A review of the design and validation of web- and computer-based 24-h dietary recall tools

Claire M. Timon¹, Rinske van den Barg¹, Richard J. Blain¹, Laura Kehoe², Katie Evans², Janette Walton², Albert Flynn² and Eileen R. Gibney¹*

¹Institute of Food and Health, University College Dublin, Belfield, Dublin 4, Republic of Ireland
²School of Food and Nutritional Sciences, University College Cork, Cork, Republic of Ireland

Abstract
Technology-based dietary assessment offers solutions to many of the limitations of traditional dietary assessment methodologies including cost, participation rates and the accuracy of data collected. The 24-h dietary recall (24HDR) method is currently the most utilised method for the collection of dietary intake data at a national level. Recently there have been many developments using web-based platforms to collect food intake data using the principles of the 24HDR method. This review identifies web- and computer-based 24HDR tools that have been developed for both children and adult population groups, and examines common design features and the methods used to investigate the performance and validity of these tools. Overall, there is generally good to strong agreement between web-based 24HDR and respective reference measures for intakes of macro- and micronutrients.

Key words: Web-based recalls: Dietary assessment: Computer-based recalls: 24-h Dietary recall tools: Validation

Introduction
Accurate measurements of dietary intake are fundamental to health and nutrition research; however, diet is inherently complicated to measure as it changes over time and varies across life stages amongst other factors(1). In recent years there have been many advances in the application of technology in the area of dietary assessment(2). Innovative dietary assessment technologies using devices such as mobile telephones(3–8), sensors, wearable cameras(9–11) and web-based platforms are common ways of collecting dietary intake data and have become increasingly popular alternatives to traditional pen-and-paper versions of dietary assessment(12). Web-based methodologies facilitate the collection of dietary intake data across many geographic locations(13–16) and in some cases are preferred by participants compared with the traditional pen-and-paper alternatives(9,11). Overall the application of technology in dietary intake assessment has been shown to reduce issues associated with traditional collection of dietary data, such as cost, participation rates (by reducing the burden associated with dietary assessment) and accuracy of data collected(18). The success of these methods has been attributed to a number of factors including the ability to collect data in a remote and neutral environment, the standardised sequence of questioning, the use of digital portion size assessment aids, automated analysis of data collected and the provision of dietary feedback(12).

The reduced participant burden and time of data collection, associated with web- and computer-based 24-h dietary recall (24HDR) tools, support the use of this method in many different populations. Web- and computer-based 24HDR tools have been used to collect data from various population groups including young children(19–22), adolescents(5,6,23) and adults(24–27). Two of the comparison and validation studies included in this review recruited elderly participants in their study population(28,29), however, the validity of these tools have not been tested explicitly in an elderly population. In general, the applications are based on the multi-pass approach described by Moshfegh et al.(30). The applications are either self-administered, whereby a participant completes the recall in the absence of a researcher(5,31–33), or interviewer administered, where the interviewer uses the application to collect/analyse dietary recall data in the presence of a participant(26,28,34,35). The software guides and prompts the participant/interviewer to recall and record food and drink consumed in the previous day (24 h). The design features, food and beverage lists, nutritional composition data, prompts and portion size assessment aids incorporated in these applications differ from model to model and are generally for target populations.

The ability of many of these tools to accurately assess dietary intakes has been investigated using different study designs, the most common of which are comparison and validation studies. A comparison study investigates the performance of a ‘test measure’ of dietary assessment against an alternative dietary

Abbreviations: 24HDR, 24-h dietary recall; ASA24, National Cancer Institute’s Automated Self-Administered 24-Hour Dietary Assessment Tool; DLW, doubly labelled water.

* Corresponding author: Eileen R. Gibney, fax +353 1716 7203, email eileen.gibney@ucd.ie
recall (such as interviewer-led 24-h recall) to ascertain if the data collected by the test measure are comparable with those using the existing method. A validation study investigates the accuracy of a test measure compared with an objective measure of intake, such as biological markers or direct observation of intake. Assessing the validity of dietary assessment tools is difficult due to the risk of correlated error between the test method and a reference method\(^{36}\). The cost and practicality associated with the direct observation of intake or the collection of biological samples for the analysis of biomarkers of intake can also be limiting factors. In the absence of biomarkers or direct observation, dietary records (preferably weighed food records) are often considered the ‘gold standard’ for the measurement of food and nutrient intake and are typically used as a reference comparison method\(^{37}\).

De Keyzer et al\(^{38}\) notes that the 24 HDR method is currently the most commonly used tool to collect dietary intake data in national surveys, usually taking multiple recalls to assess habitual intake and, in some instances, used in conjunction with an FFQ (for example, Canada and USA\(^{39,40}\)). In Europe, efforts are being made to standardise the collection of dietary intake data and the potential of computer-assisted 24 HDR tools has been recognised\(^{20,41}\). Both self-administered and interviewer-administered applications have demonstrated the ability to reach large population groups using these applications\(^{42,43}\).

At present there are numerous 24 HDR tools that have been developed, tested and validated worldwide. The aim of this review was to examine common design features amongst current 24 HDR tools and investigate the methods used to assess the comparability and validity of these tools.

Methods

A literature search was conducted using the online databases PubMed, EMBASE and Web of Science to collect information on all 24 HDR tools. The inclusion criteria were as follows: English-language publications from 2000 until 2016 reporting the development, comparison or validation of a web- or computer-based 24-h recall tool for any population group. The following search terms were used alone and in combination: ‘24h recall’, ‘24-HR’, ‘nutrition’, ‘food’, ‘intake’, ‘dietary assessment’, ‘validity’, ‘validation’, ‘comparison’ and ‘reliability’. In a similar approach described by Gemming et al\(^{44}\), abstracts were initially screened by one author (C. M. T.) for relevance and then collected and distributed to three other authors. The literature search yielded a total of forty-four relevant papers. All papers were reviewed by three investigators (C. M. T., R. v. d. B. and R. J. B.). Sixteen papers were omitted for the following reasons: the papers did not describe a 24 HDR tool (\(n = 4\)) or the papers described aspects of the study, other than the development, comparison or validation of web-based dietary recall tools (\(n = 12\)). A total of twenty-eight papers reporting twenty-one individual web- or computer-based 24 HDR tools were included in the present review.

