Cross-Validity and Reliability of the Age Scale for Assessing Activities of Daily Living among Japanese Community-Dwelling Adults Aged 75 Years or Older

Rafael Figueroa[1], Satoshi Seino[2,3], Noriko Yabushita[4], Yoshiro Okubo[1,3], Yosuke Osuka[1,3], Miyuki Nemoto[3,4], Songee Jung[1,4] and Kiyoji Tanaka[4,*]

[1]*Graduate School of Comprehensive Human Sciences, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki, 305-8577, Japan*

[2]*The Research Team for Social Participation and Community Health, Tokyo Metropolitan Institute of Gerontology, 35-2 Sakae, Itabashi, Tokyo, 173-0015, Japan*

[3]*The Japan Society for the Promotion of Science, 8 ichiban, Chiyoda, Tokyo, 102-8472, Japan*

[4]*Faculty of Health and Sport Sciences, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki, 305-8577, Japan*

Abstract: The aim of this study was to evaluate the validity and reliability of the age scale for assessing activities of daily living (ADL) among community-dwelling adults aged 75 years or older. Participants comprised 89 older Japanese: 47 men (79.1±3.8 years) and 42 women (78.9±3.7 years). The ADL age (ADLA) equation is an instrument that estimates participants' physical functioning regarding ADL using data obtained previously from 1006 subjects. The equation is as follows: ADLA for women = 0.447 (chronological age) -5.49ADLS + 44.17; and ADLA for men = 0.519CA – 4.27ADLS + 38.26. Validity was evaluated using cross-validation, and reliability was evaluated using internal consistency and test-retest methods. The correlation between the ADLA and chronological age (CA) in the cross-validation samples (r = 0.62 women, r = 0.61 men) was not significantly different from the relationship observed in the original sample (r = 0.72 women p = 0.636, and r = 0.75 men p = 0.571). Cronbach's alpha (α) value (internal consistency) for the total ADLA was 0.97 for men and 0.92 for women, whereas the intra-class correlation coefficient (ICC) value (test-retest reliability) was 0.96 for men and 0.91 for women. The results suggest that the ADLA is a reliable and valid tool for the assessment of ADL with satisfactory psychometric properties and that is applicable for persons aged 75 years and older in Japan.

Keywords: Age activities of daily living assessment, cross-validity and reliability of the age scale**,** older Japanese.

1. INTRODUCTION

The difficulties entailed in measuring health status at any age are well-known, but autonomy in daily living has become a reasonably adequate parameter of health status in the case of older adults [1]. Performance-based measures of physical function predict future incidence of disability, dependence in activities of daily living, institutionalization and death in initially non-disabled older people [2-5]. Every year approximately 10% of people over 75 are institutionalized in Japan. Although measures of physical function are accurate for many of these people, there still a need for the development of an index or tool to discriminate better between people who are of the same chronological age (CA) but differ in terms of physical and physiological functioning [6].

Cooper *et al.* (2011) have pointed out the necessity of investigating the possibility of a derived composite score representing overall lower or upper-body functioning, where an individual performance measure can contribute significantly to discerning functional status and probable adverse-health outcomes [7]. In addition, Cooper *et al.* (2010), in a systematic review and meta-analysis of the association between physical measures and all-cause mortality in community-dwelling older adults, have shown a growing body of evidence that an individual performance measure could contribute significantly to discerning functional status and adverse-health outcomes [8].

The present authors have developed the concept of activities of daily living age (ADLA) for an assessment of the functional status of people over 75 years old. ADLA is estimated using three physical performance tests. However, there is a need for valid and reliable instruments for evaluating the effects of this method [9], since it is crucial for a clinician to know whether a change in scores on functional tests is due to a real change in functioning or to measurement error [9].

The reliability of a method (i.e. when repeated measurements of an individual's performance are consistent from one time to another) [9, 10] can be described as either relative or absolute [9].

*Address correspondence to this author at the Tennoudai 1-1-1, Tsukuba, Ibaraki, 305-8574, Japan; Tel: +819021603990; E-mail: rafellfg@hotmail.co.jp

E-ISSN: 2309-6128/15

Relative reliability examines the relationship between two or more measurements and the consistency of an individual's position within the group. Absolute reliability examines variability in scores in repeated measurements. The intra-class correlation coefficient (ICC) is commonly used to evaluate relative reliability [9]. However, the ICC value is of limited use for the clinician because it is not related to the actual scale of measurement, but is dependent on the range of the individual's performance [9]. If the individual's range of scores is low, the ICC value often will show poor reliability, and vice versa [11, 12]. This means that the clinician cannot be sure whether a high ICC value obtained during an assessment of ADLA actually indicates low variability at the individual level.

