Review Article

Household food insecurity: a systematic review of the measuring instruments used in epidemiological studies

Emanuele S Marques1,*, Michael E Reichenheim1, Claudia L de Moraes1,2, Marina ML Antunes3 and Rosana Salles-Costa3

1Institute of Social Medicine, Rio de Janeiro State University (UERJ), Rua Francisco Xavier 524, 7º andar – bloco D, CEP: 20550-900, Maracanã, Rio de Janeiro-RJ, Brazil; 2Family Health Program, Estácio de Sá University, Rio de Janeiro, Brazil; 3Nutrition Institute Josué de Castro, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Submitted 18 June 2013: Final revision received 24 February 2014: Accepted 13 March 2014: First published online 25 June 2014

Abstract

Objective: To conduct a systematic review aimed at identifying and characterizing the experience-based household food security scales and to synthesize their psychometric properties.

Design: Search in the MEDLINE, LILACS and SciELO databases, using the descriptors (‘food insecurity’ OR ‘food security’) AND (‘questionnaires’ OR ‘scales’ OR ‘validity’ OR ‘reliability’). There was no limitation on the period of publication. All articles had their titles and abstracts analysed by two reviewers. The studies of interest were read in their entirety and the relevant information extracted using a standard form.

Results: The initial bibliographic search identified 299 articles. Of these, the 159 that seemed to meet the criteria for inclusion were read fully. After consultation of the bibliographic references of these articles, twenty articles and five documents were added, as they satisfied the previously determined criteria for inclusion. Twenty-four different instruments were identified; all were brief and of easy application. The majority were devised in the USA. Forty-seven references reported results of psychometric studies. The instruments that presented the highest number of psychometric studies were the Core Food Security Measurement/Household Food Security Survey Module (CFSM/HFSSM) and the Self-Perceived Household Food Security Scale.

Conclusions: There are a number of structured scales available in the literature for characterization of household food insecurity. However, despite some psychometric studies already existing about the majority of the instruments, it is observed that, except for the studies of the CFSM/HFSSM, these are still restricted to appraisal of a few aspects of reliability and validity.

In recent decades, a growing number of epidemiological studies and governmental assessments have shown interest in the evaluation of food insecurity at the household level. Worldwide research was particularly boosted after the establishment of the Millennium Development Goals, the primary goal of which was to reduce by half the proportion of people living on less than $US 1.25 per day and suffering overt hunger between 1990 and 2015(1). The increase in food prices due to the recent global economic crisis of 2008–2009 also reinforces the importance of research as a foundation for redirecting social policies to fight hunger in many countries(2).

Most of the knowledge has come from epidemiological studies. These have focused mainly on the magnitude, risk factors and consequences of household food insecurity (HFI), as well as assessments of intervention programmes aimed at reducing the effect of food insecurity. More recently, there has been a growing recognition that HFI is a difficult construct to measure due to conflicting definitions implying different forms of operationalization(3–5).

For many years, HFI has been assessed by indirect methods, such as food availability, purchasing power, consumption profile and anthropometric measurements, and the main objective was to quantify the number of individuals in a situation of food shortage or even outright hunger(4). As of the end of the 1970s, it became clear that indirect methods were insufficient to cover all dimensions of the food insecurity construct. One of the issues raised was the need to establish appropriate indicators to identify...
and monitor food insecurity in intervention studies\(^{(6)}\). Several initiatives were launched, and different tools were proposed for measuring the perception and/or the experience of families suffering from food insecurity\(^{(6)}\). Yet, despite all efforts to devise direct measures, a variety of methods and the absence of a reference standard for assessing HFI have so far precluded a consensus as to which should be the measure of choice to operationalize the concept in epidemiological studies. The purpose of the present study was to provide a systematic review of the peer-reviewed literature published in scientific journals, with an aim to identify and scrutinize direct instruments to measure HFI, in particular the experience-based household food security scales. Several features were focused on, i.e. the scale’s origin (country, place and language), main characteristics (number of items and type of response options), uses (in applied research and cross-cultural studies) and psychometric background (number and types of studies/assessed properties). This account may help in identifying the best available scales for use in decision making and research settings and/or as an aid in developing new ones.