The design characteristics of the twenty-one 24 HDR tools are described in Tables 1 and 2 and the findings of those tools whereby validity/comparability was investigated are described in Tables 3 and 4. To obtain an overview of the different tools, general characteristics were examined, such as participant age, number of food and drink items incorporated in the application, the use of prompt techniques to ensure the complete capture of data and the use of portion size assessment aids. For the assessment of techniques used to investigate validity and comparability, details such as number of recalls used, reference method, type of statistical analysis and markers of nutrient intake were compared to identify common features. Where the required information was not available in the included publications, the corresponding author was contacted via email to obtain/verify information when possible.

Results

Design characteristics

Of the twenty-one 24 HDR tools identified from the literature search, six tools were developed for use specifically with children (three developed in Europe, two in the USA and one in Canada) and thirteen tools were developed (seven in Europe, three in USA, two in Korea and one in India) specifically for adult populations and two tools (both in the UK) were developed for use both with children and adults\(^{35,51}\). All of the tools developed for use with children were self-administered, whereas four of the tools developed for adults were interviewer administered. The youngest age to test a 24 HDR tool was 7 years of age\(^{42-45}\) and the oldest age was 80+ years\(^{26,29}\) (Tables 1 and 2).

Only three 24 HDR tools reported collecting intake data at a food-group level\(^{27,29,46}\), including additional choices/questions within each food group to obtain more specific information about intakes. All other tools recorded intake by presenting lists of food and beverage items. The number of food and drink items included varied for each tool, ranging from forty\(^{45}\) to 45 000 food and beverage options\(^{51}\). The use of portion size images was the most popular aid to facilitate portion size estimation. The number of portion size images varied from tool to tool, with some investigators reporting as many as 17 000 images\(^{42}\).

The use of food prompts was prevalent across 24 HDR tools developed for both children and adults. ‘Frequently forgotten foods’ was the most common prompt function reported in 24 HDR tools developed for children. This prompt involves the presentation of a list of foods that are known to be frequently omitted from dietary recalls (for example, sugar in hot beverages, etc.)\(^{51,47}\). The Synchronised Nutrition and Activity Program (SNAP\(^{TM}\)) tool\(^{21}\) used images of foods to prompt for frequently forgotten food items rather than a list of foods. In the tools developed for adults, ‘frequently forgotten foods’ was also a common prompt, as was ‘linked foods’ (Table 2). Linked food prompts offer the participant food or drink options that are often eaten in combination with the primary food or drink item selected\(^{5,24}\). Automated data entry checks, such as those implemented by Vereecken et al\(^{48}\), for portion size information and meal gap reviews whereby any period greater than 3 h between meals was queried\(^{42}\) were also useful prompt functions incorporated into adult 24 HDR tools. Of the 24 HDR tools developed for children, two recorded supplement intake\(^{19}\).
Table 1. Design characteristics of the web- and computer-based 24-h dietary recall tools developed for children

<table>
<thead>
<tr>
<th>Author</th>
<th>Region</th>
<th>Tool name</th>
<th>Recall method</th>
<th>Age (years)</th>
<th>No. food and drink items</th>
<th>Portion size estimation method</th>
<th>Use of prompts</th>
<th>Report supplement intake</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vereecken et al. (2014)[6]</td>
<td>Belgium</td>
<td>CANAA-W (formally YANA-C)</td>
<td>24HR, MR, SA</td>
<td>11–12</td>
<td>800, 25 food groups</td>
<td>Photographs (n 2100) of increasing portion sizes presented sequentially or simultaneously. Description and entry of g weight By count (i.e. the number of times a particular food was consumed)</td>
<td>For foods eaten in combination with other items. Reliability checking: warning when extreme amounts are entered</td>
<td>No</td>
</tr>
<tr>
<td>Moore et al. (2008)[21]</td>
<td>UK</td>
<td>SNAP™</td>
<td>24HR, MR, SA</td>
<td>7–15</td>
<td>49</td>
<td>Photographs (n 14 000) of eight increasing portion sizes presented simultaneously Avatar. For frequently forgotten foods</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Baranowski et al. (2014)[57]</td>
<td>USA</td>
<td>FIRSS1</td>
<td>24HR, MR, SA</td>
<td>9–11</td>
<td>Not mentioned</td>
<td>Photographs of seven increasing portion sizes to identify the amount of food served and amount of food left over.</td>
<td>For approximate time and for location of food/meal consumption</td>
<td>No</td>
</tr>
<tr>
<td>Carvalho et al. (2014)[22]</td>
<td>Portugal</td>
<td>PAC24</td>
<td>24HR, MR, SA</td>
<td>7–10</td>
<td>332 food and 41 drinks items</td>
<td>Photographs of seven increasing portion sizes to identify the amount of food served and amount of food left over.</td>
<td>For frequently forgotten foods. Meal gap review. Final probe to ensure complete capture of data</td>
<td>Yes</td>
</tr>
<tr>
<td>Storey et al. (2012)[23]</td>
<td>Canada</td>
<td>Web-SPAN</td>
<td>24HR, MR, SA</td>
<td>11–15</td>
<td>Not mentioned</td>
<td>Photographs of increasing portion sizes to identify the amount of food served and amount of food left over. Description of the amount served and left over is entered (using household measures)</td>
<td>For foods eaten in combination with other items. Asking for additional information (i.e. brand names). ‘Same as before’ option</td>
<td>No</td>
</tr>
<tr>
<td>Foster et al. (2014)[5]</td>
<td>UK</td>
<td>Intake24</td>
<td>24HR, MR, SA</td>
<td>11–16</td>
<td>≥1600</td>
<td>Food photographs and cues regarding beverage intake Photographs of increasing portion sizes to identify the amount of food served and amount of food left over. Description of the amount served and left over is entered (using household measures)</td>
<td>For foods eaten in combination with other items. Asking for additional information (i.e. brand names). ‘Same as before’ option</td>
<td>No</td>
</tr>
<tr>
<td>Diep et al. (2015)[19]</td>
<td>USA</td>
<td>ASA24-Kids-2012</td>
<td>24HR, MR, SA</td>
<td>9–11</td>
<td>5407 (food terms)</td>
<td>Photographs (n 9759) of increasing portion sizes presented sequentially</td>
<td>For frequently forgotten foods. Meal gap review. Final probe to ensure complete capture of data</td>
<td>Yes</td>
</tr>
<tr>
<td>Albar et al. (2016)[53]</td>
<td>UK</td>
<td>Myfood24</td>
<td>24HR, MR, SA</td>
<td>11–18</td>
<td>about 45 000 individual food/drink items</td>
<td>Portion size images for 5669 food items. Portion size suggestions (household measures). Free entry of g weight by participant</td>
<td>For foods eaten in combination with other items. Asking for additional information (i.e. brand names)</td>
<td>Yes</td>
</tr>
</tbody>
</table>

