Mal J Nutr 14(2): 137 - 150, 2008

Reliability and Technical Error of Calf Circumference and Mid-half Arm Span Measurements for Nutritional Status Assessment of Elderly Persons in Malaysia

Jamaiyah H [1], Geeta A [1], Safiza MN [2], Wong NF [2], Kee CC [3], Ahmad AZ [2], Suzana S [4], Rahmah R [5], Khor GL [6], Ruzita AT [4], Chen WS [1], Rajaah M [1] & Faudzi A [2]

[1] *Clinical Research Centre, Hospital Kuala Lumpur, Malaysia*
[2] *Institute for Public Health, Kuala Lumpur, Malaysia*
[3] *Institute for Medical Research, Kuala Lumpur, Malaysia*
[4] *Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia*
[5] *Hospital Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia*
[6] *Universiti Putra Malaysia, Serdang, Selangor, Malaysia*

ABSTRACT

This study sought to examine the reliability of two measurements; Calf Circumference (CC) and Mid-half Arm Span (MHAS). A sample of 130 elderly persons aged 60 years and above seen consecutively in the Kuala Lumpur Hospital outpatient clinic during the period of December 2005 to January 2006, upon consent, were recruited to the study. There was a high degree of reliability for both inter- and intra-examiner (r close to 1). For inter-examiner, on average the CC measurements taken by the first examiner were 0.3 cm lower than that of the second examiner. The upper and lower limit of the differences were +0.4 to –0.9 cm respectively. Inter-examiner MHAS measurements on average by the first examiner were 0.2 cm lower than that of the second examiner. The limits were +1.7 to –2.1 cm. By comparison, the inter-examiner CC measurements were more reliable than the MHAS measurements. For intra-examiner, on average the CC measurements at Time 1 were consistent with Time 2 (mean difference=0) with limits of the difference at $\pm$ 0.5 cm. MHAS measurements at Time 1 were on average 0.1 cm less than at Time 2 with limits at +1.7 and –1.8 cm. The technical error of measurement (TEM) and coefficient of variation of CC and MHAS for both inter-examiner and intra-examiner measurements were within acceptable limits with the exception of MHAS TEM. This study suggests that CC and MHAS measured in elderly persons 60 years and above, using Seca Circumference Tape ® 206, Germany (0.05 cm) are reliable and can be used in a community survey.

INTRODUCTION

Anthropometry has a very important role in nutritional screening, surveillance and monitoring (Ulijaszek & Kerr, 1999). Body weight, height, various skin-fold thicknesses and circumferences and other linear dimensions characterise the nutritional status of an individual. Anthropometry is a relatively simple, quick, inexpensive means of nutritional assessment that can be used in the laboratory, clinical and community settings.

Correspondence author: Dr Jamaiyah Haniff; Email: *jamaiyah@crc.gov.my*

In assessing the nutritional status of the elderly, since direct measurement of body composition is not possible in a large number of patients, calf circumference is a pertinent marker of the nutritional status in the elderly (Bonnefoy *et al*., 2002).

However, anthropometry methods have their inherent limitations; the need for trained examiners, relatively high between-measurement technical errors and mechanical limitations (NYORC, 2006). Among various measurement methods, anthropometry techniques usually demonstrate the largest standard error and lowest correlation coefficients when compared against other techniques. Various terms are used to describe anthropometric measurement errors that include reliability and validity (Ulijaszek & Kerr, 1999). Reliability is the degree to which within-subject variability is present and is due to factors other than variance of measurement error or physiological variation. The second type of measurement error, validity, is the extent to which the 'true' value of a measurement is attained. Measurement quality and control dimension are commonly expressed using an accuracy index of technical error of measurement (TEM). TEM allows examiners to verify the degree of accuracy while performing and repeating the measurements (intra-examiner) and when it is compared with measurements from other examiners (inter-examiner) (Talita *et al.,* 2005).

While much has been written on anthropometric assessment methods and interpretation, relatively few articles have discussed reliability and TEM issues and the extent of these factors which influence both measurement and interpretation of nutritional status (Ulijaszek & Kerr 1999).

In assessing the nutritional status of the elderly, choice of mid-half arm span (MHAS) is the best practical alternative where the height measurement cannot be obtained due to spinal curvature or poor muscle tone (Pieterse, Manandhar & Suraiya, 1998). Height and demi-span (sternal notch to finger roots with arm out-stretched laterally) have been found to have significant correlations (r = 0·74) in 125 normal young to middle-aged European subjects (Bassey, 1986). The authors also deduced that demi-span is therefore a practical alternative measurement of skeletal size. Calf circumference (CC) measurement is commonly used to determine the nutritional state of elderly persons.

Since there are a few articles on the reliability of calf circumference and mid-half arm span, this paper aims to assess the reliability of these measurements which are used for nutritional status assessments for persons 60 years and above in the National Health and Morbidity Survey III (NHMS III).

