
This is the author’s final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

[http://eprints.gla.ac.uk/143321/](http://eprints.gla.ac.uk/143321/)

Deposited on: 10 July 2017
Accuracy of on-line self-reported weights and heights by young adults

Charoula Konstantia Nikolaou¹, Catherine Ruth Hankey², Michael Ernest John Lean²

Affiliations: ¹SSH/JURI - Institut pour la recherche interdisciplinaire en sciences juridiques (JUR-I), Catholic University of Louvain, Belgium, ²Human Nutrition Section, School of Medicine, University of Glasgow, UK

Corresponding Author
Charoula-Konstantia Nikolaou
Universite Catholique de Louvain
Collège Thomas More, Place Montesquieu 2
Louvain-La-Neuve, 1348
E-mail: charoulanikolaou@yahoo.co.uk

Word count: 2,618

Running title: Validation of online anthropometric data
Abstract (Word count: 249, word limit=250)

Background: e-epidemiology, a convenient and low-cost research method, is becoming increasingly popular. This study seeks to validate on-line self-reported heights and weights against objectively measured data in young adults.

Methods: Young adults self-reported heights and weights in an online lifestyle survey. These were validated using two methods; 1) measurements by staff at the primary-care clinic 2) measurements by a researcher within two weeks of distribution of the survey. Analyses were conducted to determine differences between the self-reported and measured heights and weights and to identify characteristics associated with under- or over-reporting of these.

Results: From a total of 23,010 young adults invited to the survey, 24% provided on-line data, mean age=19.2(SD3.2) years, 43% male, 91% EU citizens. Both self-reported and measured data were available for 1,446 individuals (547 men, 896 women, and mean age 19.2 (SD2.6) years); 1,278 validated using medical records, 168 by researcher measurements. Intra-class correlations between self-reported and measured parameters were: weight (r=0.99), height (r=0.98), with acceptable levels of agreement between measured and self-reported weight, height and BMI using Bland & Altman analyses.

Self-reported weight was underestimated uniformly across BMI categories, gender and ethnicity, by a mean -0.4(SD0.4) kg, (p<0.001). Height was accurately reported overall across BMI and gender: both self-reported and measured heights =1.72(SD0.01)m, p=0.783.

Discrepancies between methods caused misclassification of BMI category for 17(1.8%) of participants.
Conclusions: Engagement of young adults with on-line research is encouraging. Online self-reporting provides acceptably reliable anthropometric data for young adults, with under-reporting of weight by just 0.4 kg.

Key Words: validation, methodology, anthropometric data, weight, height
Introduction

Collecting epidemiological data using the internet (e-epidemiology), is gaining popularity for surveys because of its convenience and speed, and the greater cost of traditional research. High internet penetration across the world has made the delivery of surveys on-line an attractive and alternative way to the traditional face-to-face or paper-based surveys. Both survey methods can incur high non-response rates, for different reasons, and hence bias may be introduced. However, web-based questionnaires offer certain advantages. 1) they can reach large numbers of people simultaneously, 2) they can reach groups who can be hard-to-reach otherwise, like young adults, 3) they can be accessed by participants at any time and at minimal cost, allowing access to busy individuals and those living in remote locations, 4) Data are returned and collated automatically in real-time, 5) Data quality can be improved using automatic mechanisms such as adding logic or skipping questions in the questionnaire, ensuring that minimum essential information is provided before a submission is accepted, and 6) follow-up questionnaires and reminders can be sent easily.

Previous research indicates that various characteristics such as gender, age, and BMI can affect the accuracy of paper-based and interview-based self-reported anthropometric data, with tendencies for height to be overestimated and weight to be underestimated. This can cause misclassification of BMI category, particularly because height is squared, magnifying any errors. Self-reported data may be influenced by the mode of collection, leading to estimation bias for anthropometric data. Mail surveys may provide more accurate reporting of anthropometric data because participants are less affected by the social pressures faced at an interview for data collection. The anonymity of on-line data
collection, and possibly paper questionnaires may similarly result in more accurate self-reported data than face-to-face and telephone interviews.\(^8\)

This study validated self-reported height, weight, and calculated BMI, from data collected by an on-line survey among young adults, against objectively measured data.

**Methods**

This study was approved by the Medical, Veterinary and Life Sciences Ethics committee, of the University of Glasgow for the on-line survey (August 2012, ref FM7309), and by West of Scotland Research Ethics Service for the validation measurements (May 2012, ref 12/WS/0118).