CANAA-W, Children and Adolescents’ Nutrition Assessment and Advice on the Web (a web-based 24-h dietary recall program for children and adolescents); YANA-C, Young Adolescents’ Nutrition Assessment on Computer (a web-based 24-h dietary recall program for children and adolescents); SNAP™, Synchronised Nutrition and Activity Program (a web-based 24-h recall program which assesses energy balance-related behaviours in children and adolescents); FIRSS1, Food Intake Recording Software System (a computer-based 24-h dietary recall program for children); PAC24, Portuguese self-administered, computerised, 24-h dietary recall (a web-based 24-h dietary recall program for children); Web-SPAN, Web-Survey of Physical Activity and Nutrition (a web-based 24-h dietary recall program for children and adolescents); Intake24, a web-based 24-h dietary recall program for children and young adults; ASA24-Kids-2012, Automated Self-Administered 24-Hour Dietary Assessment Tool, Kids-2012 version; myfood24, Measure your food on One day (a web-based 24-h dietary recall program for children and adults).
Table 2. Design characteristics of the web- and computer-based 24-h dietary recall tools developed for adults

<table>
<thead>
<tr>
<th>Author Region</th>
<th>Tool name</th>
<th>Recall method</th>
<th>Age (years)</th>
<th>No. food and drink items</th>
<th>Portion size estimation method</th>
<th>Use of prompts</th>
<th>Report supplement intake</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silman et al. (2011)(26) Europe EPIC-Soft 24HR, MR, IA ≥ 15 1500–2200 individual food/drink items and 150–350 recipes Photographs (in a book). Household measures. Standard units (in g, piece) and standard portions. Weighing or volume method Photographs of seven increasing portion sizes presented.</td>
<td>Reliability checking. For frequently forgotten foods. For foods eaten in combination with other items</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Touvier et al. (2011)(17) France NutriNet-Santé 24HR, MR, SA 48–75 Not mentioned Photographs of seven increasing portion sizes presented. Free entry of g weight by participant Household measures. Photographs of fruits (small, medium and large).</td>
<td>For frequently forgotten foods. For foods eaten in combination with other items</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daniel et al. (2013)(24) India NINA-DISH 24HR, FFQ, IA, MR 35–69 910 individual food/drink items Household measures.</td>
<td>For foods eaten in combination with other items</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shin et al. (2014)(28) Korea CAPIS 24HR, MR, IA 24–85 3642 and 1886 recipes Photographs presented ½, 1 and 1½ serving sizes for each item. Free entry of g weight by participant</td>
<td>For frequently forgotten foods</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foster (2014)(5) UK Intake24 24HR, MR, SA 17–24 ≥1600 Over 2000 portion size images representing over 100 food and drink types. Description of the amount served and left over is entered (using household measures)</td>
<td>Long time periods where no food is reported. For foods eaten in combination with other items. Asking for additional information (i.e. brand names). ‘Same as before’ option</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hillier et al. (2012)(49) UK SNAPATM 24HR, MR, SA Mean 34.4 (sd 11.1) 120 individual food/drink items Report fruit and vegetables consumed as numbers of portions</td>
<td>For frequently forgotten foods</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comrie et al. (2009)(27) UK FoRC 24HR, MR, SA 18–49 121 individual food/drink items Photographs to estimate portion size. Description and free entry of g weight by participant</td>
<td>The use of layers of questioning to prompt recall</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liu et al. (2011)(29) UK Oxford WebQ IM 24HR, MR, SA 18–65 21 food groups</td>
<td>Expanding questions to prompt recall and further detail</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zoellner et al. (2005)(60) USA IMM 24HR, MR, SA 18–65 167 individual food/drink items Photographs of four increasing portion sizes presented.</td>
<td>Audio instructions</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subar et al. (2007)(24) USA ASA24 24HR, MR, SA 18–77 7200 individual food/drink items 17 000 portion size images. Foods with portion size images had eight different portion sizes (in the latest version of ASA24) An avatar gives audio instructions.</td>
<td>For frequently forgotten foods. For foods eaten in combination with other items. Long time periods where no food is reported (meal gap)</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arab et al. (2010)(25) USA DietDay 24HR, MR, SA 21–69 9349 individual food/drink items 7000 portion size images which could be modified using command buttons</td>
<td>Reliability checking</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Park et al. (2015)(36) Korea GloboDiet 24HR, MR, IA 24–68 1305 individual food/drink items A picture book of foods/dishes was prepared of food portion sizes relevant to Korean diet. Volumes (directly or as household measures or as shapes and thickness). Standard unit. Free entry of g weight by participant</td>
<td>Adaptation of EPIC-Soft</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Continued

<table>
<thead>
<tr>
<th>Author</th>
<th>Tool name</th>
<th>Region</th>
<th>Portion size estimation method</th>
<th>No. food and drink items</th>
<th>Recall method</th>
<th>Age (years)</th>
<th>No. participants</th>
<th>Portion size images</th>
<th>Use of prompts</th>
<th>Food group quick list</th>
<th>Free entry of g weight by participant</th>
<th>Recall supplement intake</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freese et al. (2015)</td>
<td>myfood24</td>
<td>UK</td>
<td>Portion size images (household measures), Portion size suggestions (individual food/drink items)</td>
<td>About 45,000 individual food/drink items</td>
<td>24HR, MR, SA</td>
<td>18-65</td>
<td>About 47 participants and used one web-based recall compared with six of the identified 24HDR tools developed for adults (Table 2).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wardenaar et al. (2015)</td>
<td>Compl-eat™</td>
<td>The Netherlands</td>
<td>Portion size images for 5669 food items, and uses statistical methods to estimate usual intake distributions; Food group quick list which expands to further subgroups and individual items on selection.</td>
<td>Not mentioned</td>
<td>24HR, MR, SA</td>
<td>18-35</td>
<td>Not mentioned</td>
<td>Free entry of g weight by participant.</td>
<td>For frequently forgotten foods.</td>
<td>No (questioned separately).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Web- and computer-based 24-h dietary recall tools comparability/validation study design

