully considered, as it influences most activities of daily living and is crucial for decreasing the risk of falls (Kocic et al., 2017; Telenius, Engedal, & Bergland, 2015). Although the FRT and TUG are relatively common in research with community-dwelling older adults (e.g., Kristensen, Bloch, Jønsson, & Jakobsen, 2019; Olson & Zareh, 2019), there is little evidence on their psychometric properties in institutionalized older adults, a population group that often shows reduced physical fitness, general frailty and cognitive deficits (Marmeleira, Ferreira, & Raimundo, 2017).

Therefore, the main objective of this study was to examine the test-retest reliability of four field motor tests (WDT and ARPT for assessing proprioception; FRT and TUG for assessing balance) in institutionalized people over 75 years old. We intend to provide meaningful information for use by health care personnel working with nursing home residents, a group of the population that frequently exhibits health and functional deficits.

2. Materials and methods

2.1. Participants

Five nursing homes were selected by convenience in the region of Évora (Portugal). With the help of the health care personnel of the nursing homes (who were informed of the study characteristics, especially the inclusion criteria), 53 potential participants were identified and invited to participate in the study. All residents who were invited agreed to participate in the study, and their eligibility was confirmed according to the following general inclusion criteria: being aged 75 years or older; living in a nursing home; and having a normal cognitive status according to the Portuguese version of the Mini-Mental State Examination (Folstein, Robins, & Helzer, 1983; Guerreiro et al., 1994). Considering the demands of each test, we also considered using a walker for ambulation as a specific exclusion criterion for the TUG and the FRT.

Table 1 shows the general characteristics of the participants. The sample included 41 women and 12 men aged 80–95 years. In general, the participants had a low educational level (54.7 % never attended school). According to the World Health Organization criteria regarding body mass index, 66 % of the participants were overweight (≥ 25 kg.m^2), and 23 % were obese (≥ 30 kg.m^2).

All participants in this study were informed about the objectives of the study and provided informed consent before participation. This study was approved by the Ethics Committee of the University of Évora and carried out in accordance with the Declaration of Helsinki (World Medical Association, 2013).

Table 1
General characteristics of the participants.

Characteristics	Mean (SD)	Min - max
Age (years)	85.9 (3.9)	80 – 95
Height (cm)	156.9 (7.2)	148.1–181.7
Weight (kg)	68.8 (6.5)	52.7–83.2
BMI (kg/m^2)	28.0 (2.7)	22.2–35.0
Education (years)	2.1 (3.1)	0 – 17
MMSE (points)	(3.1)	19 – 30

Note: MMSE, Mini-Mental State Examination; BMI, body mass index; SD, standard deviation.

2.2. Procedures

This was a test-retest reliability study, with all outcome measures collected on two occasions (10–14 days apart). The same kinesiologist administered two field tests to assess upper limb proprioception (WDT and ARPT) and two field tests to evaluate balance (FRT and TUG) at each time point.

Before collecting the data, the kinesiologist became familiar with all assessment methods by carefully studying the test protocols and applying each protocol to 5 institutionalized older adults under the supervision of a senior researcher with experience in this field. Furthermore, approximately 1 week before the initial assessment, all nursing home residents participated in a training session to become familiar with the tests to control for learning effects. All tests were performed in a quiet room at the nursing homes. The collection of all data took place between July and September 2019.

2.3. Measures

2.3.1. ARPT

This proprioception test was selected to assess the position sense of the upper limb. Participants were blindfolded and seated in a chair in front of a table. A measuring tape was fixed atop the table, paralleled to the participant's shoulders. The participant placed the index finger of the dominant hand over the measuring tape in front of the shoulder, marking the starting position. The nondominant hand rested on the leg. First, the participant's finger was slowly moved by the kinesiologist to the allocated target position, which was maintained for 5 s; afterward, the participant's finger was passively returned to the start position. Second, the participant attempted to replicate the target position actively and informed the kinesiologist when he/she considered that the previous position had been replicated. After three practice trials, the actual test consisted of 4 distances (10, 15, 25, 35 cm), two times each for a total of 8 trials. The 10- and 25-cm trials were performed to the midline of the body, and the 15- and 35-cm trials were performed to the opposite side. All subjects completed the same sequence of target positions (randomly established by the researchers). No feedback regarding performance was provided during the test. The outcome measurement was an error score calculated as the mean absolute difference between the target position and the actual position of the index finger (in cm) over the 8 target replication attempts. A similar protocol was used before with a mechanical linear device (Marmeleira et al., 2009).

