Olivares, P.R.; Hernandez-Mocholi, M.; Merellano-Navarro, E.; Gusi, N.; Collado-Mateo, D. (2019). Analysis of Age on Fitness Reliability in Elderly. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 19 (76) pp. 627-639 Http://cdeporte.rediris.es/revista/revista76/artanalisis1088.htm
DOI: 10.15366/rimcafd2019.76.005

ORIGINAL

AGE ANALYSIS ON FITNESS RELIABILITY TESTS IN THE ELDERLY

ANÁLISIS DE LA EDAD SOBRE LA FIABILIDAD DE PRUEBAS FITNESS EN MAYORES

Olivares, P.R.[1]; Hernandez-Mocholi, M.[2]; Merellano-Navarro, E.[3]; Gusi, N.[4] y Collado-Mateo, D.[5]

[1] PhD, Facultad de Educación, Psicología y Ciencias del Deporte, Universidad de Huelva, Huelva (Spain). Instituto de Actividad Física y Salud. Universidad Autonoma de Chile (Chile) olivares.pedro@gmail.com

[2] MSc, Facultad de Ciencias del Deporte, Universidad de Extremadura (Spain) mihemo04@gmail.com

[3] PhD, Instituto de Actividad Física y Salud. Universidad Autonoma de Chile (Chile) emerellanon@uautonoma.cl

[4] PhD, Facultad de Ciencias del Deporte, Universidad de Extremadura (Spain) ngusi@unex.es

[5] PhD, Facultad de Ciencias del Deporte, Universidad de Extremadura (Spain) Instituto de Actividad Física y Salud. Universidad Autonoma de Chile (Chile) danicolladom@gmail.com

Spanish-English Translator: Claudia Elena Daza Botero, claudiaelena.daza@gmail.com

Funding

The study was supported by the Internal Research Fund of the Universidad Autónoma de Chile (DIUA115-2017). The author MAHM was supported by the Scholarship (FPU3839) from the Spanish Ministry of Education, Culture, and Sport. The author DCM was supported by the Scholarship (FPU14 / 01283) from the Spanish Ministry of Education, Culture, and Sport.

Código UNESCO / UNESCO Code: 320107 Geriatría / Geriatrics
Clasificación Consejo de Europa / Council of Europe Classification: 11 Medicina del Deporte / Sport Medicine

Recibido 16 de enero de 2018 **Received** January 16, 2018
Aceptado 27 de noviembre de 2018 **Accepted** November 27, 2018

ABSTRACT

The aim of this study was to analyze if age affects the reliability of some fitness test widely used in elderly adults. Participants were 135 elderly women aged between 60 and 90 years old distributed into 5 age groups. All participants performed twice a battery of fitness tests with an interval between measurements of 1 week. The reliability indexes obtained in the bi-handgrip tests and 6 minutes walking were excellent (ICC> 0.90), while the rest were good (ICC 0.70-0.89). No statistically significant differences were found in the measurement error of these tests between age groups. It is concluded that age does not significantly affect to the reliability of the analyzed fitness tests.

KEYWORDS: Elderly, functional capacity, reliability.

RESUMEN

El presente trabajo tiene como objetivo analizar la fiabilidad de varias pruebas de condición física en adultos mayores en función de la edad. Los participantes fueron 135 mujeres mayores entre 60 y 90 años y se distribuyó en cinco grupos de edad. Todos los participantes realizaron una batería de pruebas de condición física con un intervalo entre mediciones de 1 semana. Los índices de fiabilidad obtenidos en las pruebas dinamometría bi-manual y 6 minutos caminando fueron excelentes ICC> 0,90, mientras que en el resto fueron buenos (ICC 0,70–0,89). No se encontraron diferencias estadísticamente significativas en el error de medida de estas pruebas entre grupos de edad. Se concluye que la fiabilidad de las pruebas de condición física utilizadas en población adulto mayor no varía significativamente en función de la edad.

PALABRAS CLAVE: Ancianos, capacidad funcional, fiabilidad.

