THE OFFICIAL JOURNAL OF THE BRITISH DIETETIC ASSOCIATION

Journal of Human Nutrition and Dietetics

CLINICAL NUTRITION

Reproducibility of measurements of mid-upper arm circumference in older persons

H. A. H. Wijnhoven,* M. R. de Boer,* M. J. van Maanen,* D. M. van Dongen,* S. F. Kraaij,* T. Smit* & M. Visser*†

*Department of Health Sciences, EMGO Institute for Health and Care Research, VU University, Amsterdam, The Netherlands
†Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands

Keywords
anthropometry, arm circumference, elderly, malnutrition, nutritional assessment, reproducibility.

Correspondence
H.A.H. Wijnhoven, Department of Health Sciences, Faculty of Earth and Life Sciences, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.
Tel.: +31 (0)20 5989951
Fax: +31 (0)20 5986940
E-mail: h.a.h.wijnhoven@vu.nl

How to cite this article
Wijnhoven H.A.H., de Boer M.R., van Maanen M.J., van Dongen D.M., Kraaij S.F., Smit T. & Visser M. (2013) Reproducibility of measurements of mid-upper arm circumference in older persons. *J Hum Nutr Diet*. **26**, 24–31
doi:10.1111/jhn.12010

Abstract

Background: Mid-upper arm circumference (MUAC) is used as an alternative measure for body mass index to determine thinness in older persons. However, there are limited data on the reproducibility of this measurement in an older population. The present study examined the reproducibility of MUAC measurements in older persons, as well as the influence of different body positions and clothing.
Methods: A cross-sectional reproducibility study was performed in a nursing home (n = 43; age 65–96 years) and swimming pool facilities (n = 107; age 65–88 years). A different pair of observers independently measured the MUAC of each participant in the upright position on two occasions within 1 week. In the nursing home, measurements were also performed for each participant in the laying position and with clothes covering the upper arm.
Results: Mean differences and the 95% limit of agreement for inter-observer reproducibility of MUAC were 0.0 cm (−2.6 to 2.5 cm) for the swimming pool facilities and 0.3 cm (−0.6 to 1.3 cm) for the nursing home. Intra-class correlation coefficients (ICCs) were 0.89 and 0.92, respectively. Mean differences between laying and upright positions were 0.1 cm (−2.0 to 2.2 cm) and 0.0 cm (−1.9 to 2.0 cm) for each observer, respectively (ICC 0.96–0.97). Mean differences between clothes versus bare upper arm were −2.7 cm (−6.2 to 0.7) and −2.4 (−5.6 to 0.9 cm) (ICC 0.75 and 0.78).
Conclusions: The reproducibility of the MUAC measurement in older persons is acceptable for group comparisons and, although borderline for the swimming pool facilities, remains acceptable for clinical purposes. The measurement can also be performed in the laying position but not with clothes covering the upper arm.

Introduction

Mid-upper arm circumference (MUAC) can be used as an alternative measure for body mass index (BMI) to determine thinness in older persons and is currently incorporated in two screening instruments that can be used to assess the (risk of) undernutrition in community-dwelling older persons (BAPEN, 2003; Wijnhoven *et al.*, 2012). Because of its simple assessment (i.e. it only requires a tape measure), it is especially suitable for measuring thinness in the home situation. Besides the feasibility argument, there is also evidence suggesting that a low MUAC is a more valid measure than a low BMI to define thinness in older persons. Unlike BMI, MUAC is not influenced by spinal deformities and is influenced to a lesser extent by fluid changes or limb amputations that often occur in older persons. A previous study among older persons living in sheltered accommodations found that a low MUAC was a more sensitive and specific measure than a low BMI for identifying those at risk of undernutrition based on a dietician assessment (Harris *et al.*, 2008). In line with this, another study among community-dwelling older persons showed that a low MUAC was more strongly associated with 15-year mortality than a low BMI (Wijnhoven *et al.*, 2010).

