Comparison of self-reported and directly measured weight and height among women of reproductive age: a systematic review and meta-analysis

Running Title: Self-report weight and height in women

Mariana Seijo¹*, Nicole Minckas¹§, Gabriela Cormick¹,², Daniel Comandé ³, Agustín Ciapponi³, José M. Belizan¹.

¹ Department of Mother and Child Health Research, Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Buenos Aires, Argentina.
² Department of Human Biology, Faculty of Health Sciences, University of Cape Town.
³ Argentine Cochrane Center, Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Buenos Aires, Argentina.
§: Both authors contributed equally to this work

*Corresponding author:

Department of Mother and Child Health Research, Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Buenos Aires, Argentina.

Dr. Emilio Ravignani 2024 (C1414CPV), Buenos Aires, Argentina

(+54-11) 4966-0082

Email: mseijo@iecs.org.ar
Conflict of Interest Statements: No conflict of interest was declared.
Full Abstract

Introduction: The use of self-report as a strategy for collecting data of women’s weight and height is broadly spread both in clinical practice and epidemiological studies. This study aimed to compare self-reported and directly measured weight and height among women of reproductive age. Material and methods: In July 2015 we searched MEDLINE, EMBASE, COCHRANE, CINHAL, LILACS and grey literature. We included women of reproductive age (12 to 49 years old) independently of their weight or height at the time of the study. Women with any condition that implies regular track of their weight (e.g., eating disorder) were excluded. Two reviewers independently selected, extracted and assessed the risk of bias of the studies. We used RevMan 5.3 to perform the meta-analysis. Heterogeneity was assessed using the I² statistic. Results: Following eligibility assessment, 21 studies including 18,749 women met the inclusion criteria. The results of the meta-analysis showed an underestimation of weight by -0.94kg (95%CI -1.17, -0.71kg; p<0.0001; I²=0%) in the overall sample and an overestimation of height by 0.36cm (95%CI 0.20, 0.51; p<0.0001; I²=35%) based on self-reported as compared to directly measured values. Conclusions: This review shows that self-reported weight and height of women of reproductive age is slightly different than direct measures. We consider that the magnitude at which self-reported data over or underestimates the real value is negligible regarding clinical and research use.

Key words: Self-Assessment, Body Weights and Measures, Body Weight, Body Height, Body Mass Index, Women, Reproductive Age.

List of abbreviation:

BMI: Body mass index
CI: Confidence Interval
DM: Direct measured
SD: Standard Deviation
SE: Standard Error
SR: Self-reported
WHO: World Health Organization
Key messages: Self-reported weight and height in women of reproductive age is a measure that closely estimates the real values and can be used as proxy both in clinical and research evaluations related to reproductive health.
Introduction

Body mass index (BMI) is a simple and useful indicator to classify individuals as healthy or at risk according to their weight and height \(^{(1)}\). Traditional anthropometric measures such as weight and BMI are often used in epidemiological studies to assess changes in population health and nutritional status \(^{(2)}\). Regarding women’s health, BMI prior to pregnancy requires strict attention as it can be a risk factor not only for women, but also for future generations\(^{(3)}\). Because of this, the International Federation of Gynecology and Obstetrics (FIGO) emphasizes the need to control pre-conceptional body weight and BMI to prevent abnormal values that can impact significantly on maternal and neonatal health outcomes \(^{(3)}\).

Anthropometric measures are often gathered through self-administered questionnaires. This data collection method has the advantages of being quick, easy to administer, and cost-effective when working with large samples, or when individuals are spread over large areas \(^{(4)}\). In research, the self-report of height or weight is highly used in descriptive studies to save significant amount of time and resources \(^{(5-8)}\). In clinical practice, self-reported measures of weight are also a useful strategy to determine historical weights; for example, self-report allows for estimation of pregnancy weight gain that would otherwise be difficult due to the variable stages in which the first antenatal visit occurs. Despite these advantages, the utility of self-reported measures has been questioned, particularly when it relates to anthropometric measures. There is a global preconceived idea that participants tend to overestimate their height and underestimate their weight, resulting in a lower estimate of BMI \(^{(4)}\). The greatest hazard of unreliable reporting of weight and height is the inaccurate estimation of the prevalence of overweight and obesity, which can result in unsupported decision-making \(^{(4)}\).

It is vital to have an up-to-date systematic review on this topic in order to reduce the risk of bias when reporting the results of a study. Any important difference between self-reported and directly measured data found should be taken into consideration when selecting data collection methods for future studies or clinical actions.

