

Reliability and accuracy of anthropometry performed by community health workers among infants under 6 months in rural Kenya

Martha K. Mwangome^{1,2}, Greg Fegan^{1,3}, Ronald Mbunya⁴, Andrew M. Prentice^{2,5} and James A. Berkley^{1,3}

1 KEMRI/Wellcome Trust Research Programme, Kilifi, Kenya

2 MRC International Nutrition Group, London School of Hygiene and Tropical Medicine, London, UK

3 Centre for Clinical Vaccinology & Tropical Medicine, University of Oxford, Oxford, UK

4 Ministry of Health– Kilifi District Hospital, Kilifi, Kenya

5 Medical Research Council Laboratories, MRC Keneba, Keneba, The Gambia

Abstract

OBJECTIVE To assess the inter-observer variability and accuracy of Mid Upper Arm Circumference (MUAC) and weight-for-length Z score (WFLz) among infants aged <6 months performed by community health workers (CHWs) in Kilifi District, Kenya.

METHODS A cross-sectional repeatability study estimated inter-observer variation and accuracy of measurements initially undertaken by an expert anthropometrist, nurses and public health technicians. Then, after training, 18 CHWs (three at each of six sites) repeatedly measured MUAC, weight and length of infants aged <6 months. Intra-class correlations (ICCs) and the Pitman's statistic were calculated.

RESULTS Among CHWs, ICCs pooled across the six sites (924 infants) were 0.96 (95% CI 0.95–0.96) for MUAC and 0.71 (95% CI 0.68–0.74) for WFLz. MUAC measures by CHWs differed little from their trainers: the mean difference in MUAC was 0.65 mm (95% CI 0.023–1.07), with no significant difference in variance ($P = 0.075$).

CONCLUSION Mid Upper Arm Circumference is more reliably measured by CHWs than WFLz among infants aged <6 months. Further work is needed to define cut-off values based on MUAC's ability to predict mortality among younger infants.

keywords community health workers, Mid Upper Arm Circumference, weight-for-length Z score, infants, malnutrition, Kenya

Introduction

Malnutrition underlies 35% of childhood deaths and 10–40% of hospital admissions in children under 5 years in sub-Saharan Africa (Black *et al.* 2008). Commonly, infants under 6 months are excluded from nutritional surveys resulting in miscalculation of the overall prevalence of under nutrition among the under fives (Lopriore *et al.* 2007). Furthermore, a change to new anthropometric references (WHO 2006b) has revealed a far greater burden of malnutrition among infants under 6 months than previously recognised (de Onis *et al.* 2006). It is currently estimated that worldwide 8.5 million infants under 6 months are wasted (Kerac *et al.* 2011). In poor communities, low rates of exclusive breastfeeding and the introduction of mixed feeding before the age of 3 months (Nwankwo & Brieger 2002; Fjeld *et al.* 2008) expose

infants <6 months to risks of microbial contamination and malnutrition.

In Kenya, 9.7% infants below 6 months are wasted [weight-for-length Z score (WFLz < -2)] and 11% are stunted [length-for-age Z score (LFAz < -2)] (Kenya National Bureau of statistics (KNBS) & ICF Macro (2009). The Government of Kenya has proposed a strategy in which Community Health Workers (CHWs) are trained to deliver community health services, including basic primary health care, growth monitoring (GM) and referral of critically ill patients to hospital (Ministry of Health 2006). CHWs will be expected to undertake door-to-door anthropometric screening of children and provide basic nutrition education and counselling.

At rural health facilities, weight is commonly measured in infancy. However, weight-for-age (WFA) alone does not differentiate wasting from stunting and is typically

accessible only to those attending mother and child health clinics (MCH). Weight-for-length (WFL) is recommended for the diagnosis of acute malnutrition in this age group, but it rarely routinely assessed because length boards are not usually available and length measurement is potentially unreliable (Voss *et al.* 1990).

Among children aged 6–59 months, the Mid Upper Arm Circumference (MUAC) may be used to diagnosed severe acute malnutrition (SAM). MUAC is a better predictor of mortality than WFA (Myatt *et al.* 2006) and, within this age range, is age-independent (Kikafunda *et al.* 1998). In rural communities, MUAC could be a valuable tool for use by CHWs for early detection of acute malnutrition in infants. However, reliability of MUAC measurement in early infancy is unknown, and cut-off values to determine intervention thresholds have not been defined. To address the first of these questions, we aimed to determine the inter-observer variability and accuracy of MUAC and WFLz measurements taken by CHWs among infants under 6 months in rural Kenya.

