Established in 1871

Swiss Medical Weekly

Formerly: Schweizerische Medizinische Wochenschrift
An open access, online journal • www.smw.ch

Original article | Published 21 February 2015, doi:10.4414/smw.2015.14115
Cite this as: Swiss Med Wkly. 2015;145:w14115

Accuracy of doctors' anthropometric measurements in general practice

Paul Sebo[a], Dagmar M. Haller[a,b,c], Antoinette Pechère-Bertschi[b,d], Patrick Bovier[e], François R. Herrmann[f]

[a] Primary Care Unit, Faculty of Medicine, University of Geneva, Switzerland
[b] Department of Community, Primary Care and Emergency Medicine, Geneva University Hospitals, Switzerland
[c] Department of Paediatrics, Geneva University Hospitals, Switzerland
[d] Medical Policlinic and Hypertension Unit, Division of Endocrinology, Diabetology, Hypertension and Nutrition, Geneva University Hospitals, Switzerland
[e] Primary care practice, Lausanne, Switzerland
[f] Department of Internal Medicine, Rehabilitation and Geriatrics, Geneva University Hospitals, Switzerland

Summary

PURPOSE: There is increasing pressure on general practitioners (GPs) to identify patients with abdominal obesity in order to reduce the life-threatening consequences of this condition in the population. We aimed to confirm previous findings on the inaccuracy of anthropometric measurements performed by GPs in an academic primary care clinic and to assess the effect of theoretical training to improve the quality of these measurements.
METHODS: This cross-sectional study involved 26 GPs from private practices in Geneva, Switzerland. They were asked to measure weight, height, waist and hip circumference on ten volunteers within their practice. Two trained research assistants repeated the measurementss after the GPs ("gold standard"). The GPs were then randomised to receive information detailing the correct method for taking measurements (intervention, 14 doctors) or simple information about obesity (control, 12 doctors). Measurements were repeated a few weeks later. Measurement error was computed by comparing the GPs' values with the average value of two measurements taken in turn by the research assistants, and agreement was examined by Bland-Altman plots. The GPs' skills were assessed through auto-questionnaire and direct observation.
RESULTS: All measurements except height were prone to measurement error, the least affected being weight (and therefore body mass index [BMI]). Following training, measurement errors were slightly less prominent in the intervention group. GPs' skills in measuring waist and hip circumference were frequently assessed as inadequate, but showed improvement after training.
CONCLUSIONS: Without proper training, priority should be given to using classical anthropometric measurements (i.e. weight, height and BMI determination) in daily practice.

Key words: *accuracy; anthropometric measurements; primary care*

Introduction

Worldwide prevalence of obesity is high and increases continually. It reached epidemic proportions in 2008, when 9.8% of men and 13.8% of women were obese in the world, compared with 4.8% and 7.9% in 1980 [1]. Estimates indicate more than 2 billion individuals in the world were overweight or obese in 2013 [2]. In Switzerland, the same trend is noted with 6.1% of men and 4.7% of women who were obese in 1992/1993 compared with 11.2% and 9.4% in 2012, and this figure is evenly distributed throughout the country [3–5]. This progression is alarming, since obesity is associated with serious health consequences such as cardiovascular diseases and type 2 diabetes, and is a leading cause of global burden of disease with an estimated 35.8 million (2.3%) of global disability-adjusted life years (DALYs) caused by this condition [6]. In addition, relative to normal weight, obesity is associated with significantly higher mortality and it is estimated that, worldwide, at least 2.8 million people die each year as a result of obesity [6, 7]. General practitioners (GPs) play a central role in the assessment and management of this condition and its life-threatening consequences. Recent guidelines on its management emphasise measuring abdominal as well as general obesity [8, 9], because abdominal obesity appears to be a more important risk factor than overall obesity in predicting the development of type 2 diabetes and cardiovascular diseases. Therefore, overweight patients with abdominal obesity should be urged to pursue weight reduction, even when they are not particularly at risk according to body mass index (BMI, weight in kg divided by height in squared meters [kg/m^2]) alone [10–12].

The waist circumference (WC) and the waist-to-hip ratio (WHR, WC divided by hip circumference, HC) are reliable measures for measuring abdominal adiposity [13, 14]. Anthropometric studies have shown that these newer measurement methods presented excellent intra- and inter-observer reproducibility when performed by health professionals who had been trained in anthropometrics [14–17].

However, according to a recent study, these newer measurement methods were unreliable in detectecting obesity when performed by GPs [18]. This study, however, involved only GPs working in an academic primary care clinic, and reliability was assessed using inter-observer variability. It did not involve computing measurement error, i.e. comparing doctors' measurements with those performed by trained measurers.

Our primary objective was thus to confirm previous findings from a study conducted in an academic primary care clinic on the inaccuracy of anthropometric measurements performed by GPs and to assess the effect of a theoretical training to improve the quality of these measurements. We also aimed to assess GPs' measurement knowledge and practice.

Materials and methods

Recruitment of doctors and volunteers, and randomisation

The study took place in primary care practices in Geneva, Switzerland, from September to December 2011. It was designed in two parts: blood pressure [19] and anthropometric measurements. We report here only the latter. From a convenience sample of 84 doctors who were personally invited to participate, 26 agreed (31%), slightly fewer than our target sample of 30 (see sample size estimate below). Two data collection sessions were planned. For each session GPs had to recruit ten volunteers among their patients (age >18 years, with no other selection criteria) and their next appointment was synchronised with the study visit (during which they first signed the informed consent and then participated in the study). Between the two sessions the GPs were randomised into two groups, by computer-generated random numbers and block randomisation. The first group received a training document, prepared by the authors, explaining the appropriate measurement methods according to international recommendations (intervention group, group I) [20–24]. The other group received a document which had an identical format, but only introduced the concept of obesity and its consequences (control group, group C). We chose a random allocation in order to differentiate progress related to repeat measurements throughout the study from improved technique as a result of exposure to training material. A statistician independent of the study carried out the randomisation process and sent the documents to GPs. The outcomes were triple blinded (participants, doctors and research assistants).

To take part in the study, both doctors and volunteers had to read the information sheet and sign consent forms. The research protocol was aproved by the local research ethics committee.

Data collection and training for anthropometric measurements

Two research assistants participated in theoretical and practical training sessions given by a specialist in anthropometric measurements (two 1 hour-training sessions). The training was based on international recommendations (see appendix 1) [20–24]. The research assistants were provided with a calibrated flat beam scale for mobile use (SECA 877; scale division: 100 g, capacity: 200 kg) combined with a stadiometer (SECA 217; graduation length: 1 cm, measuring range: 20–205 cm) and two measuring tapes. The 26 GPs were told that the measurements had to be performed as usual in their consultation room. After having visited the doctor, the volunteer was directed into a quiet room, close to the consultation room. While the GP took care of the next patient, the research assistants entered the room and each of them made the measurements in turn, according to the recommended procedure for which they had been trained (see appendix 1). A few days after this first session, the GPs received the training documents by post. A few weeks later, the GPs were asked to set-up the second round of measurements using the same procedure as described above.

Gold standard

The average value of the research assistants' measurements was considered as the gold standard.

Assessment of the GPs measurement technique and practice

For the first participant of each of the two sessions, the two research assistants went into the consultation room with the doctor and observed the doctors' measurement technique. They noted any pitfalls using a formal checklist; doubts and disagreements about the assessment were resolved by discussion and consensus between the research assistants. At the end of the first session, the doctors were asked to complete a questionnaire asking how frequently and in which manner they perform the measurements in their daily practice.

Sample size justification and statistical analysis

A sample size of 30 GPs allowed the simultaneous analysis of three independent variables (a group effect, i.e. comparison between two training groups; a time effect, i.e. comparison before and after intervention; and a doctor effect, i.e. comparison between the doctors) when dealing with multiple variables statistical models, as ten subjects by group are needed for each variable included in the models according to Harrell [25].

We asked the doctors to recruit only ten volunteers, since we were interested in the inter-doctor and not the intra-doctor variability.

