Public Health Nutrition: 5(4), 561–565 DOI: 10.1079/PHN2001322

Validity of self-reported height and weight in 4808 EPIC–Oxford participants

Elizabeth A Spencer*, Paul N Appleby, Gwyneth K Davey and Timothy J Key
Imperial Cancer Research Fund Cancer Epidemiology Unit, University of Oxford, Gibson Building, The Radcliffe Infirmary, Oxford OX2 6HE, UK

Submitted 13 August 2001: Accepted 14 November 2001

Abstract

Objective: To assess the validity of self-reported height and weight by comparison with measured height and weight in a sample of middle-aged men and women, and to determine the extent of misclassification of body mass index (BMI) arising from differences between self-reported and measured values.
Design: Analysis of self-reported and measured height and weight data from participants in the Oxford cohort of the European Prospective Investigation into Cancer and Nutrition (EPIC–Oxford).
Subjects: Four thousand eight hundred and eight British men and women aged 35–76 years.
Results: Spearman rank correlations between self-reported and measured height, weight and BMI were high ($r > 0.9$, $P < 0.0001$). Height was overestimated by a mean of 1.23 (95% confidence interval (CI) 1.11–1.34) cm in men and 0.60 (0.51–0.70) cm in women; the extent of overestimation was greater in older men and women, shorter men and heavier women. Weight was underestimated by a mean of 1.85 (1.72–1.99) kg in men and 1.40 (1.31–1.49) kg in women; the extent of underestimation was greater in heavier men and women, but did not vary with age or height. Using standard categories of BMI, 22.4% of men and 18.0% of women were classified incorrectly based on self-reported height and weight. After correcting the self-reported values using predictive equations derived from a 10% sample of subjects, misclassification decreased to 15.2% in men and 13.8% in women.
Conclusions: Self-reported height and weight data are valid for identifying relationships in epidemiological studies. In analyses where anthropometric factors are the primary variables of interest, measurements in a representative sample of the study population can be used to improve the accuracy of estimates of height, weight and BMI.

Keywords
Anthropometry
Height
Weight
Body mass index
Validity

Height and weight are of interest in epidemiological studies both as primary exposures and as potential confounding variables. In large studies, height and weight data are often collected by self-report. Body mass index (BMI), calculated from height and weight, is a popular and useful measure of relative weight and is often used to categorise study participants as underweight, normal weight, overweight or obese (usually defined as BMI < 20, 20–24.9, 25–29.9 and 30 kg m^{-2} and above, respectively). Of particular interest is the identification of obesity, responsible for considerable morbidity and early mortality[1].

Previous studies have reported that height and weight data are reported with acceptable accuracy[2–4]. Weight tends to be underestimated, more so by women, and height tends to be overestimated, more so by men[5–7]. However, certain population subgroups such as heavier people and older people tend to estimate their height and/or weight less accurately than others[5,8].

This study compares self-reported with measured height, weight and BMI data from participants in the Oxford cohort of the European Prospective Investigation into Cancer and Nutrition (EPIC–Oxford). It provides further evidence on factors influencing variations in the accuracy of self-reported height and weight, and examines the effect of reporting errors on the classification of subjects into standard BMI categories.

Subjects and methods

Between 1993 and 1999, 5140 middle-aged EPIC–Oxford participants who were recruited via general medical practices in England completed a diet and lifestyle questionnaire. In response to the questions 'How tall are you' and 'How much do you weigh?' subjects recorded their height and weight, in either imperial or metric units. The questionnaires were optically scanned and stored on computer. Height was rounded to the nearest cm, and

*Corresponding author: Email l.spencer@icrf.icnet.uk

weight was rounded to the nearest 0.1 kg. Within a few weeks of completing the questionnaire, the participants kept an appointment at the general practice and a nurse measured their height and weight, following a standard protocol. Height was measured without shoes and recorded to the nearest cm, weight was measured with light clothing and recorded to the nearest 0.1 kg. Discrepancies between self-reported and measured height of over 10 cm and between self-reported and measured weight of over 5 kg were checked for data-entry errors. Standard EPIC–Oxford exclusion criteria were applied to both self-reported and measured data as follows: for men, height under 100 cm or over 213 cm (one exclusion) and weight under 30 kg (no exclusions); for women, height under 100 cm or over 198 cm (two exclusions) and weight under 20 kg (two exclusions). For both men and women, measurements giving a body mass index below 15 kg m^{-2} or above 60 kg m^{-2} were excluded (one man and two women excluded). Body mass index (BMI) was calculated as weight in kg divided by the square of height in metres. BMI calculated from self-reported height and weight data is referred to throughout this paper as self-reported BMI. Standard categories of BMI were used to characterise participants as underweight (BMI < 20 kg m^{-2}), normal weight (20–24.9 kg m^{-2}), overweight (25–29.9 kg m^{-2}) or obese (30 kg m^{-2} and above). After exclusions owing to recording error, missing or extreme values as defined above, data were available for 1870 men and 2938 women. Age in years at recruitment ranged from 37 to 74 in men with a median of 55, and in women age ranged from 35 to 76 with a median of 52. Ninety-nine per cent of participants were white and 20% had a university degree or equivalent.