The objective of this review is to compare self-reported with directly measured weight and height among women of childbearing age. The purpose of these meta-analyses is to
give a summary estimate of the possible bias that can occur when using self-report as a data collection method.

**Methods**

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Statement (PRISMA Statement)\(^9,10\) and the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) statement.

**Criteria for considering studies for this review**

We selected cross-sectional and prospective or historical cohort studies that compared individual self-reported with directly measured weight and height data. We included published or unpublished studies from 2000 onward that reported at least 20-paired values of self-reported and directly measured weight or height, or data of the difference between them. No language restriction was used.

We included healthy non-pregnant women of reproductive age, independent of their weight or height. We considered reproductive age to be from 12 to 49 years old. All methods to obtain a self-reported or directly measured weight and height were accepted. We excluded women with a disease or condition that implied regular monitoring or records of their weight, such as women following dietary plans or women with eating disorders.

Studies were included only if they expressed the outcome as “mean self-reported weight or height”, “mean directly measured weight or height” or “mean difference between self-reported and directly measured weight or height”.

**Search methods for identification of studies**

**Electronic searches**

A literature search for articles published from January 1\(^{st}\), 2000 to July 14\(^{th}\), 2015 was conducted within the main international and regional databases, through generic and academic internet searches, and through meta-search engines.

We searched records from the following databases:

- CENTRAL: The Cochrane Library (last available Issue 2015)
- MEDLINE (January 2000 to July 2015)
The simplified and complete search with filters in MEDLINE is described below; these were adapted appropriately for each database (Supporting Information Appendix S1 Supplementary Methods). We also reviewed the reference lists of included studies for potential additional studies.

**Data collection and analysis**

**Selection of studies**

All phases of the study selection and processing were completed using EROS® (Early Review Organizing Software, IECS, Buenos Aires), a web-based platform designed for the process of systematic reviews. As an initial screening, pairs of reviewers (MS, NM) independently reviewed the articles, evaluating the titles and abstracts of identified studies according to pre-specified criteria. Discrepancies were resolved by consensus of the whole research team. Articles included after the initial evaluation were retrieved in full text for a second screening to determine eligibility. Finally, the same reviewers independently extracted and assessed the risk of bias of each full text article.

**Data extraction and management**

We used a web-based spreadsheet to extract the information. One reviewer extracted data from the included studies and a second reviewer double-checked it to minimize potential errors. This process was piloted on 20 papers to refine it. Discrepancies were resolved by consensus of the whole team.

The information extracted from each study included author, publication year, type of study, region and country of study, participant characteristics (age and education level), sample size, methods to obtain directly measured weight and height (stadiometer, anthropometer, or other type of measuring (tape or ruler, variety of scales), methods to obtain self-reported weight and height (long distance survey, on-site interview, self-administered questionnaire), time between collection of self-reported and directly measured data, order of measures, ethical considerations, and outcomes (mean self-reported and directly measured weight or height or mean differences between self-
reported and directly measured weight or height, and their standard deviation [SD]). Authors of studies reporting incomplete information were contacted to provide missing information. We waited for one month for the author’s answer before excluding the article.

Assessment of risk of bias and data analysis

The risk of bias of observational studies was assessed using a checklist of essential items based on the STROBE (Strengthening the Reporting of Observational studies in Epidemiology) (12). The STROBE essential checklist includes: selection of participants, control of confounders, measurement of exposure and outcome, and conflict of interest. Pairs of independent reviewers assessed the methodological quality. Discrepancies were resolved by consensus of the whole team.

The null hypothesis when comparing self-reported and directly measured weight and height stated no difference between methods (self-reported = directly measured). Those measurements expressed in pounds or inches were transformed to kilograms and centimeters, respectively, and the reported standard errors (SE) were converted to SD using the following formula: \( \sqrt{n} \times SE \). We performed a meta-analysis using the continuous outcomes of all the studies that reported mean values of weight or height using self-reported and directly measured methods. A summary estimate obtained from the meta-analysis of a mean difference not equal to 0 would indicate that the use of self-report affects positively or negatively on the measure compared to the use of direct measurements; based on either difference, self-reported values could be defined as a weak method for data collection. We used RevMan 5.3 (13) to perform the meta-analysis and to calculate the two-tailed P-values and 95% confidence intervals (CI).

We measured heterogeneity using the \( I^2 \) statistic as follows: low heterogeneity (\( I^2 \) less than 25%), moderate heterogeneity (\( I^2 \) between 25–75%), and high heterogeneity (\( I^2 \) greater than 75%).