Frequencies were used to describe the doctors' characteristics as well as their knowledge and practice. For each volunteer, we computed the difference between the doctor's measurements and the gold standard (measurement error). After checking the assumption of normality of the distribution using the Shapiro-Francia test [26], we performed unpaired t-tests to compare the doctors' and the research assistants' mean differences in measurements before and after training, and between groups. Since these values depend on the magnitude of the measurements, we also computed the relative measurement differences, in order to compare the anthropometric measurements by dividing the absolute difference by the average value of the research assistants' measurements [27]. The extent of agreement between the doctors' and the research assistants' measurements was then examined by plotting the differences between the pairs of measurements on the vertical axis, against the mean of each pair on the horizontal axis (Bland-Altman plots) [28]. Note that for data not showing a normal distribution, transformations were undertaken when appropriate.

By definition, the measurements were considered accurate when the relative measurement errors (computed by dividing the absolute errors by the average value of the research assistants' measurements) were <1% and the 95% limits of agreement (given by the mean difference between the research assistants' average value and the doctors' measurements, plus or minus twice the standard deviation of the differences) were within 2% of the research assistants' average value (example: if the average weight measured by the research assistants was 100 kg, the 95% of the doctors' values should be within 100 kg plus and minus 2 kg).

Statistical significance was set at a two-sided p-value ≤0.05. All statistical analyses were performed with STATA version 12.1.

Results

The 26 GPs who agreed to participate in the study had a mean age of 44.1 years (standard deviation [SD] 6.1, range 33–59) and 58% were women. Most of them were certified (general internal medicine 96%, no certification 4%) and relatively experienced doctors (years since certification 16.3 [SD 5.8], range 7–32).

Anthropometric measurements were performed on 259 volunteers at baseline and 250 after training. Based on the measurements made by the research assistants, their mean weight was 72.7 kg (SD 17.1) and their mean BMI 26.3 kg/m^2 (SD 5.4). The mean measurement differences between the two research assistants was very low (weight: 0.002 kg [SD 0.09]; height: 0.03 cm [SD 0.15]; WC: 0.02 cm [SD 0.14]; HC: 0.01 cm [SD 0.06]).

Table 1 shows the doctors‘ mean measurement error before and after training, first without taking into account the variance of the differences between the research assistants' and the doctors' values. The absolute measurement differences were very small for weight, height, BMI and WHR, and slightly greater for WC and HC.

We also computed the relative measurement differences: height was the most accurate measurement (before training: 0.06%, after training: 0.01%), and weight (0.19% and 0.55%), BMI (0.11% and 0.54%) and HC (0.92% and 0.91%) were also accurate, unlike WC (2.90% and 2.06%) and WHR (4.60% and 1.14%). The differences with the research assistants' measurements (data not shown) were statistically significant, except for height in the two sessions and BMI in the second session. Measurement differences improved after training for WC, HC and WHR: for HC, only in the intervention group (from 1.81% to 0.78%), whereas for WC and WHR, the intervention group improved more (WC from 2.49% to 1.40%; WHR from 5.75% to 3.45%) than the control group (WC from 3.36% to 2.81%; WHR from 3.45% to 1.14%).

The extent of agreement between the doctors' and the research assistants' measurements was analysed with Bland-Altman plots (see appendix 2). The plots confirmed the lack of systematic bias for weight, height and BMI, since the horizontal line representing the mean difference between the GPs' and the research assistants' values was very close to zero. However, when considering the 95% limits of agreement, except for height, all the anthropometric measurements were inaccurate (i.e. the 95% limits of agreement were not within 2% of the research assistants' average value), though the newer measurements methods (WC, HC and WHR) were largely more prone to measurement error than classical methods (weight and BMI).

Table 2 presents the GPs' self-report of their knowledge and usual practice in anthropometric measurements. They hardly ever used the newer measurement methods and their knowledge regarding these measurements was relatively

Table 1: Mean difference between primary care physicians' anthropometric measurements and gold standard (95% confidence intervals), before and after training, first overall, then in the intervention group (Group I) and in the control group (Group C).

	Before training			After training			p-value for difference in mean error before and after training within each group		p-value for difference in mean error between the two groups	
	Overall (n = 259)	Group I (n = 136)	Group C (n = 123)	Overall (n = 250)	Group I (n = 133)	Group C (n = 117)	Group I	Group C	Before training	After training
Weight (kg)	–0.14 (–0.25 to –0.03)	–0.07 (–0.22 to 0.09)	–0.22 (–0.38 to –0.06)	–0.40 (–0.52 to –0.29)	–0.47 (–0.62 to –0.33)	–0.32 (–0.51 to –0.13)	<0.001	0.42	0.17	0.20
Height (cm)	–0.10 (–0.23 to 0.04)	–0.09 (–0.28 to 0.10)	–0.11 (–0.31 to 0.09)	–0.01 (-0.19 to 0.17)	–0.02 (–0.21 to 0.17)	0.002 (–0.31 to 0.31)	0.62	0.56	0.89	0.90
BMI (kg/m^2)	–0.03 (–0.09 to 0.03)	–0.003 (–0.08 to 0.07)	–0.05 (–0.14 to 0.04)	–0.14 (–0.21 to –0.07)	–0.17 (–0.25 to –0.09)	–0.10 (–0.21 to 0.01)	0.003	0.47	0.44	0.31
Waist circ (cm)	2.58 (2.01 to 3.16)	2.21 (1.49 to 2.94)	2.99 (2.08 to 3.90)	1.84 (1.33 to 2.35)	1.25 (0.57 to 1.93)	2.51 (1.76 to 3.27)	0.06	0.43	0.18	0.01
Hip circ (cm)	–0.94 (–1.73 to –0.14)	–1.84 (–3.15 to –0.52)	0.06 (–0.75 to 0.87)	0.92 (0.42 to 1.42)	0.79 (0.20 to 1.37)	1.07 (0.22 to 1.91)	<0.001	0.09	0.02	0.58
WHR	0.04 (0.03 to 0.05)	0.05 (0.02 to 0.07)	0.03 (0.02 to 0.04)	0.01 (0.003 to 0.02)	0.01 (–0.003 to 0.01)	0.01 (0.003 to 0.02)	0.007	0.05	0.23	0.20

low (most did not know the correct site of tape placement, the WHR formula and the definition of abdominal obesity). GPs' knowledge and practice in anthropometrics were also assessed through direct observation by the research assistants (table 3). The findings confirmed that the waist and the hip circumferences measurements were generally not well performed. Indeed, a large number of doctors took the measurements without having removed the patient's clothes, without palpating the appropriate markers and without marking the site of measurement. Most doctors did not take the measurements in the right place. In addition, they did not measure WC at the end of a normal expiration as recommended. Finally, GPs' skills improved slightly following training in the intervention but not in the control group (data not shown).

Discussion

Summary

The current study confirms that the classical anthropometric measurements (weight, height and BMI determination), when performed by GPs within their practice, are less prone to measurement error than the newer methods (HC, WC and WHR determination), although only height measurement completely fulfilled our criteria for accuracy. The measurement errors were only slightly less important following short theoretical training. Finally, the study highlights large gaps in measurement techniques for the newer anthropometric measurements, though GPs' skills showed slight improvement after training.

Comparison with existing literature

As shown by several authors, the universality and the relatively simple procedure for measuring height and weight probably explain why these measurements are more reliable in general [18, 29, 30]. Interestingly, our study confirms that these measurements are more accurate than the newer anthropometric measurements, even when performed by untrained measurers.

Contrasting with these results, WC, HC and WHR measures are highly inaccurate, which could be related to the fact that they require specific manipulation, that different measurement sites and/or techniques are recommended with lack of standardisation [31, 32], and that they are newer concepts compared to the worldwide well-known BMI. GPs' self-reported skills were relatively low for measuring WC and HC. This was confirmed through direct observation of GPs by the research assistants.

Several authors have studied the reliability of these newer anthropometric measurements and the results tend to confirm that they might suffer from a higher measurement error than weight, height and BMI, even when they are performed by specially trained measurers [18, 30, 33–35]. To our knowledge, only Sebo et al. have so far studied the variability of these measurements when performed by GPs [18]. They showed the higher variability of these new anthropometric measurements, but improvement occurred following a short theoritical and practical training.