Tables 3 and 4 present the comparison/validation study design features, and key findings for the identified 24HDR tools. For the purposes of this review, the authors considered a comparison study to be where the test measure was compared against a traditional 24-h recall or estimated/weighted food diary. A validation study was when the test measure was compared against direct observation (of eating occasions) or biological markers of nutrient intake. According to the authors’ interpretation of a comparison and validation study, of the 24HDR tools developed for both children and adults (Foster considered one study across two age groups), nine investigators conducted comparison studies and nine investigators conducted validation studies (EPIC-Soft and NutriNet-Santé were investigated in both comparison and validation studies). The remaining three tools, of the twenty-one identified in the present review, were not evaluated in either a comparison or validation study. Of the studies included in the present review the most common reference method in comparison studies was an interviewer-led multiple-pass recall and for validation studies was direct observation of eating occasions, closely followed by the use of biological markers of intake (Table 4). The number of participants ranged from forty-one (22) to 459 for comparison and validation studies conducted with children (23). There was similar variation with the numbers that participated in the adult studies with fifty-three (27) to 1074 (43). It is important to consider the reference method used when comparing participant numbers across studies; for example Touvier et al. (37) recruited 147 participants and used one web-based recall compared with one interviewer-led recall, whereas Conrie et al. (37) asked fifty-three participants to complete a 4-d estimated food diary as a reference. The lesser burden associated with the reference used by Touvier et al. (37) may allow for a greater number of participants to be involved compared with the burden (participant training, interview on diary and analysis of data) associated with a 4-d estimated food diary (27).

A variety of statistical measures were used to investigate comparability/validation of the 24HDR tools against reference methods. Some investigators used descriptive statistics such as the number of ‘matches’, ‘intrusions’ and ‘omissions’ between the food and drink items recorded by the web-based tool compared with the reference and in some cases linear and Poisson regression analysis was used to investigate the association of matches, intrusions and omissions between the test and reference (5, 19, 20, 47). Correlation analysis (Pearson, Spearman, κ and intraclass coefficients) was the most popular statistical method used to investigate the validity of nutrient intake and in some instances food group intake (50) reported by 24HDR tools compared with the output of reference methods. A wide range of correlation coefficients was reported by investigators. De Keyzer et al. (50) reported low correlation coefficients (r 0·6; P < 0·001) for intakes of thiamine recorded by EPIC-Soft when compared with intakes from an estimated food diary.
Table 3. Main findings of the validation/comparison studies of the web- and computer-based 24-h dietary recall tools developed for children

<table>
<thead>
<tr>
<th>Author</th>
<th>Tool name</th>
<th>Study type</th>
<th>Subjects (n)</th>
<th>No. of recalls</th>
<th>Reference method(s)</th>
<th>Statistical analysis</th>
<th>Main results for food groups</th>
<th>Main results for nutrients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moore et al.</td>
<td>SNAP™</td>
<td>Comparison</td>
<td>121</td>
<td>1 d</td>
<td>Interviewer-based 24HR</td>
<td>Mean differences in frequency (defined as count) of consumption. Bland–Altman was used to evaluate agreement between reported intakes from the test and reference method.</td>
<td>The mean difference was less than 1 count for all but three categories – cakes (1.15 counts), energy-dense foods (1.52 counts) and CHO-rich foods (0.97 counts)</td>
<td></td>
</tr>
<tr>
<td>Carvalho et al.</td>
<td>PAC24</td>
<td>Validation</td>
<td>41</td>
<td>1 d</td>
<td>Direct dietary observation</td>
<td>All food items identified as matches, intrusions and omissions. Bland–Altman was used to evaluate agreement between reported intakes from the test and reference method.</td>
<td>An average match rate of 67.0 %. Levels of intrusions and omissions were 11.5 and 21.5 %, respectively. 32 % of the actual intake was underestimated by PAC24</td>
<td></td>
</tr>
<tr>
<td>Storey et al.</td>
<td>Web-SPAN</td>
<td>Comparison</td>
<td>459</td>
<td>2 d</td>
<td>3 d estimated dietary record</td>
<td>ICC was used to measure the strength of agreement between the test and reference method. Pearson’s correlations were used to assess the association between intakes reported by the test and reference method. Paired-samples t tests were used to investigate differences between both measures.</td>
<td>Of all foods listed, 79.2 % were matches, 12.5 % omissions and 6.7 % intrusions. EI underestimated on average by 3 %. LOA for energy range from 0.52 to 1.82</td>
<td>ICC ranged from 0.24 to 0.40. Pearson r values ranged from 0.33 (fat) to 0.41 (protein) and for micronutrients 0.24 (vitamin A) to 0.39 (vitamin D). No significant differences were observed for protein, fat, Fe and Zn</td>
</tr>
<tr>
<td>Foster et al.</td>
<td>Intake24</td>
<td>Comparison</td>
<td>52</td>
<td>4 d</td>
<td>Interviewer-based 24HR</td>
<td>Bland–Altman was used to evaluate agreement between reported intakes from the test and reference method. All food items identified as matches, intrusions and omissions.</td>
<td>Of all foods listed, 79.2 % were matches, 12.5 % omissions and 6.7 % intrusions. EI underestimated on average by 3 %. LOA for energy range from 0.52 to 1.82</td>
<td></td>
</tr>
<tr>
<td>Baranowski et al.</td>
<td>FIRSSt</td>
<td>Validation</td>
<td>138</td>
<td>1 d</td>
<td>Direct dietary observation and interviewer-based 24HR</td>
<td>All food items identified as matches, intrusions and omissions. Pearson’s correlations were used to assess the association between intakes reported by the test and reference method.</td>
<td>Compared with the lunch observation, FIRSSt attained 46 % matches, 24 % intrusions and 30 % omissions. FIRSSt attained 80 % matches, 15 % intrusions and 24 % omissions against the interviewer-based 24HR for all meals. Match, intrusion and omission rates were 37, 27 and 35 %, respectively (site 1) and 53, 12 and 36 %, respectively (site 2) compared with observed intakes. Percentage matches between ASA24 kids and interviewer-led method were higher</td>
<td></td>
</tr>
<tr>
<td>Diep et al.</td>
<td>ASA24-Kids-2012</td>
<td>Validation</td>
<td>69</td>
<td>1 d</td>
<td>Direct dietary observation and interviewer-based 24HR</td>
<td>All food items identified as matches, intrusions and omissions. Pearson’s correlations were used to assess the association between intakes reported by the test and reference method.</td>
<td>Match, intrusion and omission rates were 37, 27 and 35 %, respectively (site 1) and 53, 12 and 36 %, respectively (site 2) compared with observed intakes. Percentage matches between ASA24 kids and interviewer-led method were higher</td>
<td></td>
</tr>
<tr>
<td>Albar et al.</td>
<td>myfood24</td>
<td>Comparison</td>
<td>75</td>
<td>2 d</td>
<td>Interviewer-based 24HR</td>
<td>ICC was used to measure the strength of agreement between the test and reference method. Bland–Altman was used to evaluate agreement between reported intakes from the test and reference method.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SNAP™, Synchronised Nutrition and Activity Program (a web-based 24-h recall program which assesses energy balance-related behaviours in children and adolescents); 24HR, 24-h recall; CHO, carbohydrate; PAC24, Portuguese self-administered, computerised, 24-h dietary recall (a web-based 24-h dietary recall program for children); Web-SPAN, Web-Survey of Physical Activity and Nutrition (a web-based 24-h dietary recall program for children and adults); ICC, intraclass correlation coefficient; Intake24, a web-based 24-h dietary recall program for children and young adults; EI, energy intake; LOA, limits of agreement; FIRSSt, Food Intake Recording Software System (a computer-based 24-h dietary recall program for children); ASA24-Kids-2012, Automated Self-Administered 24-Hour Dietary Assessment Tool, Kids-2012 version; myfood24, Measure your food on One day (a web-based 24-h dietary recall program for children and adults).
In contrast, Touvier et al.\(^{177}\) reported a high correlation coefficient (r 0.92; P < 0.001) for niacin intakes recorded by women using NutriNet-Santé and an interviewer-led 24HDR. Liu et al.\(^{29}\) also reported a mean correlation coefficient (r 0.6; P < 0.001) for twenty-one nutrients examined to demonstrate the relationship between the test and reference measure across a range of nutrients. In addition to correlation analysis, Storey et al.\(^{25}\) and De Keyzer et al.\(^{50}\) used paired t tests and Student’s t tests to investigate significant differences in the reporting of nutrient intakes between the test and reference methods.