For those studies that only reported mean differences between methods, we performed a generic inverse-variance meta-analysis, which considered mean difference and SE. To be able to include all the studies we used RevMan’s calculator function to extract mean differences and SE for each of them. The resulting value indicated the directionality of the findings. A result under 1 indicated that the directly measured values were higher
than the self-reported ones; a result above 1 indicated that the self-reported values were higher than the direct measured ones; a value of 1 indicated no difference between methods.

Pre-specified subgroup analyses by age, time between self-reported and direct measured measurement (same day, different days), region of the study (Latin America & Caribbean, Europe, North America, Oceania, Asia), self-report method (long distance survey, self-administered questionnaire on-site, in-person interview) and women’s BMI were performed. For all the meta-analyses we used a random effect model to address possible clinical or methodological heterogeneity between studies.

We compiled the age data into three groups: 1) 12 to 18 years, 2) 19 to 35 years and 3) 36 to 49 years. For studies in which age was grouped differently and data could not be disaggregated, we based our groups on the category to which the majority of study participants belonged. BMI was classified following WHO categories (underweight less than 18.5, normal weight 18.5 to less than 25, overweight 25 to less than 30 and obesity 30 or more)\(^{(1)}\). The protocol was registered in PROSPERO, an international prospective register of systematic review protocols (Registration Number CRD42015029142).

**Results**

**Description of studies**

**Results of the search**

The search strategy retrieved 1638 references after removing duplicates. Of those, 1476 references were excluded by title and abstract, leaving 162. Two full texts were not found\(^{(14, 15)}\) and 139 studies did not meet the inclusion criteria. After assessment, 21 studies with 18 749 women were included in the review (Figure 1).\(^{(16-36)}\)

**Included studies**

Of the 21 included studies, six were from Latin America and the Caribbean (n=3470, 14.8% of the women),\(^{(18, 29, 30, 32, 35, 36)}\) nine from Europe (n=8459, 36.2% of the women),\(^{(16, 19-22, 24, 26, 28, 33)}\) and four from North America (n=8264, 35.3% of the women)\(^{(17, 23, 25, 34)}\). We only included one article from Oceania\(^{(31)}\) and one from Asia\(^{(27)}\) (n=3206, 13.7% of the women). Regarding design, two of the included studies were
prospective cohorts (19, 33) and the rest (n=19) were cross-sectional studies (Table 1) (16-18, 20-32, 34-36).

Eighteen studies reported details of the tools used for self-reported and directly measured weight and height of participants (16-22, 24-29, 31, 33-36). For directly measured data, height was most commonly measured by stadiometer, anthropometer, or some type of measuring tape or ruler with an error between 0.1 to 0.5 cm, while weight was measured by a variety of scales with an error of 0.1kg (balance beam, digital, or portable). Twelve of the 21 studies used self-administered on-site questionnaires as the self-reported method (17, 20-22, 24-27, 33-36). Three studies gathered information in an online survey or via telephone (18, 19, 31), while three other studies performed an in-person interview to obtain this data (16, 28, 29). The remaining three studies did not report the type of methods used (23, 30, 32). All the studies obtained the self-reported value prior to the directly measured data (16-36).

From the included studies, nineteen reported mean value of self-reported and directly measured weight and height (16-21, 23-28, 30-36). Two studies only reported mean difference between methods, calculated as self-reported minus directly measured values (22, 29). Only two studies showed data by women’s BMI categories (28, 29).

Risk of bias Assessment

The risk of bias assessment found six studies with high risk of bias in the selection of participants (29.0%) (24, 25, 28, 30, 33, 36) and two studies with the control of confounders (9.50%) (Supporting Information Table S1) (30, 36).

Weight

According to the meta-analysis, we found that in the overall sample, the mean difference between self-reported and direct measured data for women’s weight was -0.94kg (19 studies; 16 578 participants; 95%CI; -1.17, -0.71kg; p<0.0001; I^2=0%) (16-21, 23-28, 30-36). When analyzed by age subgroups, we found that self-reported weight was lower than directly measured weight in women between 12 and 18 years (-1.05 [95%CI; -1.32, -0.78]; p<0.0001; I^2=0%) and in women between 19 and 35 years (-1.04 [95%CI; -1.86, -0.21]; p=0.001, I^2=30%). However, in women from 36 to 49 years, there was no statistically significant difference between methods (-0.26 [95%CI; -0.99, 0.44]; p=0.49; I^2=0%) (Figure 2- Panel A).
The results by region were in the same direction in all three meta-analyzed regions (Figure 3 - Panel A). The difference between self-reported and directly measured weight was -1.14 kg (95% CI: -1.67, -0.61; p<0.0001; I²=0%) in Latin America and the Caribbean, -1.02 kg (95% CI: -1.68, -0.37; p=0.002, I²=55%) in Europe; and -1.51 kg (95% CI: -2.53, -0.48; p=0.004; I²=0%) in North America. We only found one study for Asia and one for Oceania, and they were not included in the meta-analysis (24, 28).

