International Journal of Biomedical and Advance Research
ISSN: 2229-3809 (Online); 2455-0558 (Print)
Journal DOI: 10.7439/ijbar
CODEN: IJBABN

Original Research Article

Assessment of inter-rater variability of the Senior Fitness Test in the geriatric population: A community based study

Prasanta Kumar Bhattacharya[1*], Kuldeep Deka[2] and Aakash Roy[1]

[1]*Department of General Medicine, North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences, India*
[2]*Department of Physiotherapy, Down Town Hospital, G. S. Road, Dispur, Guwahati, Assam 781036 India*

***Correspondence Info:**
Dr. Prasanta Kumar Bhattacharya
Professor & Head,
Department of General Medicine,
North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences, India
E-mail: pkbdr78@gmail.com

Abstract

Background: Maintenance of physical activity is of paramount importance in geriatric care and it is imperative to be able to objectively assess physical activity level in the geriatric population. However, majority of the established assessment protocols are designed for relatively younger adults, and are not validated for the elderly. The Senior Fitness Test (SFT) is aimed at assessing the underlying functional fitness for older adults.

Methods: A community based prospective study to assess the test-retest reliability of SFT in elderly subjects, aged 65 – 75 years. Subjects with any acute illnesses, acute musculoskeletal injury, hypertension, and with visual problems hampering test performance were excluded. The study was conducted by two observers on two days at the subject's own locality. The Intraclass Correlation Coefficient (ICC) with 95% confidence intervals (CIs) was used as a measure of relative reliability. The standard error of measurement and minimal detectable change (MDC) were used to measure absolute reliability.

Results: Thirty one participants (mean age- 68.32±2.58 years) were assessed. The ICC reflected high reliability (0.934–0.994) in all SFT items. MDC values (90% CI) for various SFT were as follows: chair sit and reach test = 1.55, back scratch test = 1.41, 30 second chair stand test = 1.15, 30-second arm curl test = 2.53, 8 foot up-and-go test = 1.47 and 2-minute step test = 1.40. Bland–Altman plots for all the tests were positively skewed and heteroscedastic.

Conclusion: The SFT showed high test–retest reliability and thus is suitable for evaluating physical fitness in elderly and can be utilized in larger studies for the geriatric population.

Keywords: Inter-rater reliability, physical fitness, elderly, community.

1. Introduction

Maintenance of physical activity and fitness is one of the issues of paramount importance in the field of clinical geriatrics care so that the elderly individuals can live independently without depending on others. A sedentary lifestyle can generate substantial losses in physical activity and predispose to the development of various disease conditions. In this regard it becomes important to be able to assess the level of physical activity in the geriatric population in order to ensure an optimum degree of healthy lifestyle.

Over the years researchers have found it problematic to obtain a comprehensive tool for physiological assessment of 'immediate level of fitness' in older adults. Majority of the established protocols are primarily designed for younger adults, and are not validated for older adults [1]. Two important tests used in adults above the age of 60 years in the United States of America are the American Alliance for Health, Physical Education, Recreation and Dance (AAPHERD) Functional Fitness Assessment for Adults [2] and the Established Populations for Epidemiologic Studies of the Elderly (EPESE) assessment tests [3]. Other researchers have established similar tests in Japan and Netherlands with regard to the geriatric population [4,5].

However, the Senior Fitness Test (SFT) developed by Rikli and Jones may be the best at assessing the underlying functional fitness for older adults [1]. The test includes six components of functional fitness including lower and upper body strength, aerobic endurance, lower and upper body flexibility, and agility/dynamic balance. Each test component of the SFT has been tested for its high content validity, criterion validity, construct validity, and reliability [6].

In this context our study was planned to estimate the relative and absolute test-retest reliability of the SFT for the Indian geriatric population. Therefore, this community based study was undertaken to determine the inter-rater reliability of SFT in the geriatric population aged between 65 to 75 years, living in Guwahati, the largest city in northeast India.

2. Materials and Methods

A community based prospective, non-experimental study was undertaken to assess the test-retest reliability of SFT in elderly subjects, aged 65 – 75, selected by a non-random, convenient sampling method from Guwahati, the largest city in the northeast of India, during the period from February 2015 to July 2015, based on the following selection criteria:

2.1 Inclusion criteria

Older adults (aged between 65 to 75 years) of either gender who had sufficient language skill to understand and respond to instruction of the test were included in the study.

