Research Article

# A community-based study to assess test–retest reliability of senior fitness test in the geriatric population in a northeastern Indian city

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

[1]Department of General Medicine, North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences, Shillong, Meghalaya, India.
[2]Department of Physiotherapy, Down Town Hospital, Guwahati, Assam, India.
Correspondence to: Prasanta Kumar Bhattacharya, E-mail: pkbdr78@gmail.com

Received November 21, 2015. Accepted December 2, 2015

## Abstract

**Background:** Maintenance of physical fitness is a cornerstone of geriatric preventive care. Senior fitness test (SFT) is a simple test for assessing physical fitness of the geriatric population. However, its reliability in the Indian setting has not been well established.

**Objective:** To determine the relative and absolute test–retest reliability of SFT as a physical fitness tool of the geriatric population in the community.

**Materials and Methods:** A community-based prospective, nonexperimental study was carried out in an urban community of Guwahati, Assam, India. Thirty-one physically active, elderly individuals (aged ≥65 and <75 years) were asked to perform the SFT, at baseline and at a second session after 2–3 days, to assess the reliability of the test. Intra-rater variability and intraclass correlation coefficients (ICC) of the SFT were assessed. Reliability of the tests was measured using the standard error of measurement (SEM), minimal detectable change (MDC), and log-transferred limits of agreement of Bland–Altman plots.

**Result:** The relative reliability of entire physical fitness test (SFT) was shown to be excellent. The ICC at the 95% confidence interval (CI) for all the tests ranged between 0.933 and 1.000. The *F* test values were significant for all the tests (0.00). The width of the CI of ICC ranged between 0.0 and 0.107 with body mass index having the smallest CI (0.0) and arm curl test (0.1.07) having the largest CI. The SEM and MDC values were small ranging from 0.43 to 0.0 and from 0.0 to 1.8 for all tests except for the 2-min step test. The SEM and MDC for the 2-min step test was 1.48 and 4.12, respectively. Bland–Altman plots for all the tests were positively skewed and heteroscedastic.

**Conclusion:** Test–retest reliability of the physical fitness test was excellent and these tests were thus applicable for cross-sectional and controlled interventional studies for the elderly population.

**KEY WORDS:** Community based, geriatric, physical fitness tests, senior fitness test, intra-rater variability, reliability

| Access this article online | |
|---|---|
| Website: http://www.ijmsph.com | Quick Response Code: |
| DOI: 10.5455/ijmsph.2016.21112015262 | |

## Introduction

The proportion of the population aged 65 or above is expected to rise globally from 12.4% in 2000 to close to 20.7% by 2050.[1] The challenge in preventive geriatric care is that physical activity decreases with age and inactive older people are at a high risk of metabolic diseases.[2] Physical activity and exercise are key interventions in improving physical function

and maintaining independence in older adult. Most protocols for determining level of physical fitness are designed for the younger population and are not appropriate for the older population, generally referred to those above 65 years of age.[3] Hence, it is difficult to make any physiological assessment of the baseline level of fitness in the elderly population, as well as the subsequent changes that occur as their age increases.

Although some fitness tests such as AAPHERD[4] (American Alliance for Health, Physical Education, Recreation, and Dance), EPESE[5] (Established Populations for Epidemiologic Studies of the Elderly), and SFT[6] (Senior Fitness Test) are used to assess fitness of elderly people, the SFT developed by Rikli and Jones[6] is one of the simple and best tools in assessing six important functional fitness parameters for older people, namely body composition, lower and upper body strength, aerobic endurance, lower and upper body flexibility, and agility/dynamic balance. Each test component of the SFT has been selected for its high reliability in a fitness facility or large community facility.[6–8] Our study was planned with the aim of finding out the relative and absolute test–retest reliability of the SFT for the older people in the context of the Indian population where, to the best of our knowledge, no such assessment has been performed. Therefore, this community-based study was undertaken to determine the intra-rater reliability of SFT in older people, aged between 65 and 75 years, living in Guwahati, the largest city in the northeast India. Based on the result of this study, this test can be further used in the community.