In the analysis by time of data collection we found that if obtained within the same day there was a -0.97 kg (95% CI: -1.37, -0.57; P<0.001; I²=15%) difference between self-reported and directly measured weight. No statistically significant difference was found when collected on separate days (-1.64 kg [95% CI: -4.30, 1.03]; p=0.23; I²=0%) (Figure 4 - Panel A).

We also evaluated the influence of the self-reported method used when compared to directly measured data (Supporting Information Figure S1 - Panel A). The analysis suggested that there was a negative difference if the information was gathered through a long-distance survey (-1.46 kg [95% CI: -2.27, -0.64]; p=0.0004; I²=0%) or a self-administered questionnaire on-site (-1.14 kg [95% CI: -1.79, -0.48]; p=0.006; I²=54%). The difference was lower when gathered during an in-person interview (-0.27 kg [95% CI: -0.80, 0.25]; p=0.74; I²=46%).

Only two studies classified their population according to BMI status of participants. We found that those who were overweight underestimated their weight by -0.39 kg ([95% CI; -0.59, -0.19]; p=0.0001; I²=0%) (28, 29). We found no statistically significant results because of the high heterogeneity between the studies for the other three BMI categories (underweight, normal weight or obesity) (Figure 5 - Panel A).

As mentioned previously, two studies only reported mean difference between methods, without specifying mean self-reported weight and mean direct measured weight (22, 29). One study included women between 15 and 18 years (22). The second study divided its population into three subgroups: 20 to 24 years, 25 to 34 years, and 35 to 44 years (29). We performed a separate analysis to evaluate if the results of these studies were consistent with the directionality of the findings previously presented. We meta-analyzed these population subgroups and found an I² of 80% (Supporting Information Figure S2 – Panel A).
According to the meta-analysis, we found that in the overall sample, the mean difference between self-reported and directly measured data for women’s height was 0.36 cm (18 studies; 13,744 participants; [95% CI; 0.20, 0.51]; p < 0.0001; I² = 35%). When analyzed by age, we found that self-reported height was higher than directly measured height in all subgroups (Figure 2 - Panel B). In the subgroup of age between 12 and 18 years the mean difference was 0.24 cm ([95% CI; 0.04, 0.44]; p = 0.02; I² = 54%); in the group between 19 and 35 years the mean difference was 0.57 cm ([95% CI; 0.25, 0.89]; p < 0.001; I² = 0%); and in the subgroup women from 36 to 49 years the mean difference was 0.50 cm ([95% CI; 0.09, 0.91]; p = 0.02; I² = 0%).

The analysis by region showed a significant mean difference between self-reported height and directly measured height of 0.63 cm ([95% CI; 0.41, 0.85]; p < 0.0001; I² = 0%) in Europe. No statistical differences were found in the Americas (North America: -0.62 cm [95% CI; -1.30, 0.06]; p = 0.08; I² = 0%), or Latin America and the Caribbean: 0.43 cm ([95% CI; -0.07, 0.92]; p = 0.09; I² = 11%) (Figure 3 - Panel B). Two studies were excluded from the meta-analysis because each was the only reference from their region (27, 31).

In the analysis by time of data collection we found that if obtained within the same day, the difference was 0.53 cm (95% CI; 0.20, 0.85; p = 0.001; I² = 43%). No significant difference was found when obtained on separate days (0.60 cm [95% CI; -0.83, 2.04]; p = 0.41; I² = 0%) (Figure 4 - Panel B).

We also evaluated the influence of the specific self-reported method used when compared to directly measured height data (Supporting Information Figure S1 - Panel B). The analysis showed a significant difference if the information was gathered through a long-distance survey (0.55 cm [95% CI; 0.00, 1.09]; p = 0.05; I² = 0%) or in an in-person interview (0.65 cm [95% CI; 0.28, 1.02]; p = 0.0005; I² = 38%). No statistically significant difference was found when the data was gathered through an on-site self-administered questionnaire (0.10 cm [95% CI; -0.68, 0.47]; p = 0.72; I² = 70%).