**Methods**

**Search strategy**

A systematic literature search was conducted in three sources: (i) a worldwide electronic database, MEDLINE (consulted through PubMed); and two comprehensive Latin American databases, (ii) LILACS and (iii) SciELO. There was no limitation on the period of publication. A search was performed in June 2011 by a single reviewer, using the following descriptors: ‘food insecurity’ OR ‘food security’) AND (‘questionnaires’ OR ‘scales’ OR ‘validity’ OR ‘reliability’). This descriptor was also used for the respective Portuguese and Spanish versions. Thereafter, two reviewers independently scrutinized the reference sections of all identified psychometric and review articles, with an aim to detect additional scientific papers not spotted in the first search round. This approach also allowed identification of some government reports introducing the main characteristics of the scales found in the first stage of the search. All of the references were filed and handled using EndNote X6\(^{TM}\).

**Selection criteria**

Only articles published in English, Spanish or Portuguese were accepted. These had to approach the subject through direct tools to measure different aspects of HFI, such as quantity, quality and/or access to food. Monographs, dissertations, academic theses, government and institutional reports, summaries of scientific events, books and articles merely expressing points of view/opinions of experts were not eligible and thus barred from further scrutiny.

The selection of the articles of interest involved a thorough scrutiny of the titles and abstracts, and was carried out by the independent reviewers mentioned before. The articles were then classified according to whether they: (i) definitely met the inclusion criteria; (ii) could possibly meet the inclusion criteria, but required full reading for confirmation; or (iii) definitely did not meet the criteria and should, therefore, be excluded. An interobserver reliability evaluation was carried out using the $\kappa$ coefficient with quadratic weighting as estimator\(^{(7,8)}\). Disagreements between the reviewers were further discussed with the other authors and settled by consensus.

**Classifying articles**

The same reviewers independently read all articles in full. These were classified into six non-mutually exclusive groups: (i) articles that made use of the measuring instruments for HFI in epidemiological studies; (ii) articles focusing on cross-cultural adaptation processes; (iii) psychometric studies; (iv) articles presenting new instruments; (v) review articles; and (vi) key documents used as guidelines in the development process of the instrument.

**Data extraction**

Extracting information from the selected articles was effected by the independent reviewers, using a purposefully designed form. The following information was sought: (i) title/name of the HFI measurement tool; (ii) the first reference introducing the instrument (authors and year of publication); (iii) country where the instrument was devised; (iv) language; (v) number of component items; (vi) types of response options; and (vii) the number of articles using the instrument in epidemiological studies and/or focusing on cross-cultural adaptation and/or evaluating psychometric properties. One of the reviewers further detailed the latter feature in tandem with two epidemiologists experienced in psychometrics. The following information was sought: (i) place and year of the study; (ii) number of items actually used in a particular analysis; (iii) sample size; (iv) method of application and reference period; and (v) psychometric features evaluated, the estimators used and respective results.

The following classification was used to synthesize information on the reliability of scales: intra-observer (or test–retest) reliability, inter-observer reliability and internal consistency\(^{(9)}\). Validity studies were classified according to Streiner and Norman\(^{(9)}\) and Terwee et al.\(^{(10)}\) in four types: (i) face or content validity; (ii) structural (dimensional) validity; (iii) criterion or concurrent validity; and (iv) construct validity. Only measurement tools with three or more items were included\(^{(11)}\).

**Results**

As presented in Fig. 1, the first stage of the systematic review identified 299 references published from 1979 to the date of search; 279 (93.3%) were indexed in
MEDLINE/PubMed. Of the total, 140 had to be excluded since they did not fulfill the eligibility criteria. This evaluation stage met with substantial agreement between the reviewers according to Shrout’s criteria(12), showing a weighted \( \kappa \) of 0.853 (95% CI 0.777, 0.917).

Once reading the remaining 159 articles in full (stage 2), another thirty-five were further excluded for not meeting the inclusion criteria. This stage also involved consulting the reference sections of forty psychometric articles and nine reviews, which added another twenty-five references. Five of those were official documents concerning the development process of some scales.

**Characteristics of the instruments**

Twenty-four instruments were identified in the 184 references evaluated. Their main features are found in Table 1. Although the majority (58.3%) were devised and developed in the USA(3,13–23), research groups in countries like Canada(24), Venezuela(25), Colombia(26), Costa Rica(27), Burkina Faso(28), Kenya(29), Iran(30), Bangladesh(17,31) and Indonesia(32) also took initiatives. Most of the instruments were originally conceived in English, are brief, and hold dichotomous responses.