Interestingly, the results of these two studies were similar, though the study populations and the statistical analyses differed greatly: the previous study by Sebo et al. was undertaken in an academic primary care clinic, the measurers were less experienced doctors (35.5 years [SD 3.2] vs 44.1

Table 2: Primary care physicians' knowledge and self reported practice in anthropometric measurements (n varies between 24 and 26 owing to missing values).

Characteristics	Weight n (%)	Height n (%)	BMI n (%)	Waist circ n (%)	Hip circ n (%)	WHR n (%)
Frequency of scale calibration						
≥1x/y	3 (11.5)					
<1x/y	13 (50)					
Never	10 (38.5)					
Frequency of measurement						
≥1x/d	18 (69.2)	9 (34.6)	11 (42.3)	2 (7.7)	0	0
≥1x/w, but <1x/d	3 (11.5)	5 (19.2)	6 (23.1)	6 (23.1)	6 (23.1)	1 (3.9)
<1x/w	5 (19.2)	6 (23.1)	8 (30.8)	6 (23.1)	6 (23.1)	1 (3.9)
Never or almost never	0	6 (23.1)	1 (3.9)	12 (46.2)	14 (53.9)	24 (92.3)
Clothing during measurement						
Clothes and shoes removed	6 (24)	4 (15.4)				
Only shoes removed	16 (64)	20 (76.9)				
Fully dressed (with clothes and shoes)	3 (12)	2 (7.7)				
Clothing covering tape placement location removed for measurement				18 (72)	12 (48)	
Correct site of measurement [1]				4 (15.4)	5 (19.2)	
Correct time of measurement [2]				7 (26.9)		
Recorded value to account for the weight of clothes [3]						
The exact reading	8 (32)					
The exact reading – 1 kg	12 (48)					
Other	4 (16)					
Knowledge of the correct formula			25 (100)			10 (38.5)
Correct definition of overweight [4]			22 (84.6)			
Correct definition of obesity [4]			24 (92.3)	8 (30.8)		2 (7.7)

[1] WC = midpoint between lowest rib and iliac crest; HC = maximum extension of buttocks
[2] End of normal expiration
[3] When clothes and/or shoes not removed
[4] overweight: 25 ≤ BMI <30 kg/m^2; obesity: BMI ≥30 kg/m^2, WC ≥102 cm (men) and ≥88 cm (women), WHR ≥0.95 (men) and ≥0.8 (women)

years [SD 6.1]), with less experience in family medicine (8.1 years [SD 3.1] vs 16.3 years [SD 5.8]). There was no gold standard in this previous study, thus reliability was assessed by using technical error of measurement (TEM) and coefficient of reliability (R). The two studies showed not only that the classical anthropometric measurements were more reliable, but also that WHR was the least accurate measurement of all. This finding could be explained by the fact that WHR represents a mixture of two circumferences (WC and HC) assessed using a similar procedure, each of which is highly prone to measurement error.

In general, GPs' skills improved slightly following a short theoretical training, which is in accordance with previous studies demonstrating the benefit of training the measurers in anthropometrics [18, 36]. Adequate training of the measurers is indeed essential to minimise measurement error and provide accurate data. Interestingly, our study showed that even a simple teaching document received by post can lead to improved measurement techniques, though the improvement was not sufficient to achieve a high level of accuracy.

Limitations

First, the GPs responded to a personal invitation to participate in the study and may thus not be representative of all GPs in the Geneva region or in Switzerland, which limits the external validity of our findings. Our results are therefore rather conservative, since the doctors who agreed to participate probably had a particular interest in the theme under study. This may have favoured the quality of the GPs involvement and measurements and contributed to the low rate of missing data. Second, as only 26 GPs agreed to participate whereas 30 were expected, the study was slightly underpowered, but this should have modest effects on the

Table 3: Results of the direct observation by the research assistants of the 26 primary care physicians' practice in anthropometric measurements (two measurements per doctor, n varies between 50 and 52 owing to missing values).

Characteristics	n (%)	Characteristics	n (%)
WEIGHT			
Type of scale used		Recorded value to account for clothes[1]	
Electronic column scale	4 (7.7)	The exact reading – 1 kg	18 (43.9)
Electronic flat scale	8 (15.4)	The exact reading – 0.5 kg	2 (4.9)
Mechanical column scale	26 (50.0)	The exact reading	21 (51.2)
Mechanical flat scale	14 (26.9)	≥2 measurements	0
Clothing during measurement		Measurement to the nearest 0.05 kg[2]	2 (3.9)
Clothes and shoes removed	10 (19.2)		
Only shoes removed	38 (73.1)		
Fully dressed (with clothes and shoes)	4 (7.7)		
HEIGHT			
Type of stadiometer used		Legs closed and stretched	50 (96.2)
Electronic fixed scale	41 (78.9)	Looking straight ahead	52 (100)
Mechanical fixed scale	11 (21.2)	Horizontal bar lowered to the head crown with hair compressed and head maintained	52 (100)
Shoes removed for measurement	48 (92.3)	Measurement at the end of deep inhalation	4/ (7.7)
Back on the side of the vertical board	46 (88.5)	≥2 measurements	0
Heels, buttocks and back touching the vertical board	45 (86.5)	Measurement to the nearest 0.1 cm[3]	6 (12.0)
Standing erect	50 (96.2)		
WAIST CIRCUMFERENCE			
Participant in standing position	52 (100)	Tape applied with correct tension to avoid slipping off and skin compression	49 (94.2)
Clothing covering tape placement location removed[4]	34 (65.4)	Measurement at the end of normal expiration	3 (5.8)
Iliac crest and lowest rib palpated	20 (38.5)	Arms at the sides	16 (31.4)
Correct site of measurement[5]	17 (32.7)	≥2 measurements	3 (5.8)
Level marked before measurement	3 (5.8)	Measurement to the nearest 0.1 cm[6]	12 (23.1)
Tape placed horizontally	49 (94.2)		
HIP CIRCUMFERENCE			
Participant in standing position	50 (96.2)	Tape applied with correct tension to avoid slipping off and skin compression	46 (88.5)
Clothing covering tape placement location removed[7]	28 (53.9)	Arms at the sides	20 (38.5)
Correct site of measurement[8]	11 (21.2)	≥2 measurements	2 (3.9)
Level marked before measurement	0	Measurement to the nearest 0.1 cm[9]	12 (24.0)
Tape paced horizontally	47 (90.4)		

[1] When clothes and/or shoes not removed (n = 41).
[2] To the nearest 0.1 kg: n = 36; 0.2 kg: n = 2; 0.5 kg: n = 9; other: n = 3.
[3] To the nearest 0.5 cm: n = 38; 1 cm: n = 6.
[4] In underwear: n = 32; measurement taken under clothing: n = 2.
[5] Midpoint between lowest rib and iliac crest.
[6] To the nearest 0.5 cm: n = 27; 1 cm: n = 13.
[7] In underwear: n = 27; measurement taken under clothing: n = 1.
[8] Maximum extension of buttocks.
[9] To the nearest 0.5 cm: n = 26; 1 cm: n = 12.

findings (such as larger standard deviations), as we were able to detect statistically significant differences. Third, because to our knowledge no training in anthropometrics has been offered to GPs in our area in recent years, we considered it unlikely that the GPs would know the usual content of such training and thus assumed that they were blinded. Yet we cannot exclude the possibility that some GPs may have been exposed to previous training in anthropometric measurements in another way. Fourth, measurements were undertaken on volunteers. This could have led to a lower than expected prevalence of overweight and obese subjects in our sample, as these could have been reluctant to participate for fear that they would be stigmatised. Our findings showed that this was not, however, the case (overweight: 33.2%, obese: 21.6%, much higher than the usual prevalence in Switzerland) [5]. Finally, as no data were collected on patients who refused to participate, we cannot exclude a selection bias. However, as our focus was on measurements and not on the prevalence of overweight and obesity, such a selection bias is unlikely to have affected our findings.

Strengths

The study conditions were kept close to daily clinical practice and the GPs had no previous training in anthropometric measurements. They only discovered the details of their task several minutes before data collection began, thus limiting the risk of bias. Further, the mean measurement difference between the two research assistants was very small, confirming that a short theoretical and practical training (two 1-hour-sessions) is enough to ensure measurers' skills. It also justified the choice of the gold standard used in our study.