The high heterogeneity found between studies in the subgroup analysis based on women’s BMI categories prevented us from obtaining an estimate difference between self-reported and directly measured height (28, 29) (Figure 5 - Panel B).
The separate analysis for the two studies (27, 31) reporting only the mean difference between methods found that the results were consistent with the findings previously presented, and showing self-reported height higher than direct measurements. There was no heterogeneity between studies ($I^2=0\%$) (Supporting Information Figure S2 – Panel B).

**Discussion**

The results of this review showed an overall underestimation of weight (-0.94 kg) and an overestimation of height (+0.36 cm) when comparing self-reported to directly measured values in women of reproductive age.

In the pre-specified subgroup analyses, the findings remained consistent. We found that women aged 12 to 35 years under-reported their weight by 0.78 kg to 1.17 kg. Older women also under-reported their weight, but the difference was not statistically significant in this age group. The underestimation of self-reported weight was found throughout all studied regions reaching a mean difference between self-reported and direct measured weight as high as 1.50 kg in North America. Few studies presented data in overweight women; the results on weight were similar to normal weight women. It was not possible to estimate the differences by underweight or obese subgroups.

We found that the underestimation of weight persisted if data was collected through an on-site self-administered questionnaire or a long-distance survey (online or via telephone); however, when self-reported data was collected by on-site in-person interviews, this underestimation was lower and not statistically significant.

Regarding height, the results showed a consistent overestimation throughout all age groups. These findings were also observed in studies from Europe and North America, but not in those from Latin America and the Caribbean. The overestimation in height persisted when collected through an on-site in-person interview or long distance survey; however, there was no statistically significant difference with directly measured values when using an on-site self-administered questionnaire. Our results confirmed the data published by Gorber et al (37) in the general population and updated by Engstrom et al
(38) results from 2002 to 2015. All these studies showed an underestimation of weight and overestimation of height. In our study, as well as that of Gorber and Engstrom’s reviews, the standard deviations were large in all included studies, suggesting significant variability between women in the accuracy of self-reported height and weight measurements.

One of the authors carried out 3 pilot tests of search strategies MEDLINE to explore the potential sensitivity and specificity of the electronic searches. We assume that the risk of publication bias is low (Supporting Information Table S1). Poor reporting of studies was the major problem found when assessing the risk of bias of included studies. To address this limitation, we contacted the primary authors of those articles with missing data.

Although large numbers of women have been studied, Asia and Oceania had little representation in the final selection of studies, with only one article from each region. Moreover, some of the included studies had a relatively small sample size.

One limitation of our review was the high heterogeneity found when the meta-analysis combined studies reporting means and those reporting only mean differences. To compensate for this limitation, we presented a separate meta-analysis for those studies reporting only a mean difference. The main strength of this review is that, by restricting the population’s inclusion criteria, we could control for the large heterogeneity between studies and calculate a reliable summary estimate that quantifies the bias that occurs when using self-reported weight and height data for women in reproductive age.

Finally, we observed that there is a difference in relation to the degree of significance in some analyzes. In this regard, the limited number of studies in some sub analysis challenged the interpretation of the results.

Conclusions

This review presents the difference of using self-reported weight and height compared to direct measurements in women of reproductive age with no eating disorders or conditions that may confound the comparison. The population selected in this study allowed us to reduce the heterogeneity between studies and to achieve a summary estimate of possible bias.
Self-reported maternal weight and height are broadly used, particularly in situations where even basic anthropometric measurements cannot be taken. Self-reported measures are used in clinical practice and in studies that relate them with pregnancy outcomes. This review shows a low bias in the estimation of weight and height using self-reported measures; for example, the BMI of a woman with a weight of 50kg and a height of 1.65mts, would differ by 2.36% (95%CI: 2.07%, 2.58 %) if measured using self-reported data. Our interpretation is that self-reported weight and height in women of reproductive age is a measure that closely estimates the real values and can be used as proxy of real values both in clinical and research evaluation.

**Funding:** The study was funded by Institute for Clinical Effectiveness and Health Policy (IECS), Buenos Aires, Argentina. The funding source had no role in study design, data collection, data analysis, data interpretation or preparation of the manuscript.

**References**


25. Leatherdale ST, Laxer RE. Reliability and validity of the weight status and dietary intake measures in the COMPASS questionnaire: are the self-reported measures of body mass index (BMI) and Canada&amp;#039;s food guide servings robust? Int J Behav Nutr Phys Act. 2013:42.


Legends

Figure 1. Flow Chart of screening and selection of studies.

Figure 2. Forest plot of self-reported vs. direct measured weight (panel A) and height (panel B) (mean difference), by age group.

Figure 3. Forest plot of self-reported vs. direct measured weight (panel A) and height (panel B) (mean difference), by region.