INTRODUCTION

Several studies have linked the fitness with the functional capacity (Merellano-Navarro, Collado-Mateo, García-Rubio, Gusi, & Olivares, 2017; Rikli & Jones, 2013; Sardinha, Santos, Marques, & Mota, 2015) and with the health-related quality of life (Chung, Zhao, Liu, & Quach, 2017; Olivares, Gusi, Prieto, & Hernandez-Mocholi, 2011). They have also been used to measure the effects of physical interventions and therapies in the elderly (Chang, Wang, Chen, & Hu, 2017; Chen *et al*., 2017; Gusi, Hernandez-Mocholi, & Olivares, 2015). Fitness assessments can play an essential role in promoting physical activity and health (Oja, 1995) since they can identify people at risk of developing chronic diseases, reduce physical fragility, and increase mobility (Chen, Lin, & Yu, 2009).

Before using any fitness test as a tool for measuring the effects of therapy, the reliability of the scores obtained and the measurement error to be assumed must be known. For analysis, several reliability studies specifically for fitness tests in the elderly have been published (Rodriguez, Valenzuela, Gusi, Nacher, & Gallardo, 1998; Shaulis, Golding, & Tandy, 1994). However, physical performance decreases with age (Chen *et al*., 2009; Krause *et al*., 2009; Rikli & Jones, 1999b) mainly due to deterioration of the aerobic endurance, flexibility, strength, speed, agility, and balance (Milanović *et al*., 2013) and this decrease in physical performance could affect their reliability indices. As far as we are concerned, not a single study analyses the effect of age on the reliability indices of the most commonly used fitness tests. This study aimed to analyze the reliability of a set of fitness in seniors according to age groups, allowing more accurate detection of real changes in physical therapies in this population. Furthermore, this information will allow a better interpretation of the normative values of these fitness tests considering the specific measurement error in each age group.

METHOD

PARTICIPANTS

Three public centers of older adults were selected and informed of the study objectives through posters and informative talks. After the first approach, 116 older women between the ages of 60 and 90 became interested in participating. All participants met the inclusion criteria: residing in a community, being functionally independent, and being free of medical conditions, physical or cognitive limitations that would prevent them from following instructions. Participants were divided into five age groups (60-64, 65-69, 70-74, 75-79, and 80 or more). The research team showed the study protocol to the participants, and they were asked to sign the informed consent as a participation requirement. All protocols were adjusted to the Declaration of Helsinki updates. The study was approved by the Biomedical Ethics Committee of the University of Extremadura (Ref: Nº028-15).

FITNESS ASSESSMENT

A five-minute warm-up was performed before taking the measurements, which consisted of neuromuscular activation (joint mobility) and stretching exercises. The total time for the measurements was approximately 45 minutes and was performed throughout the morning, between 9:00 a.m. and 1:00 p.m.

Tests used to evaluate the fitness were:

Muscular strength: Strength was measured by manual dynamometry Rodríguez *et al*., 1998. Both hands were measured using a handheld dynamometer (model TKK 5401, Tokyo, Japan) and the value of both hands summation was considered as a result.

Lower and upper extremities flexibility: the flexibility of the lower extremities was measured using the Chair Sit and Reach test (Jones, Rikli, Max, & Noffal, 1998). This test measures the distance between the middle toe and the foot tip while sitting. The Back Scratch test (Rikli & Jones, 1999a) was used to assess the flexibility of the upper extremities (shoulders). It involves a combination of shoulder abduction, adduction, and internal and external rotation, and measures the distance between (or overlapping) the middle fingers behind the back. The distance between the fingers was scored as a negative value, and the overlapping as a positive one. For calculating flexibility in both tests, the left and right were measured twice, and both averages were calculated.

The agility was assessed using the Timed up-and-go (Podsiadlo, 1991). This test involves getting up from a chair, walking 3 meters to and around a cone, and returning to the chair in the shortest time possible.

Balance: the functional reach test was used to measure the balance (Duncan, Weiner, Chandler, & Studenski, 1990). In this test, the maximum distance one can reach beyond the length of one's arms while keeping feet still is determined.

Aerobic endurance: the 6-minute walking test was applied to measure the aerobic endurance. This test involves determining the maximum walking distance in meters in 6 minutes (Rikli & Jones, 2001).