Conclusion

Our study suggests that the classical anthropometric measurements, when performed by GPs within their practice, are less prone to measurement error than the newer methods, and should therefore be favoured in daily practice, though more studies in various settings are needed to generalise these findings.

Acknowledgement: The authors warmly thank all physicians and volunteers who participated to this study.
Funding / potential competing interests: This study was supported by a grant from the SSH (Swiss Society of hypertension) and a grant from the Swiss Academy of Medical Sciences (SAMS). No other potential conflict of interest relevant to this article was reported.
Ethical approval: The research protocol was aproved by the local research ethics committee.

Correspondence: *Paul Sebo, 97–99, rue de Genève, CH-1226 Thônex / GE, Switzerland, Paulsebo[at]hotmail.com*

References

1 Finucane MM, Stevens GA, Cowan MJ, Danaei G, Lin JK, Paciorek CJ, et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9·1 million participants. Lancet. 2011;377:557–67.

2 Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2014;384:766–81.

3 Schneider H, Dietrich ES, Venetz WP. Trends and stabilization up to 2022 in overweight and obesity in Switzerland, comparison to France, UK, US and Australia. Int J Environ Res Public Health. 2010;7:460–72.

4 Marques-Vidal P, Paccaud F. Regional differences in self-reported screening, prevalence and management of cardiovascular risk factors in Switzerland. BMC Public Health. 2012;12:246.

5 Enquête suisse sur la santé 2012. Office fédéral de la statistique (OFS): Neuchâtel, 2012 http://www.bfs.admin.ch/bfs/portal/fr/index/themen/14/22/publ.html?publicationID=5353 (accessed 23 Nov2013).

6 World Health Organization. Global health risks: mortality and burden of disease attributable to selected major risks. World Health Organization: Geneva, Switzerland, 2009.

7 Flegal KM, Kit BK, Orpana H, Graubard BI. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. JAMA. 2013;309:71–82.

8 Obesity: The Prevention, Identification, Assessment and Management of Overweight and Obesity in Adults and Children. National Institute for Health and Clinical Excellence (UK): London, 2006.

9 Lau DCW, Douketis JD, Morrison KM, Hramiak IM, Sharma AM, Ur E. 2006 Canadian clinical practice guidelines on the management and prevention of obesity in adults and children [summary]. CMAJ. 2007;176:S1–13.

10 Siren R, Eriksson JG, Vanhanen H. Waist circumference a good indicator of future risk for type 2 diabetes and cardiovascular disease. BMC Public Health. 2012;12:631.

11 Lean ME, Han TS, Morrison CE. Waist circumference as a measure for indicating need for weight management. BMJ. 1995;311:158–61.

12 Pouliot MC, Després JP, Lemieux S, Moorjani S, Bouchard C, Tremblay A, et al. Waist circumference and abdominal sagittal diameter: best simple anthropometric indexes of abdominal visceral adipose tissue accumulation and related cardiovascular risk in men and women. Am J Cardiol. 1994;73:460–8.

13 Zamboni M, Turcato E, Armellini F, Kahn HS, Zivelonghi A, Santana H, et al. Sagittal abdominal diameter as a practical predictor of visceral fat. Int J Obes Relat Metab Disord. 1998;22:655–60.

14 Wang J. 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. 2003;78:902–3.

15 Chen MM, Lear SA, Gao M, Frohlich JJ, Birmingham CL. Intraobserver and interobserver reliability of waist circumference and the waist-to-hip ratio. Obes Res. 2001;9:651.

16 Moreno LA, Joyanes M, Mesana MI, González-Gross M, Gil CM, Sarría A, et al. Harmonization of anthropometric measurements for a multicenter nutrition survey in Spanish adolescents. Nutrition. 2003;19:481–6.

17 Nordhamn K, Södergren E, Olsson E, Karlström B, Vessby B, Berglund L. Reliability of anthropometric measurements in overweight and lean subjects: consequences for correlations between anthropometric and other variables. Int J Obes Relat Metab Disord. 2000;24:652–7.

18 Sebo P, Beer-Borst S, Haller DM, Bovier PA. Reliability of doctors' anthropometric measurements to detect obesity. Prev Med. 2008;47:389–93.

19 Sebo P, Pechère-Bertschi A, Herrmann FR, Haller DM, Bovier P. Blood pressure measurements are unreliable to diagnose hypertension in primary care. J Hypertens. 2014;32:509–17.

20 Assmann G, Carmena R, Cullen P, Fruchart JC, Jossa F, Lewis B, et al. Coronary heart disease: reducing the risk: a worldwide view. International Task Force for the Prevention of Coronary Heart Disease. Circulation. 1999;100:1930–8.

21 Lean ME, Han TS, Deurenberg P. Predicting body composition by densitometry from simple anthropometric measurements. Am J Clin Nutr. 1996;63:4–14.

22 National Health and Nutrition Examination Survey III (NAHNES III). Body Measurements (Anthropometry). http://www.cdc.gov/nchs/data/

nhanes/nhanes3/cdrom/nchs/manuals/anthro.pdf (accessed 21 Feb 2013).

23 Health Canada. National Nutrition Survey. PEI Nutrition Survey. Interviewer Training Manual 1995.

24 Government of Canada HC. ARCHIVED – Canadian Guidelines for Body Weight Classification in Adults. 2004. http://www.hc-sc.gc.ca/fn-an/nutrition/weights-poids/guide-ld-adult/index-eng.php (accessed 16 Nov2013).

25 Harrell FE, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med. 1984;3:143–52.

26 Shapiro SS, Francia RS. An Approximate Analysis of Variance Test for Normality. Journal of the American Statistical Association. 1972;67:215–6.

27 Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. Br J Nutr. 1999;82:165–77.

28 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1:307–10.

29 DiMaria-Ghalili RA. Medical record versus researcher measures of height and weight. Biol Res Nurs. 2006;8:15–23.

30 Sicotte M, Ledoux M, Zunzunegui M-V, Ag Aboubacrine S, Nguyen V-K. Reliability of anthropometric measures in a longitudinal cohort of patients initiating ART in West Africa. BMC Med Res Methodol. 2010;10:102.

31 Bosy-Westphal A, Booke C-A, Blöcker T, Kossel E, Goele K, Later W, et al. Measurement site for waist circumference affects its accuracy as an index of visceral and abdominal subcutaneous fat in a Caucasian population. J Nutr. 2010;140:954–61.

32 Mason C, Katzmarzyk PT. Variability in waist circumference measurements according to anatomic measurement site. Obesity. (Silver Spring) 2009;17:1789–95.

33 Dhaliwal SS, Welborn TA. Measurement error and ethnic comparisons of measures of abdominal obesity. Prev Med. 2009;49:148–52.

34 Nádas J, Putz Z, Kolev G, Nagy S, Jermendy G. Intraobserver and interobserver variability of measuring waist circumference. Med Sci Monit. 2008;14:CR15–18.

35 Panoulas VF, Ahmad N, Fazal AA, Kassamali RH, Nightingale P, Kitas GD, et al. The inter-operator variability in measuring waist circumference and its potential impact on the diagnosis of the metabolic syndrome. Postgrad Med J. 2008;84:344–7.

36 Chumlea WC, Guo S, Kuczmarski RJ, Johnson CL, Leahy CK. Reliability for anthropometric measurements in the Hispanic Health and Nutrition Examination Survey (HHANES 1982–1984). Am J Clin Nutr. 1990;51:902S–907S.

Appendix 1

Points emphasised during training of the research assistants (two 1-hour sessions) and the doctors (written training document received by the doctors in the intervention group).

Weight

The participants should remove their shoes and all clothes except underwear, and then step on the centre of the scale, remaining in a relaxed position.

Weight is recorded to the nearest 0.05 kg; if a participant refuses to remove his or her clothes, 1 kg is substracted from the measurement reading to account for the garments worn and the refusal is reported.

Height

The participants should remove their shoes and all clothes except underwear, and stand erect on the floorboard of the stadiometer with their backs on the side of the vertical board; the weight should be evenly distributed on both feet, the legs closed and stretched, the arms to the sides and the shoulders relaxed; the heels, buttocks and back should slightly touch the vertical board.

