2008-5435/13/52-82-88
INTERNATIONAL JOURNAL OF OCCUPATIONAL HYGIENE

IJOH 5: 82-88, 2013

ORIGINAL ARTICLE

Intra-observer and Inter-observer Reliability in Direct Anthropometry

NASER HASHEMI-NEJAD[1], ALIREZA CHOOBINEH[2], MOHAMMAD REZA BANESHI[3], and AKRAM JAFARI-ROODBANDI[4*]

[1]Department of Occupational Health, School of Public Health, Kerman University of Medical Sciences, Kerman, Iran; [2]Professor of Occupational Health, Ergonomics Department, School of Health and Nutrition, Research Center for Health sciences, Shiraz University of Medical Sciences, Shiraz, Iran; [3]PhD in Biostatistics, Research Center for Modeling in Health, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran; [4]MSc Student of Occupational Health, Member of Kerman Medical Student Research Committee, Kerman University of Medical Sciences, Kerman, Iran.

Received June 15, 2012; Revised August 19, 2012; Accepted August 30, 2012

This paper is available on-line at *http://ijoh.tums.ac.ir*

ABSTRACT

This study aimed to evaluate the reliability factors, and identify causes of error in direct anthropometry method. After training three beginner anthropometrists and following the instructions of anthropometric standards, 48 body dimensions of 42 male students were measured three times. In other words, the physical dimensions of each subject were measured for 9 times. All participants were wearing uniforms during anthropometry, with bare feet. Differences in values of Repeated Measurement Test were explored using SPSS software version 11. The same software was employed to evaluate, through calculating ICC index, the correlation between anthropometrists. Inter-observer repeated measurement test showed significant difference in the measurements taken in 3, 7 and 1 dimension(s) by the three anthropometrists. The average measurement was significantly different at 16 dimensions; this, however, showed no difference at 32 dimensions. Measurements taken by anthropometrist 1 had ICC values of 0.26 (Min) and 0.99 (Max); these values were 0.48 (Min) and 1.00 (Max) for anthropometrist 2 and 0.23 (Min) and 0.98 (Max) for antropometrist 3. The maximum and minimum values of ICC index in all three anthropometrists were respectively close to and above 0.98, and lower than 0.5. High value of ICC in the measured dimensions indicated high reliability of repeated measurements. The decreasing value of some indexes can be attributed to such factors as random error, poor design of measurements tool (which in turn leads to random error), the long time devoted to measurement process, high number of dimension measured, changes in posture of subjects and deviation from the standard position.

Keywords: *Measurements errors, Reliability, Direct anthropometrical*

INTRODUCTION

Anthropometry is the measurement of external morphological traits of the human body [1]. Most of body landmarks used in traditional (direct) anthropometry have to be identified through touching by an anthropometrist and then measured [2]. Anthropometry plays an important role in examining people's nutritional and health status and designing workstations and any other tools related to humans [2 - 6]. Despite many benefits of direct anthropometry including its quickness and low cost and also lack of need for sophisticated equipment, this method has some

*** Corresponding author:** Akram Jafari Roodbandi, E-mail: ergonomic.jafari@gmail.com

Table 1. Qualitative classification of inter-class correlation (ICC) values as degrees of agreement beyond chance

ICC value	Degrees of agreement (reliability) beyond chance
0	None
>0 -<0.2	Slight
0.2 - < 0.4	Fair
0.4 - < 0.6	Moderate
0.6 - < 0.8	Substantial
0.8 - 1.0	Almost perfect

Table 2. Mean of weight and age of the participant (n=42) and mean of the measurement time

Variable	Average	Standard deviation	Min	Max	SE Mean	Variance
Weight (kg)	66.3	9.4	41.6	88.7	0.8409	89.101
Age (year)	20.6	3.04	15	32	0.2713	9.274
Measurement time(min)	10.3	3.8	5	33	0.34	14.82

