Preventive Medicine 47 (2008) 389–393

Contents lists available at ScienceDirect

Preventive Medicine

journal homepage: www.elsevier.com/locate/ypmed

Reliability of doctors' anthropometric measurements to detect obesity

Paul Sebo [a,b], Sigrid Beer-Borst [c,d], Dagmar M. Haller [a], Patrick A. Bovier [a,*,1]

[a] Division of Primary Care Medicine, Geneva University Hospitals, Switzerland
[b] Community-based general internist, Geneva, Switzerland
[c] Division of Clinical Epidemiology, Geneva University Hospitals, Switzerland
[d] Section of Health, Bern University of Applied Sciences, Switzerland

ARTICLE INFO

Article history:
Available online 21 June 2008

Keywords:
Cardio-vascular risk factor
Obesity
Abdominal obesity
Anthropometric measurement
Reliability

ABSTRACT

Objective. To evaluate the reliability of anthropometric measurements (weight, height, Body Mass Index (BMI), waist and hip circumferences (WC; HC) and waist-to-hip ratio (WHR)) performed by doctors to assess obesity.

Method. Repeated anthropometric measurements were performed by 12 primary care physicians on 24 adult volunteers in Geneva, Switzerland, 2006. Volunteers (54% women, mean age 41) had a mean BMI of 28.1 (respective mean values for WC, HC and WHR: 91.4, 108.3, 0.84). Inter-observer reliability coefficient (R) and percent disagreement in categorisation of volunteers (normal weight, overweight, obesity, abdominal obesity) were computed according to these measurements.

Results. The inter-observer reliability for weight, height, and derived BMI were excellent (R>0.99), but unsatisfactory for WC (R=0.92), HC (R=0.76) and WHR (R=0.51). Based on the BMI, only 1% of the volunteers were misclassified as overweight or obese, whereas the use of WC and WHR lead to misclassification in 6% and 23% respectively. Reliability for the measurements improved after a one-hour training in anthropometric measurements (R=0.97 for WC, 0.92 for HC and 0.89 for WHR), but the proportion who were misclassified remained high despite the training session for WC (5%) and WHR (9%).

Conclusions. BMI remains the most reliable measure to detect obesity in medical practice, whereas WC, HC and WHR are less reliable. These results challenge current recommendations on obesity-related cardio-vascular risk management based on WC and WHR and underline the need for further research to improve the reliability of anthropometric measurements by doctors.

Introduction

The current progression in obesity prevalence is alarming since obesity is related with serious health consequences such as cardio-vascular disease, diabetes, osteoarthritis and some cancers (Kushner and Blatner, 2005, Morabia and Costanza, 2005, Okosun et al., 2004, WHO fact sheet 311, 2006). Doctors play an important role in the assessment and management of overweight and obesity and their associated health risks. Recent guidelines emphasise measuring abdominal as well as general obesity when assessing cardio-vascular risk (Janssen et al., 2004, Kanaya et al., 2003, National Institute for Health and Clinical Excellence, 2006, Wang, 2003), because abdominal obesity is an independent risk factor for arterial hypertension, diabetes and dyslipidaemia (Health Canada, 2003, International Task Force for Prevention of CHD, 1998, Janssen et al., 2004, Kanaya et al., 2003, Paccaud et al., 2000, Snijder et al., 2004). In particular, the presence of abdominal obesity can indicate the need for interventions in overweight patients who would otherwise not be considered at risk on the basis of body mass index (BMI, kg/m^2) alone (Booth et al., 2000, Gill et al., 2003, National Institute for Health and Clinical Excellence, 2006).