The participant is asked to look straight ahead, inhale deeply and stand fully erect while the examiner lowers the horizontal bar to the head crown with hair compressed, and takes the measurement to the nearest 0.1 cm

Waist circumference

The participants should remove their shoes and all clothes except underwear, and stand erect; they are asked to roll up the shirt/sweater and to lower the trouser/skirt waistband, so the examiner can palpate the hip area to identify the measurement reference points, and to mark the level of measurement (the midpoint between the lowest rib and the iliac crest). The measuring tape is placed horizontally, 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, the arms of the participant to the sides; if the difference between the two recorded measurements is greater than 0.5 cm, a third measurement is taken, and the mean of the two nearest values is recorded.

Hip circumference

The participants should remove their shoes and all clothes except underwear, and stand erect; they are asked to lower the trouser/skirt waistband, so the examiner can palpate the hip area. The tape is placed at the maximum extension of the buttocks, horizontal to the floor, with sufficient tension to avoid slipping off but without compressing the skin.

The measurement is made twice to the nearest 0.1 cm, the arms of the participant to the sides; if the difference between the two recorded measurements is greater than 0.5 cm, a third measurement is taken, and the mean of the two nearest values is recorded.

Cut-off values

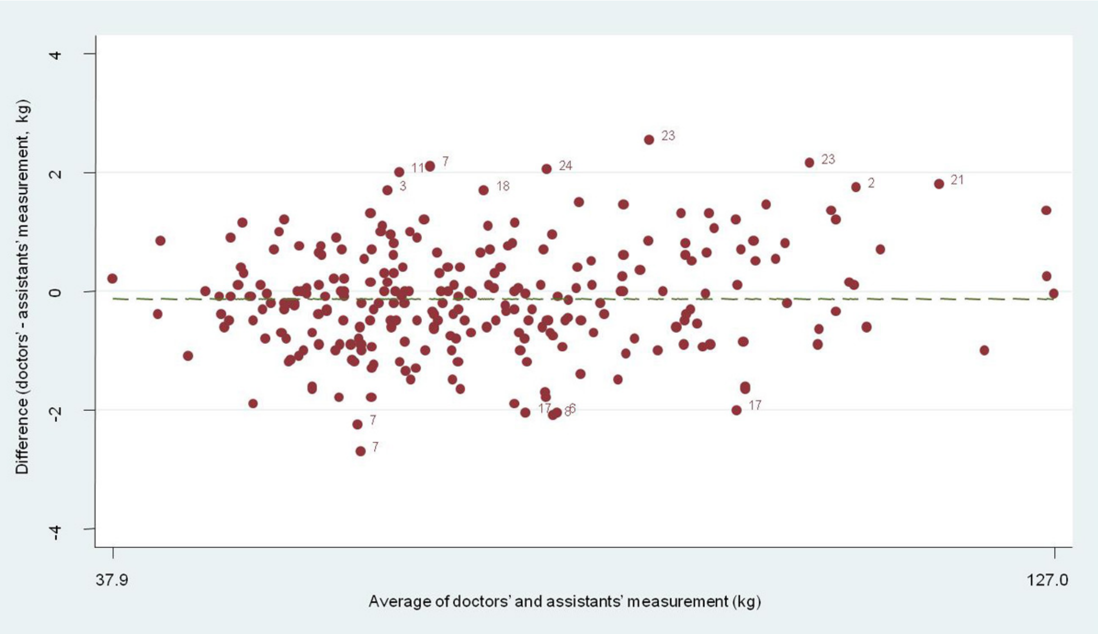
Body mass index (BMI): healthy weight when $18.5 \leq$ BMI <25 kg/m^2, overweight when $25 \leq$ BMI <30 kg/m^2, obesity when BMI ≥ 30 kg/m^2.

Waist circumference (WC) and waist-to-hip ratio (WHR): abdominal obesity in men when WC ≥ 102 cm and/or WHR ≥ 0.95; abdominal obesity in women when WC ≥ 88 cm and/or WHR ≥ 0.8.

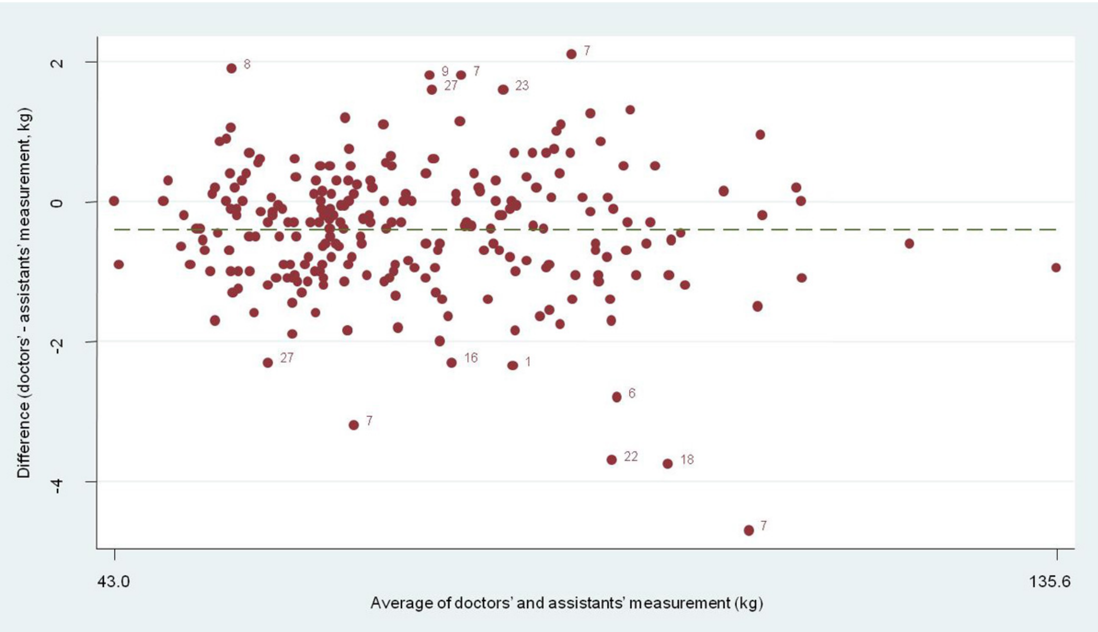
Appendix 2

Figure 1

Comparison between primary care physicians' and research assistants' weight measurement, before (upper figure[1]) and after (bottom figure[2]) training (Bland-Altman plots).



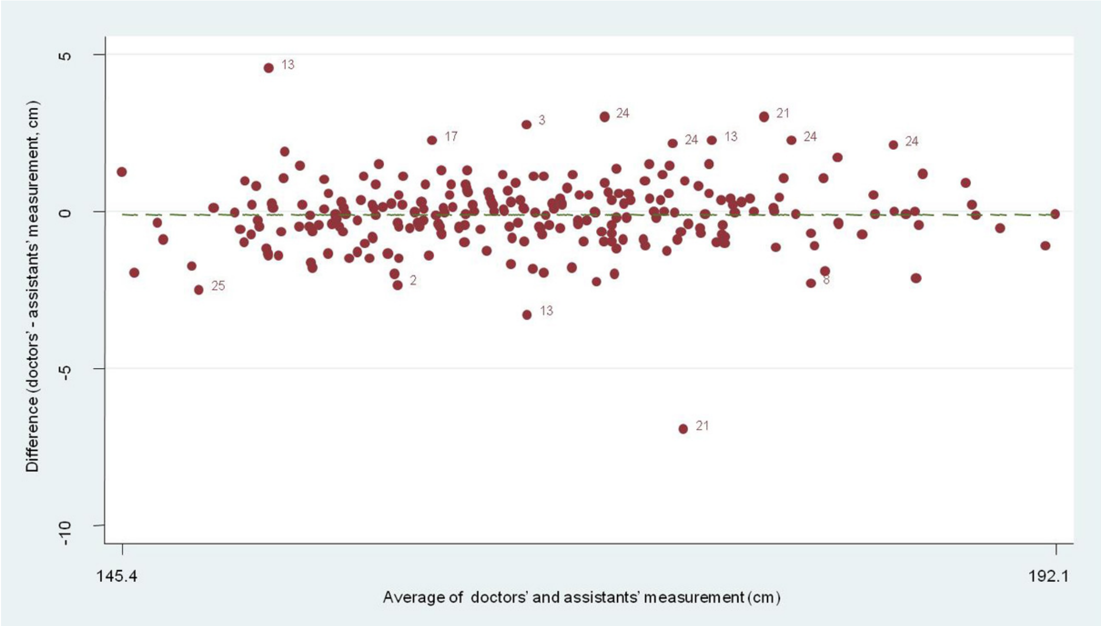
[1] Mean difference between doctors' and research assistants' measurement: 0.14 kg; 95% limits of agreement: –1.91 to 1.63 kg (note that the 15 points which are outside the limits of agreement (5.8%) are labelled by the primary care physicians' identification number).