Bland & Altman\(^{51}\) analysis was only used by six investigators, four of which investigated the agreement and bias in nutrient intakes recorded by test and reference methods\(^{5,27,52,53}\). Moore et al.\(^{24}\) investigated agreements between methods in the recording of food and drink items and Carvalho et al.\(^{22}\) investigating 95 % observed amounts consumed using Bland–Altman analysis.

The investigators of the 24HDR tools developed for children, included in this review, did not use biomarkers of nutrient intake to investigate validity. Table 4 shows the 24HDR tools developed for adults including four 24HDR tools that were validated using biological markers for nutrient intake. DietDay\(^{33}\) is the only 24HDR tool that used doubly labelled water (DLW) in validation as a biomarker of usual energy intake\(^{54}\). Results indicated that the performance of DietDay was more superior than a paper-based FFQ or diet history questionnaire. Three investigators (Wardenaar et al.\(^{52}\), Ferrari et al.\(^{43}\) and Lassale et al.\(^{55}\)) used urinary biomarkers (urinary N, urinary N and K and urinary N, K and Na, respectively) to investigate the relationship between reported and true intakes. The investigators reported moderate to strong correlations between reported nutrient intakes and their respective urinary biomarkers: r 0.37 (95 % CI 0.03, 0.70) for Na intakes reported by women\(^{55}\), r 0.65 (95 % CI 0.45, 0.79) for protein intakes reported by athletes\(^{52}\) and r 0.83 (95 % CI 0.637, 0.932) for reported N intakes\(^{43}\).

Discussion

Tool design characteristics

The range of ages involved in the testing of the 24HDR tools demonstrates the applicability of web- and computer-based dietary assessment methods to the majority of the population. Investigators did note age-specific adaptations to some of the 24HDR tools including the use of cartoon avatars\(^{56-57}\) and portion size pictures specifically developed for children\(^{56-57}\). Baranowski et al.\(^{57}\) and subsequently Diep et al.\(^{69}\) used adapted versions of the National Cancer Institute’s Automated Self-Administered 24-Hour Dietary Assessment Tool (ASA24\(^{6}\)) which had been modified for children and excluded food preparation questions deemed too difficult for children to answer\(^{68}\). However, none of the investigators noted changes to the 24HDR tools that would potentially enhance the user experience for older adults (for example, larger format or font). At present, there is no single 24HDR tool suitable for all age groups; however, some investigators such as Baranowski et al.\(^{57}\), Subar et al.\(^{52}\), Hillier et al.\(^{49}\) and Moore et al.\(^{45}\) have worked to develop separate age-appropriate versions of the same tool. Foster\(^{5}\) also tested the same tool with children and with young adults and Albar et al.\(^{55}\) proposes to test myfood24 with adults aged 18 years and over. Intuitive design features such as ‘same as before’, ‘meal gap review’ and spell check function were unique to some 24HDR tools but could potentially be useful in all web- and computer-based dietary assessment designs (whether it be diary or FFQ based).

Performance

For the majority of 24HDR tools included in this review, the investigators concluded that the performance of the respective tool was acceptable when compared with reference methods (comparison and/or validation studies). There was a variety of concluding remarks from the investigators: ‘The test measure was in good agreement with the reference method\(^{4,5,17,23,27,29}\);’  ‘The test method performed well but in some instances the reference method was more accurate\(^{19,20,47}\);’ ‘The test measure was more superior to the reference measure\(^{55}\);’ and ‘Although the findings were promising the test method required further modifications to improve accuracy\(^{22,45,49}\).’

Overall, the 24HDR tools whose performance was investigated, either in a comparison and/or validation study, demonstrated promising potential. Investigators that reported the reference method as being more accurate used direct observation as the reference method\(^{19,47,59}\). Whilst this objectively investigates the accuracy of the test measure, for example in the case of Kirkpatrick et al.\(^{47}\), it may lead to the underestimation of the performance of ASA24 if compared with how other 24HDR tools were assessed for accuracy, for example, Zoellner et al.\(^{56}\). The performance of the ASA24 may be underestimated as the bilingual interactive multimedia (IMM) dietary assessment tool was compared against an interviewer-led 24-h recall, the findings of which may agree more favourably compared with direct observation. As eating is an observable behaviour\(^{61}\), direct observation of eating occasions is an expensive and invasive measure and is usually only feasible for short periods of time (for example, lunch or dinner times) which may not accurately represent the performance of the tool for recording longer periods of intake (for example, 24-h period)\(^{19}\).