## Materials and Methods

A community-based prospective, nonexperimental study was undertaken to assess the test–retest reliability of SFT in elderly subjects, aged ≥65 and <75 years, selected by a nonrandom, convenient sampling method from Guwahati, the largest city in the northeast of India, during the period from January 2015 to June 2015, based on the following selection criteria:

1. Inclusion criteria: Older people (≥65 to <75 years) of either gender who had sufficient language skills to understand and to respond to the instruction of the test were included in the study.
2. Exclusion criteria: People with acute illnesses, acute or unstable musculoskeletal injury, elevated blood pressure, and vision problems hampering mobility or test performance were excluded from the study.

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

SFT comprises the following six components: (1) arm curl test (for upper body strength), (2) chair stand test (for lower body strength), (3) back scratch test (for upper body flexibility), (4) chair sit and reach test (for lower body flexibility), (5) eight foot up and go test (for agility/dynamic balance), and (6) 2-min step test (for aerobic endurance). Each test of the SFT was first demonstrated to the participants and if necessary, cues or gestures were provided. All participants performed the six 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 participants were assessed for weight and height, and body mass index (BMI) was calculated subsequently. Complete data were collected in preformed pro forma and entered in Microsoft Windows Excel Sheet for statistical analysis.

### Methodology of the Senior Fitness Test

1. Arm curl test: Upper body strength was measured using the arm curl test, which required the participants to perform biceps curls in 30 s holding hand weights (women: 5 pound, men: 8 pound dumbbells).
2. Chair stand test: Lower body strength was measured using the chair stand test, which required the participants to perform full stands from a seated position in 30 s with arms folded across the chest.
3. Chair sit and reach test: Lower body flexibility was assessed by chair sit and reach test where the participants were asked to be seated on a chair in such a way that he or she should come at the front edge of the seat while keeping their legs extended with ankle dorsiflexed. Second, they should reach the toes of extended legs by their hands and the measurement was taken with a ruler (in centimeters) by the examiner between the participant finger and the tips of the toes.
4. Back scratch test: Upper body flexibility using back scratch test was assessed by asking the participants to bring their hands toward the back (one hand from above the shoulder and other hand from middle of back) and asking them to touch each hand. Measurement was taken in centimeters by the examiner between the extended middle fingers of the participants.
5. Eight foot up and go test: Agility and dynamic balance was tested by the eight foot up and go test, which assessed the number of seconds required to get up from a seated position, walk 8 feet, turn, and return to seated position.
6. Two-minute step test: Aerobic endurance was tested by the 2-min step test in which the number of full steps completed in 2 min, raising each knee to midway between patella and iliac crest was counted.

### Statistical Analysis

The data were analyzed using Statistical Package for Social Survey 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 intraclass correlation coefficients (ICC). The ICC was calculated using the two-way, random, absolute agreement on single measures model with a 95% confidence interval

(CI).[9–12] 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:

$$SEM = SDx\sqrt{(1-ICC)} \qquad (SD = \text{standard deviation})$$

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

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

(SSE = the sum of squared errors in the analysis of variance [ANOVA]; $\chi^2$ *a* = 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:

$$MDC_{95} = SEM \times Z_{95}(1.96) \times \sqrt{2}$$

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.