**Study population**

All young adults, studying at a large urban university, were invited to participate in a study of lifestyle changes by responding to a questionnaire delivered through the university-wide email system. The questionnaire contained 27 questions, incorporating 23 multiple choice questions with an option of open responses for some of those about lifestyle habits (n=4). Lifestyle questions included eating and physical activity questions. The questionnaire data were handled by a commercial website (SurveyMonkey.com). Demographic information included names, gender, age, ethnicity and date of birth (DoB), to identify the participants and link the questionnaires completed at the two time-points. Respondents were asked to self-report their current height and weight in imperial or metric units, from which BMI was calculated.
The self-reported weights and heights were validated against two different sets of measured data.

**Validation 1:** Health records of students held in the university primary medical care centre were searched retrospectively and matched with the self-reported data. Weights and heights recorded in the health records had been measured by nursing staff at the surgery at the time of registration at the clinic within one month of the online self-reported weights and heights. Those who had measured data were identified on the on-line survey database using names, gender, DoB, and names, to compare the measures statistically.

**Validation 2:** Within one month of completing the questionnaire survey, weights and heights of a convenience-sample approached and identified at students’ halls were measured by a trained researcher (CN), visiting subjects at their place of residence. Residents were told that a study was being carried out on body weight and they were invited to participate. To avoid bias, participants were not asked if they had also completed the on-line questionnaire. Height was measured to within 0.1 cm by a portable stadiometer. Weight was measured to 0.1kg using a digital set of scales (SECA, UK) without shoes and heavy clothing.

**Statistical Analysis**

Data were analysed using SPPS 23 (SPSS, Chicago) and MedCal (MedCal, 2015) software. To identify errors between self-reported and measured values, the methodology of Bland & Altman was used. The intra-class correlation coefficient (ICC) was used to obtain a
summary measure of agreement between two sources of the same information measured in the same population.\textsuperscript{10} 

Obesity prevalence obtained from self-reported and measured values were compared to verify the impact of self-reported measurement errors on the magnitude of obesity. ANOVA was used to investigate whether differences between self-reported and measured weights, heights and BMI varied according to ethnic group, or weight, height and BMI quartiles. Linear regression analyses were performed to explore relationships between variables and modes of collecting anthropometric data. Outcome variables were measured weight and height: explanatory variables were self-reported weight and height, gender, and ethnic group. Separate analyses were performed for men and women.

Results

All those currently registered at the university of undergraduate degrees were invited to participate (n=23,010). Of these, 5,505 (2,367 (43%) male and 5,009 (91%) EU citizens) participated in the online survey over two consecutive years. They provided self-reported weights and heights at two time points, at the start and end of the academic year (9 months apart) for each survey year. Among these, 1,278 were seen at the GP medical centre and had clinician measures of height and weight available. In addition, the principal researcher (CN) made anthropometric validation measures for 168 subjects who participated in the online survey. Participants’ characteristics and differences between self-reported and measured data are shown in Table 1 and Table 2. Mean difference between self-reported BMI and measured BMI was -0.12 (95% CI -0.134- -0.107, IQR 0.21). Correlation between
self-reported BMI and measured BMI can be seen in Supplementary online-Figure 1. Most (78% n=1,182) of the participants were classified as of normal weight, 13% (n=187) as overweight (BMI=25-30kg/m$^2$) and 5% (n=77) as obese (BMI>30kg/m2) according to the self-reported data (n= 1,446).

Measured data revealed a BMI misclassification for 17 participants in total; 15 participants who with self-reported data were classed as healthy weight were measured as overweight, and two with self-reported data classed as overweight were measured as obese. This misclassification did not contribute to major differences in the proportions of healthy weight, overweight or obese participants [13% (n=187) vs. 14% (n=202) overweight, 5% (n=77) vs. 5.4% (n=79) obese].

**Self-reported vs. measured data from clinic records (n=1,278)**

Mean (SD) self-reported weight was 67.1(16.7) kg while measured was 67.5(16.7) kg, mean difference 0.43(0.37) kg, p<0.001. Mean (SD) self-reported and measured height was 1.72(0.01), with no difference between methods. As a result of the discrepancy between self-reported and measured weight, BMI calculated from self-reported weight was lower than from measured data, by 0.1 (0.2) kg/m$^2$, p <0.001.