2.3.2. WDT

The WDT was used to assess the muscle tension sense. The subject was seated next to a table in front of the researcher. The dominant forearm was placed on the table with the hand pointed toward the researcher. A set of weights was placed near the researcher on the table. All weights were of identical size and appearance: small black cylindric containers (2 cm diameter and 4 cm high) filled with tiny balls of lead.

For this test, a set of 16 weights was used: 11 weights ranging from 75 to 125 g in 5-gram intervals and an additional 5 weights weighing 100 g each. The weights were carefully checked with a high-precision balance (Ohaus EX324 N, Parsippany, NJ, USA). With the elbow constantly supported on the table, the participant slowly lifted each weight (with the fingertips) for 5 s and then put it down in a single, smooth motion. The participant practiced the task by lifting three different weights (90, 100, and 110 g). In the test, a standard weight (100 g) was always presented first, followed by a comparison weight, and the participant was asked to state verbally in each trial whether the comparison weight was lighter, of the same weight or heavier than the standard weight. The test was performed in two sets of 15 trials each, and the score was the total number of correct responses out of 30. In each set of trials, the same sequence of weights (randomly established by the researchers) was performed by all participants. The procedures used in this study were based on a previous protocol (Grouios, Alevriadou, & Koidou, 2001).

2.3.3. FRT

This test measures the limits of stability and the person's ability to preserve balance during a task (postural control). The functional reach is assumed to be the maximum distance (in cm) that can be reached beyond the outstretched upper limb, maintaining a fixed base of support in the standing position (Duncan et al., 1992).

The test mostly followed the original protocol (Duncan et al., 1992), with some differences in the materials used. Participants were asked to stand with their shoulder 10 cm away from a wall and to flex the upper limb at 90° with the fingers of the hand extended, as in previous studies (e.g., Bohannon, Wolfson, & White, 2017). The kinesiologist recorded this initial position (tip of middle finger) using a tape measure fixed to the wall and parallel to the ground. Then, the participants were requested to keep their feet flat on the floor and to reach in front as far as possible without losing balance or taking a step (with upper limb always outstretched). The kinesiologist registered the position reached by the tip of the middle finger when the participants stayed at least 2 s in that position. The score used was the difference (in cm) between the initial and final positions. Each participant performed the test one time for familiarization, followed two times for data collection. The best trial was used for data analysis.

2.3.4. TUG

The Timed Up and Go test (TUG) is a simple and easy-to-apply assessment that measures how long it takes for a person to stand up from a chair, walk three meters, navigate around a cone, walk back to the chair and sit again. The purpose of this test is to assess physical mobility in terms of speed, agility, and dynamic balance (Podsiadlo & Richardson, 1991). The test score corresponds to the time (in seconds) elapsed from the "starting" signal until the participant sat in the chair. There was one practice trial followed by two test trials, and the best was used for data analysis.

2.4. Statistical analysis

In the current investigation, we study the absolute and relative reliability of four tests. The relative reliability refers to the level of association (correlation) of repeated measurements; it represents the ratio of the total variability (between subjects/measurements) and individual variability (within-subjects/ measurements), which gives the coefficient of reliability (Bruton, Conway, & Holgate, 2000; Šerbetar, 2015). The test-retest relative reliability was estimated with the intraclass correlation coefficients (ICCs) using the two-way mixed-effects model analysis of variance (ANOVA). An ICC value greater than 0.90 indicates excellent reliability, 0.75 to 0.90 indicates good reliability, 0.50 to 0.75 indicates moderate reliability, and < 0.50 indicates poor reliability (Koo & Li, 2016).