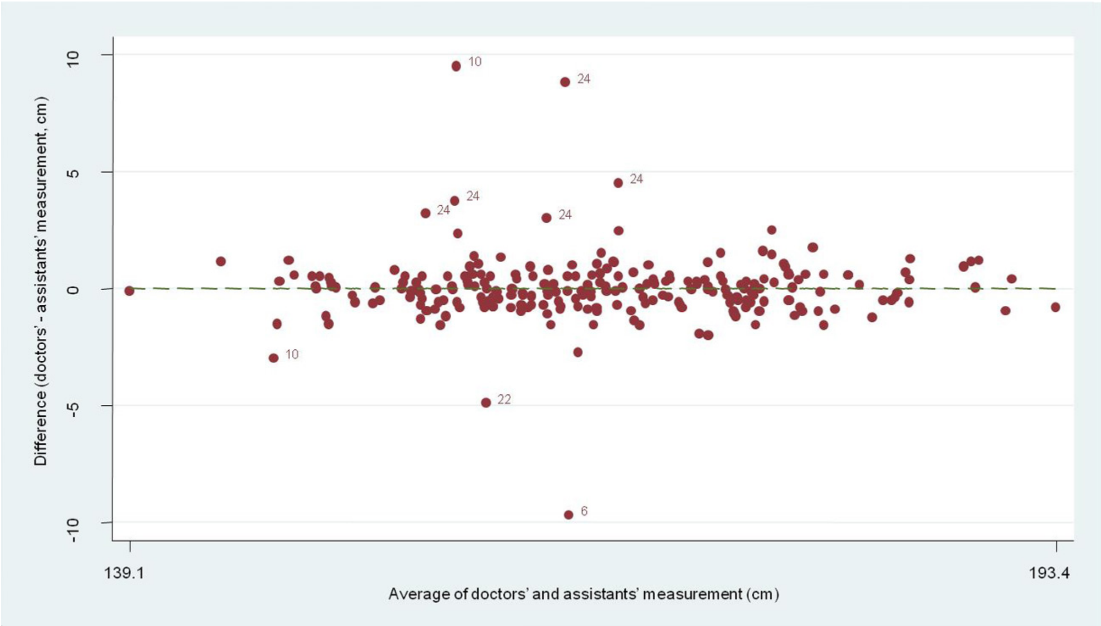
[2] Mean difference between doctors' and research assistants' measurement: 0.40 kg; 95% limits of agreement: –2.21 to 1.42 kg (note that the 14 points which are outside the limits of agreement (5.6%) are labelled by the primary care physicians' identification number).

Figure 2

Comparison between primary care physicians' and research assistants' height measurement, before (upper figure[1]) and after (bottom figure[2]) training (Bland-Altman plots).



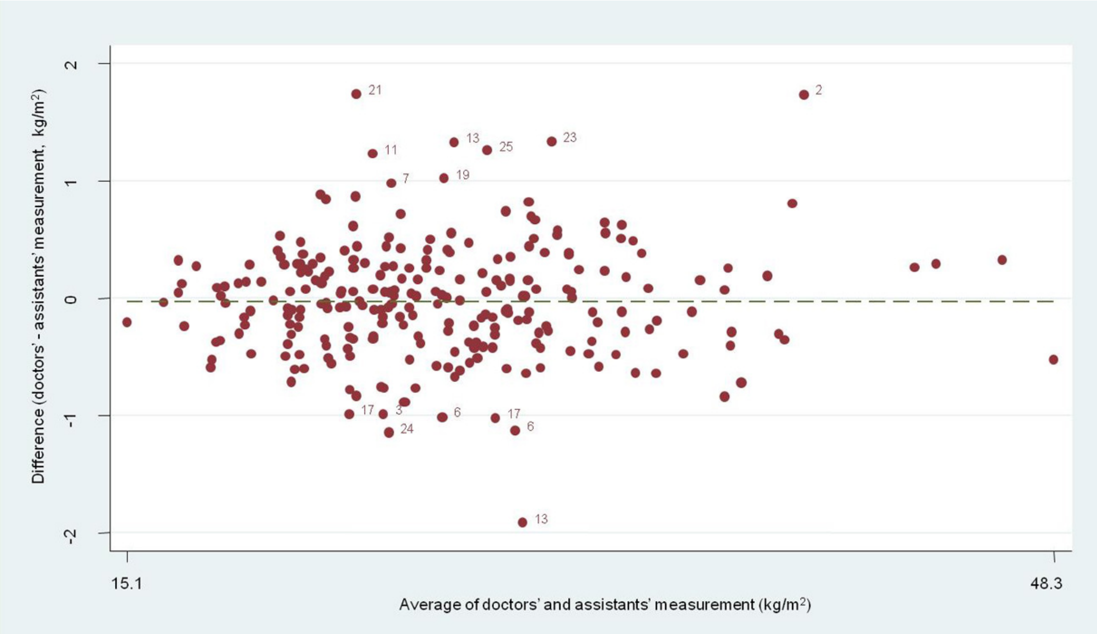
[1] Mean difference between doctors' and research assistants' measurement: 0.10 cm; 95% limits of agreement: –2.26 to 2.07 cm (note that the 14 points which are outside the limits of agreement [5.4%] are labelled with the primary care physicians' identification number).



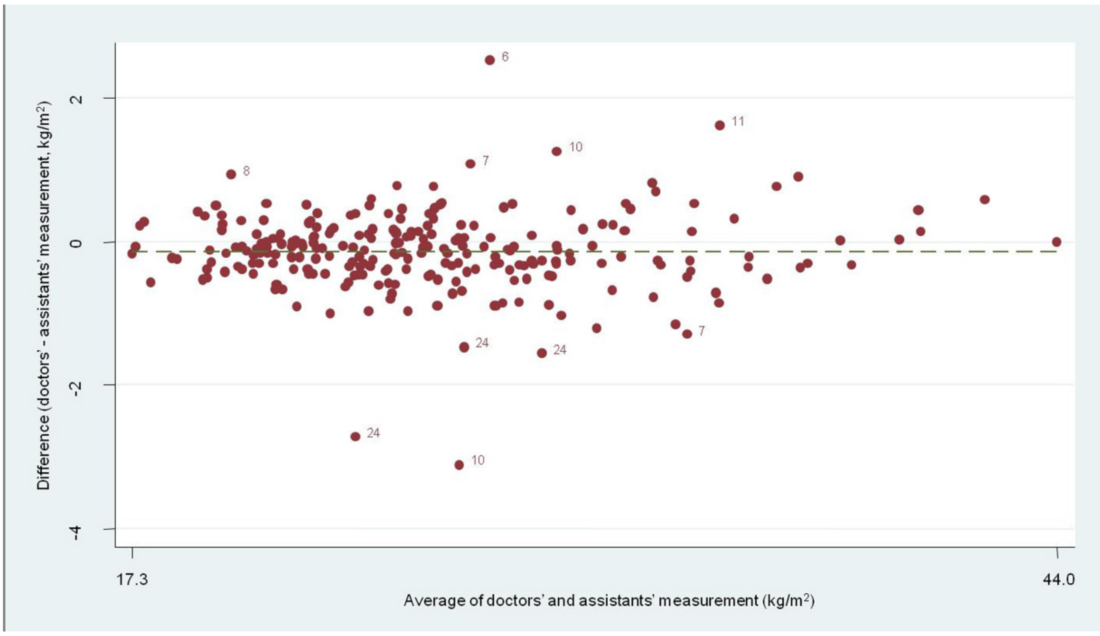
[2] Mean difference between doctors' and research assistants' measurement: 0.01 cm; 95% limits of agreement: –2.79 to 2.77 cm (note that the nine points which are outside the limits of agreement [3.6%] are labelled with the primary care physicians' identification number).