A more appropriate way of investigating the reliability of an instrument intended for use in a clinical setting is to examine absolute reliability [13]. When using absolute reliability, the assessor receives information about how much variability caused by measurement error can be expected in an individual's scores [9].

In addition, the method for the computation of functional age generally consists of several approaches previously reported in the experimental literature. Studies assessing age scale in Japanese aged 75 years or older are very limited, and an appraisal of these studies reveals that they have been performed on several different age groups [14-17].

However, from this body of evidence it can be seen that the scales developed are mostly for younger and middle-age adults in good physical condition. For the research performed within ADL, it is notable that most studies have been scale adaptations and that there is no physical performance scale specific to ADL for Japanese aged 75 years and older. Therefore, this research aims to determine the developed ADLA's validity through cross-validation and its absolute reliability, using a measure first given in 2012 and then retested in 2013. The ADLA is expected to be an enduring trait. Hence, the score in the first rating should agree with the score in the second rating.

2. METHODS

2.1. Participants

Participants included in this investigation were 100 Japanese community dwellers aged 75 years and older. The participants were recruited through poster advertisements and flyers displayed in senior centers, leisure centers, and residential retirement communities in Fukushima and Ibaraki prefectures, northeast of Tokyo. To be included, the participants needed to be community dwellers aged 75 years or older. Exclusion criteria were as follows: being unable to perform or participate in the physical test in both sessions (n = 7) or being unable to understand the instructions for the test and questionnaires (n = 4). The remaining 89 participants (42 women with a mean age of 78.9±3.7 and 47 men with a mean age of 79.1±3.8) were included in the current study. Prior to the test, recruited individuals who required nursing care, prevention programs or day-care service read and signed an informed consent form, which was approved by the institutional review board (IRB approval no. 696). This study was conducted in accordance with the guidelines proposed in the Declaration of Helsinki, and the study protocol was reviewed and approved by the ethics committee of the University of Tsukuba, Japan.

2.2. Demographic and Health-Related Information

Participants were interviewed in order to obtain demographic information, which included age, pain sites, co-morbidities and health-related information. Body height and weight were measured with minimal clothing and no shoes. Participants were asked to rate their current health status as poor, good, or very good.

2.3. Physical Performance Items

The 3 physical performance items described below formed the basis for the ADLA equation, which was used to obtain the ADLA. These items were selected based on previous research, in which they were recognized as significantly related to ADL [18]. In order to determine test-retest reliability, selected physical performance items, or tests, were assessed at the baseline in 2012 and again one year later.

2.3.1. Hand-Grip Strength

Participants stood with their arms at their sides down to the elbows, which were bent at 90 degrees to enable their hands to grip the dynamometer in front of them. The handle of the dynamometer was adjusted if needed. Participants squeezed the dynamometer at maximum isometric effort, which they maintained for about 3 seconds. No other body movement was allowed [19]. The participants were strongly encouraged to give their maximum effort with good respiration to obtain a strong result. The test was performed twice for each hand alternately. The average score of all trials, measured in kilograms, was used for analysis. The hand-grip dynamometer was a GRIP-D, T. K. K. 5401 manufactured by Takei Scientific Instruments Co. Ltd., Tokyo, Japan.

2.3.2. Five Repetitions of the Chair Sit to Stand Test (SST)

Participants were asked to stand up and sit down on a standard-height chair as quickly as possible. Specifically, the start position was seated with the knee joint angle at 90 degrees and the soles of the feet completely on the floor. The time was measured from the initial sitting position to the final fully erect position at the end of the fifth stand. The average time recorded in the two trials rounded to the nearest 0.01 seconds was used for analysis [20].

2.3.3. Timed Up and Go (TUG)

Participants were asked to sit down on a standard-height chair. After a signal, they stood up from the chair and walked forward as quickly as possible to a cone three meters away, turned 180 degrees at the cone, and then walked back to the chair and sat down. Participants were allowed to use canes or walkers. For the analysis, the average time of the two trials was rounded to the nearest 0.01 s [21].