The instruments mostly found in the scientific literature were the Core Food Security Measurement/Household Food Security Survey Module (CFSM/HFSSM; sixty-nine articles), the HFSSM Six-Item Short Form (HFSSM-6SF; seventeen articles), the Modified Radimer/Cornell Scale (a) (R/CSm_a; fifteen articles) and the Self-Perceived Household Food Security Scale (SPHFSS; ten articles).

Besides the respective original instruments, modified versions of the Community Childhood Hunger Identification Project (CCHIP), CFSM/HFSSM, HFSSM-6SF, Food Insecurity Questions of NHANES III (NHANES III_FIQ) and Radimer/Cornell Scale (R/CS) were also detected.

The instruments hold different terms/nomenclatures and underlying concepts. Terms mostly used were: (i) household food insecurity (eight instruments); (ii) food insecurity (seven instruments); (iii) hunger (six instruments); (iv) food security (four instruments); (v) food insufficiency (two instruments); (vi) food insecurity past (one instrument); and (vii) household hunger (one instrument). The concepts of food insecurity, household food insecurity, hunger and food security used by most scales were outlined by Anderson(33). The concept of household hunger that stood at the foundation for drafting the R/CS refers to three central issues: food depletion, food unsuitability and food anxiety(34). Moreover, the issue of scarce financial resources was at the core in nearly all of the instruments’ development processes.

Figure 2 shows the periods in which the instruments emerged, as well as the links between them. As early as 1977, the Agricultural Research Service of the US Department of Agriculture (USDA) pioneered the food sufficiency question in the Continuing Survey of Food Intake by Individuals (CSFII). This question fed into the Third National Health and Nutrition Examination Survey (NHANES III), which carried a variant of it along with several items from the CCHIP. This scale was developed in the early 1980s in the USA by the Community Childhood Hunger Identification Project(35). The CCHIP intended to

---

**Fig. 1** Schematic representation of the systematic review
Table 1 General characteristics of detection instruments for household food insecurity identified in the present review and their respective publications* †

