Public Health Nutrition: 16(2), 281–288 doi:10.1017/S1368980012002741

Short Communication

# Measurement error of waist circumference: gaps in knowledge

Lisanne M Verweij[1,2], Caroline B Terwee[3], Karin I Proper[1,2,*], Carel TJ Hulshof[4,5] and Willem van Mechelen[1,2]

[1]Department of Public and Occupational Health, EMGO Institute for Health and Care Research, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands: [2]Body@Work, Research Center Physical Activity, Work and Health, TNO-VU University Medical Center, Amsterdam, The Netherlands: [3]Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands: [4]Netherlands Society of Occupational Medicine, Utrecht, The Netherlands: [5]Coronel Institute of Occupational Health, Academic Medical Center (AMC), Amsterdam, The Netherlands

Submitted 7 September 2011: Final revision received 6 April 2012: Accepted 17 April 2012: First published online 25 May 2012

## Abstract

*Objective:* It is not clear whether measuring waist circumference in clinical practice is problematic because the measurement error is unclear, as well as what constitutes a clinically relevant change. The present study aimed to summarize what is known from state-of-the-art research.

*Design:* To identify the magnitude of the measurement error of waist circumference measurements from the literature, a search was conducted in PubMed from 1975 to February 2011.

*Results:* The measurement error may vary between 0·7 cm and 15 cm. Taking a realistic range of measurable waist circumference into account (60–135 cm), we argue that a short-term clinically relevant change in waist circumference of 5 % may lie between 3·0 and 6·8 cm and a maintained clinically relevant change of 3 % between 1·8 and 4·1 cm.

*Conclusions:* Based on these results, we conclude it may be difficult to distinguish clinically relevant change from measurement error in individual subjects, due to the large measurement error and unclear definition of clinically relevant change. More research is needed to address these gaps in knowledge. To minimize measurement error, we recommend using a uniform measurement protocol, training and repeated measurements.

**Keywords**
Waist circumference
Measurement error

Clinical guidelines recommend physicians to base their obesity-related cardiovascular risk management on abdominal as well as general obesity[1,2]. Abdominal obesity (as measured by waist circumference) has been conveyed as a better independent predictor of obesity related-disorders than general obesity (as measured by BMI)[3–5]. The presence of abdominal obesity can thus indicate the need for interventions in subjects who would otherwise not be considered at risk based on general obesity alone[1]. Measuring abdominal obesity using the waist circumference has been marked as feasible because it is easy to learn, takes no more time than measuring body height and weight, and requires minimal costs[6].

Waist circumference is now used more often to monitor changes as a result of interventions, not only by trained researchers but also by clinicians in primary care settings as well[7–9]. However, physicians report they find it hard to measure[10–12]. Moreover, studies showing good reliability of waist circumference measurements are mainly performed by health professionals trained in anthropometrics (1·3 to 6·5 cm)[13–15] while studies in which measurements are conducted by physicians show larger variability (0·7 to 12 cm)[16–18].

The consequences of variability for clinical practice are not yet clear. This depends on whether the variability is so large that clinically relevant changes within subjects or clinically relevant differences between subjects cannot be measured reliably. Three problems can be identified here. First, only a few reliability studies are available. Second, many studies report reliability (e.g. the intra-class correlation coefficient), but not an absolute measurement error (e.g. in centimetres). This information is needed to interpret

*Corresponding author: Email ki.proper@vumc.nl

change scores in individual subjects in clinical practice[19]. Third, it is not clear what a clinically relevant change in waist circumference is, because there is insufficient evidence on the dose–response relationship between reductions in waist circumference and obesity-related morbidity and mortality[11,20,21]. Thus, it is necessary to summarize what is known from state-of-the-art research and identify gaps in knowledge. The aims of the present study were therefore to: (i) explain the difference between reliability and measurement error and highlight why it is important to determine the measurement error of waist circumference; (ii) discuss what is known about the measurement error of waist circumference and which factors may cause this; (iii) discuss what is known about clinically relevant changes in waist circumference; (iv) discuss how knowledge about clinically relevant changes can help interpret the magnitude and importance of the measurement error of waist circumference; and finally (v) provide recommendations for future research and clinical practice on the measurement error of waist circumference.

