Article:

Amoutzopoulos, B, Page, P, Roberts, C et al. (8 more authors) (2020) Portion size estimation in dietary assessment: a systematic review of existing tools, their strengths and limitations. Nutrition Reviews, 78 (11). pp. 885-900. ISSN 0029-6643

https://doi.org/10.1093/nutrit/nuz107

Portion size estimation in dietary assessment: a systematic review of existing tools, their strengths and limitations

INTRODUCTION

Assessment of dietary intake is critical for nutrition research and surveillance programs to inform public health policy and for evaluating and measuring the effectiveness of interventions across populations. A well-recognized challenge in dietary assessment is the accurate estimation of portion sizes[1-5]. Traditional dietary assessment methods such as 24 hour (24hr) recalls, food-frequency questionnaires (FFQs), and unweighed food records are subject to measurement error resulting from various factors such as food matrix, demographic characteristics (e.g. gender, age, ethnicity, education and income) and the nature of the dietary assessment instrument, especially if there is a need to recall consumed amounts of foods from memory[6-10]. Other factors that may add to measurement error are nutrition policies, religion[11], food familiarity, hunger status and the expected filling capacity of the food[12], meal type and its energy density[10,13,14]. Misreporting of consumed amounts is a fundamental issue affecting the capture of accurate habitual dietary intake data[6]. A study conducted in Finland[15] showed that only about 50% of the study population (n=146) estimated the amount of 52 foods correctly, with underreporting being common for bread, spreads, cold cuts and dishes and over-reporting observed for cereals, snacks, vegetables and fruit. However, the use of portion size estimation elements (PSEEs) such as food models, household utensils, photographs or diagrams can aid respondents to report their food intake more accurately and reduce respondent burden associated with weighing of foods[16], although their effectiveness is subject to individual use and customary dietary patterns[11].

Some studies[4,17-21] have explored ways to improve portion size estimation in dietary assessment. In some cases, the performance of these instruments depended heavily on the shape and texture of foods[22,23]. In general, the literature suggests that solid foods are better estimated than liquids and these are estimated better than amorphous foods (i.e. those which take the shape of their container e.g. pasta)[24]. A review[25] of efficacy studies concluded that food type, shape, size and previous training all affected tool performance. Portion size estimation continues to be a key factor in dietary assessment error[17,26-28], and there is a lack of knowledge on the strengths and limitations of different PSEEs and quality studies testing validation and efficiency[5,29-31]. To our knowledge only three reviews[4,6,25] have focused on PSEEs, two[6,25] of which analyzed only a limited number of studies according to broad categories and covered short time periods i.e. 2005 to 2016. The other review[4] was a parallel review led by the authors of this present review which specifically focused on PSEEs for minority ethnic groups[4]. Two studies[6,31] highlighted the need for additional information and guidelines on the validity of PSEEs. This is because some PSEE validation studies suffer from issues such as failing to use an approach to isolate the specific effects of PSEEs from the effects of dietary assessment methods; lack of testing of the tools in relation to different sociodemographic, age, gender, and cultural groups; using non-objective comparators instead of using premeasured amounts of foods; conducting data analyses which are focused on nutrients rather than food amounts; and providing sufficient level of detail on PSEE descriptions and their use (such as dimensions of PSEE, modifications, how they were presented to the subject, methods used in accuracy measurement). The lack of clear, descriptive, and complete research on the validation of PSEEs makes it currently challenging to establish a consensus on the validity of PSEEs, to draw conclusions about guidelines for the use of PSEEs and to comment of the validity of PSEEs[6,31]. Finding solutions to these issues would improve the accuracy of dietary assessment methods, knowledge of the diet-disease relationships, nutritional monitoring of populations and ability to educate the public on portion size measurement for healthy eating[31]. To address these gaps, the present review focused on characterizing and assessing the validity of existing PSEEs applicable to dietary assessment instruments. For the purpose of this work, PSEEs were defined as a component of the dietary assessment instrument designed to help quantify the amount of food reported as consumed including: portion size estimation aids (PSEA) (e.g. photographs, everyday reference objects, household utensils, food models); categorical size estimates (e.g. small, medium, large); household utensil measures, unit food amounts (e.g. 1 slice, 1 egg), standard units of measurement (e.g. grams, ounces, milliliters) and any other quantifying component.

The objectives of this review were, first, to explore the range of existing PSEEs applicable to dietary assessment methods and categorize them according to their applications. Second, to assess the quality of existing studies validating PSEEs and to develop a tool to assess the quality of the studies validating PSEEs. Third, to explore the relative efficacy of tools tested through validation and comparison studies, as well as addressing the limitations in these studies. This information may inform the design of future nutrition surveys and prospective cohort studies, support dietary assessment in clinical practice and research and may also contribute to reduce misreporting by guiding researchers on selecting high quality PSEEs, and may improve the quality of validation and comparison studies.

METHODS

Database searches

A systematic review of the literature was first conducted in 2016 and an update was undertaken between March and June 2018, based on standard systematic review guidelines[32,33], for records published between 1910 and 2018 (see the PRISMA[34] diagram (Figure 1) and checklist in Appendix 1). The study protocol is available by contacting the authors. We selected studies for review using PICOS (population, intervention, comparison group, outcome, study design) criteria (Table 1).

Studies were excluded if they reported on the use of a dietary assessment instrument without a portion size measuring element (e.g. non-quantitative FFQs); or when the PSEE was not described in full; or it was not applicable for dietary assessment. Titles with no accessible abstracts, editorials, and commentary or opinion pieces, review papers with no relevant references and non-English language papers were also excluded.

In total, 20 medical, social and economic databases (See Figure 1) were searched. In addition all the references from a relevant PhD thesis[35] and a previous review[6] were screened. Title search was complemented through cross-reference and the authors' own knowledge.

A search pathway containing keywords and combinations for the searches was designed and pre-piloted by two of the authors (CG and EAR) (see the Search Pathway Form (key words) in Appendix 2 in the online Supporting Information). Searches were structured in blocks containing descriptors for portion size estimation elements. The following block themes were used: portion size; tool; measures; assessment; quantity; dietary; electronic; foods; texture; and target population characteristics. Each block consisted of at least 3 descriptors. For instance, the block 'portion' consisted of 'portion OR serving OR helping'; the block 'tool' consisted of 'Tool* OR utensil* OR appliance* OR guide* OR instrument*', and so on. We then searched 19 different combinations of the above descriptor blocks. In order to reduce the number of ineligible hits in combinations producing more than 1000 hits, we excluded abstracts where the words “portion” and “size” were not within 3 words of each other.

Title and abstract screening and data extraction were conducted by five investigators (CG, EAR, DY, TH and RB). A subsample of abstracts was screened in duplicate to assess consistency between reviewers. Disagreements were discussed within the team to reach consensus and, when necessary, further information was sought from authors. If a paper´s abstract did not contain sufficient information to confirm eligibility, the whole paper was reviewed.

Data extraction of eligible abstracts and papers involved extracting information on the instrument description (i.e. name of PSEE, dimension, format of usage and dietary assessment instrument to which the PSEE belonged, plus purpose of PSEE in the study) and the population (i.e. country, nationality, age, setting and health status) where the PSEE was applied. Risk of bias in individual studies was assessed by looking at the study design; outcomes and analysis; plus other strengths or limitations using adapted versions of published resources[36,37]. Analysis of risk of bias across studies was not applicable as this review is meant to inform decisions across a variety of settings[36].

Selection of category and grouping criteria

PSEEs were categorized based on dimension and format of usage to reflect their different measuring scope. For dimension, PSEEs were categorized as follows:

- One- dimensional tools included image-free, non-volumetric tools such as lists of portion size options including numerical values as well as categorical size estimates (i.e. large, medium, small), lists of open-ended questions where an amount was requested, lists of household units (e.g. number of tablespoons) as part of questionnaires or food guides, portion information in text form on food packaging, or till receipts and voice recordings.

- Two- dimensional tools included all image-based tools such as paper-based or electronic/computer-based food images (i.e. photos, diagrams or other pictorial representations), and images of; hand-based portion measurements, non-food objects, food models, food replicas and measuring utensils.
- Three- dimensional or volumetric tools included food models and replicas (i.e. models imitating the color, shape and texture of foods, and non-food objects such as sticks, boards, circles and cartons); measuring utensils (i.e. tablespoon, measuring cup, measuring jug, ruler; food scales) and hand measures (e.g. size of palm and width of fingers).