Figure 4. Forest plot of self-reported vs. direct measured weight (panel A) and height (panel B) (mean difference), by time between self-reported and direct measured.

Figure 5. Forest plot of self-reported vs. direct measured weight (panel A) and height (panel B) (mean difference), by Body Mass Index (BMI).

Appendix S1. Supplementary Methods: MEDLINE search strategy. We include search terms (Mesh and others) and description of how they were combined.

Table S1. Assessment of risk of bias by article. The findings of the present study that the risk of bias assessment found that there was a high risk of bias in the selection of participants in six studies and in the control of confounder in two.

Figure S1. Forest plot of self-reported vs. direct measured weight (panel A) and height (panel B) (mean difference), by self-report method. The findings study suggested that there was a negative difference if the information is gathered in through a long-distance survey (-1.46kg [95%CI; -2.27, -0.64]; p=0.0004; I2=0%) or in a self-administered
questionnaire on-site (-1.14kg [95%CI -1.79, -0.48]; p=0.006; I²=54%) for weight. For 
Height a significant difference if the information is gathered through a long-distance 
survey (0.55cm [95%CI 0.00, 1.09]; p=0.05; I²=0%) or in an in-person interview 
(0.65cm [95%CI; 0.28, 1.02]; p=0.0005; I²=38%). No important difference was found 
when the data is gathered through an on-site self-administered questionnaire.