2.2 Exclusion criteria

People with acute illnesses, acute or unstable musculoskeletal injury, elevated blood pressure, vision problems hampering mobility or test performance were excluded from the study.

Appropriate ethical clearance was taken from the institutional ethics committee of Gauhati Medical College, and written informed consent was obtained from all subjects prior to enrolment in the study. Upon satisfying the inclusion and exclusion criteria a total of 31 participants were included for the study.

The study was carried out by two physiotherapists who administered the test as per the standardised instructions of the test manual of senior fitness test for older adult by Rikli and Jones [1]. The methodology of the individual components of the SFT has been elaborated inn Table 1.

Table 1: Test methodology for Senior Fitness Test [1]

Assessment category	Test item	Test description
Lower body flexibility	Chair sit-and-reach test	To be seated in such a way that he/she should come at the front edge of the seat while keeping their legs extended with ankle dorsiflexed. Secondly to reach the toes of extended legs by their hands and the measurement taken with a ruler (in centimetres) by examiner between the participant finger and the tips of the toes.
Upper body flexibility	Back scratch test	Hands to be brought towards the back (one hand from above the shoulder and other hand from middle of back) and to touch each hand. Measurement taken in centimetres by examiner between the extended middle fingers of the participants.
Lower body strength	30-second chair stand test	Number of full stands in 30 second with arms folded across chest
Upper body strength	30-second arm curl test	Number of bicep curls in 30 second holding hand weight (women 5 pound; men 8 pound)
Agility/dynamic balance	8-foot up-and-go test	Number of seconds required to get up from seated position, walk 8 foot, turn, and return to seated position on chair
Aerobic endurance	2-minute step test	Number of full steps completed in 2 minutes, raising each knee to point midway between patella and iliac crest

Each test of the SFT was first demonstrated to the participants and if necessary, cues or gestures were provided. All participants performed the 6 physical tests in the same sequence at baseline and at the second session scheduled 2-3 days later. Test scores were taken three times and the average of the all three readings was taken into consideration. The two physiotherapists scored the test performance independently without knowledge of each other rating (blind) on two days at the subject's own place of locality. The participants were assessed for weight and height and body mass index (BMI) was calculated subsequently. Complete data was collected in pre-formed proformas and entered in Microsoft Windows Excel Sheet for statistical analysis.

2.3 Statistical Analysis

The data were analyzed using Statistical Package for Social Survey (SPSS) for Windows, version 16.0. Demographic data were expressed in terms of mean and standard deviation. Mean and standard deviation were also calculated for the individual tests. Relative test-retest reliability was calculated with the intra class correlation (ICC). The ICC was calculated using the 2-way, random, absolute agreement on single measures model with a 95% CI (confidence interval) [7-10]. We also considered their absolute reliability, which we calculated using the Bland-Altman 95% limits of agreement (LoA) and Standard error of measurement (SEM). The following formula was used:

SEM =sd $x\sqrt{(1-\text{ICC})}$ (sd= standard deviation)

The 95% CIs for the SEM were calculated as described by using the following formula:

$$\text{SEM}= \frac{SSE}{\chi^2\alpha,errordf}, \frac{SSE}{\chi^2 1-\alpha,errordf}$$

(SSE= the sum of squared errors in the analysis of variance (ANOVA), x^2 *a, error df*=the chi-square value for probability level *a*; and error *df*=the degrees of freedom of the SSE provided in the ANOVA table).

Finally, to be able to interpret changes in test scores, the minimal detectable change (MDC) with 95% CI was calculated using the formula:

$MDC_{95} = SEM \times Z_{95}$ (1.96) x √2.

The MDC is the required magnitude of observable change that exceeds the anticipated measurement error and within-subject variability. In other words, if a participant's

score exceeds the value of the MDC, it can be said to reflect a true change in performance with 95% CI.

For a visual inspection of the similarity between the two measurements, Bland-Altman plots were created with the LoA (Limits of agreement).

