Author’s response to reviews

Title: Accuracy of anthropometric measurements by general practitioners in overweight and obese patients

Authors:

Paul Sebo (paulsebo@hotmail.com)

Dagmar Haller (dagmar.haller-hester@hcuge.ch)

François Herrmann (francois.herrmann@hcuge.ch)

Version: 2 Date: 18 May 2017

Author’s response to reviews:

Reviewer reports:

Jennifer L Kuk, PhD (Reviewer 1): Calling height, weight and bmi classical and waist, hip and WHR 'newer' is somewhat confusing as none of these measures are really new. I would suggest overall and regional obesity measures or do not distinguish groups at all.

We changed “classical and newer anthropometric measures” to “general and abdominal obesity anthropometric measures” or to “general and abdominal obesity measures”.

The percent difference values are tough to interpret, particularly as there are clear sex differences. Adjusted mean differences in cm should be given instead.

Instead of absolute measurement differences, we used percent differences, i.e. relative differences (by dividing the absolute differences between the GPs’ measurements and the gold standard (=average values of the measurements undertaken by the two research assistants)). They allowed us to make comparisons between BMI subgroups and other anthropometric measurements, as relative values do not depend on the magnitude of these measurements.

Indeed, 1kg measurement error for a child weighing 10kg is not equivalent to 1kg measurement error for an obese adult weighing 100kg (10% error for the child, 1% error for the obese adult). In addition, it would be tricky to compare 1kg weight with 1cm height measurement error.
However, as suggested by the reviewer, we propose to add the absolute measurement differences as supplementary material. We changed the following sentence “overall, the mean relative measurement differences were not associated with BMI subgroups…” into “overall, the mean relative measurement differences, computed from the absolute measurement differences (see appendix#2), were not associated with BMI subgroups…”

Note that we did not record the age and gender of the volunteers and thus we cannot provide adjusted mean differences. Assessing the role of gender on measurement error of anthropometric measurements could be an interesting topic for a new study. We added this point as a limitation of the study in the discussion section: “Fourth, we did not record the age and gender of the volunteers, so we cannot provide adjusted mean differences.”

Also, the more important thing to know would be if the error caused misclassification and thus improper clinical decision making as a result.

In the current study, we did not record the age and gender of the volunteers. Therefore, we cannot compute the % of misclassification, as the definition of abdominal obesity depends on the gender. However, in a previous study (ref: Sebo P, Beer-Borst S, Haller DM, Bovier PA. Reliability of doctors’ anthropometric measurements to detect obesity. Prev Med. 2008;47: 389–393), we showed that only 1% of the volunteers were misclassified when the measurements were based on the BMI, compared to 6% when using WC measurement, and 23% when using WHR determination.

We slightly modified the introduction section: “we previously showed in two studies conducted in primary care settings that, in contrast to BMI, obesity was not accurately detected by these abdominal obesity measurement methods and that these measures led to frequent diagnostic misclassification”.

The clinical utility of the study is unclear given that waist hip and whr are not used to dictate weight management.

In fact, as stated in the introduction section, recent guidelines on the management of obesity emphasise measuring abdominal in addition to general obesity, because the pattern of fat
distribution has been shown to have a large influence on cardiometabolic risk and, as a consequence, abdominal obesity seems to predict the development of cardiovascular diseases better than overall obesity. Thus, overweight patients with abdominal obesity should lose weight, even when they are not particularly at risk according to BMI.

The waist circumference (WC) and the waist-to-hip ratio (WHR), which is determined by dividing the WC by hip circumference (HC), have become widely accepted measures for assessing abdominal obesity [see for example ref. 9 and 10 in our paper].

We think that our study is interesting, because it suggests that the anthropometric measurements of abdominal obesity are particularly inaccurate when GPs use these to assess overweight and obese patients. Therefore, as stated in our conclusion, we recommend that GPs essentially use BMI determination in daily practice, particularly when assessing overweight or obese patients, though recent guidelines emphasize measuring WC, HC and/or WHC.

Were the physicians and techs told to measure at the same sites (i.e. Minimal waist versus iliac crest)

The GPs were asked to perform the measurements as usual, within the consultation. Thus, the study was undertaken within the normal conditions of day-to-day clinical practice and involved GPs with no particular training in anthropometric measurements.