## The difference between reliability and measurement error

The terms 'reliability' and 'measurement error' are part of the concept term 'reproducibility', as both address whether measurement results are reproducible in test–retest situations[19]. Reliability refers to how well subjects can be distinguished from each other in populations, despite the measurement error. This information is required for instruments that are used for discriminative purposes (e.g. to characterize individual differences between subjects in order to establish their clinical status and therapeutic needs, such as for discriminating between overweight and obese subjects). Measurement error assesses exactly how close values of repeated measurements within subjects are. This information is required for instruments that are used for evaluative purposes (e.g. to register change over time). The difference between reliability and measurement error is important for clinical practice because when studies present the reliability (e.g. intra-class correlation coefficient) of waist circumference measurements, a clinician is informed whether the instrument is able to discriminate between (e.g. overweight and obese) subjects in a sample. The clinician is not informed whether the instrument is suitable for monitoring waist circumference of individual subjects over time. In the latter case, the absolute measurement error around a single measurement of a single change score is important[19]. This measurement error is expressed in for example the standard error of measurement or the limits of agreement[22]. Moreover, measurement error provides an important advantage over reliability for clinical interpretation as it is expressed on the actual scale of measurement (e.g. centimetres), and not as a dimensionless value between 0 and 1. While information on both reliability and measurement error is necessary for clinical practice, reliability is generally high[15] but the magnitude of measurement error is not clear. In the present paper we focus on this absolute measurement error, which influences measurements in individual persons[22].

## Measurement error of waist circumference

To identify the magnitude of the measurement error of waist circumference measurements from the literature, a systematic search was conducted in PubMed from 1975 to February 2011. Search terms for measurement error were selected from a search filter that was developed for finding studies on measurement properties[23] and combined with the text word 'waist circumference'. Studies using self-reported measurements and those among children or adolescents were excluded because these are associated with higher measurement error[15]. Data were extracted on the smallest detectable change (SDC) or smallest detectable difference (SDD), which reflect the smallest change or difference in waist circumference of an individual subject that can be detected beyond measurement error[19]. The search resulted in 559 studies, of which nine reported on the intra- or inter-observer measurement error of waist circumference (e.g. repeated measurements on the same subjects by one observer or by different observers, respectively; Table 1). The methodological quality of studies was assessed by two authors (L.M.V. and C.B.T.) using the COSMIN checklist for grading studies on measurement properties (Box C)[24]. An overall methodological quality score was obtained by taking the lowest rating of the eleven items ('worst score counts') from the following ratings: excellent, good, fair or poor. For example, if for a study one item scored poor, the overall methodological quality of that study was rated as poor.

The selected studies included between seven and 9279 participants, consisting of healthy adults to employees or patients. The outcome assessors were physicians in three studies and other health professionals in six studies. The outcome assessors were trained in advance in seven studies or between repeated measurements in two studies. All followed a standard (although different) protocol. Participants were measured in standing position, except for one study that measured participants in supine position. These measurements were carried out midway between the lower rib and the iliac crest in five studies, at the narrowest point between the rib cage and the iliac crest in one study, and at the uppermost limit of the ileum in another study. One study examined the effect of measurement site (lower rib, iliac crest or midway) on the measurement error[25]. Finally, the overall methodological quality of the studies was fair or poor.

Overall, the intra-observer measurement error varied from 0·7 cm to 9·2 cm. The inter-observer measurement