Methods

Study site

The study was conducted from February 2008 through August 2009 in Kilifi District, a rural district on the Kenyan Coast. Kilifi is the second poorest district in Kenya with an estimated 67% of the people living in poverty (World Bank 2008). Approximately 500 cases of SAM are admitted to Kilifi District hospital (KDH) every year.

Study participants

We recruited three cadres of participant: (i) an expert in anthropometry with >30 years of experience of anthropometry training and conducting nutritional assessments in Kenya; (ii) health professionals (HPs; nurses and public health officers) in-charge of Mother and Child Health (MCH) clinics and (iii) CHWs based at health centres throughout the district.

Study design

We employed a cross-sectional repeatability design. We use the term ‘reliability’ in a statistical sense to mean the interobserver variability.

Measuring equipment

For applicability within the health system, we used measuring equipment normally in use in government

facilities in Kenya, with an exception of the infantometer, which is usually not available. Weight was measured to the nearest 100 g using a ‘hanging’ scale [Salter model number 235-65 (25 kg × 100 g); Salter Brecknell, UK] costing 103 USD. The machine was quality controlled every morning using standard weighing stones certified by the Kenya Bureau of Standards. Length was taken using a professional infantometer [SECA model number 416-1821009 (33–100 cm), SECA, Germany] calibrated to the nearest 1 mm and costing 443 USD. MUAC was measured on the left arm of the child using TALC insertion tape marked to the nearest 2 mm (TALC, ST Albans, UK) costing 0.25 GBP. All measures followed procedures indicated in the United Nations (1986) guidelines.

Sample size

We calculated sample sizes separately for infants older and younger than 90 days according to the method described by Walter *et al.* (1998; Bonett 2002). We defined ‘complete unreliability’ as an intra-class correlation (ICC) of <0.4 and estimated the number of infants required for 90% power to distinguish ICC values of 0.6, which we defined as minimum reliability, from 0.4. This gave a required sample size of 71 infants in each age group for three observers. To allow for possible dropout from the study, we aimed to recruit at least 75 infants for each age group. Sample size for accuracy was 15 infants for every 75 infants recruited.

Study procedure

The study was then undertaken in three stages. First, to establish intrinsic reliability, the expert anthropometrist measured weight, length and MUAC among infants visiting the MCH at KDH. Measurements were repeated after each cohort of 10 children. First and second set of measurements were recorded on separate forms.

At the second stage, a training manual was produced for training HPs and CHWs following guidelines from the United Nations 1986 on anthropometry. The expert and the first author (PhD student with >3 years experience in anthropometry) trained six HPs on anthropometry, safety procedures when handling infants and quality control of the measuring equipment. This 2-day training program included a practical assessment. After the training, the HPs were divided into two groups of three each and repeatedly measured weight, length and MUAC of 150 infants (75 infants above and 75 below 90 days old). Each child was measured once by each of the three HPs. For every 5th child, the expert took measurements to determine accuracy. Further training was given to address issues arising in the second stage to establish HPs as trainers for the CHWs.

At the third stage, 18 CHWs were recruited, three from each of six sites: a district hospital, a peri-urban health centre, two rural health centres and two rural dispensaries. HPs conducted a 1-day practical training on anthropometry, safety procedures when handling infants and equipment quality control. This aimed to replicate the type of training that could be provided operationally. Then, each group of three CHWs independently measured MUAC, weight and length among 150 infants (75 under and 75 over 90 days old) at their health facilities. A single MUAC, weight and length measurement were taken once by each CHW on each infant. Measures were blinded from each other. At each facility, one HP took measures on every 5th infant to estimate accuracy.

Statistical analysis

Data were double-entered in a database and transferred to STATA 11 (Stata Corp., TX, USA) for analysis. Data were excluded if incomplete (missing variables). Z scores including weight-for-length (WFLz), weight-for-age (WFAz) and MUAC-for-age (MAUCz) were computed using the WHO growth reference standards STATA macro (WHO 2010). Because of limitations of the growth reference standards, MAUCz could only be computed for infants older than 90 days. The anthropometric distribution of infants measured by CHWs was estimated using the mean measure of the three observers.

Intra- or inter-observer reliability was estimated using the ICC coefficient (ICC). Pooled ICCs by age and by site were calculated by meta-analysis using fixed effect models because there was no anticipated reason for heterogeneity between sites.