For format of usage, PSEEs were categorized as follows:

- Stand-alone: PSEEs that were used individually as part of the dietary assessment method, for example a list of portion size options as part of a semi-quantitative FFQs and a set of measuring spoons as part of a 24hr recall.
- Related set: PSEEs used in combination within the same dietary assessment method and measuring the same food dimension (e.g. image or volume), for example a set of measuring spoons used together with a set of measuring cups.
- Combined: PSEEs used in combination within the same dietary assessment method and measuring different dimensions, for example a one-dimensional tool (food packaging information) alongside a three- dimensional tool (set of measuring cups).

Analysis of validation studies

Validation studies were identified from the whole sample of publications. Validation studies were defined as studies comparing the portion size estimates made using a

PSEE against actual weights for the purpose of evaluation or validation of a PSEE. The validation studies were then examined through a quality scoring tool (Table 2) which was established in this study through the investigation of previously published quality scoring systems used for dietary assessment tools[38-40]. An initial version of the quality scoring tool was piloted twice by three independent investigators (EAR, BA, EV) and the first and last versions were evaluated for content validity against expert opinion within the team (JC, MR). A scoring system was applied to standardize data collection (Appendix 3 in the Supporting Information online). Finally, two other investigators (RB and BA) scored the validation studies using the quality scoring tool.

Levels of agreement between PSEEs and actual weights were not reported in a consistent way across the validation studies. Studies used various approaches to explain the differences between actual and estimated weight, such as “within 10%, 25% or 33% of true weight”, “percent estimation error”, “differences in mean weight” or “percent of participants making correct estimations”. To overcome this issue, we compared the validity of PSEEs using the accuracy parameters reported in these studies. This was done by grouping the quantitative results of similar accuracy measures (e.g. estimation error) reported in validation studies for each PSEE category (e.g. range of estimation error reported for food models) and comparing these results among each PSEE category (e.g. comparing the range of estimation error reported for food models to the range or estimation error reported for food atlases).

Analysis of comparison studies

In the whole sample, comparison studies were identified as studies which were comparing different PSEEs to each other in terms of efficacy, usability or accuracy. The full text of comparison studies were then examined through extracting data on population characteristics, possible confounders, the context in which PSEEs were compared, statistical tests and any other study outcomes in relation to PSEE performance. The data extraction process was piloted by EAR and then three investigators (TH, RB and BA) extracted the whole data. The studies that were not clearly comparing PSEEs were not included in comparison, for example studies comparing overall dietary assessment tools through measuring difference in nutrient intakes.

Due to the nature of the data, meta-analysis was not suitable for this review, therefore a narrative synthesis of outcomes is presented and the findings are combined in tables and figures when appropriate. There was great variation in study designs and accuracy parameters across the studies, therefore it was not possible to quantify the differences between the accuracy of PSEEs; however an overall assessment of PSEEs was conducted by investigating the accuracy parameters reported in these studies.

RESULTS

Results of all searches (whole sample)

Number of records and PSEEs

In total 16,801 records were identified from initial searches from which a total of 334 records covering 542 PSEEs were selected (Figure 1). The records were published between 1975 and 2018 (with an average of 8.5 records per year). The greatest number of publications was published between 2009 and 2014 (yearly average of 23). Most of the studies were published in the US (n=126) and UK (n=75) followed by Canada (n=15). A list of the 542 PSEEs in the whole sample is given in the Table S1 online Supporting Information.

PSEE categories

The 542 PSEEs identified were categorized according to the format of usage (Table 3). The two most common PSEEs were three (n=263) and two- dimensional (n=249), and these mainly included household utensils and photographic atlases. The one-dimensional PSEEs (n=30) were the least used and they mostly included portion lists and food guides (Table 3).Overall, food photos constituted 18.6% of PSEEs and electronic images and devices constituted 20% of the PSEEs. Of the electronic images and devices, 40% were combined with food records and 17% were combined with 24hr recalls. Among the electronic PSEEs, 26% and 54% were applicable to children and adults respectively and 98% were applicable to developed countries (mainly USA and UK). Two studies[41,42] showed that the digital image assistance improved the overall accuracy of dietary assessments Figure 2 gives information on (a) the purpose of studies using PSEEs, (b) the format of the usage of PSEEs and (c) the dietary assessment instruments of which PSEEs were part of.

Distribution of PSEEs by study purpose

In terms of the purpose of studies (Figure 2a) the most common purpose was the evaluation studies (46%) including validation, comparison and usability testing of dietary assessment tools or PSEEs. Next were the development studies (16%) which included development of tools for estimation of dietary intake or portion size estimation. Of the population studies (12%), 52% and 48% focused on portion size estimation and dietary intake, respectively. Nutrition surveys accounted for 19% of population studies.

Distribution of PSEEs by format of usage

In terms of the format of usage (Figure 2b), most PSEEs (70%) were used in studies as a stand-alone tool which mainly included image-based tools such as food atlases. Within population studies, only one study[43] was using food models for portion size estimation as a stand-alone PSEE whereas 31% and 62% of studies were using food photos and scales as standalone tools, respectively (data not shown). The 60% of stand-alone scales were applied to children, adolescents and older adults (data not shown).

Distribution of PSEEs by the dietary assessment instrument to which they are applied

In terms of the dietary assessment instrument (Figure 2c), not all PSEEs were linked to a dietary assessment instrument (e.g. some studies just measured the serving sizes served at a restaurant or portion sizes served as a school meal). Some PSEEs were identified as a commercial item such as portion size guide book. Food records, including estimated and weighed diet diaries, were the most popular dietary assessment instruments related to PSEEs (21%).

Population distribution across studies

Figure 3 describes the characteristics of populations which used PSEEs in the study settings. In terms of the population origin (Figure 3a), the predominant populations were North Americans (34%) and Europeans (28%), 50% of which were British and Irish. The smallest proportion of PSEEs (3%) were tested in Arab, Eastern and African populations. Some PSEEs (7%) were used in studies focused on specific ethnic groups such as African Americans, South Americans and South Asians living in UK, USA, Canada and Norway (for further details see the parallel review on ethnic PSEEs[4]) (Figure 3a). Only 3% of all PSEEs were identified as being applied to low-income and middle-income countries and only 0.3% of PSEEs were tested in low-income countries (data not shown). Only 3% of PSEEs were tested in South American and South Asian populations (native origin) and 1% of PSEEs were tested in African populations (native origin) (Figure 3a) of which 60% were based on food images[23,44,45] (data not shown).

Age distribution across studies

In terms of age (Figure 3b), adults (54%) were the most dominant age group; of these, 20% were gender-specific. Of the PSEEs applicable to children (16%), 40% and 8% were specifically applied to adolescents and preschool children, respectively (Figure3b). The most popular PSEEs applicable to children in various ages (15%) were food images (27%) and electronic PSEEs (23%). On the other hand, household utensils were not popular in children as much as in adults (data not shown).

Health status distribution across studies

In terms of health (Figure 3c), the majority of PSEEs were tested on healthy people (66%), followed by people with chronic diseases (3%), obesity/overweight (3%) and other health issues (2%). Of all PSEEs, 1% were specifically used with pregnant women.

Study setting distribution across studies

In terms of the study setting (data not shown), most PSEEs were used in free living settings (58%) followed by school or university settings (13%) and institutionalized settings (e.g. care homes) (1%) and other settings (3%) such as dialysis and metabolic units, general practitioners or work. In hospitalized and institutionalized settings, common practice for portion size estimation was weighing the served portion size and the left overs.

Validation studies (Absolute validity)

Quality assessment of validation studies

A total of 21 validation studies were identified from the whole sample of records (n=334). The results of the quality assessment, focused on the validation studies, are shown in Table S2 online Supporting Information. Two validation studies using image-based tools[23,46] achieved the highest score (22 out of 25). In total, six validation studies scored over 19 (equivalent to more than 75% of criteria being met) while the rest of the validation studies (n=16) were scored between 14 and 19. Validation studies that scored over 19 (n=6), were relatively well designed as the sample size was at least 50 participants, study population was representative of the reference population and sufficient detail was provided on population characteristics, plus the comparator involved foods being weighed by investigators as an objective measure. Out of 21 validation studies, six studies included piloting of PSEEs, five assessed the reliability of PSEEs and seven tested agreement by using tests such as Bland-Altman. Most validation studies (20 out of 21) were rated versatile as they tested PSEEs using a good range of food textures, for example solid, semi-solid, liquid and amorphous foods in accordance with the focus of PSEEs. Most of the validation studies (n=15) scored high for potential for long term efficacy as the likelihood of future use or user preference of PSEEs were discussed or implied in the study. In all validation studies, 17 tested two-dimensional PSEEs whereas only three studies tested three- dimensional tools (e.g. tennis ball) and one study tested a one- dimensional tool (till receipt). The median score for food atlases (n=6) was higher than for digital images (n=3) followed by 3D PSEEs (n=3) and other food photos (n=8), 19.7, 19, 18 and 17.8 respectively. Food atlases were paper based books which showed long lists of photos for many foods usually representing staple foods consumed by populations. Other food photos were those not in atlas format with a limited number of photos usually for a selected list of foods.