The ADLA used for women was obtained through the following process: Statistical analysis of the data began with a calculation of the arithmetic means and standard deviations (SDs), along with a correlation matrix among the eight physical tests. The Spearman rank-order correlation coefficient was calculated to determine the associations among the Barthel index score and each of the variables, and these were submitted to principal component analysis. The first principal component was used as the best single descriptor of total ADL performance. Three variables were the highest Spearman rank-order correlated (Hand grip strength, 5SST and TUG). The first principal component score was used as a unitary index ADL score (ADLS). To calculate individual ADLS, each score was first standardized and then summed across tests in a weighted manner, using the coefficients of the principal component scores obtained from the principal component analysis. The ADLS was converted into an age scale where the average and SD of the CA was used. During this process, it was taken into consideration that ADLs were distributed with a mean of 0 and an SD of 1.0 [22].

In order to compute the ADLS for each subject, we calculated principal scores as Σ ai xi where ai is the factor loading of the three test items and xi is an individual's standard score on the three test items.

The following equations were obtained for the ADLS:

Women's ADLS = $0.075\ X_1 - 0.082\ X_2 - 0.063\ X_3 + 0.124$

Men's ADLS = $0.051\ X_1 - 0.105\ X_2 - 0.099\ X_3 + 0.249$

Where X_1 = hand-grip strength (kg), X_2 = five-chair SST (s), X_3 = TUG (s).

To transform individual ADLS into an age scale, a T-scale was used, taking into consideration that these are distributed with a mean of 0 and an SD of 1.0. First the scores were standardized using the average and SD of the CA.

Then the following equation for ADLA was derived:

Women's ADLA = −5.493 ADLS + 79.90

Men's ADLA = −4.272 ADLS + 79.57

The above figures for women, −5.493 and 79.90, and for men, −4.272 and 79.57, are respectively the mean and SD of the CA of our sample of 1006 participants.

Next, there was a need to correct the distortion of the ADLA at the regression edges that was a function of the CA and the disagreement between the slopes of both regression lines. We calculated these correction terms by following Dubina *et al.*'s (1984) method [22]: the correction is calculated as $Z = (1 - b)(Y_i - Y)$, where "Y_i" is the CA of an individual, "Y" is the mean CA, and "b" is the coefficient of simple linear regression that expresses the relation between ADLA and CA. Finally, the corrected ADLA was obtained by summing the Z in the second equation [23].

The equations after this correction were as follows:

Z = (1 − 0.553) (CA − 79.90)

Women's ADLA = 0.447CA − 5.49ADLS + 44.17

Z = (1 − 0.480) (CA − 79.57)

Men's ADLA = 0.519CA − 4.27ADLS + 38.26

2.4. Statistical Analyses

Statistical analysis was carried out using the IBM SPSS Statistics software, version 21 (SPSS Inc., Chicago, IL, USA), with the level of statistical significance set at 5%. The reliability of the ADLA was measured by analyzing internal consistency using Cronbach's α coefficient [24]. Inter-rater and intra-rater reliability was calculated by means of the ICC [25].

The ICC was chosen because it provides a powerful and flexible method to examine the reliability of the ADLA [26]. Coefficients below 0.50 indicate poor reliability, those between 0.50 and 0.75 indicate

moderate reliability, and those above 0.75 indicate good reliability [27].

The following two aspects of the ADLA's reliability were assessed: (1) test-retest reliability was determined for the total score of the three components and (2) internal consistency was determined using Cronbach's α coefficient for all tasks within the three motor domains. The model was two-way mixed and the type was absolute agreement. In the present study, we expected the two ADLA scores to match.

Cross-validation sample: In order to assess the stability of the ADLA equation for a different sample of subjects, ADLA scores were computed using the original regression equation for this new sample of 89 participants. Correlations between ADLA and CA were determined in order to assess the stability of the ADLA-CA relationship across the two sample groups.

3. RESULTS

The baseline characteristics of the 89 study participants are summarized by gender in Table **1**. All participants claimed no known neuromuscular, musculoskeletal or cardiovascular pathology that would affect their ambulatory capacity to perform the tasks in the current study. Also, most of the participants rated their current health status as good to very good.