**Table 1** Studies reporting intra- or inter-observer measurement error of waist circumference measurements

| First author, year, reference and country | Design and population | Outcome assessor and training | Number of measurements and protocol | Measurement site | Intra- and inter-observer measurement error (SDC and SDD) | Quality |
|---|---|---|---|---|---|---|
| Bosy-Westfahl, 2010[25]<br>Germany | Cross-sectional.<br>16 lean and obese adults | 4 nutritionists.<br>Well-educated and trained, regular comparison of results | 3 non-consecutive measurements, one measurement per site.<br>Protocol: refs (1, 38–40).<br>Horizontal tape, after normal expiration, non-elastic plastic tape, standing, no compression of skin | (i) Lowest rib (distal border)<br>(ii) Iliac crest (superior border)<br>(iii) Midway between both | Intra-observer:<br>(i) 3·3 cm<br>(ii) 6·1 cm<br>(iii) 5·5 cm<br>Inter-observer:<br>(i) 6·7 cm<br>(ii) 15·0 cm<br>(iii) 14·1 cm | Poor |
| Dhaliwal 2009[47]<br>Australia | Cross-sectional.<br>9279 subjects from the Australian Risk Factor Prevalence Study (1989).<br>Aged 20–69 years, 93 % Australian, UK, European and 5 % Asian or African | 2 observers (*n* unknown at 8 survey sites).<br>Trained | 2 consecutive measurements.<br>Protocol: refs (41, 42).<br>Horizontal tape, after full expiration, metal tape, no compression of skin, no clothing around the waist, front of subjects | Narrowest point between the rib cage and iliac crest | Intra-observer:<br>1·8 cm (men)<br>1·7 cm (women) | Fair |
| Geeta 2009[43]<br>Malaysia | Cross-sectional.<br>130 working adults at selected office setting (2005).<br>Aged 18–64 years, mean age 36 (SD 11) years, 67 % female, 83 % Malaysian.<br>Non-pregnant, ≥2 months postnatal, no physical disability or body deformation to stand upright | 2 public health nurses.<br>Trained, unaware of objective, previous measurement blinded | 3 consecutive measurements.<br>Protocol: clinical manual of NHMS III[45]<br>Horizontal tape, after normal expiration, Seca S201 tape, front of subjects | Midway between lowest rib margin and iliac crest | Intra-observer:<br>1·2 cm (lower, upper limit: −1·1, 1·3 cm)<br>Inter-observer:<br>2·1 cm (lower, upper limit: −1·9, 2·3 cm) | Fair |
| Nordhamn 2000[14]<br>Sweden | Cross-sectional.<br>25 lean and 26 overweight (BMI <26 and ≥26 kg/m$^2$) students, university staff and company employees, 50 % men | 5 observers from the metabolic unit.<br>Trained | 3 consecutive measurements, 1 measurement after 1–3 weeks.<br>Protocol: unknown.<br>Supine position | Midway between lower rib and iliac crest | Intra-observer:<br>≈5·3 cm (all)<br>≈3·8 cm (lean)<br>≈6·4 cm (overweight)<br>Intra-observer:<br>≈2·5 cm (all)<br>≈3·2 cm (lean)<br>≈2·1 cm (overweight) | Fair |
| Panoulas 2008[17]<br>UK | Cross-sectional.<br>102 patients in an outpatient department of a hospital for the preliminary study, 28 new patients matched to 28 original patients for the clinical practice study (2006) | Preliminary study: 4 doctors, 4 nurses, 2 physiotherapists.<br>Trained.<br>Clinical practice study: 3 medical students.<br>Untrained and after written instructions | Preliminary study: 3 measurements by the 9 other participants.<br>Clinical study: ?<br>Protocol: ref. (39).<br>During expiration, standard measuring tape, no clothing around the waist | Midway between the palpated lowest rib margin and iliac crest in the midaxillary lines | <u>Preliminary study</u><br>Intra-observer:<br>0·7 cm<br>Inter-observer:<br>1·4 cm<br><u>Clinical study, untrained</u><br>Inter-observer:<br>2·6 cm (*n* 102), 2·5 cm (*n* 28)<br><u>Clinical study, trained</u><br>Inter-observer:<br>3·3 cm (*n* 28) | Poor |

