Comparison of Known Food Weights with Image-Based Portion-Size Automated Estimation and Adolescents’ Self-Reported Portion Size

Christina D. Lee,1 Junghoon Chae, B.S.,2 Tusa Rebecca E. Schap, M.Sc., R.D.,3 Deborah A. Kerr, Ph.D.,4 Edward J. Delp, Ph.D.,2 David S. Ebert, Ph.D.,2 and Carol J. Boushey, Ph.D., M.P.H., R.D.3,5

Abstract

Background:
Diet is a critical element of diabetes self-management. An emerging area of research is the use of images for dietary records using mobile telephones with embedded cameras. These tools are being designed to reduce user burden and to improve accuracy of portion-size estimation through automation. The objectives of this study were to (1) assess the error of automatically determined portion weights compared to known portion weights of foods and (2) to compare the error between automation and human.

Methods:
Adolescents (n = 15) captured images of their eating occasions over a 24 h period. All foods and beverages served were weighed. Adolescents self-reported portion sizes for one meal. Image analysis was used to estimate portion weights. Data analysis compared known weights, automated weights, and self-reported portions.

Results:
For the 19 foods, the mean ratio of automated weight estimate to known weight ranged from 0.89 to 4.61, and 9 foods were within 0.80 to 1.20. The largest error was for lettuce and the most accurate was strawberry jam. The children were fairly accurate with portion estimates for two foods (sausage links, toast) using one type of estimation aid and two foods (sausage links, scrambled eggs) using another aid. The automated method was fairly accurate for two foods (sausage links, jam); however, the 95% confidence intervals for the automated estimates were consistently narrower than human estimates.

Conclusions:
The ability of humans to estimate portion sizes of foods remains a problem and a perceived burden. Errors in automated portion-size estimation can be systematically addressed while minimizing the burden on people. Future applications that take over the burden of these processes may translate to better diabetes self-management.

Introduction

In 2010, there were approximately 25.8 million people or 8.3% of the U.S. population affected by diabetes. Noninsulin-dependent diabetes, adult-onset diabetes, or type II diabetes that is now referred to as type 2 diabetes composes about 90–95% of those with diabetes. Currently, type 2 diabetes is an epidemic in the United States and worldwide, affecting both children and adults largely due to the rising rates of obesity. There are many long-term complications that come with diabetes, such as retinopathy, nephropathy, and peripheral neuropathy, which are associated with high medical costs. More importantly, diabetes was the seventh leading cause of death by diseases in 2010.

Diabetes self-management is an integral component of diabetes care and requires considerable attention to diet, exercise, and medication use. For most persons with diabetes, managing diet and exercise routines can be challenging. Therefore, to enhance self-management of diabetes, industry and researchers have turned to technology. A growing body of literature suggests that using information technology such as computers, the Internet, multimedia, and mobile devices could improve and positively affect the process of care for people with diabetes.

Diet is a critical element of self-management and probably the most challenging to assess and monitor. To date, few mobile technologies that address dietary adherence have been developed. Individuals with diabetes described recording food intake using a smartphone, personal computer, and a blog as motivating; however, several usability problems were identified as reducing the likelihood of sustained use of these tools. Specifically, the steps needed to identify the foods and estimate the portion and serving sizes were described as limited and lacking flexibility. An emerging area of research is in the use of images for dietary records using small mobile devices with embedded cameras (e.g., mobile telephone). These tools are being designed to identify foods and beverages and their portion sizes accurately through automation. The objectives of this article are to (1) assess the error surrounding the mean estimate of automatically determined portion weights based on images taken by adolescents compared with known portion weights of selected foods and beverages, (2) compare the error between automation and human estimation of selected foods, and (3) envision future application of these automated, image-based methods for diabetes management and prevention.

Methods

Recruitment and Study Design

A convenience sample of healthy adolescents between 11 and 18 years was recruited from the local community. On the day of the study, participants were transported to a university campus early in the morning prior to consuming any food or beverages. The participants were served all meals at set times and snacks were provided ad libitum over a 24 h period while being closely monitored. Between eating occasions, camp-like activities were provided. At the end of the day, all individuals estimated the portion sizes of their breakfast foods. The study methods were approved by the Purdue University Institutional Review Board, and informed assent and consent were obtained from the volunteers and their parents, respectively.