Even with high ICCs, the consistency between assessment trials can be poor, especially when the study population may present very heterogeneous values (Bruton et al., 2000; Weir, 2005), and we also studied the absolute reliability. The absolute reliability depends on the variability of the scores from trial to trial (within-subjects/measurement), and this value is not a sample-dependent quantity since the range of individual scores is not considered (Šerbetar, 2015). The absolute reliability index was analyzed through the standard error of the mean (SEM) to estimate the minimal detectable change (MDC) at the 95 % confidence level. The formulas used to analyze SEM and MDC95 were as follows: $SEM = SD \times \sqrt{(1\text{-}ICC)}$ and $MDC_{95} = SEM \times \sqrt{2} \times 1.96$. SD corresponds to the average of standard deviations from test and retest sessions (Šerbetar, 2015). Relative to SEM, we take into consideration the criterion SEM < SD/2 for acceptable or nonacceptable measurement precision (Bautmans, Jansen, Van Keymolen, & Mets, 2011; Pontes & Griffiths, 2015; Šerbetar, 2015). MDC_{95} refers to the minimum change that must exist to be considered a real effect of an intervention or training program (Šerbetar, 2015; Weir, 2005). To facilitate the interpretation of the values obtained, we report SEM and MDC_{95} data in the same unit of measurement as the corresponding evaluation protocols.

Additionally, to detect possible systematic bias between sessions, after checking for data normality with the Shapiro–Wilk test, we conducted the paired-samples *t*-test or the Wilcoxon signed-ranks test. The level of significance was established to be $p < 0.05$. Considering the suggestion of Walter, Eliasziw, and Donner (1998), we calculated a required sample size of 39 to identify the desired ICC of 0.8, with a lower CI of 0.60, given an $\alpha = 0.05$ and $\beta = 0.20$. All data were analyzed using the SPSS statistical program (version 20.0, Inc., Chicago, IL, USA).

3. Results

In total, we collected data from 53 persons, most of whom completed all four physical performance tests. In the FRT and the TUG, six participants were not considered for analysis because they used a walker for ambulation.

Table 2 shows the mean scores and standard deviations (SDs) of the tests and their relative and absolute reliability values. Regarding the relative reliability, the WDT, ARPT, and FRT showed good ICCs (between 0.75 and 0.90), and the TUG showed an excellent ICC (0.99).

We also considered the absolute reliability of the tests. In Table 2, we present the SEM and MDC_{95}. According to the criterion SEM < SD/2, all tests have acceptable measurement precision (WDT = 1.0 < 1.3; ARPT = 0.3 < 0.5; FRT = 1.5 < 1.9, and TUG = 0.5 < 2.7).

No significant differences were found between test and retest scores for the WDT, the ARPT, or the FRT ($p < 0.05$). In the TUG, there were statistically significant changes between the two evaluations ($p < 0.05$).

Table 2

Test-retest reliability of four field motor tests in older nursing home residents.

Test item		Mean (SD)					
	n	Test	Retest	Difference	ICC (95 %)	SEM	MDC_{95}
WDT (points)	53	19.2 (2.6)	19.5 (2.5)	0.3	0.84 (0.74 – 0.91)	1.0	2.8
ARPT (cm)	53	2.7 (1.0)	2.7 (0.9)	0.1	0.87 (0.79 – 0.92)	0.3	0.9
FRT (cm)	47	19.9 (4.0)	20.0 (3.6)	0.1	0.85 (0.75 – 0.92)	1.5	4.0
TUG (s)	47	14.6 (5.3)[a]	14.8 (5.4)[a]	0.2	0.99 (0.99–1.00)	0.5	1.5

[a] Statistically significant differences between groups according to the Wilcoxon signed-ranks test. Abbreviations: WDT, Weight Detection Test; ARPT, Arm Ruler Positioning Test; FRT, Functional Reach Test; TUG, Timed Up and Go test.

The scores obtained in sessions one and two were very similar for all tests, and the paired sample *t*-test or the Wilcoxon signed-rank test confirmed that they were not significantly different except for the TUG. Hence, the mean difference in proprioception scores between the two test sessions was 0.3 points (1.4 %; w = -1.37, p = 0.17) for the WDT and 0.1 cm (-0.74 %; t = 0.41, p = 0.68) for the ARPT. In the balance tests, the mean score difference was 0.1 cm (0.45 %; w = -0.33, p = 0.74) for the FRT and 0.2 s (1.37 %; w = -2.28, p = 0.02) for the TUG.

4. Discussion

The main purpose of this study was to evaluate the absolute and relative reliability of 4 physical function tests (focusing on proprioception and balance) in older adults (80–95 years) living in nursing homes.

4.1. Proprioception tests

For both tests focused on proprioception, the results showed good relative reliability: ICCs of 0.84 and 0.87 in the WDT and the ARPT, respectively. There are few previous studies with which to compare our findings since most proprioception studies have focused on the lower limbs, especially knee-joint position sense (Henry & Baudry, 2019; Ko et al., 2015).