**Table 1** *Continued*

| First author, year, reference and country | Design and population | Outcome assessor and training | Number of measurements and protocol | Measurement site | Intra- and inter-observer measurement error (SDC and SDD) | Quality |
|---|---|---|---|---|---|---|
| Sebo 2008[16] Switzerland | Cross-sectional. 24 healthy adult volunteers. Mean age 41 (SD 14) years, 54 % women, 21 % overweight, 42 % obese | 12 doctors. Untrained and after a 1 h training session (theory, demonstration, practice) | 2 consecutive measurements by all 12 doctors, repeated after 1 week. Protocol: refs (3, 44–46). Horizontal tape, end of normal expiration, standard measuring tape, standing, no clothing around the waist, no compression of skin | Midpoint between lowest rib and iliac crest | Inter-observer: 12·0 cm (untrained) 7·2 cm (trained) | Poor |
| Sicotte 2010[18] Canada | Cross-sectional. 12 (at 3 months) and 7 other (at 18 months) HIV-patients in Mali (year unknown) | 1 doctor, 2 health professionals, without prior experience. Trained, written instructions, practice every 2 weeks during 3 months | 2 times on 2 consecutive days by the same observer on 2 study occasions. Protocol: unknown. Marks on location, horizontal tape, non-stretchable Gulick tape | Uppermost limits of the ileum | Intra-observer: 3·4–9·2 cm Inter-observer: 5·5–6·5 cm | Poor |
| Ulijaszek 1999[15] UK | Review: (i) 2 studies between 1987 and 1995 (ii) 8 studies between 1987 and 1997 | Health professionals | Not specified | Not specified | Inter-observer: (i) 3·6 cm (range: 2·8–4·4 cm) Intra-observer: (ii) 6·5 cm (range: 1·7–11·7 cm) | Fair |
| Wang 2010[28] Taiwan | Cross-sectional. 76 participants from in-patient wards (year unknown). 33 % men, mean age 47 (SD 14) years, underweight (*n* 5), normal (*n* 27), overweight (*n* 21), obese (*n* 23) | 2 research assistants, without prior experience. Trained, instructions, 20-min practice | 2 times with 10-min intervals. Protocol: unknown. Horizontal tape, end of normal expiration, inelastic plastic measuring tape, standing, feet apart and arms hanging freely | Midpoint between lowest rib and iliac crest | Intra-observer: 3·0 cm (all) 1·7 cm (underweight) 1·7 cm (normal), 2·2 cm (overweight) 4·7 cm (obese) | Fair |

ref./refs = reference number(s) in the References list.

error varied from 1·4 cm to 15 cm (Table 1). In most studies that measured both, the intra-observer measurement error was smaller[15]. Moreover, smaller intra- and inter-observer measurement errors were found in larger studies. No notable differences in relation to the measurement error were observed according to participant characteristics, outcome assessor, measurement protocol, effects of training or methodological quality. However, greater measurement error was reported from measuring at the iliac crest or midway, compared with the lower rib, possibly because the latter is most easily located[25].

Based on the small number of studies and the many differences between the studies, we conclude that it is difficult to draw conclusions on the magnitude of measurement error. Moreover, the variation in measurement error may be caused by a number of other factors not mentioned in Table 1 such as muscle mass, bone structure, lean tissues, looseness of abdominal muscles, posture, phase of respiration and time since the last meal[26,27]. Additionally, measurement error may be larger among overweight and obese subjects compared with normal-weight subjects due to difficulty in locating anatomical landmarks[14,15,27,28].

## Clinically relevant change in waist circumference

Whether the measurement error is problematic in clinical practice can only be judged if there is a clear conception of the magnitude of change in, or the difference between, waist circumference that is considered important. In other words, we need to identify a minimal important change (MIC) within subjects or a minimal important difference (MID) between subjects in waist circumference[29]. While several studies suggest that a reduction of waist circumference may be associated with benefits across a wide range of health outcomes, there is limited evidence for what constitutes a minimal important change or difference in waist circumference[30–33]. The National Institutes of Health stated in 1998 that a sustained reduction of 4 cm may be clinically relevant[34]. More recently it has been suggested that, similar to body weight and BMI, a reduction in waist circumference of >5 % may be considered a clinically relevant change for individual subjects in the short term and a maintained waist circumference of >3 % from initial waist circumference may be considered clinically relevant for individual subjects in the long term[35] (I Lemieux and R Ross, personal communications). No clear definitions were provided on what short-term change and long-term maintenance are[35]. Following that recommendation, for an overweight woman with a waist circumference of 80 cm this corresponds to a waist reduction of at least 4·0 cm and a maintained reduction in waist circumference of at least 2·4 cm. For an obese woman with a waist circumference of 110 cm, this corresponds to a waist reduction of at least 5·5 cm and a maintained reduction in waist circumference of at least 3·0 cm. This shows that for subjects with a larger waist circumference, a larger reduction in waist circumference is necessary for change to be clinically relevant. Taking a realistic range of measurable waist circumference, for example 60–135 cm, this implies that a short-term change between 3·0 and 6·8 cm and a maintained change between 1·8 and 4·1 cm may be clinically relevant.

