The reliability and validity of the PortionSize™ and MyFitnessPal apps
(Study 2: Semi-controlled, free-living evaluation)

NCT05166226

August 24, 2022

Title: The reliability and validity of the PortionSize™ and MyFitnessPal apps (Study 2: Semi-controlled, free-living evaluation).

Principal Investigators: Corby K. Martin, Ph.D.; John W. Apolzan, Ph.D.

Co-Investigator: Stephanie T. Broyles, Ph.D.

Protocol Version Date: August 25, 2022

Background

Accurately quantifying food intake is vital to promoting health and reducing chronic disease risk. Food intake encompasses energy intake, nutrient intake (macronutrients, micronutrients, vitamins, minerals), and intake of various food groups (e.g., fruits, vegetables), and thus reflects the nutritional status of individuals. Nutrition affects disease risk, including risk of developing obesity, [1] diabetes, [2] and cancer, [3] all of which negatively affect the United States (U.S). Nonetheless, accurate assessment of food and nutrient intake has remained challenging, despite an improvement in methods. Self-report methods, namely food records, are a mainstay of nutritional epidemiology research, [4] with food recall being another popular method. [5]These methods rely on the participant to accurately estimate portion size and, for food recall, remember what was consumed. The accuracy of these methods have been questioned [6-9] and the problems with human recall have been comprehensively outlined. [10] As a result, there remains a significant need for methods that are sufficiently accurate to provide researchers with good outcome data and to guide health promotion efforts.

The PortionSize™ app was designed by our laboratory to overcome the limitations outlined above, and to guide users to follow specific diets. PortionSize™ relies on users capturing images of their food selection and waste. Food intake data are immediately provided since the user relies on built in tools, including templates, to estimate portion size. However, despite promising early indications, the PortionSize™ app has yet to be validated. Determining the accuracy and reliability of PortionSize™ is vital before the app can be used by people to obtain immediate feedback about their food intake. We accordingly aim to test the reliability and validity of PortionSize™ in a semi-controlled free-living environment using food provision. We will also test the reliability and validity of the MyFitnessPal app in a semi-controlled free-living environment. The accuracy of the apps will also be compared to each other.

Methods

Study Design

We will recruit adults to use the PortionSize™ app and the MyFitnessPal apps in free-living conditions, where participants will consume pre-weighed food from a cooler, which provides a test of nutrient intake in free-living conditions. Participants will consume pre-prepared and pre-weighed foods from a cooler for two non-consecutive three-day periods that are separated by approximately one week. In a within subjects and randomized, counterbalanced design, participants will use PortionSize™ during one three-day period and MyFitnessPal during the other period. Following our previous methods,[11] participants will pick up a cooler in the morning and, at approximately the same time, return the cooler from the previous day, allowing us to weigh food waste (i.e., determine food intake since the foods were weighed when they were put in the cooler). This procedure allows us to test the accuracy of the apps at measuring energy and nutrient intake in free-living conditions. Participants will be trained in our clinic/laboratory to use the apps following the procedures of Study 1.

Participants and Recruitment

We will recruit 45 adults (approximately equal proportion of men and women) who will be randomly assigned to use PortionSize™ or MyFitnessPal. We expect to enroll an approximately equal number of participants with body mass indexes (BMI; kg/m^2) between 18.5 to 25, 25 to 30, and 30 – 50kg/m^2. Recruitment will be stratified based on sex and body mass index.

- Inclusion criteria include:
  - Male or female, age 18-62 years
  - Body mass index (BMI) 18.5-50 kg/m^2
  - Subjects who are capable of giving informed consent and complying with all study procedures/requirements.
  - Ownership of an iPhone model 6s or later, which the participant is willing to use for the study
  - Access to Apple ID, password, and email address and willing to use them during the study
  - Willing to use data and any accompanying charges as part of study participation
  - Willing to complete all study procedures and adhere to study visit timelines
  - Willing to be re-contacted for future research and/or follow-up
- Exclusion criteria include
  - Active cancer diagnosis (excluding some melanomas)
  - Have been 6-month weight unstable (gain/loss >10 lbs in last 6 months)
  - Undergoing weight loss treatment

- PBRC employee, as previous reviewers argued that they are not representative of the community
- Women who are currently pregnant or breastfeeding (self-reported)
- Diagnosed with an uncontrolled thyroid disorder (controlled = 3 months of medication)
- Have had or plan to have weight loss surgery *(gastric band removal may be allowed at PI discretion)*
- Allergic to or unwilling to eat study foods provided *(exceptions made at the discretion of the PI)*
- Consume >28 alcoholic beverages per week
- Anyone severely immunocompromised
- Any condition or circumstance that in the judgement of the PI could interfere with study participation

Procedures

Screening visit

At the screening visit, an individual or group presentation related to study procedures will be presented. For participants interested in participating thereafter, a written informed consent will be obtained. Once consented, the following will occur:

- Completion of study questionnaires including the Lifestyle and Demographics Questionnaire and a Dietary Screening Questionnaire
- Anthropometrics: height, weight (non-metabolic) and BMI
- Determine eligibility to proceed with Day 0.

