



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.

© 2008 Elsevier Inc. All rights reserved.

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; Paccoud 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.

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 $\times$ 12 doctors $\times$ 24 volunteers $\times$ 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 ($\text{TEM}/\text{mean} \times 100$), 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 ($R = 1 - (\text{TEM}^2/s_{\text{is}}^2)$, 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 ($18.5 \text{ kg}/\text{m}^2 \leq \text{BMI} < 25 \text{ kg}/\text{m}^2$), overweight ($25 \text{ kg}/\text{m}^2 \leq \text{BMI} < 30 \text{ kg}/\text{m}^2$), obesity ($\geq 30 \text{ kg}/\text{m}^2$) and abdominal obesity ($\text{WC} \geq 102 \text{ cm}$ (men) and $\geq 88 \text{ cm}$ (women) and/or $\text{WHR} \geq 0.95$ (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,

Download English Version:

<https://daneshyari.com/en/article/3101482>

Download Persian Version:

<https://daneshyari.com/article/3101482>

[Daneshyari.com](https://daneshyari.com)