Additionally, three tests were measured to describe body composition. Waist-hip circumferences, as well as height and weight, were measured to calculate the body mass index (IMC, kg / m^2) and waist-hip ratio (WHR), respectively. Body fat percentage (% BF) was calculated using a portable impedance analyzer according to the manufacturer's instructions (Omron BF306, Omron Healthcare Europe BV, Hoofddorp, The Netherlands) (Deurenberg *et al*., 2001).

PROCESS

Participants were assessed in two measurement sessions one week apart by three evaluators graduated from Sports Sciences and with previous experience in assessing fitness in the elderly. The evaluators received a test manual developed by the project directors describing all the procedures and protocols of each one of the trials; they completed three sessions of four hours each to standardize the evaluation methods and reduce the internal and external error.

In the first session, besides the fitness assessment, a general questionnaire was applied to collect demographic data including age, marital status, educational level, and physical activity level per week. For safety, all participants were evaluated using the Physical Activity Preparation Questionnaire (Thomas, Reading, & Shephard, 1992); also, all participants were evaluated using the Physical Activity Preparation Questionnaire. Those who answered "yes" to any PAR-Q question or who had blood pressure higher than 160/100 mmHg were excluded from the study.

All assessments were conducted in senior centers with a large indoor area, such as a multipurpose room or gym. Participants were instructed to wear appropriate clothing and footwear, eat a light meal approximately 1 hour before the tests, avoid drinking alcoholic beverages within the previous 24 hours, and not engage in vigorous physical activity the day before the evaluation.

Each participant was individually assessed. First, weight, height, waist and hip circumference, and body fat percentage were measured. Afterward, participants did general warm-up exercises before starting the tests and received the same instructions for each test procedure: do the best they could, but never push themselves to the point of excessive exertion or beyond what they thought was safe for them. Participants completed one or two rehearsals to become familiar with the procedures of each test, except for the 6-minute walking one. Tests were handled in the following sequence to minimize the effects of fatigue: Balance (functional reach), Flexibility (chair sit-and-reach and back scratch), Strength (handgrip), y Agility (time up and go). After a five-minute break, the 6-min-walking test was held on a 20-meter circuit.

Statistical Analysis

The mean (±SD) was used. Relative reliability was determined by calculating the Intraclass Correlation Coefficient ($ICC_{2,2}$) and its 95 % confidence interval between the two measured days (Shrout & Fleiss, 1979).

Absolute reliability was determined by calculating the indices *Standard Error Measurement* (SEM) [SEM= SD√(1-ICC) where SD is the standard deviation of day 1 and day 2] and the *Minimum Real Change* (SRD) [SRD= 1.96 x √2xSEM] (Weir, 2005). Both the SEM and the SRD were calculated in absolute terms and as a percentage to facilitate their interpretation.

All analyses were performed separately by age groups: 60-64, 65-69, 70-74, 75-79, and 80 or more. An ANOVA with Games-Howel post-hoc test was performed to analyze the disparities between the different age groups. An effect size (ES) analysis was carried out to analyze the magnitude of the differences in the measurement error obtained between the different age groups. In general, an ES ratio of 0.80 or higher is considered large; an ES of about 0.50 is considered moderate, and an ES of 0.20 or less is considered small (Tomas & Nelson, 2005).

RESULTS

The socio-demographic characteristics of the participants are described in Table 1. 43.1 % of the participants did not receive an education. The percentage of women living alone rises in the older age groups, which is consistent with the increasing proportion of widowed, separated or divorced women. Regarding the level of physical activity, most older women reported more than 3 hours per week.