By contrast, the research assistants have been trained prior to the start of the study. A specialist in anthropometric measurements provided a theoretical and practical training to the two research assistants. The training was based on international guidelines. The document in appendix explains in detail the points emphasized during this training (see appendix#1 for the sites used for measuring WC and HC). The two research assistants used the same procedure (including the same sites for measuring WC and HC).

Why are there significant errors in the measurement of body weight within each BMI group. That would suggest a systematic difference in how weight is measured. This should be discussed.
The reviewer is right. The results show statistically significant results regarding the mean relative differences between GPs’ weight measurement and the gold standard. However, these differences are clinically non relevant (0.36% error for normal group (absolute difference: 0.22kg), 0.44% for overweight group (absolute difference: 0.30kg) and 0.37% for obese group (absolute difference: 0.34kg)).

We completed the results section: “For weight, although the differences were small, they were statistically significant (0.36% measurement error for the normal group (absolute difference: 0.22kg), 0.44% for the overweight group (absolute difference: 0.30kg) and 0.37% for the obese group (absolute difference: 0.34kg))”.

We also completed the discussion section: “We showed statistically significant results regarding the mean relative differences between GPs’ weight measurement and the gold standard. However, these differences are not clinically relevant (0.36% measurement error for the normal group (absolute difference: 0.22kg), 0.44% for the overweight group (absolute difference: 0.30kg) and 0.37% for the obese group (absolute difference: 0.34kg)).

All analyses should be stratified by sex.

As stated above, we did not record the gender of the volunteers. We completed the discussion section: “We did not record the gender of the volunteers. Therefore, we cannot compute the % of misclassification, as the definition of abdominal obesity differs by gender. However, in a previous study (ref: Sebo P, Beer-Borst S, Haller DM, Bovier PA. Reliability of doctors’ anthropometric measurements to detect obesity. Prev Med. 2008;47: 389–393), we showed that only 1% of the volunteers were misclassified when the measurements were based on the BMI, compared to 6% when using WC measurement, and 23% when using WHR determination”.

In addition, we completed the limitations section: “Fourth, we did not record the age and gender of the volunteers, so we cannot provide adjusted mean differences”.

Masaharu Kagawa (Reviewer 2): This study aimed to determine measurement error of general practitioners (GPs) in their measurements of various anthropometric variables such as height, weight, waist and hip circumferences. The authors requested 26 GPs to measure ten adult
patients in two occasions and compared the results with the values measured by two trained research assistants. From the results, the authors concluded that weight, height, and hip circumference measurements were more accurate than waist circumference measurement. Also the body mass index (BMI) was found to be more accurate compared to waist-to-hip ratio (WHR). In addition, measurement error was increased as increase in BMI of patients.

Understanding measurement error in clinical practice is an important topic and therefore the present study has a significant importance. However, the present study has a number of considerations and insufficient information in their study design and methodology. Most important consideration is uncertainty of acceptable measurement error. While previous studies and existing protocols indicated certain criteria for acceptable measurement errors such as technical error of measurement (TEM), percentage technical error of measurement (%TEM) or intra-class correlation coefficient (ICC) (Stewart et al. 2012; Wang et al. 2000; Ulijaszek and Kerr, 1999), the present study have not indicate such particular cut-off point for acceptable measurement errors. In fact, the authors did not calculate both inter- and intra-TEM of each GP.

We used TEM, %TEM and coefficient of reliability (R) in another study (ref: Sebo P, Beer-Borst S, Haller DM, Bovier PA. Reliability of doctors’ anthropometric measurements to detect obesity. Prev Med. 2008;47: 389–393) where we aimed to evaluate the reliability of anthropometric measurements. In that study 12 doctors were asked to measure 24 volunteers. Measurement errors could not be calculated, because there were no gold standards. Instead of measurement errors, we used TEM, %TEM and R.

By contrast, we believe that TEM, %TEM and ICC should probably not be used in the current study, because each volunteer was measured only once and only by one GP, and this value was compared to a gold standard (the measurements of the two research assistants, with the mean of these two measures being considered as the “real” exact value).