## Result

A total of 31 elderly participants comprising 21 men and 10 women with a mean age of 68.67 ± 2.35 and 67.60 ± 3.03 years, respectively, were assessed for the physical fitness tests using the SFT. The overall mean age of the study population was 68.32 ± 2.59 years [Table 1]. The intra-rater variability between observations for all the physical fitness tests was calculated with reference to the individual maxima and minima and mean scores [Table 2]. The ICC of the seven physical fitness tests was calculated at the 95% CI and levels of significance were assessed, the results of which are shown in Table 3. The relative reliability of all the tests was highly significant (ICC >0.90). The width of the CI of ICC ranged between 0 and 0.107 with the BMI having the smallest CI and arm curl test having the largest CI. The absolute reliability was measured for all the tests. The SEM and MDC values for all tests except the 2-min step test were small ranging from 0.43 to 0.0 and from 0 to1.8, as shown in Table 4. The SEM and MDC for the 2-min step test were 1.48 and 4.12, respectively. Bland–Altman plots were calculated for the seven tests for the total group (Figures 1–7). The data of the entire test were positively skewed and heteroscedastic, with higher mean yielding higher variability, as is reflected by the wider LoAs.

**Table 1:** Demographic characteristics of the elderly subjects

| Sex | Number of participants | Age (years) | |
|---|---|---|---|
| | | Range | Mean ± SD |
| Male | 21 | 65–73 | 68.67 ± 2.35 |
| Female | 10 | 65–74 | 67.60 ± 3.03 |
| Total | 31 | 65–74 | 68.32 ± 2.59 |

SD, standard deviation.

## Discussion

Aging is a slow but dynamic process, which involves many internal and external influences, including genetic programming, physical, and social environment.[13] Physical activity including activities of daily living refers to bodily movement involving muscle contractions and increased energy expenditure whereas exercise refers to planned, structured, and repetitive movements aimed at improving various components of physical fitness and encompasses strength, flexibility, balance, and endurance.[14] Protocols for fitness tests are mostly designed for the younger population and not appropriate for the geriatric population. Though there are a few fitness tests available for the geriatric population, the SFT is perhaps one of the simplest and easiest to administer.

The main goal of our study was to evaluate the relative and absolute reliability of the SFT in the elderly (65–74 years) population of this region.

### Relative Reliability

The result of all physical fitness tests of SFT was excellent, as the testing procedure was very simple, easy, structured, and required very less instrument. Even the tests were comprehensive and usable in the field settings and did not require training by the therapist for administering the tests. So the chances of high reliability were well established. For the arm curl test and chair stand test, ICC values were 0.933 and 0.934, respectively. The rest all physical fitness tests showed ICC more than 0.97. This may be due to the fact that the generation of body strength by a person may not remain same all the time and it may vary based on psychological issues of day-to-day variation of a person. All these fitness tests were studied in diseased population such as fibromyalgia, post total knee arthroplasty, and dementia, where it was highly significant in terms of reliability.[15,16] As the entire test exceeded the threshold of high reliability, they do seem suitable for use in cross-sectional and controlled intervention studies.

### Absolute Reliability

The absolute reliability of a test provides an estimate of the precision of its outcome score on repeated testing. The SEM and MDC are easy to interpret because they are expressed in the same units as the original measures and as such are useful for clinicians to determine individual

**Table 2:** Intra-rater variability of the SFT in the elderly subjects

| Parameter (units) | *n* | Variable | Test result | | |
|---|---|---|---|---|---|
| | | | Minimum value | Maximum value | Mean ± SD |
| Body mass index (kg/m$^2$) | 31 | Tester 1 | 16 | 29 | 22.39 ± 2.81 |
| | | Tester 1A | 16 | 29 | 22.39 ± 2.81 |
| Arm curl test (number of times) | 31 | Tester 1 | 5 | 15 | 11.16 ± 2.35 |
| | | Tester 1A | 7 | 15 | 11.32 ± 2.02 |
| Chair stand test (number of times) | 31 | Tester 1 | 7 | 15 | 10.74 ± 1.69 |
| | | Tester 1A | 8 | 15 | 10.55 ± 1.82 |
| Back scratch test (number of times) | 31 | Tester 1 | 5 | 34 | 15.34 ± 7.29 |
| | | Tester 1A | 5 | 34 | 15.81 ± 6.97 |
| Chair sit and reach test (number of times) | 31 | Tester 1 | 10 | 30 | 15.71 ± 5.45 |
| | | Tester 1A | 10 | 31 | 15.65 ± 5.41 |
| 8 foot up and go test (number of times) | 31 | Tester 1 | 7.89 | 17.84 | 11.26 ± 2.41 |
| | | Tester 1A | 7.8 | 19 | 11.57 ± 2.67 |
| 2-min step test (number of times) | 31 | Tester 1 | 48 | 80 | 64.13 ± 9.69 |
| | | Tester 1A | 48 | 86 | 65.00 ± 9.81 |