Comparison of the accuracy levels of validated PSEEs

For food photos, 42%[47] and 55%[22] were reported as accurately estimated For food atlases, 68%[23] and 77%[44] of all portions were accurately estimated, although there was no uniform or clear definition of accuracy in these studies. These findings suggest that for photographic PSEEs, food atlases have greater accuracy compared with food photos. Another study testing a food atlas[45] reported that 54.2% of participants made accurate estimations, again there was no clear definition of accuracy.

One study[48] reported average estimation error as 2.3% for the food atlas. The estimation error for food atlas was also reported as a range of (−36·8) to 17·1g[49], and a mean of 137.6kcal[46]. For food photos, the range of percent of the difference was reported as (-9.9) to 18.6%[50], (-4.1) to 28.6%[20], (-10.7) to 5.3%[51] and 1.0 to 39%[52].

For digital pictures[53] mean relative error was reported as −2.8%, and the estimation error was reported as a range of (−13) to 4g[54] with a mean of 56.7 kcal[54]. These results suggest that the digital images are comparable to the printed food photos[46,49]. Digital images tested on children[55] reported the average estimation error as 32% which was generally higher than food photos tested on adults ((-10.7) to 39%)[46,53,54]. These findings suggest that when food photos are tested only on children the estimation error can be expected to be lower than the PSEEs tested on adults.

Three studies[56-58] tested 3D tools in children or young adults. For manipulative props (crinkled paper strips, clay, water and glass)[56] average estimation error was 58%. For cups and spoons[57] and modeling clay[57] the estimation error was 53.1% and 33.2%, respectively. These findings indicate that the estimation errors for 3D PSEEs ranged from 33.2 to 58% and they were not more accurate than food photos ((-10.7) to 39%)[20,48,50,51]. However a clear comparison was not possible as 3D PSEEs were tested only in children and young adults, and food photos tested in mixed age groups including adults. In addition, a tennis ball[58] achieved a total score of 7.4 (out of 12) where 1 point was assigned to each estimate which was within 33% of the actual size.

Overall, 18 of the 21 validation studies concluded that the tested PSEEs were providing a level of validity or accuracy for the tested population. The present study shows that there was more evidence indicating the validity of food photos compared to food models and household utensils, and this evidence was stronger for the food atlases compared to other food photos.

Comparison studies (Relative validity)

A total of 13 comparison studies were identified from the whole sample of records (n=334). A summary of extracted data is given in Table S3 online Supporting Information. The comparison studies were identified as those comparing different PSEEs to each other in terms of efficacy, usability or accuracy.

Reported average estimation error (compared with actual intake) was 2.3% for photographic food atlas, 56.9% for household utensils (measuring cups) and 32% for food models[48]. For household utensils (cups and spoons)[57], a food model (modelling clay)[57], manipulative props (paper strips, clay, water and glass)[56] and visual food models (e.g. drawings of glasses)[56], estimation error (compared with actual intake) was reported as 53.1%, 33.2%, 58% and 32.8% respectively. One study[59], reported the estimation error (compared with actual intake) as 18.9% for the international food unit (a $64cm^2$ cube divided into 2cm cubes), 87.7% for measuring cup and 44.8% for a food model (modelling clay), however the estimation error for international food unit was large for some foods. The usability of the international food unit was tested and the participants perceived it as the easiest, particularly for foods with geometric shapes. Overall the comparison studies showed that the estimation error ranged from 53.1 to 87.7% for household utensils (including cups and spoons) and ranged from 32 to 44.8% for food models (including modelling clay). This indicates that food models perform better than household utensils. Considering that the validation studies (section 2. Validation studies) showed the estimation error for photographic PSEEs ranged from (-

10.7) to 39%, it can be suggested that photographic PSEEs are more accurate than food models.

One study[60] reported that, for geometrically shaped foods (e.g. cheese and cake) and liquids, 80% of estimations made with hands (finger width, fist, fingertip) and 29% of estimations made with the household utensils (cups and spoons) were within ±25 % of actual weight, and 13% of estimations made with hands and 8% of estimations made with household utensils were within ±10 % of actual weight. However the same study showed that for more irregularly shaped foods (e.g. chicken breast), estimations made with both hand and household utensils were above 50% of actual weight and for amorphous foods no estimations made with hands were within ±10 % of actual weight whereas three estimations made with household utensils were within ±10% of actual weight. Another study[21] showed that 15% of estimations made with digital images were within ±10% of actual weight. These results support the finding that food images may perform better than household utensils and estimations made with hands. For the foods that closely resemble a geometric shape, the hand method can perform better than household utensils; however this is the opposite for irregularly shaped foods.

Four studies[19,56,58,61] compared the images or drawings of non-food objects (e.g. pictures of tennis ball, drawings of glasses) vs actual non-food objects (e.g. tennis ball, glass) and reported no difference in estimation for most of the foods tested. This finding suggests that there is no convincing evidence that there is a difference in accuracy between the images of non-food objects / household utensils and their actual forms.

No differences were found in accuracy between digital images and printed images in two studies[62,63] comparing the same number of images and foods. This may suggest that printed images are comparable to digital images. Another study[21] found no difference between different types (aerial vs angled vs mounds vs household measures), presentations (sequential vs simultaneous) and sizes (large vs small) of digital images. Participant preference supported simultaneous presentation vs sequential and large vs small. Although not statistically significant, the use of 8 digital photos for the portions size estimation of one food achieved higher accuracy than using 4 digital photos. The accuracy results showed that the mean absolute gram weight differences between weighed and reported amounts ranged from 5.8 to 35.6g for 8 photos and from 8.4 to 47.5g for 4 photos.

DISCUSSION

Findings across all studies

In total 542 PSEEs were identified in this review across 334 publications and compiled in a database that represents the first complete inventory of portion size estimation tools to date. The most common type of PSEEs were 3D tools (49%) followed by 2D (46%) with 1D PSEEs (5%) being much less common. The household utensils (41%) were the most popular PSEEs within 3D tools, whereas, similar to the previous research[25], image-based PSEEs (e.g. photographic food atlas) (37%) were the most popular PSEEs within 2D PSEEs. This is probably due to their practicality as household utensils are easily available tools, and food photos are easy to use across populations and able to represent a range of foods[15,46,49,64,65]. Although there is an individual variation in the portion estimation of different foods[14,20,65,66], food photos were also judged as an appropriate PSEE to estimate portions at a population level[22] and they are also a key tool in national nutrition surveys[67,68]. Food photos, especially food atlases developed through systematic procedures[4,45] can also be a valuable instrument for low-middle income settings[22]. If an image based PSEE is going to be used in a different location, food images should be adapted to foods consumed in the targeted region[69].

Similar to food photos, food models and replicas can help respondents in visualizing their portion sizes however their disadvantage is the risk of limiting portion size choices subject to their food range[70]. In this current review, within 63 population studies identified, only one study[43] used food models for portion size estimation as a stand-alone PSEE. This finding may suggest that the food models and replicas are not very suitable for population studies unless they are combined with other PSEEs. Combined PSEE usage has been endorsed by two previous reviews[4,25]. In terms of food scales, these are deemed to be laborious PSEEs in practice[71]. From the studies identified in this review, most of the stand-alone scales were applied to children, adolescents and older adults perhaps to improve the accuracy of portion size estimations in these age groups.