Besides feasibility and validity, good reproducibility is also an important prerequisite of MUAC measurement. There are, however, limited data available on the inter-observer reproducibility of the MUAC measurement in an older population. To the best of our knowledge, there is only one previous reproducibility study, which was conducted among 50 hospitalised patients [mean (SD) age 67 (15) years, range 16–91 years] (Burden *et al.*, 2005). The present study found a good intra- and inter-observer reproducibility for MUAC based on intra-class correlation coefficients of 0.99 and 0.98, respectively, and a mean difference within and between observers of ±0.15 cm with 95% limits of agreement from ±−2 to 2 cm. Furthermore, the influence of body position (i.e. laying position versus the standard upright position) or clothing (i.e. clothes covering the upper arm versus the standard measurement with bare upper arm) on the reproducibility of the MUAC measurement is not unknown.

The present study therefore aimed to (i) examine the intra- and inter-observer reproducibility of MUAC measurements in community-dwelling and institutionalised older persons and (ii) examine the influence of different body positions and clothing on MUAC measurements in institutionalised older persons.

Materials and methods

Design and setting

This cross-sectional reproducibility study consists of two sub-studies that were performed 1 year apart in two different settings by two different observer pairs. We chose to perform our reproducibility study in two different settings to enhance the generalisability of our findings. Sub-study 1 was performed in 2009 in swimming pool facilities in the Amsterdam area (The Netherlands) that were available to senior citizens. Sub-study 2 was performed in 2010 in a nursing home facility in Amsterdam, the Netherlands. The study was approved by the Ethics Review Board of the VU University Medical Center.

Study populations

For sub-study 1, participants were recruited through four public swimming pools in Amsterdam and the surrounding area, during hours that were especially reserved for swimming for senior citizens (>55 years). Information about the study was spread through posters and flyers that were distributed in the swimming pool facility 1 week before the study and on the day of the first measurement round. In total, 149 persons were recruited for the first measurement round. After excluding those aged <65 years, 107 older persons (aged $\geq$ 65 years) were included in the present study. For the second measurement round that took place 1 week later, 59 out of these 107 persons were present and participated in the second measurement round. We did not assess the influence of different body positions and presence of clothing in this group for practical reasons; clothes were put in lockers and beds were not readily available.

For sub-study 2, participants were recruited through a nursing home (somatic and rehabilitation ward) in Amsterdam, The Netherlands. Of the 77 nursing home residents who were initially approached, 49 were willing to participate. Reasons for nonparticipation were: not being interested in participating (*n* = 25), illness (*n* = 1) or not being able to stand or sit (*n* = 2). Another four respondents were excluded because they were aged <65 years and two respondents were excluded because they were unable to stand or sit during the measurements. In total, 43 older persons were included in the sub-study.

Measurements

The observers were trained by a single trainer to perform the MUAC measurement (and height and weight in sub-study 1) using a standardised protocol. For sub-study 1, two observers independently assessed MUAC, body height and weight during two measurement rounds 1 week apart. The results of both measurement rounds were noted on a separate sheet to blind the observers to the result of their previous measurement. Participants all wore swimming clothes when they were assessed. Height was measured to the nearest 0.001 m using a stadiometer. Weight was measured to the nearest 0.5 kg using a calibrated bathroom scale (Seca, model 100; Lameris, Utrecht, The Netherlands). BMI was calculated as body weight (kg) divided by height (m) squared. MUAC was measured to the nearest 0.005 m using a nonstretch tape measure at the left (bare) arm with the participant in standing position (Lohman *et al.*, 1988). Circumference was measured midway between the lateral projection of the acromion process of the scapula and the inferior margin of the olecranon process of the ulna. This point was determined (but never marked) with the elbow bend; the circumference itself was determined with the arm hanging loose. Day of birth and sex were recorded during the first measurement round.

