Reproducibility of a functional assessment test of the dynamic balance and agility of elderly people

Adriana Angarita Fonseca[1], Claudia Lucía Ariza García[2], Milena del Carmen Boneth Collante[3], Javier Parra Patiño[4], Rosa Nathaly Rojas Santisteban[5], Yurley Tatiana Angarita Carrascal[6]

SUMMARY

Background: The *8 foot up & go* test assesses the dynamic balance and agility in elderly people. Its reproducibility has been evaluated in American population, but it is unknown whether it would work similarly in a different population like the Colombian.

Objective: To evaluate the test-retest reliability and agreement level of the *8 foot up & go* test in a sample of older adults from Bucaramanga, Colombia.

Materials and methods: An evaluation of diagnostic tests was done in 114 elderly individuals. In the analysis, we assessed the test-retest reliability of the *8 foot up & go* test by the Intraclass Correlation Coefficient (ICC 2.1) with their respective confidence intervals at 95% (95% CI). The agreement level was established by the Bland-Altman method.

Results: The test-retest reliability of the *8 foot up & go* test was very good (ICC: 0.98; 95% CI: 0.98- 0.99). The agreement was good in females (mean difference [MD] = 0.04 seconds and limits of agreement [LA]: -1.27; 1.36 seconds), and in elderly institutionalized (MD = 0.04 seconds [LA]: -3.18; 3.27 seconds).

Conclusion: The *8 foot up & go* test has very good reliability and good agreement in Colombian local elderly population.

KEY WORDS

Elderly; 8 foot up & go test; Postural Balance; Reliability

[1] Physical Therapist, Epidemiology Advisor, Universidad de Santander, Bucaramanga, Colombia.
[2] Physical Therapist, Professor, Universidad de Santander, Bucaramanga, Colombia.
[3] Physical Therapist, Physical Therapy Internships Coordinator, Universidad de Santander (UDES), Bucaramanga, Colombia.
[4] Physician, Universidad Industrial de Santander, Bucaramanga, Colombia.
[5] Physical Therapist, Fundación El Principio de una Esperanza, Cúcuta, Colombia.
[6] Physical Therapist, Coorpoces Espíritu Santo. ESE Imsalud and IPS Karisalud, Cúcuta, Colombia.
Contact person: Adriana Angarita Fonseca; adriangarita@hotmail.com

Received: 10 March 2013
Accepted: 19 November 2013

INTRODUCTION

The 8 foot up & go test assesses agility and dynamic balance. Agility is a physical quality, defined by Sheppard and Young as a rapid movement of the entire body, changing speed or direction as a response to a stimulus. It is related to other physical qualities that can be trained, such as strength, power and technique, as well as cognitive components, such as the speed of visual exploration and anticipation (1).

Balance is the ability to maintain the body's vertical centre of gravity on the base of support (2). This concept is important in terms of maintaining a static posture, which is the resistance of the body to the destabilising influence of gravity (static balance) and to respond to the active stimuli, regardless whether they are internal or external (dynamic balance) (2). It is worth noting that keeping balance integrates the information from visual, vestibular, somatosensory and neuromuscular systems.

In elderly population it is possible to see some physiological changes, triggered by ageing, that affect the balance: the visual field and acuity, deep perception and proprioception decrease; the number of hair cells villous cells of the vestibular system and neurons decrease; the type of muscular fibre changes; the reaction time of the sensitive component increases, among other (4-6). In summary, with age, the functional deterioration of all the systems that participate in maintaining the balance can be seen, with the aggravating factor that, in spite of these changes, healthy elderly people have to be able to maintain their balance when performing their daily activities (5).

So far, there is no field test to assess the agility and the static and dynamic component of balance (7); however, falls in the elderly occur while they move when walking, turning, and going up and down stairs, so it is necessary to assess dynamic balance (8,9). In order to do so, there are several tests, among which is the *8 foot & go* test, which is a modified version of *Timed Get up & go test*, in which potentially destabilising activities, such as getting up from a chair, walking, turning, going back and sitting again (10), are carried out.

