How reliable is internet-based self-reported identity, socio-demographic and obesity measures in European adults?

Article

Accepted Version

Celis-Morales, C., Livingstone, K. M., Woolhead, C., Forster, H., O'Donovan, C. B., Macready, A., Fallaize, R., Marsaux, C. F. M., Tsirigoti, L., Efstathopoulou, E., Moschonis, G., Navas-Carretero, S., San-Cristobal, R., Kolossa, S., Klein, U. L., Hallmann, J., Godlewska, M., Surwiłło, A., Drevon, C. A., Bouwman, J., Grimaldi, K., Parnell, L. D., Manios, Y., Traczyk, I., Gibney, E. R., Brennan, L., Walsh, M. C., Lovegrove, J., Martinez, J. A., Daniel, H., Saris, W. H. M., Gibney, M. and Mathers, J. C. (2015) How reliable is internet-based self-reported identity, socio-demographic and obesity measures in European adults? Genes and Nutrition, 10 (5). pp. 476-487. ISSN 1555-8932 doi: https://doi.org/10.1007/s12263-015-0476-0 Available at http://centaur.reading.ac.uk/40725/

How reliable is internet-based self-reported identity, socio-demographic and obesity measures in European adults?

Article

Accepted Version

VALIDATION OF INTERNET-BASED SELF-REPORTED ANTHROPOMETRIC, DEMOGRAPHIC DATA AND PARTICIPANT IDENTITY IN THE FOOD4ME PAN-EUROPEAN STUDY

RUNING TITLE – Validation of internet –based self-reported data

AUTHOR NAMES

Carlos Celis-Morales[1*], Katherine M Livingstone[1*], Clara Woolhead[2], Hannah Forster[2],Clare B O'Donovan[2], Anna L Macready[3], Rosalind Fallaize[3], Cyril F M Marsaux[4], Lydia Tsirigoti[5], Eirini Efstathopoulou[5], George Moschonis[5], Santiago Navas-Carretero[6], Rodrigo San-Cristobal[6], Silvia Kolossa[7], Ulla L Klein[7], Jacqueline Hallmann[7], Magdalena Godlewska[8], Agnieszka Surwiłło[8], Christian A Drevon[9], Jildau Bouwman[10], Keith Grimaldi[11], Laurence D Parnell[12], Yannis Manios[5], Iwona Traczyk[8], Eileen R Gibney[2], Lorraine Brennan[2], Marianne C Walsh[2], Julie A Lovegrove[3], J Alfredo Martinez[6], Hannelore Daniel[7], Wim H M Saris[4], Mike Gibney[2], John C Mathers[1], on behalf of the Food4Me Study.

AUTHOR AFFILIATIONS

[1] Human Nutrition Research Centre, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK.

[2] UCD Institute of Food and Health, University College Dublin, Belfield, Dublin 4, Republic of Ireland.

[3] Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research, University of Reading, Reading, UK.

[4] Department of Human Biology, NUTRIM School for Nutrition, Toxicology and Metabolism, Maastricht University Medical Centre, Maastricht, The Netherlands

[5]Department of Nutrition and Dietetics, Harokopio University, Athens, Greece.

[6]Department of Nutrition, Food Science and Physiology, University of Navarra; CIBER Fisiopatología Obesidad y Nutrición (CIBERobn), Instituto de Salud Carlos III, Spain (SN-C & JAM)

[7] ZIEL Research Center of Nutrition and Food Sciences, Biochemistry Unit, Technische Universität München, Germany.

[8] National Food & Nutrition Institute (IZZ), Poland.

[9] Department of Nutrition, Faculty of Medicine, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway.

[10] TNO, Microbiology and Systems Biology Group, Zeist, the Netherlands.

[11]Eurogenetica Ltd, 7 Salisbury Road, Burnham-on-Sea, UK

[12]Nutrition and Genomics Laboratory, Jean Mayer US Department of Agriculture Human Nutrition Research Center on Aging at Tufts University, Boston, Massachusetts, USA

* These authors contributed equally to this work.

CORRESPONDING AUTHOR

Professor John C. Mathers

Human Nutrition Research Centre

Institute of Cellular Medicine

Newcastle University

Biomedical Research Building,

Campus for Ageing and Vitality

Newcastle upon Tyne

NE4 5PL

UK

john.mathers@newcastle.ac.uk

Tel: +44 (0) 1912081133

Fax: +44 (0) 1912081101

ABSTRACT

Purpose

In e-health intervention studies, there are concerns about the reliability of internet-based, self-reported (SR) data and about the potential for identity fraud. This study introduced and tested a novel procedure for assessing the validity of internet-based, SR identity and validated anthropometric and demographic data via measurements performed face-to-face in a validation study (VS).

Methods

Participants (n=140) from seven European countries, participating in the Food4Me intervention study which aimed to test the efficacy of personalised nutrition approaches delivered via the internet, were invited to take part in the VS. Participants visited a research centre in each country within two weeks of providing SR data via the internet. Participants received detailed instructions on how to perform each measurement. Individual's identity was checked visually and by repeated collection and analysis of buccal cell DNA for 33 genetic variants.

Results

Validation of identity using genomic information showed perfect concordance between SR and VS. Similar results were found for demographic data (age and sex verification). We observed strong Intra Class Correlation coefficients between SR and VS for anthropometric data (height 0.990, weight 0.994 and BMI 0.983). However, internet-based SR weight was under-reported (Δ -0.70 kg [-3.6 to 2.1], $p<0.0001$) and, therefore, BMI was lower for SR data (Δ -0.29 $kg.m^2$ [-1.5 to 1.0], $p<0.0001$). BMI classification was correct in 93% of cases.

CONCLUSION

We demonstrate the utility of genotype information for detection of possible identity fraud in e-health studies and confirm the reliability of internet-based, SR anthropometric and demographic data collected in the Food4Me study.