For sub-study 2, two observers (other observers than in sub-study 1) independently assessed MUAC on three occasions within a period of 8 days. The observers were trained to perform the measurements using the standardised protocol for MUAC measurement, as described above. Besides the upright measurement, one measurement was performed with the participant in the laying position and another measurement was performed with the participant in the upright position but with clothes (whatever they were wearing) covering the upper arm. If participants were unable to stand (n = 41; 35 in a wheelchair), the measurement was performed in the sitting position. All three measurements (upright; laying; with clothes) were performed in random order, at least 1 day apart, by each observer independently. The results of each measurement were noted on a separate sheet to blind the observers to the result of their previous measurement. Day of birth and sex were recorded during the first measurement round.

Statistical analysis

Inter- and intra-observer reproducibility were examined by two commonly used measures of reproducibility (De Vet, 1998; de Winter *et al.*, 2004): (i) the method of Bland & Altman (1986) and (ii) intra-class correlation coefficients (ICC). Bland & Altman scatterplots were created by calculating the difference in MUAC between (or within) two observers (for each participant) and plotting this against the mean value of the two measurements (for each participant). Subsequently, the mean (SD) difference between (or within) observers was calculated (for the whole study population). The 95% limits of agreement were calculated as the mean difference ± 1.96 SD of this difference. These limits represent the range in which 95% of the differences between (or within) two observers are expected to fall. Because the choice of extracting observer A from observer B or vice versa is arbitrary, the largest limit of agreement (either upper or lower) is the most relevant. In addition, the frequency of agreement of the observers (within of between) within 1 and 2 cm, respectively, was calculated. Because no clear criteria for the acceptable degree of inter-observer agreement of MUAC are available, we decided that differences <2 cm would be acceptable for clinical purposes.

The ICC was derived from a random-effects two-way analysis of variance. The variation of measurements is partitioned into potential sources of variation: differences between (or within) observers, differences between participants and random error. The ICC is defined as the ratio of variance between participants over the total variance. The ICC can range from 0 to 1, with a higher value indicating a smaller between (or within) observer variance and thus a higher degree of reproducibility. ICCs are strongly influenced by the heterogeneity of the study population (i.e. in a population with a wider range of values of MUAC measurement, the ICC will be higher). An ICC of at least 0.70 is considered satisfactory for group comparisons, and a value of 0.90–0.95 is considered satisfactory for individual comparisons (Scientific Advisory Committee of the Medical Outcome Trust, 2002).

Results

Baseline characteristics of the study participants are presented in Table 1. More women participated than men, especially for the swimming pool facility (sub-study 1). On average, participants of the nursing home facility (sub-study 2) were older and had a lower mid-upper arm circumference.

Table 2 summarises the results of the inter-observer reproducibility of the MUAC measurement. In both sub-studies, there was no systematic difference between the two observers (mean difference 0.0 and 0.3 cm) when MUAC was measured in the upright position. The corresponding 95% limits of agreement varied between −2.6 and 2.5 cm for sub-study 1 and between −0.6 and 1.3 cm for sub-study 2. The percentage of agreement within 2 cm was very high (96% and 100%, respectively), as was the percentage of agreement within 1 cm (79% and 98%, respectively). ICCs were 0.89 and 0.99, respectively.

Tables 3–5 show the results of the intra-observer reproducibility of the MUAC measurement for sub-study 1 and the influence of different body positions and clothing (intra-observer) for sub-study 2. MUAC measured in the upright position did not systematically differ from MUAC when measured in the laying position (mean difference <0.2 cm for both observers). The largest limit of agreement was 2.2 cm. When MUAC was measured with clothes cov-

Table 1 Baseline characteristics of the study participants

	Sub-study 1: swimming pool (n = 107)		Sub-study 2: nursing home (n = 43)	
	Men (n = 21)	Women (n = 86)	Men (n = 16)	Women (n = 27)
Age (years), mean (SD)	74.8 (4.6)	73.7 (5.8)	79.3 (7.4)	84.7 (6.9)
MUAC* (cm), mean (SD)	30.8 (2.8)	30.8 (2.7)	29.1 (3.1)	27.4 (4.2)
BMI* (kg/m^2), mean (SD)	26.0 (3.5)	26.9 (3.4)	Not assessed	

*Based on the mean value of the two observers during the first measurement round.