Table 1: Characteristics of the Study Participants (n = 89)

Variables	mean ± standard deviation or n (%)					
	women (n = 42)			men (n = 47)		
Age, years	78.9	±	3.7	79	±	3.9
Height, cm	146.1	±	5.9	157.1	±	6.2
Weight, kg	49.9	±	7.4	57.3	±	9.0
Body mass index, kg/m^2	23.5	±	3.9	23.2	±	2.9
Chronic disease, n (%)						
Hypertension	22	(51.2)		15	(31.9)	
Stroke	2	(4.7)		4	(8.5)	
Heart disease	5	(11.6)		5	(10.6)	
Diabetes mellitus	5	(11.6)		2	(4.3)	
Self-rated health, n (%)						
Excellent to good	41	(95.3)		47	(100.0)	
Fair to poor	2	(4.7)		0	-	
Alcohol drinking status, n (%)						
Current	20	(46.5)		22	(46.8)	
No drink	23	(53.5)		26	(55.3)	
Body pain, n(%)						
Waist	20	(46.5)		11	(23.4)	
Shoulder joint	4	(9.3)		3	(6.4)	
Elbow joint	0	-		1	(2.1)	
Hip joint	0	-		1	(2.1)	
Knee joint	15	(34.9)		10	(21.3)	
Feet	4	(9.3)		1	(2.1)	

Table **2** shows a comparison between CA and ADLA in the validation and cross-validation groups. For both sexes, no differences were found between either CA or ADLA. The correlation between the ADLA and CA in the cross-validation samples (r = 0.62 women, r = 0.61 men) was not significantly different from the relationship observed in the original sample in women (r = 0.72 women p = 0.636 and r = 0.75 men P = 0.571).

Table 2: Comparison of Chronological Age and ADL Age in Validation and Cross-Validation Groups (n = 89)

Variables	mean ± standard deviation or n (%)											
	Validation group (n = 1006)						Cross-validation group (n = 89)					
	women (n = 694)			men (n = 312)			women (n = 42)			men (n = 47)		
Chronological age, years	79.9	±	5.4	79.6	±	4.3	78.9	±	3.7	79.1	±	3.8
ADLA	79.8	±	6.0	79.6	±	7.4	85.4	±	3.8	79.9	±	4.5
Correlation ADLA/CA	r = 0.72			r = 0.75			r = 0.62			r = 0.61		
							NS			*NS*		

r: Correlation between ADLA and CA.

p = NS: No significance diference between ADLA and CA among the original and the cross-validation samples.

ADLA: Activities of daily living age.

CA: Chronological age.

As shown in Table **3**, the internal consistency of the scale was good for both sexes with a Cronbach's α coefficient of 0.919 for women and 0.968 for men. Inter-rater reliability (test-retest) was analyzed using ICC, where the average measure was rated for all of the items independently, obtaining 0.906 for women and 0.958 for men.

Table 3: Intraclass Correlation Coefficients of the Study Participants (n = 89) by Gender

Variables	Intraclass correlation	
	women (n = 42)	men (n = 47)
single measure	0.829	0.920
average measure	0.906	0.958
Cronbach`s Alpha	0.919	0.968

Table 4: Internal Consistency and Intra-Rater Raliability of the ADLA (89)

Subscales	women (n = 42)		ICC	95%CI		C.α	men (n = 47)		ICC	95%CI		C.α
	Rating 1	Rating 2		lower	upper		Rating 1	Rating 2		lower	upper	
	mean (SD)	mean (SD)					Mean (SD)	Mean (SD)				
1. Hand-grip, kg	19.7 (41)	19.3 (3.3)	0.803[a]	0.634	0.894	0.801	30.1 (6.6)	29.0 (6.2)	0.947[a]	0.898	0.972	0.953
2. 5-chair SST, s	8.4 (3.1)	8.7 (3.0)	0.798[a]	0.625	0.891	0.796	11.1 (4.7)	11.5 (5.1)	0.960[a]	0.928	0.978	0.961
3. Timed up and go, s	7.8 (2.5)	7.9 (3.0)	0.912[a]	0.836	0.953	0.910	7.4 (2.7)	7.6 (3.9)	0.908[a]	0.835	0.949	0.907

ADLA, activities of daily living age; SD, standard deviation; ICC, Intraclass correlation coefficient; CI, confidence intreral; C.α, Cronbach`s alpha.
SST, sit to stand test; [a] correlation is significant at the .01 level.

Similarly, inter-rater reliability and intra-rater reliability (test-retest) among the different items was analyzed using ICC, and all items were found to have values higher than 0.8 (Table **4**).