Table 1. Characteristics of Participants

	All n (%)	60-64 n (%)	65-69 n (%)	70-74 n (%)	75-79 n (%)	≥80 n (%)
Marital status						
Single	4 (3.45)	3 (12.5)	0 (0)	1 (3.3)	0 (0)	0 (0)
Married	74 (63.79)	12 (50)	18 (66.6)	23 (76.7)	15 (62.5)	6 (42.9)
Widow / Separated / Divorced	38 (32.76)	6 (37.5)	9 (33.3)	6 (20)	9 (37.5)	8 (57.1)
Education						
No education	50 (43.10)	9 (42.9)	11 (40.7)	14 (46.7)	7 (30.4)	9 (64.3)
Elementary school	66 (56.90)	12 (57.1)	16 (59.3)	16 (53.3)	17 (69.6)	5 (35.7)
High School	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
University	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Lives with ...						
Alone	34 (29.31)	5 (23.8)	7 (25.9)	7 (23.3)	8 (33.3)	7 (50)
With my partner	74 (63.79)	12 (57.1)	18 (66.7)	23 (76.7)	15 (62.5)	6 (42.9)
With family (brother/sister/son/daughter)	8 (6.90)	4 (19.0)	2 (7.4)	0 (0)	1 (4.2)	1 (7.1)
Physical activity per week						
0 hours/week	12 (10.34)	5 (23.8)	2 (7.4)	2 (6.7)	0 (0)	3 (21.4)
<3 hours/week	13 (11.21)	3 (14.3)	2 (7.4)	5 (16.79)	1 (4.2)	2 (14.3)
≥3 hours/week	91 (78.45)	13 (61.9)	23 (85.2)	23 (76.7)	23 (95.8)	9 (64.3)
(N=116)						

Table 2 shows the ICC, the standard error of measurement and the SRD values obtained in both tests. In the bi-manual dynamometry tests and the 6-minutes walking, the reliability was excellent (ICC> 0.90), while for the other tests it was good (ICC 0.70-0.89). The lowest reliable age group was 75-79, and the best age group was 60-64 years.

Table 2. Reliability of fitness tests (test-retest) according to age group (n=116)

Age group	Test	Mean 1 (±SD)	Mean 2 (±SD)	ICC (95% CI)	SEM	%SEM	SRD	%SRD
All (n=116)								
	Bi-manual handgrip (Kg)	41.2±7.9	41.3±8.3	0.93 (0,90 to 0.95)	2.2	5.2	6.0	14.6
	Chair Seat-and-reach (cm)	7.1±10.6	7±10.5	0.88 (0.83 to 0.92)	3.0	30.3	8.4	84.0
	Functional reach (cm)	28.7±6.8	29.5±6.4	0.82 (0.76 to 0.88)	3.2	10.9	8.8	30.2
	Timed up-and-go (s)	6.9±1.3	6.9±1.3	0.87 (0.82 to 0.91)	0.5	6.8	1.3	18.8
	6-minute walking (m)	454.1±79.8	455.7±77.8	0.95 (0.93 to 0.96)	17.8	3.9	49.4	10.8
60-64 (n=21)								
		Mean 1 (±SD)	Mean 2 (±SD)	ICC (95% CI)	SEM	%SEM	SRD	%SRD
	Bi-manual handgrip (Kg)	45.2±7.4	45.2±8.2	0.89 (0.72 to 0.95)	2.64	5.84	7.32	16.20
	Chair Seat-and-reach (cm)	7.5±10.4	6.6±11.7	0.93 (0.83 to 0.97)	2.44	25.80	6.77	71.52
	Functional reach (cm)	29.5±5.8	30.1±6.3	0.75 (0.46 to 0.89)	3.56	11.94	9.88	33.11
	Timed up-and-go (s)	6.3±1.3	6.4±1.1	0.92 (0.81 to 0.97)	0.35	5.53	0.97	15.33
	6-minute walking (m)	476.9±66.8	480.2±69.4	0.95 (0.89 to 0.98)	14.93	3.12	41.37	8.64
65-69 (n=27)								
		Mean 1 (±SD)	Mean 2 (±SD)	ICC (95% CI)	SEM	%SEM	SRD	%SRD
	Bi-manual handgrip (Kg)	40.3±4.5	40.6±7.7	0.92 (0.82 to 0.96)	1.77	4.38	4.91	12.14
	Chair Seat-and-reach (cm)	6.6±12.7	8.1±10.5	0.87 (0.73 to 0.93)	2.94	25.18	8.14	69.80
	Functional reach (cm)	29.1±4.2	30.4±4.7	0.71 (0.46 to 0.86)	2.39	8.03	6.62	22.25
	Timed up-and-go (s)	6.8±1.3	6.7±1.2	0.85 (0.70 to 0.93)	0.47	7.01	1.31	19.43
	6-minute walking (m)	463.9±83.6	461.5±74.7	0.93 (0.86 to 0.97)	20.64	4.46	57.21	12.36
70-74 (n=30)								
		Mean 1 (±SD)	Mean 2 (±SD)	ICC (95% CI)	SEM	%SEM	SRD	%SRD
	Bi-manual handgrip (Kg)	41.5±6.8	42.2±7.6	0.95 (0.89 to 0.98)	1.66	3.96	4.59	10.97
	Chair Seat-and-reach (cm)	11.4±8.8	10.8±9.3	0.86 (0.72 to 0.93)	3.97	34.64	11.01	96.00
	Functional reach (cm)	29.5±7.2	30.5±6.6	0.87 (0.73 to 0.94)	3.23	10.76	8.95	29.84
	Timed up-and-go (s)	6.7±1.2	6.7±1.3	0.88 (0.76 to 0.94)	0.44	6.63	1.22	18.39
	6-minute walking (m)	473.7±65.7	472±66.1	0.94 (0.87 to 0.97)	16.81	3.55	46.58	9.85
75-79 (n=24)								
		Mean 1 (±SD)	Mean 2 (±SD)	ICC (95% CI)	SEM	%SEM	SRD	%SRD
	Bi-manual handgrip (Kg)	41.1±9.6	40.4±9.4	0.94 (0.86 to 0.97)	2.37	5.82	6.57	16.13
	Chair Seat-and-reach (cm)	5.3±7.5	5.3±7.7	0.85 (0.69 to 0.93)	2.48	33.51	6.88	92.87