BMI, body mass index; MUAC, mid-upper arm circumference.

Table 2 Inter-observer reproducibility: mean SD of the mid-upper arm circumference measurement, followed by the mean difference between observers, 95% limits of agreement, frequency of agreement within 1 and 2 cm, and intra-class correlation coefficient

	Observer A, mean (SD)	Observer B, mean (SD)	Mean (SD) difference for observers A and B	95% limits of agreement	Agreement ≤ 1 cm (%)	Agreement ≤ 2 cm (%)	ICC
Sub-study 1: swimming pool							
MUAC (cm) upright position (*n* = 107)	30.8 (2.8)	30.8 (2.8)	−0.0 (1.3)	−2.6 to 2.5	79	96	0.89
Sub-study 2: nursing home							
MUAC (cm) upright position (*n* = 43)	28.2 (3.9)	27.8 (3.9)	0.3 (0.5)	−0.6 to 1.3	98	100	0.99
MUAC (cm) laying position (*n* = 34)*	28.3 (4.1)	28.0 (4.0)	0.3 (0.7)	−1.1 to 1.6	88	100	0.98
MUAC (cm) clothes covering upper arm (*n* = 42)[†]	31.0 (4.4)	30.2 (4.2)	0.8 (1.2)	−1.5 to 3.1	79	86	0.95

*Nine participants were unable to lay down for the measurement.
[†]One extreme value was excluded (difference bare arm and with clothes of 12.5 cm).
ICC, intra-class correlation coefficient; MUAC, mid-upper arm circumference.

Table 3 Intra-observer reproducibility: mean and SD of the mid-upper arm circumference measurement, followed by the mean difference between measurements, 95% limits of agreement, frequency of agreement within 1 and 2 cm, and intra-class correlation coefficient

	Measurement 1, mean (SD)	Measurement 2, mean (SD)	Mean (SD) difference for means 1–2	95% limits of agreement	Agreement ≤ 1 cm (%)	Agreement ≤ 2 cm (%)	ICC
Sub-study 1, swimming pool: intra-observer reproducibility							
Method used to measure MUAC	*Upright position*	*Upright position*					
MUAC (cm), observer A (*n* = 59)	30.6 (2.9)	30.7 (2.8)	−0.1 (0.5)	−1.0–0.9	100	100	0.99
MUAC (cm), observer B (*n* = 59)	30.7 (2.8)	30.7 (2.7)	0.0 (0.6)	−1.2–1.1	98	100	0.98

ICC, intra-class correlation coefficient; MUAC, mid-upper arm circumference.

Table 4 The influence of different body positions: mean and SD of the mid-upper arm circumference measurement, followed by the mean difference between measurements, 95% limits of agreement, frequency of agreement within 1 and 2 cm, and intra-class correlation coefficient

	Measurement 1, mean (SD)	Measurement 2, mean (SD)	Mean (SD) difference for means 1–2	95% limits of agreement	Agreement ≤ 1 cm (%)	Agreement ≤ 2 cm (%)	ICC
Sub-study 2, nursing home: influence of body position							
Method used to measure MUAC	*Upright position*	*Laying position*					
MUAC (cm), observer A (*n* = 34)*	28.4 (3.7)	28.3 (4.1)	0.1 (1.1)	−2.0–2.2	82	94	0.96
MUAC (cm), observer B (*n* = 34)*	27.9 (3.8)	28.0 (4.0)	−0.0 (1.0)	−1.9–2.0	79	97	0.97

*Nine participants were unable to lay down for the measurement.
ICC, intra-class correlation coefficient; MUAC, mid-upper arm circumference.