In stages II and III of the study, accuracy was estimated using Bland–Altman plots and the Pitman's test, based on calculating the correlation between the difference and the mean of paired variables. At both stages, individual measurements by each of the three observers were paired to the one measurement by the trainer on the same child and were used for the accuracy test. Mean differences were calculated to indicate systematic bias among observers. The Pitman's test was used to test of the null hypothesis of no difference in variance between the pairs (Kirkwood & Sterne 2003).

Ethical considerations

Informed consent was obtained from the expert, HPs and the CHWs for their participation and verbal consent was sought from the caregivers of the infants. Only infants of caregivers who consented participated in this study. The

study was approved by the Kenya National Ethical Review Committee (SCC number 1334).

Results

Stage I: Expert

We recruited 147 infants for MUAC and length measurements, 71 were aged <90 days, 32 (45%) male and 76 infants aged 90 days or more, 37 (49%) male. Because of a delay in equipment availability, weight measurement was conducted in a different group of 164 infants, of whom 50% were <90 days and 46% male; and 50% were ≥90 days and 54% male. WFLz could therefore not be calculated at stage 1.

The median MUAC was 130 mm, the median weight 5.8 kg and median length 58.0 cm. The median and interquartile range for LFAz was -1.4 (-2.1 to -0.6); 41 (32%) infants were stunted (LFAz < -2). The median WFAz was -0.37 (-1.0 to 0.5); 14 (9%) of the infants were underweight (WFAz < -2).

The ICCs pooled by age for all measurements undertaken by the expert were at least 0.92 (Figure 1). The ICC for MUAC was 0.97 (95% CI 0.97–0.98). There was no significant difference in ICCs for infants younger or older than 90 days.

Stage II: HPs

We recruited 155 infants but one was dropped owing to an obvious error in recording age. Analysis was carried out on 154 infants, of whom 77 were <90 days and 58% male; and 77 were ≥90 days and 42% male.

Among infants <90 days, the median MUAC was 126 mm, weight was 5.2 kg and length was 55.3 cm. Ten (13%) infants were stunted (LFAz < -2) and three (4%) underweight (WFAz < -2). Among infants >90 days, the median MUAC was 141 mm, weight was 6.9 kg and length was 62.6 cm. Stunting and underweight levels did not differ from those of the younger infants.

The pooled ICC for MUAC was 0.88 (95% CI 0.83–0.92) and for WFLz was 0.60 (95% CI 0.52–0.68) (Figure 2).

A total of 87 accuracy measurements for MUAC, weight and length were recorded. The mean differences in HPs measures compared to experts for MUAC, weight, length and WFLz were +3.5 mm (95% CI 2.5–4.4 mm), 0.004 kg (95% CI -0.04 to $+0.05$ kg), +0.81 cm (95% CI 0.59–1.0 cm) and -0.41 z scores (95% CI -0.57 to -0.24 z scores), respectively. We found no evidence that the variance of the paired measures by HPs differed from that of the expert (Pitman's statistic *P*-value was >0.05 for all measures).