Testing Period 1 & Period 2

Eligible participants who choose to enroll will complete 2 testing periods (Period 1 and Period 2) where they will be provided with prepared foods from the Metabolic Kitchen to consume over 3 days. Once Period 1 is complete, participants will have a washout period before starting Period 2. Participants will be randomized to use either the PortionSize™ app or MyFitnessPal app for Period 1 or Period 2. Period 1 will include a baseline visit to randomize the participant to app order and to train the participant how to use the app that they are to use first. Period 2 will include a Day 0 for app training. Each period will have the following visit days: Day 1, 2, 3, and 4.

Visits for each period are ideally consecutive days. For example, Day 1 is +1 day from Baseline/Day 0, Day 2 is +1 day from Day 1, etc. This is not possible, however, for participants whose data collection spans a weekend or a holiday, for example. Additionally, circumstances may arise that prohibit participants' or the Center's ability to collect data across 4 consecutive days (inclusive of the run-in day and Day 4 when participants return their last cooler). Thus, efforts will be made for Day 1 to

occur the day after Baseline/Day 0, though the PIs may approve changes to this schedule if it is not possible due to an unforeseen event. Similarly, the PIs may approve the collection of data on non-consecutive days between Days 1 and 2, and 2 and 3 due to unforeseen events, illness, holidays, etc. In rare cases where many days elapse between either Day 1 and 2, or 2 and 3, the PIs may request that the participant complete an abbreviated booster training and run-in for the app being used during that period. In no cases shall more than 14 days elapse between days of data collection. Lastly, in cases where Days 1 and 2, and/or 2 and 3, are not sequential, up to two additional visits to the Center are possible since the participant cannot drop off and pick up their next cooler in the same visit.

After the completion of Period 1, a washout of approximately 1 week will occur before initiating Period 2. The washout period may be expanded up to 30-45 days as needed for scheduling and unexpected circumstances.

Food pick-up and drop-off visits can be repeated at the PI's discretion to provide participants and staff flexibility. Should extenuating circumstances occur, this will allow for proper data collection and quality in the face of unforeseen events.

Baseline (and Day 0)

Participants will attend an in-person visit at Pennington for Baseline, which occurs during Period 1, to be randomized and trained how to use the app that they will use first. During Day 0, which occurs during Period 2, participants will be trained to use the app that they will use second. These visits will last ~45 to 60 minutes. The training will teach participants how to report food intake and estimate portion size per our standard procedures. Participants will receive training by study staff members on the app that they will use for that period. That is, individuals randomized to use PortionSize™ first will be trained on PortionSize™, whereas those randomized to use MyFitnessPal first will be trained on MyFitnessPal. This first day is usually a partial day of data collection and will serve as a run-in/familiarization period for the participants, during which they will be instructed to use the app for the remainder of the day. These data will not be analyzed as part of the 3-day use period.

Day 1

Participants will return to the lab before breakfast in the morning to be weighed (non-metabolic) and be provided with the first cooler of food to consume for the remainder of the day. The coolers will contain about 1.3 times the participant's overall estimated energy needs, with ~20%, 30%, 40%, and 10% of energy provided as breakfast, lunch, dinner, and snacks, respectively. Staff will attempt to provide a day's worth of food that reflects 50% carbohydrate, 20% protein, and 30% fat (+/- 5%). The Mifflin St. Jeor Equation will be used to calculate basal metabolic rate (BMR), and total energy requirements will be calculated by multiplying BMR by an activity factor of 1.5. This value then will be multiplied by 1.3 to provide 30% more energy than the person requires. Participants will be instructed to only consume foods and beverages (with the

exception of water and non-caloric beverages such as black coffee or plain tea) from the cooler, but they can eat/drink as much or as little as they wish. Participants who report drinking diet, regular sodas, or sweet teas will be provided with their preferred beverage in their cooler, and efforts will be made to provide these beverages in typical store-bought containers (e.g., 12-ounce cans). All food and beverage items will be pre-weighed and weighed immediately upon return per our previous methods.[11] We may include dietary supplement(s) on certain days, though the supplement may vary among participants (e.g., multivitamin, energy bar or shake, etc.).