Validation studies (Absolute validity)

In validation studies (n=21) most PSEEs were 2D (81%) and only a few were 3D (19%), although in the whole inventory we found similar proportions of 2D and 3D tools. Another review[25] also showed that 2D PSEEs, especially food photos, were the most common PSEEs assessed for their validity. In this present review, across the validation studies, there was more evidence on the validity of food photos compared to food models and household utensils and therefore food photos were deemed to be the most accurate. The comparisons do not clearly establish if one photographic PSEE is better than the other but there was stronger evidence of the validity of food atlases compared to other food photos. The comparisons also indicate that there is no difference between digital images and printed food photos in terms of accuracy. This is empowering for electronic PSEEs which mostly rely on digital images. However even with tools that demonstrated greater accuracy there can still be very significant errors in intake estimates compared with 'true' intakes. Besides, a tool that has not been validated may still be well developed and useful.

The quality of the PSEE is an important part of any dietary assessment tool as inaccurate portion size reporting can increase the measurement error in dietary intake[4,15]. This review showed that very few studies were of high quality according to the quality scoring tool that was established in this study. Similarly, a previous review[6] focusing on the validity of PSEEs has found that research on the reporting of PSEEs and the investigation of their accuracy was lacking clarity and completeness. In this current review, lower quality was mainly due to small sample size, inadequate description of study population, poor representation of reference population and PSEEs not being piloted and tested for reliability. These are the areas in which future validation studies may need attention.

There was great variation among studies in terms of the approaches used to estimate the level of accuracy. Studies used real time or recall approach and various numbers of food types and portions for portion size estimations or presented foods using self-served vs pre-served portions. The type of foods used in studies also varied greatly as some studies focused on particular foods like bread and cheese whereas others covered full food categories such as beverages, meats etc. Accuracy was not reported in a consistent way across all studies. All these issues have caused a difficulty in comparing the validity of different PSEEs. These issues are similar to the ones raised in a study published in 1995 which highlighted the lack of standards in methodologies used for the validation of portion size estimation methods, and the lack of certainty and comparability across studies with regard to accuracy.

Comparison studies (Relative validity)

Comparison studies varied greatly in their design and methodology which should be taken into account. Overall according to comparison studies, photographic PSEEs performed better than food models and food models performed better than household utensils in terms of the agreement between portion size estimations. These results are in line with a comprehensive comparison study[48] conducted on 463 adolescents using 163 foods which suggested that photographic food atlas perform better than food models and household utensils. One of the advantages of food photos is that they can be equivalent to various food types and portion sizes, whereas food models and cups represent limited numbers of food types and portion sizes. While this review did not compare the number of images, one study, using online 24hr recalls, showed improved accuracy using 8 images compared to 4 for each food item[21]. This finding suggests that future users should consider increasing the number of photos on display. Two studies[62,63] comparing the same number of images showed that there is no difference in accuracy between digital images and printed images. These results further support the use of computer-based PSEEs which have the added advantage of being able to increase the number of food photos relatively easily.

The assessment of three comparison studies[19,56,58] indicated that there is no difference in accuracy between the images vs actual forms of non-food objects and household utensils. Friedman et al[57] comparing three 3D PSEEs (household utensils vs household objects vs modelling clay) identified significant interactions between food and PSEE type (e.g. solid, liquid, amorphous) which may be due to different food types being estimated with different degrees of accuracy according to the type of PSEEs. This was raised as a potential issue for studies using a single PSEE for the estimation of different food types as it could add to estimation error[57], therefore a combination of 3D PSEEs may be more appropriate.

International food unit cube[59] and hand measures[60] (using the width of the fingers as a 'ruler' to measure the dimensions of foods) were two original ideas which performed better than other 3D tools such as household measures. Although their performance was not adequate for some food types (foods with less geometric shape e.g. chicken breast and amorphous food), they have the potential to be used as a reference object for estimating the amount of certain food types (foods with geometric shape e.g. cheese, cake) considering the practicality, especially hands being readily available[59,60].

The acceptability and usability of the PSEEs is another factor to be considered in nutrition studies, given the challenges in achieving effective participation rates[72]. In studies, comparing the efficacy and usability of PSEEs[59,71,73], participants' preference differed according to food type which highlights the fit-for-purpose approach in the selection of PSEEs.

Strengths and limitations of the review

To our knowledge, this work represents the most current, comprehensive review on portion size estimation methodologies applicable to dietary assessment. The parallel review led by one of the authors of this current study focused on portion size estimation instruments for minority ethnic groups[4]. Two other previous literature reviews[6,25] on the validity and effectiveness of portion size estimation methods were limited in terms of inclusion criteria (e.g. targeting tools used in validation studies only, mostly 2D or 3D aids rather than the complete spectrum of PSEEs), the small number of studies being tested (ranged from 5 to 27) and the coverage of short time periods (e.g. 1980-1994[6] and 2005-2016[25]). Whereas, this study categorized PSEEs according to a comprehensive variety of categories (e.g. dimensions, tool descriptions) in relation to various parameters (e.g. population characteristics, setting, dietary assessment methods and the format of usage) and analyzed 334 studies published across a large time period, 1910-2018. In addition we have developed a scoring tool for validation studies which could be useful for future studies. A review[25], evaluating the validity of PSEEs, included studies some of which were focused on the validation of dietary assessment tools instead of portion size estimation, and therefore did not use food weights as a comparator. In contrast the present study investigated validation studies which specifically validated PSEEs using actual food weights as the comparator. Therefore this study reports its findings specific to the validity of the tools.

The comprehensive data extraction process in this review enabled the identification of 16,801 records; however some relevant studies may still not have been captured as PSEEs are often not the focus of dietary assessment studies. A previous review[6] published in 1997 indicated a lack of evidence on the quality, validity, comparability and effectiveness of PSEEs, and 20 years later, the same limitations have been observed in this review. Therefore the outcomes of this present review are limited to the evidence available in the literature and it is possible that there is a tool which is not identified in the quality assessment of PSEEs as it lacks validation but nonetheless performs well.

The focus being the effectiveness of portion size estimation tools, this study has not looked into which foods were more inclined to be under-estimated or over-estimated by PSEEs. Future studies should further explore these issues using a systematic approach.

Considerations for the utilization of PSEEs

Misreporting of portion size is an important contributor to error associated with dietary intake measurement[6,74]. Such error may have serious implications for the analysis of health-outcome data in particular. Accurate dietary intake data is key to evaluate the impact of intervention studies measuring the efficiency of public health policies, nutritional or lifestyle approaches, on disease risk, as well as dietary interventions assessing portion control strategies[18]. However measurement error will always exist as it depends on an individuals' ability to perceive, conceptualize and remember the foods consumed[62]. Some individuals can make underestimates up to 40% or overestimates up to 60% and greater even when different PSEEs are being tested on the same foods and drinks[62,75]. As well as underestimation, over estimation is also a problem in dietary assessment[62]. Various factors affect peoples' ability to estimate portion size accurately such as age of the respondent, training provided to respondents, food type and size, and the visual perception and cognitive skills of individuals[4,10,62]. These are summarized below. Selection of appropriate PSEEs to be used in all these contexts requires careful consideration. This review identified four main areas to be considered when selecting a PSEE, presented in Table 4.

Individuals' age

The age of the respondent and the type of the dietary assessment instrument may impact on the estimation accuracy especially if there is a need to recall amounts from memory[5,30,76].

Children are more prone to portion size estimation errors than adults due to their shorter memory span[77]. Age-specific food atlases[27,78] or portion size lists[78] are options developed to enhance accuracy of portions size estimation in young children. Mobile technology can also offer particular advantages for portion size estimation by adolescents as there is less interference from adults and less urge to change normal habits[79,80]. As research and technology progresses, new strategies are expected to improve the recall of food amounts[21,29,81].

Individuals' visual perception

Research has shown that foods presented as unit foods, such as foods consisting of a single item (e.g. one sausage roll) or presented as one homogenous mixture (e.g. a meal of macaroni and cheese) tend to be underestimated[10,12,62,82]. Food size (e.g. small vs large) may also affect the portion size estimation such as the assumption of portion estimation getting more difficult for individuals as the portion size increases[31].

As the visual measurement tools have a cognitive influence on portion estimation, similarity in size and shape between the way a food is consumed and the PSEE could improve the accuracy of portion estimations[62].

Training

Training respondents may improve the accuracy of estimations[25,83,84]. For the dietary assessment of young children, training parents may also improve accuracy[83]. While training is an advantage to improve accuracy of estimations, when trying to assess the validity of a tool, it may be better to use a population without previous experience with PSEEs as this may improve the accuracy of portion size estimations[58]. Choosing PSEEs that are tested or validated in populations similar to the target population is also recommended[4,85]. Future research should explore if emerging technologies or other novel methods are suitable to decrease the influence of individual characteristics and skills on estimation accuracy.