In the case of the ARPT, we found only two related studies that have examined the reliability of position sense measures of the upper limb (Hoseini, Sexton, Kurtz, Liu, & Block, 2015; Rinderknecht et al., 2018). Hoseini et al. (2015) proposed a method for assessing proprioception using a tablet in which participants (18 and 82 years old) were asked to identify the position of the index finger with respect to a marker (line). The authors reported an ICC of 0.62, far below the ICC value for the ARPT (ICC = 0.87). Rinderknecht et al. (2018) designed an automated robot-assisted assessment of proprioception for the metacarpophalangeal joint. In the procedure used, the participants were asked to identify the difference between angular joint positions, as in the protocol used in the current study. The ICC value reported by Rinderknecht et al. (2018) (ICC = 0.73) was relatively lower than the ICC we found in the present study. This difference is probably because the participants in the former study were people with stroke (which could compromise hand function), while the participants in our study were older adults without evident physical or cognitive impairment.

Our results suggest acceptable measurement precision for ARPT, as the SEM (0.3 cm) was below the criterion SEM < SD/2. We also found that, according to the MDC_{95} values, the changes between two measurements (for instance, pre- and postintervention scores) should be at least 0.9 cm for the ARPT to be considered meaningful in older adults living in nursing homes.

The present study fills a void by providing evidence that a test focus on muscle tension sense, namely the WDT, has good relative reliability (ICC = 0.84) in older people, specifically in those living in institutionalized settings. This test has the advantage of being very easy to administer and to score. Surprisingly, we have not found studies about the psychometric properties of specific assessment methods of the muscle tension sense. Furthermore, we have identified only a small number of studies that measured such ability, and all focused on specific groups of the population, namely people with Parkinson's disease, schizophrenia, or visual impairment (Chang & Lenzenweger, 2005; Grouios et al., 2001; Maschke, Tuite, Krawczewski, Pickett, & Konczak, 2006; Ritzler & Rosenbaum, 1974).

For the WDT, the SEM (1.0 points) was below SD/2 (i.e., 2.5 points), indicating that this proprioception test has acceptable measurement precision. Furthermore, the MDC_{95} was 2.8 points. Unfortunately, as stated before, we did not find any studies focused on the reliability of similar methods for measuring muscle tension sense. For instance, Chang and Lenzenweger (2005) used a similar weight discrimination test in patients with schizophrenia but did not report any data on its reliability.

Finally, regarding the two proprioception tests, the paired-samples *t*-test or the Wilcoxon signed-ranks test showed that there was no systematic bias in the ARPT or the WDT, respectively. This adds evidence to the robustness of both tests.

4.2. Balance tests

For the two tests focused on balance, our results showed that the relative reliability was good for the FRT (ICC = 0.85) and excellent for the TUG (ICC = 0.99). The reliability value of the FRT was in line with previous research (Galen et al., 2015; Merchán-Baeza, González-Sánchez, & Cuesta-Vargas, 2014; Olson & Zareh, 2019). Additionally, Martins, de Menezes, de Sousa, de Araujo Barbosa, and Costa (2012) carried out a study with a similar protocol. Those investigators reported reliable (> 0.82) inter- and intrarater values for people (28–80 years old) with and without motor disability (chronic hemiparesis). Galen et al. (2015) proposed a novel FRT protocol using new technology (Optotrak Certus 3D motion-capture system and specific software) and concluded that the FRT has good relative reliability (ICC = 0.79) in people with mild balance deficit (55.8 ± 13.5 years). Finally, a recent study (Mullin et al., 2018) in adults with neurofibromatosis (16–66 years) also reported intrarater reliability (ICC = 0.90) comparable to what we found in the present study.

Our results show that the TUG has excellent relative reliability (ICC = 0.99). This test has been widely used in different populations, including community-dwelling older adults and people with cognitive (e.g., dementia) or physical disabilities (e.g., balance and gait limitations) (Botolfsen et al., 2008; Cadore et al., 2015; Chang & Lenzenweger, 2005; Kristensen, Foss, & Kehlet, 2009). Several studies have confirmed that the TUG is a reliable test in older adults comparable in age to the participants in the present study, although not living in nursing homes. For instance, Ries et al. (2009) studied the reliability of the TUG in a group of older adults (average of 80.7 years; 66.7 % women) with cognitive impairment and reported ICC values (0.985 − 0.988) very close to those in the present study. Similarly, the results obtained by Kristensen et al. (2019) in older adults (average of 80 years) recovering from hip fracture confirm that the TUG has very high relative reliability (ICC = 0.98).