KEY WORDS – Internet-based, validation, identity, anthropometrics, personalised nutrition and randomized controlled trial.

TRIAL REGISTRATION - NCT01530139 (http://clinicaltrials.gov/show/NCT01530139)

INTRODUCTION

Non-communicable diseases (NCD) account for over half of global deaths [30], with 4 million deaths annually attributed to cardiovascular diseases (CVD) alone [19]. Because modifiable risk factors, notably diet, smoking and physical activity (PA), account for more than 80% of deaths from CVD and cerebrovascular diseases [30], effective lifestyle-based interventions are important for minimising NCD burden. However, current strategies to improve diet and PA result in relatively modest behavioural changes [9, 15] and may have limited ability to reduce NCD-related mortality. Traditionally, face-to-face interventions have been used to promote behavioural changes. By 2015, 85% of the EU population are predicted to be internet users [6] and internet-based interventions are increasing. The degree of behavioural change achievable via internet-based interventions is similar to [26, 29], or potentially greater than [28], those conducted face-to-face.

The advantages of administrating nutritional interventions via the internet include scalability, efficient and cost-effective collection of data, and lower respondent and researcher burden[5]. On the other hand, intervention studies conducted remotely via the internet may incur problems of fidelity in the self-reported (SR) data and in the collection of biological samples, the provenance of which may be uncertain or unreliable. Furthermore, SR anthropometric data may be prone to respondent biases and measurement errors. Validation studies (VS) in which trained researchers repeat measurements in a sub-sample of the population are integral to ensure the quality of data collected in internet-based interventions and provide some reassurance [27]. However, verification of participant identity appears to have been neglected in previous validation studies [1, 3, 17, 22]. Using the internet to recruit participants into intervention studies delivered remotely provides opportunities for participant mis-representation (identity fraud i.e. pretending to be who they are not) which may undermine the objectives and findings of the study.

The Food4Me study, an internet-based randomized controlled trial (RCT) conducted across seven European countries, was designed to test the efficacy of personalised nutrition (PN) approaches on health-related outcomes [4]. Using data from the Food4Me study, the present paper introduces a novel approach for validating participant identity and describes outcomes from a VS to assess the validity of internet-based, SR anthropometric, demographic and identity data, compared with standardized measurements performed face-to-face.

METHODS

The present VS was performed in a subsample of the Food4Me PoP study, a four-arm, internet-based RCT conducted across seven European countries on the efficacy of PN approaches on health-related outcomes[4].

Design of the Proof of Principle study

The Food4Me PoP study protocol has been described in detail [4]. In brief, participants across seven European countries were recruited via the internet to emulate an internet-based PN service. Recruitment was aided by local and national advertising via the internet, radio advertisements, posters, e-flyers, the use of social media and word of mouth. Identical standardised protocols for recruitment were used in the seven European countries, aiming for 1540 participants (i.e. 220 participants per country). The PoP study recruitment sites were: University College Dublin, (Ireland); Maastricht University, (The Netherlands); University of Navarra, (Spain); Harokopio University, (Greece); University of Reading, (United Kingdom); National Food and Nutrition Institute, (Poland); Technische Universität München, (Germany).

Eligibility criteria

Participants aged ≥ 18 years were included in the study. To keep the cohort representative of the adult population, a minimal set of exclusion criteria were applied: a) pregnancy or lactation; b) no or limited access to the internet; c) following a prescribed diet for any reason, including weight loss, in the last 3 months; d) insulin dependent diabetes, celiac disease, Crohn's disease, or any metabolic disease or condition that alters nutritional requirements e.g. food intolerances or allergies.

PoP study measures

Participants consented to report their measurements via the internet and to return self-collected biological samples (Dried Blood Spot Cards and Buccal swabs) by post, using pre-paid stamped addressed envelopes. To ensure that procedures were similar in all recruiting centres, standardised operating procedures were prepared for all measurements, and researchers underwent centralised training. In addition, to enable participants to collect and report the required information and to collect, process and dispatch the biological samples correctly,

participants were given printed detailed instructions, and video demonstrations of key procedures were available online. All instructions were provided in the local language.

Collection of demographic and anthropometric data

An online screening questionnaire collected detailed SR information about demographic, food choices, health and anthropometric data. Body weight, height and upper thigh, waist and hip circumferences were self-measured and reported by participants via the internet. Participants were instructed to measure body weight after an overnight fast, without shoes and wearing light clothing using a home or commercial scale, and to measure height, barefoot, using a standardised measuring tape provided by Food4Me[4].

Genotypic analyses

Buccal cell samples were collected from participants at baseline using Isohelix SK-1 DNA buccal swabs and Isohelix dried-capsules and posted to each recruiting centre for shipment to LCG Genomics (Hertfordshire, United Kingdom). LCG Genomics extracted DNA and genotyped 33 loci using KASP™ genotyping assays to provide bi-allelic scoring of single nucleotide polymorphisms (SNP) and insertions and deletions at specific loci[8].

Validation study design

To validate the SR demographic (identity, age and sex) and anthropometric (height, weight and estimated BMI) data, an intervention arm-balanced sub-sample of 140 participants (approximately 20 participants per country) from the PoP intervention study were randomly selected and invited to take part in the VS. Whereas participants for the intervention study were recruited nationally, for logistic reasons, participants living near research centres participated in the VS. Upon completion of the PoP online survey and measurements, participants attended a measurement session at their national research centre. To minimize variations in body mass due to time lags between the completion of SR measures online and the appointment at the research centre, participants were instructed to visit the centre within 2 weeks of their last completed online measurements.

At the research centre, researchers measured height and weight, assessed sex visually, confirmed participant's age and collected buccal cell samples which were sent to LGC Genomics to replicate genotyping of the 33 loci previously genotyped in baseline samples of the PoP study. Concordance between both sets of genotypic data was used to confirm participant identity.