3. Results

A total of 31 elderly participants comprising of 21 males and 10 females with a mean age of 68.66 ± 2.35 yrs and 67.60 ± 3.02 years respectively, were assessed for the physical fitness tests using the SFT. The overall mean age of the study population was 68.32 ± 2.58 yrs (Table 2).

Table 2: Demographic characteristics of the subjects

Sex	Number of participants	Age (years)	
		Range	Mean ± SD
Male	21	65–73	68.66 ± 2.35
Female	10	65–74	67.60 ± 3.02
Total	31	65–74	68.32 ± 2.58

The inter-rater variability between observations for all the physical fitness tests was calculated with reference to the individual maxima and minima and mean scores (Table 3).

Table 3: Inter-rater variability of Senior Fitness Test in the subjects

Parameter (units)	n	Variable	Test result		
			Minimum value	Maximum value	Mean ± SD
Chair sit and reach test (in centimetres)	31	Tester 1	10	30	15.71±5.45
		Tester 2	10	31	15.84±5.40
Back scratch test(in centimetres)	31	Tester 1	5	34	15.34±7.29
		Tester 2	5	33	15.26±7.26
30 Second Chair stand test (Number of times)	31	Tester 1	7	15	10.74±1.69
		Tester 2	8	16	10.94±1.97
30 second Arm curl test(Number of times)	31	Tester 1	5	15	11.16±2.35
		Tester 2	5	16	11.65±2.23
8 foot up and go test(Number of Seconds)	31	Tester 1	7.89	17.84	11.26±2.41
		Tester 2	7.55	18	11.44±2.50
2 minutes step test (Number of times)	31	Tester 1	48	80	64.13±9.69
		Tester 2	49	83	64.45±9.64

N= number of participants; Tester 1=Testing of participants by investigator 1 for inter-rater reliability; Tester 2=Testing of participants by investigator 2 for inter-rater reliability; SD=standard deviation.

The Intra-class co-relation (ICC) of the physical fitness tests were calculated at the 95% confidence interval (CI) and levels of significance were assessed, the results of which are shown in Table 4.

Table 4: Intra-class correlation of Senior Fitness Test in the subjects

Test		95% Confidence interval			F test with true value 0			
		Intra-class Correlation	Lower Bound	Upper Bound	Value	Df1	Df2	Sig
Chair sit and reach test (in centimetres)	Average measures	0.991	0.983	0.996	114.845	33.0	33	0
Back scratch test (in centimetres)	Average measures	0.994	0.988	0.977	170.929	33.0	33	0
30 Second Chair stand test (Number of times)	Average measures	0.934	0.864	0.968	15.249	30.0	30	0
30 second Arm curl test (Number of times)	Average measures	0.943	0.883	0.973	17.688	30.0	30	0
8 foot up and go test (Number of Seconds)	Average measures	0.972	0.943	0.986	35.108	33.0	33	0
2 minutes step test (Number of times)	Average measures	0.991	0.982	0.996	112.609	33.0	33	0

Df1= degree of freedom for test by first investigator; Df2 = degree of freedom for test by second investigator; Sig =significance

The relative reliability of the entire test were highly significant (ICC >0.90). The width of the CI of ICC ranged between 0.04 and 0.104. The 8 foot up and go test with smallest CI value and 30-second chair stand test had the largest CI value.

The absolute reliability was measured for all the tests. The SEM and MDC values for all tests are shown in Table 5.

Table 5: Reliability measures* of the Senior Fitness Test in the subjects

	R-square	KT	F value	Pr > F	ICC	SEM	MDC
Chair sit and reach test (in centimetres)	0.967	29.977	849.83	<.0001	0.991(C)	0.562357	1.558775
Back scratch test(in centimetres)	0.9774	30.2994	1256.41	<.0001	0.994(c)	0.512289	1.419993
30 Second Chair stand test (Number of times)	0.7863	24.3753	106.7	<.0001	0.934(c)	0.41833	1.159552
30 second Arm curl test(Number of times)	0.7998	24.7938	115.83	<.0001	0.943(c)	0.914531	2.53495
8 foot up and go test(Number of Seconds)	0.906	28.086	279.47	<.0001	0.972(c)	0.532405	1.475752
2 minutes step test (Number of times)	0.8762	27.1622	205.23	<.0001	0.991(c)	0.506102	1.402844