Figure S2. Forest plot of mean difference between self-reported and direct measured 
weight (panel A) and height (panel B) in studies that only reported mean differences. 
The study founding two studies only reported mean difference between methods, 
without specifying mean self-reported weight and mean direct measured weight. We 
found that, self-reported height was higher than direct measured height
<table>
<thead>
<tr>
<th>First Author - Year</th>
<th>Country</th>
<th>Type of study</th>
<th>Population characteristic</th>
<th>Database analysed</th>
<th>Age range or mean +SD</th>
<th>Sample size</th>
<th>Reported outcome</th>
<th>SR* method</th>
<th>Time Lag between SR* and DM**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brettschneider 2011</td>
<td>Germany</td>
<td>Cross sectional</td>
<td>General population</td>
<td>KiGGS</td>
<td>14-17</td>
<td>948</td>
<td>MW,MH,MBMI,DW,DH,DBMI</td>
<td>Personal interview</td>
<td>Same day</td>
</tr>
<tr>
<td>Brunner 2007</td>
<td>USA</td>
<td>Cross sectional</td>
<td>General population</td>
<td>CHIC</td>
<td>26-35</td>
<td>138</td>
<td>MW,MH,MBMI,DW,DH,DBMI</td>
<td>Self-administered survey</td>
<td>Same day</td>
</tr>
<tr>
<td>Carvalho 2014</td>
<td>Brazil</td>
<td>Cross sectional</td>
<td>General population</td>
<td>ISA-Capital</td>
<td>12-19</td>
<td>32</td>
<td>MW,MH,MBMI</td>
<td>Long distance survey (Telephonic)</td>
<td>Different Days (non-specified)</td>
</tr>
<tr>
<td>Ekstrom 2015</td>
<td>Stocolm</td>
<td>Prospective cohort</td>
<td>General population</td>
<td>-</td>
<td>16.5+0.4</td>
<td>889</td>
<td>MW,MH,MBMI,DW,DH,DBMI</td>
<td>Long distance survey (Online)</td>
<td>NR</td>
</tr>
<tr>
<td>Elgar 2005</td>
<td>Wale</td>
<td>Cross sectional</td>
<td>High school students</td>
<td>HBSC Study</td>
<td>15-17</td>
<td>211</td>
<td>MW,MH,MBMI,DW,DH,DBMI</td>
<td>Self-administered survey</td>
<td>Same day</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Country</td>
<td>Study Design</td>
<td>Sample Description</td>
<td>Method</td>
<td>Age Range</td>
<td>Sample Size</td>
<td>Body Measurement Indices</td>
<td>Data Collection Format</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>--------------</td>
<td>--------------------</td>
<td>--------</td>
<td>-----------</td>
<td>------------</td>
<td>------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Fonseca 2009</td>
<td>2009</td>
<td>Portugal</td>
<td>Cross sectional</td>
<td>High school students</td>
<td>HBSC Study</td>
<td>14 +1.8</td>
<td>233</td>
<td>MW, MH, MBMI, DW, DH, DBMI</td>
<td>Self-administered survey</td>
</tr>
<tr>
<td>Galán 2001</td>
<td>2001</td>
<td>Spain</td>
<td>Cross sectional</td>
<td>High school students</td>
<td>-</td>
<td>15-18</td>
<td>1810</td>
<td>DW, DH, DBMI</td>
<td>Self-administered survey</td>
</tr>
<tr>
<td>Himes 2001</td>
<td>2001</td>
<td>USA</td>
<td>Cross sectional</td>
<td>General population</td>
<td>NHANES III</td>
<td>12-16</td>
<td>876</td>
<td>MW, MH, MBMI, DW, DH, DBMI</td>
<td>NR</td>
</tr>
<tr>
<td>Larsen 2007</td>
<td>2007</td>
<td>Netherlands</td>
<td>Cross sectional</td>
<td>University students</td>
<td>-</td>
<td>20.9+2.40</td>
<td>209</td>
<td>MW, MH, MBMI</td>
<td>Questionnaire on-site</td>
</tr>
<tr>
<td>Leatherdale 2013</td>
<td>2013</td>
<td>Canada</td>
<td>Cross sectional</td>
<td>High school students</td>
<td>-</td>
<td>14-15</td>
<td>65</td>
<td>MW, MH, MBMI, DW, DH, DBMI</td>
<td>Self-administered survey</td>
</tr>
<tr>
<td>Legleye 2014</td>
<td>2014</td>
<td>France</td>
<td>Cross sectional</td>
<td>General population</td>
<td>ESCAPAD</td>
<td>17-18</td>
<td>140</td>
<td>MW, MH, MBMI, DW, DH, DBMI</td>
<td>Self-administered survey</td>
</tr>
<tr>
<td>Lo 2011</td>
<td>2011</td>
<td>China</td>
<td>Cross sectional</td>
<td>High school students</td>
<td>HKSOS project</td>
<td>16.29 +0.98</td>
<td>1275</td>
<td>MW, MH, MBMI</td>
<td>Questionnaire on-site</td>
</tr>
<tr>
<td>Marrodan 2013</td>
<td>2013</td>
<td>Spain</td>
<td>Cross sectional</td>
<td>General population</td>
<td>-</td>
<td>18-24</td>
<td>181</td>
<td>MW, MH, MBMI</td>
<td>Personal interview</td>
</tr>
<tr>
<td>Study</td>
<td>Location</td>
<td>Study Design</td>
<td>Population</td>
<td>Age Range</td>
<td>Sample Size</td>
<td>Measurements</td>
<td>Data Collection Method</td>
<td>Data Collection Timeframe</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>--------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------</td>
<td>--------------</td>
<td>------------------------</td>
<td>--------------------------</td>
<td></td>
</tr>
<tr>
<td>Peixoto 2006&lt;sup&gt;(29)&lt;/sup&gt;</td>
<td>Brazil</td>
<td>Cross sectional</td>
<td>General population</td>
<td>35-44</td>
<td>1876</td>
<td>DW,DH,DBMI</td>
<td>Personal interview</td>
<td>Same day</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20-24</td>
<td>97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25-34</td>
<td>184</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35-44</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregolato 2009&lt;sup&gt;(30)&lt;/sup&gt;</td>
<td>Brazil</td>
<td>Cross sectional</td>
<td>University students</td>
<td>28.3+11</td>
<td>549</td>
<td>MW,MH, MBMI</td>
<td>NR</td>
<td>Same day</td>
<td></td>
</tr>
<tr>
<td>Pursey 2014&lt;sup&gt;(31)&lt;/sup&gt;</td>
<td>Australia</td>
<td>Cross sectional</td>
<td>General population</td>
<td>18-35</td>
<td>93</td>
<td>MW, MH</td>
<td>Long distance survey (Online)</td>
<td>Different Days (&lt;1 month)</td>
<td></td>
</tr>
<tr>
<td>Rodrigues 2013&lt;sup&gt;(32)&lt;/sup&gt;</td>
<td>Brazil</td>
<td>Cross sectional</td>
<td>High school students</td>
<td>14-19</td>
<td>40</td>
<td>MW,MH,MBMI,DW,DH,DBMI</td>
<td>NR</td>
<td>Same day</td>
<td></td>
</tr>
<tr>
<td>Savane 2013&lt;sup&gt;(33)&lt;/sup&gt;</td>
<td>Spain</td>
<td>Prospective cohort</td>
<td>University students</td>
<td>18-37</td>
<td>476</td>
<td>MW,MH,MBMI,DW,DH,DBMI</td>
<td>Self-administered survey</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Shin 2014&lt;sup&gt;(6)&lt;/sup&gt;</td>
<td>USA</td>
<td>Cross</td>
<td>General</td>
<td>NHANES 8</td>
<td>16-25</td>
<td>1252</td>
<td>MW</td>
<td>Self-</td>
<td>Same day</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Type</td>
<td>Age Range</td>
<td>Population</td>
<td>Measurement</td>
<td>Survey Administration</td>
<td>Survey Timing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>---------</td>
<td>----------------</td>
<td>-----------</td>
<td>------------</td>
<td>--------------</td>
<td>-----------------------</td>
<td>---------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unikel Santocini 2009(35)</td>
<td>Mexico</td>
<td>Cross sectional</td>
<td>26-35</td>
<td>592</td>
<td>-</td>
<td>MW, MH, DW, DH</td>
<td>Self-administered survey</td>
<td>Same day</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>36-44</td>
<td>599</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitale 2013(36)</td>
<td>Argentina</td>
<td>Cross sectional</td>
<td>15-18</td>
<td>61</td>
<td>MW, MH, MBMI</td>
<td>Self-administered survey</td>
<td>Same day</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Flow Chart of screening and selection of studies.