The waist circumference (WC) and the waist-to-hip ratio (WHR, WC divided by hip circumference, HC) have been proposed as reliable measures of abdominal adiposity (Wang, 2003, Zamboni et al., 1998). These measures, together with the assessment of the BMI, have the potential to help physicians in their assessment of their patients' obesity-related cardio-vascular risk and are also believed to be easy to perform. Anthropometric studies have shown that the intra-observer reproducibility (reproducibility of the measurement by the same observer) and the inter-observer reproducibility (reproducibility of the measurement by two or more observers) for these measurements were excellent (Chen et al., 2001, Moreno et al., 2003, Nordhamn et al., 2000, Ulijaszek and Kerr, 1999, Wang et al., 2003). To date, however, little consideration has been given to the fact that the studies

* Corresponding author. Division of Community and Primary Care Medicine, Geneva University Hospitals, 24 rue Micheli-du-Crest, CH-1211 Geneva 14, Switzerland. Fax: +41 22 372 9600.
E-mail address: patrick.bovier@post.harvard.edu (P.A. Bovier).
[1] Current address: 2 Grand-Rue, CH-1095 Lutry, Switzerland.

0091-7435/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ypmed.2008.06.012

assessing the reproducibility of these measurements involved only health professionals who had been trained in anthropometrics. Yet, if doctors are to provide appropriate guidance to their patients based on anthropometric measurements, they must perform them in a reliable way. But so far no data on the reproducibility of these measurements when performed by doctors have been published.

The aim of this study was to assess the reliability of anthropometric measurements in a group of doctors working in a teaching hospital. In addition, we aimed to explore whether the reliability of doctors' measurements could be improved by a short training session in anthropometrics.

Materials and methods

Recruitment of doctors and volunteers

The study took place at the Division of Primary Care Medicine, Geneva University Hospitals, Switzerland. Twenty doctors presently or formerly affiliated with the Division were personally invited to participate and 12 agreed. The participating doctors had a mean age of 35.5 years (range: 28–39, standard deviation (SD) 3.2) and 50% were males. They were predominantly experienced doctors (on average 8.1 years (SD 3.1) since graduation) but with limited experience in family medicine (mean experience in family medicine 2.7 years (SD 2.4)). They were given minimal indications of what the study was about. They were only told that they would have to perform a limited clinical examination on a group of adult volunteers twice over a period of 3 weeks, and attend a one-hour training session. Healthy adult volunteers ($N=24$) were recruited through advertisements. They had a mean age of 40.6 years (SD 14.1) and 54% (13/24) were women. The research protocol was accepted by the hospital's research ethics committee.

Data collection and training in anthropometrics sessions

The first measurement session was performed in 12 consultation rooms of the Division of Primary Care Medicine. The rooms were equipped with standardised, calibrated beam balances, stadiometers and measuring tapes. In each of the rooms, a completely dressed volunteer awaited the doctors. All volunteers were aware of the study procedure and told not to influence or help the doctors.

The 12 doctors were instructed just minutes before the study began. Each doctor was given 4 min to measure weight, height, WC and HC as per their standard practice. Each doctor started in one of the 12 consultation rooms and then moved on to the next room in a pre-established order. The volunteers always stayed in the same room to minimise measurement errors related to the measuring instruments.

When the doctors had completed the measurements on the first 12 volunteers, 12 other volunteers took place in the consultation rooms and the doctors started a new round of measurements after a 15 min break. Thus each of the 24 volunteers had their anthropometric measurements taken by each of the 12 doctors.

At the end of this first session, the doctors were asked to complete a questionnaire asking how frequently and how they performed the different anthropometric measurements in their daily practice (e.g. patient dressed or not, site of measurement for WC and HC) and how they calculated and interpreted the resulting indices (threshold values to define normal body weight, overweight, obesity and abdominal obesity).

One week after the first session, the doctors attended a one-hour training session in anthropometrics, conducted by a nutritional scientist (SBB). The training manual was based on international guidelines (Health Canada, 1995, Health Canada, 2003, International Task Force for Prevention of CHD, 1998, Lean and Han, 1996, National Health and Nutrition Examination Survey III (NAHNES III), 1988). After a short theoretical introduction and demonstration of the appropriate measurement methods, the doctors practiced the measurement method on each other (information about the measurement protocol is described in Appendix 1). The potential pit-falls and sources of systematic measurement errors were then discussed with the participants.