Recovery biomarkers measure dietary intake with little error and have been successfully used to validate some of the studies included in this review; however, these biomarkers are limited to certain nutrients and cannot validate all nutrient estimates from a test dietary assessment measure. Based on the findings of this review, the most appropriate use of reference measures for investigating the validity of a 24HDR tool is the combination of an estimated dietary record in conjunction with biological markers of intake or direct observation of eating occasions as undertaken by Arab et al.\(^{55}\) and Baranowski et al.\(^{20}\), respectively.

In the assessment of the performance of 24HDR tools, acceptability of the tool is an aspect that is not always considered. In the comparison of ASA24 with an interviewer-led 24-h recall, Thompson et al.\(^{62}\) noted that a substantial proportion of study participants preferred the self-administered ASA24 approach compared with the interviewer-led recall.
Table 4. Main findings of the validation/comparison studies of the web- and computer-based 24-h dietary recall tools developed for adults

<table>
<thead>
<tr>
<th>Author</th>
<th>Tool name</th>
<th>Study type</th>
<th>Subjects (n)</th>
<th>No. of recalls</th>
<th>Reference method</th>
<th>Summary of statistical analysis</th>
<th>Main results for food groups</th>
<th>Main results for nutrients</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Keyzer et al. (2011)(50)</td>
<td>EPIC-Soft</td>
<td>Comparison</td>
<td>127</td>
<td>2 d</td>
<td>5 d estimated dietary record</td>
<td>Student’s t test was used to investigate whether intakes reported by the test and reference measure differed significantly. Spearman and ρ correlations were used to assess the association between intakes reported by the test and reference method</td>
<td>Macronutrient correlation coefficients ranged from 0.47 (protein) to 0.62 (CHO) and for micronutrients 0.16 (thiamine) to 0.56 (Fe) Significant differences between methods for the reporting of ten nutrients, for example, fat</td>
<td></td>
</tr>
<tr>
<td>Ferrari et al. (2009)(42)</td>
<td>EPIC-Soft</td>
<td>Validation</td>
<td>1072</td>
<td>1 d</td>
<td>24-h urinary biomarkers</td>
<td>Pearson correlations were used to assess the association between intakes reported by the test and the respective biological value</td>
<td>Correlation coefficients between means of urinary and 24HR measurements were 0.838 and 0.756 for N and K intakes, respectively</td>
<td></td>
</tr>
<tr>
<td>Touvier et al. (2011)(17)</td>
<td>NutriNet-Santé</td>
<td>Comparison</td>
<td>147</td>
<td>1 d</td>
<td>Interviewer-based 24HR</td>
<td>ICC was used to measure the strength of agreement between the test and reference method. Pearson’s correlations were used to assess the association between intakes reported by the test and reference method</td>
<td>ICC ranged from 0.5 for fats/sauces (both sexes), breakfast cereals, cakes/biscuits/pastries and dairy food (women only) to 0.9 for fruits, pulses (both sexes), breakfast cereals, alcoholic drinks and meat (men only)</td>
<td>The range for energy-adjusted Pearson’s coefficients for macronutrients was 0.80 (protein intake recorded by men) to 0.88 (protein intake recorded by women) and for micronutrients was 0.54 (retinol recorded by men) to 0.92 (niacin recorded by women)</td>
</tr>
<tr>
<td>Lassale et al. (2015)(50)</td>
<td>NutriNet-Santé</td>
<td>Validation</td>
<td>199</td>
<td>3 d</td>
<td>24-h urinary biomarkers</td>
<td>ICC was used to measure the strength of agreement between the test and reference method. Pearson’s correlations were used to assess the association between intakes reported by the test and reference method</td>
<td>Correlations between reported and true intakes were 0.61, 0.78 and 0.47 for men and 0.64, 0.42 and 0.37 for women for protein, K and Na, respectively</td>
<td>Correlations between reported and true intakes were 0.61, 0.78 and 0.47 for men and 0.64, 0.42 and 0.37 for women for protein, K and Na, respectively</td>
</tr>
<tr>
<td>Foster (2014)(5)</td>
<td>Intake24</td>
<td>Comparison</td>
<td>167</td>
<td>4 d</td>
<td>Interviewer-based 24HR</td>
<td>Bland–Altman was used to evaluate agreement between reported intakes from the test and reference method. All food items identified as matches, intrusions and omissions</td>
<td>For all foods listed, 82 6 % were matches, 9 % omissions and 7.5 % intrusions. Energy intake under-estimated on average by 1 %. LOA for energy range from 0.50 to 1.97</td>
<td>LOA for CHO, fat, vitamin C, Fe and alcohol were 0.51−1.99, 0.43–3.21, 0.18–5.75, 0.43–2.24 and 0.09–10.91, respectively</td>
</tr>
<tr>
<td>Hillier et al. (2012)(49)</td>
<td>SNAPATM</td>
<td>Validation</td>
<td>77</td>
<td>5 d</td>
<td>Direct dietary observation on 4 d</td>
<td>Pearson’s correlation was used to assess the association between intakes reported by the test and reference method. Pearson’s correlations were used to assess the association between intakes reported by the test and reference method</td>
<td>The mean match rate was 81.7 %, with an intrusion rate of 5.6 %. Pearson’s correlations ranged from 0.39 to 0.56 for percentage fat and fruit and vegetable intake respectively</td>
<td></td>
</tr>
<tr>
<td>Comrie et al. (2009)(27)</td>
<td>FoRC</td>
<td>Comparison</td>
<td>53</td>
<td>4 d</td>
<td>4 d estimated dietary record</td>
<td>Spearman’s correlation was used to assess the association between intakes reported by the test and reference method. Wilcoxon signed-rank test was used to investigate whether intakes reported by the test and reference measure differed significantly. Bland–Altman was used to evaluate agreement between reported intakes from the test and reference method</td>
<td>Intakes of fat, NSP and bread were similar between the two methods. FoRC recorded significantly lower intakes of energy and alcohol and higher intakes of fruit and vegetables and cereals</td>
<td>Correlation coefficients at a food-group level ranged from 0.4 for alcohol to 0.76 for bread. For the few nutrients investigated, all r values were statistically significant and were greater than 0.5 (except for percentage fat)</td>
</tr>
<tr>
<td>Liu et al. (2011)(29)</td>
<td>Oxford WebQ</td>
<td>Comparison</td>
<td>116</td>
<td>1 d</td>
<td>Interviewer-based 24HR</td>
<td>The percentage differences in energy and nutrient intake between the two methods were calculated. Spearman’s correlation was used to investigate agreement between estimates for each nutrient. Comparison of tertile intakes from the test and reference measure to identify misreporters</td>
<td>Mean Spearman’s correlation for the 21 nutrients was 0.6. Macronutrient r values ranged from 0.57 (fat) to 0.66 (CHO) and for micronutrients 0.37 (vitamin E) to 0.72 (Mg). The mean differences in intake were less than 10 % for all nutrients except carotene and vitamins B12 and D</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Continued