This visit will last approximately 15-30 minutes.

Day 2

Participants will return to the lab before breakfast in the morning to be weighed (non-metabolic) and to return their cooler from Day 1. They will then pick up a new cooler containing their food and beverages for Day 2. The cooler content and procedures are the same as described above though the menus of foods in the cooler may differ each day. This visit will last approximately 15-30 minutes.

Day 3

Participants will return to the lab before breakfast in the morning to be weighed (non-metabolic) to return their cooler from Day 2. They will then pick up a new cooler containing their food and beverages for Day 3. The cooler content and procedures are the same as described above though the menus or foods in the cooler may differ each day. This visit will last approximately 15-30 minutes.

Day 4

After 3 days of study food tracking, participants will have their weight measured a final time and return the cooler to the laboratory. They will then complete the following questionnaires:

- The Computer System Usability Questionnaire (CSUQ) for MyFitnessPal or PortionSize™ (the CSUQ that is completed will correspond to the app used during that period). The CSUQ is a standardized, reliable, and valid questionnaire originally designed to evaluate computer programs[12, 13] that has been used to quantify the usability of mobile phone apps.[14, 15]
- The PortionSize™ User Satisfaction Survey or the MyFitnessPal User Satisfaction Survey (the survey that is completed will correspond to the app used during that period).
- PortionSize and MyFitnessPal-specific Dietary Health Questionnaires (DHQ III) are modified such that participants will report the presence of vitamins and supplements during the days of food provision. Thus, only those items (vitamins and supplements) of the DHQ-III will be administered, and the instructions will ask participants to note if those items were present in the coolers.

- Period 2 only, the User Preference Survey.

The visit will last approximately 30-60 minutes for Period 2 with the addition of the Preference Survey.

Outcome measures

The **primary analyses** assessing equivalence between PortionSize™ and the criterion measure (and MyFitnessPal and the criterion measures) will rely on equivalence testing using the Two One-side T-test (TOST)[16, 17] method. The null hypothesis of an equivalence tests is that the means are nonequivalent, i.e., the confidence interval of the mean of the test condition exceeds the pre-specified error bound. The alternative hypothesis is that the means are equivalent, and the mean difference between the experimental method and gold standard is hypothesized to be zero. Differences in error from PortionSize™ compared to error from MyFitnessPal will be assessed using dependent samples t-tests, with error calculated as the difference between each experimental method and the criterion variable (e.g., energy intake values from PortionSize™ minus energy intake values from weighed foods). These procedures will also produce results indicating if the error from each method differs from zero. Subsequent regression-based **exploratory analyses** will determine if error from the apps varies by body mass (kg, BMI), age, or sex; and will evaluate differences in user satisfaction between the apps. Bland-Altman analysis[18] will be used to test for differences in error variance over levels of the variable being measured (e.g., food intake).

Endpoints will be:

- Energy intake (kcal)
- User satisfaction and usability of the apps
- Nutrients (e.g., fat, carbohydrate, and protein as grams and percent energy from each macronutrient) and food groups (PortionSize™ only since it has the ability to provide output in food groups due to the link with the FNDDS database).

Other exploratory endpoints include human rater analysis of captured food images and the modified Dietary History Questionnaire III (DHQIII) as comparators for food and possible dietary supplement intake.

Schedule of procedures

The schedule of procedures is outlined below:

Procedure	Screening Visit	Baseline/Day 0	Day 1	Day 2	Day 3	Day 4
Informed Consent	X					
Lifestyle and Demographics Questionnaire; Anthropometrics; Dietary Screening Questionnaire, Eligibility Evaluation	X					
Anthropometric: Weight only (Non-Metabolic)			X	X	X	X
Randomization (to use PortionSize™ or MyFitnessPal first)		X				
PortionSize™ or MyFitnessPal app training and ~1 day run-in		X				
Use of app during free-living conditions			X	X	X	
Usability Surveys (MyFitnessPal & PortionSize™ specific CSUQ)						X*
PortionSize and MyFitnessPal-specific Modified Dietary Health Questionnaire (DHQIII[19])						X*
PortionSize and MyFitnessPal-specific						X*
User Preference Survey						X**

** MyFitnessPal or PortionSize™ specific survey will occur on Day 4 in the respective period that app was utilized*

***Day 4 in Period 2 will include a User Preference survey to rate the two apps*

Power calculation

We expect no more than ~8% drop-out and will recruit at least 45 participants; hence, power analyses were conducted with a sample size of at least 42.