Food characteristics

There is solid evidence that portion size estimation efficacy strongly depends on the food's characteristics including its physical form (i.e. amorphous vs. defined shape; liquid vs. solid)[18,27]. Portion size of some specific food types such as mixed dishes[4,15,26,71], foods in small unit size (e.g. under 20g)[23,84,86], pieces of sliced meat (e.g. cold cuts and fried beef)[15,23],foods served in sauce or gravy[49], light but voluminous foods [86] and non-staple foods[86] are particularly difficult to estimate. Whether the portion sizes of some foods are always under- or over-estimated is uncertain as some studies indicate low error rate for certain foods like milk[1,31] and others indicate no food-related differences[87]. Friedman et al[57] identified significant interactions between food and PSEE type as different food characteristics (e.g. solid, liquid) can be estimated with different degrees of accuracy depending on the type of PSEE used. Studies may consider using a combination of tools that can be applied to a range of food textures and can be flexible in estimating portion sizes of composite dishes[4,25]. For example, specifically for the portion size estimation of meats with more than one irregular dimension (e.g. ribs), large scale studies can consider using PSEEs such as the finger width method[60] or categorical size estimates (e.g. small, medium, large)[73] in addition to other PSEEs. Food photos are also highly efficient PSEEs as they can cover a wide range of food textures and shapes[25,54].

Unanswered questions and future research

PSEEs used in low-income settings

PSEEs have challenges in terms of the quality, particularly in low-income settings[4,88,89]. It is promising that a project called the International Dietary Data Expansion (INDDEX)[89] is aiming to facilitate data collection and processing in low-income countries by using a combination of direct weighing of actual foods, food replicas, household measures and food images alongside 24hr recalls collected by tablet technology[89,90].

PSEEs for shared plate eating practices

In some low and lower-middle income countries, shared plate eating is a common habit[11,91,92] which is a difficulty for portion size assessment. Although there are various approaches that have been explored[11,22,69,93,94], shared-plate issue continues to remain as a complexity in dietary assessment[4,92]. A narrative review[92] identified two key factors to be considered to improve the accuracy of assessment of dietary intake from shared plate eating, these were accurate assessment of the dietary intake of staple foods and the requirement to use combined approaches for portion size estimation.

PSEEs applicable to children

In this review, the most popular PSEEs applicable to children were food images and electronic PSEEs whereas household utensils were not popular. Friedman et al[57] reported that utensils used for portion size estimation in children performed worse than modelling clay and household objects. Therefore utensils may not be very applicable to children, however, as shown by Foster et al[95], using age-appropriate photos of children portion sizes can be more applicable to children.

PSEEs applicable to individuals with difficulty to estimate portion sizes

People with low literacy skills, chronic illnesses and living in institutionalized settings may have difficulty to estimate portion sizes[76,96]. Only a limited number of PSEEs were tested in these population groups. The addition of photo assistance[76] and the use of specifically developed mobile applications[97-99] were suggested as potential approaches to improve portion size estimation in these groups.

Technological PSEEs

Previous research[25,100,101] reported that the use of technological PSEEs especially those using digital food images hold good potential as they can reduce respondent burden, especially when combined with 24hr recalls[25]. This might be the reason why, in this review, 17% of electronic devices and digital images were combined with 24hr recalls. Some studies[41,42] also indicated that digital image assistance improves the overall dietary assessment as researchers reviewing images can correct the errors in food records. However, the interpretation error in estimating intake by the assessor is an area that may need to be addressed in future studies[42].

Overall, electronic images and devices constituted 20% of the PSEEs in the whole sample. However some issues around technologic dietary assessment methods were poor image quality, burdensome image review process, the lack of cost information, discomfort of users (especially for wearable cameras), change in eating behavior and low motivation of users. As the use of technology in dietary assessment advances it is expected that electronic PSEEs will be preferred over manual systems in future[25,100], however there is an ongoing need for technical improvements and more focus on portion size assessment in these areas to move this line of research forward[101].

CONCLUSION

Across an inventory of more than 500 PSEEs, photographic PSEEs such as food atlas were identified as the most widely applicable PSEEs to populations living in various settings and from different countries, plus versatile enough to cover various food shapes and textures. Based on the validation and comparison studies, photographic PSEEs were more accurate than food models and household utensils, and food models were more accurate than household utensils. There were no differences in accuracy between digital images and printed food images. Electronic PSEEs, especially digital images used in web-based 24hr recalls have been increasingly used for portion size estimation and further opportunities exist to improve and develop technology-assisted PSEEs.

Selection of appropriate PSEEs needs careful consideration of key elements which include sociodemographic factors such as age, sex, the level of education and geographic location. Also precautions to reduce measurement error need to be taken into account such as; using age appropriate photos for children, using PSEEs (maybe in combination) applicable to various food textures, composite dishes, customary portions and servings; using PSEEs that have high user preference, training respondents and providing additional assistance to people having difficulty in portion size estimation.

There is a lack of validation of PSEEs and the field would benefit from increased evaluation of tools, perhaps as part of surveys and studies, to add to the available literature about PSEEs and their efficacy in different population settings. Validation studies testing PSEEs should include sufficient sample size, perform a validation against weighed amounts in preference to a comparison study against relative amounts (estimated), measure reliability and agreement using appropriate statistics and employ a sufficiently wide range of foods. When selecting a PSEE it is advisable to choose one that has been properly validated or compared (the criteria included in the scoring tool developed in this study could be used as a guideline) and incorporating ongoing evaluation and validation as PSEEs are adopted and further evolved.

Acknowledgements

The authors are grateful to Eleni Vantana for conducting the preliminary scoring of validation studies. *Author contributions*. B.A. coordinated the study, checked the data extracts, designed the quality scoring tool, conducted the data analysis and wrote the article in collaboration with E.A.R. E.A.R. designed the study and advised on all the steps of the project. E.A.R., T.H., C.G., R.B. and D.Y. performed the data searches and data extraction. T.H. and R.B. scored the validation studies and prepared the first draft of some tables. J.C. and M.R. advised on the development of quality scoring tool and critically reviewed the draft paper and content. J.C. advised on the development of quality scoring tool. C.R., P.P. and T.S. critically reviewed the paper and advised on the content. All authors contributed to the editing of the manuscript and share responsibility for the final content. All authors read and approved the final manuscript.

Funding/support. B.A., EAR, P.P. and C.R. were supported by the Medical Research Council programme number U105960384. EAR was also supported by the Centre for Nutrition Research, University of Navarra.

Declaration of interest. The authors have no relevant interests to declare.

Supporting information

Appendix 1 **PRISMA 2009 Check List**

Appendix 2 **Search Pathway**

Appendix 3 **Quality Scoring Form**

Table S1 **Whole sample of portion size estimation elements (PSEEs) identified across 334 records**

Table S2 **The results of the quality assessment of 21 studies validating portion size estimation elements (PSEEs)**

Table S3 **Characteristics and main findings of 13 studies comparing multiple portion size estimation elements (PSEEs) to each other**

REFERENCES

1. Guthrie HA. Selection and quantification of typical food portions by young adults. *J Am Diet Assoc.* 1984;84:1440-1444.
2. Blake AJ, Guthrie HA, Smiciklas-Wright H. Accuracy of food portion estimation by overweight and normal-weight subjects. *J Am Diet Assoc.* 1989;89:962-964.
3. Riboli E, Elmstahl S, Saracci R, et al. The Malmo Food Study: validity of two dietary assessment methods for measuring nutrient intake. *Int J Epidemiol.* 1997;26 Suppl 1:S161-173.
4. Almiron-Roig E, Aitken A, Galloway C, et al. Dietary assessment in minority ethnic groups: a systematic review of instruments for portion-size estimation in the United Kingdom. *Nutr Rev.* 2017;75:188-213.
5. Subar AF, Freedman LS, Tooze JA, et al. Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. *J Nutr.* 2015;145:2639-2645.
6. Cypel YS, Guenther PM, Petot GJ. Validity of portion-size measurement aids: a review. *Journal of the Amercan Dietetic Association.* 1997;97:289-292.
7. Baranowski T, Baranowski JC, Watson KB, et al. Children's accuracy of portion size estimation using digital food images: effects of interface design and size of image on computer screen. *Public Health Nutrition.* 2011;14:418-425.
8. Nelson M, Atkinson M, Darbyshire S. Food photography I: the perception of food portion size from photographs. *British Journal of Nutrition.* 1994;72:649-663.
9. Lanerolle P, Thoradeniya T, de Silva A. Food models for portion size estimation of Asian foods. *Journal of Human Nutrition and Dietetics.* 2013;26:380-386.
10. Almiron-Roig E, Solis-Trapala I, Dodd J, et al. Estimating food portions. Influence of unit number, meal type and energy density. *Appetite.* 2013;71:95-103.
11. Ngo J, Gurinovic M, Frost-Andersen L, et al. How dietary intake methodology is adapted for use in European immigrant population groups - a review. *British Journal of Nutrition.* 2009;101:86-94.