	Functional reach (cm)	29.1±7.7	29.4±6.4	0.88 (0.74 to 0.95)	3.10	10.60	8.58	29.38
	Timed up-and-go (s)	7.1±1.0	7.2±1.0	0.82 (0.62 to 0.92)	0.64	8.94	1.77	24.79
	6-minute walking (m)	455.4±52.0	460.4±58.0	0.87 (0.71 to 0.94)	20.20	4.41	56.00	12.23
≥80 (n=14)		Mean 1 (±SD)	Mean 2 (±SD)	ICC (95% CI)	SEM	%SEM	SRD	%SRD
	Bi-manual handgrip (Kg)	36.6±6.3	36.4±7.4	0.93 (0.78 to 0.98)	1.82	4.98	5.04	13.81
	Chair Seat-and-reach (cm)	1.1±12.5	0.7±12.9	0.85 (0.60 to 0.95)	2.66	28.81	7.38	79.86
	Functional reach (cm)	24.4±8.5	24.5±7.3	0.82 (0.52 to 0.94)	4.03	16.52	11.18	45.78
	Timed up-and-go (s)	8.1±1.3	8.2±1.3	0.92 (0.77 to 0.97)	0.37	4.48	1.02	12.42
	6-minute walking (m)	356.4±94.4	365±93.7	0.98 (0.93 to 0.99)	13.95	3.87	38.67	10.72

SD: Standard Deviation; ICC: Intraclass Correlation Coefficient; CI: 95% Confidence Interval; SEM: Standard Error Measurement, SRD: Minimum Real Change

Table 3. Differences between test and retest according to age group (n=116). ANOVA analysis with Games-Howel post-hoc.

Test	All (n=116)	60-64 (n=21)	65-69 (n=27)	70-74 (n=30)	75-79 (n=24)	≥80 (n=14)	F	P
Bi-manual dynamometry (Kg)	0.2±0.3	0.0±0.7	0.3±3.2	0.8±0.7	-0.8±-0.2	-0.2±1.1	.479	.751
Chair Seat-and-reach (cm)	0.0±-0.1	-0.8±1.3	1.5±-2.2	-0.6±0.5	0.0±0.2	-0.4±0.4	1.167	.330
Functional reach (cm)	0.8±-0.4	0.6±0.5	1.4±0.5	1.0±-0.6	0.3±-1.4	0.1±-1.1	.303	.875
Timed up-and-go (s)	0.0±0.0	0.0±-0.2	-0.1±-0.1	0.0±0.1	0.0±0.0	0.1±0.1	.219	.928
6-minute walking (m)	1.7±-2.0	3.3±2.6	-2.4±-8.9	-1.7±0.3	5.0±6.0	8.6±-0.7	.688	.602

Values as mean ± standard deviation
There were no significant differences between the groups using the Games-Howell post hoc test.