Even though the reproducibility of 8 foot up & go test in American population has been assessed and intraclass correlation coefficients (ICC) have been found to be above 0.94 (10,11), it is not known whether this test works in similar ways in another population, such as the Colombian. For this reason, it was put forward assessing the test-retest reproducibility and the level of agreement between measurements of 8 foot up & go test in the elderly population of the city of Bucaramanga.

MATERIALS AND METHODS

An assessment study of diagnostic tests was performed in a population of elderly people in Bucaramanga, Colombia, classified by convenience sampling and that were affiliated to Fundación Albeiro Vargas and Ángeles Custodios (n = 20), Asilo San Rafael (n = 44), Grupo de Adultos Mayores del Café Madrid (n = 45) and Fundación Hogar Geriátrico Luz de Esperanza (n = 7). The three first were from the northern commune and the fourth of the central commune of Bucaramanga. The three institutions are non-profit entities that provide care for the elderly population that do not receive any help from the State. Grupo de Adultos Mayores is a group that meets to mainly carry out recreational and sport activities.

People over 60, volunteers, functionally independent, without any risk of getting sick when doing aerobic exercises or authorised by a doctor in case they have positive risk screening, were included. People with flu or a common cold at the time the test was made were excluded, as well as those using prosthetics or orthopaedic devices to move, or with evidence of any organic disorder, cardiovascular or cardio-respiratory condition putting their health at risk, revealed before the test, and also, people that were not affiliated to the social security system.

The information was collected between January and May 2011. The study variables were age, gender, education, marital status, socioeconomic strata, personal background, dynamic balance and agility (measured by means of 8 foot up & go tests), and institutionalisation. The elderly people with domicile at an institution were considered institutionalised and the one that received day care but does not sleep there was considered not-institutionalised.

Procedure

In the first place, the institutions from where the population of elderly people were going to be selected for assessment were identified, and then there was a meeting to inform about the research and the elderly people were motivated to participate. Then there was the informed consent formality and queries about sociodemographic aspects and personal background. Also, the screening to determine risk of getting sick due to physical exercise was carried out by means of questionnaire PAR-Q&YOU (12). The participants who responded affirmatively to some of the questions of said questionnaire were referred to the doctor.

Assessments were made after a 5-minute warming-up, between 7 and 11 in the morning, at roofed premises of the institution. The participants were convened between 4 and 8 days after the first assessments for the second part of the test, under the same conditions of the first time. The same person assessed the elderly people both times, but the records from the first assessment were filed in such a way so that the second time the tester did not have the previous results. The tester tester was a physiotherapy senior year student, previously trained and supervised by the principal investigator. Also, a chronometer and a wooden ruler, duly calibrated, were used.

Pilot study

Seven elderly people, who live in the municipality of Girón, 45 minutes away from Bucaramanga, were tested. As a result of the pilot study, the exercises to be made and the warming-up times were unified. The need to use a whistle to indicate the beginning of the test was observed.

Protocol to be applied

Based on the protocol proposed by Rikli and Jones in Senior Fitness test battery, a 42.5 cm (17 inches) chair was placed leaning firmly against the wall; opposite to it, a reflective cone exactly 2.44 metres (8 feet), measured from the front of the chair, was placed. A 1.22 metres (4 feet) space was left beyond the cone to allow the participant to surround it back to the chair (10).

The tester explained the test, made a demonstration and allowed one trial to ensure its correct execution. Then the participant, sitting on the chair, with their hands on their thighs and their feet on the floor, one slightly in front of the other, got up from the chair at the sound of the whistle, walked towards the cone to surround it and then went back to sit. The time spent from the whistle signal to sitting again was timed (in seconds and tenths of seconds). In order to interpret the test it is necessary to bear in mind that the less dynamic balance and agility the greater the time performing it. Also, it is important to clarify that, unlike the original Rikli and Jones protocol (10), in this study the data obtained from just one test were analysed, only one trial was made before the measurement and a whistle was used.