Figure 3

Comparison between primary care physicians' and research assistants' body mass index determination, before (upper figure[1]) and after (bottom figure[2]) training (Bland-Altman plots).



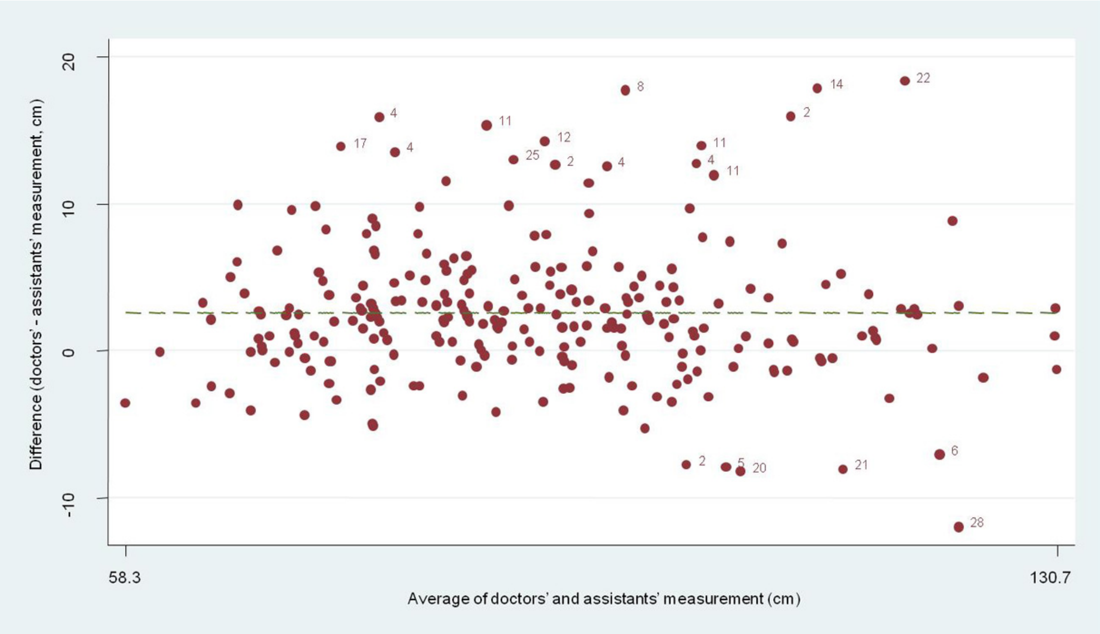
[1] Mean difference between doctors' and research assistants' measurement: 0.03 kg/m^2; 95% limits of agreement: –0.98 to 0.93 kg/m^2 (note that the 15 points which are outside the limits of agreement [5.8%] are labelled with the primary care physician's identification number).



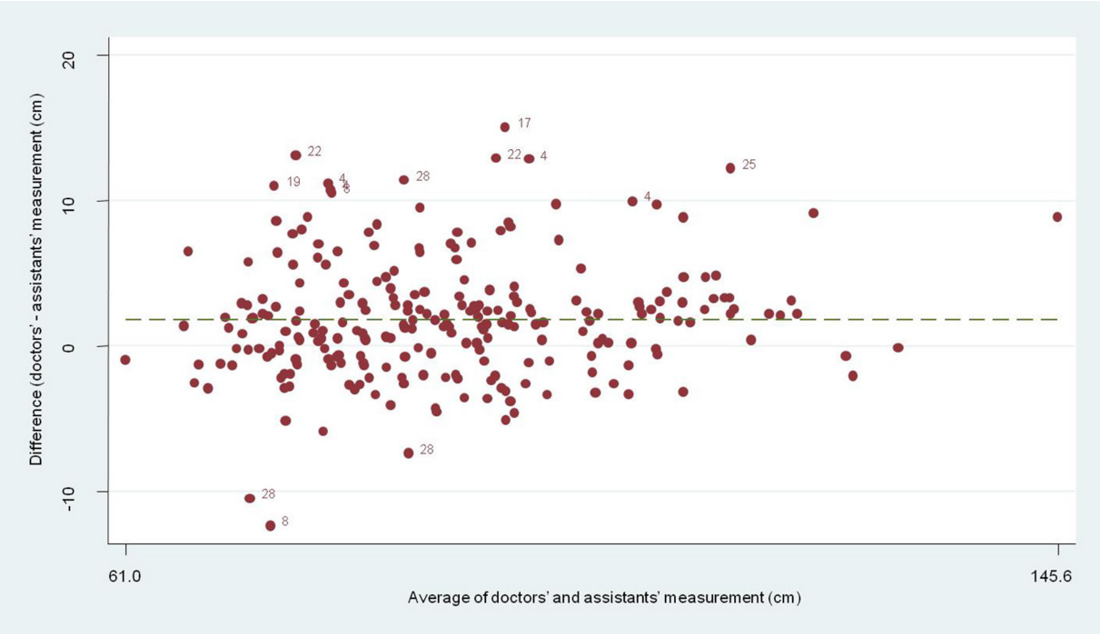
[2] Mean difference between doctors' and research assistants' measurement: 0.14 kg/m^2; 95% limits of agreement: –1.20 to 0.93 kg/m^2 (note that the 10 points which are outside the limits of agreement [4.0%] are labelled with the primary care physician's identification number).

Figure 4

Comparison between primary care physicians' and research assistants' waist circumference measurement, before (upper figure[1]) and after (bottom figure[2]) training (Bland-Altman plots).



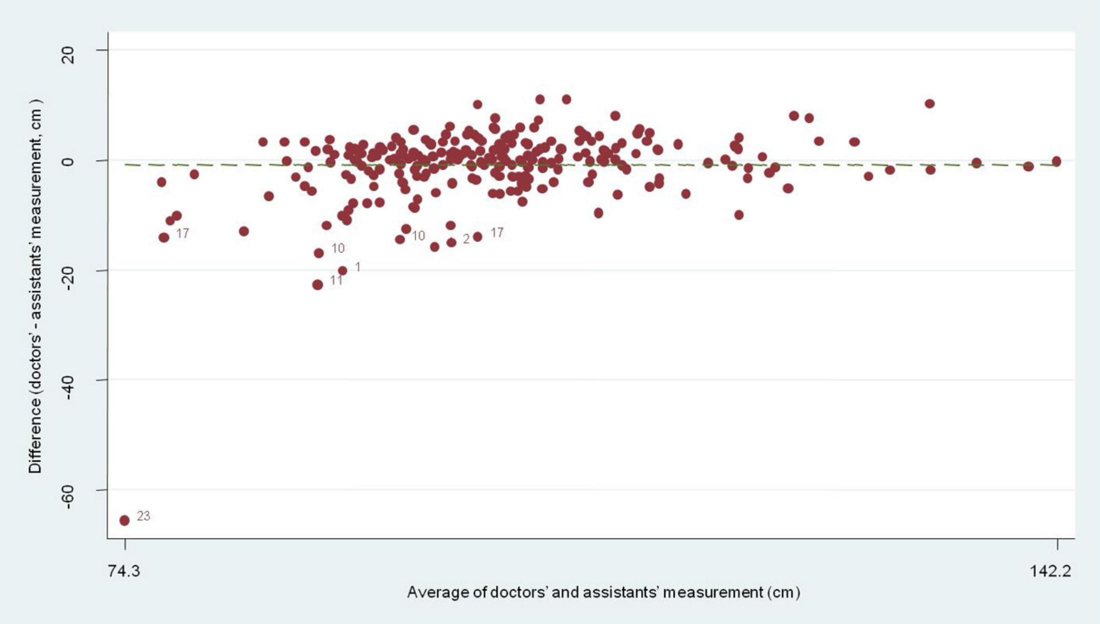
[1] Mean difference between doctors' and research assistants' measurement: 2.58 cm; 95% limits of agreement: –6.61 to 11.78 cm (note that the 21 points which are outside the limits of agreement [8.1%] are labelled with the primary care physician's identification number).