Estimated food intake values from the PortionSize™ and MyFitnessPal apps will be compared to directly weighed values, which serve as the criterion measure. Equivalence tests that follow the Two One-sided T-test (TOST)[16, 17] method will be used

(error reflects the difference between weighed intake and estimates from either PortionSize or MyFitnessPal). Based on a new power analysis and variance estimates from a recently completed pilot study, equivalence is set as within ± 17.9% of the directly weighed value. The power assumed: 42 participants, energy (kcal) intake as the outcome variable, alpha = 0.05, power >0.80 as acceptable, and mean weighed intake = 2,360 kcal with an SD (difference) = 917 kcal. We will also test for equivalence at the original bound of ± 9%, as well as a new bound of ± 14%, though we recognize that these tests will likely be underpowered. Finally, with at least 42 participants completing the PortionSize and MyFitnessPal condition, we have 0.80 power to detect an R2 of 0.17, which is acceptable as studies with Bland-Altman analysis find that poor methods of measurement frequently have R2 $\geq$ 0.16.[20]

With 42 participants in each group, power is limited for the tests to examine differences in error between PortionSize™ and MyFitnessPal; therefore, these comparisons are exploratory.

The reliability of each apps' food intake estimates will be quantified with intraclass correlation coefficients (ICCs), and reliability will be supported if the ICC is significant and $\geq$ 0.50. A power analysis indicates that, with 2 observations per subject, an ICC of 0.54 can be detected with 23 subjects at 80% power. The power analysis was conducted in R software using the 'ICC.Sample.Size' package [21] based on Zou. [22] Regression-based exploratory analyses are powered to detect an R^2 of 0.127 and thus have ample power.

Provisions to Monitor the Data to Ensure the Safety of Subjects

Adverse events will be monitored at each assessment visit. The PIs and their co-investigators will review all data continuously to ensure the safety of each subject.

Withdrawal of Subjects

Subjects may be withdrawn from the study at the PI's discretion for, for example, missing study visits and failing to follow study procedures. Subjects will be notified of their withdrawal via telephone, email, or mail. If a subject voluntarily withdraws from the study, no additional data will be collected, and they will be considered dropouts in the study.

Risks to Subjects

This study involves no greater than minimal risk. The main risk is breach of confidentiality, and the PBRC team will work to minimize this during data collection, handling, and analysis.

Because of the way our meals are prepared for research and the possibility that the ingredients in the foods we get from commercial vendors could change at any time without our knowledge, it cannot be guaranteed that allergens will be identified and

removed from the foods used in our research studies. If a participant has a food allergy and they are participating in a study where foods are provided, there is a risk that a participant could have an allergic reaction. All participants with known life-threatening food allergies will be asked to inform staff of their allergies.

Potential Benefits to Subjects

Participants may benefit by an increased awareness of their portion sizes and the amount of food consumed in various food groups.

Sharing of Results with Subjects

The participants may receive information such as body weight. Such results will be provided to the participant at the end of the study if requested. Food intake data are readily apparent in both the PortionSize™ app and the MyFitnessPal app; hence, those data will be available during participation.

Setting

The Ingestive Behavior, Weight Management, and Health Promotion Laboratory, and Metabolic Kitchen at Pennington Biomedical Research Center.

Compensation

Participants will receive up to $250 for the successful completion of the study. Participants can receive up to $125 for completing Period 1 and up to $125 for completing Period 2 and final measures.

Provisions to Protect the Privacy Interests of Subjects

All attempts will be made to maintain a subject's privacy. Safeguards such as password-protected computers and networks have been put in place to limit access to subject data. Subjects will be given ample time to read over the consent, ask questions, and agree to participate in the research study. Subjects may decline to answer questions with which they are not comfortable. Each procedure will be explained to the subject before it is performed. We will always ensure the privacy of the subjects.

Participants will be advised to read the current version of the Privacy Statement for MyFitnessPal provided as a link within the consent form to understand how the company will use their information. The current version can be found here:

https://account.underarmour.com/en-us/privacy

Compensation for Research-Related Injury

No compensation will be provided for research-related injury.

Economic Burden to Subjects

All study-related tests and procedures will be at no cost to the subject. The subject will incur transportation costs in getting to Pennington Biomedical Research Center. It is also possible that the subject will incur data charges by using their own smartphone during the study. The chances of this occurring can be mitigated by connecting their smartphone to WiFi while at the Center and while downloading and using the applications.