Statistical analysis

Central tendency (average or median) and dispersion (standard deviation or interquartile range, IQR) were applied, as well as position or percentage measures according to the nature and distribution of the variables. Their normality was measured using the Shapiro Wilk test. The comparison of the time needed for 8 foot up & go test between the first and the second measurements was made by applying the Wilcoxon signed-rank test for paired data; the comparison of the time needed for the test between men and women and between institutionalised and non-institutionalised individuals was made by applying the Two-sample Wilcoxon rank-sum (Mann-Whitney) test.

The test-retest reproducibility was assessed by using the intra-class correlation coefficient (ICC [2.1]) (13) and their respective confidence intervals of 95% (CI95%); the agreement level between the first and the second assessments was established by applying the Bland and Altman methodology (14), with which the average difference between measurements and the upper (UL) and lower (LL) agreement of 95% limits are obtained.

The ICC was interpreted by way of the Altman classification: poor reproducibility (≤ 0.20); acceptable (0.21-0.40); moderate (0.41-0.60); good (0.61-0.80) and very good (0.81-1.00). The database was subjected to double typing and validated in the Epidata 3.1. Programme. The different statistic calculations and the processing of variables were made in the STATA/SE v. 11.0. Programme. An alpha level of <0.05 was considered for all the tests.

According to Resolution 8430 of 1993 of the Ministry of Health of Colombia, this study was considered free of risk (16). A written consent was requested from the

participants and the study was approved by the Research Ethics Committee of Universidad de Santander in Bucaramanga.

RESULTS

Characteristics of the population

Full information for 114 out of the 116 individuals who entered the study was obtained; 70 of them (61.4%) were women. The socio-demographic characteristics are summarised in table 1. Statistically significant differences were detected per gender ($p < 0.05$). Men had a greater age average (74.0 ± 7.4 years versus 68.4 ± 6.6 years in women; $p = 0.0001$); a greater percentage without a permanent partner (81.8% versus 64.3%; $p = 0.044$) and of socioeconomic stratum one (97.7% versus 80,0%; $p = 0.006$). Also, the percentage of men with respiratory diseases was greater than women's (20.5% versus 4.3%; $p = 0.006$) (table 1).

Table 1. Characteristics of the population subject to the study (n = 114)

Variable		Female n = 70 (61.4%)	Male n = 44 (38.6%)	Total n = 114 (100%)
Age *(DE)*		68.4 ± 66	74.0 ± 7.4	70.6 ± 7.4*
Education n (%)	None	48 (68.6)	31 (73.8)	79 (70.5)
	Primary	22 (31.4)	11 (26.2)	33 (29.5)
With permanent partner n (%)	No	45 (64.3)	36 (81.8)	81 (71.0)*
	Yes	25 (35.7)	8 (18.2)	33 (29.0)
Socioeconomic stratum n (%)	One	56 (80.0)	43 (97.7)	99 (86.8)*
	Two	14 (20.0)	1 (2.3)	15 (13.2)
Personal background n (%)	Heart disease	1 (1.4)	3 (6.8)	4 (3.5)
	Musculoskeletal alteration	6 (8.6)	6 (13.6)	12 (10.5)
	Respiratory disease	3 (4.3)	9 (20.5)	12 (10.5)*
	Vision alterations	37 (52.9)	24 (54.6)	61 (53.5)
	Smoking	1 (1.4)	2 (4.6)	3 (2.6)