Table 3. ICC maximum, minimum, and standard deviation for each anthropometrist

Inter-class Correlation Coefficient	Minimum	Maximum	Mean	SD
First anthropometrist Intra-observer	0.26	0.99	0.8	0.16
Second anthropometrist Intra-observer	0.48	0.99	0.84	0.10
Third anthropometrist Intra-observer	0.23	0.98	0.83	0.16
Inter- observer	0.91	1	0.80	0.20

inherent limitations such as the need to train anthropometrists and high error between the measurements and the mechanical constraints [1, 7]. If these errors are not considered in applying anthropometric data, such data will lose their validity; evidently, designing based on these data results in physical imbalance between the user and the product. Other consequences would be human error, reduced efficiency and musculoskeletal disorders [8, 9].

Maximum standard error and minimum correlation efficient is the method proposed for measuring direct anthropometry [1]. Reliability of measurements directly affects the quality of the data obtained [10]. Besides, finding absolutely reliable clinical methods of measurements is a difficult task in that measuring tools differ and those who observe or measure are not free from instability or error; people's degree of responsibility may also be different [11]. Therefore, the repeated measurements of a quantitative amount in a case study may not be similar. The reasons may be the changing posture of the case during measurement or a difference in the process of measurement [12].

Reliability and validity are two terms used for describing measurement error [7]. Reliability is considered as the degree of consistency and reproducibility of measurements which is used in various situations. Validity refers to the degree of closeness of the measured value and the actual value of the variable [11, 13, 14].

The error is actually the different between the measured and the actual value and is statically attributed to all those resources that one cannot characterize using independent variables [11]. Generally, measurement errors are the total sum of systematic errors and random errors. The former refers to predictable errors biased toward a specific direction. Such errors do not affect the reliability of measurements but rather impact the validity of these measurements because in such a situation the measured value will be different from the actual value [11, 15]. Random errors simply occur by chance; they are unpredictable and accordingly constitute the real meaning of reliability [11].

Some researchers have undertaken the issue of validity factors in nutritional and health studies, but none of these studies have addressed body dimensions in Ergonomics. Therefore, the objective of the present study was to evaluate the reliability and its affecting factors. Furthermore, it tried to identify causes of error in direct anthropometry.

MATERIALS AND METHODS

In order to evaluate the method of direct anthropometry, a repeated cross-sectional study was undertaken. In this study three anthropometrist including two BSc and one MSc students of Occupational Health from Kerman University of Medical Sciences did the measurement after receiving appropriate training.

Each of these beginner anthropometrists measured 48 body dimensions of 4 individuals for three times so as to master the method. In order to ensure the compatibility of the performance of these beginner anthropometrists with the anthropometric standards, they were allowed to consult with one another; Additionally, in cases where there was a mismatch in measurements, identification of anatomic features and the use of tool or when measures were not taken in the

Published online: April 8, 2013

Table 4. The results of calculated intra-group correlation coefficient and the significant difference between intra-observer and inter-observer for each anthropometrist