Ethical approval and participant consent

The Research Ethics Committees at each centre administering the intervention granted ethical approval for the VS. Before participation, all participants signed two online consent forms, which were automatically directed to study investigators to be counter-signed and archived. All Ethical Committees accepted an online informed consent procedure, with the exception of The Netherlands and Germany whose ethics committees requested additional hard copy consent forms, which were posted to the respective recruitment centres. The Maastricht University Ethics Committee specified that an extra 10% of the participants should be invited to participate to confirm their demographic SR data (age and sex). This check was performed by teleconference.

Data analysis

SR and VS data are presented as means ± SD for continuous variables and as percentages for categorical variables. Kolmogorov-Smirnov tests for normal distribution were used for continuous variables. Differences between SR and measured height, weight and calculated BMI were assessed using paired t tests. Simple and multiple regression analyses were used to investigate determinants of differences between SR and measured values. General Linear Models were used to investigate differences between SR and measured values by age group, sex and country.

Intra-class Correlation Coefficients (ICC) were used to quantify associations and Bland Altman analyses to investigate the degree of agreement between SR and measured height, weight and BMI [2]. Cohen's kappa statistics and the corresponding 95% confidence interval (CI) for classification were used to assess the concordance of sex, age group and BMI status (underweight, normal weight, overweight and obesity) derived from SR and measured values. The degree of agreement between measured and SR overweight and obesity was assessed as follows: $\kappa < 0$ was none/poor; $0 \leq \kappa \leq 0.20$ was slight; $0.21 \leq \kappa \leq 0.40$ was fair; $0.41 \leq \kappa \leq 0.60$ was moderate; $0.61 \leq \kappa \leq 0.80$ was substantial; and $0.81 \leq \kappa \leq 1.0$ was almost perfect [14]. The sensitivity and

specificity of correctly classified BMI based on the SR data were assessed by ROC analysis. Data analyses were performed using STATA/SE v.13 (StataCorp. College Station, TX, USA) and MedCalc v.12 (Ostend, Belgium)

RESULTS

Participant characteristics

Table 1 summarises characteristics of the 1607 Food4Me participants, and the sub-sample in the VS (n=140). Of 194 participants invited to take part in the VS, 43 were unable to visit the research centre because of location, time constraints or personal reasons and 11 invitees did not respond. The baseline characteristic of these participants who did not take part in the VS were similar to those who accepted to take part in the VS (age 41.3 ± 13.9; weight 72.8 ± 15.6; BMI 25.3 ± 4.7). Demographic and anthropometric characteristics of VS participants were similar to those of the Food4Me PoP Study participants (Table 1).

Validity and reliability of self-reported data

SR weight was slightly lower than measured weight (Δ-0.70 kg SD 1.5, range -6.0 to 5.9, P<0.0001) but there was no significant difference between SR and measured height (Δ0.19 cm SD 1.2, range -3 to 5, P=0.066). Thus, BMI calculated from SR height and weight was slightly lower (Δ-0.29 $kg.m^{-2}$ SD 0.6, range -2.2 to 1.7, P<0.0001) than measured values. There were no significant differences between SR and measured values by age group (<45 and ≥45 years) but men overestimated whereas women underestimated height (Table 2). Overweight and obese participants showed higher levels of under-reporting of body mass compared with normal weight participants (P<0.0005). Results stratified by country are presented in supplementary material (Table S1).

Strong correlations (ICC) were observed between SR and measured values for height (0.990 [95%CI: 0.987 to 0.993], P<0.0001), weight (0.994 [0.991 to 0.995], P<0.0001) and BMI (0.983 [0.977 to 0.988], P<0.0001) (Table 2).

Self-reported and measured values

Outcomes of Bland-Altman analyses of SR v. measured values for height, weight and BMI with the corresponding lower and higher level of agreement (LOA) showed a small systematic under-reporting bias for SR weight (Δ -0.70kg [LOA: -3.6 to 2.1], P<0.0001) and BMI (Δ -0.29 $kg.m^{-2}$ [LOA: -1.5 to 1.0], P<0.0001) compared with the measured values (Figure 1, Table 3). We noted trends for greater under-reporting with increasing body weight and BMI. Bland-Altman results stratified by country are presented in supplementary material (Table S2).

Concordance of demographic and BMI classification

There was a strong concordance for BMI classification (underweight, normal, overweight and obese), estimated from SR and measured height and weight, weighted kappa 0.94 (95% CI 0.89 to 0.99). Five overweight participants (3.5%) were incorrectly classified as being normal weight by the SR method. Of those who were obese, just one participant (0.7%) was incorrectly classified as overweight using SR values, leading to a sensitivity of 94.1% and a specificity of 87.8% (Table 4).

Validation of identity

To validate the identity of the participants, the 33 SNPs genotyped previously for the intervention study were re-genotyped and the two datasets were compared. At the VS visit, we collected new buccal cell samples (n=140) from which we obtained reliable genotypes for 135 (33 SNP x 135 individuals = 4455 genotypes). For the remaining five samples, the poor DNA quality precluded informative analysis. There was perfect genotype concordance between original and repeat samples for all but 4 participants, who had a total of four instances at two distinct SNPs (rs2282679, rs4680) where genotypes did not agree. This mismatch incidence is very low, 4/4455 = 0.09% and falls within accepted values for this technology [24]. To explore possible reasons for the apparent genotype mismatches, DNA sequences in the neighbourhoods of these two SNPs were examined for possible copy number variants (CNVs). This analysis revealed that the two SNPs mapped to known CNVs. Participant sex and age showed perfect concordance between SR data and researcher assessed data.