Consent Process

The PI or one of the designated staff will obtain informed consent during the Screening Visit. Ample opportunity will be given for the subject to review the consent and ask any questions prior to signing the consent form. If subjects wish, they can take the form home and return at a different visit. We are aware that consent is an ongoing process.

References

1. Swinburn BA, Caterson I, Seidell JC, James WP. Diet, nutrition and the prevention of excess weight gain and obesity. *Public Health Nutr*. Feb 2004;7(1A):123-46.
2. Ley SH, Hamdy O, Mohan V, Hu FB. Prevention and management of type 2 diabetes: dietary components and nutritional strategies. *Lancet*. Jun 7 2014;383(9933):1999-2007. doi:10.1016/S0140-6736(14)60613-9
3. Huang T, Yang B, Zheng J, Li G, Wahlqvist ML, Li D. Cardiovascular disease mortality and cancer incidence in vegetarians: a meta-analysis and systematic review. *Ann Nutr Metab*. 2012;60(4):233-40. doi:10.1159/000337301
4. Willett WC. *Nutritional epidemiology*. Oxford University Press; 1998.
5. Lytle LA, Nicastro HL, Roberts SB, et al. Accumulating Data to Optimally Predict Obesity Treatment (ADOPT) Core Measures: Behavioral Domain. *Obesity (Silver Spring)*. Apr 2018;26 Suppl 2:S16-S24. doi:10.1002/oby.22157
6. Beaton GH, Burema J, Ritenbaugh C. Errors in the interpretation of dietary assessments. *Am J Clin Nutr*. Apr 1997;65(4 Suppl):1100S-1107S.
7. Tran KM, Johnson RK, Soultanakis RP, Matthews DE. In-person vs telephone-administered multiple-pass 24-hour recalls in women: validation with doubly labeled water. *J Am Diet Assoc*. Jul 2000;100(7):777-83.
8. Schoeller DA, Bandini LG, Dietz WH. Inaccuracies in self-reported intake identified by comparison with the doubly labelled water method. *Can J Physiol Pharmacol*. Jul 1990;68(7):941-9.
9. Bandini LG, Schoeller DA, Cyr HN, Dietz WH. Validity of reported energy intake in obese and nonobese adolescents. *Am J Clin Nutr*. Sep 1990;52(3):421-5.
10. Archer E, Pavela G, Lavie CJ. The Inadmissibility of What We Eat in America and NHANES Dietary Data in Nutrition and Obesity Research and the Scientific Formulation of National Dietary Guidelines. *Mayo Clin Proc*. Jun 5 2015;doi:10.1016/j.mayocp.2015.04.009

11. Martin CK, Han H, Coulon SM, Allen HR, Champagne CM, Anton SD. A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. *Br J Nutr*. Feb 2009;101(3):446-56. doi:10.1017/S0007114508027438
12. Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International*

Journal of Human-Computer Interaction. 1995;7(1):57-78.
13. Lewis JR. Usability: Lessons Learned ... and Yet to Be Learned. *International Journal of Human-Computer Interaction*. 2014;30(9):663-684.
14. Al Ayubi SU, Parmanto B, Branch R, Ding D. A Persuasive and Social mHealth Application for Physical Activity: A Usability and Feasibility Study. *JMIR Mhealth Uhealth*. May 22 2014;2(2):e25. doi:10.2196/mhealth.2902
15. English LL, Dunsmuir D, Kumbakumba E, et al. The PAediatric Risk Assessment (PARA) Mobile App to Reduce Postdischarge Child Mortality: Design, Usability, and Feasibility for Health Care Workers in Uganda. *JMIR Mhealth Uhealth*. Feb 15 2016;4(1):e16. doi:10.2196/mhealth.5167
16. Limentani GB, Ringo MC, Ye F, Bergquist ML, McSorely EO. Beyong the t-test: statistical equivalence testing *Analyitical Chemistry*. 2005:221-226.
17. Walker E, Nowaki AS. Understanding Equivalence and Noninferiority Testing. *J Gen Intern Med*. 2010;26(2):192-196.
18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. Feb 8 1986;1(8476):307-10.
19. Diet History Questionnaire, Version 3.0. National Institute of Health, Applied Research Program. https://epi.grants.cancer.gov/dhq3/
20. Bray GA, DeLany JP, Volaufova J, Harsha DW, Champagne C. Prediction of body fat in 12-y-old African American and white children: evaluation of methods. *Am J Clin Nutr*. Nov 2002;76(5):980-90. doi:10.1093/ajcn/76.5.980
21. *Calculation of Sample Size and Power for ICC*. 2015.
22. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Stat Med*. Dec 20 2012;31(29):3972-81. doi:10.1002/sim.5466