		Intra-class correlation coefficients (ICC)				**P value Repeated Measurement Test**			
No	Measured dimensions	*Observer1*	*Observer2*	*Observer3*	*Inter observer*	Observer1	Observer2	Observer3	Inter observer
1	Height	0.93	0.99	0.97	0.99	0.67	0.0001	0.14	0.32
2	Eye height, standing	0.97	0.98	0.98	0.99	0.44	0.48	0.41	0.32
3	Shoulder height(acromion) standing	0.99	0.99	0.75	0.99	0.05	0.17	0.2	0.32
4	Elbow height, standing	0.94	0.92	0.85	0.99	0.16	0.18	0.64	0.32
5	Wrist height, standing	0.85	0.78	0.92	0.99	0.58	0.68	0.38	0.32
6	Hip height(trochanter), standing	0.88	0.83	0.85	0.99	0.92	0.25	0.48	0.32
7	Tibia height	0.65	0.8	0.23	0.99	0.6	0.98	0.35	0.32
8	Kuckle height, standing	0.8	0.84	0.87	0.99	0.12	0.63	0.21	0.32
9	Fingertip height, standing	0.93	0.86	0.96	0.99	0.41	0.44	0.14	0.32
10	Sitting height	0.95	0.96	0.94	0.99	0.20	0.98	0.75	0.31
11	Eye height, sitting	0.67	0.87	0.68	0.94	0.43	0.37	0.35	0.32
12	Elbow height, sitting	0.52	0.62	0.62	0.97	0.35	0.56	0.62	0.32
13	Thickness of thigh	0.58	0.88	0.94	0.99	0.88	0.03	0.96	0.32
14	Shoulder height, sitting	0.66	0.82	0.88	0.97	0.64	0.4	0.1	0.32
15	buttock –knee depth, stting	0.79	0.88	0.77	0.98	0.77	0.97	0.73	0.32
16	buttock –popliteal depth, stting	0.63	0.82	0.68	0.99	0.56	0.48	0.93	0.32
17	Deep abdominal	0.9	0.93	0.62	0.99	0.88	0.84	0.54	0.15
18	Deep posterior abdominal, sitting	0.94	0.96	0.96	0.98	0.18	0.53	0.08	0.38
19	Knee height, sitting	0.39	0.92	0.77	0.98	0.81	0.5	0.21	0.51
20	Popliteal height	0.61	0.8	0.78	0.94	0.06	0.28	0.2	0.39
21	Shoulder breadth (bideltoid)	0.9	0.87	0.96	0.93	0.37	0.28	0.44	0.14
22	Shoulder breadth (biacromial)	0.63	0.73	0.82	0.99	0.78	0.58	0.13	0.02
23	Hip breadth	0.87	0.88	0.8	0.99	0.58	0.7	0.18	0.04
24	Chest depth	0.8	0.8	0.78	0.99	0.38	0.5	0.099	0.09
25	Hip depth	0.8	0.75	0.93	0.99	0.78	0.72	0.49	0.57
26	Abdominal depth	0.93	0.93	0.96	0.99	0.34	0.04	0.88	0.047
27	shoulder-elbow length	0.66	0.79	0.87	0.99	0.91	0.93	0.29	0.0001
28	Elbow-fingertips length	0.82	0.66	0.92	0.98	0.80	0.39	0.76	0.0001
29	Upper extremity length	0.8	0.48	0.39	0.99	0.03	0.30	0.28	0.4
30	shoulder-clutch length	0.78	0.88	0.5	1	0.05	0.62	0.15	0.13
31	Head length	0.87	0.88	0.89	0.99	0.48	0.61	0.82	0.3
32	Head breadth	0.71	0.64	0.68	0.98	0.28	0.032	0.17	0.4
33	Hand length	0.67	0.76	0.9	0.99	0.58	0.54	0.31	0.0001
34	Hand breadth	0.69	0.76	0.82	0.99	0.5	0.003	0.72	0.0001
35	Metatarsus length	0.26	0.93	0.94	0.99	0.44	0.01	0.97	0.04
36	Metatarsus breadth	0.71	0.77	0.82	0.95	0.31	0.5	0.01	0.0001
37	Neck circumference	0.96	0.85	0.84	0.99	0.3	0.19	0.45	0.35
38	Arm circumference	0.98	0.98	0.96	0.96	0.001	0.08	0.1	0.001
39	Chest circumference	0.95	0.69	0.97	0.91	0.2	0.67	0.24	0.014
40	Waist circumference	0.98	0.96	0.97	0.99	0.58	0.44	0.25	0.0001
41	Wrist circumference	0.92	0.89	0.87	1	0.81	0.46	0.88	0.0001
42	Thigh circumference	0.89	0.87	0.97	0.96	0.25	0.04	0.27	0.64
43	Calf circumference	0.97	0.95	0.96	0.99	0.97	0.54	0.44	0.01
44	Span	0.56	0.87	0.81	0.99	0.16	0.8	0.22	0.44
45	Elbow span	0.89	0.9	0.9	0.97	0.35	0.59	0.18	0.004
46	overhead grip reach, standing	0.85	0.96	0.95	0.95	0.16	0.31	0.37	0.09
47	overhead grip reach, sitting	0.88	0.86	0.91	0.99	0.15	0.12	0.77	0.42
48	forward grip reach, sitting	0.89	0.8	0.8	0.96	0.71	0.9	0.6	0.005