A week after the training session, the 12 doctors repeated the anthropometric measurements on the same 24 subjects, following exactly the same procedure as described above. This data collection scheme resulted in 2304 planned measures (4 measures × 12 doctors × 24 volunteers × 2 rounds).

Statistical analyses

Inter-observer reliability

We assessed the inter-observer variability by computing the technical error of measurement (TEM). TEM is the square root of measurement error variance, also called imprecision. It is obtained from replicate measurements on the same subjects taken within a short span of time by two or more observers (Moreno et al., 2003, Ulijaszek, 1994, Ulijaszek and Kerr, 1999, WHO Multicentre Growth Reference Study Group, 2006). The main sources of imprecision are random imperfections in the measuring instruments or in the measuring and recording techniques. In our study, lack of precision due to the measuring instruments was minimal, thus values of TEM provided information predominantly on measuring and/or recording errors. Values of TEM can be computed using a formula based on the difference between measurements and the number of individuals measured (Ulijaszek and Kerr, 1999). Due to the positive association between TEM and measurement size (large mean values of measurement are associated with high TEM and small ones with low TEM), it is pointless to compare TEMs directly. Instead, a measure of the coefficient of variation of TEM, the relative TEM or %TEM (TEM / mean × 100), is used to facilitate comparisons between different anthropometric measures or indices (Marks et al., 1989, Moreno et al., 2003, Ulijaszek and Kerr, 1999, WHO Multicentre Growth Reference Study Group, 2006). Another measure of measurement error is the coefficient of reliability ($R = 1 - (\mathrm{TEM}^2 / s_{\mathrm{is}}^2)$, where s_{is}^2 = total inter-subject variance), which reflects how much of the between-subject variance is free from measurement error. It ranges from 0 to 1. If $R = 0.9$, other factors than the measurement error are responsible for 90% of the total variance. By definition, inter-observer variability is excessive when $R < 0.95$ (Ulijaszek and Kerr, 1999).

For this study, we computed TEM, %TEM and R for all measurements and indices performed by the doctors (i.e. weight (kg), height (m), WC (cm), HC (cm), the derived BMI (kg/m^2) and WHR) before and after the one-hour specific training in anthropometrics.

Disagreement in diagnostic categorisation following anthropometrical measurements

Based on the measurements performed by each doctors, we assessed in which diagnostic category they would have classified each volunteer. The following categories were used: normal weight (18.5 kg/m^2 ≤ BMI < 25 kg/m^2), overweight (25 kg/m^2 ≤ BMI < 30 kg/m^2), obesity (≥ 30 kg/m^2) and abdominal obesity (WC ≥ 102 cm (men) and ≥ 88 cm (women) and/or WHR ≥ 0.95 (men) and ≥ 0.8 (women)) (WHO Technical Report, 2000). We then computed the proportion of disagreement between physicians as the number of measurements that led to a different classification of a subject compared to the majority, over the total number of measurements (see Appendix 2).

All statistical analyses were performed with SPSS (Statistical Package for Social Sciences, version 12.0) and Microsoft Excel version 9.0.

Results

Almost 100% of the planned measures (99.9%, 2302 out of 2304) had been performed; 2 doctors did not fill one measurement on their data collection sheet. Based on the measurements made by the doctors after training, the mean weight of the volunteers was 79.8 kg (range: 53.5–102.9, SD 14.0), the mean height 169.3 cm (range: 147.7–179.6, SD 9.4) and their mean BMI was 28.1 kg/m^2 (range: 19.9–39.6, SD 5.9). The mean WC was 91.4 cm (range: 65.1–117.7, SD 15.1), the mean HC 108.3 cm (range: 94.2–128.7, SD 9.0) and the mean WHR 0.84 (range: 0.67– 0.99, SD 0.09). Based on the BMI, 37.5% had a normal weight, 20.8% were overweight and 41.7% obese.