*R-square = co-efficient of determination implying overall goodness of fit; KT = Koenker test for heteroscedasticity; Pr > F = p value associated with the F statistic of the given test; MSE = mean squared error; ICC = intra-class correlation; SEM=standard error of mean; MDC = minimal detectable change. *Calculation over log-transferred data*

Bland-Altman plots were calculated for the 6 tests for the total group (Figures 1-6). The data of the entire test were positively skewed and heteroscedastic, with higher means yielding higher variability, as is reflected by the wider LoAs.

Figure 1: Chair sit and reach test in the elderly subjects

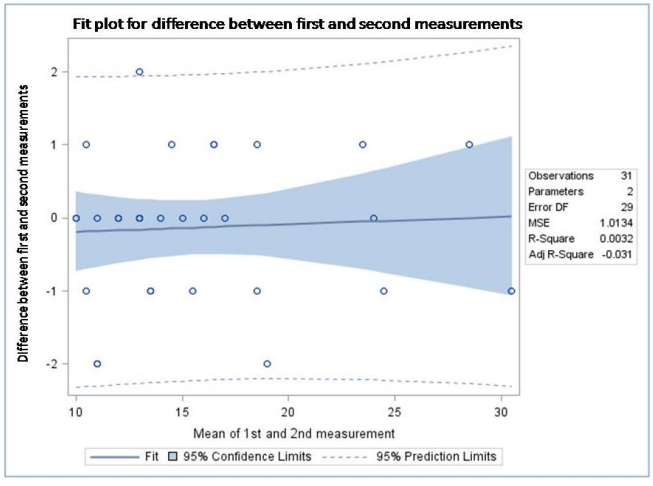


Figure 2: Back scratch test in the elderly subjects

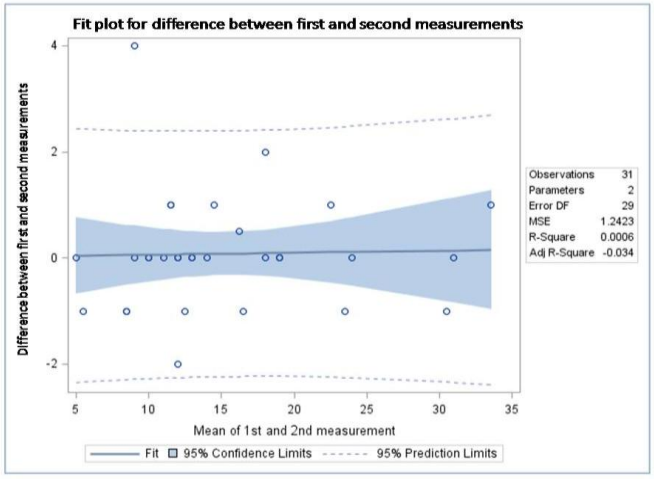


Figure 3: 30-second chair stand test in the elderly subjects

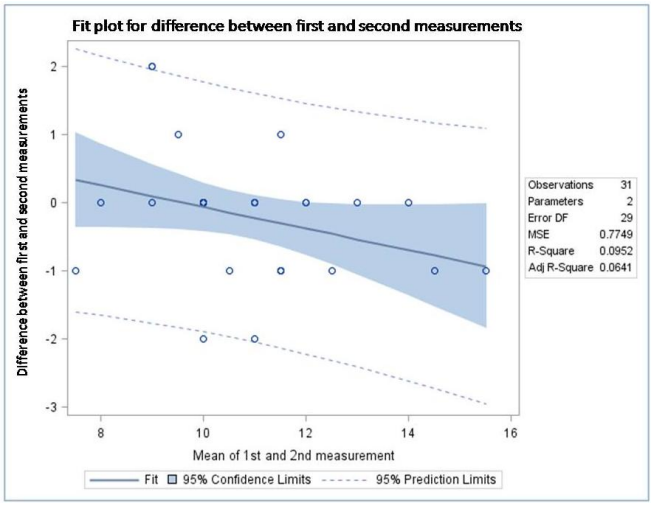


Figure 4: 30-second Arm curl test in the elderly subjects

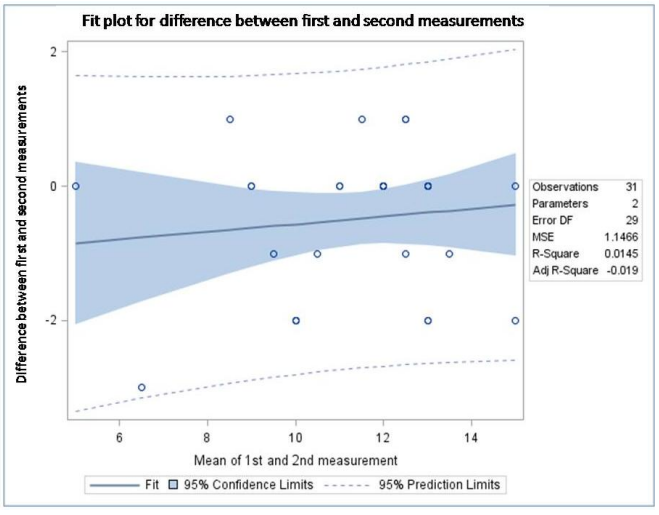