Eating Occasions and Use of the Mobile Telephone Application

Foods served represented common foods reported by adolescents. For each eating occasion, all foods and beverages were preweighed separately to one-tenth of a gram prior to plating. Participants received instruction for using a mobile telephone to capture images of each eating occasion. In order to obtain an image useful for image analysis, participants were instructed to include in each image (1) all food and beverages and (2) the fiducial marker, a small credit card-sized item used for color correction and volume (Figure 1A). Participants were instructed to eat to satiation and to request seconds, if desired. The procedures for capturing images were then repeated for any additional portions.

HTC p4351 mobile telephones (HTCAmerica, Bellevue, WA) running Windows Mobile 6.0 (Microsoft, Redmond, WA) were used. The user was given a choice to retake the image or save the image. Once the user was satisfied with the image, the mobile telephone prompted the user to eat before proceeding to the next screen as shown in Figure 1A. After eating, the user was prompted to take an image of the place setting regardless of whether food and beverages remained. Figure 1B is an example of the final screen with before-and-after images.
**Self-Reported Portion-Size Estimation of Breakfast Foods**

While viewing an image of their breakfast meal, participants were asked to estimate the amount of each food item consumed 14 h after the breakfast meal. Numerous methods for estimating portion size are available so at least two of the most common methods of portion-size estimation were used: (1) multiple measurement descriptors (MDes) pertinent to each specific food from the What’s In The Foods You Eat Search Tool, 3.0. (http://www.ars.usda.gov/Services/docs.htm?docid=17032) and (2) two-dimensional (2D) food portion visual with 2D images of standard-sized plates and bowls with cubes depicting ¼ cup, ½ cup, 1 cup, and 2 cups (Block Dietary Data Systems, Berkeley, CA; http://www.nutritionquest.com). Self-reporting of beverage portions was not done. Participants were randomly divided into two groups. One group ($n = 8$) used the 2D portion estimation aid and the other group ($n = 7$) used the MDes portion estimation aid. For portion evaluation, the self-reported intake of each food was converted into grams.

**Automated Portion-Size Estimation of Foods and Beverages Served**

Images from the mobile telephone were sent to a server for image analysis. Methods for automatic identification of food using image analysis have been described previously. Methods used to automatically estimate volume of foods served using computer algorithms have been previously described. Briefly, once a food was identified automatically, volume was estimated using camera location, orientation, and other parameters for use in 3D reconstruction of the food volume from an image. A fiducial marker (e.g., checkerboard square in Figure 1B) was used for size and spatial location of food on the plate.

The system partitioned the space of the food objects into two geometric shapes, cylinders and squares, each with their own set of parameters. The spherical approximation models drew upon spheres and prismatic approximation models. For the foods and beverages served at meals, the automated volume was estimated as cubic centimeters and converted into weight (g) using density values derived from rapeseed volumeter measures of duplicate plates of each meal.

**Data Analysis**

Means of the gram weights of each food and beverage actually served during meals and consumed were computed. Means of the estimated portions, i.e., automated and self-report, were computed. Accuracy of the portion-size estimates was assessed by plotting the mean of the ratio of the estimates, i.e., automated and self-reports, against the actual weights of the foods and beverages, as well as, the 95% confidence intervals (CIs) of the ratios. Thus, a ratio >1 would indicate overestimation and a ratio <1 would indicate underestimation. Descriptive analysis included frequencies and percents. When needed, differences between quantitative variables were assessed using a paired $t$-test. All analyses were conducted using SPSS 17.0.

**Results**

A total of 15 adolescents (3 girls, 12 boys) participated in this study and took images of their meals. Each before image contained the foods and beverages served. A total of 19 unique foods were served over the three meals. Milk was served at each of the meals, thus 45 images that included milk were available for image analysis. Soda was served at lunch and dinner, thus 30 images were used for image analysis. For the 17 remaining foods served, only 15 images (i.e., 1 for each participant) contained those foods. The number of children self-reporting food portions using two different methods for the breakfast foods only is shown in Table 1. An example of a before image from the breakfast meal as taken by a participant is also shown in Table 1. The comparison of known gram weights to automatic gram weights by all foods is shown in Table 2.