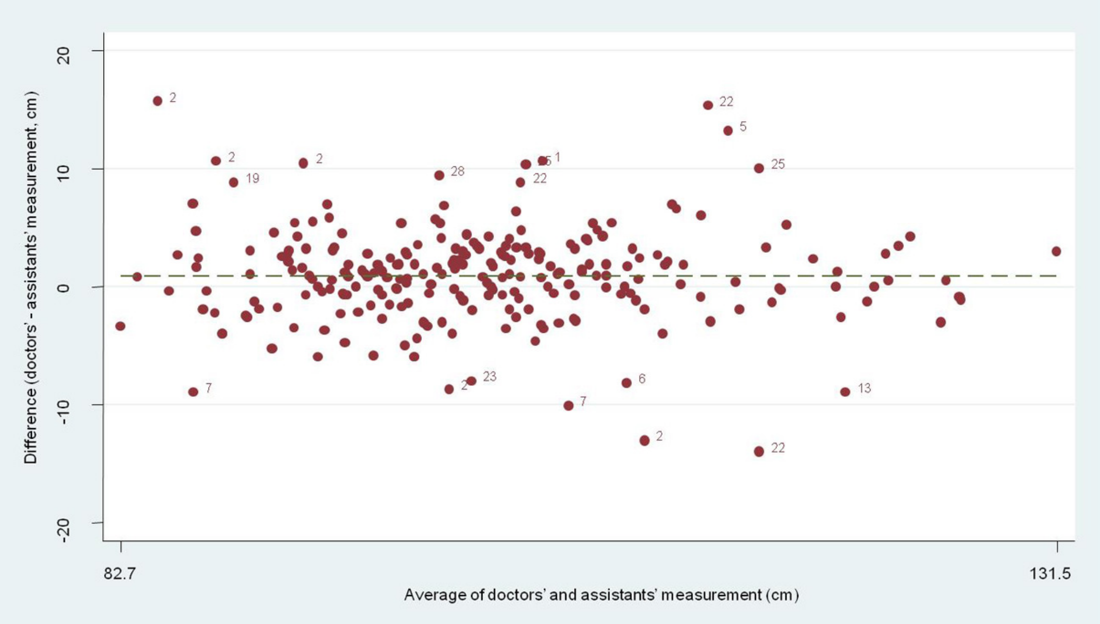
[2] Mean difference between doctors' and research assistants' measurement: 1.84 cm; 95% limits of agreement: –6.15 to 9.83 cm (note that the 14 points which are outside the limits of agreement [5.6%] are labelled with the primary care physician's identification number).

Figure 5

Comparison between primary care physicians' and research assistants' hip circumference measurement, before (upper figure[1]) and after (bottom figure[2]) training (Bland-Altman plots).



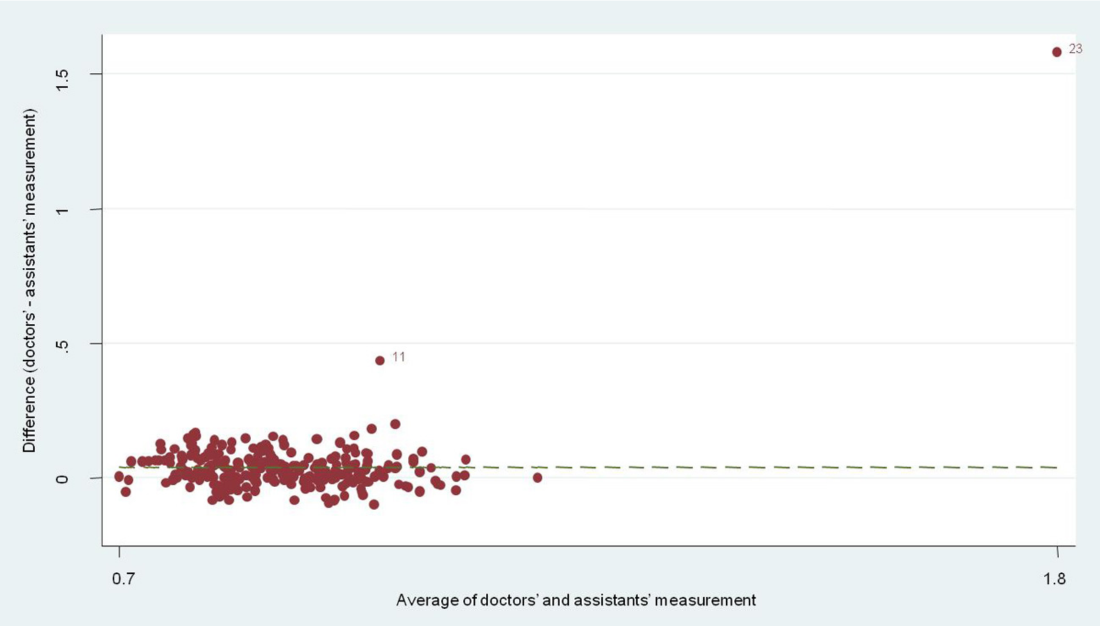
[1] Mean difference between doctors' and research assistants' measurement: 0.94 cm; 95% limits of agreement: –13.65 to 11.77 cm (note that the nine points which are outside the limits of agreement [3.5%] are labelled with the primary care physician's identification number).



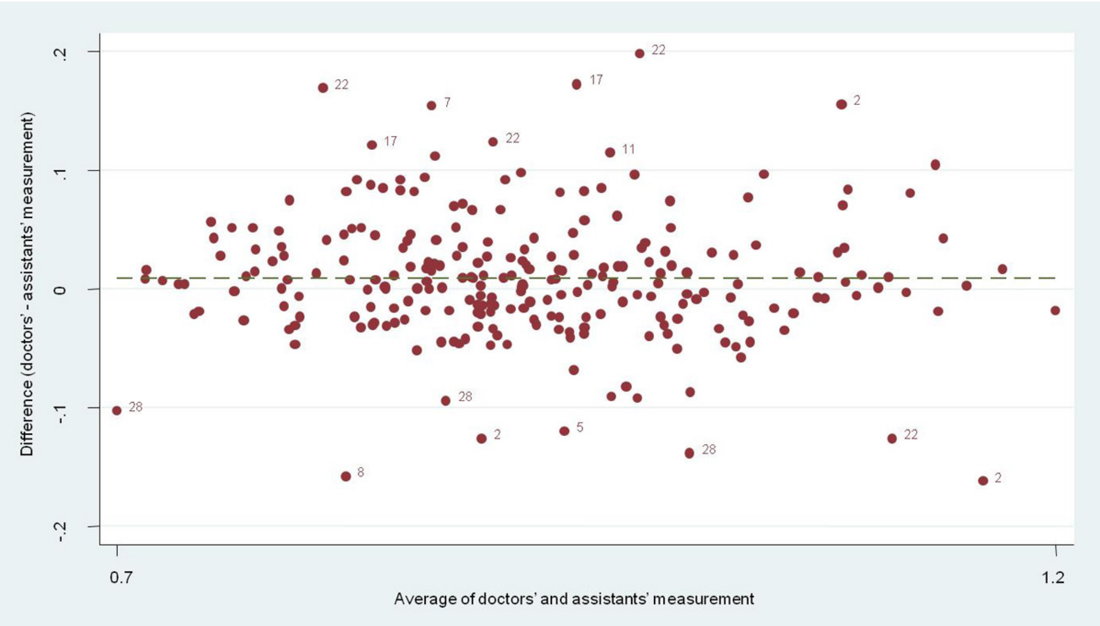
[2] Mean difference between doctors' and research assistants' measurement: 0.92 cm; 95% limits of agreement: –6.94 to 8.77 cm (note that the 19 points which are outside the limits of agreement [7.6%] are labelled with the primary care physician's identification number).

Figure 6

Comparison between primary care physicians' and research assistants' waist-to-hip determination, before (upper figure[1]) and after (bottom figure[2]) training (Bland-Altman plots).



[1] Mean difference between doctors' and research assistants' measurement: 0.04; 95% limits of agreement: –0.19 to 0.26 (note that the two points which are outside the limits of agreement [0.8%] are labelled with the primary care physician's identification number).



[2] Mean difference between doctors' and research assistants' measurement: 0.01; 95% limits of agreement: –0.09 to 0.11 (note that the 16 points which are outside the limits of agreement [6.4%] are labelled with the primary care physician's identification number).