## The relationship between measurement error and clinically relevant change

In order to distinguish clinically relevant change from measurement error, the measurement error (SDC) should be smaller than the clinically relevant change (MIC; see Fig. 1a). In this case, changes as large as the clinically

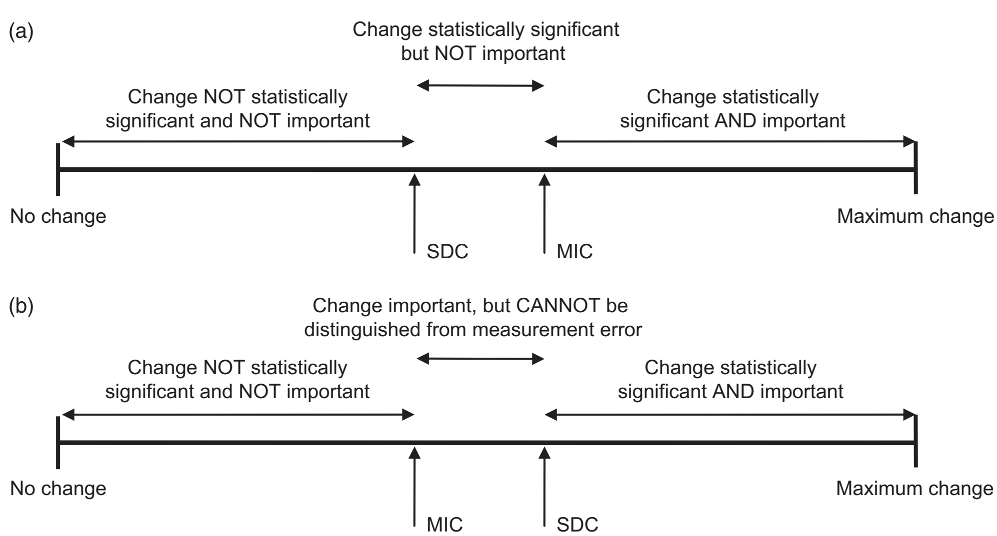


**Fig. 1** Interpretation of change: (a) the smallest detectable change (SDC) as a parameter of measurement error is smaller than the minimal important change (MIC); and (b) the smallest detectable change (SDC) as a parameter of measurement error is larger than the minimal important change (MIC). Adapted from Terwee *et al.*[29]

https://doi.org/10.1017/S1368980012002741 Published online by Cambridge University Press

relevant change will be statistically significant[29]. Thus the smaller is the measurement error, the smaller the change that can be detected beyond the measurement error. But if the measurement error (SDC) is larger than the clinically relevant change (MIC), this change cannot be distinguished from measurement error (see Fig. 1b).

The range of measurement error presented in Table 1 (0·7–15 cm) indicates that we are probably able to detect a short-term clinically relevant change of 4·0 cm (5 % for a women of 80 cm) or 5·5 cm (5 % for a women of 110 cm), as the intra-observer measurement error is smaller than 4 cm in all but one study. However, the probability to detect a long-term clinically relevant change of 2·4 cm (3 % for a women of 80 cm) or 3·0 cm (3 % for a women of 110 cm) is much lower, as the intra-observer measurement error is larger than 3 cm in more than half of the studies. Across the realistic range of waist circumference measurements (60–135 cm), many relevant short-term changes (between at least 3·0 and 6·8 cm) and maintained changes (between at least 1·8 and 4·1 cm) probably cannot be distinguished from measurement error. Interestingly, the measurement error of waist circumference seems equally problematic for normal-weight, overweight or obese subjects. Although the measurement error is larger among overweight or obese subjects, a larger reduction in waist circumference is also necessary to obtain a clinically relevant change.

## Recommendations for future research and clinical practice

To summarize, we have shown that there are two important gaps in knowledge. First, the assessment of measurement error identified a wide range (0·7–15 cm) of measurement errors, due to the small number of fair and poor quality studies and many differences between studies. Second, no clear definition of clinically relevant change could be extracted from the literature. Taking a realistic range of a measurable waist circumference (60–135 cm) into account, we argue that a proposed clinically relevant change in waist circumference of 5 % in the short term (approximately 3·0–6·8 cm) may be detectable, but a proposed maintenance of 3 % (approximately 1·8–4·1 cm) may not be detectable, because it cannot be distinguished from measurement error. Although the current paper does not provide practising clinicians with empirical insight into the application and interpretation of waist circumference measurements in the clinical setting, the results do highlight that more attention should be paid to reducing measurement error, in order for clinicians and researchers to accurately measure real change in waist circumference rather than measurement error.