SFT, senior fitness test; *n*, number of participants; tester 1, first testing of participants by investigator for intra-rater variability; tester 1A, second testing of participants by investigator for intra-rater variability; SD, standard deviation.

**Table 3:** Intraclass correlation of the SFT in the elderly subjects

| Test | | 95% Confidence interval | | | *F* test with true value 0 | | | |
|---|---|---|---|---|---|---|---|---|
| | | Intraclass correlation | Lower bound | Upper bound | Value | df1 | df2 | Sig |
| Body mass index | Average measures | 1.000 | 1.000 | 1.000 | 0 | 30.0 | 0 | 0.00 |
| Arm curl test | Average measures | 0.933 | 0.861 | 0.968 | 14.968 | 30.0 | 30 | 0.00 |
| Chair stand test | Average measures | 0.934 | 0.864 | 0.968 | 15.254 | 30.0 | 30 | 0.00 |
| Back scratch test | Average measures | 0.991 | 0.982 | 0.996 | 113.182 | 30.0 | 30 | 0.00 |
| Chair sit and reach test | Average measures | 0.987 | 0.973 | 0.994 | 76.150 | 30.00 | 30 | 0.00 |
| 8-Foot up and go test | Average measures | 0.974 | 0.946 | 0.988 | 38.610 | 30.00 | 30 | 0.00 |
| 2-Min step test | Average measures | 0.977 | 0.953 | 0.989 | 44.017 | 30.00 | 30 | 0.00 |

SFT, senior fitness test; df1, degree of freedom for test 1; df2, degree of freedom for test 2; Sig, significance.

**Table 4:** Reliability measures[a] of the various test components of the SFT in the elderly subjects

| | R-square | KT | *F* value | Pr > *F* | MSE | ICC | SEM | MDC |
|---|---|---|---|---|---|---|---|---|
| Body mass index | 1 | 31 | ∞ | <0.0001 | 0 | 1 | 0 | 0 |
| Arm curl test | 0.7829 | 24.2699 | 104.55 | <0.0001 | 0.91928 | 0.933 | 0.522864 | 1.449305 |
| Chair stand test | 0.7733 | 23.9723 | 98.92 | <0.0001 | 0.77924 | 0.934 | 0.467566 | 1.296028 |
| Back scratch test | 0.9672 | 29.9832 | 854.14 | <0.0001 | 1.65187 | 0.991 | 0.661232 | 1.832842 |
| Chair sit and reach test | 0.9489 | 29.4159 | 538.47 | <0.0001 | 1.54562 | 0.987 | 0.616835 | 1.709779 |
| 8 foot up and go test | 0.911 | 28.241 | 296.94 | <0.0001 | 0.65387 | 0.974 | 0.430525 | 1.193353 |
| 2-min step test | 0.9132 | 28.3092 | 305.26 | <00.01 | 8.63394 | 0.977 | 1.48776 | 4.123861 |

SFT, senior fitness test; R-square, coefficient 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, intraclass co-relation co-efficient; SEM, standard error of mean; MDC, minimal detectable change.
[a]Calculation over log-transferred data.