[a] DE: desviación estándar

[b] $p < 0,05$

Table 2. Time (seconds) spent in *8 foot up & go test*

Variable		n	First assessment (E1)[a]	Second assessment (E2)[a]	p[b]
Gender	Female	70	7.1 (6.5-8.3)	7.1 (6.5-8.1)	0.5503
	Male	44	7.0 (6.4-8.1)	6.9 (6.0-7.6)	0.0020
Institutionalised	Yes	47	7.1 (6.4-7.8)	7.0 (6.2-7.6)	0.0094
	No	67	7.1 (6.5-85)	7.0 (6.4-8.2)	0.3024
Global		114	7.1 (6.5-8.3)	70 (6.3-8.0)	0.0128

[a]Median (Interquartilic Range); [b]Wilcoxon signed-rank test for paired data

Time spent on testing

The time spent on testing was statistically longer in the first assessment, in comparison with the second, both in the global population and in the sub-groups: males and institutionalised adults. Upon comparing the assessment per sex and institutionalisation, no statistically significant differences were found.

Reproducibility test-retest

The test had very good reproducibility, with an ICC of 0.98 (CI95%: 0.98-0.99). The reproducibility was greater in men and in institutionalised individuals, with ICC, in both cases, of 0.99 (CI95%: 0.98-0.99) (table 3).

Table 3. Reproducibility test-retest and level of agreement between the first and the second assessment of the *8 foot up & go*

	Variables	n	ICC (CI95%)	B&A (LL-UL)
Gender	Females	70	0.91 (0.86- 0.94)	0.04 (-1.27; 1.36)
	Males	44	0.99 (0.98- 0.99)	0.07 (-3.27; 3.41)
Institutionalised	Yes	47	0.99 (0.98-0.99)	0.04 (-3.18; 3.27)
	No	67	0.91 (0.86-0.94)	0.06 (-1.29; 1.41)
Global		114	0.98 (0.98- 0.99)	0.05 (-2.25; 2.35)

ICC = Intra-class B&A correlation coefficient: Bland and Altman analysis. LL= Lower limit. UL= Upper limit

Level of agreement between the first and the second assessments

Bland and Altman agreement analysis between the first and the second test assessments (table 3) showed an average of the differences close to zero (0.053) and close agreement limits, upper and lower, of -2.25 and 2.35. Also, it was observed that in men and institutionalised individuals (figure 1), although the average of the differences was close to zero, the agreement limits were broader compared to women and non-institutionalised individuals (figure 2).

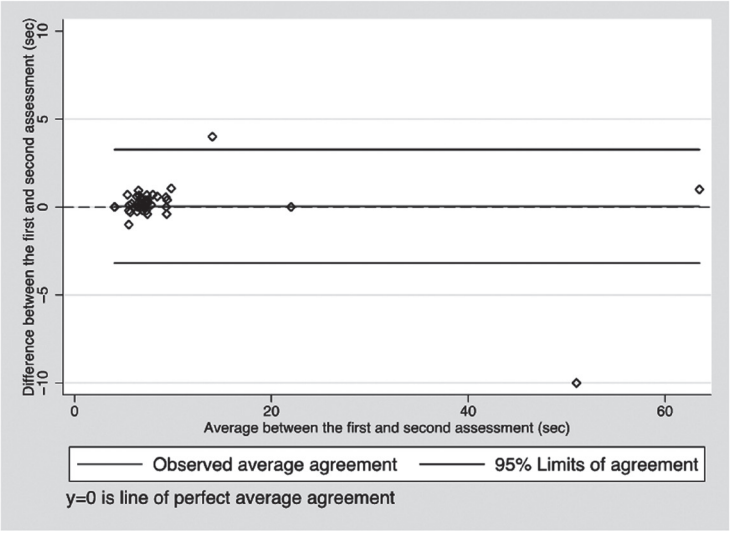

Figure 1. Level of agreement between the first and the second assessment of the *8 foot up & go* test in the institutionalised elderly people