Table 5 The influence of clothing: mean and SD of the mid-upper arm circumference measurement, followed by the mean difference between measurements, 95% limits of agreement, frequency of agreement within 1 and 2 cm, and intra-class correlation coefficient

	Measurement 1, mean (SD)	Measurement 2, mean (SD)	Mean (SD) difference for means 1–2	95% limits of agreement	Agreement ≤ 1 cm (%)	Agreement ≤ 2 cm (%)	ICC
Sub-study 2, nursing home: influence of clothing (bare arm versus arm clothes covering upper arm)							
Method used to measure MUAC	*Upright bare arm*	*Upright with clothes*					
MUAC (cm), observer A (n = 42)*	28.3 (3.8)	31.0 (4.4)	−2.7 (1.8)	−6.2–0.7	21	50	0.75
MUAC (cm), observer B (n = 43)	27.8 (3.9)	30.2 (4.1)	−2.4 (1.6)	−5.6–0.9	30	56	0.78

*One extreme value was excluded (difference bare arm and with clothes of 12.5 cm).
ICC, intra-class correlation coefficient; MUAC, mid-upper arm circumference.

ering the upper arm, the outcome was systematically higher (2.7 and 2.4 cm, respectively, for observers A and B) than without clothing. The 95% limits of agreement varied widely between −0.7 and 6.2 cm and −0.9 and 5.6 cm, respectively. The other measures of reproducibility showed similar results (i.e. poorest agreement for the measurement with clothes covering the upper arm).

Bland and Altman plots of the results are shown in Fig. 1 and Fig. 2. The plots show that the error of measurement did not depend on the magnitude of the MUAC.

Discussion

The present study shows that the reproducibility of the MUAC measurement in older persons when performed by different observers is satisfactory for group comparisons and acceptable for clinical purposes in a nursing home setting, as well as borderline acceptable in a swimming pool setting. No systematic differences were observed between observers, and the largest limit of agreement was 2.6 cm in the swimming pool setting (ICC = 0.89) and 1.3 cm in the nursing home setting (ICC = 0.99). The measurement can also be performed in the laying position (no systematic difference, largest limit of agreement 2.2 cm, ICC = 0.96) but, as expected, not with clothes covering the upper arm because this causes a systematic overestimation of more than 2 cm and a large random variation in assessment.

Most previous studies that have examined the reproducibility of the MUAC were performed in adults and not in older persons. They found good inter-observer reproducibility statistics for MUAC, with coefficients of variations ranging between 1.0% and 4.7% (Heymsfield *et al.*, 1982; Harries *et al.*, 1985; Harris *et al.*, 2008). The results of the present study are supported by one previous study that was performed among 50 hospitalised patients with a mean (SD) age of 66.7 (SD) 15.2 years (range 16–91 years) (Burden *et al.*, 2005). In the present study, the mean difference between and within observers was also negligible (0.16 and 0.14 cm, respectively) and the 95% limits of agreement varied from −1.7 to 2.0 cm (between observers) and from −1.6 to 1.8 cm (within observers). The absence of systematic differences between (or within) observers indicates that the total variation in the MUAC measurements is not caused by a systematic difference in technique between observers but merely caused by random variation. This implies that, with proper training, this random variation may be reduced further.

The random variation between observers (95% limits of agreement) was somewhat larger in the swimming pool setting than in the nursing home setting and slightly exceeded the limit of 2 cm that was set as acceptable for clinical purposes. Although this limit of 2 cm is a matter of clinical judgement, the conclusions are consistent with the conclusions based on the ICC, which was 0.89 (also slightly lower than 0.90 set *a priori* as acceptable for individual comparisons). One explanation is that the participant characteristics influence the reproducibility of the measurements. Participants in the swimming pool were younger, more often female and had a higher MUAC. However, the Bland and Altman plots showed that the error of measurement does not depend on the magnitude of the MUAC and stratifying the analysis by age and sex (and setting) did not reveal differences in reproducibility between these groups (data not shown). Most likely, the differences in MUAC reproducibility between the two settings are explained by differences between