Weight remained significantly misreported by approximately 0.4kg for all subgroups of BMI category, gender and ethnicity (Range:0.1-1.3kg). Self-reported and measured weight differed significantly in both healthy weight participants (BMI 18.5-24.9kg/m$^2$, p=0.01 n=1,037) and in overweight/obese participants (BMI >25.0 kg/m$^2$ p<0.001, n=241). Self-reported weight was significantly under-reported by both males (p=0.02) and females (p=0.01).
There were no differences overall between self-reported and measured heights for either males (p=0.84) or females (p=0.66) or by BMI category (p=0.55). However, a subgroup of male individuals of Scottish origin (n=279) underreported their height by 0.1cm (p=0.03).

The ICC between self-reported and clinic-measured data was 0.998 for height, 0.993 for weight, and 0.985 for BMI. Bland-Altman plots for the average versus mean difference in self-reported and actual measurements showed narrow limits of agreement which suggested no bias across the ranges of anthropometric variables (Figure 1).

**Self-reported data vs. data measured by study researcher (n=168)**

A total of 171 students were approached and 168 agreed to have their height and weight measured, a response rate of 98%. Mean (SD) self-reported weight was 66.9(17.7)kg while measured weight was significantly lower at 67.5(16.7)kg, mean difference= -0.6(0.54) kg, p<0.001. Mean (SD) self-reported height and measured height were the same 1.71(0.09)m and 1.71(0.07)m, respectively. As a result of the discrepancy between self-reported and measured weight, BMI calculated from self-reported height and weight was significantly lower than measured, by 0.2(0.2) kg/m² p<0.001).

Self-reported and measured weight differed significantly in both healthy-weight participants (BMI 18.5-24.9kg/m², p=0.03 n=145) and in overweight/obese participants (BMI >25.0 kg/m² p=0.03, n=23). Self-reported weight was significantly underreported by both females and males (p=0.02). Height was the same for self-reported and measured methods for all groups by BMI and gender.
The ICC between self-reported and researcher-measured data was 0.9968 for height, 0.990 for weight, and 0.9992 for BMI.

Bland Altman plots for the average versus mean difference in self-reported and actual measurements showed narrow limits of agreement, with little bias across the range of variables (Figure 2).

**Discussion**

Recruitment to this study using electronic methods was convenient and less time-consuming compared to traditional research. A recent systematic review exploring recruitment methods specific to young adults for lifestyle programmes aimed at the prevention of weight-gain, suggested social media/electronic approaches held promise, but suggested research in this area was scare. This study has established that online self-reported height and weight is generally reliable in a young adult population, if it is accepted that weight was under-reported by around 0.4kg (1 pound) in self-reported data compared to objectively measured data in a largely normal weight population. There was no difference between methods for height, so BMI estimates from self-reported data were only affected by the small under-reporting of weight. This underestimation changed the BMI category classification for only 17 (1.3%) of the 1,443 participants who had objectively measured data. Measured and self-reported data for height, weight, and BMI were all strongly positively correlated, with good agreement across the ranges. Concern is widely expressed over the reliability of self-reported data in general, and in particular the risk of under-reporting body weight and over-estimating height leading to exaggerated underestimates of BMI, particularly among overweight and obese individuals.
Encouraging results are from a reasonably large number of individuals whose height and weight distributions were rather similar to those in the entire on-line study, so these results appear generalizable. With conventional survey methods, heavier adults are more prone to underreport. The discrepancies between on-line self-reported and measured weight in the current study were comparable with, or smaller than, those reported by the few published on-line studies validating weight and height measurements. The only previous validation of on-line data of young adults, in 117 Australians with mean age 23.7 (3.9) years and mean BMI of 24.18 kg/m², also found on-line weight under-reporting by 0.55 kg. In that study, participants over-reported height by 1.36 cm. In older subjects, Lassale et al. found under-reporting of 0.40 kg by men and 0.52 kg by women, among 815 adults in France with mean age 53 years and mean BMI of 24.1 kg/m² with 67.4% of participants being of normal weight. Bonn et al. found greater under-reporting, of 1.2 kg among 149 normal weight individuals (76.5% of participants were of normal weight, mean BMI not reported) however those aged <30 years (77 (51.7%)) under-reported by only 0.7 kg, while those >30 years under-reported weight by 1.7 kg. It therefore appears that on-line self-reporting of body-weight is less reliable in older subjects. Another on-line study with a validation sample of 140 adult participants from seven European countries (20 participants from each country) found weight underreporting by 0.7 kg and correct BMI classification in the 93% of the cases (Age range: 18-60 years old, Mean BMI=24.9 kg/m², 56.4% healthy weight participants). In a weight loss study with 277 participants (Mean BMI=36 kg/m²), weight reported on-line was underreported by 0.5 kg at 6 months and by 1.1 kg at 24 months. In a study of 1,698 adolescents (approximately 16 years old, Mean BMI=21.1 kg/m²) weight was underreported on-line by 1.1 kg.
No discrepancies in height reporting were found in this study, except among Scottish males, whose self-reported height was 0.1 cm lower than the true value. This is a very small discrepancy and may merely reflect measurement errors. Interestingly, this unexpected finding is similar to that of Bolton-Smith et al who reported underreporting of height in older Scottish adults by a mean of 1.3 cm.²¹