TEM is used when the exact value is unknown and we aim at calculating the imprecision of repeated measurements. Indeed, TEM, that is the square root of measurement error variance, is “obtained by carrying out a number of repeat measurements on the same subject, either by the same observer, or by two or more observers, taking the differences and entering them into an appropriate equation” (ref: Ulijaszek and Kerr, 1999).
As a measurement difference of <3% is unlikely to be clinically relevant (ref: Verweij 2013), we used this cut-off value to consider measurement errors as being acceptable or not. Though this point was explained in the results section, it is true that we did not define accurate measurements in the methods section. We completed the methods section: “As a measurement difference of <3% is unlikely to be clinically relevant (ref: Verweij 2013), we decided to consider that the measurements were accurate when the relative measurement errors were <3%”.

Rather than calculating TEM of each GP, the authors aimed to compare the results measured by GPs and those from two trained research assistants. However, again the authors have not provided TEM of each research assistant on the manuscript. In addition, there is no description about how much experience each research assistant had at the time of generating this paper.

Please see above for the first part of the reviewer’s comment.

As stated in our paper, the research assistants were already skilled before the training, because they had previously been trained in anthropometrics and were asked to regularly take anthropometric measurements on many individuals living in Geneva as part of the Bus Santé study (an annual population-based survey collecting data on cardiovascular risk factors in Geneva).

In addition, a specialist in anthropometric measurements provided a theoretical and practical training to the two research assistants. The training was based on international guidelines (we used the guidelines from NAHNES III, PEI nutrition survey and the Canadian guidelines for body weight classification). The document in the appendix explains in detail the points emphasized during this training (see appendix#1).

However, we agree with this reviewer that it could be of importance to readers to know more about the inter-observer variability between the two research assistants. We thus included in the results section the inter-observer variability of the two research assistants’ measures by computing the technical error of measurement (TEM) and %TEM for each anthropometric measurement.

We also completed the methods section: “First, we assessed the inter-observer variability between the two research assistants by computing the technical error of measurement (TEM) for each anthropometric measurement. We also computed %TEM (TEM / mean x 100), a measure of
the coefficient of variation of TEM, because it is difficult to compare TEMs directly, due to the positive association between TEM and measurement size”.

Furthermore, as two trained research investigators involved in the study, the present study should include their inter-tester TEM in order to clarify that these two research assistants have comparable measurement skill.

As stated in the results section, “the mean differences between the research assistants was very small (for weight: 0.002 kg (SD 0.09); for height: 0.03 cm (SD 0.15); for WC: 0.02 cm (SD 0.14); for HC: 0.01 cm (SD 0.06))”.

As suggested by the reviewer, we included in the results section the inter-observer variability of the two research assistants’ measures by computing the technical error of measurement (TEM) and %TEM for each anthropometric measurement.

In addition to no information on skill level of research assistants whose results were utilized as criterion, the authors did not provide any information on experience of anthropometric measurements by GPs and their frequency of measurements.

The GPs were asked to perform the measurements as usual, within the consultation. Thus, the study was undertaken within the normal conditions of day-to-day clinical practice and involved GPs with no particular training in anthropometric measurement at the start of the study.

However, as stated in the methods section, “in the original study doctors were randomly assigned to two separate groups for the second measurement session. The intervention group received special training in anthropometric measures, the other acted as control. Since the intervention did not appear to be associated with a significant improvement in GPs’ measurement accuracy, measurements from both groups and both sessions were pooled for the present analysis”.

We completed the methods section: “...the intervention group received training in anthropometric measures (the doctors received a training document, prepared by the authors, explaining the
appropriate measurement methods according to international recommendations, the other acted as control.”

GPs’ knowledge and self reported practice in anthropometrics (including frequency of measurement) were presented and discussed in our previous paper (Sebo, Swiss Med Wkly 2015). In the discussion section, it is stated: “We recently reported that, in contrast to general obesity measurements, GPs’ perceived skills were relatively low for WC and HC measurements, as well as for WHC assessment and interpretation”.

We slightly modified this paragraph: “We discussed the role of GPs’ knowledge and their usual practice in anthropometrics in our previous paper (Sebo, SMW 2015). We showed that, compared to weight, height and BMI, a majority of GPs hardly ever used the abdominal obesity measures and their knowledge regarding these measurements was relatively low”.