DISCUSSION

To the best of our knowledge, this is the first ADLA assessment for Japanese people aged 75 years and older. Moreover, our results have shown that the physical performance assessed by ADLA can indicate with high confidence low physical function in older subjects. The results suggest that the ADLA is a highly valid and reliable tool for the assessment of ADL.

The correlation between the ADLA and CA in the cross-validation sample of 89 participants was not significantly different from the relationship observed in the original sample which consisted of 1006 subjects. Obtained results thus indicate a satisfactory validation of ADLA assessment and correspond well with previous findings [28, 29]. For example, a study of Korean women by Kim and Tanaka (1995) assessing functional age using ADL has also shown that there were no differences between the original sample (r = 0.77) and the cross-validation group (r = 0.68) [28]. In addition, Yabushita *et al.* (2004) demonstrated that there was no significant difference between the first sample and a second sample 10 months later regarding the relationship between CA and the study's physical fitness assessment for older Japanese [29]. Similarly, our results imply that the ADLA-CA relationship remained stable across the two samples.

Satisfactory test-retest reliability was also indicated by the fact that the ICC of the items that compose the ADLA assessment ranged from 0.80 to 0.91 for women and 0.91 to 0.96 for men. In the present study, Cronbach's α for internal consistency also ranged from 0.80 to 0.91 for women and 0.91 to 0.96 for men.

Furthermore, the ICC of the ADLA was 0.91 for women and 0.96 for men, and Cronbach's α coefficient was 0.92 for women and 0.97 for men. Obtained results therefore indicate that the ADLA assessment has adequate reliability and this corresponds well with previous findings for community-dwelling older individuals [30]. Shigematsu *et al.* (1998) analyzed the reliability and objectivity of test items that assessed the functional fitness required for performing ADL among older adult Japanese women (n = 207; people aged 60-91 years), and they obtained excellent test-retest reliability for hand-grip strength (ICC = 0.91) [30]. Schaubert *et al.*'s (2005) study (n = 10; mean age = 75.5 [5.8] years) found adequate test-retest reliability for 5SST (ICC = 0.82); this study used the MicroFET 2 hand-held dynamometer (knee extension strength), the Jamar dynamometer (grip strength), and the 5SST. Mobility was tested using TUG and a timed-walk test. Intra-class correlation coefficients, which were used to characterize the reliability of the strength tests, ranged from 0.807 to 0.981. Pearson correlations between the lower extremity strength measures and the TUG and gait speed ranged from 0.635 to -0.943. Their examination of these three measures thus adds to the previous evidence of the stability of these strength measures; and this further justifies the use of hand-held dynamometry and the 5STS test when investigating limitations in mobility, as the current study does [31]. In addition, Tiedemann *et al.* (2008) and Bohannon *et al.* (2007) obtained adequate test-retest reliability for the 5SST with ICC = 0.890 and ICC = 0.957, respectively [32, 33]. Similarly, Nordin *et al.* (2004), in a study of the Timed "Up and Go" (TUG) Test investigated the expected variability of TUG scores among older dependent subjects (n = 78; mean age 84.8±5.7). The TUG assessments were performed on 3 different days. The intra-class correlations were greater than 0.90 and were similar within and between raters [34].

Previous research has also shown that the ICC is strongly affected by the range of scores used to calculate the coefficient: the ICC is high when the difference in scores between measurements is small in

comparison with the range of scores among the studied participants [35-37].

It is also important to recognize that this study has some limitations. First, our sample of the population might not be representative of the entire Japanese older population, because we recruited our participants at community centers, thus limiting participation in our study to these centers' visitors. Second, our participants were a convenient sample and not randomly selected. The sample consisted only of relatively healthy older Japanese aged 75 years and over who were sufficiently mobile to commute to our study center, and thus tended to participate with, which had a positive effect in the strata. Third, it must be noted that there is no universal agreement on the interpretation of correlation coefficients. A variety of guidelines are suggested in the literature: > 0.75 equals "excellent reliability", [38] ≥ 0.80 is "very reliable", [39] and > 0.75 indicates "good reliability" [40]. Therefore, several factors may have been involved in these results, especially since there was an average of one year between test and retest.

In conclusion, this study found further evidence that the ADLA is a reliable and valid tool for the assessment of ADL for people 75 years and older. In this study, the Cronbach's α coefficient values were good, and there were high values for test-retest reliability for almost all the test items. The good construct results therefore indicate that the ADLA can be used with confidence. However, a validation process should be confirmed by more than one approach and using multiple techniques [41]. Future research efforts are thus required to provide more evidence for the reliability and validity of ADLA assessment in different groups of older Japanese people, with and without movement difficulties, and for all advanced age groups.