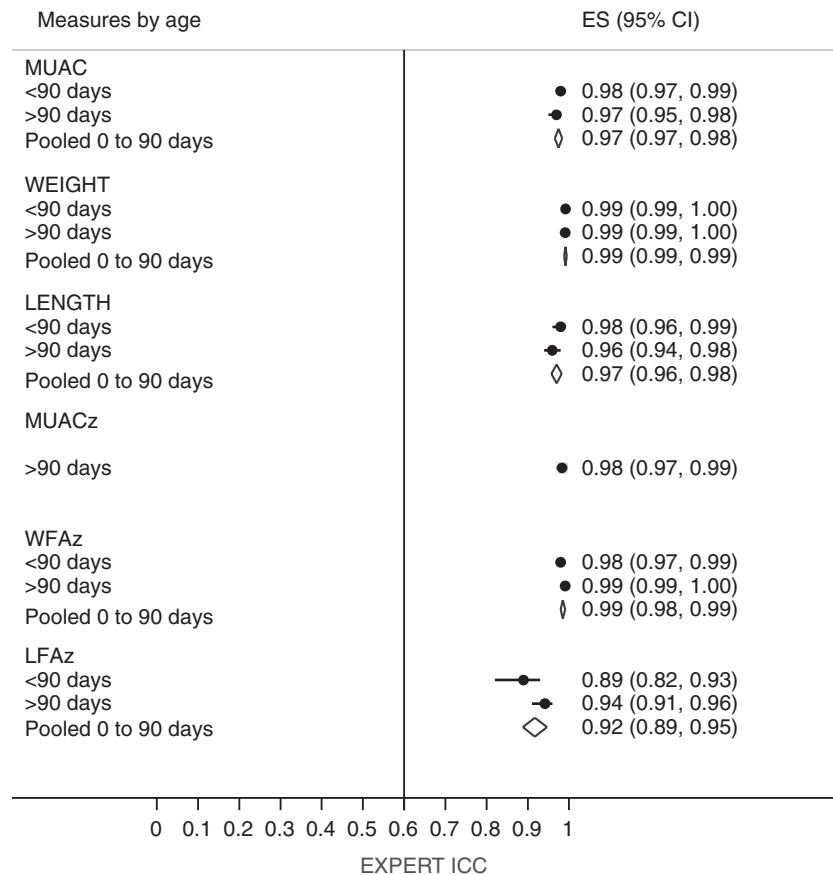


Figure 1 Intra-Class Correlation Coefficient (ICC) measures by expert anthropometrist (Line at ICC = 0.6 shows minimal acceptable reliability).

Stage III: CHWs

Eighteen CHWs measured 924 infants at six sites, median age 84 days, 480 (52%) were <90 days and 478 (52%) were male.

The mean (SD) estimate for MUAC was 127 mm (1.75), for weight 5.6 kg (1.6) and for length 58.0 cm (6.1). The mean WFAz was -0.29 (1.1), WFLz was 0.37 (1.2) and LFAz was -0.71 (1.36). One hundred and sixty-seven (18%) infants had an MUAC < 110 mm, 18 (2%) were wasted (WFLz < -2), 132 (14%) were stunted (LFAz < -2) and 52 (6%) were underweight (WFAz < -2).

The ICCs, pooled across the six sites, were 0.96 (95% CI 0.95–0.96) for MUAC and 0.71 (95% CI 0.68–0.74) for WFLz (Figure 3).

At each site, 30 infants were measured for accuracy by each CHW-HP pair, giving 540 measurements. The mean differences for MUAC, weight, length and WFLz were 0.65 mm (95% CI 0.02–1.07 mm), 0.006 kg (95% CI -0.019 to 0.031 kg), 0.17 cm (95% CI -0.001 to +0.34 cm) and -0.034 Z score (95% CI -0.13 to +0.06 z scores), respectively. We found no strong evidence that the

variance of the paired MUAC measures by CHWs differed from that of their trainers (Pitman's statistic P -value = 0.075). The Bland–Altman plot for both MUAC and WFLz as taken by CHWs are presented in Figure 4.

Discussion

We aimed to determine if MUAC has acceptable inter-observer variability (reliability) and accuracy among CHWs measuring infants aged <6 months. With 1 day of training, CHWs measured MUAC with high reliability: the overall ICC was 0.96 and at all sites it was >0.90. The mean difference between the CHWs and their trainers in MUAC was 0.65 mm (95% CI 0.023–1.07), with no significant difference in variance. We believe this is not of major clinical importance as the larger discrepancies in MUAC occurred with larger children (Figure 4).

Reliability

In routine practice, anthropometric measurement errors are common and such errors influence interpretation

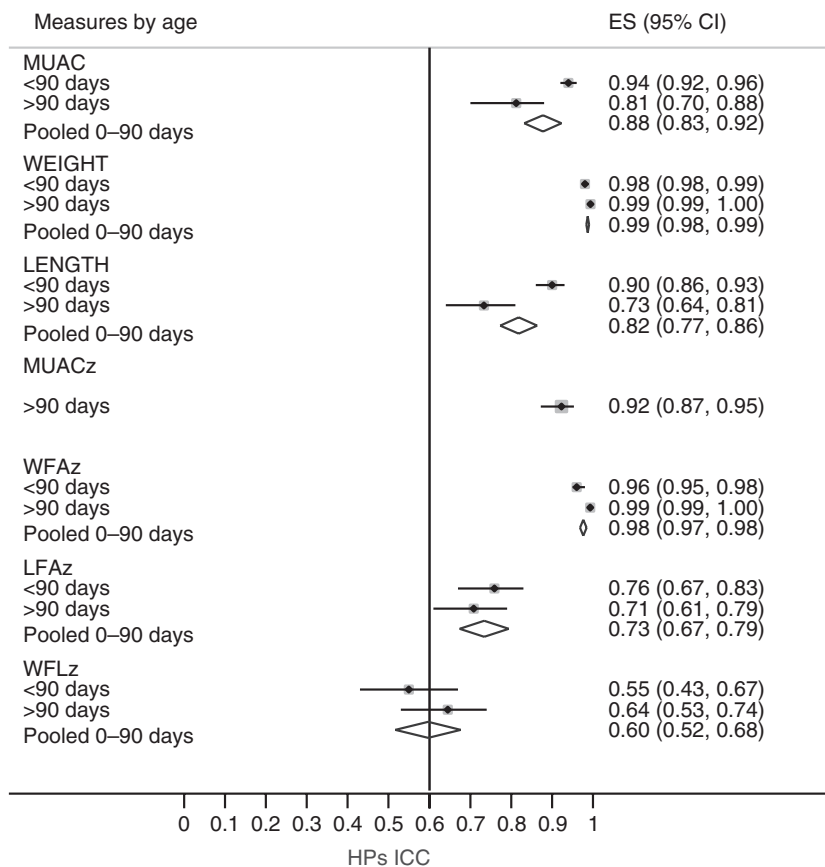
M. K. Mwangome *et al.* Reliability of MUAC in infants under 6 month