Published online: April 8, 2013

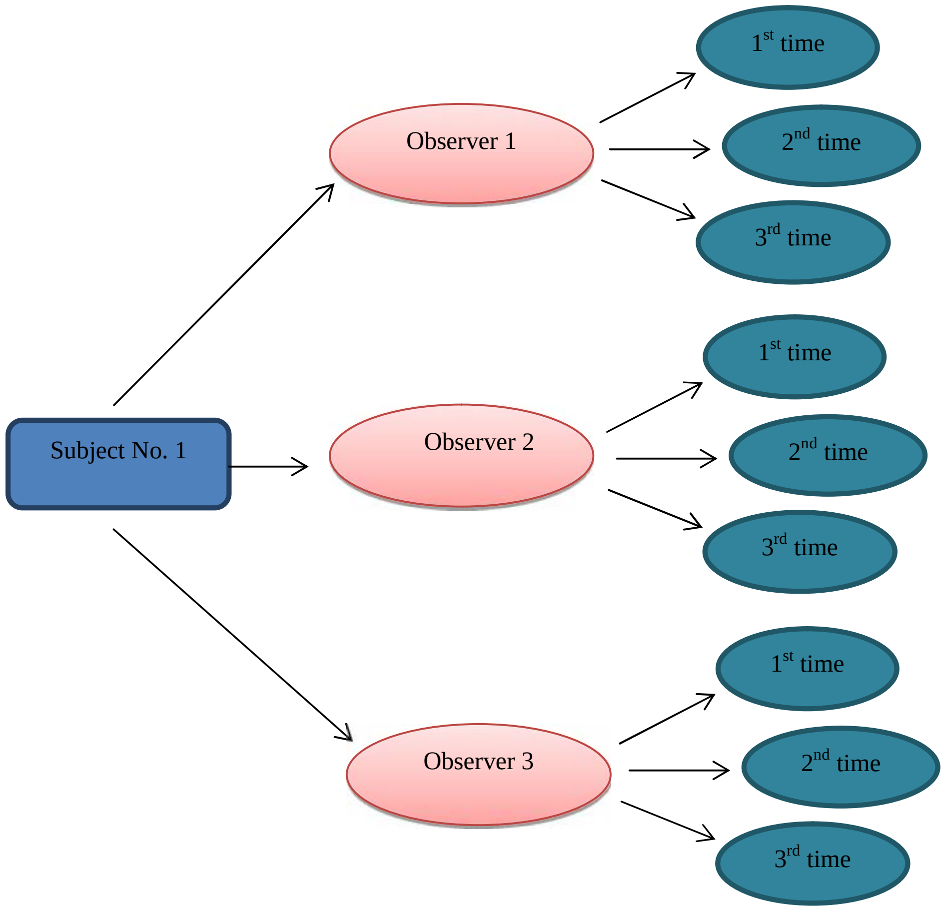

Fig 1. Pattern of measurement of body dimensions of each subject by observers

standard body positions, they were provided with the necessary instructions. These measures were not included in the final analysis of the data. Having completed this stage, reading errors and errors in correction of body position were eliminated; besides, the significant differences in measurements were disregarded.

In this study, 48 body dimensions of 42 randomly selected male students of Kerman University of Medical Sciences and high school of Kerman from 15 to 32 who did not suffer from any musculoskeletal disorder were measured.

Each observer evaluated 48 body dimensions for 3 times. In fact, body dimensions of subject were measured 9 times based on the pattern show in Fig 1.

All the measurements were carried out in Ergonomic Laboratory of Occupational Health Department of Kerman University of Medical Sciences. The tools used in the present study included stadiometers, tape measures and calipers with 1 mm of precision. A digital scales with then precision of 0.1 kg was to measure the subjects' weight.