The closeness of the self-reported measures to those taken by others may reflect the greater availability of accurate scales and height measures. This anecdotal association is supported by the presence of scales in gyms and leisure facilities, and on a pay-per-use basis in many retail outlets.

The main strength of the current study is the fact that the self-reported data were validated against objectively measured data which were collected routinely for another purpose, independently from the on-line study. This reduced risk of self-selection bias towards including more motivated volunteers than in the general population. Participants were unaware that the heights and weights they provided on-line would be compared with those measured when registering at the general practice health centre. The second set of measured data was collected by trained researcher (CN) approaching young adults who lived in university halls, and therefore not random samples of students. The response rate of about 25% for completing the questionnaire survey was above the average response rate observed in similar on-line research²⁰, indicating a willingness among young adults, studying in an urban setting, to report their heights and weights on-line. However, the results are for a university population, and they were perhaps reassured that their data was going to a reputable source (university researchers). A low response rate does not inevitably introduce bias, but people willing to volunteer for research may not be representative of
those who decline. The study population was young adults attending higher education. About 50% of school leavers now go on to college or university in UK\textsuperscript{21} so these are no longer a minor elite group. The prevalence of BMI > 30 obesity (5.3\%) in this sample were comparable with that reported in the population-based Scottish Health Survey (SHS) for young adults, which showed that 16-24 year-olds that 7\% were obese at this age.\textsuperscript{22}

A limitation of this study is that while all the students agreed to provide height and weight measurement when registering at the health service clinic of the university as a requirement for registration, so these data were unselective, we did not get data from the other health clinics outside the university. Hence, our sample was not random and we cannot say whether the students registering with health services outside the university were any different.

Some participants may habitually weigh themselves regularly, others more rarely. The time elapsing between self-reported and measured data is therefore important, particularly among young adults whose weights can change rapidly.\textsuperscript{23,24} Greater time elapsed (5-6 weeks) between the self-reported and the researcher-measured data, possibly explaining the greater discrepancy than with the clinic measurements. Measurement bias is also possible, between the equipment used by participants and the calibrated equipment used by the principal researcher and clinic. Under reporting of weight is established in the obese and overweight, so our data, with a huge majority of those of a healthy weight have a lower risk of under reporting.

Importantly, the weight discrepancy between self-reported and measured values is small, and not likely to be of clinical importance. This is very encouraging in an often difficult to
engage population of young adults studying for further qualifications and prone to
unwanted weight-gain.24

To conclude with, there is very strong agreement between on-line self-reported and
measured anthropometric data in young adults studying at an urban university. Self-
reported weight was under-reported by about 0.4kg, across genders, and BMI categories.

There was no bias in self-reported height. These findings suggest that online self-reporting
can be considered a valid method for collecting anthropometric data, provided a consistent
small underestimate is accepted. Response rates of around a quarter of the sample are
encouraging and suggest on-line data collection offers promise.
**Key Points**

- Online collection of anthropometric data is a convenient and low cost research method
- There is very strong agreement between on-line self-reported and measured anthropometric data in young adults
- Online self-reporting can be considered a valid method for collecting anthropometric data

**Acknowledgements**

We would like to acknowledge IT staff and those of NHS medical centres for facilitating data collection.

**Conflict of interest**

The authors declare that they have no competing financial interests

**Funding**

The first author received a scholarship from the State scholarship foundation of Greece. The funding body had no role in the design, collection or analysis and interpretation of data.
References


22. Scottish Health Survey 2014.


Table 1: Means of measured and reported weight, height and Body Mass Index (BMI), and intra-class correlation between reported and taken measurements by health records according to gender.