Figure 5: 8-foot up and go test in the elderly subjects

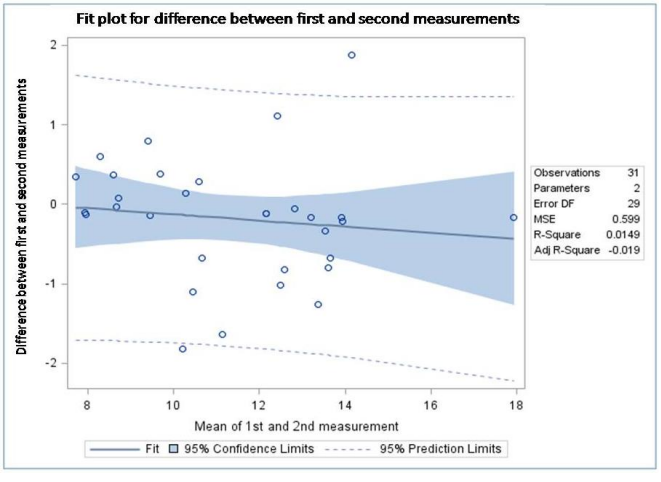
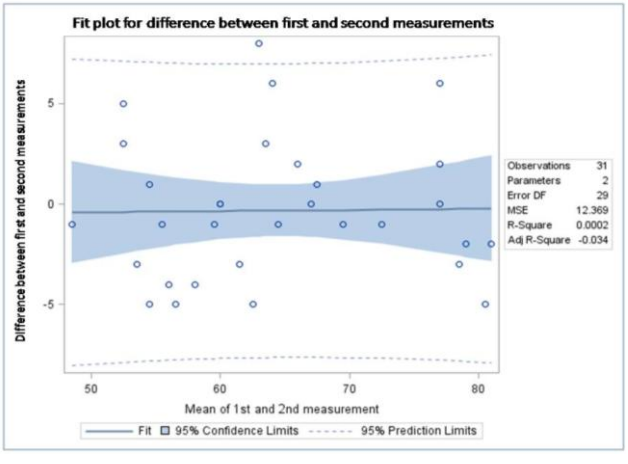


Figure 6: 2-minute step test in the elderly subjects



4. Discussion

General fitness establishes the quality of life, but it often decreases with age. Even chronological age does not correlate perfectly with functional age i.e. two person may be of same age, but differ in their mental and physical capacities. Therefore, it becomes important to find out the individual fitness level through appropriate assessment in the geriatric population. But the testing procedure needs to be simple, comprehensive and easy to administer. The Senior Fitness Test perhaps is one of the simplest of such fitness tests but

its implementation in fitness testing for clinical purpose in developing countries is not known [1,6]. Therefore, the main goal of our study was to evaluate the relative and absolute inter-rater reliability in the elderly (65-74 years) population in this region.