DISCUSSION

Main findings

A novel aspect of this study was the application of genotype analysis using DNA from buccal cell samples to validate the identity of participants recruited via the internet. By replicating the analysis of 33 genetic variants, we showed 99.9% concordance between patterns of genotypic variants in DNA collected in the VS and those observed in DNA obtained from previous, self-collected buccal cell samples. This demonstrates the utility of this novel approach for identity checking - a potentially sensitive aspect of internet-based interventions delivered remotely which has not been investigated in earlier studies. In addition, our findings provide further evidence that SR data via internet for height, weight and BMI showed a high degree of reliability compared with face-to-face measurements made by experienced researchers using standard protocols. Concordance for BMI classification between SR and measured data was strong and we observed perfect agreement for SR sex and age with that assessed in the VS.

Validation of participant identity

Administrating lifestyle-based interventions via the internet offers advantages of scale, efficiency and cost-effective data collection [5, 31]. Nevertheless, internet-based intervention studies conducted remotely may result in problems of reliability in the recruitment of participants and in the collection of biological samples. To the best of our knowledge, the issue of validation of participant identity appears to have been overlooked in previous validation studies. Inevitably, the use of internet to recruit participants to intervention studies provides undesirable opportunities for participant mis-representation, which may undermine the study objectives. In the current VS, we replicated the analysis of 33 genetic variants as a proxy of validation of identity. We found strong agreement for over 99.9% of participant genotypes, with just four examples showing disagreement. As our results showed a perfect concordance for age and sex verification, these minor mismatches represent technical errors during genotyping or may reflect the presence of copy number variants (CNVs), which complicate genotyping. LGC Genomics reports that the average genotyping error in positive control DNA samples using Kompetitive Allele Specific PCR, or KASP™ is between 0.7 to 1.6% and the assay design success rate is between 98 to 100% [23]. We conclude that it is likely that we had perfect agreement in participant identity between samples collected remotely during the Food4Me study and those collected in the VS. Furthermore, we suggest that this novel genotype-based approach to validation of participant identity may be used in many internet-based observational and intervention studies.

Comparison with other studies

The magnitude of differences between SR and measured height (0.19 cm SD 1.2), weight (-0.70 kg SD1.5), and BMI (-0.29 $kg.m^{-2}$ SD 0.6) observed here is similar to findings from previous internet-based studies in adult populations. NutriNet-Sante,[17] a French internet-based prospective cohort study including a VS in a sub-sample of 815 adults, found that height was over-reported by 0.56 cm (SD 2.4) and that weight and BMI were under-reported by 0.49 kg (SD 1.4) and 0.34 $kg.m^{-2}$ (SD 1.5), respectively. A study conducted in 177 adults (aged 18-35 years) in Australia [22] observed a larger over-reporting bias for height (1.36 cm SD 1.9), and a similar under-reporting bias for weight (-0.55 kg SD 2.0) and BMI (-0.56 $kg.m^{-2}$ SD 0.08) compared with the present study. In contrast, an internet-based study conducted in 149 adults in Sweden[3], reported larger differences between SR and measured weight (1.2 kg SD 2.6) compared with our results. A systematic review [7] of validation of SR anthropometric data found that height was over-reported by 0.6 to 7.5 cm whereas weight and BMI were under-reported by -0.1 to 6.5 kg and 0 to -2.2 $kg.m^{-2}$ respectively. It should be noted that under-reporting of body weight is quite common particularly among overweight and obese subjects [11, 17, 18, 25].

In agreement with some [18, 20, 25] but not all previous studies [3, 17], men in the Food4Me study were more likely to over-report height. Although women appeared more likely to under-report weight than men, this difference was not significant in our study. Previous studies have observed that women were significantly more likely to under-report their weight compared with men [17, 18, 25]. Whilst height was more likely to be over-reported with increasing age in previous studies [1, 13, 17], we did not find any effect of age on differences between SR and measured height.

In addition to sex and age, BMI was a strong predictor of differences between SR and measured methods. As a consequence of mis-reporting of the primary measurements of height and weight, differences in under-reporting of calculated BMI was 4.8 times higher in both overweight and obese individuals compared with normal weight participants (Δ -0.12, -0.54 and -0.53 $kg.m^{-2}$ for normal, overweight and obese participants respectively). Our results confirm previous findings of under-reporting of BMI by 0.16, 0.36 and 0.63 $kg.m^{-2}$ for normal weight, overweight and obese participants respectively [17]. However, we found smaller differences in weight mis-reporting between BMI categories than those observed by another internet-based study [22] in which under-reporting among overweight and obese participants was -1.36 kg compared with -0.31 kg in those of normal

BMI. A possible explanation for the greater degree of mis-reporting of body weight by overweight and obese individuals lies in the social desirability concept, which argues that perceptions are influenced by desires to conform to perceived societal norms and that, with respect to body weight, such pressures apply more strongly in obese participants [16]. However, the estimated proportion of subjects for whom SR height, weight and calculated BMI was within 5% of the measured values were 100% (n=140) for height, 96% (n=135) for weight, and 92% (n=129) for estimated BMI, respectively. This suggests that most Food4Me participants provided reliable measures of their anthropometrics.

Concordance of BMI classification

One of the main concerns arising from data collection, either SR via the internet or with paper-based forms, is the validity and accuracy of the data provided and its utility as a basis for provision of health-related advice. Several studies have reported greater under-estimation of weight (and BMI) with remote SR collection methods than with face-to-face interviews [10]. However, we observed a good agreement between the BMI classifications derived from SR and measured height and weight (kappa 0.939), with just six participants being wrongly classified when SR data were used. There were no differences in the proportions of those classified as underweight, and only small differences in the proportions of normal weight (3.6%), overweight (-2.9%) and obese participants (-0.7%). These results are comparable with previous findings reporting a kappa of 0.97 for BMI classification and prevalence differences between SR and measured values of 0.6 and 0.7% for overweight and obese participants, respectively [17]. Similarly, Pursey et al. reported that the prevalence of overweight was 2.6% lower when using SR compared with measured values, but there was no difference for obesity prevalence [22].