Ratios of the automatic estimation to actual weights and the 95% CIs for breakfast foods are shown in Figure 2. Accuracy of the self-reported portions, as depicted by
Comparison of Known Food Weights with Image-Based Portion-Size Automated Estimation and Adolescents’ Self-Reported Portion Size

Lee

www.journalofdst.org
431

Table 1.
Foods Served to Adolescents (n = 15, 11–18 years) at Breakfast and the Number of Self-Reported Food Portions

<table>
<thead>
<tr>
<th>Breakfast</th>
<th>Food descriptions</th>
<th>Self-report per food</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2D&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Scrambled eggs</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Sausage links</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>White toast</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Margarine</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Strawberry jam</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Orange juice</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2% milk</td>
<td>—</td>
</tr>
</tbody>
</table>

<sup>a</sup> Example of an image of foods and beverages taken by one of the adolescents.

<sup>b</sup> 2D portion-size estimation aid used for self-reported food portions (2% milk and orange juice not estimated).

<sup>c</sup> MDes is multiple descriptors of common household and weight measures used, e.g., cups, teaspoon, etc. (2% milk and orange juice not estimated).

Table 2.
Automated Volume Analysis Converted to Weight (g) and Energy (kcal) Compared with Known Weight and Energy Based on Images Taken by 15 Adolescents (11–18 years) of Foods During a 24 h Period under Controlled Conditions

<table>
<thead>
<tr>
<th>FNDDS food code</th>
<th>Brief name</th>
<th>Mean weight (g ± SD)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Ratio of estimate to known&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Energy (kcal ± SD)&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Known</td>
<td>Estimate measures</td>
<td>n</td>
</tr>
<tr>
<td>11112110</td>
<td>2% Milk</td>
<td>45</td>
<td>220.0 ± 0.0</td>
<td>208.7 ± 9.8&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>25221660</td>
<td>Sausage links</td>
<td>15</td>
<td>46.5 ± 1.0</td>
<td>41.5 ± 2.8&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>32105000</td>
<td>Scrambled eggs</td>
<td>15</td>
<td>61.5 ± 0.7</td>
<td>108.5 ± 27.4&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>51101010</td>
<td>Toast</td>
<td>15</td>
<td>47.7 ± 3.4</td>
<td>80.0 ± 17.9&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>51121040</td>
<td>Garlic bread</td>
<td>15</td>
<td>41.1 ± 3.0</td>
<td>119.8 ± 15.3</td>
</tr>
<tr>
<td>53106050</td>
<td>Chocolate cake w/ icing</td>
<td>15</td>
<td>81.5 ± 12.5</td>
<td>105.7 ± 17.5&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>53241500</td>
<td>Sugar cookie</td>
<td>15</td>
<td>27.8 ± 1.9</td>
<td>31.6 ± 3.2&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>58132310</td>
<td>Spaghetti w/ sauce, cheese</td>
<td>15</td>
<td>240.3 ± 2.6</td>
<td>214.5 ± 60.9&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>61202220</td>
<td>Orange juice</td>
<td>15</td>
<td>124.0 ± 0.0</td>
<td>128.6 ± 10.3&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>63135140</td>
<td>Peaches</td>
<td>15</td>
<td>69.3 ± 9.9</td>
<td>116.1 ± 18.4&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>63137170</td>
<td>Pear halves</td>
<td>15</td>
<td>75.6 ± 4.9</td>
<td>138.9 ± 20.7&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>71401020</td>
<td>French fries</td>
<td>15</td>
<td>70.5 ± 4.3</td>
<td>204.7 ± 31.1&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>74401010</td>
<td>Ketchup</td>
<td>15</td>
<td>15.5 ± 0.4</td>
<td>17.0 ± 7.1&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>75114000</td>
<td>Lettuce (salad)</td>
<td>15</td>
<td>48.3 ± 4.8</td>
<td>220.3 ± 35.5&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>81103040</td>
<td>Margarine</td>
<td>15</td>
<td>27.8 ± 0.6</td>
<td>40.4 ± 12.4&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>83202020</td>
<td>French dressing</td>
<td>15</td>
<td>35.7 ± 1.0</td>
<td>32.7 ± 1.5&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>91402000</td>
<td>Strawberry jam</td>
<td>15</td>
<td>21.1 ± 1.1</td>
<td>21.4 ± 5.3</td>
</tr>
<tr>
<td>92410310</td>
<td>Coke</td>
<td>30</td>
<td>227.2 ± 2.3</td>
<td>305.7 ± 27.6&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>99999999&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Cheeseburger sandwich</td>
<td>15</td>
<td>198.8 ± 11.5</td>
<td>187.2 ± 34.5</td>
</tr>
</tbody>
</table>