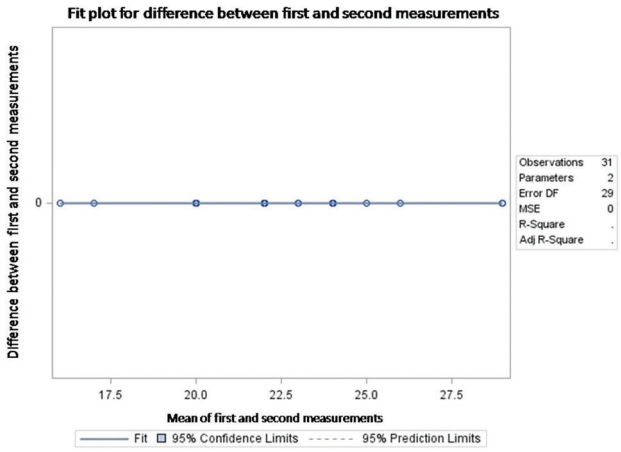


**Figure 1:** Body mass index of the elderly subjects.

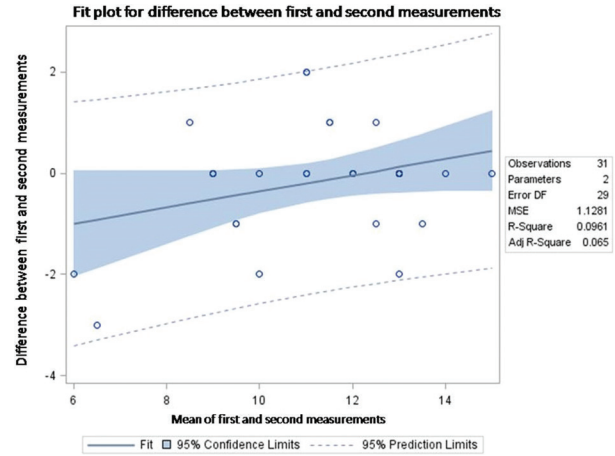


**Figure 2:** Arm curl test in the elderly subjects.

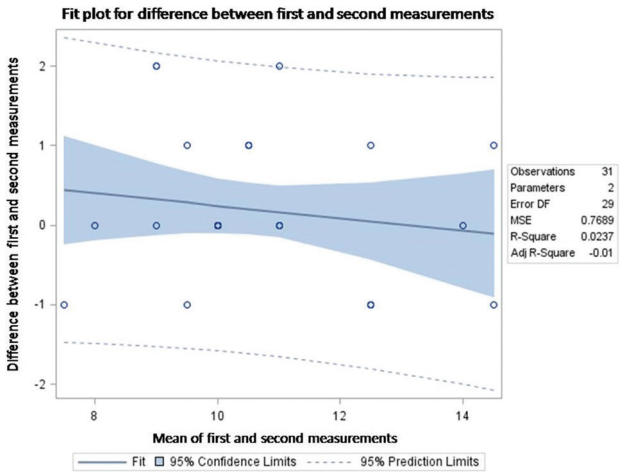


**Figure 3:** Chair stand test in the elderly subjects.

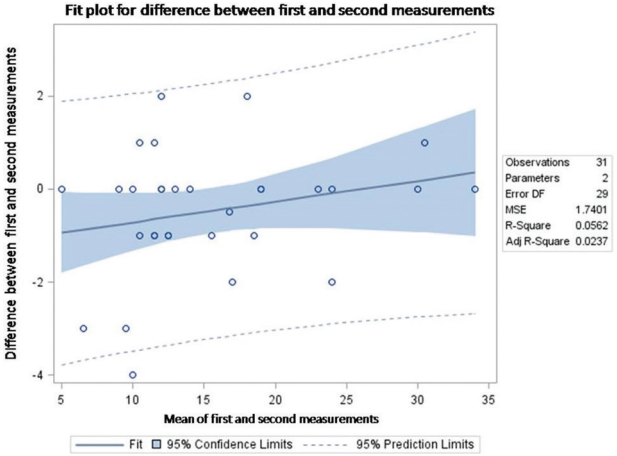


**Figure 4:** Back scratch test in the elderly subjects.

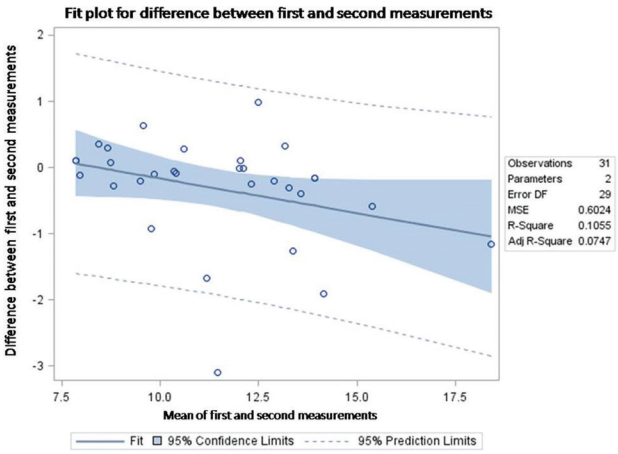


**Figure 5:** Chair sit and reach test in the elderly subjects.

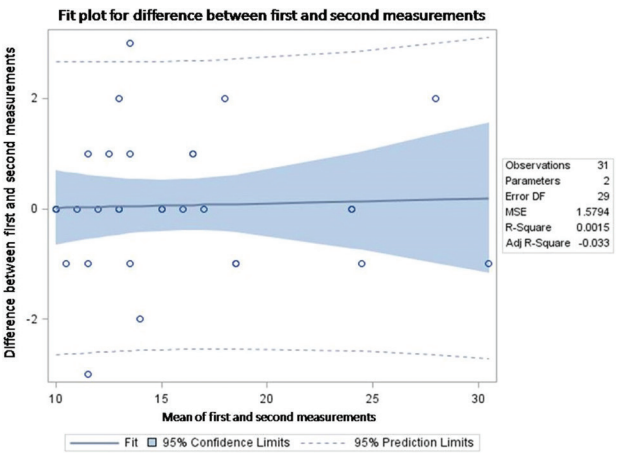


**Figure 6:** Eight foot up and go test in the elderly subjects.

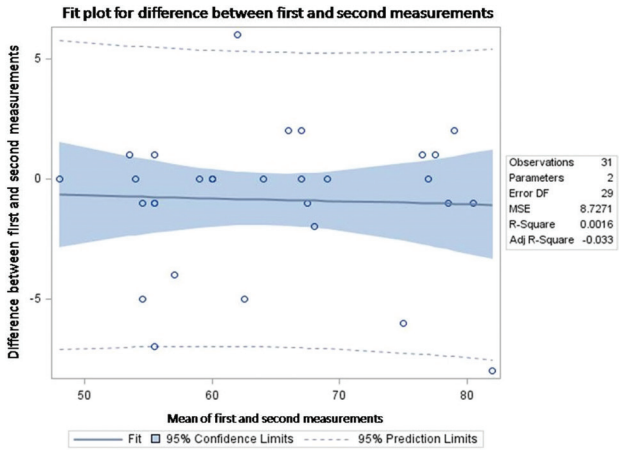


**Figure 7:** Two-minute step test in the elderly subjects.

improvement. The MDC is based on the SEM, but is more conservative (~2.7 SEMs). If the score change is larger than the MDC, this difference is not caused by measurement error or participant variability.[17]

With the exception of the 2-min step test, all the other components of the SFT showed lower SEM with low variability in SEM and very low values of MDC. The exception in the 2-min step test can be partly explained by the fact that this test has greater cardiovascular implications than the other components of the SFT and such cardiovascular comorbidities are more common in the elderly people. This may act as a confounding factor with respect to the 2-min step test in the elderly subjects. The sample for the study was selected by the nonrandom convenient sampling method and the sample size was small, selected from a limited geographical region of the city, which is a limitation of the study.