Data were analysed separately for men and women. Spearman rank correlation coefficients between self-reported and measured height, self-reported and measured weight and self-reported and measured BMI were calculated, and Student's paired samples *t*-test was used to compare the means of self-reported and measured values. Analysis of variance was used to examine how the differences (self-reported minus measured values) varied with sex-specific quartile of measured height and weight, with standard categories of BMI as above, and with age at recruitment (35–49, 50–59 and 60–76 years). *F*-tests were used to assess the statistical significance of the heterogeneity in the mean differences across groups. We also cross-tabulated self-reported BMI with measured BMI, both in the standard categories and in quartiles, in order to determine the extent of misclassification of BMI that would arise from the use of self-reported height and weight.

To simulate the effect of measuring height and weight in a random sample of subjects from an epidemiological cohort, and using the measured values to correct for biases in the self-reported values, we selected a random sample of 10% of subjects from our study population. Simple regression equations were derived from the sample data for measured height and measured weight in men and women separately, each as a function of age and self-reported height and weight. The equations were then used to predict measured height and weight from age and self-reported height and weight in the remaining 90% of subjects, and the predicted values used to calculate a 'predicted' BMI. We then cross-tabulated predicted and measured BMI in standard categories in the full cohort. Statistical analyses were performed using the Stata statistical package[9].

Table 1 Mean (standard deviation (SD)) self-reported and measured anthropometric measurements and their differences

Variable	Self-reported	Measured	Difference
Men			
Height (cm)	177.20 (6.90)	175.98 (6.90)	1.23 (2.57)
Weight (kg)	80.58 (11.94)	82.46 (12.41)	−1.85 (2.92)
BMI (kg m^{-2})	25.64 (3.34)	26.60 (3.54)	−0.96 (1.24)
Women			
Height (cm)	163.10 (6.56)	162.50 (6.18)	0.60 (2.68)
Weight (kg)	66.48 (12.06)	67.88 (12.42)	−1.40 (2.45)
BMI (kg m^{-2})	24.99 (4.34)	25.71 (4.57)	−0.72 (1.27)

Results

Based on their measured height and weight, 0.9% of men were underweight, 33.4% were of normal weight, 50.8% were overweight and 14.9% were obese. In women, 4.4%

Table 2 Mean differences between self-reported and measured height by quartile of measured height, age group and quartile of measured weight

Category	*n*	Difference (95% CI)
Men		
≤ 172 cm	553	1.82 (1.58, 2.07)
173–176 cm	428	1.05 (0.81, 1.30)
177–181 cm	509	1.25 (1.07, 1.42)
182+ cm	380	0.53 (0.27, 0.78)
35–49 years	600	0.80 (0.60, 1.01)
50–59 years	629	1.15 (0.95, 1.35)
60+ years	641	1.69 (1.50, 1.89)
≤ 74.0 kg	483	1.35 (1.09, 1.60)
74.1–81.0 kg	466	1.05 (0.84, 1.26)
81.1–89.0 kg	458	1.20 (0.96, 1.45)
89.1+ kg	463	1.30 (1.08, 1.52)
Women		
≤ 158 cm	760	0.82 (0.62, 1.01)
159–162 cm	737	0.52 (0.36, 0.69)
163–166 cm	713	0.44 (0.25, 0.64)
167+ cm	728	0.62 (0.46, 0.83)
35–49 years	1168	0.15 (−0.01, 0.30)
50–59 years	1039	0.65 (0.50, 0.81)
60+ years	731	1.26 (1.07, 1.46)
≤ 60.0 kg	823	0.20 (0.00, 0.39)
60.1–65.8 kg	646	0.40 (0.22, 0.59)
65.9–74.0 kg	774	0.77 (0.60, 0.94)
74.1+ kg	695	1.09 (0.87, 1.30)

were underweight, 47.8% were of normal weight, 33.1% were overweight and 14.7% were obese.