Reliability of anthropometrical measures

The inter-observer reliability prior to training was excellent for weight, height, and derived BMI, but unsatisfactory for WC, HC and WHR (Table 1). Based on the BMI determination, less than 1% (n/N 2/288) of measurements led to the volunteers being misclassified, whereas the use of WC alone and WHR lead to misclassification in 6% (16/288) and 23% (65/288) respectively. Reliability improved after the one-hour training, but a high variability for the HC measure persisted, leading to a poor inter-observer reliability for WHR. Percent disagreement in diagnostic categorisation of abdominal obesity also improved, but the proportion of volunteers who were misclassified remained high despite the training session (5%, 14/288 when based on WC vs. 9%, 26/288 when based on WHR). Overall misclassification for WC occurred as an underestimation of the risk category in most cases (67%, 20/30), whereas overestimation was more frequent for WHR (64%, 58/91).

Practitioners' knowledge and practice in anthropometrics

The participants reported assessing patients' BMI more frequently than WC or WHR (Table 2). Most doctors knew the correct BMI cut-off values to define normal weight, overweight and obesity, whereas only a minority reported the correct WC thresholds for abdominal obesity. None knew the correct cut-off values for the WHR. Similarly, most doctors reported correctly how to measure height and weight and calculate BMI, but most had poor knowledge of how WC and HC measurements should be performed.

Discussion

Our study confirms the reliability of height and weight measurements and BMI calculation by doctors for the assessment of obesity, whereas the reliability of WC and HC doctors' measures was unsatisfactory to accurately determine abdominal obesity. However the reliability of these measurements improved after a short one-hour training in anthropometrics.

As shown by other authors (Block et al., 2003), both the universality of height and weight measurements and the simpler, uncomplicated thresholds for BMI interpretation probably explain the high reliability of these measurements. Contrasting with these results, the reliability of WC and HC was poor, contradicting previous studies of the reproducibility of these measurements (Chen et al., 2001, Moreno et al., 2003, Nordhamn et al., 2000, Ulijaszek and Kerr, 1999, Wang et al., 2003). The lack of reliability for WC and HC measurements can be explained in several ways. First, the participating doctors were not familiar with these measurements. Their responses to the pre-training questionnaire indicated they had only poor knowledge of the appropriate measurement techniques and reference points for these measurements. Second, as different measurement sites and techniques are recommended in the scientific literature (Wang et al., 2003), this lack of standardisation can also contribute to the risk of measurement errors. Third, WC and WHR are also more recent concepts, with thresholds that vary with sex, ethnicity and authors (Wang, 2003), that are still rarely known and used by doctors. Finally, as WC and HC measurements require specific manipulation, this may also increase the risk for measurement errors.

Following a one-hour training in anthropometrics, the reliability of WC measurements improved, but not of HC. This is in line with previous studies, which showed greater reliability for WC than WHR (Chen et al., 2001, Nordhamn et al., 2000). This is also well in accordance with recent guidelines in which WC measurements are favoured over WHR determination because they are less complicated to perform (National Institute for Health and Clinical Excellence, 2006). Our one-hour training appeared to fill the gap in knowledge and practice for WC measurements, but did not seem sufficient to improve the performance of doctors to appropriately determine HC.

Finally the lack of inter-observer reproducibility of measurements in our study would be without clinical importance if the risk assessment of participants was not affected by these variations. Disagreement regarding the categorisation of patients, due to these variations, was important for WC (1 case for every 20 measurements) and very high for WHR (1 case for every 10 measurements).