Figure 2 Intra-Class Correlation Coefficient (ICC) measures by HPs. HPs, health professionals (Line at ICC = 0.6 shows minimal acceptable reliability).

(Ulijaszek & Kerr 1999). Previous studies of reliability have used a variety of methods to recruit study participants, train observers and analyse data in their attempt to assess and quantify these errors. The WHO multicentre growth reference study group applied a process of rigorous training and monitoring under research conditions, rather than among typical health workers. They calculated ‘coefficients of reliability’, reporting these to be high (>0.95) for all anthropometric measures undertaken in newborns, infants and older children, except skinfold thickness (WHO 2006a). Among ‘minimally trained’ health workers in Guatemala, the ‘reliability coefficients’ of all anthropometric measures were <0.90 among children of 12–60 months (Velzeboer *et al.* 1983). Fewer and smaller errors were made for arm circumference than for WFLz, which concurs with our findings.

We found that absolute measures of MUAC, weight and length were more reliable than calculated Z scores. Among the absolute measures, length was the least reliable when measured by HPs; pooled ICC was 0.82. Overall, WFLz was the least reliable anthropometric index: the overall

ICC was 0.71, and at one site, WFLz met our criteria for ‘completely unreliable’. The most likely explanation is that WFLz is very sensitive to changes in the absolute measurements. To investigate this, we used the WHO anthropometric calculator (WHO 2011) to examine the effect of hypothetical errors in length and weight measurement on WFLz score. For a female child weighing 6 kg and measuring 65 cm in length (WFLz –1.88), we found that 1 cm change in length measurement (a 1.5% change) resulted in a 21% change in WFLz to –2.25 z scores. A 100 g change in weight (1.67% change) results in a 10% change in WFLz to –1.68 z scores; a change that could easily be attained immediately after feeding or passing urine. Additionally, our results indicate higher likelihood of variation in absolute length compared to weight measurements which is similar to findings from other studies (Velzeboer *et al.* 1983).

We did not study errors in looking up or calculating z scores. However, in recalculating per cent weight-for-height from a single dataset within a 1-month period, dietitians in the UK had wide intra-examiner variation