METHODS

This was a cross-sectional study. A convenient sample of elderly persons, seen consecutively in the Kuala Lumpur Hospital outpatient clinic during the period of December 2005 to January 2006, were chosen. The sample criteria were those aged 60 years and above and who gave consent to participate. The exclusion criteria were elderly persons with obvious physical disabilities and body deformation that inhibits ability to stand upright, including kyphosis. The sample size was determined (Walter, Eliasziw & Donner, 1998), with two replicates per subject, the expected reliability coefficient (r) of at least 0.8 (H_1: ρ_1=0.8), the reliability of 0.7 (H_0: ρ_0=0.7) or higher to be minimally acceptable, α=0.05 and β=0.2 (corresponds to 80% power); this would require a total number of 117.1 subjects. Using a 10% over-estimate to account for poor response, the final sample size was 130.

Two students from a local university undergoing a Dietetics Masters programme, conducted the measurements of CC and MHAS. There were no special reasons for choosing only two examiners instead of more. A pair was the minimum number of examiners needed for inter-examiner reliability. The decision was made entirely

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

ICC value	*Degree 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

(Source: MUSC, 2006)

based on logistics convenience. To avoid many measurements on the same subjects, the decision to assess intra-examiner reliability in only one of the two examiners, was deliberately made. Selection of one of the examiners for intra-examiner reliability testing was also not influenced in any way.

Each subject was examined thrice for the same measurements; 3 times for CC and 3 times for MHAS. The order was as follows: examiner #1 measures the subject for CC and MHAS, then the subject goes to examiner #2 who would measure the subject again for CC and MHAS. The subject would then return to examiner #1 for the last CC and MHAS measurements. Both examiners were not part of the research team and were therefore blind as to the exact aim of the study. They were requested not to recall their previous readings. The data capture form was designed in such a way that recordings of previous readings were obscured immediately after each recording was made to minimise recall bias.

Study specific procedure

Both the examiners were trained to adhere to the standard procedure as in the Technical Manual of NHMS III (NHMS III, 2006). Calf circumference was measured to 0.05 cm accuracy (WHO, 1995). Mid-half Arm Span was measured to 0.05 cm accuracy (WHO, 1995). Both these measurements were made using the Seca Circumference Tape ® model 206 ® Germany. The examiners were given a tape each.

Statistical analysis

For reliability, findings of statistical analyses are reported using absolute mean difference, correlation coefficient (*r*) and Bland and Altman plot (Bland & Altman, 1986). Absolute mean difference is a crude way of checking for difference or agreement between two readings. We also tested the difference for significance using independent *t*-test and paired *t*-test for absolute means. Correlation coefficient (*r*) was used as a more objective way of assessing reliability. Correlation coefficient was computed using intra-class correlation to demonstrate the strength of the relationship (similarities) between two measurements. The values for reliability coefficient range from 0 to 1. A coefficient of 0 indicates 'no reliability' and 1 indicates 'perfect reliability' (Table 1).

A Bland and Altman plot was used to provide an illustration of the spread of differences in readings, the mean difference and the upper and lower limit of agreement both for inter- as well as intra-examiner reliability. There is no 'acceptable' range for Bland and Altman plots.

The technical error of measurement (TEM) is an accuracy index and represents

Table 2. Relative TEM values considered as acceptable

Type of analysis		*Beginner anthropometrist* (≤)	*Skilful anthropometrist* (≤)
Intra-examiner	Skin folds	7.5%	5.0%
	Other measures	1.5%	1.0%
Inter-examiner	Skin folds	10%	7.5%
	Other measures	2.0%	1.5%

Norton K & Olds T (eds)(2000). *Antropometrica Argentina Biosystem.* In Talita AP *et al.* (2005).

the measurement quality and control dimension. It is the most common way to express the margin of error in anthropometry and has been adopted by the International Society for Standardization Advancement in Kinanthropometry (ISAK). It is being used by anthropometrists in Australia (Talita *et al.,* 2005). It is essentially the standard deviation between repeated measures. The TEM index allows anthropometrists to verify the degree of accuracy when performing and repeating anthropometric measurements (intra-examiner) and when comparing their measurement with measurements from other anthropometrists (inter-examiner). The acceptable ranges for Relative TEM are as shown in Table 2.

Inter-examiner reliability refers to how consistently readings from the two examiners on the same subjects are in agreement. Intra-examiner reliability refers to how consistently readings from the same examiner on the same subjects, but at two different time points, are in agreement.

In addition to these, the coefficient of variation (CV) is calculated to further determine the precision of methods of measurements. The CV provides a general 'feeling' about the performance of a method. CVs of 5% or less generally give us a feeling of good method performance, whereas CVs of 10% and higher are bad (Zady, 1999).

In order to compare the variability of the two methods, CC and MHAS, the percentage of coefficient of variation (% CV) was calculated using the data from both intre- as well as the intra-examiner. Percentage of coefficient of variation is therefore a good indicator to use when comparing methods (Martin, 2006).