Although social desirability may drive differences between SR and measured values [12], we found very good agreement between the internet-based SR and validation measures for the key anthropometric variables height and weight suggesting that, in an internet-based setting, participants may be less prone to social desirability bias. This apparently enhanced truthfulness may result from the greater feeling of anonymity when using the web rather than other media such as the telephone [12]. However, the reliability of more difficult self-measurements such as waist and hip circumferences need to be explored in future studies.

Strengths and limitations

To our knowledge, this is the first internet-based study that has validated participant identity using genotypic analysis. Our findings of the utility, and practicability, of this approach to validation of participant identity provide proof of concept for remotely-conducted, e.g. internet-based, studies in which participant mis-representation is a potentially major, and often ignored, concern. A particular strength of this study was the collection of data via a novel internet-based server in European countries from a relatively large sample of the adult population with a wide range of ages and BMIs. Our ability to obtain reliable SR anthropometric data was enhanced by the use of standardized protocols by study participants. Protocols were provided in text format with pictures, but also as a series of online videos. In addition, during the VS, trained researchers collected the anthropometric data using the same standardised protocols. An additional strength of our study was the short period of time (i.e. up to 2 weeks) between the collection of internet-based SR data and direct measurement by the researchers. Furthermore, to ensure independence of measurements in the subsequent VS, subjects were invited to participate in the VS only after they had completed their internet-based measures.

A potential limitation of our study is that the participants in the Food4Me study were recruited from those showing interest in an intervention study on PN. As a result, we may have recruited those with a particular interest in lifestyle-based interventions but we have no reason to believe that this interest influenced the truthfulness of SR data. In addition, the BMI distribution among Food4Me participants was comparable with the prevalence of normal weight, overweight and obesity in the adult European population [21].

In conclusion, we introduced and tested, a simple genotype-based approach for validation of the identity of study participants recruited to internet-based studies. This approach is simple and robust and, given the low costs of genotyping we envisage that it may have wide utility for identity validation in the many types of studies (including internet-based studies) where participant recruitment and sample data collection are conducted remotely. Although overall agreement between SR and measured values was excellent, under-reporting of weight was more common among overweight and obese individuals, and such SR data should be interpreted with caution when adiposity is an important outcome. Overall, our findings clearly demonstrate the reliability of internet-based, SR anthropometric and demographic data collected in the Food4Me study.

ACKNOWLEDGEMENTS

The Food4Me study was supported by the European Commission under the Food, Agriculture, Fisheries and Biotechnology Theme of the 7th Framework Programme for Research and Technological Development, grant number 265494. We acknowledge the design contribution of Anette Karlsen and Ingrid Gjelstad at early stages of the validation study.

Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The USDA is an equal opportunity provider and employer.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Bes-Rastrollo M, Sabate J, Jaceldo-Siegl K, Fraser GE (2011) Validation of self-reported anthropometrics in the Adventist Health Study 2. BMC Public Health 11. doi 10.1186/1471-2458-11-213.
2. Bland JM, Altman DG (2010) Statistical methods for assessing agreement between two methods of clinical measurement. International Journal of Nursing Studies 47:931-936. doi 10.1016/j.ijnurstu.2009.10.001.
3. Bonn SE, Lagerros YT, Balter K (2013) How Valid are Web-Based Self-Reports of Weight? Journal of Medical Internet Research 15. doi 10.2196/jmir.2393.
4. Celis-Morales C, Livingstone KM, Marsaux CFM, Forster H, O'Donovan CB, Woolhead C, Macready AL, Fallaize R, Navas-Carretero S, San-Cristobal R, Kolossa S, Hartwig K, Tsirigoti L, Lambrinou CP, Moschonis G, Godlewska M, Surwiłło A, Grimaldi K, Bouwman J, Daly EJ, Akujobi V, O'Riordan R, Hoonhout J, Claassen A, Hoeller U, Gundersen TE, Kaland SE, Matthews JNS, Manios Y, Traczyk I, Drevon CA, Gibney ER, Brennan L, Walsh MC, Lovegrove JA, Alfredo Martinez J, Saris WHM, Daniel H, Gibney M, Mathers JC (2014) Design and baseline characteristics of the Food4Me study: a web-based randomised controlled trial of personalised nutrition in seven European countries. Genes Nutr 10:1-13. doi 10.1007/s12263-014-0450-2.
5. Celis-Morales CL, J; Mathers, JC; (2014) Personalising nutritional guidance for more effective behaviour change. Proceedings of the Nutrition Society:1-9. doi 10.1017/S0029665114001633
6. Eurostat (2012) Life Online In:
7. Gorber SC, Tremblay M, Moher D, Gorber B (2007) Diagnostic in obesity comorbidities - A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. Obesity Reviews 8:307-326. doi 10.1111/j.1467-789X.2007.00347.x.
8. He C, Holme J, Anthony J (2014) SNP genotyping: the KASP assay. Methods in molecular biology (Clifton, N.J.) 1145:75-86. doi 10.1007/978-1-4939-0446-4_7.
9. Hobbs N, Godfrey A, Lara J, Errington L, Meyer TD, Rochester L, White M, Mathers JC, Sniehotta FF (2013) Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? A systematic review and meta-analysis. BMC Medicine 11. doi 10.1186/1741-7015-11-75.
10. Hood K, Robling M, Ingledew D, Gillespie D, Greene G, Ivins R, Russell I, Sayers A, Shaw C, Williams J (2012) Mode of data elicitation, acquisition and response to surveys: a systematic review. Health Technology Assessment 16:1-+. doi 10.3310/hta16270.