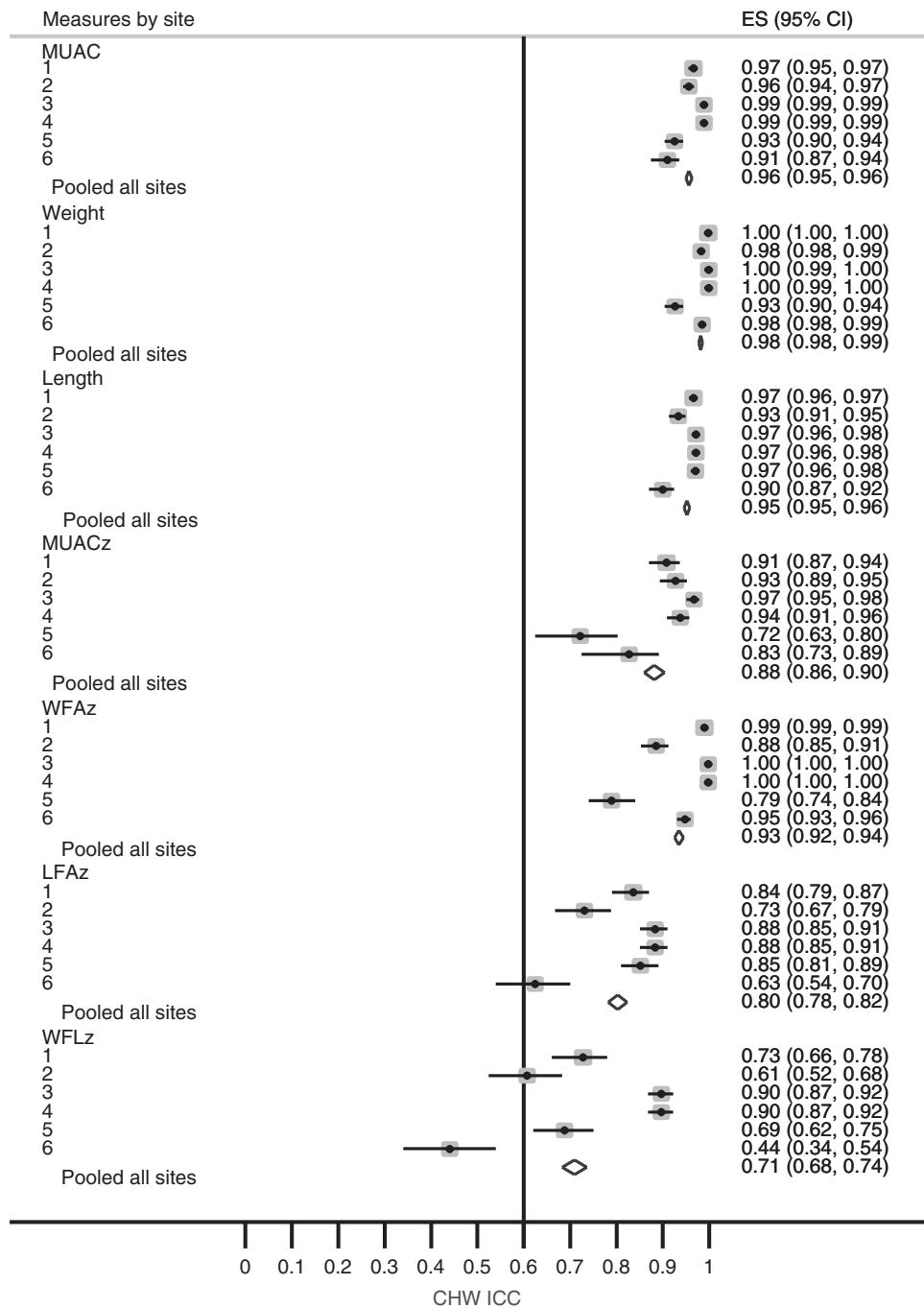


Figure 3 Intra-Class Correlation Coefficient (ICC) measures by (CHWs). CHWs, Community Health Workers (Line at ICC = 0.6 shows minimal acceptable reliability).

(between 13% and 24% differences) (Poustie *et al.* 2000). Inter-examiner estimates varied by 16.5–40%, suggesting that even among experienced staff, such calculation can be unreliable.

Accuracy

There are few published data on the accuracy of anthropometry in early infancy. None have used a systematic