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Acronym</th>
<th>Country</th>
<th>Language</th>
<th>No. of items</th>
<th>Response option</th>
<th>Variants‡: language and country</th>
<th>Use studies</th>
<th>CCA studies</th>
<th>Psychometric studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community Childhood Hunger Identification Project(35)</td>
<td>CCHIP</td>
<td>USA</td>
<td>English</td>
<td>8</td>
<td>Dichotomous</td>
<td></td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Modified CCHIP(16)</td>
<td>CCHIP-m</td>
<td>USA</td>
<td>English</td>
<td>7</td>
<td>Dichotomous</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Self-Perceived Household Food Security Scale and its linguistic and semantic variants(29)</td>
<td>SPHFS</td>
<td>Venezuela</td>
<td>Spanish</td>
<td>12</td>
<td>Dichotomous</td>
<td></td>
<td>2</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Household Food Security of Bangladesh (a)(17)</td>
<td>HFSB_a</td>
<td>Bangladesh</td>
<td>Bengali</td>
<td>11</td>
<td>Varies with item</td>
<td>Spanish: Colombia(68,69)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Food Security of Bangladesh (b)(31)</td>
<td>HFSB_b</td>
<td>Bangladesh</td>
<td>Bengali</td>
<td>–</td>
<td>Not defined</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latin American Food Security Measurement Scale(26)</td>
<td>ELCSA</td>
<td>Colombia</td>
<td>Spanish</td>
<td>17</td>
<td>Dichotomous</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Experience-Based Measurement of Household Food Insecurity(28)</td>
<td>EbMHFI</td>
<td>Burkina Faso</td>
<td>French</td>
<td>9</td>
<td>Dichotomous</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Measurement of Household Food Insecurity(27)</td>
<td>MHFI</td>
<td>Costa Rica</td>
<td>Spanish</td>
<td>14</td>
<td>Ordinal scale</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Household Food Security Access Scale and its linguistic and semantic variants(21)</td>
<td>HFIAS</td>
<td>USA</td>
<td>English</td>
<td>9</td>
<td>Dichotomous</td>
<td>Amharic: Ethiopia(70)</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Core Food Security Measurement/ Household Food Security Survey Module and its linguistic and semantic variants(35)</td>
<td>CFSM/HFSSM</td>
<td>USA</td>
<td>English</td>
<td>18</td>
<td>Dichotomous</td>
<td>Swahili: Tanzania(71)</td>
<td>37</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Food Insecurity by Elders(67)</td>
<td>FIE</td>
<td>USA</td>
<td>English</td>
<td>14</td>
<td>Ordinal scale</td>
<td>Spanish: Argentina(73), Bolivia(74), Ecuador(75), Mexico(76,77), Peru(78)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Items of Food Insecurity(29)</td>
<td>IFI</td>
<td>Kenya</td>
<td>English</td>
<td>4</td>
<td>Ordinal scale</td>
<td>English: Canada(81)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>HFSSM-6SF and its linguistic and semantic variants(14)</td>
<td>HFSSM-6SF</td>
<td>USA</td>
<td>English</td>
<td>6</td>
<td>Dichotomous</td>
<td>Spanish: Argentina(73)</td>
<td>10</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Modified HFSSM Six-Item Short Form(22)</td>
<td>HFSSM-6SFm</td>
<td>USA</td>
<td>English</td>
<td>5</td>
<td>Dichotomous</td>
<td>English: Trinidad &amp; Tobago(86)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tool to Assess Past Food Insecurity(20)</td>
<td>TAPFI</td>
<td>USA</td>
<td>Spanish</td>
<td>7</td>
<td>Varies with item</td>
<td>Portuguese: Brazil(89)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Insecurity Questions of NHANES II(39)</td>
<td>NHANES III_FIQ</td>
<td>USA</td>
<td>English</td>
<td>12</td>
<td>Varies with item</td>
<td>Spanish: Mexico(90)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modified Food Insecurity Questions of NHANES III(35)</td>
<td>NHANES III_FIQm</td>
<td>USA</td>
<td>Spanish</td>
<td>4</td>
<td>Varies with item</td>
<td>Farsi/Persian: Iran(91)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radimer/Cornell Scale(34)</td>
<td>R/CS</td>
<td>USA</td>
<td>English</td>
<td>12</td>
<td>Varies with item</td>
<td>Russian: Russian Federation(93)</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Modified Radimer/Cornell Scale (a)(33)</td>
<td>R/CSm_a</td>
<td>USA</td>
<td>English</td>
<td>10</td>
<td>Ordinal scale</td>
<td>English: Vanuatu(94)</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Korean: Korea(95)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>English: UK(96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Malaysian: Malaysia(97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
evaluate the prevalence of hunger among low-income families with children up to the age of 12 years. New initiatives from US government institutions and academia subsequently gained prominence. Driven by the NHANES III and Cornell University, this movement culminated in the development of important indicators for the area. NHANES III enabled an instrument with questions related to the sufficiency of food and hunger to be used at the national level\(^{(36)}\). At the same time, researchers from Cornell University proposed an instrument based on women’s reports of their perception and experience of hunger, as well as their difficulties in obtaining adequate food\(^{(34,37)}\).

The stimulus generated by research and publications in the 1980s, pushed by a growing awareness for the need to develop new measurement tools to address food insecurity adequately, bore fruit early in the next decade. The Conference on Food Security Measurement and Research was one of the most significant initiatives in this period. The event convened many researchers and experts in the area and sought to develop a consensus on a conceptual basis for a measure of food insecurity and hunger to be used throughout the USA and to debate how to operationalize this concept\(^{(38)}\). A year later, the USDA opted to measure food insecurity through a new instrument, the CFMS\(^{(39)}\). This measurement tool had the CCHIP\(^{(35)}\) and the R/CS\(^{(34,37,40)}\) at its core, as well as other scales focusing on the experience of HFI caused by limitations of economic resources in meeting basic needs of the individual and families\(^{(41)}\).

Despite the efforts, it was not until the 2000s that the development of most scales occurred. Besides the development of entirely new instruments, adaptation of existing ones also took place. This was the case with the CFMS scale. Updated in 2000 by the Food and Nutrition Service of the USDA, it became known as the HFSSM\(^{(3)}\). Since then, the CFMS and HFSSM are recognized as one because they hold exactly the same items. They differ only with respect to the order with which the questions are presented\(^{(3)}\). The CCHIP, R/CS, CFMS/HFSSM form the base of the majority of the instruments available nowadays.