In summary, we conclude that SFT for the elderly population is safe, effective, highly reliable, and feasible, and can be applied as a clinical assessment for assessing physical fitness. Future research should focus on inter-rater reliability for the elderly population in a modified environment.

## Conclusion

This study was aimed to assess the test–retest reliability of SFT in the geriatric population in an urban community setting so that this test can be utilized for testing the physical fitness of the elderly population in the Indian community setting. The test–retest reliability of the SFT was excellent and these tests are thus applicable for cross-sectional and controlled interventional studies for elderly population. The inferences of this study would be able to help in assessing the physical fitness of the elderly population in the community by using the SFT to specifically measure their physical fitness. This study was undertaken in a limited sample in the age group of 65–74 years of the geriatric population. In this context, large-scale studies may be undertaken in different ages of the geriatric population to substantiate the reliability and usefulness of this fitness test in the different geriatric age groups.

## References

1. US Census Bureau. *U.S. Interim Projections by Age, Sex, Race, and Hispanic Origin*. 2004. Available at: https://www.census.gov/hhes/www/laborfor/Working-Beyond-Retirement-Age.pdf (last accessed on November 18, 2015).
2. Loh DA, Hairi NN, Choo WY, Hairi FM, Peramalah D, Kandiben S, et al. MultiComponent Exercise and theRApeutic lifeStyle (CERgAS) intervention to improve physical performance and maintain independent living among urban poor older people—a cluster randomised controlled trial. BMC Geriatr 2015;15(1):8–14.
3. Ogden LL, Richards CL, Shenson D. Clinical preventive services for older adults: the interface between personal health care and public health services. Am J Public Health 2012;102(3):419–25.
4. Osness WH. *Functional Fitness Assessment for Adults Over 60 Years (A Field Based Assessment)*. Reston, VA: AAHPERD Publication Sales Office, 1990. pp. 220–91.
5. 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. J Gerontol 1994;49(2):85–94.
6. Rikli RE, Jones CJ. Development and validation of a functional fitness test for community residing older adults. J Aging Phys Act 1999;7(4):127–59.
7. Rikli RE, Jones CJ. Assessing physical performance in independent older adults: issues and guidelines. J Aging Phys Act 1997;5(7):244–61.
8. Rikli RE, Jones CJ. Physical activity level, fitness, and functional ability of community residing older adults. Med Sci Sports Exercise 2000;32(S2):218–34.
9. Bruton A, Conway JH, Holgate ST. Reliability: what is it, and how is it measured? Physiotherapy 2000;86(2):94–9.
10. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1(8476):307–10.
11. Domholdt E. *Rehabilitation Research: Principles and Applications*, 3rd edn. St Louis, MO: Elsevier Saunders, 2005.
12. Flechner L, Tseng TY. Understanding results: P-values, confidence intervals, and number need to treat. Indian J Urol 2011;27(4):532–5.
13. Stuart-Hamilton I. *The Psychology of Ageing: An Introduction*. Jessica Kingsley Publishers, London and Philadelphia, 2012.
14. Haskell WL, Leel-M, Pate AA, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Circulation 2007;116(9):1081–93.
15. Carbonell-Baeza A, Alvarez-Gallardo IC, Segura-Jiménez V, Castro-Piñero J, Ruiz JR, Delgado-Fernández M, et al. Reliability

and feasibility of physical fitness tests in female fibromyalgia patients. Int J Sports Med 2015;36(2):156–62.
16. 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.
17. Blankevoort CG, van Heuvelen MJG, Scherder EJA. Reliability of six physical performance tests in older people with dementia. Phys Ther 2013;93(1):69–78.

**How to cite this article:** Bhattacharya PK, Deka K, Roy A. A community-based study to assess test–retest reliability of senior fitness test in the geriatric population in a northeastern Indian city. Int J Med Sci Public Health 2016;5:1606-1612

**Source of Support:** Nil, **Conflict of Interest:** None declared.