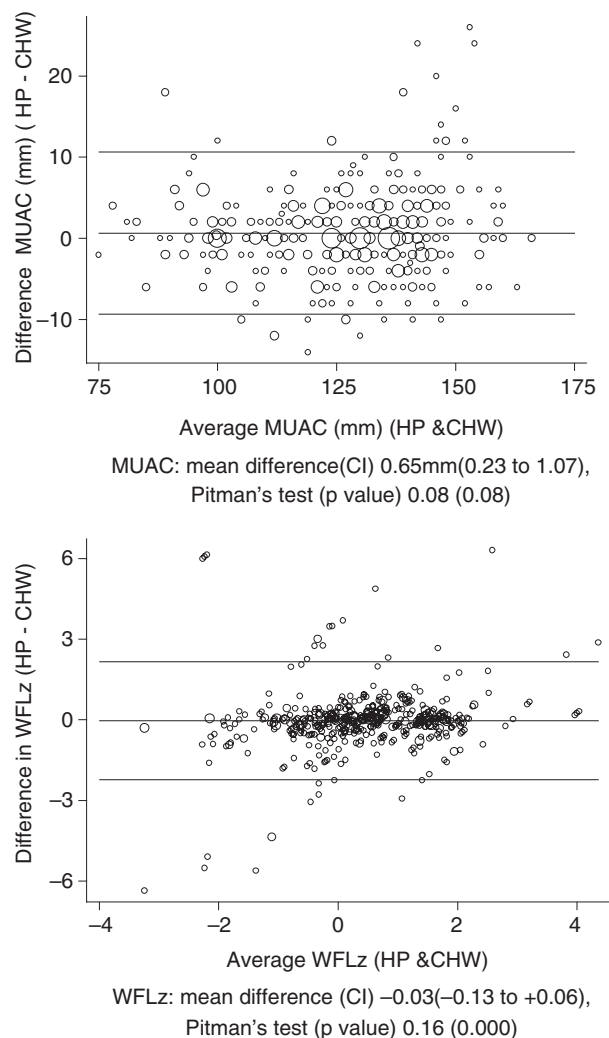


Figure 4 Bland–Altman plots for MUAC (mm) & WFLz by (CHWs). CHWs, Community Health Workers; MUAC, Mid Upper Arm Circumference; WFLz, weight-for-length Z score.

approach and included CHWs as used in this study. In our study, there was better concordance between CHWs and their trainers for MUAC than for WFLz (Figure 4). In the WHO growth reference study, there was a comparable technical error of measurement (TEM) between experts and observers for all measurements in all sites; however, it also reported an average bias in length; trained observers tended to underestimate length by -0.21 to -0.37 cm relative to the expert (WHO 2006a). No average bias in MUAC was found, which is consistent with our findings.

In other studies, MUAC achieved high specificity ($>95\%$) and varied sensitivity (48–58%) in identifying infants with severe malnutrition (WFLz < -3) (Fernandez *et al.* 2010) and low-birth-weight (weight < 2500 g) (Ramaiya *et al.*

1994). In such studies, however, there is a trade-off between specificity and sensitivity, and therefore reported levels of sensitivity and specificity should be interpreted within the study context and the cut-off used.

This study evaluated CHWs across a representative range of sites in a rural district in a realistic setting after a typical practical training. Such an arrangement is likely to be similar to proposed changes by the Government of Kenya's Ministry of Health when aiming to better identify infants *in situ* within their community at risk of malnutrition.

Evaluation in the context of research may have resulted in a better than normal environment for measuring infants in that there was more time with fewer interruptions and greater supervision. This may limit the application of these findings to rapid assessment or a door-to-door visit scenario. Secondly, the measuring equipment was in good condition and regularly calibrated. This may not normally be the case in rural public health facilities in Kenya. Thirdly, owing to a delay in equipment availability, the expert anthropometrist was unable to take weight and length measurements in the same group of infants. We therefore were unable to calculate and estimate expert's intra-observer variation of WFLz. Finally, the majority of the infants involved in this study were recruited from the routine GM clinics and not randomly selected within the community, thus they were relatively healthy. But because anthropometry is a non-invasive practical skill, CHWs should be able to replicate similar levels of reliability and accuracy among unhealthy infants. Further studies of the generalisability of our findings in other settings and to assess the relationship of MUAC with mortality and illness to establish appropriate cut-off values for MUAC use among infants under 6 months old are needed.

Conclusion

Community health workers can be trained to take absolute MUAC, weight and length measurements accurately and reliably among infants age < 6 months. However, the length-based Z score indices, LFAz and WFLz, are the least reliable anthropometric measures. With appropriate cut-off values, and further studies of its relationships with mortality, MUAC could be used by minimally trained non-professionals for community-based screening of SAM in infancy.

Acknowledgements

This work was supported by a strategic award and personal fellowship from the Wellcome Trust and is published with the permission of the Director, KEMRI.

We acknowledge the expert anthropometrist Joyce Nduku of KEMRI, CPHR-Nairobi, the support from KEMRI, CGMR-Coast, the Ministry of Health-Kenya for permission to access health facilities and staff, the study participants and community for their co-operation.