Relative reliability:

All the component tests of SFT were excellent as the testing procedures were very simple and the elderly participants could understand the test with very few instruction and verbal cues. Even the therapist did not require a formal training process for administrating the test. So the overall agreements of the tests as evidenced by the high ICC values and reliability scores were well established during the course of our study. The ICC values of the 30-second chair stand test and 30-second arm curl test were 0.934 and 0.943 respectively. The rest of the fitness tests showed ICC of more than 0.97. This minor variability in case of the 30-second chair stand test and the 30-second arm curl test in comparison to the other tests may be because of the fact that strength generation by muscle may not remain same at all the time and may be subject to change depending upon individual physical capacities. The ICC value for chair sit and reach test and back scratch test were 0.991 and 0.994 respectively. This may be due to the reason that muscle length variation does not occur on a shorter duration. All the tests were able to show a high degree of reliability in patients who had associated co-morbidities including cognitive impairment, fibromyalgia and post-opertaive states, thus indicating its potential use for both clinical and research purposes throughout the entire spectrum of the geriatric population [11].

Absolute reliability:

Estimation of the precision of its clinical outcome can be obtained with absolute reliability testing. The SEM and MDC are relatively easy measures to interpret. The MDC is based on SEM but is more conservative (equivalent to 2.7 SEMs). If the score change is larger than the MDC, this difference is not caused by measurement error or participant variability [8].

All the other components of the SFT showed low SEM with low variability in SEM and low values of MDC. The 30-second chair stand test showed a much lower SEM value of 0.41 and very low value of MDC on comparison to other test. The chances of minimal error is quite possible with this testing procedure as variation on muscle strength does not occur drastically over days unless there is some specified neurological, metabolic or nutritional causes . Even studies on subjects with total knee replacement 30-second chair stand test showed high to very high reliability indicating the test seems to be reliable and feasible in both disease and non-disease state [12].

5. Conclusion

Our study was aimed to assess the inter-rater reliability of Senior Fitness Test in the geriatric population in an urban community setting so that the test can be readily used to the geriatric community of India as a whole. The inter-rater reliability of the SFT was excellent and these can perhaps be utilized in larger studies for the geriatric population. Though this study was conducted in a comparatively small sample size, which is a limitation of the study, the results of the study indicate a promising prospective to assess the functionality of the test in a larger population in the community.

References

[1] Rikli RE, Jones CJ. Development and validation of a functional fitness test for community-residing older adults. *Journal of aging and physical activity*.1999; 7:129-61.

[2] Osness WH. Functional Fitness Assessment for Adults Over 60 Years (A Field Based Assessment). AAHPERD Publication Sales Office, 1900 Association Drive, Reston, VA 22091. 1990.

[3] Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al.A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *Journal of gerontology*. 1994; 49(2):M85-94.

[4] Tanaka K, Kim H, Yang H, Shimamoto H, Kokudo S, Nishijima T. Index of assessing functional status in elderly Japanese men. *Applied Human Science.* 1995; 14(2):65-71

[5] Van Heuvelen MJ, Kempen GI, Ormel JO, Rispens PI. Physical fitness related to age and physical activity in older persons. *Medicine and Science in Sports and Exercise*. 1998; 30(3):434-41

[6] Roberta E.. Rikli, Jones CJ. Senior fitness test manual. Human Kinetics; 2001.

[7] Bruton A, Conway JH, Holgate ST. Reliability: what is it, and how is it measured? *Physiotherapy*. 2000; 2:94 – 99.

[8] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; 1:307–310.

[9] Domholdt E. Rehabilitation Research: Principles and Applications. 3rd ed. St Louis, MO: Elsevier Saunders; 2005.

[10] Flechner L, Tseng TY. Understanding results: P-values, confidence intervals, and number need to treat. *Indian J Urol.* 2011; 27:532–535

[11] Hesseberg K, Bentzen H and Bergland A .Reliability of the Senior Fitness Test in Community-dwelling Older People with Cognitive Impairment. *Physiotherapy Research International* 2015 March; 20(1): 37–44.

[12] Unver B, Kalkan S, Yuksel E, Kahraman T, Karatosun V. Reliability of the 50 foot walk test and 30 second chair stand test in total knee arthoplasty. *Acta Ortop Bras* 2015; 23(4):184–7.