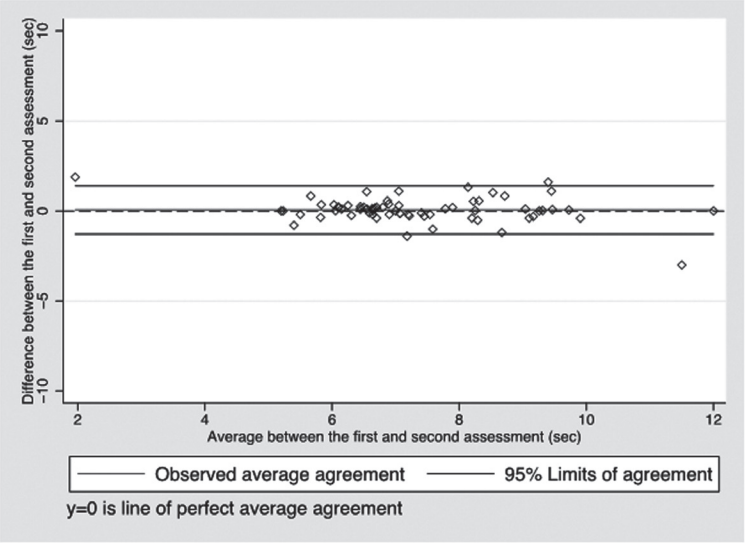

Figure 2. Level of agreement between the first and the second assessment of the *8 foot up & go* test in the non-institutionalised elderly people

DISCUSSION

This study assessed the reproducibility of the 8 foot up & go test in a population of elderly people of Bucaramanga. The main findings were: good level of agreement and very good reproducibility.

The Bland & Altman analysis showed an average of the differences close to zero (0.053), which suggests that this test may detect small changes in the functionality of the elderly people that depend on dynamic balance. As a complement, the lower (-2.25) and upper (2.35) agreement limits show that an adult that takes 20 seconds in the initial assessment, needs to take more than 22.4 seconds in the second assessment to consider that there was a real deterioration in the dynamic balance. When performing the analysis per subgroups (institutionalised *versus* non-institutionalised and men *versus* women), a lower level of agreement was observed in institutionalised individuals and in men, as the analysis of Bland and Altman indicates that a difference of 3.3 seconds between the first and the second assessments is required to consider that there was indeed a change, whilst in women and in non-institutionalised individuals a change of approximately 1.4 seconds is required.

The favourable results obtained in this research are probably due to the reduction of variability. First the variability of the tester was reduced, as only one previously trained person applied the test on the participants; in contrast, in the Rikly and Jones (10) study several testers participated and obtained very good reproducibility. It is worth highlighting that, in clinical practice and in population research, more than one tester participates.

On the other hand, the results obtained in the subgroup analysis can be explained because institutionalised elderly people are characterised for having low levels of physical activity and, also, they have a more deteriorated health condition, less strength and less muscular resistance (17), take more medicines and are more likely to get sick (18). Also, the men in this group were older and had more respiratory diseases than women (table 1). These multi-systemic alterations, that may be included within one fragility condition, were probably the cause case of more variability in the results.

Another reason for the variability may be the test itself. In order to complete the test, apart from the static and dynamic balance, other physical and physiologi-

cal qualities are required: muscular strength, flexibility and intra and intermuscular co-ordination to assume the standing position from the sitting position; flexibility, speed, muscular strength and co-ordination to accelerate and decelerate when walking 2.44 metres, as well as the proprioceptive sensitivity and space orientation to go back to the initial position.

If it is understood that each of said activity requires multiple physiological processes that are performed in different manners, depending on the previous health, physical and functional conditions, there will be a greater probability of an elderly person doing the test with differences in their performance as long as they execute it more than one time. In this respect, the use of a whistle may increase the possibility of the participant to start the test in a standardised manner in both assessments, reducing part of the variability resulting from the test.