Regarding absolute reliability, our results showed that both the FRT and the TUG have acceptable measurement precision (SEM = 1.5 cm and 0.5 s, respectively), as these values are below the criteria on SEM < SD/2.

For comparison with the absolute reliability scores of the TUG and the FRT, we choose to consider mainly studies that examined the absolute reliability of both tests. Muir-Hunter, Graham, and Odasso (2015) studied the absolute reliability of the TUG and FRT in a group of older community residents (with Alzheimer's disease) with an average age (85.9 ± 3.9 years) relatively close to that in the present study (80.2 ± 5.0 years). The authors reported higher values for the TUG (SEM = 1.24 s MDC_{95} = 3.44 s) and for the FRT (SEM = 4.56 cm; MDC_{95} = 12.64 cm), but it should be noted that, in contrast to our study, the sample consisted of people with cognitive deficits. Other authors (Huang et al., 2011) studied the TUG to estimate the MDC_{95} in people (average age 67.5 years) with Parkinson's disease. The authors reported an MDC_{95} of 3.5 s, slightly higher than the MDC_{95} that we found. We also found a study (Mullin et al., 2018), of the interrater reliability of the FRT and the TUG in adults with neurofibromatosis, which reported slightly higher absolute reliability measures than we found in our study for the FRT (SEM = 2.92 cm, MDC_{95} = 8.08 cm) and for the TUG (SEM = 1.03 s and MDC_{95} = 2.86 s).

We also studied the systematic bias of the two balance tests. Thus, the Wilcoxon signed-rank test showed that there was no systematic bias in the FRT, but in the TUG, there were significant changes from session 1 to session 2 (w = -2.28, p = 0.02). This is probably due to changes in the rank position of some participants between sessions since the group means performance was very similar between the two time points. Muir-Hunter et al. (2015) also found statistically significant differences between sessions ($p = 0.002$) for a group of community-dwelling adults.

5. Limitations and strengths of the study

This study had some limitations. First, the sample included more women than men, although it should be noted that in Portugal, there are more older women than men, especially at advanced ages (as is the case for participants in the present study). Second, we did not examine the participant's level of frailty or their ability to perform activities of daily living.

Conversely, this study has some important strengths. It provides evidence that two field-usable tests measuring two different components of upper-limb proprioception (sense of muscle tension and sense of positioning) have highly satisfactory reliability in older adults living in nursing homes. Notably, previous related studies have focused mostly on the lower limb and the joint position sense component of proprioception. Additionally, although there have been some previous studies on the reliability of the TUG and the FRT in older adults (Huang et al., 2011; Kristensen et al., 2019), almost no evidence was available in older adults living in nursing homes.

6. Conclusion

This study showed that four field-usable motor tests (the WDT and ARPT for assessing proprioception; the FRT and TUG for assessing balance/agility) had good to excellent test-retest reliability and acceptable measurement precision in older nursing home residents. Thus, these tests could be valuable clinical tools for assessing proprioception and balance among institutionalized older adults.

Funding

This study was supported by the National Funds through FCT - Portuguese Foundation for Science and Technology (SFRH/BD/140669/2018) – and the European Fund for regional development (FEDER). This study was also supported by Horizon - Portugal – Alentejo 2020 (ALT20-03-0145-FEDER-000007) – ESACA. The sponsors had no role in the preparation of this manuscript.

CRediT authorship contribution statement

Luis Galhardas: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization, Supervision, Project administration, Funding acquisition. **Armando Raimundo:** Validation, Formal analysis, Writing - review & editing, Funding acquisition. **José Marmeleira:** Conceptualization, Methodology, Formal analysis, Writing - review & editing, Supervision.

Declaration of Competing Interest

None.

Acknowledgments

The authors thank the participants and the nursing home personnel for their contributions to this work.

References

Bautmans, I., Jansen, B., Van Keymolen, B., & Mets, T. (2011). Reliability and clinical correlates of 3D-accelerometry based gait analysis outcomes according to age and fall-risk. *Gait & Posture, 33*(3), 366–372. https://doi.org/10.2522/ptj.20080258.

Bohannon, R. W., Wolfson, L. I., & White, W. B. (2017). Functional reach of older adults: Normative reference values based on new and published data. *Physiotherapy, 103*(4), 387–391. https://doi.org/10.1016/j.physio.2017.03.006.