Spearman rank correlations between self-reported and measured height, self-reported and measured weight and self-reported and measured BMI were all high ($r > 0.9$, $P < 0.0001$). Mean values of self-reported and measured height, weight and BMI and their differences are shown in Table 1. Both sexes overestimated their height, underestimated their weight and, consequently, underestimated their BMI. The differences were all highly statistically significant ($P < 0.0001$). Based on the 5th and 95th percentiles of the differences, 90% of the values of self-reported height lay in men within −2 and +5 cm of the measured value and in women within −3 and +5 cm. In men, 90% of the self-reported weights lay within −6.0 and +1.9 kg of the measured value and in women 90% lay within −5.0 and +1.5 kg of the measured value.

Mean differences in self-reported and measured height by quartile of measured height, age and quartile of measured weight are shown in Table 2. In men, the extent of overestimation of height decreased with increasing measured height ($P < 0.0001$), but did not vary significantly with measured weight. Conversely, in women, there was only a weak association between the extent of overestimation of height and measured height ($P < 0.05$), but height overestimation increased significantly with measured weight ($P < 0.0001$). There was a significant effect of age on the extent of overestimation of height in both men and women, with greater overestimation at older ages (both $P < 0.0001$).

Mean differences in self-reported and measured weight by quartile of measured weight, age and quartile of measured height are shown in Table 3. In both men and women, the extent of underestimation of weight increased with increasing measured weight (both $P < 0.0001$). However, the extent of underestimation of weight did not vary significantly with age or with measured height for either sex.

The mean differences between self-reported and measured BMI by measured BMI in standard categories and by age are shown in Table 4. Underweight men and women tended to overestimate their BMI, but there was no significant difference between self-reported and measured BMI in this category. However, men and women in the normal weight, overweight and obese categories significantly underestimated their BMI, the extent of underestimation increasing with increasing measured BMI (both $P < 0.0001$). In both men and women, the extent of underestimation of BMI increased with increasing age (both $P < 0.0001$).

Self-reported BMI and measured BMI are cross-tabulated in standard categories in Table 5. Based on their self-reported height and weight, 29.4% of underweight men and 16.3% of underweight women would have been classified as being of normal weight or above. Conversely, 40.9% of obese men and 27.0% of obese women would have been classified as being overweight at most. Men and women of normal weight were least likely to have been allocated to the wrong BMI category. Overall, 22.4% men and 18.0% of women would have been allocated to the wrong BMI category based on their self-reported height and weight. Cross-classification of BMI by quartiles of self-reported and measured BMI (rather than the standard categories) gave a similar percentage misclassification (results not shown).

When we compared predicted BMI (calculated using the predictive equations for height and weight derived from a random sample of the cohort) with measured BMI,

Table 3 Mean differences between self-reported and measured weight by quartile of measured weight, age group and quartile of measured height

Category	n	Difference (95% CI)
Men		
≤ 74.0 kg	483	−0.86 (−1.04, −0.670)
74.1–81.0 kg	466	−1.69 (−1.93, −1.44)
81.1–89.0 kg	458	−2.06 (−2.29, −1.83)
89.1+ kg	463	−2.86 (−3.21, −2.52)
35–49 years	600	−1.76 (−2.00, −1.53)
50–59 years	629	−1.88 (−2.12, −1.64)
60+ years	641	−1.91 (−2.13, −1.70)
≤ 172 cm	553	−1.67 (−1.89, −1.44)
173–176 cm	428	−1.78 (−2.07, −1.48)
177–181 cm	509	−1.98 (−2.23, −1.72)
182+ cm	380	−2.05 (−2.36, −1.75)
Women		
≤ 60.0 kg	823	−0.66 (−0.80, −0.52)
60.1–65.8 kg	646	−1.19 (−1.38, −1.01)
65.9–74.0 kg	774	−1.61 (−1.75, −1.46)
74.1+ kg	695	−2.23 (−2.45, −2.01)
35–49 years	1168	−1.35 (−1.48, −1.21)
50–59 years	1039	−1.41 (−1.55, −1.26)
60+ years	731	−1.47 (−1.66, −1.28)
≤ 158 cm	760	−1.35 (−1.51, −1.20)
159–162 cm	737	−1.43 (−1.59, −1.27)
163–166 cm	713	−1.36 (−1.56, −1.16)
167+ cm	728	−1.45 (−1.64, −1.26)

Table 4 Mean differences between self-reported and measured BMI by measured BMI category and age group