One of the elderly people assessed had a difference of -10 seconds between the first and the second assessments (figure 1) and it was considered an outlier, because the rest of the values varied between -3 and 4. This can be explained by the intra-subject variability; this institutionalised elderly person, who in fact showed a poor dynamic balance because he took an average of 50 seconds to do the test, was able get to the second assessment in different conditions than the first one, which made him take 10 seconds more to perform it.

The reproducibility results (ICC = 0.98 [CI95%: 0.98-0.99]) agree with Rikli and Jones (10)study, authors of *Senior Fitness* battery of test, who found, in 84 elderly people (71.8 ± 6.9 years) a global ICC of 0.95 (CI95%: 0.92-0.97). They also agree with the results of Miotto et al (11), who, in 79 physically active or sedentary elderly people (68.4 ± 5.5 years) found very good reproducibility, with an ICC of 0.94. In the analysis per sub-group, contrary to what was found in the Bland & Altman analysis, the ICC was greater in men and in institutionalised individuals. This may be accounted for by the influence of heterogeneity of the data about the ICC: the greater the dispersion of the data the greater the ICC (19). In this study, the duration of the test for men and institutionalised individuals varied between 4.1 and 64 seconds, whilst in women and non-institutionalised subjects, it varied between 1 and 13 seconds (data not shown).

It is important to consider that, in this study, the data obtained from a single test in each assessment were analysed in order to reduce the period necessary to collect information; whilst in the original Rikli and Jones (10) study, the shorter time of two tests was taken, and Miotto et al (11) worked with the less time spent in three tests. The test was submitted to a previous trial in all the studies. In the measurement theory it has been established that the use of averages is more reproducible, but Rikly and Jones (10) suggested using the best score as the results were highly reproducible and this saves time when recording the data. With the results of this study the reader has to consider the use of only one measurement; the one with the best score of two or three measurements, or the one of their average, highlighting that no reproducibility studies using the average of several measurements of 8 foot up & go test have been reported.

Regarding the number of assessments, in this study two were made with 4 to 8 days difference. In turn, Miotto et al (11) made three assessments within two weeks. In both studies, the results of the first and second assessments were similar; however, Miotto et al (11) found that, in the third assessment, the time average was lower compared to the time of the second one. The possible explanation is that, in the last assessment, the participants were able to make an extra effort once they felt comfortable with the test and they were sure they would not slip or fall; another plausible explanation is that, in the two first assessments, they could develop balance strategies that led to a more efficient and quick performance of the test during the last assessment, which could be interpreted as the effects of learning and practice may influence this test (11).

In this sample, the duration median of the test was 7.1 (IQR = 6.5-8.3) and 7.0 (IQR = 6.3-8.0) seconds in the first and the second assessments; these results are high compared with those of Rikli and Jones (10) and Miotto et al (11). The first found that the time was 5.2 ± 0.6 for elderly people between 60 and 69 years old, and 7.1 ± 2.0 seconds for those between 80 and 89 years old (10). In the second work, the times were 4.96 ± 1.02 seconds for those that were physically active, and 5.71 ± 1.01 seconds for the sedentary ones (10). The results of this study showed that women tend to have better dynamic balance than men; on the

contrary, the normative data of the USA show that the duration of the test is, average, greater in women (6.2 ± 1.9 seconds) than men (5.6 ± 1.8 seconds) (20). Also, it was observed that the women of this study spend more time performing the test compared to a population of women of Brazil (5.2 ± 1.2 seconds for women between 60 and 64 years old and 7.2 ± 2.2 seconds for those between 80 and 84 years old) (21). Unfortunately, there are no reference data to compare in Colombia. For this reason, it could be thought that the population of elderly people from Colombia has lower functionality limits than those of the populations of other countries.

It was found that, among the limitations of this study, other variables that may affect the performance of the test, such as obesity or regular exercise were not assessed.