11. Johansson L, Solvoll K, Bjorneboe GEA, Drevon CA (1998) Under- and overreporting of energy intake related to weight status and lifestyle in a nationwide sample. American Journal of Clinical Nutrition 68:266-274.
12. Joinson A (1999) Social desirability, anonymity, and Internet-based questionnaires. Behavior Research Methods Instruments & Computers 31:433-438. doi 10.3758/bf03200723.
13. Kuczmarski MF, Kuczmarski RJ, Najjar M (2001) Effects of age on validity of self-reported height, weight, and body mass index: Findings from the third National Health and Nutrition Examination Survey, 1988-1994. Journal of the American Dietetic Association 101:28-34. doi 10.1016/s0002-8223(01)00008-6.
14. Landis JR, Koch GG (1977) Measurement of observer agreement for categorical data. Biometrics 33:159-174. doi 10.2307/2529310.
15. Lara J, Hobbs N, Moynihan PJ, Meyer TD, Adamson AJ, Errington L, Rochester L, Sniehotta FF, White M, Mathers JC (2014) Effectiveness of dietary interventions among adults of retirement age: a systematic review and meta-analysis of randomized controlled trials. BMC Medicine 12. doi 10.1186/1741-7015-12-60.
16. Larson MR (2000) Social desirability and self-reported weight and height. International Journal of Obesity 24:663-665. doi 10.1038/sj.ijo.0801233.
17. Lassale C, Peneau S, Touvier M, Julia C, Galan P, Hercerg S, Kesse-Guyot E (2013) Validity of Web-Based Self-Reported Weight and Height: Results of the Nutrinet-Sante Study. Journal of Medical Internet Research 15. doi 10.2196/jmir.2575.
18. Merrill RM, Richardson JS (2009) Validity of self-reported height, weight, and body mass index: findings from the National Health and Nutrition Examination Survey, 2001-2006. Preventing Chronic Disease 6.
19. Nichols M, Townsend N, Scarborough P, Rayner M (2013) European Cardiovascular Disease Statistics 4th edition 2012: EuroHeart II. European Heart Journal 34:3007-3007. doi 10.1093/eurheartj/eht379.
20. Niedhammer I, Bugel I, Bonenfant S, Goldberg M, Leclerc A (2000) Validity of self-reported weight and height in the French GAZEL cohort. International Journal of Obesity 24:1111-1118. doi 10.1038/sj.ijo.0801375.
21. OECD (2012) Health at a Glance: Europe 2012. In: Publishing. O (ed)OECD
22. Pursey K, Burrows TL, Stanwell P, Collins CE (2014) How Accurate is Web-Based Self-Reported Height, Weight, and Body Mass Index in Young Adults? Journal of Medical Internet Research 16. doi 10.2196/jmir.2909.
23. Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Molecular Breeding 33:1-14. doi 10.1007/s11032-013-9917-x.
24. Smith CE, Arnett DK, Corella D, Tsai MY, Lai CQ, Parnell LD, Lee YC, Ordovas JM (2012) Perilipin polymorphism interacts with saturated fat and carbohydrates to modulate insulin resistance. Nutrition Metabolism and Cardiovascular Diseases 22:449-455. doi 10.1016/j.numecd.2010.09.003.
25. Spencer EA, Appleby PN, Davey GK, Key TJ (2002) Validity of self-reported height and weight in 4808 EPIC-Oxford participants. Public Health Nutrition 5:561-565. doi 10.1079/phn2001322.
26. Steele RM, Mummery WK, Dwyer T (2009) A Comparison of Face-to-Face or Internet-Delivered Physical Activity Intervention on Targeted Determinants. Health Education & Behavior 36:1051-1064. doi 10.1177/1090198109335802.
27. Thorndike FP, Ritterband LM, Saylor DK, Magee JC, Gonder-Frederick LA, Morin CM (2011) Validation of the Insomnia Severity Index as a Web-Based Measure. Behavioral Sleep Medicine 9:216-223. doi 10.1080/15402002.2011.606766.
28. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM (2004) The effectiveness of Web-based vs. non-Web-based interventions: A meta-analysis of behavioral change outcomes. Journal of Medical Internet Research 6:67-84.
29. Weigold A, Weigold IK, Russell EJ (2013) Examination of the Equivalence of Self-Report Survey-Based Paper-and-Pencil and Internet Data Collection Methods. Psychological Methods 18:53-70. doi 10.1037/a0031607.
30. WHO (2010) Global status report on noncommunicable diseases 2010. In:World Health Organization
31. Wright KB (2005) Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and Web survey services. Journal of Computer-Mediated Communication 10.

455 **Table 1.** Demographic and anthropometric characteristics of the Food4Me Proof of Principle (PoP) Study and

456 Validation Study participants.

	Food4Me PoP Study participants	**Validation Study participants**	**P-value**
Demographic			
Total (n)*	1607	140	-
Sex - female (%)	60.9	56.4	0.719
Age (years)	39.8 ± 13.1	42.6 ± 13.6	0.018
Age range (years)	18 to 79	18 to 68	-
Anthropometrics			
Height (cm)	171.1 ± 9.4	170.1 ± 9.1	0.227
Weight (kg)	74.6 ± 15.8	72.3 ± 14.2	0.089
BMI ($kg.m^{-2}$)	25.5 ± 5.2	24.9 ± 3.9	0.173
Weight status categories (%)			
Underweight: BMI <18.5	2.7	0.7	0.171
Normal weight: BMI ≥18.5 to ≤24.9	51.2	56.4	0.244
Overweight: BMI ≥25 to ≤29.9	30.3	30.7	0.926
Obese: BMI ≥30.0	15.8	12.2	0.252

457 Data represent means ± SD for continuous variables and percentages for categorical variables. Differences for

458 continuous variables were analysed using independent t-test and Chi-square for categorical variables.

459 *Sex and age were verified by teleconference in an additional 21 participants in The Netherlands.

460 **Table 2.** Summary statistics and correlation coefficients for self-reported and measured height, weight and BMI.