Study limitations and strengths

This study has several limitations. First it was undertaken in an academic primary care clinic and involved only twelve physicians. The

Table 1
Inter-observer variability and reliability of weight, height, waist circumference (WC) and hip circumference (HC) measurements and of derived body mass index (BMI) and waist-to-hip ratio (WHR), before and after a one-hour training in anthropometrics, Geneva, Switzerland, 2006

	Mean (SD)		TEM		TEM%		R	
	Before	After	Before	After	Before	After	Before	After
Weight (kg)	80.5 (13.7)	79.8 (13.7)	0.33	0.40	0.41	0.50	0.999	0.999
Height (cm)	169.1 (9.3)	169.3 (9.3)	0.61	0.49	0.36	0.29	0.996	0.997
BMI (kg/m^2)	28.4 (5.7)	28.1 (5.8)	0.21	0.22	0.73	0.78	0.999	0.999
WC (cm)	93.2 (15.2)	91.4 (15.0)	4.33	2.58	4.65	2.82	0.915*	0.971
HC (cm)	105.4 (10.0)	108.3 (9.2)	4.93	2.61	4.68	2.41	0.702*	0.916*
WHR	0.88 (0.10)	0.84 (0.09)	0.06	0.03	6.88	3.58	0.506*	0.892*

TEM=technical error of measurement (square root of measurement error variance).
%TEM=relative technical error of measurement (TEM / mean × 100); this is a measure of the coefficient of variation of TEM, which allows direct comparison of different anthropometric measures.
R=coefficient of reliability (1−(TEM^2/s_{is}^2), where s_{is}^2 is the total inter-subject variance, including measurement error); it represents the proportion of the variance free of measurement error.
* Excessive variability (R<0.95).

Table 2
Knowledge and practice of the doctors in anthropometrics (N=12), Geneva, Switzerland, 2006

	n	(%)
Frequency of body mass index (BMI) assessment		
Daily	4	(33.3)
Weekly	6	(50.0)
Never or almost never	2	(16.7)
Frequency of waist circumference (WC) assessment		
Daily	0	(0.0)
Weekly	3	(25.0)
Never or almost never	9	(75.0)
Frequency of waist-to-hip ratio (WHR) assessment		
Daily	0	(0.0)
Weekly	2	(16.7)
Never or almost never	10	(83.3)
Knowledge of the appropriate measurement procedure and/or formula for		
Weight, height, and BMI	11	(91.7)
WC	1	(8.3)
WHR	0	(0)
Correct definition of overweight and obesity by physicians according to BMI[a]	10	(83.3)
Correct definition of abdominal obesity by physicians according to WC[b]	3	(25)
Correct definition of abdominal obesity by physicians according to WHR[c]	0	(0)

[a] BMI 25–29.9 kg/m^2 for pre-obesity (overweight) and ≥30 kg/m^2 for obesity.
[b] WC≥102 cm (men) and 88 cm (women) for abdominal obesity.
[c] WHR≥0.95 (men) and 0.8 (women) for abdominal obesity.

situation we describe may thus not entirely be transferable to primary care in general. Practical/logistic considerations dictated the choice of the number of practitioners involved. To avoid physiological variations, it was essential to perform all measurements on the same volunteer within a short time span. This would not technically have been possible had more doctors been involved. As the physicians had all recently been trained in an academic context, this may to some extent have reduced differences in the observed measurements. Our findings regarding inter-observer reliability may indeed be more conservative than is the case in the reality of primary care, but this assumption only further supports our conclusions. Due to their academic affiliation, the doctors were expected to be well aware of recent recommendations in relation to anthropometric measurements and indices. Analyses of the questionnaire showed that this was only true for BMI, but not for WC or WHR. A one-hour training in anthropometrics appeared to improve the reliability of doctors' measurements. In the absence of a control group, we were unable to exclude other reasons for this improvement (such as through increased practice or personal study, for example).

Despite these limitations, this study has many strengths. Indeed the study conditions were kept close to daily clinical practice and participating doctors had no previous training in anthropometrics. They only discovered the details of their task minutes before data collection began, thus limiting the risk for bias due to differential knowledge of the participating doctors. In addition, variability unrelated to measurement errors was largely reduced by having subjects stay in the same consulting-room in order to eliminate variability related to measuring instruments (i.e. the subject in a given room was always measured with the same measuring instruments). Measurements were taken within a short time span (maximum 45 min between the first and the 12th doctor) in order to eliminate physiological variations (also called undependability) (Marks et al., 1989, Mueller, 1998, Ulijaszek and Kerr, 1999). A large number of measurements and well validated anthropometric methods were used to calculate reliability. Finally, the study provided information on the possible impact of training, especially for WC and HC, on the reproducibility of the anthropometric measurements.