RESULTS

Sample characteristics

The mean age of 130 elderly persons involved in this study is 66 (± 5.41) years. The median is also 66 years and the range is 60 to 85 years old (see Table 3). Two-thirds were male. Almost half (48%) belonged to the Chinese ethnic group followed by Malays (31%) and Indians 20%. From Figure 1, it is noted that a small *p*-value (<0.01) is obtained from the Shapiro-Wilk test, which indicates that the age distribution is not normally distributed. Generally there are more 'younger-old' compared to 'older-old' in the sample.

Reliability

1. Inter-examiner reliability

ˇAbsolute mean difference

The mean, median, range of measurements and absolute mean difference for examiner #1 and examiner #2 are illustrated in Table III. The absolute difference in mean was found to be very small; 0.3 for CC and 0.1 for MHAS. This indicates good agreement between the two examiners.

Table 3. Characteristics of sample

Characteristics	*Total=130*	
Age (year)		
Mean (SD)	66 (5.41)	
Median (Min ,Max)	66 (60, 85)	
	No	%
Gender		
Male	88	67.69
Female	42	32.31
Ethnic group		
Malay	40	30.77
Chinese	62	47.69
Indian	26	20.00
Others	2	1.54
Education level		
No education	12	9.23
Primary	73	56.15
Secondary	41	31.54
Tertiary	4	3.08

ˇCorrelation coefficient
Figures 1 and Figure 2 illustrate graphically the strong correlation between the readings from the two examiners for CC and MHAS respectively. For both CC and MHAS, the results indicate a high degree of reliability between the two examiners as the *r* is very close to 1.

• Bland and Altman plot
Figure 3 shows that on average, the CC measurement taken by the 2nd examiner is 0.3 cm lower than that for the 1st examiner with the limits of agreement ranging from –0.9 cm to +0.4 cm.

Figure 4 shows that on average, the MHAS measurement taken by the 2nd examiner is 0.2 cm lower than that for the 1st examiner and the upper limit of the agreement is about 2.1 cm while the lower limit agreement is 1.7 cm.

2. Intra-examiner reliability

ˇAbsolute mean difference
For within examiner analysis, the mean (SD), median and absolute mean differences are as shown in Table 5. For MHAS the difference in the mean is very small and for CC there is no difference at all.

ˇCorrelation coefficient
Figures 5 and Figure 6 illustrate the strong correlation between the readings from time 1 and time 2 by an examiner for CC and MHAS respectively. For both CC and MHAS, the results indicate a high degree of reliability within examiner with the *r* being very close to 1.

ˇBland and Altman plot
Figure 7 shows that on average, there is no mean difference in the CC measurement taken at time 1 and time 2; they are within the limits of agreement ranging from –0.4 cm to +0.5 cm.

Figure 8 shows that on average, the MHAS measurement taken from time 2 is 0.1 cm lower than that for time 1. The upper limit of the agreement is –1.8 cm while the lower limit agreement is 1.7 cm.

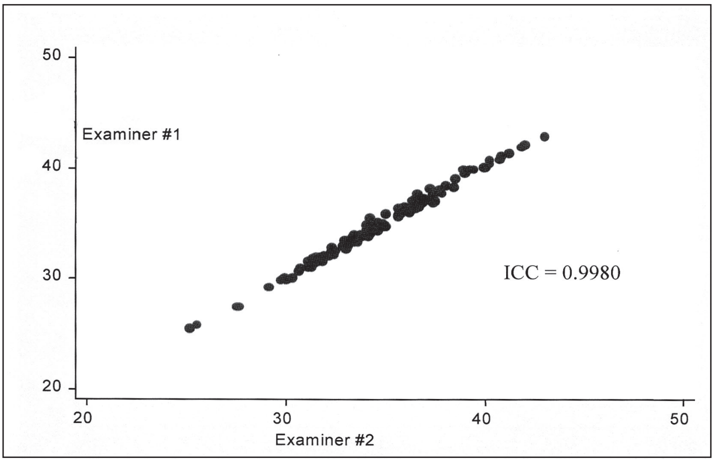


Figure 1. Scatter plot of the calf circumference measurements of examiner #1 versus examiner #2