References

- Black RE, Allen LH, Bhutta ZA *et al.* (2008) Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* **371**, 243–260.
- Bonett DG (2002) Sample size requirements for estimating intraclass correlations with desired precision. *Statistics in Medicine* **21**, 1331–1335.
- de Onis M, Onyango AW, Borghi E, Garza C & Yang H (2006) Comparison of the World Health Organization (WHO) Child Growth Standards and the National Center for Health Statistics/WHO international growth reference: implications for child health programmes. *Public Health Nutrition* **9**, 942–947.
- Fernandez MA, Delchevalerie P & Van Herp M (2010) Accuracy of MUAC in the detection of severe wasting with the new WHO growth standards. *Pediatrics* **126**, e195–e201.
- Fjeld E, Siziya S, Katepa-Bwalya M, Kankasa C, Moland KM & Tylleskar T (2008) ‘No sister, the breast alone is not enough for my baby’ a qualitative assessment of potentials and barriers in the promotion of exclusive breastfeeding in southern Zambia. *International Breastfeeding Journal* **3**, 26.
- Kenya National Bureau of statistics (KNBS) & ICF Macro (2009) Kenya demographic and health survey 2008–09 [Online]. Government of Kenya: Central Bureau of Statistics. <http://www.measuredhs.com/pubs/pdf/FR229/FR229.pdf>.
- Kerac M, Blencowe H, Grijalva-Eternod C *et al.* (2011) Prevalence of wasting among under 6-month-old infants in developing countries and implications of new case definitions using WHO growth standards: a secondary data analysis. *Archives of Disease in Childhood* **96**, 1008–1013.
- Kikafunda JK, Walker AF, Collett D & Tumwine JK (1998) Risk factors for early childhood malnutrition in Uganda. *Pediatrics* **102**, E45.
- Kirkwood BR & Sterne JAC (2003) *Medical Statistics*. Blackwells, Oxford.
- Lopriore C, Dop MC, Solal-Celigny A & Lagnado G (2007) Excluding infants under 6 months of age from surveys: impact on prevalence of pre-school undernutrition. *Public Health Nutrition* **10**, 79–87.
- Ministry of Health (2006) *Taking the Kenya Essential Package for Health to the Community: A Strategy for the Delivery of Level One Services*. Publisher by Ministry of Health, Afya house, Nairobi, Kenya.
- Myatt M, Khara T & Collins S (2006) A review of methods to detect cases of severely malnourished children in the community for their admission into community-based therapeutic care programs. *Food and Nutrition Bulletin* **27**, S7–S23.
- Nwankwo BO & Brieger WR (2002) Exclusive breastfeeding is undermined by use of other liquids in rural southwestern Nigeria. *Journal of Tropical Pediatrics* **48**, 109–112.
- Poustie VJ, Watling RM, Ashby D & Smyth RL (2000) Reliability of percentage ideal weight for height. *Archives of Disease in Childhood* **83**, 183–184.
- Ramaiya C, Msamanga G, Massawe S, Mpanju W & Ngwalle E (1994) Newborn’s arm circumference as a screening tool of low birth weight in Temeke District, Dar es Salaam, Tanzania. *Tropical and Geographical Medicine* **46**, 318–321.
- Ulijaszek SJ & Kerr DA (1999) Anthropometric measurement error and the assessment of nutritional status. *British Journal of Nutrition* **82**, 165–177.
- United Nations (1986) *Summary Procedures of, “How to Weigh and Measure Children: Assessing the Nutritional Status of Young Children in Household Surveys”*. The United Nations Department of Technical Co-operation for Development and Statistics, New York.
- Velzeboer MI, Selwyn BJ, Sargent F 2nd, Pollitt E & Delgado H (1983) The use of arm circumference in simplified screening for acute malnutrition by minimally trained health workers. *Journal of Tropical Pediatrics* **29**, 159–166.
- Voss LD, Bailey BJ, Cumming K, Wilkin TJ & Betts PR (1990) The reliability of height measurement (the Wessex Growth Study). *Archives of Disease in Childhood* **65**, 1340–1344.
- Walter SD, Eliasziw M & Donner A (1998) Sample size and optimal designs for reliability studies. *Statistics in Medicine* **17**, 101–110.
- WHO (2006a) Reliability of anthropometric measurements in the WHO multicentre growth reference study. *Acta Paediatrica. Supplement* **450**, 38–46.
- WHO (2006b) *The WHO Child Growth Standards* [Online]. World Health Organisation. <http://www.who.int/childgrowth/standards/en/>.
- WHO (2010) *WHO Growth Standards STATA Macro* [Online]. World Health Organisation. <http://www.who.int/childgrowth/software/en/>.
- WHO (2011) *WHO Anthro (Version 3.2.2, January 2011) and macros* [Online]. <http://www.who.int/childgrowth/software/en/>.
- World Bank (2008) *Kenya Poverty and Inequality Assessment* [Online]. Poverty Reduction and Economic Management Unit Africa Region. <http://siteresources.worldbank.org/INTAFRREGTOPGENDER/Resources/PAKENYA.pdf> [Accessed 44190-KE].

Corresponding Author Martha K. Mwangome, KEMRI/Wellcome Trust Research Programme, PO Box 230, Kilifi 80108, Kenya.
E-mail: mmwangome@kilifi.kemri-wellcome.org