Category	n	Difference (95% CI)
Men		
< 20.0 kg m^{-2}	17	0.61 (−0.55, 1.77)
20.0–24.9 kg m^{-2}	625	−0.60 (−1.10, −0.93)
25.0–29.9 kg m^{-2}	949	−1.02 (−1.09, −0.95)
30.0+ kg m^{-2}	279	−1.66 (−1.87, −1.44)
35–49 years	600	−0.78 (−0.84, −0.73)
50–59 years	629	−0.95 (−1.01, −0.89)
60+ years	641	−1.13 (−1.18, −1.07)
Women		
< 20.0 kg m^{-2}	129	0.19 (−0.07, 0.45)
20.0–24.9 kg m^{-2}	1404	−0.44 (−0.50, −0.38)
25.0–29.9 kg m^{-2}	972	−0.96 (−1.04, −0.89)
30.0+ kg m^{-2}	433	−1.35 (−1.51, −1.20)
35–49 years	1168	−0.55 (−0.59, −0.50)
50–59 years	1039	−0.74 (−0.78, −0.70)
60+ years	731	−0.97 (−1.02, −0.92)

Table 5 Cross-tabulation of measured and self-reported BMI

Self-reported BMI (kg m^{-2})	Measured BMI (kg m^{-2})				Total
	<20.0	20.0–24.9	25.0–29.9	30.0+	
Men					
< 20.0	12 (70.6%)	17 (2.7%)	0 (0.0%)	0 (0.0%)	29
20.0–24.9	4 (23.5%)	589 (94.2%)	249 (26.2%)	2 (0.7%)	844
25.0–29.9	1 (5.9%)	18 (2.9%)	685 (72.2%)	112 (40.1%)	816
30.0+	0 (0.0%)	1 (0.2%)	15 (1.6%)	165 (59.1%)	181
Total	17 (100%)	625 (100%)	949 (100%)	279 (100%)	1870
Women					
< 20.0	108 (83.7%)	70 (5.0%)	1 (0.1%)	0 (0%)	179
20.0–24.9	20 (15.5%)	1292 (92.0%)	265 (27.3%)	2 (0.5%)	1579
25.0–29.9	1 (0.8%)	38 (2.7%)	694 (71.4%)	115 (26.6%)	848
30.0+	0 (0.0%)	4 (0.3%)	12 (1.2%)	316 (73.0%)	332
Total	129 (100%)	1404 (100%)	972 (100%)	433 (100%)	2938

rates of misclassification of BMI fell from 22.4% to 15.2% in men and from 18.0% to 13.8% in women.

Discussion

This study shows high rank correlations between self-reported and measured height and weight in a large sample of subjects from the EPIC–Oxford cohort, demonstrating that these self-reported data are valid for detecting associations between height and weight and disease in epidemiological studies.

However, this study has also highlighted systematic errors in self-reported height and weight and identified groups likely to make greater errors. Height was overestimated by both men and women, more so by men and especially by shorter men. Heavier women overestimated their height more than lighter women, and older men and women overestimated their height more than younger men and women. Weight was underestimated by both men and women, with a clear trend towards greater underestimation with increasing weight, although the extent to which weight was underestimated did not vary with age or with height.

In contrast to some previous studies[6,10], our study suggests that women report their weight more accurately on average than men. Because women also reported their height more accurately than men, the resulting discrepancy in BMI was smaller in women than in men.

Some of the differences noted may be due to genuine discrepancies between measured and self-reported values. For example, height may vary by up to 2.4 cm over the course of a day[11], whilst weight may vary from day to day, especially among dieters. In addition, clothes may weigh up to 1 kg and this may account for some of the apparent underestimation in self-reported weight, since self-weighing is likely to be done in minimal clothing. However, it is unlikely that these factors could account for all of the differences noted here.

Errors in reporting anthropometric data have been evaluated in a number of previous studies. Although there is some variation in the direction of the errors, the majority of studies report deviation towards a 'preferred' body size[12]. In particular, height is generally overestimated, with greater overestimation in shorter individuals, especially men[11], and weight is generally underestimated, with greater underestimation in heavier individuals[6], although weight may be overestimated by underweight men[7]. Increasing age has been shown to be associated with increasing error in reported height[5]. It is well known that height declines with age in later life[11], yet this height loss may not be perceived by the individual.

The magnitude of the differences between self-reported and measured values varies between studies. The mean differences between self-reported and measured height were less in our study than in some other studies[2,6] but greater than in others[4,5,7]. The mean difference between self-reported and measured weight in men was greater in our study than in other studies[2,3,7,13,14], but the mean difference for women in our study was lower than in three other studies[2,7,14]. Population characteristics such as age distribution, health status, motivation and variations in study protocol may account for these differences.