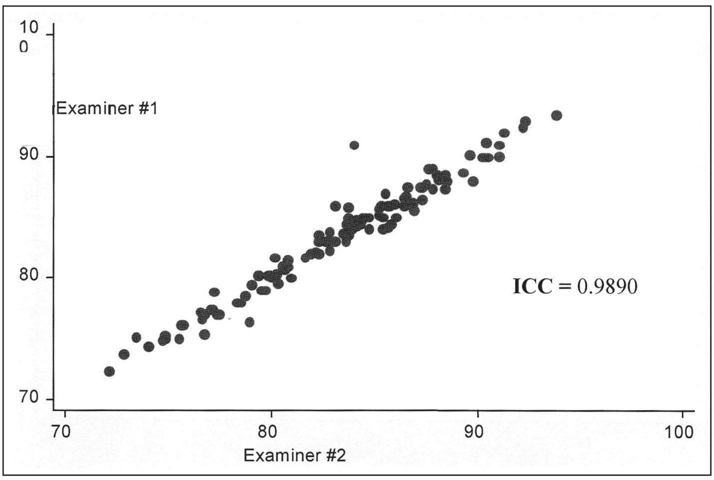


Figure 2. Scatter plot of the mid-half arm span measurements of examiner #1 versus examiner #2

Technical Error of Measurement (TEM)

The results for the TEM are tabulated in Table 6. The relative TEMs for inter- and intra-examiner for CC were 0.88% and 0.49% respectively. The relative TEMs for inter- and intra-examiner for MHAS were 0.83% and 6.35%.

Measurement comparison

Measurement comparison results are displayed in Table 7. We found that the CVs of MHAS are close to 5%, which is favourable, while that of CC are close to 10%, which is not favourable. However, it is to be noted that there is no difference in the CV in terms of inter- and intra-examiner.

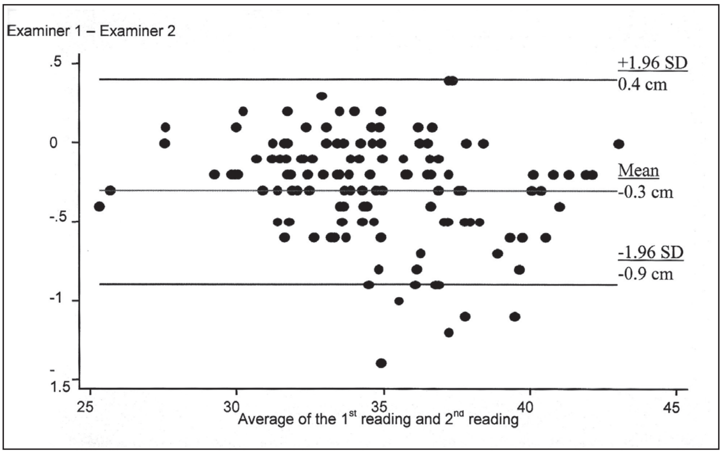


Figure 3. Bland Altman plot of the calf circumference measurements between examiners

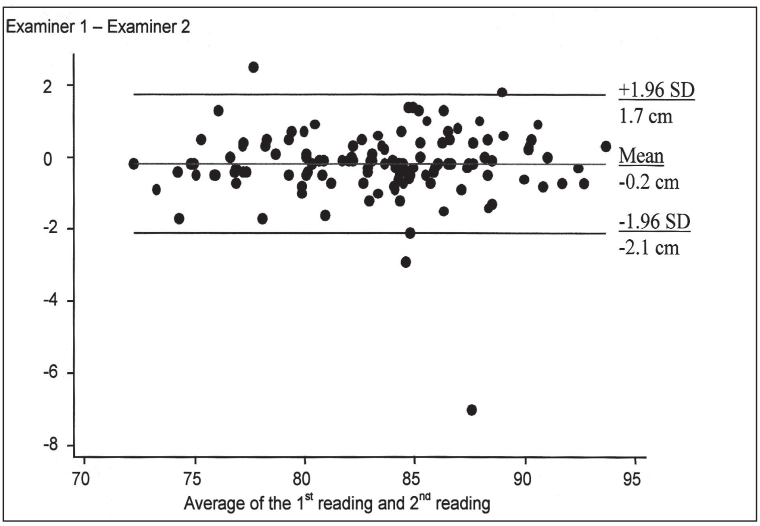


Figure 4. Bland Altman plot for the mid-half arm span measurements between examiners

DISCUSSION

Anthropometric measurements have different types of errors. This study in particular looks at imprecision as one of the components of reliability. Imprecision is the variability of repeated measurements and due to intra- and inter-examiner measurement differences (Ulijaszek, 1999). There are a few indices which are often used to assess the intra- and inter- examiner variability. These include technical error of measurement (TEM), coefficient variation (CV), coefficient of reliability (R), intra-class correlation coefficient (ICC) and Bland & Altman plot. In this study we report on all the above except for coefficient of reliability (R).

There are only a handful of published articles on measurement errors. There is one

Table 4. Summary statistics: mean (SD), median and absolute mean difference for inter-examiner reliability

Summary statistics	*Examiner 1(1)*	*Examiner 2(2)*	*Absolute mean diff (1)-(2)*	*P value*
Calf circumference, cm				
N	130	130		
Mean (SD)	34.8 (3.3)	34.5 (3.2)	0.3	0.407
Median (Min, max)	34.6 (25.5, 43.0)	34.2 (25.1, 43.0)		
Mid-half arm span, cm				
N	130	130		
Mean (SD)	83.4 (4.6)	83.2 (4.6)	0.1	0.698
Median (min, max)	84.0 (72.3, 93.5)	83.7 (72.1, 93.8)		

*P value is obtained by independent *t*-test.