In addition to no information on anthropometric skill and calculation of TEM, the authors failed to provide whether they have standardized measurement protocols, including landmarking protocol, equipment used, the method of measurements.

The GPs were asked to perform the measurements as usual, with their own material. By contrast, the measurements taken by the two research assistants were standardised.

We completed the methods section: “the training of the research assistants was based on international guidelines (see appendix#1). They were asked to take the measurements with a calibrated flat beam scale for mobile use (SECA 877, scale division: 100g, capacity: 200kg), a stadiometer (SECA 217, graduation length: 1cm, range: 20-205cm) and a measuring tape. After having been measured by the GP in his/her consultation room, and while the GP was taking care of the next patient, each volunteer was measured in turn by the two research assistants in a quiet room, close to the consultation room. The research assistants took the measurements according to the recommended procedure for which they had been trained.”

While the authors stated that they undertook training session to a group of GPs base on an international guideline, they failed to clearly indicate which protocol they followed (e.g. the WHO protocol, the ISAK protocol, the IAEA protocol).
Three references have been added. We used the guidelines from NAHNES III, PEI nutrition survey and the Canadian guidelines for body weight classification.

In addition, there is insufficient information how the training course was conducted to the intervention group (e.g. skill level of the instructor, a number of GPs attended the session at once, an hour of hand-on session, presence of practical exam). As a result, although the authors stated that the intervention did not appear to be associated with a significant improvement in GPs's measurement accuracy, readers cannot justify if the intervention was appropriately conducted.

The instructor of the research assistants was a senior attending physician at Geneva University Hospitals, in the department of therapeutic education for obesity and chronic diseases.

The GPs in the intervention group received a training document, prepared by the authors, explaining the appropriate measurement methods according to international recommendations. We completed the methods section, adding this information.

Furthermore, few minor points need to be clarified:

- Waist and hip circumferences and subsequent waist-to-hip ratio (WHR) are not very "new" anthropometric variables. The authors should clarify their intention of calling these anthropometric variables as "new".

  We changed “classical and newer anthropometric measures” to “general and abdominal obesity anthropometric measures” or to “‘general and abdominal obesity measures”.

- While the authors described that every GPs have measured 20 volunteers, the text in the manuscript stated that measured only 10 volunteers but at two occasions. Measuring 20 different volunteers and conducting repeated measures on 10 volunteers are different and therefore strongly request the authors to correct the sentence.
The GPs took measurements on 10 volunteers in the first session, and then they took measurements on 10 other volunteers in the second session. Therefore, each GP took measurements on 20 volunteers.

We changed the following sentence “the GPs were asked to recruit ten adult volunteers among their patients, for each of two pre-planned measurement sessions” to “the GPs were asked to recruit twenty adult volunteers among their patients, ten for each of two pre-planned measurement sessions”.

- It is unclear where the authors obtained ethical approval from. Since the present study involved a multiple sites to conduct a data collection, it is anticipating that ethical approval may need to be obtained from each hospital/clinics as well as the affiliated institutions of the authors.

All participating community-based GPs were practicing in the canton of Geneva, Switzerland. We obtained ethical approval by the Geneva Research Ethics Committee (Commission d’éthique pour la recherche clinique en ambulatoire): protocol 11-04, which is valid for the entire canton.

If improvements to the English language within your manuscript have been requested, you should have your manuscript reviewed by someone who is fluent in English. If you would like professional help in revising this manuscript, you can use any reputable English language editing service. We can recommend our affiliates Nature Research Editing Service (http://bit.ly/NRES_BS) and American Journal Experts (http://bit.ly/AJE_BS) for help with English usage. Please note that use of an editing service is neither a requirement nor a guarantee of publication. Free assistance is available from our English language tutorial (https://www.springer.com/gb/authors-editors/authorandreviewertutorials/writinginenglish) and our Writing resources (http://www.biomedcentral.com/getpublished/writing-resources). These cover common mistakes that occur when writing in English.
Editorial Policies

Please read the following information and revise your manuscript as necessary. If your manuscript does not adhere to our editorial requirements this will cause a delay whilst the issue is addressed. Failure to adhere to our policies may result in rejection of your manuscript.

In accordance with BioMed Central editorial policies and formatting guidelines, all submissions to BMC Obesity must have a Declarations section which includes