Three ways to potentially reduce measurement error in clinical practice are: (i) adopting a standard protocol; (ii) training; and (iii) repeating measurements[15]. Two papers studied the influence of using different measurement protocols on waist circumference measurements. The first found that using different measurement protocols influenced the association between waist circumference, all-cause and CVD mortality, CVD and diabetes[36]. However these protocols were only compared on measurement site. The second study found that the type of protocol significantly influenced waist circumference measurements by comparing the measurement of waist circumference in eleven different ways (by anatomical site, posture, respiratory phase and time since the last meal)[27]. However, as we have shown, other factors may also influence measurement error and smaller measurement errors are required in order to detect (smaller) changes beyond measurement error. For clinicians, no standard protocol was advised as best. To overcome this gap in knowledge, we support the worldwide request for a uniform measurement protocol, decided upon by an expert team[10,25,37,47].

A second way to reduce measurement error is by training. Measurement error is likely to be larger if measurements are carried out by poorly (often recently) trained individuals[15]. Training may thus reduce measurement error by quality control across time and by minimizing the number of observers[15]. Unfortunately, it is unclear how (much) training is needed to decrease measurement error, nor whether the effect of training is sustained over time[16,17].

A third way to reduce measurement error is to repeat waist circumference measurements. If the same measurement is repeated for example two or three times and the average value is taken, the measurement error of this average value is much smaller (by a factor $\sqrt{k}$, with $k$ being the number of repeated measurements)[48]. For example, taking the realistic short-term (approximately 3·0–6·8 cm) and long-term (approximately 1·8–4·1 cm) clinically relevant change, two repeated measurements would result in an average measurement error of 2·1–4·8 cm for short-term clinically relevant change and 1·3–2·9 cm for long-term clinically relevant change. Three repeated measurements would result in an average measurement error of 1·7–3·9 cm for short-term clinically relevant change and 1·0–2·4 cm for long-term clinically relevant change. Thus, two measurements seem to be sufficient for detecting short-term changes, but three measurements seem to be necessary to distinguish long-term change from measurement error.

## Conclusions

Four gaps in knowledge have been identified. First, the magnitude of measurement error in waist circumference is unclear. Second, the definition of clinically relevant change in waist circumference is unclear. We therefore caution clinicians and researchers when interpreting individual changes in waist circumference, as clinically relevant changes in waist circumference may not be distinguished from measurement error. Third, consensus is needed on adopting a uniform protocol for measuring waist circumference. Fourth, there is a lack of knowledge on the effects of training on measurement error in waist circumference. Considering these gaps in knowledge, it is clear that there is a need for more good quality research and for action. Until then, we recommend consistently using one standard protocol, quality control as part of training and minimizing the number of observers, outsourcing measurements to well-trained clinicians and repeating measurements at least two, but preferably three times. Ultimately, by reducing measurement error, smaller changes in waist circumference may be detected by clinicians beyond measurement error. This is necessary for accurately monitoring changes in waist circumference of individual subjects over time.

## Acknowledgements

This study was funded by The Netherlands Organization for Health Research and Development (ZonMw, project 120510007). The authors declare that they have no conflict of interest. L.M.V. wrote the initial manuscript. C.B.T. provided intellectual input and wrote sections of the manuscript. K.I.P., C.T.J.H. and W.v.M. provided intellectual input and had a role in supervision. All authors have read and approved the final version of the manuscript.