Table 5. Summary statistics: mean (SD), median, range and absolute mean difference for intra-examiner reliability in Examiner #1.

Summary statistics	*Examiner # 1*		*Absolute mean difference (1) - (2)*	*P value **
	1st Measurement(1)	*2nd Measurement(2)*		
Calf circumference, cm				
N	130	130		
Mean (SD)	34.8 (3.3)	34.8 (3.3)	0.0	0.264
Median (Min, max)	34.6 (25.5, 43.0)	34.6 (25.8, 43.4)		
Mid-half arm span, cm				
N	130	129**		
Mean (SD)	83.4(4.6)	83.3 (4.6)	0.1	0.524
Median (Min, max)	84.0 (72.3, 93.5)	83.8 (72.4, 94.1)		

*P value is obtained by paired *t*-test.
** *1* subject refused to participate in the measurement at Time 2.

landmark review article (Ulijaszek, 1999) that captured 13 anthropometric measurements amongst which are CC and MHAS, with each having a range of 1 - 40 studies for inter-examiner reliability and 1 – 29 studies for intra-examiner reliability.

This study examined the reliability for CC in inter- and intra-examiner. As expected, the error between examiners was higher (0.3) than that of within examiner (0.0), since systematic between-examiner bias would contribute more measurement differences. For CC measurements, the absolute TEM found was 0.30 (inter) and 0.17 (intra). These results were then compared with two other studies, HHANES and NHES III, which found TEM values of 0.52 and 0.34 (inter) and 0.85 and 0.87 (intra) respectively. These values are not the same probably because of different age groups; 12-73 years in

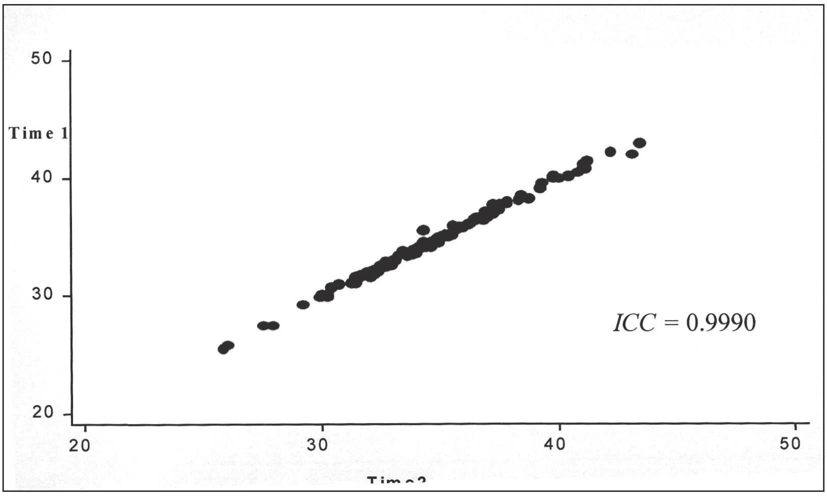


Figure 5. Scatter plot of the calf circumference measurements in examiner #1 at time1 versus time 2

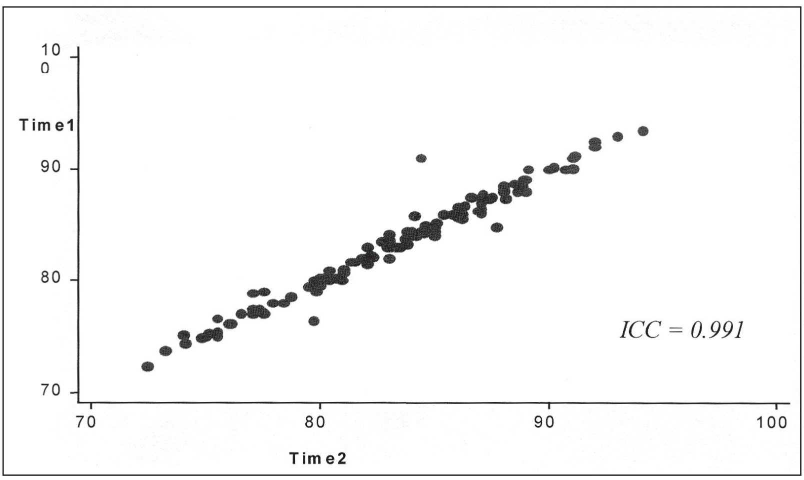
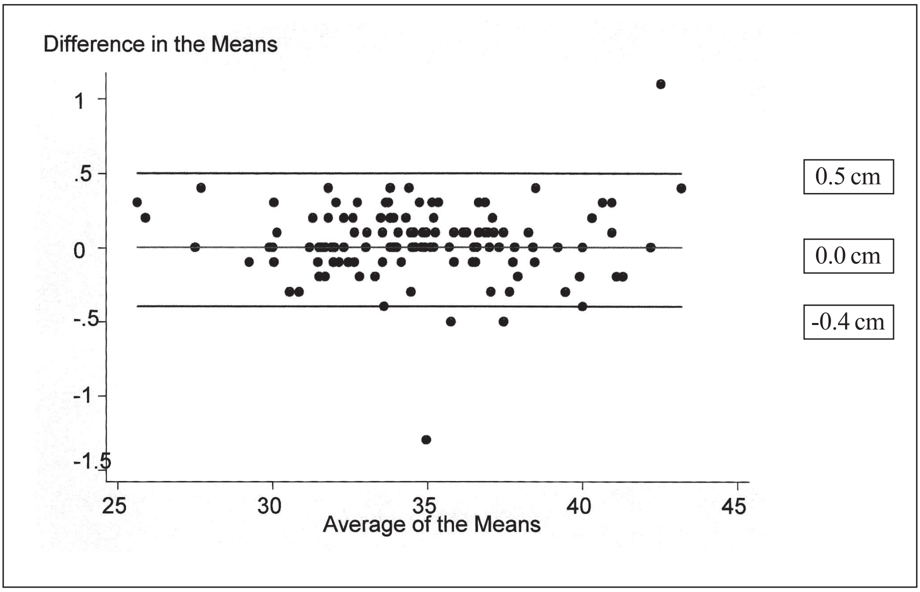