**Psychometric properties of the instruments**

Table 2 summarizes forty-seven studies evaluating the psychometric properties of the instruments. The CFMS/HFSSM (eighteen articles) is the most often evaluated instrument according to the consulted peer-reviewed literature. The SPHFSS follows with seven. The HFSSM-6SF shows five studies. Six out of twenty-four identified instruments did not have any psychometric study: the Modified CCHIP (CCHIP-m); the Household Food Security of Bangladesh (a) (HFSB_a); the Household Food Security of Bangladesh (b) (HFSB_b); the Modified HFSSM Six-Item Short Form (HFSSM-6SFm); the Modified Food Insecurity Questions of NHANES III (NHANESIII-FIQm); and the Modified Radimer/Cornell Scale (c) (R/CSn_c).
The psychometric studies date back to the 1990s and have been mostly carried out in the USA. The majority are restricted to some aspects of the reliability and validity. Internal consistency estimated through Cronbach’s $\alpha$ coefficient has been the most analysed facet of reliability. A few studies also report results from structural (dimensional) and construct validity evaluations.

In all, the CFSM/HFSSM, PHFSS and HFSSM-6SF have thus far been the most thoroughly assessed instruments. Their psychometric evaluations were mostly carried out in the last 15 years or so. According to the reviewed literature, the CFSM/HFSSM and HFSSM-6SF have been studied in many countries worldwide, whereas the PHFSS only in Latin America (Colombia and Venezuela). Internal consistency has been mostly supported; eleven studies on the CFSM/HFSSM and HFSSM-6SF showing Cronbach’s $\alpha$ coefficients ranging from 0.73 to 0.95, and six studies on the PHFSS with $\alpha$ varying from 0.82 to 0.94. A few studies also assessed other facets of reliability, such as test–retest reliabilities (CFSM/HFSSM: $r = 0.66$; $r = 0.75$ and PHFSS: $r = 0.98$).

All three instruments had some form of construct (structural) validity supported. Regarding the CFSM/HFSSM, among the fourteen validity studies detected in the eligible literature, nine assessed structural validity. Eight of those employed a complete Rasch analysis, which on the whole supported a one-dimensional structure and the appropriateness of the component items. Six out of seven psychometric studies on the PHFSS evaluated structural validity employing exploratory factor analyses, Rasch analyses and/or full structural equation models. These studies suggested a two-factor structure. Convergent validity was also sustained in all three instruments, with several studies showing some association between the food insecurity measure and predictor variables as expected. Moreover, acceptable screening capability of the HFSSM-6SF was demonstrated by two criterion validity studies, which showed sensitivities ranging from about 85% to almost 100%, and specificities from just below 80% to nearly 100%.