ACKNOWLEDGEMENTS

We would like to thank the large number of open sources and related Japanese community support projects that had a critical role in facilitating this work, and we are especially grateful to all the study participants. Finally, the authors would like to sincerely thank the Yamaha Motor Foundation for Sports (YMFS), which provided advice and financial support. This study was also supported by the Japan Society for the Promotion of Science (JSPS) (Grant-in-Aid for scientific research [A]) #19200047.

REFERENCES

[1] World Health Organization: The uses of epidemiology in the study of the elderly. Report of a WHO scientific Group on the Epidemiology of Aging. Geneva 1984; pp. 1-84.

[2] Gill TM, Williams CS and Tinetti ME. Assessing risk for the onset functional dependence among older adults: the role of physical performance. J Am Geriatr Soc 1995; 43: 603-609.

[3] Guralnik JM, Ferrucci L, Simonsick EM, Salive ME and Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Eng J Med 1995; 332: 556-561. http://dx.doi.org/10.1056/NEJM199503023320902

[4] Shibata H, Haga H, Nagai H, Suyama Y, Koyano W and Suzuki T. Predictors of all-cause mortality between ages 70 and 80: the Koganei study. Arch Gerontol Geriatr 1992; 14: 283-97. http://dx.doi.org/10.1016/0167-4943(92)90028-3

[5] Sonn U. Longitudinal studies of dependence in daily life activities among elder persons. Scand J Rehab Med 1996; 34(suppl): 2-33.

[6] Tanaka K, Kim HS, Yang JH, Shimamoto H, Kokudo S and Nishijima T. Index assessing functional status in elderly Japanese men. Appl Human Sci 1995; 14(2): 65-71.

[7] Cooper R, Kuh D, Cooper C, Gale CR, Lawlor DA, Mattews F, *et al*. Objective measures of physical capability and subsequent health: a systematic review. Age Ageing 2011; 40: 14-23. http://dx.doi.org/10.1093/ageing/afq117

[8] Cooper R, Kuh D and Hardy R. Mortality Review Group; FALCon and HALCyon Study Teams. Objectively measured physical capability levels and mortality: systematic review and meta-analysis. Br Med J 2010; 341: c4467. http://dx.doi.org/10.1136/bmj.c4467

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

[10] Vincent WJ. Statistics in Kinesiology. 2nd ed. Champaign, IL: Human Kinetics Inc 1999.

[11] Bland JM and Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. Comput Biol Med. 1990; 20: 337-340. http://dx.doi.org/10.1016/0010-4825(90)90013-F

[12] Rankin G and Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. Clin Rehabil 1998; 12: 187-199. http://dx.doi.org/10.1191/026921598672178340

[13] Bland JM and Altman DG. Measurement error. BMJ 1996; 313: 744. http://dx.doi.org/10.1136/bmj.313.7059.744

[14] Chodzko-Zajko WJ and Ringel RL. Physiological fitness measures and sensory and motor performance in aging. Exp Gerentol 1987; 22: 317-328. http://dx.doi.org/10.1016/0531-5565(87)90030-1

[15] Lee MS, Matsuura Y and Tanaka K. Assessment of physical fitness age in middle-aged and elderly men. Jpn J Phys Fit Sport 1993; 42: 59-68.

[16] Tanaka K, Matsuura Y, Nakadomo F and Nakamura E. Assessment of vital age of Japanese women by principal component analysis. Japan J Phys Ed 1990; 35: 121-131.

[17] Nakamura E, Miyao K and Ozeki T. Assessment of biological age by principal component analysis. Mech Ageing Dev 1988; 46: 1-18. http://dx.doi.org/10.1016/0047-6374(88)90109-1

[18] Bravo G, Gauthier P, Roy PM, Tessier D, Gaulin P, Dubois MF, *et al*. The functional fitness assessment battery: Reliability and validity data for elderly women. J Aging Phys Act 1994; 2: 67-79.

[19] Shinkai S, Kumagai S, Fujiwara Y, Amano H, Yoshida Y, Watanabe S, *et al*. Predictor for the onset of functional decline among initially non-disabled older people living in a community during a 6-year follow-up. Geriatr Gerentol Int 2003; 3(Suppl 1): S31-S39. http://dx.doi.org/10.1111/j.1444-0594.2003.00094.x

[20] Guralnik JM, Ferrucci L, Simonsick EM and Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Eng J Med 1995; 332: 556-561.