FNDDS, food and nutrient database for dietary studies; SD, standard deviation.

<sup>a</sup> Paired t-tests were used to evaluate differences between known and estimated values.

<sup>b</sup> Ratio of estimated weight to known weight. A value >1 indicates an overestimation. A value <1 indicates an underestimation.

<sup>c</sup> p < .001.

<sup>d</sup> p < .01.

<sup>e</sup> 99999999 represents a combination food composed of six separate foods with food codes within FNDDS.
Comparison of Known Food Weights with Image-Based Portion-Size Automated Estimation and Adolescents’ Self-Reported Portion Size

Lee

www.journalofdst.org


A common problem in using self-reported dietary intake data is the amount of error present, which reduces the ability to find statistically significant associations between diet and health outcomes. In certain cases (i.e., sausage links, toast, scrambled eggs), the adolescents estimated portion size better than the automated estimate. However, in all cases, the degree of error, or spread around the means for each food and beverage, was substantially less for the automated estimates over the human estimated amounts. The potential for reduced error in portion-size estimation due to investment in automation is promising. In the future, data generated from automation may mimic the results of nutrition studies in controlled clinical studies.

Using simple shapes, the automated volume estimation differs substantially from human estimation. Human estimation is biased by social desirability or lack of knowledge about sizes, whereas the automated volume estimation is not influenced by these factors. Challenges for the automated system include foods with void space or porosity, reflection from containers, and containers. Therefore, a food such as margarine may be overestimated, an unlikely occurrence among humans who are self-reporting portions. At this point, the results of the correlation between the true energy served and energy based on image analysis for volume still need improvement. The development of an expanded selection of shape templates may enhance results, as well as the benefits accrued from the rapid advances in technology.

Portion-size estimation is difficult for people and regarded as burdensome. Several attempts have been made to provide creative methods to improve portion-size estimation. For example, Matheson and colleagues provided 8–12-year-old girls with clay to mold and shape a known quantity of a bread stick and crinkled paper strips for spaghetti with sauce and salad with dressing. This method worked well for some foods. Correlations between actual and estimated intakes with two portion-size measurement aids were high for three foods ($r = 0.56–0.79$, all $p < .001$) and low for the bread ($r = 0.16$, $p = .43$). Translating this method into practice, i.e., providing individuals with clay and paper for quickly estimating dietary intakes, is questionable. Despite attempts to improve human’s abilities to estimate portion sizes, accuracy remains elusive. Conversely, advancements in technology that can equate to improvements in automation can

Discussion

To our knowledge, this is the first report of automatic volume estimation of food in meals over an entire day using images taken by adolescents. The weights of some foods and beverages were estimated fairly close and others were not. Of the 19 foods and beverages that represented commonly consumed foods, about 50% were estimated within an acceptable range, i.e., within 15% of truth (2% milk, sausage links, sugar cookie, spaghetti with sauce and cheese, orange juice, ketchup, French dressing, strawberry jam, cheeseburger sandwich). On the other hand, a food, such as lettuce, which has a large amount of void space, was overestimated by over 400%. Because of the low energy density of lettuce, the average energy difference between known weight and estimated weight was less than 30 kcal, which is small for an entire day. The gram weights were overestimated for at least 12 of the 20 foods, but depending on the food, the error in energy varied. For all foods combined, the mean error was overestimated, suggesting that the energy dense foods seem to be most affected by any error. Thus, energy density of a food can help guide the focus of which foods need the most attention with regard to improving the accuracy of automated estimation.