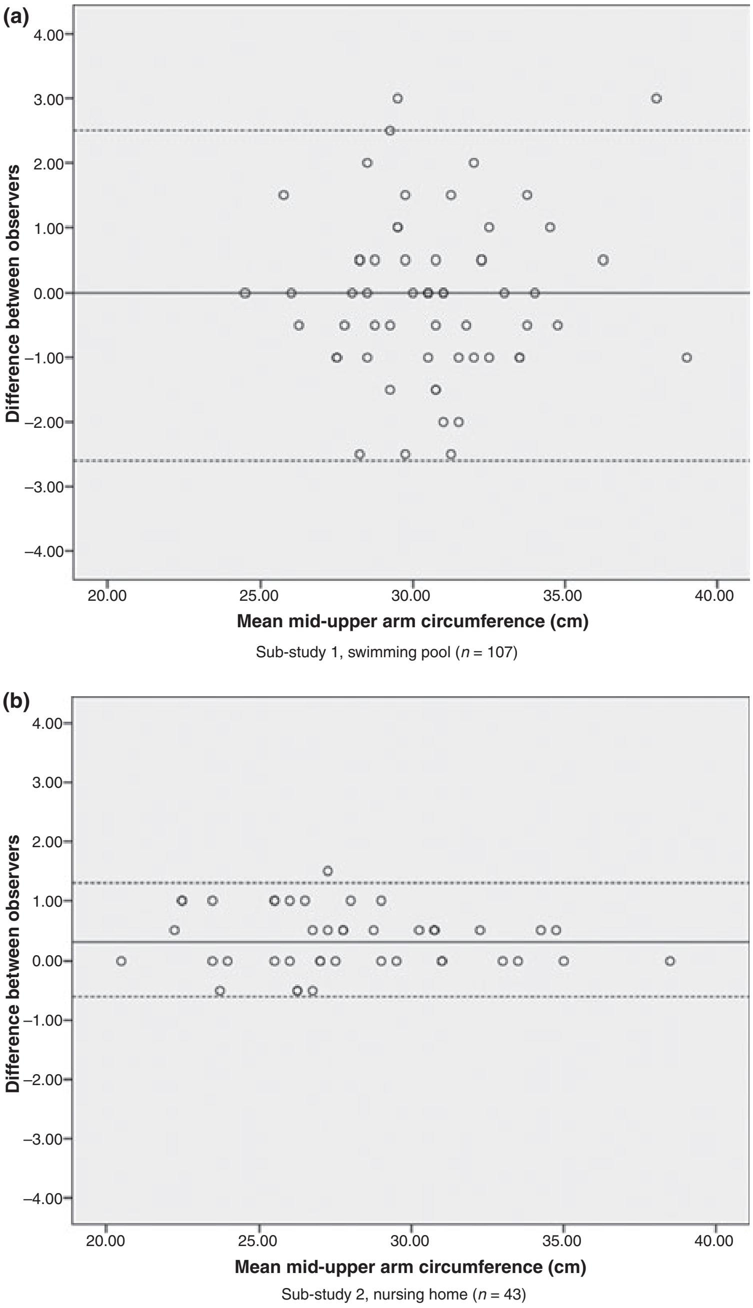

Figure 1 Differences in mid-upper arm circumference measurements between observers plotted against the mean value of both observers for each participant, for sub-study 1 (a) and sub-study 2 (b)

observers and thus are an indication of how precisely the individual observers performed the measurements, and perhaps are also a result of the different circumstances under which the measurements were performed (public swimming pool versus nursing home).