Botolfsen, P., Helbostad, J. L., Moe-nilssen, R., & Wall, J. C. (2008). Reliability and concurrent validity of the Expanded Timed Up-and-Go test in older people with impaired mobility. *Physiotherapy Research International, 13*(2), 94–106. https://doi.org/10.1002/pri.394.

Bruton, A., Conway, J. H., & Holgate, S. T. (2000). Reliability: What is it, and how is it measured? *Physiotherapy, 86*(2), 94–99. https://doi.org/10.1016/S0031-9406(05)61211-4.

Cadore, E. L., Casas-Herrero, A., Zambom-Ferraresi, F., Martínez-Ramírez, A., Millor, N., Gómez, M., ... Izquierdo, M. (2015). Do frailty and cognitive impairment affect dual-task cost during walking in the oldest old institutionalized patients? *AGE, 37*(6), 124. https://doi.org/10.1007/s11357-015-9862-1.

Chang, B. P., & Lenzenweger, M. F. (2005). Somatosensory processing and schizophrenia liability: Proprioception, exteroceptive sensitivity, and graphesthesia performance in the biological relatives of schizophrenia patients. *Journal of Abnormal Psychology, 114*(1), 85–95. https://doi.org/10.1037/0021-843X.114.1.85.

Duncan, P. W., Weiner, D. K., Chandler, J., & Studenski, S. (1992). Functional reach: Predictive validity in a sample of elderly male veterans. *Journal of Gerontology, 47*(3), M93–M98. https://doi.org/10.1093/geronj/47.3.m93.

Folstein, M. F., Robins, L. N., & Helzer, J. E. (1983). The mini-mental state examination. *Archives of General Psychiatry, 40*(7), 812. https://doi.org/10.1001/archpsyc.1983.01790060110016.

Galen, S. S., Pardo, V., Wyatt, D., Diamond, A., Brodith, V., & Pavlov, A. (2015). Validity of an interactive functional reach test. *Games for Health Journal, 4*(4), 278–284. https://doi.org/10.1089/g4h.2015.0002.

González-Vaca, J., de la Rica-Escuín, M., Silva-Iglesias, M., Arjonilla-García, M. D., Varela-Pérez, R., Oliver-Carbonell, J. L., & Abizanda, P. (2014). Frailty in INstitutionalized older adults from ALbacete. The FINAL Study: Rationale, design, methodology, prevalence and attributes. *Maturitas, 77*(1), 78–84. https://doi.org/10.1016/j.maturitas.2013.10.005.

Grouios, G., Alevriadou, A., & Koidou, I. (2001). Weight-discrimination sensitivity in congenitally blind and sighted adults. *Journal of Visual Impairment & Blindness, 95*(1), 30–39. https://doi.org/10.1177/0145482X0109500104.

Guerreiro, M., Silva, A. P., Botelho, M. A., Leitão, O., Castro-Caldas, A., & Garcia, C. (1994). Adaptação à população portuguesa da tradução do Mini Mental State Examination (MMSE). *Revista Portuguesa de Neurologia, 1*(9), 9–10.

Han, J., Waddington, G., Adams, R., Anson, J., & Liu, Y. (2016). Assessing proprioception: A critical review of methods. *Journal of Sport and Health Science, 5*(1), 80–90. https://doi.org/10.1016/j.jshs.2014.10.004.

Henry, M., & Baudry, S. (2019). Age-related changes in leg proprioception: Implications for postural control. *Journal of Neurophysiology, 122*(2), 525–538. https://doi.org/10.1152/jn.00067.2019.

Hillier, S., Immink, M., & Thewlis, D. (2015). Assessing proprioception: A systematic review of possibilities. *Neurorehabilitation and Neural Repair, 29*(10), 933–949. https://doi.org/10.1177/1545968315573055.

Hoseini, N., Sexton, B. M., Kurtz, K., Liu, Y., & Block, H. J. (2015). Adaptive staircase measurement of hand proprioception. *PloS One, 10*(8), e0135757. https://doi.org/10.1371/journal.pone.0135757 Retrieved from.

Huang, S. L., Hsieh, C. L., Wu, E. M., Tai, C. H., Lin, C. H., & Lu, W. S. (2011). Minimal detectable change of the timed "up & go" test and the dynamic gait index in people with parkinson disease. *Physical Therapy, 91*(1), 114–121. https://doi.org/10.2522/ptj.20090126.