Variables	Collection method		P-value[a]	Correlation coefficient
	Self-reported	Measured		ICC (95%CI)[b]
All (n=140)				
Height (cm)	170.3± 9.4	170.1 ± 9.1	0.066	0.990 (0.986 to 0.993)*
Weight (kg)	71.6 ± 13.9	72.3 ± 14.3	<0.0001	0.993 (0.991 to 0.995)*
BMI (kg.m^{-2})	24.6 ± 3.8	24.9 ± 3.9	<0.0001	0.983 (0.977 to 0.988)*
By sex:				
Women(n=79)				
Height (cm)	164.2 ± 6.4	164.3 ± 6.1	0.084	0.974 (0.960 to 0.983)*
Weight (kg)	64.8 ± 10.7	65.5 ± 11.1	0.0004	0.987 (0.981 to 0.992)*
BMI (kg.m^{-2})	24.1 ± 3.9	24.3 ± 4.1	0.005	0.982 (0.972 to 0.988)*
Men (n=61)				
Height (cm)	178.1 ± 6.4	177.6 ± 6.3	0.0002	0.985 (0.975 to 0.981)*
Weight (kg)	80.4 ± 12.6	81.2 ± 13.0	<0.0001	0.993 (0.988 to 0.995)*
BMI (kg.m^{-2})	25.3 ± 3.5	25.7 ± 3.6	<0.0001	0.983 (0.973 to 0.990)*
By age group				
<45 years (n=71)				
Height (cm)	171.2± 8.9	171.2 ± 8.4	0.136	0.990 (0.985 to 0.994)*
Weight (kg)	70.0 ± 13.6	70.5 ± 13.8	0.009	0.992 (0.988 to 0.996)*
BMI (kg.m^{-2})	23.7 ± 3.6	23.9 ± 3.7	0.005	0.981 (0.970 to 0.988)*
≥45 years (n=69)				
Height (cm)	169.3± 9.8	169.1 ± 9.7	0.236	0.990 (0.984 to 0.993)*
Weight (kg)	73.3 ± 14.1	74.2 ± 14.5	<0.0001	0.994 (0.990 to 0.996)*
BMI (kg.m^{-2})	25.4 ± 3.7	25.8 ± 3.9	<0.0001	0.983 (0.973 to 0.989)*
By BMI categories				
Normal weight (n=80)				
Height (cm)	169.6 ± 9.0	169.5 ± 8.7	0.719	0.992 (0.987 to 0.994)*
Weight (kg)	63.1 ± 8.5	63.4 ± 8.4	0.053	0.984 (0.976 to 0.990)*
BMI (kg.m^{-2})	21.9 ± 1.7	22.0 ± 1.7	0.071	0.937 (0.903 to 0.959)*
Overweight (n=43)				
Height (cm)	171.1± 9.5	170.6 ± 10.0	0.017	0.987 (0.977 to 0.993)*
Weight (kg)	78.8 ± 9.2	79.9 ± 9.3	<0.0001	0.986 (0.975 to 0.992)*
BMI (kg.m^{-2})	26.8 ± 1.5	27.40 ± 1.3	<0.0001	0.839 (0.722 to 0.909)*
Obese (n=17)				
Height (cm)	171.8 ± 9.0	171.8 ± 9.0	0.984	0.991 (0.970 to 0.997)*
Weight (kg)	93.3 ± 10.4	94.8 ± 10.3	0.002	0.974 (0.934 to 0.990)*
BMI (kg.m^{-2})	31.5 ± 1.7	32.1 ±1.6	0.006	0.864 (0.672 to 0.948)*

461 Data represent means ± SD for self-reported and measured values. [a] Paired t-test was used for assessing
462 differences between means of both methods. [b] Intraclass correlation coefficient (ICC) and [c] Pearson Product
463 correlation coefficient (r) and their corresponding 95% confident intervals were used to assess the level of
464 reliability between methods. *All P-values for ICC and Pearson correlation were significant at <0.0001.

Table 3. Bland-Altman analyses for self-reported and measured height, weight and BMI.