<table>
<thead>
<tr>
<th>Author</th>
<th>Tool name</th>
<th>Study type</th>
<th>Subjects (n)</th>
<th>No. of recalls</th>
<th>Reference method</th>
<th>Summary of statistical analysis</th>
<th>Main results for food groups</th>
<th>Main results for nutrients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zoellner et al.</td>
<td>IMM</td>
<td>Comparison</td>
<td>80</td>
<td>1 d</td>
<td>Interview-based 24 HR</td>
<td>ANOVA tests were performed to determine the effect of substituting standardised portion sizes for reported portion sizes. Unadjusted and energy-adjusted correlations were calculated to measure the strength of relationship between the IMM recall and the interview-administered recall.</td>
<td>Unadjusted correlation coefficients were 0.6 and energy-adjusted correlations were lower. Macronutrient unadjusted r values ranged from 0.44 (fat) to 0.78 (CHO) and for micronutrients 0.29 (folate) to 0.7 (Ca)</td>
<td></td>
</tr>
<tr>
<td>Kirkpatrick et al.</td>
<td>ASA24</td>
<td>Validation</td>
<td>81</td>
<td>3 d</td>
<td>Feeding study, in which the true intake for three meals was known</td>
<td>Linear regression models were used to examine the association between recall mode and the proportion of items truly consumed for which a match was reported. Poisson regression was used to assess the association between recall mode and the number of items reported but not truly consumed.</td>
<td>Average match rate of 80% between the recall and the reference method. Mean number of intrusions for all eating occasions combined was 2.6 for the ASA24.</td>
<td></td>
</tr>
<tr>
<td>Arab et al.</td>
<td>DietDay</td>
<td>Validation</td>
<td>261</td>
<td>6 d</td>
<td>Double-labelled water and diet history questionnaire</td>
<td>The ratio of reported intakes to TEE measurements was calculated to identify reporting bias and misreporters. Pearson correlation for agreement between recall and biomarker. Attenuation factor was estimated by using the regression coefficient from the regression model.</td>
<td>Attenuation factors were 0.30 for blacks and 0.26 for whites. Adjusted correlations between true energy intake and the recalls were 0.50 and 0.47 for blacks and whites, respectively.</td>
<td></td>
</tr>
<tr>
<td>Wardenaar et al.</td>
<td>Compl-eat™</td>
<td>Validation</td>
<td>47</td>
<td>3 d</td>
<td>24-h urinary N</td>
<td>Paired t tests were used to determine difference between biomarker value and reported intake. Pearson correlations were used to rank individuals according to intake. Attenuation factors and CI (calculated using slope of the linear regression of log transformed biomarker values and reported intakes). Bland–Altman was used to evaluate agreement between reported intakes and biological value.</td>
<td>Estimated mean dietary protein intake was 109.6 (±33.0) g/day according to the test measure compared with 141.3 (±38.2) g/day as determined by N excretion. Reasonably good association of intakes for protein estimation (0.65, 95% CI 0.45, 0.79). Under-reporting of protein intakes was larger with higher intakes of protein v. lower intakes</td>
<td></td>
</tr>
</tbody>
</table>

EPIC-Soft, European Prospective Investigation into Cancer and Nutrition Software (a standardised computerised 24-h dietary recall method); CHO, carbohydrate; 24HR, 24-h recall; NutriNet-Santé, a web-based 24-h dietary recall program for adults; ICC, intraclass correlation coefficient; Intake24, a web-based 24-h dietary recall program for children and young adults; LOA, limits of agreement; SNAPA™, Synchronised Nutrition and Activity Program for Adults (a web-based 24-h dietary recall program for adults); FoRC, Food Recall Checklist (121-item food recall list which has been used to measure diet in undergraduate students); Oxford WebQ, a web-based 24-h dietary recall program for adults; IMM, a bilingual interactive multimedia dietary assessment tool based on the 24-h dietary recall method (a web-based 24-h dietary recall program for adults); ASA24, Automated Self-Administered 24 Hour Dietary Assessment Tool (web-based 24-h dietary recall tool for adults, developed by the National Cancer Institute, USA); DietDay, a web-based 24-h dietary recall program for adults; TEE, total energy expenditure; Compl-eat™, a computer-based 24-h dietary recall program for adults.
However, Vereecken et al.\textsuperscript{(48)} highlighted the opposite, when parents preferred a traditional approach to recording diet (i.e. a written record) compared with a web- or computer-based tool. Therefore, to ensure prolonged usage of web- and computer-based dietary assessment methods, research into the acceptability of these methods should be considered in the comparison/validation study design so that improvements can be made if warranted.

**Accuracy/precision**

Although the accuracy of the 24HDR tools is difficult to compare across the board, one can consider the results reported within the individual studies. It is known that one statistical measure in isolation cannot determine the validity of any method of dietary assessment\textsuperscript{(360)}. However, due to the variation and combination of statistical measures used across studies, for the purposes of this review statistical measures are compared in isolation across studies. Correlation coefficients are commonly reported across studies investigating the accuracy of 24HDR tools. Comparing the ranges of correlation coefficients for macronutrients and micronutrients reported across all 24HDR tools, NutriNet-Santé\textsuperscript{(17)} had the strongest \( r \) values for intakes derived from the web-based tool compared with a reference measure, with \( r \) values ranging for macronutrient intake as high as 0·88 (protein intakes recorded by women) and for micronutrient intake an \( r \) value of 0·92 (niacin recorded by women). As with all studies investigating the comparability/validity of a test measure of dietary assessment compared with a reference, it is important to consider the study design. Touvier et al.\textsuperscript{(17)} compared intakes of NutriNet-Santé against intakes derived from an interviewer-led dietary recall at a single time point, whereas De Keyzer et al.\textsuperscript{(50)} compared intakes across two separate time points and Conrie et al.\textsuperscript{(27)} compared intake across four separate time points. It is, therefore, difficult to make conclusions about accuracy of one tool compared with another, particularly using just one type of statistical analysis. The number of recalls administered across the comparison/validation studies varied from a single dietary recall\textsuperscript{(241)} to six recalls\textsuperscript{(33)}. The difference in the number of recalls administered compared with the number of days of dietary recording by the reference method is important to consider when investigating the agreement between methods in their assessment of nutrient intake.