Ko, S.-U., Simonsick, E., Deshpande, N., & Ferrucci, L. (2015). Sex-specific age associations of ankle proprioception test performance in older adults: Results from the Baltimore Longitudinal Study of Aging. *Age and Ageing, 44*(3), 485–490. https://doi.org/10.1093/ageing/afv005.

Kocic, M., Stojanovic, Z., Lazovic, M., Nikolic, D., Zivkovic, V., Milenkovic, M., & Lazarevic, K. (2017). Relationship between fear of falling and functional status in nursing home residents aged older than 65 years. *Geriatrics & Gerontology International, 17*(10), 1470–1476. https://doi.org/10.1111/ggi.12897.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Kristensen, M. T., Bloch, M. L., Jønsson, L. R., & Jakobsen, T. L. (2019). Interrater reliability of the standardized Timed Up and Go Test when used in hospitalized and community-dwelling older individuals. *Physiotherapy Research International, 24*(2), e1769. https://doi.org/10.1002/pri.1769.

Kristensen, M. T., Foss, N. B., & Kehlet, H. (2009). Factors with independent influence on the 'timed up and go' test in patients with hip fracture. *Physiotherapy Research International, 14*(1), 30–41. https://doi.org/10.1002/pri.414.

Lee, Y. C., Yu, W. H., Hsueh, I. P., Chen, S. S., & Hsieh, C. L. (2017). Test-retest reliability and responsiveness of the Barthel Index-based Supplementary Scales in patients with stroke. *European Journal of Physical and Rehabilitation Medicine, 53*(5), 710–718. https://doi.org/10.23736/S1973-9087.17.04454-9.

Marmeleira, J., Pereira, C., Cruz-Ferreira, A., Fretes, V., Pisco, R., & Fernandes, O. M. (2009). Creative dance can enhance proprioception in older adults. *The Journal of Sports Medicine and Physical Fitness, 49*(4), 480–485. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20087310.

Marmeleira, J., Ferreira, S., & Raimundo, A. (2017). Physical activity and physical fitness of nursing home residents with cognitive impairment: A pilot study. *Experimental Gerontology, 100*, 63–69. https://doi.org/10.1016/j.exger.2017.10.025.

Martínez-Amat, A., Hita-Contreras, F., Lomas-Vega, R., Caballero-Martínez, I., Alvarez, P. J., & Martínez-López, E. (2013). Effects of 12-Week proprioception training program on postural stability, gait, and balance in older adults. *Journal of Strength and Conditioning Research, 27*(8), 2180–2188. https://doi.org/10.1519/JSC.

0b013e31827da35f.

Martins, E. F., de Menezes, L. T., de Sousa, P. H. C., de Araujo Barbosa, P. H. F., & Costa, A. S. (2012). Reliability of the Functional Reach Test and the influence of anthropometric characteristics on test results in subjects with hemiparesis. *Neurorehabilitation, 31*(2), 161–169. https://doi.org/10.3233/NRE-2012-0786.

Maschke, M., Tuite, P. J., Krawczewski, K., Pickett, K., & Konczak, J. (2006). Perception of heaviness in Parkinson's disease. *Movement Disorders, 21*(7), 1013–1018. https://doi.org/10.1002/mds.20876.

Merchán-Baeza, J. A., González-Sánchez, M., & Cuesta-Vargas, A. I. (2014). Reliability in the parameterization of the functional reach test in elderly stroke patients: A pilot study. *BioMed Research International, 2014*, 1–8. https://doi.org/10.1155/2014/637671.

Muir-Hunter, S. W., Graham, L., & Odasso, M. M. (2015). Reliability of the Berg balance scale as a clinical measure of balance in community-dwelling older adults with mild to moderate alzheimer disease: A pilot study. *Physiotherapy Canada, 67*(3), 255–262. https://doi.org/10.3138/ptc.2014-32.

Mullin, R. L., Golding, J. F., Smith, R., Williams, V., Thomas, M., & Ferner, R. E. (2018). Reliability of functional outcome measures in adults with neurofibromatosis 1. *SAGE Open Medicine, 6*. https://doi.org/10.1177/2050312118786860 2050312118786860.