**Discussion**

Since the end of the 1970s, efforts have been directed to define better and characterize HFI worldwide. Triggered...
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Place and year</th>
<th>No. of items*</th>
<th>Sample size</th>
<th>Application method (recall period)</th>
<th>Psychometric results (type of study in italics)</th>
</tr>
</thead>
</table>
| CFSM/HFSSM and its linguistic and semantic variants | USA, 2000<sup>(96)</sup> | 18            | 1664        | Telephone (12 months)             | • Test–retest reliability: Pearson correlation coefficient: $r = 0.75$ ($P < 0.01$)  
                                                                                                                  Face validity: Unverified  
                                                                                                                  Construct validity: Partially corroborated → factorial analysis via the Rasch model. Findings presented for GOI of the respondents indicated an acceptable rate of misfit. This did not vary with family status or with any ethnic group except the Samoans  
                                                                                                                  Structural validity: Partly corroborated → factorial analysis via the Rasch model. This did not vary with family status or with any ethnic group except the Samoans  
                                                                                                                  Internal consistency: $\alpha = 0.91$  
                                                                                                                  Face validity: Unverified  
                                                                                                                  Construct validity: Partially corroborated → factorial analysis via the Rasch model. Findings presented for GOI of the respondents indicated an acceptable rate of misfit. This did not vary with family status or with any ethnic group except the Samoans  
                                                                                                                  Structural validity: Partially corroborated → factorial analysis via GLM and Rasch model. Findings presented for GOI of the respondents indicated an acceptable rate of misfit. This did not vary with family status or with any ethnic group except the Samoans  
                                                                                                                  Internal consistency: $\alpha = 0.91$  
                                                                                                                  Face validity: Uncorroborated  
                                                                                                                  Construct validity: Corroborated → responses with only one affirmative response to item ‘Only worried about FI’ exhibited behaviours consistent with FI (decreased vegetable intake, greater reliance on alternative food resources and low-cost foods)  
                                                                                                                  Structural validity: Corroborated → factorial analysis via GLM and Rasch model. Findings presented for GOI of the respondents indicated an acceptable rate of misfit. This did not vary with family status or with any ethnic group except the Samoans  
                                                                                                                  Internal consistency: $\alpha = 0.91$  
                                                                                                                  Face validity: Not presented for the population of Hawaii  
                                                                                                                  Construct validity: Corroborated → respondents with only one affirmative response to item ‘Only worried about FI’ exhibited behaviours consistent with FI (decreased vegetable intake, greater reliance on alternative food resources and low-cost foods)  
                                                                                                                  Structural validity: Corroborated → factorial analysis via GLM and Rasch model. Findings presented for GOI of the respondents indicated an acceptable rate of misfit. This did not vary with family status or with any ethnic group except the Samoans  
                                                                                                                  Internal consistency: $\alpha = 0.91$  
                                                                                                                  Face validity: Uncorroborated  
                                                                                                                  Construct validity: Corroborated → respondents with only one affirmative response to item ‘Only worried about FI’ exhibited behaviours consistent with FI (decreased vegetable intake, greater reliance on alternative food resources and low-cost foods)  
                                                                                                                  Structural validity: Corroborated → factorial analysis via GLM and Rasch model. Findings presented for GOI of the respondents indicated an acceptable rate of misfit. This did not vary with family status or with any ethnic group except the Samoans  
                                                                                                                  Internal consistency: $\alpha = 0.91$  
                                                                                                                  Face validity: Uncorroborated  
                                                                                                                  Construct validity: Corroborated → respondents with only one affirmative response to item ‘Only worried about FI’ exhibited behaviours consistent with FI (decreased vegetable intake, greater reliance on alternative food resources and low-cost foods)  
                                                                                                                  Structural validity: Corroborated → factorial analysis via GLM and Rasch model. Findings presented for GOI of the respondents indicated an acceptable rate of misfit. This did not vary with family status or with any ethnic group except the Samoans  
                                                                                                                  Internal consistency: $\alpha = 0.91$  
                                                                                                                  Face validity: Uncorroborated  
                                                                                                                  Construct validity: Corroborated → respondents with only one affirmative response to item ‘Only worried about FI’ exhibited behaviours consistent with FI (decreased vegetable intake, greater reliance on alternative food resources and low-cost foods)  
                                                                                                                  Structural validity: Corroborated → factorial analysis via GLM and Rasch model. Findings presented for GOI of the respondents indicated an acceptable rate of misfit. This did not vary with family status or with any ethnic group except the Samoans  
                                                                                                                  Internal consistency: $\alpha = 0.91$  
                                                                                                                  Face validity: Uncorroborated  
                                                                                                                  Construct validity: Corroborated → respondents with only one affirmative response to item ‘Only worried about FI’ exhibited behaviours consistent with FI (decreased vegetable intake, greater reliance on alternative food resources and low-cost foods)  
                                                                                                                  Structural validity: Corroborated → factorial analysis via GLM and Rashe
### Table 2 Continued

