

## Reliability of 4-m and 6-m walking speed tests in elderly people with cognitive impairment

Carmen L. Muñoz-Mendoza<sup>a</sup>, M. José Cabañero-Martínez<sup>a</sup>, José C. Millán-Calenti<sup>b</sup>, Julio Cabrero-García<sup>a</sup>, Rosalía López-Sánchez<sup>b</sup>, Ana Maseda-Rodríguez<sup>b</sup>

<sup>a</sup> Department of Nursing, University of Alicante, Carretera de San Vicente s/n 03690 (San Vicente del Raspeig), Alicante, Spain

<sup>b</sup> Gerontology Research Group, Department of Medicine, Faculty of Health Sciences, University of A Coruña, Ed. Universitario de Oza, Oza s/n 15006-A Coruña, Spain

### Abstract

The purpose of this study was to evaluate the interrater and test–retest reliability of 4-m and 6-m walking speed tests in elderly people with cognitive impairment. 50 subjects aged 65 and over with cognitive impairment were selected from an adult day-care centre and a nursing home. To examine interrater reliability, 21 people were evaluated independently by two researchers who administered the 4-m and 6-m walking tests in each evaluation. For test–retest reliability, two observers administered the tests to the same 29 subjects, with a time interval of one week. Intraclass correlation coefficients (ICCs) were calculated to examine interrater and test–retest reliability. The ICCs for interrater reliability reached values of 0.96 and 0.88 for the 4-m and 6-m walking tests, respectively. In the test–retest study, the time interval was  $7.4 \pm 1.17$  days. The ICCs were 0.91 for the 4-m test and 0.86 for the 6-m test. The results support the use of walking tests in elderly people with cognitive impairment.

*Keywords:* Walking speed test Elderly people Cognitive impairment Reliability

### 1. Introduction

Walking speed tests are the most widely used objective physical performance-based tests to evaluate mobility in elderly people (Imms and Edholm, 1981; Buchner et al., 1995; Guralnik and Ferrucci, 2003; Wang et al., 2009). Longitudinal studies have demonstrated that walking speed tests are useful to predict important adverse results such as: hospitalization, dependence, frailty and mortality (Cesari et al., 2005; Onder et al., 2005; Ostir et al., 2007; Kim et al., 2009); however, numerous studies have reported high values for test–retest and interrater reliability (Hoeymans et al., 1997; Jette et al., 1999; Ostchega et al., 2000; Cabrero-García et al., 2007). Tests over 4m and 6 m are most commonly used in studies with elderly people (Muñoz-Mendoza et al., 2010). Although they are both short-distance tests, Guralnik et al. (2000) consider that 4 m is an adequate distance to evaluate walking speed; furthermore, it is more viable for use both in the home context and at the clinic. In the literature, few studies evaluate the reliability of walking tests in subjects with cognitive impairment. Tappen et al. (1997) reported reliability values between 0.57 and 0.97 for the 25-foot test (7.5 m) and Thomas and Hageman (2002) obtained an ICC of 0.92 for the 6-m test. Evidence of reliability for this type of population is still scarce. Therefore, the purpose of this study was to evaluate the reliability of 4-m and 6-m walking speed tests on elderly people with cognitive impairment.

### 2. Subjects and methods

#### 2.1. Participants

The sample included people aged 65 and over with different degrees of cognitive impairment, selected from a population of 119 people belonging to the Gerontological Complex “La Milagrosa”, which includes an adult day-care centre and a nursing home (A Coruña, Spain). So as not to burden the subjects, they were divided randomly into two groups, one to examine interrater reliability and the other to determine test–retest reliability. People with cardiac or respiratory pathologies which could be aggravated with exercise were excluded from the study, as were those with visual or hearing problems that impeded their evaluation, people who were unable to walk at least 10 m without assistance, people incapable of understanding simple verbal orders to walk, and people who scored over 24 in the mini-mental state examination (MMSE) (Folstein et al., 1975; Blesa et al., 2001). Before commencing the study, informed consent was gained from all the participants, either autonomously or through their legal representatives.