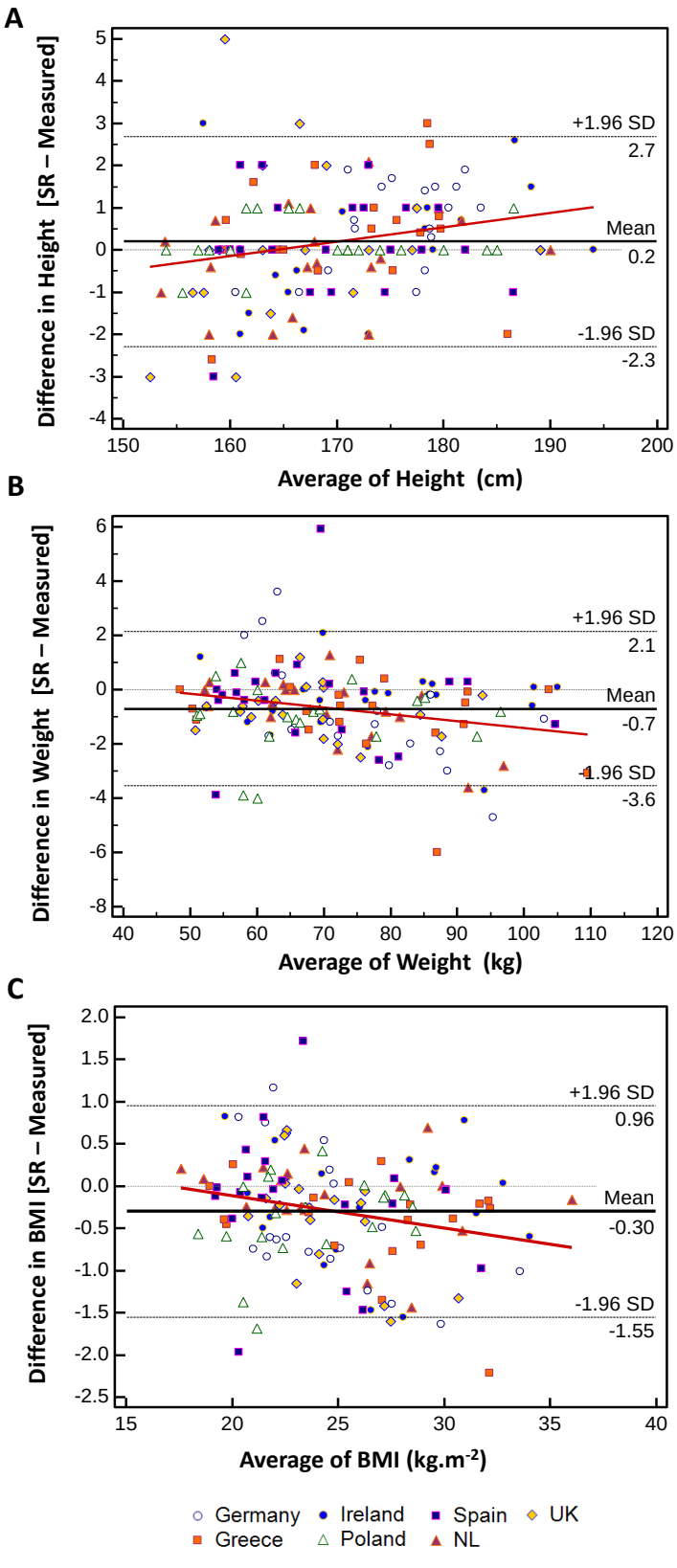
Variables	Bland-Altman		P-value*
	Absolute mean differences (LOA)	(%) Relative mean differences (LOA)	
All (n=140)			
Height (cm)	0.19 (-2.3 to 2.7)	0.11 (-1.4 to 1.6)	0.066
Weight (kg)	-0.70 (-3.6 to 2.1)	-0.93 (-4.9 to 3.1)	<0.0001
BMI ($kg.m^{-2}$)	-0.29 (-1.5 to 1.0)	-1.14 (-6.2 to 4.0)	<0.0001
By sex:			
Women(n=79)			
Height (cm)	0.03 (-2.8 to 2.7)	0.02 (-1.7 to 1.7)	0.084
Weight (kg)	-0.65 (-3.7 to 2.4)	-0.94 (-5.6 to 3.7)	0.0004
BMI ($kg.m^{-2}$)	-0.23 (-1.6 to 1.2)	-0.89 (-6.7 to 4.9)	0.005
Men (n=61)			
Height (cm)	0.49 (-1.4 to 2.4)	0.28 (-0.8 to 1.4)	0.0002
Weight (kg)	-0.81 (-3.3 to 1.8)	-0.90 (-3.9 to 2.1)	<0.0001
BMI ($kg.m^{-2}$)	-0.38 (-1.4 to 0.6)	-1.45 (-5.3 to 2.4)	<0.0001
By age group:			
<45 years (n=71)			
Height (cm)	0.21 (-2.1 to 2.5)	0.11 (-1.3 to 1.5)	0.136
Weight (kg)	-0.50 (-3.6 to 2.6)	-0.69 (-5.3 to 3.9)	0.009
BMI ($kg.m^{-2}$)	-0.23 (-1.5 to 1.1)	-0.91 (-6.6 to 4.8)	0.005
>45 years (n=69)			
Height (cm)	0.18 (-2.1 to 2.5)	0.10 (-1.5 to 1.7)	0.236
Weight (kg)	-0.91 (-3.5 to 1.6)	-1.16 (-4.4 to 2.0)	<0.0001
BMI ($kg.m^{-2}$)	-0.37 (-1.5 to 0.8)	-1.37 (-5.7 to 3.0)	<0.0001
By BMI categories			
Normal weight (n=80)			
Height (cm)	0.04 (-2.1 to 2.2)	0.02 (-1.3 to 1.3)	0.719
Weight (kg)	-0.32 (-3.1 to 2.5)	-0.52 (-5.0 to 4.0)	0.053
BMI ($kg.m^{-2}$)	-0.12 (-1.3 to 1.0)	-0.56 (-5.9 to 4.7)	0.071
Overweight (n=43)			
Height (cm)	0.56 (-2.4 to 3.5)	0.32 (-1.5 to 2.1)	0.017
Weight (kg)	-1.08 (-3.2 to 1.0)	-1.37 (-3.9 to 1.2)	<0.0001
BMI ($kg.m^{-2}$)	-0.54 (-1.7 to 0.7)	-2.01 (-6.4 to 2.4)	<0.0001
Obese (n=17)			
Height (cm)	0.01 (-2.4 to 3.2)	0.01 (-1.4 to 1.4)	0.984
Weight (kg)	-1.56 (-3.8 to 1.4)	-1.70 (-5.6 to 2.2)	0.002
BMI ($kg.m^{-2}$)	-0.53 (-1.8 to 0.7)	-1.68 (-6.1 to 2.8)	0.006

Data represent absolute and relative mean differences [SR - measured values] with their corresponding limits of agreements (LOA ±1.96 SD). * Paired t-test was used for assessing absolute differences between means of SR and measured values.

Table 4. Validity and concordance of weight classification estimated from self-reported and measured values.

BMI categories	SR	Measured	Number misclassified
Underweight	1 (0.7%)	1 (0.7%)	0
Normal	84 (60.0%)	79 (56.4%)	5 (3.5%)
Overweight	39 (27.9%)	43 (30.7%)	4 (2.9%)
Obese	16 (11.4%)	17 (12.1%)	1 (0.7%)
*Kappa	0.939 (0.891 to 0.988)		

Data represent count (and percentages) for measured and self-reported (SR) values. *A weighted Kappa value and its corresponding 95% CI were estimated to measure the level of concordance between both methods.



472

473

474 Figure 1. Bland-Altman plots illustrating the agreement between self-reported (SR) and measured (a) height, (b)

475 weight, (c) BMI, and the corresponding means estimated by the two methods across all countries. Solid lines are

476 mean differences and dotted lines are the lower and upper 95% limits of agreements; red lines illustrate the
477 regression line for differences in measurements against the mean of both SR and VS measurements.