## References

1. National Institute for Health and Clinical Excellence (2006) *Obesity: Guidance on the Prevention, Identification, Assessment and Management of Overweight and Obesity in Adults and Children. NICE Clinical Guideline* no. 43. London: NICE.
2. Ministry of Health, Clinical Trials Research Unit (2009) *Clinical Guidelines for Weight Management in New Zealand Adults*. Wellington: Ministry of Health.
3. Assmann G, Cullen P, Jossa F *et al.* (1999) Coronary heart disease: reducing the risk: the scientific background to primary and secondary prevention of coronary heart disease. A worldwide view. *Arterioscler Thromb Vasc Biol* **19**, 1819–1824.
4. Lean ME, Han TS & Morrison CE (1995) Waist circumference as a measure for indicating need for weight management. *BMJ* **311**, 158–161.
5. Yusuf S, Hawken S, Ounpuu S *et al.* (2005) Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case–control study. *Lancet* **366**, 1640–1649.
6. Wang J (2003) Waist circumference: a simple, inexpensive, and reliable tool that should be included as part of physical examinations in the doctor's office. *Am J Clin Nutr* **78**, 902–903.
7. Smith SC Jr & Haslam D (2007) Abdominal obesity, waist circumference and cardio-metabolic risk: awareness among primary care physicians, the general population and patients at risk – the Shape of the Nations survey. *Curr Med Res Opin* **23**, 29–47.
8. Pajak A, Szafraniec K, Janion M *et al.* (2010) The impact of the Polish national Programme of Cardiovascular Disease Prevention on the quality of primary cardiovascular disease prevention in clinical practice. *Kardiol Pol* **68**, 1332–1340.
9. Giorda CB, Guida P, Avogaro A *et al.* (2009) Association of physicians' accuracy in recording with quality of care in cardiovascular medicine. *Eur J Cardiovasc Prev Rehabil* **16**, 722–728.
10. Mason C & Katzmarzyk PT (2009) Variability in waist circumference measurements according to anatomic measurement site. *Obesity (Silver Spring)* **17**, 1789–1795.
11. Lee SK (2010) Should waist circumference be used to identify metabolic disorders than BMI in South Korea? *Eur J Clin Nutr* **64**, 1373–1376.
12. Verweij LM, Proper KI, Hulshof CTJ *et al.* (2011) Process evaluation of an occupational health guideline aimed at preventing weight gain among employees. *J Occup Environ Med* **53**, 722–729.
13. Moreno LA, Joyanes M, Mesana MI *et al.* (2003) Harmonization of anthropometric measurements for a multicenter nutrition survey in Spanish adolescents. *Nutrition* **19**, 481–486.
14. Nordhamn K, Sodergren E, Olsson E *et al.* (2000) Reliability of anthropometric measurements in overweight and lean subjects: consequences for correlations between anthropometric and other variables. *Int J Obes Relat Metab Disord* **24**, 652–657.
15. Ulijaszek SJ & Kerr DA (1999) Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr* **82**, 165–177.
16. Sebo P, Beer-Borst S, Haller DM *et al.* (2008) Reliability of doctors' anthropometric measurements to detect obesity. *Prev Med* **47**, 389–393.
17. Panoulas VF, Ahmad N, Fazal AA *et al.* (2008) The inter-operator variability in measuring waist circumference and its potential impact on the diagnosis of the metabolic syndrome. *Postgrad Med J* **84**, 344–347.
18. Sicotte M, Ledoux M, Zunzunegui MV *et al.* (2010) Reliability of anthropometric measures in a longitudinal cohort of patients initiating ART in West Africa. *BMC Med Res Methodol* **10**, 102.
19. de Vet HC, Terwee CB, Knol DL *et al.* (2006) When to use agreement versus reliability measures. *J Clin Epidemiol* **59**, 1033–1039.
20. Huxley R, Barzi F, Lee CM *et al.* (2007) Waist circumference thresholds provide an accurate and widely applicable method for the discrimination of diabetes. *Diabetes Care* **30**, 3116–3118.
21. Tuan NT, Adair LS, Stevens J *et al.* (2010) Prediction of hypertension by different anthropometric indices in adults: the change in estimate approach. *Public Health Nutr* **13**, 639–646.
22. de Vet HC, Terluin B, Knol DL *et al.* (2010) Three ways to quantify uncertainty in individually applied 'minimally important change' values. *J Clin Epidemiol* **63**, 37–45.
23. Terwee CB, Jansma EP, Riphagen II *et al.* (2009) Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* **18**, 1115–1123.
24. Terwee CB, Mokkink LB, Knol DL *et al.* (2012) Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* **24**, 651–657.
25. Bosy-Westphal A, Booke CA, Blocker T *et al.* (2010) Measurement site for waist circumference affects its accuracy as an index of visceral and abdominal subcutaneous fat in a Caucasian population. *J Nutr* **140**, 954–961.
26. Misra A, Wasir JS & Vikram NK (2005) Waist circumference criteria for the diagnosis of abdominal obesity are not applicable uniformly to all populations and ethnic groups. *Nutrition* **21**, 969–976.