**Table 1**

Baseline characteristics of the samples, n, mean± S.D., range, n (%).

Characteristics	Interrater reliability sample (n = 21)		Test–retest reliability sample (n = 29)	
Age (years)	84.4+/-6.9	65.0=95.0	83.3+/-5.5	73.0=92.0
Gender				
Male	3 (14.3)		7 (24.1)	
Female	18 (85.7)		22 (75.9)	
Height (cm)	154.6+/-7.8	140.0=168.5	153.9+/-7.6	141.2=173.4
Weight (kg)	66.2+/-13.0	50.6=93.6	65.6+/-11.9	48.8=96.4
BMI (kg/m <sup>2</sup> )	27.7+/-5.1	20.0=38	4 27.8+/-4.8	17.9=38.4
Level of education				
Unable to read or write	1 (4.8)		0	
Incomplete primary school	7 (33.3)		12 (41.4)	
Completed primary school	9 (42.9)		14 (48.3)	
High school or university	4 (19.0)		3 (10.3)	
Marital status				
Single	2 (9.5)		3 (10.3)	
Married	4 (19.2)		4 (13.8)	
Widowed	15 (71.4)		22 (75.9)	
MMSE score	16.4+/-4.5	10.0=23.0	16.7+/-5.0	6.0=24.0
BI score	78.6+/-19.8	25.0=100.0	76.4+/-14.1	35.0=100.0
4-m time (s)	8.6+/-5.4	4.1=28.4	7.8+/-2.9	3.4=16.0
6-m time (s)	11.9+/-5.	8 6.2=29.2	11.0+/-4.5	5.8=24.8

Notes: BI= Barthel index; BMI = body mass index; MMSE = mini-mental state examination; S.D. = standard deviation; TUG = timed-up-and-go test.

### 2.2.1. Walking speed tests

The walking speed tests were adapted from the protocols used in previous studies (Guralnik et al., 1994; Cabrero-García et al., 2007). In this study, the protocol instructions were simplified. Participants were told to stand with their feet touching the starting line and were then given the following instruction ‘‘I want you to walk until you cross that line, at your normal speed, as if you were walking down the corridor’’, and they began walking when they received the simple order: ‘‘start walking’’. Each test was carried out twice, recording the shortest time (in seconds) required to cover each distance. They were allowed to use any technical means of assistance required but not the help of another person.

To examine interrater reliability, 25 people were evaluated independently by two previously trained researchers, with a rest period of 10 min between evaluations. The researchers administered the 4-m and 6-m tests alternately in each evaluation.

To evaluate test–retest reliability, the same two observers administered the 4-m and 6-m tests to the same participants in the test and retest, with a time interval of one week.

### 2.2.2. Other study measures

In the baseline evaluation, the level of physical disability was measured using Barthel index (BI)(Mahoney and Barthel, 1965; Bazta’net al., 1993) and cognitive functioning was assessed using the MMSE (Folstein et al., 1975; Blesa et al., 2001). Additionally, standardized measures of weight and size were recorded and used to calculate body mass index (BMI) as an approximate measurement of body fat.

## 2.3. Statistical analysis

The characteristics of the sample were analyzed using descriptive statistics and frequency distributions. Interrater and test–retest reliability were calculated using the ICC, the recommended statistic for this type of study (SACMOT, 2002; Terwee et al., 2007). To calculate the ICCs, a two-way random effects

## 3. Results

Of the 119 subjects admitted to the adult day-care centre or nursing home, 55 met the eligibility criteria. Of these, 21 took part in the interrater reliability study and 29 completed the test–retest reliability study. The reasons for losses during the data gathering process were: hospitalization (n = 2), immobility (n = 1), refusal to take part (n =1) and leaving the centre (n = 1).