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Place and year</th>
<th>No. of items*</th>
<th>Sample size</th>
<th>Application method</th>
<th>Psychometric results (type of study in italics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil, 2008&lt;sup&gt;(106)&lt;/sup&gt;</td>
<td>15</td>
<td>108,606</td>
<td>Interview (3 months)</td>
<td>• <em>Structural validity:</em> Corroborated → factorial analysis via Rasch model. Scale presented good fitness and most item GOF values were within adequate range, being practically identical when comparing female and male responses. None of the items presented substantial DIF</td>
<td></td>
</tr>
<tr>
<td>Brazil, 2008&lt;sup&gt;(107)&lt;/sup&gt;</td>
<td>15</td>
<td>847</td>
<td>Interview (3 months)</td>
<td>• <em>Structural validity:</em> Corroborated → factorial analysis via the Rasch model. Except for one item, GOF statistics were within a range considered adequate, indicating a common phenomenon being measured with approximately equal discrimination</td>
<td></td>
</tr>
<tr>
<td>Iran, 2009&lt;sup&gt;(83)&lt;/sup&gt;</td>
<td>15</td>
<td>2004</td>
<td>– (12 months)</td>
<td>• <em>Structural validity:</em> Corroborated → factorial analysis via the Rasch model. Most Item GOF statistics were near unity, and none exceeded 1.20</td>
<td></td>
</tr>
<tr>
<td>Brazil, 2009&lt;sup&gt;(108)&lt;/sup&gt;</td>
<td>15(?)</td>
<td>458</td>
<td>Interview (3 months)</td>
<td>• <em>Internal consistency:</em> $\alpha = 0.91$</td>
<td></td>
</tr>
<tr>
<td>USA, 2009&lt;sup&gt;(109)&lt;/sup&gt;</td>
<td>15</td>
<td>242</td>
<td>Interview (3 months), 12 months before and during pregnancy</td>
<td>• <em>Internal consistency:</em> $\alpha = 0.91$ (before pregnancy) and $\alpha = 0.88$ (during pregnancy)</td>
<td></td>
</tr>
<tr>
<td>Peru, 2010&lt;sup&gt;(78)&lt;/sup&gt;</td>
<td>15</td>
<td>300</td>
<td>Interview (12 months)</td>
<td>• <em>Internal consistency:</em> $\alpha = 0.86$ (overall); $\alpha = 0.76$ (anxiety about food supply); $\alpha = 0.86$ (anxiety about food quality); $\alpha = 0.84$ (food ingestion by adults); $\alpha = 0.75$ (food ingestion by children)</td>
<td></td>
</tr>
<tr>
<td>USA, 2011&lt;sup&gt;(110)&lt;/sup&gt;</td>
<td>15</td>
<td>63</td>
<td>Self-administered (12 months)</td>
<td>• <em>Internal consistency:</em> $\alpha = 0.73$</td>
<td></td>
</tr>
<tr>
<td>SPHFSS and its linguistic and semantic variants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Venezuela, 1999&lt;sup&gt;(25)&lt;/sup&gt;</td>
<td>12</td>
<td>238</td>
<td>Interview (6 months)</td>
<td>• <em>Internal consistency:</em> $\alpha = 0.82$</td>
<td></td>
</tr>
<tr>
<td>Colombia, 2005&lt;sup&gt;(68)&lt;/sup&gt;</td>
<td>12</td>
<td>193</td>
<td>Interview (6 months)</td>
<td>• <em>Internal consistency:</em> $\alpha = 0.90$</td>
<td></td>
</tr>
<tr>
<td>Colombia, 2005&lt;sup&gt;(69)&lt;/sup&gt;</td>
<td>12</td>
<td>1624</td>
<td>Interview (6 months)</td>
<td>• <em>Internal consistency:</em> $\alpha = 0.95$ for the first factor and $\alpha = 0.89$ for the second</td>
<td></td>
</tr>
<tr>
<td>Venezuela, 2007&lt;sup&gt;(111)&lt;/sup&gt;</td>
<td>12</td>
<td>924</td>
<td>Interview (6 months)</td>
<td>• <em>Internal consistency:</em> $\alpha = 0.89$</td>
<td></td>
</tr>
</tbody>
</table>

*Note: GOF = Goodness of Fit, SPHFSS = Socio-Political Health Food Security Survey, HFI = Household Food Insecurity, FI = Food Insecurity.*
Table 2 Continued