A limitation of the present study is that we did not examine the validity of the MUAC (e.g. for example by relating it to a gold standard). It might thus be that both assessors in the present study systematically over- or underestimated the MUAC, and that this was not expressed in our reliability indices. However, we do not expect that the MUAC measured with a measure tape using a standard procedure would systematically differ from a gold standard measure, such as circumference measures determined by computerised tomography. Although we are unaware of studies conducted in older persons, in adults, no systematic difference was observed between MUAC (or $MUAC^2/2\pi$) assessed by a measure tape or by computerised tomography (Heymsfield *et al.*, 1979, 1982). An important finding of the present study is

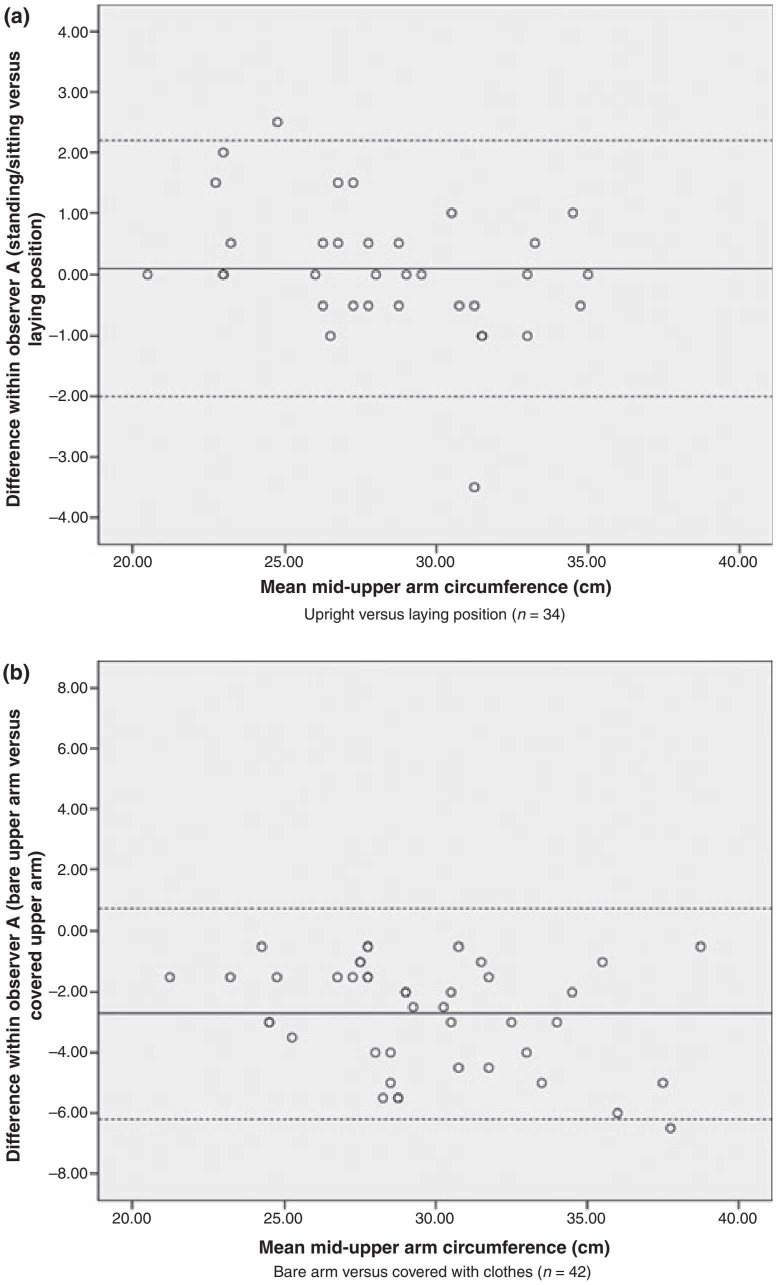

Figure 2 Differences in mid-upper arm circumference measurements due to different body position (a) and clothing (b) for one observer (A), plotted against the mean value for each participant; sub-study 2 nursing home

that the MUAC measurement can reliably be performed in thye laying position as well. This is a major advantage in comparison with BMI, which is very difficult to assess can only be assessed in the laying position, enhancing its applicability in nursing home and rehabilitation centres.