Table 1 summarizes the baseline characteristics of the two samples. The mean age of the participants in the interrater reliability study was  $84.4 \pm 6.9$  years. There was a higher proportion of female to male. The mean BMI was  $27.7 \pm 5.1$ . The mean scores on the MMSE and BI were, respectively,  $16.4 \pm 4.5$  and  $78.6 \pm 19.8$ . As regards the characteristics of the sample that took part in the test–retest reliability study, the mean age of the subjects was  $83.3 \pm 5.5$  years. 75.9% were female. The BMI was  $27.8 \pm 4.8$ . The mean MMSE score was  $16.7 \pm 5.0$  and  $76.4 \pm 14.1$  on the BI.

As for reliability data, the Bland–Altman graphic method revealed no significant findings in terms of trends and only one outlying value, which was excluded from the analysis of interrater reliability (figure not shown, provided on request). Hence, the ICCs reached values of 0.96 (95% CI = 0.88–0.98) and 0.88 (95% CI = 0.70–0.96) for walking speed tests over four and 6m, respectively (Table 2).

In the test–retest study, the time interval was  $7.4 \pm 1.17$  days. The ICCs were 0.91 (95% CI = 0.81–0.96) for the 4-m test and 0.86 (95% CI = 0.71–0.94) for the 6-m test (Table 2).

**Table 2**  
Interrater and test–retest reliability of the 4 and 6-m walking tests.

	n ICC (95% CI) <sup>b</sup>	Pearson correlation	Mean $\pm$ S.D	
			Observation 1	Observation 2
Interrater				
4m	16 0.96 (0.88–0.98) 0.96*		7.51 $\pm$ 2.55	7.06 $\pm$ 2.66
6m	15 0.88 (0.70–0.96) 0.93*		10.85 $\pm$ 3.98	11.54 $\pm$ 5.31
Test–retest				
4m	29 0.91 (0.81–0.96) 0.93*		7.79 $\pm$ 2.88*	8.03 $\pm$ 3.63
6m	26 0.86 (0.71–0.94) 0.88*		10.97 $\pm$ 4.49	11.87 $\pm$ 5.17

a People who agreed to take part in the two evaluations.

b 2-way random effects ANOVA, definition of absolute agreement.

\*  $p = 0.01$ .

The Pearson correlation coefficients were similar to the ICCs, both in terms of interrater reliability and test–retest reliability (Table 2).

#### 4. Discussion

This study has examined interrater and test–retest reliability on this population. The values for interrater and test–retest reliability were high for both the 4-m and 6-m test, according to the quality standards of the SACMOT (2002) and Terwee et al. (2007), which give a value of 0.70 for the minimum reliability required to make group decisions. The 4-m test achieved values higher than those recommended for individual decisions (0.90).

Although one limitation of this study is the small sample size, the literature on walking speed tests concurs with the results obtained in this study, both for similar and longer distances, and with smaller sample sizes. Thomas and Hageman (2002), on a sample of 10 patients with dementia, reported an ICC of 0.92 for the 6-m test. Similarly, on a sample of 33 patients with Alzheimer’s Disease, Tappen et al. (1997) presented ICC data ranging from 0.76 to 0.99 for interrater and test–retest reliability on the 6-min walk test and lower values for the 25-foot walking test (0.57–0.97). These tests are also comparable to those obtained for elderly people with their cognitive functioning intact (Studenski et al., 2003; Cabrero-García et al., 2007).

Another limitation deriving from the small sample size is the impossibility of stratifying the sample according to the level of cognitive impairment, and so we have not been able to examine whether the level of reliability is homogeneous within the range of cognitive impairment examined. However the Bland–Altman plot revealed no trend associated with the MMSE score.

Owing to the fact that the centre’s eligibility criteria are determined by the level of cognitive impairment, behavioral disorders and physical capacity, since the rates are established on the basis of these factors, in many cases, the existence of dementia has not yet been determined, or the type should it exist.

Tappen et al. (1997) suggest that the 6-min walk test is more appropriate than the 25-foot test for use on elderly people with Alzheimer’s disease. However, the results obtained here suggest that a distance of 4 m would be sufficient to evaluate physical performance in people with cognitive impairment. Furthermore, bearing in mind that these people display alterations in their execution, memory and attention functions (Perry and Hodges, 1999) and, therefore, can be easily distracted and forget instructions quickly, the use of short distances would contribute to the viability of this test

on this population.