However, when comparing correlation coefficients from studies that have similar study designs (for example, Liu et al.\textsuperscript{(29)} (r 0·37 to 0·72); Touvier et al.\textsuperscript{(17)} (r 0·54 to 0·92); Zoellner et al.\textsuperscript{(60)} (r 0·29 to 0·78)), all of whom used a single time point of data collection in their investigation and used an interviewer-led recall as a reference, Touvier et al.\textsuperscript{(17)} performed the best based on this type of analysis. Another type of measure that was used to investigate validity of 24HDR tools was the incidence of ‘Matches, Intrusions and Omissions’ whereby food and drink items recorded using the test measure are compared with the reference measure. Again, there was variation across studies depending on the reference measure used; for example comparing the output of the web-based tool against direct observation, the number of percentage matches was not as high (ranging from 46 to 67 \%\textsuperscript{(20,22)}) when compared with percentage matches identified when comparing the web-based tool with interviewer-led recalls (79·2 and 82·6 \% both\textsuperscript{(5)}).

Precision (the degree to which the same method produces the same value on repeated measures) was less frequently reported. Based on the limited number of investigators that calculated the limits of agreement in their analysis, it is not possible to compare the precision of the methods included in this review. This type of analysis was included in a report on the performance of INTAKE24 (a web-based 24-h dietary recall program for children and young adults\textsuperscript{(50)}. Foster\textsuperscript{(5)} assessed the precision of INTAKE24 compared with other web- and computer-based methods of dietary assessment, which had included these findings in their analysis, by calculating the width of the limits of agreement. The lack of this statistical interpretation in dietary assessment validation studies was also noted by Lombard et al.\textsuperscript{(360)}.

Another factor affecting the design of the comparison/validation study was the proximity and sequence of the test and reference measure. Margest & Nelson\textsuperscript{(63)} suggested that the sequence (i.e. that the test measure be administered before the reference method) and the proximity of the test and reference measure need to be carefully considered to avoid raising the apparent level of agreement between the measures. The cross-over design is a popular design for testing the performance of 24HDR tools\textsuperscript{(5,24,60)}. The design involves the use of the test and reference method in the same day. Although this limits the impact of variation of diet on the results, completing both methods in the same day may introduce a ‘learning effect’. Having acknowledged this in the study design, Foster\textsuperscript{(5)} attempted to limit this effect by weighting the order in which the test measure was administered to participants so that the impact of a ‘learning effect’ could be investigated retrospectively. Interestingly, although the weighed food record is considered the ‘gold standard’ (in the absence of unbiased measures) of dietary assessment\textsuperscript{(60)}, none of the investigators included this as a reference method. Alternatives may have been used instead due to the high respondent burden and cost associated with weighed food diaries\textsuperscript{(64)}.

Biological markers of nutrient intake can serve as an objective validation of dietary assessment methods as they reflect nutritional status, but are independent of dietary intake assessment\textsuperscript{(65)}. Of all the tools included in this review, Arab et al.\textsuperscript{(33)} was the only investigator to use DLW to assess the validity. Whilst DLW is considered the ‘gold standard’ recovery biomarker of nutrient intake, urinary N\textsuperscript{(60)} and urinary K\textsuperscript{(67)} are also useful biomarkers in the validation of dietary assessment tools. Biomarkers of nutrient intake from blood were not used by any of the investigators included in this review. Biomarkers of nutrient intake from plasma and serum samples (for example, plasma vitamin C and plasma carotenoids) have been used in other validation studies concerning novel dietary assessment methodologies\textsuperscript{(13,68,69)}. In the studies investigating the validity of 24HDR tools in children (included in the present review), none of the investigators used biological markers of nutrient intake. A reason for this may be that biological samples are difficult to obtain from children\textsuperscript{(70)}. Baranowski et al.\textsuperscript{(200)} used other strategies such as obtaining a bogus pipeline hair sample from children using the FIRSSt tool to encourage more accurate
reporting of dietary intake. Using this method, the researchers informed participants that the hair sample provided would reveal what they had eaten through chemical analysis. Collecting this sample appeared to reduce the level of omissions when intakes from FIRSS1 were compared with an interviewer-administered recall(120).

**Conclusion**

Web- and computer-based 24HDR tools are cost-effective, useful methods for assessing dietary intake from different populations. The overall accuracy of these methods is difficult to determine because, in many cases, direct comparisons cannot be made. Across-the-board agreement between these 24HDR tools for macro- and micronutrient intake is generally strong when compared against references. Few studies used biomarkers in the validation of 24HDR tools; however, the investigators that did use biomarkers showed that these methods provide reliable estimates of protein intake, moderate estimates of energy intakes and reliable estimates for micro-nutrient intake such as K and Na. This review highlights some findings which may be applied when designing and investigating the performance of 24HDR tools in the future. Age-specific adaptations have been shown to be of benefit in younger populations, and may prove beneficial for older adult populations also. In the assessment of performance and validity, direct observation is a useful reference method; however, it is often limited to short periods of intake, whereas estimated records in conjunction with biological markers of intake may be more feasible methods of investigating validity. It is important to be mindful of the type of statistical analysis used to investigate validity. Lombard et al.(560) noted that in some cases more than three statistical measures are required to truly assess the validity of dietary assessment measures. Lastly, this review demonstrates that although a standardised dietary assessment methodology is preferential for nutrition surveillance across countries, valuable technology-based 24HDR tools exist and could be used to collect intermittent data for continual health and nutrition research purposes.

**Acknowledgements**

The present review was completed as part of a Food Institutional Research Measure (FIRM) project, DietIreland, funded by the Department of Agriculture, Food and Marine (13/F/424).

The present review was completed as part of a project conceived by E. R. G. and A. F.; C. M. T. and E. R. G. devised the study and collected the data. The present manuscript was written by E. R. G. and A. F.; C. M. T. drafted the manuscript. All authors approved the final content of the manuscript. The authors declare no conflict of interest.

**References**