Table 3 shows the mean difference between the test measurement and the retest for all age groups. Statistically significant differences in age were not obtained.

Table 4 shows the ES of test-retest differences between age groups. All ES were low, except for the time up-and-go test in which it was moderate.

Table 4. Size of error difference effect obtained by test-retest between age groups.

	Age group	65-69 (n=27)	70-74 (n=30)	75-79 (n=24)	≥80 (n=14)
Bi-manual dynamometry (Kg)	60-64 (n=21)	0.2	1.6	-3.2	-0.3
	65-69 (n=27)		0.3	-0.6	-0.2
	70-74 (n=30)			-3.6	-1.1
	75-79 (n=24)				0.9
Chair Seat-and-reach (cm)	60-64 (n=21)	1.3	0.2	1.1	0.5
	65-69 (n=27)		-1.6	-1.3	-1.5
	70-74 (n=30)			1.7	0.4
	75-79 (n=24)				-1.3
	60-64 (n=27)	-0.2	0.9	0.2	1.2
Functional reach (cm)	60-64 (n=21)	0.3	0.1	-0.1	-0.2
	65-69 (n=27)		-0.7	-1.2	-1.6
	70-74 (n=30)			-0.7	-1.1
	75-79 (n=24)				-0.2
Timed up-and-go (s)	60-64 (n=21)	-0.7	0.0	0.0	0.7
	65-69 (n=27)		1.0	2.0	2.0
	70-74 (n=30)			0.0	1.0
	75-79 (n=24)				2.0
6-minute walking (m)	60-64 (n=21)	-1.0	-2.8	0.4	3.2
	65-69 (n=27)		0.4	1.0	2.3
	70-74 (n=30)			1.8	18.6
	75-79 (n=24)				0.1

DISCUSSION

Overall, the reliability of tests in different age groups was good, and the ANOVA analysis showed that aging does not affect the values of *test* and *retest*.

The fitness gradually decreases with age, accelerating progressively after age 60 (Milanović *et al*., 2013). Many studies provide normative values of fitness and how these values decline as people age (Rikli & Jones, 2013; Sardinha *et al*., 2015; Vagetti *et al*., 2015), however there are few precedents from studies analyzing the reliability of fitness instruments in the elderly (Dewhurst & Bampouras, 2014; Steffen, Hacker, & Mollinger, 2002). Furthermore, none of them analyze whether the reliability values of these tests are modified according to the age of the people evaluated. The results of this study do not show significant differences based on age in the measurement error calculated by *test-retest*.

Regarding both the *test and retest* application, previous studies have analyzed the reliability of several fitness tests performed at different times on the same day. This study shows that the 6-minute walking speed test, chair sit and reach test and functional reach have excellent reliability. Since the current study assesses reliability between one week separated sessions, it seems reasonable that the reliability reported in this previous study is higher than the one obtained in the latter.

For calculating reliability, the CCI value was used as it is a widely used index. However, interpretation can be problematic given the existence of diverse literature with different threshold uses. In this case, the classification proposed by Munro *et al*. was used. (Munro, Visintainer, & Page, 1986) It considers as moderate reliability values between 0.50 and 0.69, high values between 0.70 and 0.89, and excellent values higher than 0.90.

Results show that the reliability for all age groups evaluated was high in the sit-and-reach, functional reach, and timed up-and-go tests following this classification; while the one obtained in bi-manual dynamometry and 6-minutes walking was excellent. These results may indicate that all the tests assessed are highly reliable in all age groups. When analyzing differences according to age, the functional reach test is the one with the most changes in its reliability among the analyzed groups (CCI 0.71-0.88). The ANOVA analysis did not show any statistically significant differences according to age in any of the tests assessed. As for the differences in effect size between the different age groups, all tests have a low size except for the time up-and-go test which has a moderate size. This result indicates that there are no relevant changes by age group regarding the magnitude of the mean difference obtained between the *test* and the *retest*.