Conclusion

In this study, we have found that the inter-observer reliability for weight, height, and derived BMI were excellent ($R>0.99$), but unsatisfactory for WC ($R=0.92$), HC ($R=0.76$) and WHR ($R=0.51$). As emphasised in current guidelines, BMI determination remains the most reliable measure for the assessment of obesity, and should thus remain the cornerstone of anthropometric measurements in primary care (National Institute for Health and Clinical Excellence, 2006). A recent paper has also shown that BMI tends to perform even better than more sophisticated measures like fat-free mass or adipose tissue-free mass estimated by dual-energy X-ray absorptiometry (Heymsfield et al., 2007). Although WC and WHR have been proposed as additional measures of obesity-related health risks, we have found that they had a low inter-observer reliability, which could result in frequent diagnostic misclassification. Based on the BMI, only 1% of the volunteers were misclassified, whereas the use of WC and WHR lead to misclassification in 6% and 23% respectively. Reliability for the measurements improved after a one-hour training in anthropometric measurements ($R=0.97$ for WC, 0.92 for HC and 0.89 for WHR), but the proportion of volunteers who were misclassified remained high for WC (5%) and WHR (9%) despite the training session. The theoretical advantages of more sophisticated measurements for estimating an individual risk may be more than outweighed by the disadvantage of losing reliability in repeated measurements over time. These results challenge the current recommendations on obesity-related cardio-vascular risks management in primary care based on WC and WHR and underline the need for further research to improve the reliability of anthropometric measurements by doctors.

Conflict of interest statement

The authors have no conflict of interest to declare.

Acknowledgments

The authors thank all physicians and volunteers who participated to this study, Professors Hans Stalder, Alain Golay and Dr. Michel Boulvain for their thoughtful advices when designing the study, and the reviewers for their constructive comments.

Support: Institutional funds (Department of Community Medicine, Geneva University Hospitals, Prof. H. Stalder).

Appendix 1. Points emphasised during anthropometrics training of doctors, Geneva, Switzerland, 2006

Weight (mechanical beam balance, maximum weight 150 kg)
- The patient should empty his pockets, remove his shoes and all outer ware and then step on the centre of the balance, remaining in a relaxed vertical position
- Weight is recorded to the nearest 0.1 kg. 1 kg is subtracted from the measurement reading to account for the weight of garments worn

Standing height (stadiometer, maximum height 220 cm)
- The patient is dressed as for the weight measurement and stand erect on the floorboard of the stadiometer. The weight is evenly distributed on both feet, arms to the sides, shoulders relaxed. Shoulder blades, buttocks and heels slightly touch the measuring rod
- The patient is asked to look straight ahead (Frankfort Horizontal Plane position of head), inhale deeply and stand fully erect while the examiner lowers the horizontal bar to the crown of the head and takes the measure to the nearest 0.5 cm

Waist circumference (standard tailor measuring tape, maximum length 150 cm)
- The patient is dressed as for the weight measurement and is standing erect. He/she is asked to roll up the shirt/sweater, to undo the belt and/or open and lower the trouser/skirt waistband, so the examiner can palpate the hip area to identify the measurement reference points
- The measure is taken at the midpoint between the lowest rib and the iliac crest. The measuring tape is placed perpendicular to the long axis of the body and horizontal to the floor, with sufficient tension to avoid slipping off but without compressing the skin
- The measurement is made at the end of a normal expiration, twice to the nearest 0.1 cm. If the difference between the two recorded measurements is >0.5 cm, a third measurement is taken, and the mean of the two nearest values is recorded