The errors in self-reported height and weight are compounded in the derived BMI variable and this is particularly important when subjects are classified into standard categories of BMI. As may be expected, BMI was underestimated in both men and women in all BMI categories except for those classified as underweight (BMI $< 20\,\mathrm{kg\,m^{-2}}$). The extent of underestimation of BMI increased from the normal BMI category through the overweight category to a maximum in the obese category (BMI of $30\,\mathrm{kg\,m^{-2}}$ and above). Men and women of normal weight (BMI $= 20–24.9\,\mathrm{kg\,m^{-2}}$) were least likely to be incorrectly allocated to another BMI category, and obese participants were most likely to be incorrectly classified. Older men and women underestimated their BMI more than younger men and women. Misclassification of overweight and obese subjects as belonging to a lower BMI category would bias relative risks of diseases associated with increasing BMI and, for this reason, relative risks associated with standard BMI categories where BMI is calculated from self-reported height and weight should be interpreted with caution.

The interpretation of errors in self-reported anthropometric data has varied from concluding that they are acceptably small[2–4] to noting a systematic bias[5,8] and suggesting that it may be necessary to make adjustment for this error in epidemiological studies relying on self-reported height[11]. We have shown that deriving simple predictive equations for measured height and weight using data from a random sample of the cohort and applying them to the whole cohort improves the accuracy of BMI estimates. An alternative approach would be to perform analyses by ranking study participants in quantiles of BMI calculated from self-reported height and weight; the 'true' mean or median value of BMI in each quantile can then be calculated from the measured values for a random sample of participants and used to quantify the association of BMI with disease risk.

Self-reported height and weight data have been shown here to be valid for identifying associations in epidemiological studies. In analyses where anthropometric factors are the primary variables of interest, measurements in a random sample of the study population can be used to improve the accuracy of estimates of height, weight and BMI.

Acknowledgements

We thank all participants in EPIC–Oxford and all scientists and clerical staff who have worked on these studies. EPIC is supported by the Europe Against Cancer Programme of the Commission of the European Communities. This analysis was supported by the Imperial Cancer Research Fund and the Medical Research Council.

References

1 National Audit Office. *Tackling Obesity in England*. London: The Stationery Office, 2001.
2 Stewart AL. The reliability and validity of self-reported weight and height. *J. Chronic Dis.* 1982; **35**(4): 295–309.
3 Stunkard AJ, Albaum JM. The accuracy of self-reported weights. *Am. J. Clin. Nutr.* 1981; **34**(8): 1593–9.
4 Weaver TW, Kushi LH, McGovern PG, Potter JD, Rich SS, King RA, Whitbeck J, Greenstein J, Sellers TA. Validation study of self-reported measures of fat distribution. *Int. J. Obes. Relat. Metab. Disorders* 1996; **20**(7): 644–50.
5 Nieto-Garcia FJ, Bush TL, Keyl PM. Body mass definitions of obesity: sensitivity and specificity using self-reported weight and height. *Epidemiology* 1990; **1**(2): 146–52.
6 Palta M, Prineas RJ, Berman R, Hannan P. Comparison of self-reported and measured height and weight. *Am. J. Epidemiol.* 1982; **115**(2): 223–30.
7 Pirie P, Jacobs D, Jeffery R, Hannan P. Distortion in self-reported height and weight data. *J. Am. Diet. Assoc.* 1981; **78**(6): 601–6.
8 Rowland ML. Self reported height and weight. *Am. J. Clin. Nutr.* 1990; **52**: 1125–33.
9 Stata Corporation. *Stata Statistical Software: Release 5.0*. College Station, TX: Stata Press, 1997.
10 Wing RR, Epstein LH, Ossip DJ, LaPorte RE. Reliability and validity of self-report and observers' estimates of relative weight. *Addict. Behav.* 1979; **4**(2): 133–40.
11 Giles E, Hutchinson DL. Stature- and age-related bias in self-reported stature. *J. Forensic Sci.* 1991; **36**(3): 765–80.
12 Ziebland S, Thorogood M, Fuller A, Muir J. Desire for the body normal: body image and discrepancies between self reported and measured height and weight in a British population. *J Epidemiol. Community Health* 1996; **50**(1): 105–6.
13 Millar WJ. Distribution of body weight and height: comparison of estimates based on self-reported and observed measures. *J. Epidemiol. Community Health* 1986; **40**(4): 319–23.
14 Rimm EB, Stampfer MJ, Colditz GA, Chute CG, Litin LB, Willett WC. Validity of self-reported waist and hip circumferences in men and women. *Epidemiology* 1990; **1**(6): 466–73.