In line with Hauer and Oster (2008), the authors of this paper believe that it is advisable to modify the test administration protocols (e.g., shorter instructions, demonstration of the task, repeating instructions). Although these modifications could make it difficult to compare with studies on populations with their cognitive functioning intact and with standard versions of these tests (Thomas and Hageman, 2002; Van Iersel et al., 2007), it should increase understanding and, therefore, the possibility of including people with cognitive impairment in other types of studies, especially in relation to physical rehabilitation.

Finally, it would be very useful to obtain information about cutoff points depending on the different levels of cognitive impairment, attempting to establish the threshold at which it would not be possible to administer physical performance-based measures.

## Conflicts of interest statement

None

## Acknowledgement

Our sincere thanks go to the Gerontological Complex “La Milagrosa” for giving us access to their elderly residents.

## References

- Baztán, J.J., Pérez, J., Alarcón, T., San Cristóbal, E., Izquierdo, G., Manzarbeitia, J., 1993. Índice de Barthel: instrumento válido para la valoración funcional de pacientes con enfermedad cerebrovascular. *Rev. Esp. Geriatr. Gerontol.* 28, 32–40 (in Spanish).
- Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1, 307–310.
- Blesa, R., Pujol, M., Aguilar, M., Santacruz, P., Bertrán-Serra, I., Hernández, G., Sol, J.M., Peña-Casanova, J., 2001. Clinical validity of the “Mini-Mental State” for Spanish speaking communities. *Group NORMACODEM Neuropsychologia* 39, 1150–1157.
- Buchner, D.M., Guralnik, J.M., Cress, M.E., 1995. The clinical assessment of gait, balance and mobility in older adults. In: Rubinstein, L.Z., Wieland, D., Bernabei, R. (Eds.), *Geriatric Assessment Technology: The State of the Art*. Editrice Kurtis, Milan, pp. 75–88.
- Cabrero-García, J., Reig-Ferrer, A., Muñoz-Mendoza, C.L., Cabañero-Martínez, M.J., Ramos-Pichardo, J.D., Richart-Martínez, M., González-Llopis, L., Sanjuán-Quiles, A., Gómez-Moreno, N., García-Beviá, S., 2007. Reproducibilidad de la batería EPESE de desempeño físico en atención primaria. *Anal. Modif. Conduct.* 33, 67–83.
- Cesari, M., Kritchevsky, S.B., Penninx, B.W.H.J., Nicklas, B.J., Simonsick, E.M., Newman, A.B., Tyllavsky, F.A., Brach, J.S., Satterfield, S., Bauer, D.C., Visser, M., Rubin, S.M., Harris, T.B., Pahor, M., 2005. Prognostic value of usual gait speed in well functioning older people: results from the health aging and body composition study. *J. Am. Geriatr. Soc.* 53, 1675–1680.
- Folstein, M.F., Folstein, S., McHugh, P.R., 1975. Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198.
- Guralnik, J.M., Ferrucci, L., 2003. Assessing the building blocks of function: utilizing measures of functional limitation. *Am. J. Prev. Med.* 25, 112–121.
- Guralnik, J.M., Simonsick, E.M., Ferrucci, L., Glynn, R.J., Berkman, L.F., Blazer, D.G., Scherr, P.A., Wallace, R.B., 1994. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J. Gerontol.* 49, M85–M94.
- Guralnik, J.M., Ferrucci, L., Pieper, C.F., Leveille, S.G., Markides, K.S., Ostir, G.V., Studenski, S., Berkman, L.F., Wallace, R.B., 2000. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *J. Gerontol. A: Biol. Sci. Med. Sci.* 55, M221–M231.
- Hauer, K., Oster, P., 2008. Measuring functional performance in persons with dementia. *J. Am. Geriatr. Soc.* 56, 949–950.
- Hoeymans, N., Wouters, E.R., Feskens, E.J., Van den Bos, G.A., Kromhout, D., 1997. Reproducibility of performance-based and self-reported measures of functional status. *J. Gerontol. A: Biol. Sci. Med. Sci.* 52, M363–M368.
- Imms, F.J., Edholm, O.G., 1981. Studies of gait and mobility in the elderly. *Age Ageing* 10, 147–156.
- Jette, A.M., Jette, D.U., Ng, J., Plotkin, D.J., Bach, M.A., 1999. Are performance based measures sufficiently reliable for use in multicenter trials? Musculo-skeletal Impairment (MSI) Study Group. *J. Gerontol. A: Biol. Sci. Med. Sci.* 54, M3–M6.
- Kim, M.-J., Yabushita, N., Kim, M.-K., Nemoto, M., Seino, S., Tanaka, K., 2009. Mobility performance tests for discriminating high risk of frailty in community-dwelling older women. *Geriatr. Arch. Gerontol.* doi:10.1016/j.archger.2009.10.007.
- Mahoney, F.I., Barthel, D.W., 1965. Functional evaluation: the Barthel index. *Md. State Med. J.* 14, 61–65.
- Muñoz-Mendoza, C.L., Cabrero-García, J., Reig-Ferrer, A., Cabañero-Martínez, M.J., 2010. Evaluation of walking speed tests as a measurement of functional limitations in elderly people: a structured review. *Int. J. Clin. Health Psychol.* 10, 359–378.
- Onder, G., Penninx, B.W., Ferrucci, L., Fried, L.P., Guralnik, J.M., Pahor, M., 2005. Measures of physical performance and risk for progressive and catastrophic disability: results from the Women’s Health and Aging Study. *J. Gerontol. A: Biol. Sci. Med. Sci.* 60, 74–79.
- Osthege, Y., Harris, T.B., Hirsch, R., Parsons, V.L., Kington, R., Katzoff, M., 2000. Reliability and prevalence of physical performance examination assessing mobility and balance in older persons in the US: Data from the Third National Health and Nutrition Examination Survey. *J. Am. Geriatr. Soc.* 48, 1136–1141.
- Ostir, G.V., Kuo, Y.F., Berges, I.M., Markides, K.S., Ottenbacher, K.J., 2007. Measures of lower body function and risk of mortality over 7 years of follow-up. *Am. J. Epidemiol.* 166, 599–605.
- Perry, R.J., Hodges, J.R., 1999. Attention and executive deficits in Alzheimer’s disease: a critical review. *Brain* 122, 383–404.
- Scientific Advisory Committee of the Medical Outcomes Trust, 2002. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual. Life Res.* 11, 193–205.