12. Brogden N, Almiron-Roig E. Estimated portion sizes of snacks and beverages differ from reference amounts and are affected by appetite status in non-obese men. *Public Health Nutr.* 2011;14:1743-1751.
13. Gretarsdottir S, Baas AF, Thorleifsson G, et al. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. *Nat Genet.* 2010;42:692-697.
14. Haraldsdottir J, Tjonneland A, Overvad K. Validity of individual portion size estimates in a food frequency questionnaire. *Int J Epidemiol.* 1994;23:786-796.
15. Ovaskainen ML, Paturi M, Reinivuo H, et al. Accuracy in the estimation of food servings against the portions in food photographs. *Eur J Clin Nutr.* 2008;62:674-681.
16. Eddington J TM, Geekie M, Ball M & Mann J. Assessment of nutritional intake using dietary record with estimated weight. *Journal of Human Nutrition and Dietetics.* 1989;61:407-414.
17. Chambers E, Godwin SL, Vecchio FA. Cognitive strategies for reporting portion sizes using dietary recall procedures. *Journal of the Amercan Dietetic Association.* 2000;100:891-897.
18. Almiron-Roig E, Navas-Carretero S, Emery P, et al. Research into food portion size: methodological aspects and applications. *Food Funct.* 2018;9:715-739.
19. Godwin SL, Chambers Et, Cleveland L. Accuracy of reporting dietary intake using various portion-size aids in-person and via telephone. *J Am Diet Assoc.* 2004;104:585-594.
20. Robson PJ, Livingstone MB. An evaluation of food photographs as a tool for quantifying food and nutrient intakes. *Public Health Nutr.* 2000;3:183-192.
21. Subar AF, Crafts J, Zimmerman TP, et al. Assessment of the accuracy of portion size reports using computer-based food photographs aids in the development of an automated self-administered 24-hour recall. *J Am Diet Assoc.* 2010;110:55-64.
22. Huybregts L, Roberfroid D, Lachat C, et al. Validity of photographs for food portion estimation in a rural West African setting. *Public Health Nutr.* 2008;11:581-587.
23. Venter CS, MacIntyre UE, Vorster HH. The development and testing of a food portion photograph book for use in an African population. *Journal of Human Nutrition and Dietetics.* 2000;13:205-218.
24. Yuhas JA, Bolland JE, Bolland TW. The impact of training, food type, gender and container size on the estimation of food portion sizes *Journal of the Amercan Dietetic Association.* 1989;89:1473-1477.
25. Sharma V, Chadha R. Effectiveness of food portion size estimation aids for diet assessment: A systematic review. *International Journal of Food Science and Nutrition.* 2017;2:106-112.
26. Anderson AS, Barton K, Craigie A, et al. *Exploration of Adult Food Portion Size Tools.* Edinburgh: NHS Scotland;2008.
27. Foster E, Adamson AJ, Anderson AS, et al. Estimation of portion size in children's dietary assessment: lessons learnt. *Eur J Clin Nutr.* 2009;63 Suppl 1:S45-49.
28. Livingstone MB, Black AE. Markers of the validity of reported energy intake. *J Nutr.* 2003;133 Suppl 3:895S-920S.
29. Rangan A, Allman-Farinelli M, Donohoe E, et al. Misreporting of energy intake in the 2007 Australian Children's Survey: differences in the reporting of food types between plausible, under- and over-reporters of energy intake. *J Hum Nutr Diet.* 2013.
30. Murakami K, Livingstone MB. Prevalence and characteristics of misreporting of energy intake in US children and adolescents: National Health and Nutrition Examination Survey (NHANES) 2003-2012. *British Journal of Nutrition.* 2015:1-11.
31. Young LR, Nestle MS. Portion sizes in dietary assessment: issues and policy implications. *Nutr Rev.* 1995;53:149-158.
32. Academy of Nutrition and Dietetics. Evidence Analysis Manual: Evidence Analysis Process. In: Development RaSB, ed. Chicago: Academy of Nutrition and Dietetics; 2012.
33. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011.

34. Moher D, Liberati A, Tetzlaff J, et al. Reprint--preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Physical therapy.* 2009;89:873-880.

35. Lewis HB. Food portion size and implications for appetite control and obesity: MRC Human Nutrition Research, University of Cambridge; 2013.

36. EPOC Resources for review authors. Effective Practice and Organisation of Care (EPOC). Suggested risk of bias criteria for EPOC reviews. . 2015 http://epoc.cochrane.org/epoc-specific-resources-review-authors. Accessed 4 Aug 2015.

37. Critical Appraisal Skills Programme (CASP). 2013; http://www.casp-uk.net/.

38. Bryant M, Ashton L, Brown J, et al. Systematic review to identify and appraise outcome measures used to evaluate childhood obesity treatment interventions (CoOR): evidence of purpose, application, validity, reliability and sensitivity. *Health Technol Assessment.* 2014;18:129-139.

39. Nutrition Epidemiology Group. Making the best use of new technologies in the National Diet and Nutrition Survey: a review. University of Leeds and NatCen Social Research; 2014.

40. Serra-Majem L, Frost Andersen L, Henrique-Sanchez P, et al. Evaluating the quality of dietary intake validation studies. *British Journal of Nutrition.* 2009;102 Suppl 1:S3-9.

41. Ptomey LT, Willis EA, Goetz JR, et al. Digital photography improves estimates of dietary intake in adolescents with intellectual and developmental disabilities. *Disabil Health J.* 2015;8:146-150.

42. Pettitt C, Liu J, Kwasnicki RM, et al. A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate. *Br J Nutr.* 2016;115:160-167.

43. Murkin E, Cole DC, Kearney JP, et al. Fish consumption practices among frequent consuming fishers of five Ontario Great Lakes areas of concern (AOCs). *Journal of Great Lakes Research.* 2003;29:436-447.

44. Amougou N, Cohen E, Mbala L, et al. Development and validation of two food portion photograph books to assess dietary intake among adults and children in Central Africa. *British Journal of Nutrition.* 2016;115:895-903.

45. Boateng GP. The Development of A Photographic Food Atlas with Portion Sizes of Commonly Consumed Carbohydrate Foods in Accra, Ghana. Accra, Ghana: The School of Allied Health Sciences, The University of Ghana, Legon; 2014.

46. Harris-Fry H, Paudel P, Karn M, et al. Development and validation of a photographic food atlas for portion size assessment in the southern plains of Nepal. *Public Health Nutr.* 2016;19:2495-2507.

47. Biltoft-Jensen A, Holmgaard Nielsen T, Hess Ygil K, et al. Accuracy of food photographs for quantifying food servings in a lunch meal setting among Danish children and adults. *J Hum Nutr Diet.* 2018;31:131-140.

48. Bernal-Orozco MF, Vizmanos-Lamotte B, Rodriguez-Rocha NP, et al. Validation of a Mexican food photograph album as a tool to visually estimate food amounts in adolescents. *Br J Nutr.* 2013;109:944-952.

49. Tueni M, Mounayar A, Birlouez-Aragon I. Development and evaluation of a photographic atlas as a tool for dietary assessment studies in Middle East cultures. *Public Health Nutr.* 2012;15:1023-1028.

50. Bouchoucha M, Akrout M, Bellali H, et al. Development and validation of a food photography manual, as a tool for estimation of food portion size in epidemiological dietary surveys in Tunisia. *Libyan J Med.* 2016;11:32676.

51. Badari SAZ, Arcot J, Sulaiman N. An exploratory study of visual aids life-sized photographs of serve/portion sizes of foods and their effectiveness in recording dietary intakes. *MAJCAFE.* 2015;18:50-71.