Hip circumference (standard tailor measuring tape, maximum length 150 cm)
- The patient is dressed as for the weight measurement. The measure is taken with the pants/skirt on
- The patient stands erect, the weight evenly distributed on both feet
- The tape is placed at the maximum extension of the buttocks, horizontal to the floor, with sufficient tension to avoid slipping off. With thick garments, the tape is held a bit tighter but without compressing the buttocks
- The measurement is made twice to the nearest 0.1 cm. If the difference between the two recorded measurements is>0.5 cm, a third measurement is taken, and the mean of the two nearest values is recorded
- Cut-off values (according to: WHO Technical Report, 2000. Obesity: preventing and managing a global epidemic—Report of a WHO consultation: WHO.)
- Body Mass Index (BMI)
  - normal weight $18.5 \le BMI < 25$ kg/m^2
  - overweight $25 \le BMI < 30$ kg/m^2
  - obesity $BMI \ge 30$ kg/m^2
- Waist circumference (WC) and waist-to-hip ratio (WHR)
  - abdominal obesity in men: $WC \ge 102$ cm and/or $WHR \ge 0.95$
  - abdominal obesity in women: $WC \ge 88$ cm and/or $WHR \ge 0.8$

Appendix 2. Example illustrating how the proportion of disagreement was calculated

Example: 12 doctors have measured height and weight in 2 volunteers A and B:
- Measurements in volunteer A: on the basis of their weight and height measurements, 10 doctors find a $BMI \ge 30$ kg/m^2, whereas 2 doctors find a BMI between 25 and 29.9 kg/m^2
- Measurements in volunteer B: on the basis of their weight and height measurements, 8 doctors find a BMI between 25 and 29.9 kg/m^2 and 4 doctors find a $BMI < 25$ kg/m^2
- 2 doctors' measurements disagree for A, 4 disagree for B: 2+4=6 disagreeing measurements of a total of 24 measurements taken=25% disagreement

References

Block, J.P., DeSalvo, K.B., Fisher, W.P., 2003. Are physicians equipped to address the obesity epidemic? Knowledge and attitudes of internal medicine residents. Prev. Med. 36, 669–675.

Booth, M.L., Hunter, C., Gore, C.J., Bauman, A., Owen, N., 2000. The relationship between body mass index and waist circumference: implications for estimates of the population prevalence of overweight. Int. J. Obes. Relat. Metab. Disord. 24, 1058–1061.

Chen, M.M., Lear, S.A., Gao, M., Frohlich, J.J., Birmingham, C.L., 2001. Intraobserver and interobserver reliability of waist circumference and the waist-to-hip ratio. Obes. Res. 9, 651.

Gill, T., Chittleborough, C., Taylor, A., Ruffin, R., Wilson, D., Phillips, P., 2003. Body mass index, waist hip ratio, and waist circumference: which measure to classify obesity? Soz Praventivmed 48, 191–200.

Health Canada, 1995. National Nutrition Survey, PEI Nutrition Survey, Interviewer Training Manual Health Canada Publications Centre, Ottawa, Canada.

Health Canada, 2003. Canadian Guidelines for Body Weight Classification in Adults. Health Canada Publications Centre, Ottawa, Canada.

Heymsfield, S.B., Gallagher, D., Mayer, L., Beetsch, J., Pietrobelli, A., 2007. Scaling of human body composition to stature: new insights into body mass index. Am. J. Clin. Nutr. 86, 82–91.

International Tack Force for Prevention of CHD, 1998. Nutrition, Metabolism and Cardiovascular Diseases—Coronary Heart Disease: Reducing the Risk: The Scientific Background for Primary and Secondary Prevention of Coronary Heart Disease—A Worldwide View.

Janssen, I., Katzmarzyk, P.T., Ross, R., 2004. Waist circumference and not body mass index explains obesity-related health risk. Am. J. Clin. Nutr. 79, 379–384.