Figure 6. Scatter plot of the MHAS measurements in examiner #1 at time1 versus time 2

HHANES and 12-17 years in NHES III and also because of smaller sample sizes; n= 82 in HHANES and n= 301 in NHES III (Chumlea, 1990).

We also found that the %TEM was within the acceptable limits (≤ 1.5%), 0.88 for inter-examiner and even better for intra-examiner at 0.49. One study using 369 subjects found the %TEM to be 0.56 (Ross, 1994) which is consistent with our findings. In general, CV for circumference measurements would have much smaller values (Ulijaszek, 1999). However, we found that CV for CC measurement is almost twofold of the acceptable value (5%) at 9.5% for both inter- and intra-examiners. Ross *et al.* as in Ulijaszek & Kerr's (1999) paper also found a high CV of 7.32% for CC. The authors are unsure of the reason(s) behind this but suspect that it could be due to examiner fatigue.



Note: On average, the measurement taken from the 1st observer is consistent at two different time points, the 95% limits of agreement ranges from –0.4 cm to 0.5 cm.

Figure 7. Bland Altman plot on the calf circumference measurements within examiner 1

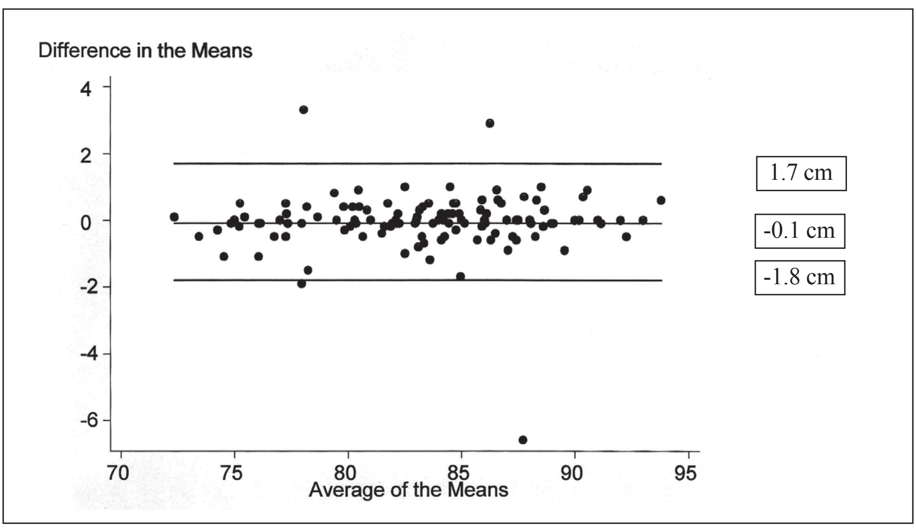


Figure 8. Bland Altman plot on the mid-half arm span measurements within examiner 1

Table 6. Inter- and intra-examiner TEM$^{\infty}$ for CC and MHAS measurements

Indices	*CC*		*MHAS*	
	Inter-examiner	*Intra examiner*	*Inter examiner*	*Intra examiner*
Absolute TEM* (m)	0.0030	0.0017	0.0069	0.0528
VAV □″	34.63	34.79	83.28	83.03
Relative TEM(%) **	0.88	0.49	0.83	6.35