52. Szenczi-Cseh J, Horvath Z, Ambrus A. Validation of a food quantification picture book and portion sizes estimation applying perception and memory methods. *Int J Food Sci Nutr.* 2017;68:960-972.

53. Jia W, Chen H-C, Yue Y, et al. Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera. *Public health nutrition.* 2014;17:1671.

54. Lazarte CE, Encinas ME, Alegre C, et al. Validation of digital photographs, as a tool in 24-h recall, for the improvement of dietary assessment among rural populations in developing countries. *Nutr J.* 2012;11:61.

55. Carvalho MA, Baranowski T, Foster E, et al. Validation of the Portuguese self-administered computerised 24-hour dietary recall among second-, third- and fourth-grade children. *J Hum Nutr Diet.* 2015;28:666-674.

56. Matheson DM, Hanson KA, McDonald TE, et al. Validity of children's food portion estimates: a comparison of 2 measurement aids. *Arch Pediatr Adolesc Med.* 2002;156:867-871.

57. Friedman A, Bennett TG, Barbarich BN, et al. Food portion estimation by children with obesity: the effects of estimation method and food type. *J Acad Nutr Diet.* 2012;112:302-307.

58. Byrd-Bredbenner C, Schwartz J. The effect of practical portion size measurement aids on the accuracy of portion size estimates made by young adults. *J Hum Nutr Diet.* 2004;17:351-357.

59. Bucher T, Weltert M, Rollo ME, et al. The international food unit: a new measurement aid that can improve portion size estimation. *Int J Behav Nutr Phys Act.* 2017;14:124.

60. Gibson A, Hsu SH, Rangan M, et al. Accuracy of hands v. household measures as portion size estimation aids. *Journal of nutritional science.* 2016;5.

61. Posner BM, Smigelski C, Duggal A, et al. Validation of two-dimensional models for estimation of portion size in nutrition research. *J Am Diet Assoc.* 1992;92:738-741.

62. Hernandez T, Wilder L, Kuehn D, et al. Portion size estimation and expectation of accuracy. *Journal of Food Composition and Analysis.* 2006;19:S14-S21.

63. Liu YC, Chen CH, Lee CW, et al. Design and usability evaluation of user-centered and visual-based aids for dietary food measurement on mobile devices in a randomized controlled trial. *J Biomed Inform.* 2016;64:122-130.

64. Subasinghe AK, Thrift AG, Evans RG, et al. Novel dietary intake assessment in populations with poor literacy. *Asia Pac J Clin Nutr.* 2016;25:202-212.

65. Turconi G, Guarcello M, Berzolari FG, et al. An evaluation of a colour food photography atlas as a tool for quantifying food portion size in epidemiological dietary surveys. *Eur J Clin Nutr.* 2005;59:923-931.

66. Lillegaard IT, Overby NC, Andersen LF. Can children and adolescents use photographs of food to estimate portion sizes? *Eur J Clin Nutr.* 2005;59:611-617.

67. Centers for Disease Control and Prevention (CDC). Measuring guides for the dietary recall interview. *National Health and Nutrition Examination Survey* 2010; http://www.cdc.gov/nchs/nhanes/measuring_guides_dri/measuringguides.htm. Accessed November 23, 2018.

68. Roberts C, Steer T, Maplethorpe N, et al. *National Diet and Nutrition Survey, Results from Years 7 and 8 (combined) of the Rolling Programme (2014/2015 to 2015/2016).* London, England2018.

69. Aglago EK, Landais E, Nicolas G, et al. Evaluation of the international standardized 24-h dietary recall methodology (GloboDiet) for potential application in research and surveillance within African settings. *Global Health.* 2017;13:35.

70. Wrieden WL, Momen NC. Workshop 3: Novel approaches for estimating portion sizes. *Eur J Clin Nutr.* 2009;63 Suppl 1:S80-81.

71. Faulkner GP, Livingstone MB, Pourshahidi LK, et al. An evaluation of portion size estimation aids: precision, ease of use and likelihood of future use. *Public Health Nutrition.* 2016;18:2874-2880.

72. Keeble C, Baxter PD, Barber S, et al. Participation Rates In Epidemiology Studies And Surveys: A Review 2007–2015. *The Internet Journal of Epidemiology.* 2017;14:1-14.

73. Godwin S, McGuire B, Chambers IV E, et al. Evaluation of portion size estimation aids used for meat in dietary surveys. *Nutrition Research.* 2001;21:1217-1233.

74. Jonnalagadda SSM, D.C.; Smiciklas-Wright, H.; Kris-Etherton, P.M.: Karmally, W.; VanHeel, N. Portion Size Estimation: A Source of Error in Diet Assessment Studies. *Journal of the Academy of Nutrition & Dietetics.* 1995;95:A21.

75. Lucas F, Niravong M, Villeminot S, et al. Estimation of food portion size using photographs: validity, strengths, weaknesses and recommendations. *Journal of Human Nutrition and Dietetics.* 1995;8:65-74.

76. Ptomey LT, Herrmann SD, Lee J, et al. Photo-assisted recall increases estimates of energy and macronutrient intake in adults with intellectual and developmental disabilities. *J Acad Nutr Diet.* 2013;113:1704-1709.

77. Foster E, Adamson A. Challenges involved in measuring intake in early life: focus on methods. *Proc Nutr Soc.* 2014;73:201-209.

78. Wrieden WL, Longbottom PJ, Adamson AJ, et al. Estimation of typical food portion sizes for children of different ages in Great Britain. *Br J Nutr.* 2008;99:1344-1353.

79. Simpson E, Bradley J, Poliakov I, et al. Iterative Development of an Online Dietary Recall Tool: INTAKE24. *Nutrients.* 2017;9.

80. Boushey CJ, Kerr DA, Wright J, et al. Use of technology in children's dietary assessment. *Eur J Clin Nutr.* 2009;63 Suppl 1:S50-57.

81. Rennie KL, Coward A, Jebb SA. Estimating under-reporting of energy intake in dietary surveys using an individualised method. *Br J Nutr.* 2007;97:1169-1176.

82. Rolls BJ. The Supersizing of America: Portion Size and the Obesity Epidemic. *Nutr Today.* 2003;38:42-53.

83. Small L, Lane H, Vaughan L, et al. A systematic review of the evidence: the effects of portion size manipulation with children and portion education/training interventions on dietary intake with adults. *Worldviews Evid Based Nurs.* 2013;10:69-81.

84. Slawson DL, Eck LH. Intense practice enhances accuracy of portion size estimation of amorphous foods. *J Am Diet Assoc.* 1997;97:295-297.

85. Cade JE, Burley VJ, Warm DL, et al. Food-frequency questionnaires: a review of their design, validation and utilisation. *Nutrition Research Reviews.* 2004;17:5-22.

86. Gittelsohn J, Shankar AV, Pokhrel RP, et al. Accuracy of estimating food intake by observation. *J Am Diet Assoc.* 1994;94:1273-1277.

87. Thompson CH, Head MK, Rodman SM. Factors influencing accuracy in estimating plate waste. *J Am Diet Assoc.* 1987;87:1219-1220.

88. Walls HL, Johnston D, Mazalale J, et al. Why we are still failing to measure the nutrition transition. *BMJ Glob Health.* 2018;3:e000657.

89. Coates JC, Colaiezzi BA, Bell W, et al. Overcoming Dietary Assessment Challenges in Low-Income Countries: Technological Solutions Proposed by the International Dietary Data Expansion (INDDEX) Project. *Nutrients.* 2017;9.

90. Tufts University. International Dietary Data Expansion Project. https://inddex.nutrition.tufts.edu/. Accessed November 23, 2018.

91. Landais E, Bour A, Gartner A, et al. Socio-economic and behavioural determinants of fruit and vegetable intake in Moroccan women. *Public Health Nutr.* 2015;18:809-816.

92. Burrows T, Collins C, Adam M, et al. Dietary Assessment of Shared Plate Eating: A Missing Link. *Nutrients.* 2019;11.

93. Kigutha HN. Assessment of dietary intake in rural communities in Africa: experiences in Kenya. *Am J Clin Nutr.* 1997;65:1168S-1172S.

94. Hudson GJ. Food intake in a west African village. Estimation of food intake from a shared bowl. *Br J Nutr.* 1995;73:551-569.