<table>
<thead>
<tr>
<th></th>
<th>Mean Measured</th>
<th>95% CI</th>
<th>Mean Reported</th>
<th>95% CI</th>
<th>ICC (Absolute agreement)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All (n=1,278)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>67.5(16.7)</td>
<td>66.5-68.4</td>
<td>67.1(16.7)</td>
<td>66.1-68.0</td>
<td>0.9993</td>
<td>0.9976 - 0.9996</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.72(0.01)</td>
<td>1.71-1.72</td>
<td>1.71(0.01)</td>
<td>1.71-1.72</td>
<td>0.9979</td>
<td>0.9976-0.9981</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.6(4.6)</td>
<td>22.4-22.9</td>
<td>22.5(4.6)</td>
<td>22.3-22.8</td>
<td>0.9985</td>
<td>0.9976-0.9989</td>
</tr>
<tr>
<td><strong>Male (n=478)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>76.1(16.3)</td>
<td>74.6-77.6</td>
<td>75.7(16.3)</td>
<td>74.3-77.2</td>
<td>0.9990</td>
<td>0.9988-0.9992</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.81(0.08)</td>
<td>1.79-1.81</td>
<td>1.8(0.08)</td>
<td>1.79-1.8</td>
<td>0.9970</td>
<td>0.9964-0.9975</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.1(4.5)</td>
<td>21.8-22.5</td>
<td>23.3(4.4)</td>
<td>22.9-23.7</td>
<td>0.9981</td>
<td>0.9975-0.9989</td>
</tr>
<tr>
<td><strong>Female (n=800)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>62.2(14.2)</td>
<td>61.2-63.2</td>
<td>61.8(14.2)</td>
<td>60.9-62.8</td>
<td>0.9994</td>
<td>0.9774-0.9998</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.67(0.08)</td>
<td>1.66-1.68</td>
<td>1.67(0.07)</td>
<td>1.67-1.68</td>
<td>0.9984</td>
<td>0.9981-0.9986</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.1(4.5)</td>
<td>21.8-22.5</td>
<td>22.0(4.8)</td>
<td>21.7-22.3</td>
<td>0.9988</td>
<td>0.9985-0.9993</td>
</tr>
</tbody>
</table>

All data are mean and SD
Table 2: Means of measured and reported weight, height and Body Mass Index (BMI), and intra-class correlation between reported and taken measurements by trained researcher according to gender.

<table>
<thead>
<tr>
<th></th>
<th>Mean Measured</th>
<th>95% CI</th>
<th>Mean Reported</th>
<th>95% CI</th>
<th>ICC (Absolute agreement)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All (n=168)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>67.5(17.6)</td>
<td>64.7-70.1</td>
<td>66.9(17.7)</td>
<td>64.2-69.5</td>
<td>0.9990</td>
<td>0.9939-0.9996</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.71(0.09)</td>
<td>1.7-1.72</td>
<td>1.71(0.8)</td>
<td>1.69-1.72</td>
<td>0.9968</td>
<td>0.9956-0.9976</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.9(5.0)</td>
<td>22.1-23.7</td>
<td>22.7(5.0)</td>
<td>21.9-23.4</td>
<td>0.9992</td>
<td>0.9990-0.9995</td>
</tr>
<tr>
<td><strong>Male (n=69)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>67.9(16.0)</td>
<td>64-71.7</td>
<td>67.3(16.1)</td>
<td>63.4-71.2</td>
<td>0.9988</td>
<td>0.9907-0.9996</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.72(0.09)</td>
<td>1.7-1.74</td>
<td>1.72(0.08)</td>
<td>1.7-1.74</td>
<td>0.9887</td>
<td>0.9803-0.9935</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.8(4.7)</td>
<td>21.6-24.0</td>
<td>22.6(4.7)</td>
<td>21.5-23.7</td>
<td>0.9983</td>
<td>0.9981-0.9987</td>
</tr>
<tr>
<td><strong>Female (n=99)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>67.1(18.7)</td>
<td>63.3-70.9</td>
<td>66.6(18.8)</td>
<td>62.8-70.4</td>
<td>0.9989</td>
<td>0.9938-0.9996</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.7(0.08)</td>
<td>1.68-1.72</td>
<td>1.7(0.08)</td>
<td>1.68-1.72</td>
<td>0.9962</td>
<td>0.9937-0.9976</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.9(5.1)</td>
<td>21.9-24.0</td>
<td>22.8(5.2)</td>
<td>21.7-23.8</td>
<td>0.9978</td>
<td>0.9973-0.9983</td>
</tr>
</tbody>
</table>

All data are mean and SD
**Figure 1:** Bland-Altman plots for agreement between self-reported data and measurements by clinic nursing staff.
**Figure 2:** Bland Altman plots for agreement between self-reported and measurements by trained researcher.
Supplementary Figure 1: Scatter plot for self-reported BMI and measured BMI