In conclusion, the reproducibility of the MUAC measurement in older persons is acceptable for group comparisons and, although borderline for the swimming pool facility, is acceptable for clinical purposes as well. The measurement can also be performed in the laying position but not with clothes covering the upper arm.

Conflict of interests, source of funding and authorship

The authors declare that they have no conflicts of interest.
No funding is declared.

HAHW was responsible for the conception and design of the study, data analysis, interpretation of the data, and writing of the paper; MRdB was responsible for advice on data analysis, interpretation of the data, and critically reviewing the paper; MV was responsible for the conception and design of the study, interpretation of the data, and critically reviewing the paper; MJvM, DMvD, SFK and TS contributed to data analysis and drafting of the paper, critical reviewing of the paper, and the collection of data. All authors critically reviewed the manuscript and approved the final version submitted for publication.

References

BAPEN (2003). *Malnutrition Universal Screening Tool (MUST)*. Redditch: British Association of Enteral and Parenteral Nutrition.

Bland, J.M. & Altman, D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.

Burden, S.T., Stoppard, E., Shaffer, J., Makin, A. & Todd, C. (2005) Can we use mid upper arm anthropometry to detect malnutrition in medical inpatients? A validation study. *J. Hum. Nutr. Diet.* **18**, 287–294.

De Vet, H.C. (1998) Observer reliability and agreement. In *Encyclopedia of Biostatistics*, Vol. 4. eds P. Armitage & T. Colton, pp. 3123–3128. Boston, MA: John Wiley & Sons Ltd.

Harries, A.D., Jones, L.A., Heatley, R.H., Newcombe, R.G. & Rhodes, J. (1985) Precision of anthropometric measurements: the value of mid-arm circumference. *Clin. Nutr.* **4**, 77–80.

Harris, D.G., Davies, C., Ward, H. & Haboubi, N.Y. (2008) An observational study of screening for malnutrition in elderly people living in sheltered accommodation. *J. Hum. Nutr. Diet.* **21**, 3–9; quiz 10–12.

Heymsfield, S.B., Olafson, R.P., Kutner, M.H. & Nixon, D.W. (1979) A radiographic method of quantifying protein-calorie undernutrition. *Am. J. Clin. Nutr.* **32**, 693–702.

Heymsfield, S.B., McManus, C., Smith, J., Stevens, V. & Nixon, D.W. (1982) Anthropometric measurement of muscle mass: revised equations for calculating bone-free arm muscle area. *Am. J. Clin. Nutr.* **36**, 680–690.

Lohman, T.G., Roche, A.F. & Martorall, R. (1988) *Anthropometric Standardization Reference Manual*. Champaign, IL: Human Kinetics Books.

Scientific Advisory Committee of the Medical Outcome Trust. (2002) Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual. Life Res.* **11**, 193–205.

Wijnhoven, H.A., van Bokhorst-de van der Schueren, M.A., Heymans, M.W., de Vet, H.C., Kruizenga, H.M., Twisk, J.W. & Visser, M. (2010) Low mid-upper arm circumference, calf circumference, and body mass index and mortality in older persons. *J. Gerontol. A Biol. Sci. Med. Sci.* **65**, 1107–1114.

Wijnhoven, H.A., Schilp, J., van Bokhorst-de van der Schueren, M.A., de Vet, H.C., Kruizenga, H.M., Deeg, D.J., Ferrucci, L. & Visser, M. (2012) Development and validation of criteria for determining undernutrition in community-dwelling older men and women: The Short Nutritional Assessment Questionnaire 65+. *Clin. Nutr.* **31**, 351–358.

de Winter, A.F., Heemskerk, M.A., Terwee, C.B., Jans, M.P., Deville, W., van Schaardenburg, D.J., Scholten, R.J. & Bouter, L.M. (2004) Inter-observer reproducibility of measurements of range of motion in patients with shoulder pain using a digital inclinometer. *BMC Musculoskelet Disord.* **5**, 18.