95. Foster E, Matthews JN, Nelson M, et al. Accuracy of estimates of food portion size using food photographs--the importance of using age-appropriate tools. *Public Health Nutr.* 2006;9:509-514.

96. Chaudhry BM, Schaefbauer C, Jelen B, et al. Evaluation of a food portion size estimation interface for a varying literacy population. Paper presented at: The 2016 CHI Conference on Human Factors in Computing Systems2016; San Jose, USA
97. Paiva I, Amaral T, Barros H. Influence of individually estimated portion size on the assessment of nutritional risk in colorectal cancer in Portugal. *J Hum Nutr Diet.* 2004;17:529-536.
98. Chaudry BM, Connelly K, Siek KA, et al. Formative evaluation of a mobile liquid portion size estimation interface for people with varying literacy skills. *J Ambient Intell Humaniz Comput.* 2013;4:779-789.
99. Donald H, Franklin V, Greene S. The use of mobile phones in dietary assessment in young people with Type 1 diabetes. *Journal of Human Nutrition and Dietetics.* 2009;22:256-257.
100. Amoutzopoulos B, Steer T, Roberts C, et al. Traditional methods v. new technologies - dilemmas for dietary assessment in large-scale nutrition surveys and studies: a report following an international panel discussion at the 9th International Conference on Diet and Activity Methods (ICDAM9), Brisbane, 3 September 2015. *J Nutr Sci.* 2018;7:e11.
101. Boushey CJ, Spoden M, Zhu FM, et al. New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. *Proc Nutr Soc.* 2017;76:283-294.

Table 1 **Description of research question components by population, interventions, comparisons, outcomes, and study designs (PICOS): Systematic review on portion size estimation elements (PSEEs)**

Criteria	Description
Population	Human subjects Subgroups (e.g. age, gender, ethnicity, health condition, setting)
Intervention	Whenever a PSEE is described to allow quantification of dietary intake
Comparison	Any comparisons, especially between PSEEs
Outcome	Population/individual dietary intake; method development; method validation or comparison; any other health or diet-related outcome where a PSEE is described
Study design	Any study design describing a PSEE; review papers with relevant references; websites of health professional/non-government organizations; academic and industry reports. Excluded outcomes: editorial, commentary and opinion pieces; review papers with no relevant references

Table 2. **Scoring tool developed in this study for the analysis of quality and relative efficacy parameters including study design, validity, reliability and agreement of dietary instruments including portion size estimation elements (PSEEs)**

Criteria	Criterion is met (min-max)
Section 1. Study design	
1. Sample size is adequate	1-4

2. Population is adequately described and representative	1-4
Section 2. Validity	
3. Comparator is appropriate (tool tested for validity)	1-3
4. Tool is versatile	1-2
5. Tool was piloted	1-2
Section 3. Reliability	
6. Reliability measured using appropriate techniques	1-2
Section 4. Agreement	
7. Agreement measured using appropriate techniques	1-4
Section 5: Future application	
8. There is potential for long-term efficacy	1-4
Score calculation	Add up all points (Maximum score 25)

Table 3. **Categories of the 542 portion size estimation elements (PSEEs) identified**

Tool dimension	Tool description	n
1D (n=30)	Portion lists	11
	Food guide	10
	Label/food packaging	6
	Voice recording	2
	Till receipt	1
2D (n=249)	Food photos (e.g. photographic atlas)	101
	Electronic image-based method	61
	Electronic device, computer-based method	49
	Food diagrams/ drawings[a]	14
	Non-food object image[b]	15
	Utensil image	5
	Hand image	4
3D (n=263)	Household utensils including measuring utensils	107
	Food scale	68
	Food replica and food model[c]	79
	Ruler	7
	Hands	2
Total		542

Abbreviations: 1D: One- dimensional, 2D: Two- dimensional, 3D: Three- dimensional. [a]"Food diagrams/ drawings" includes tools such as drawings of bread, images of rectangles with thickness grid and diagrams of rectangles. [b]"Non-food object image" includes the images of food models and replicas such as tennis ball and food mound. [c]"Food replica and food model" includes items such as golf ball, deck of cards, plastic meat pieces, thickness sticks, modelling clay bean bag and wedge.

Table 4 **Areas to consider when selecting a portion size estimation tool (PSEE)**

Area	Things to consider

Validation	• PSEE has been validated using appropriate methods • Sample size of the validation study was large enough • PSEE has been piloted in the population of interest • Reliability of PSEE has been tested • Comparator was weighed data (not measured by respondents) • Agreement has been tested
Efficiency	• High user preference • High feasibility/low complexity • No burden and limitation to implement • Cost effective • Easy to use
Specificity	• PSEE is applicable to different food textures and shapes including mixed dishes, amorphous foods and irregularly shaped foods, if not PSEE may need to be combined with other complementary PSEEs. • PSEE is culturally appropriate (e.g. use of customary utensils) • PSEE is applicable to traditional eating habits (e.g. the ways of serving, sharing dishes, eating by hand, customary portions) • PSEE is age appropriate (e.g. children specific portion sizes)
Implication	• If a PSEE will be applied to a different country/region it need to be adapted to this setting (e.g. removing unrelated cutlery photos) • Consider to provide training and use photo or interview assistance to participants especially those having difficulty with portion size estimation

Figure 1 **Literature search process.** The following databases were searched for publications reporting the use of a portion size estimation element (PSEE) applicable to dietary assessment methods (based on the PRISMA statement[34]): University of York Centre for Reviews and Dissemination (three databases with coverage of health care and services, health economics and health technology), Cochrane Library, EBSCO, NHS Evidence, Ovid, Oxford journals, Scopus, SocINDEX, Sociological Abstracts, Econlit, Web of Knowledge, Wiley Online Library, Google, Google scholar, EthOS, University of Birmingham e-Theses, University of Chesters' online research.

Figure 2 **Distribution of the 542 portion size estimation elements (PSEEs) identified in this review. (A)** Distribution by purpose of the study. "Data collection" includes studies collecting portions size data such as portion weights of food served in schools. "Interventions" includes experiments, for example examining whether increasing the portion size of a meal is affecting energy intake. "Population studies" includes studies such as cross-sectional food consumption studies. "Training" includes educational material such as measuring guides. "Development" includes studies developing a tool for dietary assessment or portion estimation. "Evaluation" includes validation, comparison or usability testing of tools for dietary assessment of portion estimation. **(B)** Distribution by the format of the usage of PSEE. “Stand alone” includes tools consisting of only one PSEE. “Related set” includes tools consisting of more than one PSEE applied within the same dietary assessment method measuring the same dimension (e.g. two-dimensional). “Combined set’ includes tools consisting of more than one PSEE measuring various dimensions (e.g. measuring cups and images) and applied to the same dietary assessment instrument. **(C)** Distribution by type of dietary assessment instrument into which the PSEE was integrated. “Food record” includes both weighed and estimated records such as diet diaries. “Not available” refers to no specific instrument (e.g. in comparison studies using only PSEE without being part of a particular dietary assessment instrument). “Multiple methods” refers to the combination of more than one dietary assessment instrument. “Questionnaire” refers to questionnaires other than FFQs. “Dietary guide” includes dietary guides and only food pyramid was identified in this category. *Abbreviations*: Commercial p., “Commercial products” which includes portion control tools such as pasta portioner. 24hr R, 24 hour recall; FFQ, food frequency questionnaires.

Figure 3 **Portion-size estimation elements (PSEEs) by population origin across the 334 publications analyzed in this review. (A)** Population distribution across all studies. “Native or mixture” refers to people living in a certain region including both the immigrant and native populations. “Ethnic group” refers to a specific ethnic group such as ethnic minorities and immigrants living in a region, for example African Americans living in USA and South Asians living in UK. “Eastern” includes Israeli and Lebanese. “Asian” includes Chinese, Japanese and Taiwanese. “American” refers to Americans living in the USA. “American (ethnic)” refers to Americans with other ethnic origins such as African Americans and American Indians. **(B)** Age distribution across all studies. “Children” includes all ages up to 18 years old including pre-school ages and adolescence. “Mixed” includes combination of more than one age group such as adult and children. “Elderly” includes people over 60 years old, although the use of definition varies across studies analyzed in this review. **(C)** Health status distribution across all studies. “Other” includes other health issues than obesity and chronic diseases such as undernutrition, eating disorders and disabilities. “Chronic diseases” *Abbreviations*: N/A, not available; Mixed, more than one ethnic group such as African American and Chinese population living in USA.