Records identified through database searching (n = 2017)
- Pubmed: 1347
- Cochrane library: 172
- EMBASE: 323
- LILACS: 56
- CINAHL: 119

Records after duplicates removed (n = 1638)

Records screened (n = 1638)

Full-text articles assessed for eligibility (n = 162)

Records excluded by title and abstract (n = 1476)

Full-text articles excluded, with reasons (n = 141)
- Exclusion Criteria: 69
- No Outcomes: 32
- Inadequate Design: 20
- Duplicated: 8
- Others: 6

Studies included (n = 21)

Studies included in quantitative synthesis (n = 21)
Figure 3. Forest plot of self-reported vs. direct measured weight (panel A) and height (panel B) (mean difference), by region.

**Panel A: self-reported vs. direct measured weight**

**Panel B: self-reported vs. direct measured height**
Figure 4. Forest plot of self-reported vs. direct measured weight (panel A) and height (panel B) (mean difference), by time between self-reported and direct measured.

Panel A: self-reported vs. direct measured weight

Panel B: self-reported vs. direct measured height
Figure 5. Forest plot of self-reported vs. direct measured weight (panel A) and height (panel B) (mean difference), by Body Mass Index (BMI).

### Panel A: self-reported vs. direct measured weight

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Mean Difference (SE)</th>
<th>Weight</th>
<th>Mean Difference (SE)</th>
<th>Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5.1 BMI: Low weight</td>
<td>0.61 (0.3)</td>
<td>0.10 (0.8)</td>
<td>0.61 (0.3)</td>
<td>0.10 (0.8)</td>
</tr>
<tr>
<td>1.5.1 BMI: Normal weight</td>
<td>3.37 (0.4)</td>
<td>0.16 (1.0)</td>
<td>3.37 (0.4)</td>
<td>0.16 (1.0)</td>
</tr>
<tr>
<td>1.5.1 BMI: Overweight</td>
<td>4.32 (0.6)</td>
<td>0.18 (1.2)</td>
<td>4.32 (0.6)</td>
<td>0.18 (1.2)</td>
</tr>
<tr>
<td>1.5.1 BMI: Obese</td>
<td>5.52 (0.7)</td>
<td>0.20 (1.4)</td>
<td>5.52 (0.7)</td>
<td>0.20 (1.4)</td>
</tr>
</tbody>
</table>

### Panel B: self-reported vs. direct measured height

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Mean Difference (SE)</th>
<th>Weight</th>
<th>Mean Difference (SE)</th>
<th>Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5.1 BMI: Low weight</td>
<td>3.08 (0.4)</td>
<td>0.15 (1.1)</td>
<td>3.08 (0.4)</td>
<td>0.15 (1.1)</td>
</tr>
<tr>
<td>2.5.1 BMI: Normal weight</td>
<td>3.72 (0.6)</td>
<td>0.18 (1.2)</td>
<td>3.72 (0.6)</td>
<td>0.18 (1.2)</td>
</tr>
<tr>
<td>2.5.1 BMI: Overweight</td>
<td>4.34 (0.7)</td>
<td>0.20 (1.4)</td>
<td>4.34 (0.7)</td>
<td>0.20 (1.4)</td>
</tr>
<tr>
<td>2.5.1 BMI: Obese</td>
<td>5.52 (0.8)</td>
<td>0.20 (1.4)</td>
<td>5.52 (0.8)</td>
<td>0.20 (1.4)</td>
</tr>
</tbody>
</table>