[21] Podsiadlo D and Richardson S. The timed "up and go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39: 142-148.

[22] Dubina TL, Mints AY and Zhuk EV. Biological age and its estimation. III. Introduction of a correction to the multiple regression model of biological age and assessment of biological age in cross-sectional and longitudinal studies. Exp Gerontol 1984; 19: 133-143. http://dx.doi.org/10.1016/0531-5565(84)90016-0

[23] Figueroa R, Seino S, Yabushita N, Okubo Y, Osuka Y, *et al*. Age scale for assessing activities of daily living. Intl Jour Stats Med Res 2015; 4; 48-56. http://dx.doi.org/10.6000/1929-6029.2015.04.01.5

[24] Cronbach LJ and Warrington WG. Time-limit test: Estimating their reliability and degree of speeding. Psychometrika 1951; 16: 167-188. http://dx.doi.org/10.1007/BF02289113

[25] Lord FM and Novick MR. Statistical theories of mental test scores. Reading, MA: Addison-Wesley 1968.

[26] Ottenbacfhef JK and Tomcheck SD. Reliability analysis in therapeutic research: Practice and procedure. Am J Occup Ther 1993; 47(1): 10-16. http://dx.doi.org/10.5014/ajot.47.1.10

[27] Portney LG and Watkins MP. Foundations of clinical research: Applications to practice. Norwalk, Connecticut: Appleton and Lange 1993.

[28] Kim HS and Tanaka K. The assessment of Functional age using "activities of daily living" Performance tests: A study of Korean women. J Aging Phys Act 1995; 3: 39-53.

[29] Yabushita N, Kikkawa K, Sakai T, Nakamura Y and Tanaka K. Assessment of physical fitness status in older Japanese men using a physical fitness age. JSHHE 2004; 70(5): 195-206.

[30] Shigematsu R, Kim H, Kim H and Tanaka K. Reliability and objectivity of the test items to assess functional fitness required for performing activities of daily living in older adult Japanese women. JSPA 1998; (3): 13-18.

[31] Schaubert KL and Bohannon RW. Reliability and validity of three strength measures obtained from community-dwelling elderly persons. J Strength Cond Res 2005; 19(3): 717-720.

[32] Tiedemann A, Shimada H, Sherrington C, Murray S and Lord S. The comparative ability of eight functional mobility tests for predicting falls in community-dwelling older people. Age Ageing 2008; 37(4): 430-435. http://dx.doi.org/10.1093/ageing/afn100

[33] Bohannon RW, Shove ME, Barreca SR, Masters LM and Sigouin CS. Five-repetition sit-to-stand test performance by community-dwelling adults: A preliminary investigation of times, determinants, and relationship with self-reported physical performance. Isokinetics and exercise science 2007; 15(2): 77-81.

[34] Nordin E, Rosendahl E and Lundin-Olsson L. Timed "Up and Go" Test: Reliability in Older People Dependent in Activities of Daily Living-Focus on Cognitive State. Phys Ther 2006; 86: 646-655.

[35] Bland JM and Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. Comput Biol Med 1990; 20: 337-340. http://dx.doi.org/10.1016/0010-4825(90)90013-F

[36] Domholdt E. Physical Therapy Research: Principles and Application. Philadelphia, Pa: WB Saunders Co 2000; 221-238.

[37] Atkinson G and Nevill A. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998; 26: 217-238. http://dx.doi.org/10.2165/00007256-199826040-00002

[38] Shrout PE and Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979; 86: 420-428. http://dx.doi.org/10.1037/0033-2909.86.2.420

[39] Currier DP. Elements of Research in Physical Therapy. 3rd ed. Baltimore, Md: Williams and Wilkins 1990; 167.

[40] Portney LG and Watkins MP. Foundations of Clinical Research: Applications to Practice. Upper Saddle River, NJ: Prentice-Hall Inc 2000: 557-586.

[41] Nunnally JC and Bernstein IH. Psychometric theory (3rd ed.). New York: McGraw Hill 1994.

Received on 30-12-2014 Accepted on 24-01-2015 Published on 31-12-2015

DOI: http://dx.doi.org/10.12974/2309-6128.2015.03.01.3