Kanaya, A.M., Vittinghoff, E., Shlipak, M.G., et al., 2003. Association of total and central obesity with mortality in postmenopausal women with coronary heart disease. Am. J. Epidemiol. 158, 1161–1170.

Kushner, R.F., Blatner, D.J., 2005. Risk assessment of the overweight and obese patient. J. Am. Diet Assoc. 105, S53–S62.

Lean, M.E., Han, T.S., 1996. Predicting body composition by densitometry from simple anthropometric measurements. Am. J. Clin. Nutr. 63, 4–14.

Marks, G.C., Habicht, J.P., Mueller, W.H., 1989. Reliability, dependability, and precision of anthropometric measurements. The Second National Health and Nutrition Examination Survey 1976–1980. Am. J. Epidemiol. 130, 578–587.

Morabia, A., Costanza, M.C., 2005. The obesity epidemic as harbinger of a metabolic disorder epidemic: trends in overweight, hypercholesterolemia, and diabetes treatment in Geneva, Switzerland, 1993–2003. Am. J. Public. Health 95, 632–635.

Moreno, L.A., Joyanes, M., Mesana, M.I., et al., 2003. Harmonization of anthropometric measurements for a multicenter nutrition survey in Spanish adolescents. Nutrition 19, 481–486.

Mueller, M., 1998. Reliability and accuracy of measurement. In: Lohman, T.G., Roche, A.F., Martorell, R. (Eds.), Anthropometric Standardisation Reference Manual. Human Kinetics Books, Champaign, pp. 83–86.

National Health and Nutrition Examination Survey III (NAHNES III), 1988. Body Measurements (Anthropometry). Westat Inc., 1659 Research Boulevard, Rockville, MD 20859.

National Institute for Health and Clinical Excellence, 2006. Obesity: Guidance on the Prevention, Identification, Assessment and Management of Overweight and Obesity in Adult and Children. NICE Clinical Guideline 43.

Nordhamn, K., Sodergren, E., Olsson, E., Karlstrom, B., Vessby, B., Berglund, L., 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.

Okosun, I.S., Chandra, K.M., Boev, A., et al., 2004. Abdominal adiposity in U.S. adults: prevalence and trends, 1960–2000. Prev. Med. 39, 197–206.

Paccaud, F., Schluter-Fasmeyer, V., Wietlisbach, V., Bovet, P., 2000. Dyslipidemia and abdominal obesity: an assessment in three general populations. J. Clin. Epidemiol. 53, 393–400.

Snijder, M.B., Zimmet, P.Z., Visser, M., Dekker, J.M., Seidell, J.C., Shaw, J.E., 2004. Independent and opposite associations of waist and hip circumferences with diabetes, hypertension and dyslipidemia: the AusDiab Study. Int. J. Obes. Relat. Metab. Disord. 28, 402–409.

Ulijaszek, S.J., 1994. Intra- and inter-observer error in anthropometric measurement. In: Ulijaszek, S.J., Mascie-Taylor, C.G.N. (Eds.), Anthropometry: the Individual and the Population. Cambridge University Press, Cambridge, pp. 30–55.

Ulijaszek, S.J., Kerr, D.A., 1999. Anthropometric measurement error and the assessment of nutritional status. Br. J. Nutr. 82, 165–177.

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.

Wang, J., Thornton, J.C., Bari, S., et al., 2003. Comparisons of waist circumferences measured at 4 sites. Am. J. Clin. Nutr. 77, 379–384.

WHO fact sheet 311, 2006. Obesity and Overweight. WHO.

WHO Multicentre Growth Reference Study Group, 2006. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. Acta Paediatrica 450, 38–46.

WHO Technical Report, 2000. Obesity: Preventing and Managing a Global Epidemic—Report of a WHO Consultation. WHO.

Zamboni, M., Turcato, E., Armellini, F., et al., 1998. Sagittal abdominal diameter as a practical predictor of visceral fat. Int. J. Obes. Relat. Metab. Disord. 22, 655–660.