Studies such as this one focused on analyzing the reliability of evaluation instruments, help professionals who use these tests in their daily work. Specifically, knowing the reliability data shown in this study contributes to understanding the accuracy in evaluating tests that may be used to classify the risk of low physical function in the elderly (Merellano-Navarro *et al*., 2017; Rikli & Jones, 2013; Sardinha *et al*., 2015). It is also useful to know the minimum necessary difference to achieve in therapies aimed at improving the fitness of this population to ensure that the effect found is greater than the error to be assumed during measurement. Additionally, in research settings, an essential use of the reliability parameters of the tests to be used is the estimation of the sample size in experimental studies (Hopkins, 2000).

LIMITATIONS

The sample in this study does not consider older men, which means that the results cannot be generalized to this population. Another limitation to bear in mind is that participants from the older groups are few, as well as that all participants are self-valid elderly living in the community; therefore the results do not represent those with functional capacity problems.

CONCLUSIONS

The fitness tests analyzed in this study show good reliability in older women. The *test-retest* analysis by age group showed that the reliability in this population is not affected by the increasing age.

REFERENCES

Chang, Y., Wang, J., Chen, H., & Hu, S. (2017). Aerobic-synergized exercises may improve fall-related physical fitness in older adults. *The Journal of sports medicine and physical fitness, 57*(5), 660-669. https://doi.org/10.23736/s0022-4707.16.05728-5

Chen, H., Lin, C., & Yu, L. (2009). Normative physical fitness scores for community-dwelling older adults. *J Nurs Res, 17*(1), 30-41. https://doi.org/10.1097/JNR.0b013e3181999d4c

Chen, T., Tseng, W., Huang, G., Chen, H., Tseng, K., & Nosaka, K. (2017). Superior Effects of Eccentric to Concentric Knee Extensor Resistance Training on Physical Fitness, Insulin Sensitivity and Lipid Profiles of Elderly Men. *Frontiers in physiology, 8*, 209. https://doi.org/10.3389/fphys.2017.00209

Chung, P., Zhao, Y., Liu, J., & Quach, B. (2017). A canonical correlation analysis on the relationship between functional fitness and health-related quality of life in older adults. *Archives of gerontology and geriatrics, 68*, 44-48. https://doi.org/10.1016/j.archger.2016.08.007

Deurenberg, P., Andreoli, A., Borg, P., Kukkonen-Harjula, K., de Lorenzo, A., van Marken Lichtenbelt, W. D., . . . Vollaard, N. (2001). The validity of predicted body fat percentage from body mass index and from impedance in samples of five European populations. *Eur J Clin Nutr, 55*(11), 973-979. https://doi.org/10.1038/sj.ejcn.1601254

Dewhurst, S., & Bampouras, T. (2014). Intraday reliability and sensitivity of four functional ability tests in older women. *American journal of physical medicine & rehabilitation, 93*(8), 703-707. https://doi.org/10.1097/PHM.0000000000000078

Duncan, P., Weiner, D., Chandler, J., & Studenski, S. (1990). Functional reach: a new clinical measure of balance. *Journal of gerontology, 45*(6), M192-197. https://doi.org/10.1093/geronj/45.6.M192

Gusi, N., Hernandez-Mocholi, M. A., & Olivares, P. R. (2015). Changes in HRQoL after 12 months of exercise linked to primary care are associated with fitness effects in older adults. *European journal of public health, 25*(5), 873-879. https://doi.org/10.1093/eurpub/ckv079

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Med, 30*(1), 1-15. https://doi.org/10.2165/00007256-200030010-00001

Jones, C. J., Rikli, R. E., Max, J., & Noffal, G. (1998). The reliability and validity of a chair sit-and-reach test as a measure of hamstring flexibility in older adults. *Res Q Exerc Sport, 69*(4), 338-343. https://doi.org/10.1080/02701367.1998.10607708

Krause, M. P., Januário, R. S. B., Hallage, T., Haile, L., Miculis, C. P., Gama, M. P. R., . . . da Silva, S. G. (2009). A comparison of functional fitness of older Brazilian and American women. *J Aging Phys Act, 17*(387-397). https://doi.org/10.1123/japa.17.4.387

Merellano-Navarro, E., Collado-Mateo, D., García-Rubio, J., Gusi, N., & Olivares, P. R. (2017). Criterion-Referenced Fitness Standards associated with maintaining functional capacity in Chilean Older Adults. *Rejuvenation Research*. https://doi.org/10.1089/rej.2016.1913

Milanović, Z., Pantelić, S., Trajković, N., Sporiš, G., Kostić, R., & James, N. (2013). Age-related decrease in physical activity and functional fitness among elderly men and women. *Clinical interventions in aging, 8*, 549. https://doi.org/10.2147/CIA.S44112

Munro, B. H., Visintainer, M. A., & Page, E. B. (1986). *Statistical methods for health care research*. Philadelphia: JB Lippincott.