∞TEM : Technical error of measurement; ∞ ∞ VAV : Variable average value

** Absolute TEM = $\sqrt{\dfrac{\sum di^2}{2n}}$

where

$\sum di^2$ = summation of deviations raised to the second power; n= number of volunteers measured, i= the number of deviations;

** **VAV** =∑ measurements, where ∑ measurements= summation of measurements

*** **Relative TEM** = $\dfrac{TEM}{VAV}\times 100$, where TEM= Technical error of measurement expressed in %

VAV= Variable average value

Table 7. Coefficient of variation of CC and MHAS using data from inter-examiner and intra-examiner

Variables	*N*	*Mean (SD)*		*Coefficient of variation (%) **	
		Inter examiner	*Intra examiner*	*Inter examiner*	*Intra examiner*
Calf circumference, cm	130	34.6(3.3)	34.8(3.3)	9.5	9.5
Mid half arm span, cm	130	83.3(4.6)	83.4(4.6)	5.5	5.5

* CV% = (SD/mean)100. *Source*: Zady MF (1999).

The ICC is an estimate of the proportion of the combined variance for the true biological value for any anthropometric measure and for the measurement errors associated with it. We found ICC values close to 1; 0.9980 for inter-examiner and 0.9990 for intra-examiner, indicating a low variability between repeated measures on the same subject, which is good. In the inter-examiner Bland & Altman plot for CC, the mean differences across all the range of readings was -0.3 cm which indicates some evidence of systematic bias between examiners. At 1SD, examiner 1 can under-estimate up to almost 1 cm compared to examiner 2 or vice versa; examiner 2 can over-estimate up to almost 1 cm. There is also some 'funneling effect' found (Figure 3) in inter-examiner CC measurement, as the larger the calf circumference, the more dispersed the measurements are from the mean.

We also examined reliability for MHAS in inter- and intra-examiner results. The errors between examiners and within examiner were the same with a value of 0.1 while the absolute TEM found was 0.69

(inter) and 5.28 (intra). This was compared with a study by Kaur & Singh (1994), as in Ulijaszek & Kerr's (1999) paper, found TEM values of 0.1 and 0.5 (inter) and 0.3 (intra) respectively. These values are not the same probably because of different age groups and sample sizes. It was found that the TEM value in this study was much higher than the acceptable range (Table 2) for intra-examiner. This could be due to lack of training in the technique of measuring MHAS by that particular examiner. Though obtaining accurate MHAS in older people can be troublesome, this problem can be controlled with good training (Ulijaszek, 1999). We also found that the %TEM was within the acceptable limits ($\leq$ 1.5%), 0.69 for inter-examiner.

In our study, we found that CV for MHAS measurement is almost at the acceptable value (5%) of 5.5% for both inter- and intra-examiners. We also found ICC values close to 1; 0.9890 for inter-examiner and 0.9910 for intra-examiner, indicating that there is low variability between repeated measurements on the same subject, which is good. In the inter- examiner Bland & Altman plot for CC, the mean differences across the range of readings was -0.2 cm which indicates some evidence of systematic bias between examiners. At 1SD, examiner 1 can under-estimate up to almost 2 cm compared to examiner 2 or vice versa.

Although the inter-examiner correlation coefficient showed a high degree of reliability (r= 0.9890), the Bland and Altman Plot (Figure 5), however, suggests that there is some evidence of random error between the examiners in their MHAS measurements, in fact by almost $\pm$ 2 cm, a bigger range than that for CC. This problem could be due to the difficulty in ascertaining the mid-point. Hence, thorough training before and during the conduct of data collection itself becomes critical. The smaller number of examiners seems to be the better option especially for MHAS.

Intra-examiner reliability was only assessed in examiner #1. Figure 6 and Figure 7 depict the strong correlation between the readings at time 1 and time 2 of examiner #1 for CC and MHAS respectively. Validity of the two methods cannot be tested because presently there are no reference/gold standard tests.

Study limitations

One of the reasons why differences are observed in this study could probably be explained by the fact that there were no target training values imposed upon the examiners prior to the conduct of the study. Training in itself is essential as it influences the degree of measurement error and interpretation especially if there is a high inter-examiner variation. Besides training, some targets need to be set *a priori* and failure to achieve that would disqualify a person from being an examiner. This study also did not have the benefit of a trained anthropometrist, first, to test reliability of measurements and second, to take charge of the training. The fact that only two examiners were employed for this study and who were, not actual NHMS III examiners, was also another limiting factor. However, this limitation was due to the fact that the field team for the actual NHMS III study had not been assembled at the time of this study.

CONCLUSION

Anthropometric measurement errors are unavoidable and should be minimised by giving close attention to every aspect of the data collection process including selection of examiners. The intra-examiner correlation coefficient for CC and MHAS were almost perfect at 0.9990 and 0.991 respectively. Similarly the inter-examiner correlation coefficient both for CC and MHAS were also almost perfect at 0.9960 and 0.9890 respectively.