Ogard, W. K. (2011). Proprioception in sports medicine and athletic conditioning. *Strength and Conditioning Journal, 33*(3), 111–118. https://doi.org/10.1519/SSC.0b013e31821bf3ae.

Olson, L., & Zareh, A. (2019). Reliability of the Functional Reach Test using a mobile pole versus the traditional fixed ruler. *Journal of Acute Care Physical Therapy, 10*(1), 31–35. https://doi.org/10.1097/JAT.0000000000000088.

Podsiadlo, D., & Richardson, S. (1991). The timed "up & go": A test of basic functional mobility for frail elderly persons. *Journal of the American Geriatrics Society, 39*(2), 142–148. https://doi.org/10.1111/j.1532-5415.1991.tb01616.x.

Pontes, H. M., & Griffiths, M. D. (2015). Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. *Computers in Human Behavior, 45*, 137–143. https://doi.org/10.1016/j.chb.2014.12.006.

Proske, U., & Gandevia, S. C. (2012). The proprioceptive senses: Their roles in signaling body shape, body position and movement, and muscle force. *Physiological Reviews, 92*(4), 1651–1697. https://doi.org/10.1152/physrev.00048.2011.

Ries, J. D., Echternach, J. L., Nof, L., & Gagnon Blodgett, M. (2009). Test-retest reliability and minimal detectable change scores for the Timed "Up & Go" Test, the six-minute walk test, and gait speed in people with alzheimer disease. *Physical Therapy, 89*(6), 569–579. https://doi.org/10.2522/ptj.20080258.

Rinderknecht, M. D., Lambercy, O., Raible, V., Büsching, I., Sehle, A., Liepert, J., & Gassert, R. (2018). Reliability, validity, and clinical feasibility of a rapid and objective assessment of post-stroke deficits in hand proprioception. *Journal of Neuroengineering and Rehabilitation, 15*(1), 47. https://doi.org/10.1186/s12984-018-0387-6.

Ritzler, B., & Rosenbaum, G. (1974). Proprioception in schizophrenics and normals: Effects of stimulus intensity and interstimulus interval. *Journal of Abnormal Psychology, 83*, 106–111. https://doi.org/10.1037/h0036477.

Rosa, M. V., Perracini, M. R., & Ricci, N. A. (2019). Usefulness, assessment and normative data of the Functional Reach Test in older adults: A systematic review and meta-analysis. *Archives of Gerontology and Geriatrics, 81*, 149–170. https://doi.org/10.1016/j.archger.2018.11.015.

Šerbetar, I. (2015). Establishing some measures of absolute and relative reliability of a motor test. *Croatian Journal of Education - Hrvatski Časopis Za Odgoj i Obrazovanje, 17*(1), 37–48. https://doi.org/10.15516/cje.v17i0.1484.

Sohn, J., & Kim, S. (2015). Falls study: Proprioception, postural stability, and slips. *Biomedical Materials and Engineering, 26*(s1), S693–S703. https://doi.org/10.3233/BME-151361.

Tabue-Teguo, M., Kelaiditi, E., Demougeot, L., Dartigues, J.-F., Vellas, B., & Cesari, M. (2015). Frailty index and mortality in nursing home residents in France: Results from the INCUR study. *Journal of the American Medical Directors Association, 16*(7), 603–606. https://doi.org/10.1016/j.jamda.2015.02.002.

Telenius, E. W., Engedal, K., & Bergland, A. (2015). Long-term effects of a 12 weeks high-intensity functional exercise program on physical function and mental health in nursing home residents with dementia: A single blinded randomized controlled trial. *BMC Geriatrics, 15*(1), 158. https://doi.org/10.1186/s12877-015-0151-8.

Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine, 17*(1), 101–110. https://doi.org/10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength and Conditioning Research, 19*(1), 231. https://doi.org/10.1519/15184.1.

World Health Organization (2015). *World report on ageing and health*Retrieved fromhttps://www.who.int/ageing/events/world-report-2015-launch/en/.

World Medical Association (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA, 310*(20), 2191. https://doi.org/10.1001/jama.2013.281053.

Zou, L., Han, J., Li, C., Yeung, A. S., Hui, S. S., Tsang, W. W. N., ... Wang, L. (2019). Effects of tai chi on lower limb proprioception in adults aged over 55: A systematic review and meta-analysis. *Archives of Physical Medicine and Rehabilitation, 100*(6), 1102–1113. https://doi.org/10.1016/j.apmr.2018.07.425.