- Studenski, S., Perera, S., Wallace, D., Chandler, J.M., Duncan, P.W., Rooney, E., Fox, M., Guralnik, J.M., 2003. Physical performance measures in the clinical setting. *J. Am. Geriatr. Soc.* 51, 314–322.
- Tappen, R.M., Roach, K.E., Buchner, D., Barry, C., Edelstein, J., 1997. Reliability of physical performance measures in nursing home residents with Alzheimer's disease. *J. Gerontol. A: Biol. Sci. Med. Sci.* 52, M52–M55.
- Terwee, C.B., Bot, S.D., De Boer, M.R., Van der Windt, D.A., Knol, D.L., Dekker, J., Bouter, L.M., De Vet, H.C., 2007. Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol.* 60, 34–42.
- Thomas, V.S., Hageman, P.A., 2002. A preliminary study on the reliability of physical performance measures in older day-care center clients with dementia. *Int. Psychogeriatr.* 14, 17–23.
- Van Iersel, M.B., Benraad, C.E., Olde, M.G., 2007. Validity and reliability of quantitative gait analysis in geriatric patients with and without dementia. *J. Am. Geriatr. Soc.* 55, 632–634.
- Wang, C.-Y., Yeh, C.-J., Hu, M.-H., 2009. Mobility-related performance tests to predict mobility disability at 2-year follow-up in community-dwelling older adults. *Arch. Gerontol. Geriatr.* doi:10.1016/j.archger.2009.11.001.