Oja, P. (1995). *Eurofit for Adults: Assessment of Health-Related Fitness.* Tampere, Finland: Council of Europe Publishing.

Olivares, P., Gusi, N., Prieto, J., & Hernandez-Mocholi, M. (2011). Fitness and health-related quality of life dimensions in community-dwelling middle aged and older adults. *Health and quality of life outcomes, 9*, 117. https://doi.org/10.1186/1477-7525-9-117

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*, 142-148. https://doi.org/10.1111/j.1532-5415.1991.tb01616.x

Rikli, R., & Jones, C. (1999a). Development and validation of a functional fitness test for community-residing older adults. *J Aging Phys Act, 7*(2), 129-161. https://doi.org/10.1123/japa.7.2.129

Rikli, R., & Jones, C. (1999b). Functional fitness normative scores for community-residing older adutls, ages 60-94. *J Aging Phys Act, 7*(3), 60-179. https://doi.org/10.1123/japa.7.2.162

Rikli, R., & Jones, C. (2001). *Senior Fitness Test Manual* Champaign, IL.

Rikli, R., & Jones, C. (2013). Development and validation of criterion-referenced clinically relevant fitness standards for maintaining physical independence in later years. *The Gerontologist, 53*(2), 255-267. https://doi.org/10.1093/geront/gns071

Rodríguez, F. A., Gusi, N., Valenzuela, A., Nácher, S., Nogués, J., & Marina, M. (1998). [Evaluation of health-related fitness in adults (I): background and protocols of the AFISAL-INEFC battery]. *Apunts. Educ. Fís. Deportes, 52*, 54-75.

Rodriguez, F. A., Valenzuela, A., Gusi, N., Nacher, S., & Gallardo, I. (1998). [Evaluation of health-related fitness in adults (II): reliability, applicability and normative values of the AFISAL-INEFC battery]. *Apunts. Educ. Fís. Deportes, 54*, 54-65.

Sardinha, L., Santos, D., Marques, E., & Mota, J. (2015). Criterion-referenced fitness standards for predicting physical independence into later life. *Experimental gerontology, 61*, 142-146. https://doi.org/10.1016/j.exger.2014.12.012

Shaulis, D., Golding, L., & Tandy, R. (1994). Reliability of the AAHPERD functional fitness assessment across multiple practice sessions in older men an women. *J Aging Phys Act, 2*(3), 273-279. https://doi.org/10.1123/japa.2.3.273

Shrout, P., & Fleiss, J. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin, 86*(2), 420-428. https://doi.org/10.1037/0033-2909.86.2.420

Steffen, T. M., Hacker, T. A., & Mollinger, L. (2002). Age-and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. *Physical therapy, 82*(2), 128-137. https://doi.org/10.1093/ptj/82.2.128

Thomas, S., Reading, J., & Shephard, R. J. (1992). Revision of the Physical Activity Readiness Questionnaire (PAR-Q). *Can J Sport Sci, 17*(4), 338-345.

Tomas, J., & Nelson, J. (2005). *Research Methods in Physical Activity* (5 ed.). Champaign, IL: Human Kinetics.

Vagetti, G. C., Barbosa Filho, V. C., de Oliveira, V., Mazzardo, O., Moreira, N. B., Gomes, A. C., & de Campos, W. (2015). Functional fitness in older women from southern brazil: normative scores and comparison with different countries. *Brazilian Journal of Kinanthropometry and Human Performance, 17*(4), 472-484. https://doi.org/10.5007/1980-0037.2015v17n4p472

Weir, J. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of strength and conditioning research, 19*(1), 231-240. https://doi.org/10.1519/15184.1

Número de citas totales / Total references: 31(100%)
Número de citas propias de la revista / Journal's own references: 0 (0%)