Participants were asked to wear uniforms so as to eliminate the effect of individuals' clothing and to make it easy to detect anatomic features. In addit, measurements were carried out with bare feet (Fig 2 and 3).

In order to eliminate the effects of changes in weight and body dimensions, all measurements were done manually and participants were requested to refrain from eating and drinking between measuring sessions. Considering the fact that observers might use their short memories and consequently change the data to make them homogenous, they were asked to measure a series of 48 body dimensions consequently and only then repeat their measurements. In other words, observers were not allowed to measure a single body dimension three times, one after another.

Each observer read the numbers loudly and a co-observer recorded it in the measurement sheet. The co-observer read the measurement loudly to rule out any chance of misunderstanding.

After measurement stage, two persons entered data into SPSS software to ensure the optimum reliability; one read the numbers and the other entered them into the software. In addition, outlier data were compared with measurement sheet to be corrected in case of any mismatch.

Published online: April 8, 2013

Fig 2. Standard body position in study (sitting)

Fig 3. Standard body position in study (standing)

Statistical methods

When a variable is measured for more than two times or such measures are done by more than one group, Repeated Measurement Test (RMT) is used which is the variance analysis of successive measurements [16]. What makes the result of this study significant is the difference either in the measured values of each anthropometrist (intra-observer) or in the measured values of all three anthropometrists (inter-observer).

ICC (Inter-Class Correlation Coefficient) is an index for the evaluation of reliability. ICC is used for identifying the relation between two quantitative variables in one group or class. A value of 0.95 indicates that 95% of variance in measures is attributed to the real variance among the participants, and the remaining 5% is related to either the error of measurements or the variance between participants and observers (Table 1) [17].

RESULTS

In this study, examining 42 male students, the mean weight was found to be 66.33 kg (SD=9.43) where the mean age was 20.63 (SD=3.04). The average time for a single measurement of 48 body dimensions amounted to 10:33 (SD=3.85) minutes with the max of 33 minutes and minimum of 5 minutes (Table 2).

The average, maximum and minimum values of ICC indexes are presented in Table 3. As shown in this table, the value of ICC for the first anthropometrist is between 0.26 and 0.99, for the second anthropometrist, this value is between 0.48 and 1.00 and for the third one is between 0.23 and 0.98. The maximum value for each anthropometrist was higher than 0.98; the minimum value for the three anthropometrists equaled to 0.1. The standard deviation of the observer of ICC index turned out to be less than the standard deviation of ICC in the observers of all three anthropometrists (ICC=0.01) (Table 3).

ICC index value of below 0.5 for the first anthropometrist in knee height and the metatarsal length were 0.93 and 0.26 respectively; for the second anthropometrist in the dimension of upper extremity length ICC was equal to 0.48; finally for the third anthropometrist, the lowest ICC index values for tibia height and upper extremity length were 0.23 and 0.39 respectively.

The present study examined 48 body dimensions of 42 participants three times. The result of RMT showed that the three measurements done by the observer 1 regarding dimensions of the length of scapula-clutch, upper extremity length and arm circumference have had a significant difference ($p<0.05$).Observer 3 measured 7 dimensions, including height, deep abdominal, head width, hand width, metatarsus length and thigh circumference; the results of this observer showed a significant difference. Measures taken by observer 3 showed a significant difference only in metatarsus length ($p<0.05$). An RMT in the mean measures of the three observers illustrated that the measures of 16 body dimensions were significantly different while the remaining 32 dimensions didn’t show any significant difference (Table 4) ($p<0.05$).

DISCUSSION

This study aimed at evaluation of intra/inter-observer reliability and measured 48 body dimensions by three beginner anthropometrists. It sought to bring into light the fact that statistical indexes of each single body dimension was unique. In other words, the performance of each anthropometrist which is different from both his previous performances and his colleagues' performances would have an effect on the way body dimensions are measured. It also affects the measurement tools.

Published online: April 8, 2013

Fig 6. Deviation from the standard position during anthropomethry

Most of the previous studies measured indexes of growth, nutrition and physiotherapy, yet none had addressed reliability and factors involved in measuring reliability of anthropometry in Ergonomics.