27. Agarwal SK, Misra A, Aggarwal P *et al.* (2009) Waist circumference measurement by site, posture, respiratory phase, and meal time: implications for methodology. *Obesity (Silver Spring)* **17**, 1056–1061.
28. Wang CY, Liu MH & Chen YC (2010) Intrarater reliability and the value of real change for waist and hip circumference measures by a novice rater. *Percept Mot Skills* **110**, 1053–1058.
29. Terwee CB, Roorda LD, Knol DL *et al.* (2009) Linking measurement error to minimal important change of patient-reported outcomes. *J Clin Epidemiol* **62**, 1062–1067.
30. Cardona-Morrell M, Rychetnik L, Morrell SL *et al.* (2010) Reduction of diabetes risk in routine clinical practice: are physical activity and nutrition interventions feasible and are the outcomes from reference trials replicable? A systematic review and meta-analysis. *BMC Public Health* **10**, 653.
31. Berentzen TL, Jakobsen MU, Halkjaer J *et al.* (2010) Changes in waist circumference and mortality in middle-aged men and women. *PLoS One* **5**, e13097.
32. Voruganti VS, Cai G, Klohe DM *et al.* (2010) Short-term weight loss in overweight/obese low-income women improves plasma zinc and metabolic syndrome risk factors. *J Trace Elem Med Biol* **24**, 271–276.
33. Stevens VJ, Obarzanek E, Cook NR *et al.* (2001) Long-term weight loss and changes in blood pressure: results of the Trials of Hypertension Prevention, phase II. *Ann Intern Med* **134**, 1–11.
34. National Institutes of Health (2000) *The Practical Guide to the Identification, Evaluation and Treatment of Overweight and Obesity in Adults. NIH Publication.* no. 00-4084. Bethesda, MD: NIH.
35. Stevens J, Truesdale KP, McClain JE *et al.* (2006) The definition of weight maintenance. *Int J Obes (Lond)* **30**, 391–399.
36. Ross R, Berentzen T, Bradshaw AJ *et al.* (2008) Does the relationship between waist circumference, morbidity and mortality depend on measurement protocol for waist circumference? *Obes Rev* **9**, 312–325.
37. Park JY, Mitrou PN, Keogh RH *et al.* (2011) Effects of body size and sociodemographic characteristics on differences between self-reported and measured anthropometric data in middle-aged men and women: the EPIC-Norfolk study. *Eur J Clin Nutr* **65**, 357–367.
38. World Health Organization (1999) *Definition, Diagnosis and Classification of Diabetes Mellitus and Its Complications: Report of a WHO Consultation.* Geneva: WHO.
39. World Health Organization (1995) *Physical Status: The Use and Interpretation of Anthropometry. Report of a WHO Consultation. WHO Technical Report Series* no. 854. Geneva: WHO.
40. Alberti KG, Zimmet P & Shaw J (2005) The metabolic syndrome – a new worldwide definition. *Lancet* **366**, 1059–1062.
41. Alexander H & Dugdale A (1990) Which waist–hip ratio? *Med J Aust* **153**, 367–368.
42. Boyle CA, Dobson A, Egger G *et al.* (1993) Waist-to-hip ratios in Australia: a different picture of obesity *Aust J Nutr Diet* **50**, 57–64.
43. Geeta A, Jamaiyah H, Safiza MN *et al.* (2009) Reliability, technical error of measurements and validity of instruments for nutritional status assessment of adults in Malaysia. *Singapore Med J* **50**, 1013–1018.
44. Lean ME, Han TS & Deurenberg P (1996) Predicting body composition by densitometry from simple anthropometric measurements. *Am J Clin Nutr* **63**, 4–14.
45. Westat Inc. (1988) *National Health and Nutrition Examination Survey III (NHANES III). Body Measurements (Anthropometry).* Rockville, MD: Westat Inc.
46. Health Canada (2003) *Canadian Guidelines for Body Weight Classification in Adults.* Ottawa: Health Canada Publications Centre.
47. Dhaliwal SS & Welborn TA (2009) Measurement error and ethnic comparisons of measures of abdominal obesity. *Prev Med* **49**, 148–152.
48. Streiner DL & Norman GR (2008) *Health Measurement Scales. A Practical Guide to Their Development and Use.* New York: Oxford University Press.