CONCLUSIONS

The 8 foot up & go test has very good reproducibility. The level of agreement shows that 3.4 seconds are needed by men and institutionalised individuals, and 1.4 seconds by women and non-institutionalised individuals to observe actual changes in the dynamic balance of 60-year old elderly people who are functionally independent. It is demonstrated that 8 Foot up & go tests can be used as a field test; also, it is easy to apply, without the need of using many supplies.

REFERENCIAS BIBLIOGRÁFICAS

1. Sheppard JM, Young WB. Agility literature review: classifications, training and testing. J Sports Sci. 2006;24(9):919–32.
2. Nashner LM. Practical biomechanics and physiology of balance. In: Jacobson G P, Newman CW, Kartush JM. Handbook of balance function testing. Saint Louis: Mosby Year Book; 1993. p. 261–79.
3. Berg K. Balance and its measure in the elderly: a review. Physiother Can. 1989; 41(5): 240-6.
4. Nikolić M, Vranid TS, Arbanas J, Cvijanović O, Bajek S. Muscle loss in elderly. Coll Antropol. 2010;34 Suppl 2:105–8.
5. Sturnieks DL, St George R, Lord SR. Balance disorders in the elderly. Neurophysiol Clin. 2008;38(6):467–78.
6. Laughton CA, Slavin M, Katdare K, Nolan L, Bean JF, Kerrigan DC, et al. Aging, muscle activity, and balance control: physiologic changes associated with balance impairment. Gait Posture. 2003;18(2):101–8.
7. Mancini M, Horak FB. The relevance of clinical balance assessment tools to differentiate balance deficits. Eur J Phys Rehabil Med. 2010;46(2):239–48.
8. Yasumura S, Haga H, Niino N. Circumstances of injurious falls leading to medical care among elderly people living in a rural community. Arch Gerontol Geriatr.1996; 23(2):95–109.
9. Bleijlevens MHC, Diederiks JPM, Hendriks MRC, van Haastregt JCM, Crebolder HFJM, van Eijk JTM. Relationship between location and activity in injurious falls: an exploratory study. BMC Geriatr. 2010;10:40.
10. Rikli RE, Jones CJ. Development and Validation of a Functional Fitness Test for Community-Residing Older Adults. J Aging Phys Act.1999; 7(2):129-61.
11. Miotto JM, Chodzko-Zajko WJ, Reich JL, Supler MM. Reliability and validity of the Fullerton Functional Fitness: An independent replication study. J Aging Phys Act. 1999;7(4):339-53.
12. Adams R. Revised Physical Activity Readiness Questionnaire. Can Fam Physician. 1999;45(2):992, 995, 1004–5.
13. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979 ;86(2):420–8.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307–10.
15. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
16. RESOLUCIÓN Nº 008430 de 1993, por la cual se establecen las normas científicas, técnicas y administrativas para la investigación en salud. Ministerio de Salud, Decreto 2164 de 1992 y la Ley 10 de 1990, (4-10-1993)
17. Król-Zielińska M, Kusy K, Zieliński J, Osiński W. Physical activity and functional fitness in institutionalized vs. independently living elderly: a comparison of 70-80-year-old city-dwellers. Arch Gerontol Geriatr. 2011; 53(1):e10–6.

18. Boneth M, Ariza CL, Angarita A, Parra J, Monsalve A, Gómez E. Reproducibilidad de las pruebas ArmCurl y Chair Stand para evaluar resistencia muscular en población adulta mayor. Rev Cienc Salud 2012; 10 (2):179-193.

19. Keating J, Matyas T. Unreliable inferences from reliable measurements. Aust J Physiother. 1998;44(1):5–10.

20. Rikli RE, Jones CJ. Functional Fitness normative scores for community-residing older adults, ages 60–94. J Aging Phys Act. 1999;7(2):162–81.

21. Krause MP, Januário RSB, Hallage T, Haile L, Miculis CP, Gama MPR, et al. A comparison of functional fitness of older Brazilian and American women. J Aging Phys Act. 2009;17(4):387–97.