Measuring the index of height in Malaysian children under 2 years old demonstrated a high level of intra/inter-observer reliability in measuring height. ICC index was equal to 0.99. Statistical tests showed no significant difference in repetitive measurements ($p<0.05$) [18]. In this study, the intra-observer ICC indexes of three anthropometrists were 0.93, 0.99 and 0.97, respectively. The inter-observer ICC index was 0.99. The RMT showed no significant difference in each of the measurements done by the three anthropometrists ($p=0.23$). 1

In Kenya, anthropometric dimensions of children under the age of six months were measured, and ICC index=0.6 was considered as the minimum acceptable reliability [19]. The lowest value of ICC for arm circumference measured by experienced anthropometrists was 0.97. This value was equal to 0.88 for the beginner anthropometrists in our study. The value obtained by the latter group was 3.5 mm (95% ICC: 2.5-4.4) more than that of the former group. This difference was not, however, statistically significant [19]. The ICC index was 0.96 which was similar to the results of Mwangoma et al. [19].

Still in another study, the reliability of body dimensions of 130 people over the age of 60 were measured; calf circumference was one of these dimensions. The results marked a high correlation between these observers in their measurements, in a way that ICC index amounted to 0.99 [20]. ICC indexes of anthropometries carried out by intra-observer were 0.97, 0.95, and 0.96; the ICC index obtained by inter-observer was 0.99. The data seems to resemble the aforementioned study.

Nordhman et al. measured ICC index to examine the reliability of measurements done by two observers. The observers measured dimensions of height and waist circumference as nutritional indicators. Values of ICC index in these dimensions were equal to 1 and 0.97, respectively [21]. In the present study, both of these values were 0.99.

In some of the measured dimensions, despite the high value of ICC index, RMT showed a significant difference. The mean values of three anthropometries indicated that in Ergonomics, it is not possible to determine the reliability of each anthropometry, though a 0.5 cm difference has made the result of the test significant.

The special design of the stadiometer used in this study poses the possibility of reading error from dials. For example, as seen in Fig 1, the closeness of numbers and arrows leads anthropometrist to misread the actual numbers (anthropometrist may be puzzled what the numbers refer to); that is while a single point in the line could make anthropometrist more certain in reading numbers (Fig 4).

One other source of reading error was the inverse numbering on the calipers; as an example one may mistake 9 for 6. Two and three-digit numbers make the situation even worse. It weakens the reliability of one's measurements and makes it different from those of his/her colleagues (Fig 5).

Deviation from the standard position is another factor that insidiously influences reliability of direct anthropometry. As it is illustrated in Fig 6, when measuring a dimension like that of middle finger, the participant leans to one side in order to be able to see

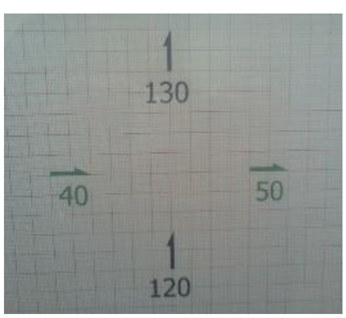

Fig 4. The way numbers are shown in the stadiometer used in this study

Fig 5. The way caliper used in this study is graduated

anthropometrist, which may have an erroneous effect on the actual value.

Another factor affecting an anthropometrist's repetitive measurements is tiredness. Three times measuring of 48 body dimensions which was carried out in this study has certainly been a cause of tiredness.

Generally in the most measured dimensions, high value of ICC indicated high reliability in repeated measurements. However, the decreasing amount of some indexes can be due to random error, poor design of measurement tools causing reading error, time-consuming measurement process, high number of measurement dimensions, changes in posture of participants and deviation from the standard position.

CONCLUSION

Different studies have adopted different terminologies and frequency indexes. Further researches may be needed to analyze these indexes to propose a more valid index by which to determine the reliability of intra/inter-observers' performances. This study measured many body dimensions which might lead to anthropometrists' tiredness; for this reason it could be suggested that less body dimensions be measures for such studies.