A common problem in using self-reported dietary intake data is the amount of error present, which reduces the ability to find statistically significant associations between diet and health outcomes. In certain cases (i.e., sausage links, toast, scrambled eggs), the adolescents estimated portion size better than the automated estimate. However, in all cases, the degree of error, or spread around the means for each food and beverage, was substantially less for the automated estimates over the human estimated amounts. The potential for reduced error in portion-size estimation due to investment in automation is promising. In the future, data generated from automation may mimic the results of nutrition studies in controlled clinical studies.

Using simple shapes, the automated volume estimation differs substantially from human estimation. Human estimation is biased by social desirability or lack of knowledge about sizes, whereas the automated volume estimation is not influenced by these factors. Challenges for the automated system include foods with void space or porosity, reflection from containers, and containers. Therefore, a food such as margarine may be overestimated, an unlikely occurrence among humans who are self-reporting portions. At this point, the results of the correlation between the true energy served and energy based on image analysis for volume still need improvement. The development of an expanded selection of shape templates may enhance results, as well as the benefits accrued from the rapid advances in technology.

Portion-size estimation is difficult for people and regarded as burdensome. Several attempts have been made to provide creative methods to improve portion-size estimation. For example, Matheson and colleagues provided 8–12-year-old girls with clay to mold and shape a known quantity of a bread stick and crinkled paper strips for spaghetti with sauce and salad with dressing. This method worked well for some foods. Correlations between actual and estimated intakes with two portion-size measurement aids were high for three foods ($r = 0.56–0.79$, all $p < .001$) and low for the bread ($r = 0.16$, $p = .43$). Translating this method into practice, i.e., providing individuals with clay and paper for quickly estimating dietary intakes, is questionable. Despite attempts to improve human’s abilities to estimate portion sizes, accuracy remains elusive. Conversely, advancements in technology that can equate to improvements in automation can
Comparison of Known Food Weights with Image-Based Portion-Size Automated Estimation and Adolescents’ Self-Reported Portion Size

Lee

offer substantially more benefits, e.g., low burden and more precision, than pursuing methods that involve an unreasonable amount of time on the part of individuals.

Besides the necessity to improve portion-size estimation prior to launching a system that can provide immediate feedback to a user, other technology challenges need to be addressed. The computational capacity of mobile telephones is still incapable of running the image analysis for automatic food identification and portion estimation. Thus, either the computational capacity of mobile telephones or speed of data transfer must improve substantially. However, given the proliferation of technological advances, the introduction of these tools for use in the future is assured.

Conclusions

Environmental factors such as diet and exercise play significant roles in the prevention of type 2 diabetes\textsuperscript{32} and in self-management of diabetes. These lifestyle factors are potentially modifiable.\textsuperscript{33} Progress in mobile technologies holds promise to reduce health care costs and move medical care more toward preventive care.\textsuperscript{34} Figure 3 shows an example of the type of message that could be displayed on the screen of a mobile telephone in the future that would immediately inform individuals about their diets. The information regarding the amount of carbohydrate or other key nutrients can assist with modifications in food intake and medication levels. Although individuals with diabetes found recording food using smartphones to be motivating, they were less likely to continuously use these applications that required them to manually identify foods and estimate portion sizes.\textsuperscript{14} Applications that take over the burden of these processes, as described in this article, may in the future encourage regular daily use, which may translate to better self-management.

Funding:

Support for this work comes from the National Cancer Institute (U01CA130784-01) and the National Institute of Diabetes, Digestive, and Kidney Disorders (R01-DK073711-01A1). TusaRebecca E. Schap is supported through the NIH/NCRR—Indiana Clinical and Translational Sciences Institute—TL1 Program (TL1 RR025759).

References:


Figure 3. A futuristic example of immediate dietary feedback using automated food identification and volume estimation running on a mobile application.