This study found that, although both methods had a high degrees of reliability, at both the intra- and inter-examiner level, the reliability was higher for CC than for MHAS.

The CV was also reassuring for both methods as it was found to be consistently below 5%. The same trend was found in the technical error of measurement findings in which the intra-examiner result was more acceptable for CC compared to MHAS.

Calf circumference measurements seem to have higher intra-examiner reliability both as evidenced by the lower absolute mean differences, higher correlation coefficient and smaller range of limits of agreement.

The findings of this study suggest that calf circumference and mid-half arm span measurement in the elderly aged 60 years and above were reliable in this pair of examiners. However, to be used in a community survey, the reliability of the actual examiners has to be re-assessed.

RECOMMENDATIONS

For NMHS III, which uses CC measurement, we would like to suggest that these limitations be acknowledged. We also recommend that a liberal correction factor of +1 cm be used for the actual study on the circumference measurements.

We would like to stress upon the critical importance of (a) training both before and during the course of data collection in surveys, especially for MHAS measurements in the elderly, and (b) to minimise potential errors by limiting to a single/minimum number of examiners where possible, to reduce inter-examiner differences (Ulijaszek, 1999).

Since anthropometry is usually regarded as less troublesome compared to other nutritional studies, measurement is often delegated to lower qualified staff. Thus, to rectify this, examiner(s) would preferably be trained and later be assessed by qualified anthropometrists against some predetermined target training values. Even during the course of the data collection, it is recommended that to maintain quality of measurement, some on-site assessment be done on a random basis.

The working environment for data collection should also be arranged so that it is adequate for measurement purposes. Too cramped or poorly lit stations may cause additional error as a consequence of crowding and misreporting.

In addition, this study recommends similar studies be conducted for other anthropometry measurements (example height and weight in elderly persons). Besides, we also recommend that a study be done in the field while the actual study is ongoing, just like in HHANES and NHES II. It is important to carry out a pilot study before engaging in studies involving nutritional anthropometry.

ACKNOWLEDGMENTS

We would like to thank the Director of Institute of Public Health, Dr Nirmal Singh for his permission to deploy the Institute's staff for the conduct of the study, Director of Clinical Research Centre, Dr Lim Teck Onn for guidance, Dr Yusof bin Ahmad and the team from the Outpatient Department Kuala Lumpur Hospital, for assisting us in conducting our study at their site. Special appreciation goes to Ms Siti Nurasyura Mohd Adzman, Ms Siti Junara Mohd Halim, Ms Nurul Naquiyah Kamaruzzaman and Ms Noor Akma Hassim for their assistance in data collection and manuscript editing. We also thank all those whose names are not mentioned here for services rendered during data collection and data entry.

REFERENCES

Bassey EJ (1986). Demi-span as a measure of skeletal size. Ann Hum Biol 13 : 499-502. <*http://www.informaworld.com/smpp/title~content=t713723502~db=all~tab=issueslist~branches=13 - v13*>Accessed on 7 Oct 2006.

Bland MJ & Altman DG (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* i : 307-310.

Bonnefoy M, Jauffret M, Kostka T & Jusot JF (2002). Usefulness of calf circumference measurement in assessing the nutritional state of hospitalized elderly people. *Gerontology* 48: 162-169.

Chumlea WC, Guo S, Kuczmarski RJ, Johnson CL & Leahy CK (1990). Reliability for anthropometric measurements in the Hispanic Health and Nutrition Examination Survey (HHANES 1982-1984). *Am J Clin Nut* 51: 902S-907S

Bland, Martin (2006). How could I calculate a within-subject coefficient of variation? <*http://www-users.york.ac.uk/~mb55/meas/cv.htm*> Accessed on 7 October 2006.

Medical University of South Carolina (MUSC)(2006). <*http://www.musc.edu/dc/icrebm/index.html*>. Accessed on 20 Sept 2006.

National of Health and Morbidity Survey III (NHMS III) (2006). Technical Manual of Measurements. Institute of Public Health, Malaysia.

Pieterse S, Manandhar M & Suraiya I (1998). The nutritional status of older Rwandan refugees. *Public Health Nutr* 1 (4) : 259-164.

Talita AP, Glauber LO, Juliana SO & Fatima PO (2005). Technical error of measurement in anthropometry. *Rev Bras Med Esporte 11* : 86-90.

The New York Obesity Research Centre. Body Composition Unit. (2006). <*www.nyorc.org/bcu/labs /anthropometry.html*> Accessed on 16 Sept 2006.

Ulijaszek S & Kerr D (1999). Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr* 82: 165-177.

Walter SD, Eliasziw M & Donner A (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine* 17: 101-110.

WHO (1995). Physical status: The use and interpretation of anthropometry. Technical Report Series, Geneva, Switzerland.

Zady MF (1999). Z-stats 4: mean, standard deviation, and coefficient of variation. <*http://www.westgard.com/lesson34.htm#coefficient*> Accessed on 7 October 2006.