ACKNOWLEDGMENT

This article was entranced from the thesis written by Miss Jafari Roodbandi Msc Student of Occupational Health and was financially supported by Kerman Medical Student Research Committee. Authors wish to express their deepest thanks to all the students who participated in this study. The authors declare that there is no conflict of interests.

REFERENCES

1. Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr* 1999; 82(03):165-77.
2. Kouchi M, Mochimaru M. Errors in landmarking and the evaluation of the accuracy of traditional and 3D anthropometry. Applied Ergonomics. 2011; 42(3):518-27.
3. Norton KI, Olds T. Anthropometrica: a textbook of body measurement for sports and health courses: NewSouth Publishing; 1996.
4. Karwowski W. International encyclopedia of ergonomics and human factors. Vol. 3: Taylor & Francis US; 2006.
5. Micozzi MS. Applications of anthropometry to epidemiologic studies of nutrition and cancer. *American Journal of Human Biology* 1990; 2(6):727-39.
6. Clarke G, Whittemore AS. Prostate cancer risk in relation to anthropometry and physical activity: the National Health and Nutrition Examination Survey I Epidemiological Follow-Up Study. *Cancer Epidemiology Biomarkers & Prevention* 2000; 9(9):875-81.
7. Geeta A, Jamaiyah H, Safiza M, Khor G, Kee C, Ahmad A, et al. Reliability, technical error of measurements and validity of instruments for nutritional status assessment of adults in Malaysia. *Singapore Medical Journal* 2009; 50(10):1013.
8. Sadeghi NH, Heidaripour M. Kansei engineering and ergonomic design of products. *Int J Occ Hyg* 2011; 3(2), 81-84.
9. Mirmohammadi SJ, Mehrparvar AH, Jafari S, Mostaghaci M. An Assessment of the Anthropometric Data of Iranian University Students. *Int J Occ Hyg* 2011; 3(2), 85-89.
10. Onis M. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. *Acta Paediatrica* 2006; 95(S450):38-46.
11. Bruton A, Conway JH, Holgate ST. Reliability: what is it, and how is it measured? *Physiotherapy* 2000; 86(2):94-9.
12. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet* 1986; 327(8476):307-10.
13. Harris EF, Smith RN. Accounting for measurement error: a critical but often overlooked process. Archives of Oral Biology. 2009; 54:S107-S17.
14. Saw S, Ng T. The design and assessment of questionnaires in clinical research. *Singapore Medical Journal* 2001; 42(3):131.
15. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Medicine 1998; 26(4):217-38.
16. Chehrei A, Haghdoost AA, Fereshtehnejad SM, et al. Statistical method in medical science researches using SPSS software. 1st ed, Pejvak Elm Arya Publishing Co.,Tehran, Iran,2011.[Persian]
17. Lewsey J. Medical Statistics: A Guide to Data Analysis and Critical Appraisal. Annals of The Royal College of Surgeons of England 2006; 88(6):603.
18. Jamaiyah H, Geeta A, Safiza M, Khor G, Wong N, Kee C, et al. Reliability, technical error of measurements and validity of length and weight measurements for children under two years old in Malaysia. *Med J Malays* 2010; 65:131-7.
19. Mwangome MK, Fegan G, Mbunya R, Prentice AM, Berkley JA. Reliability and accuracy of anthropometry performed by community health workers among infants under 6 months in rural Kenya. *Trop Med Int Health* 2012; 17(5):622-9.
20. Jamaiyah H Jr, Geeta A, Safiza MN, Wong NF, Kee CC, Ahmad AZ, et al. Reliability and technical error of Calf Circumference and Mid-half Arm Span measurements for nutritional status assessment of elderly persons in Malaysia. Malaysian Journal of Nutrition 2008. 14(2): p. 137-150.
21. Nordhamn K, Södergren E, Olsson E, Karlström B, Vessby B and Berglund L, et al. Reliability of anthropometric measurements in overweight and lean subjects: consequences for correlations between anthropometric and other variables. *Int J Obes* 2000. 24: p. 652-657.