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Place and year</th>
<th>No. of items*</th>
<th>Sample size</th>
<th>Application method (recall period)</th>
<th>Psychometric results (type of study in italics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colombia, 2008&lt;sup&gt;(112)&lt;/sup&gt;</td>
<td>12</td>
<td>1319</td>
<td>Interview (30 d)</td>
<td><strong>Structural validity</strong>: Corroborated → factorial analysis via EFA and Rasch model. Most CHFSS items presented good fitness with most GOF values within the adequate range. Only two adult items exhibited DIF between urban and rural households</td>
<td></td>
</tr>
<tr>
<td>Colombia, 2009&lt;sup&gt;(113)&lt;/sup&gt;</td>
<td>12</td>
<td>2784</td>
<td>Interview (30 d)</td>
<td><strong>Construct validity</strong>: Corroborated → statistically significant associations were found between household food insecurity and diagnoses of children’s diarrhoea, respiratory infections and parasitosis. Risk for child stunting and underweight increased in a dose-response way as FI became more severe</td>
<td></td>
</tr>
<tr>
<td>Colombia, 2009&lt;sup&gt;(114)&lt;/sup&gt;</td>
<td>12</td>
<td>108</td>
<td>Interview (30 d)</td>
<td><strong>Internal consistency</strong>: ( \alpha = 0.92 ) in the 1st application and ( 0.94 ) in the 2nd</td>
<td></td>
</tr>
<tr>
<td>HFSSM-6SF and its linguistic and semantic variants</td>
<td>US, 1999&lt;sup&gt;(14)&lt;/sup&gt;</td>
<td>6</td>
<td>44 647</td>
<td>–</td>
<td><strong>Criterion validity</strong>: In order to determine the HFI in general, in households without children and in households with children, Sens = 92.0 %; 99.7 % and 85.9 % and Spec = 99.4 %, 99.3 % and 99.5 %, respectively. To determine the FI with hunger, Sens = 84.7 %, 90.3 % and 78.4 %, and Spec = 99.6 %, 99.9 % and 99.2 %, respectively</td>
</tr>
<tr>
<td>Trinidad and Tobago, 2004&lt;sup&gt;(115)&lt;/sup&gt;</td>
<td>6</td>
<td>531</td>
<td>Interview (12 months)</td>
<td><strong>Internal consistency</strong>: ( \alpha = 0.87 )</td>
<td></td>
</tr>
<tr>
<td>Trinidad and Tobago, 2005&lt;sup&gt;(115)&lt;/sup&gt;</td>
<td>6</td>
<td>1903</td>
<td>Self-administered (12 months)</td>
<td><strong>Internal consistency</strong>: ( \alpha = 0.77 )</td>
<td></td>
</tr>
<tr>
<td>Iran, 2007&lt;sup&gt;(91)&lt;/sup&gt;</td>
<td>6</td>
<td>300</td>
<td>Interview (12 months)</td>
<td><strong>Criterion validity</strong>: using hunger screening as a reference standard → Sens = 97.8 %, Spec = 85.5 %. Using hidden hunger screening as a reference standard → Sens = 23.5 %, Spec = 96.9 %</td>
<td></td>
</tr>
<tr>
<td>Australia, 2009&lt;sup&gt;(116)&lt;/sup&gt;</td>
<td>6</td>
<td>49</td>
<td>Self-administered –</td>
<td><strong>Internal consistency</strong>: ( \alpha = 0.90 )</td>
<td></td>
</tr>
<tr>
<td>R/CSm_a</td>
<td>US, 1995&lt;sup&gt;(13)&lt;/sup&gt;</td>
<td>10</td>
<td>193</td>
<td>Interview –</td>
<td><strong>Internal consistency</strong>: ( \alpha = 0.84 ) (household level/factor); ( \alpha = 0.86 ) (individual level/factor) and ( \alpha = 0.85 ) (child level/factor)</td>
</tr>
<tr>
<td>R/CSm_b</td>
<td>Indonesia, 2001&lt;sup&gt;(32)&lt;/sup&gt;</td>
<td>9</td>
<td>1356</td>
<td>Interview (varied)</td>
<td><strong>Structural validity</strong>: Corroborated → factorial analysis via EFA. Confirmed that hunger and FI are experienced differently at family level and at individual level. There are distinct aspects related to the child and the adult</td>
</tr>
<tr>
<td>R/CSm_d</td>
<td>US, 2004&lt;sup&gt;(19)&lt;/sup&gt;</td>
<td>8</td>
<td>381</td>
<td>Telephone/interview (12 months)</td>
<td><strong>Internal consistency</strong>: ( \alpha = 0.94 )</td>
</tr>
</tbody>
</table>

<sup>*</sup> Items presented good fitness with most GOF values within the adequate range.
### Table 2 Continued

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Place and year</th>
<th>No. of items</th>
<th>Sample size</th>
<th>Application method (recall period)</th>
<th>Psychometric results (type of study in italics)</th>
</tr>
</thead>
</table>
| R/CSm_e          | Iran, 2007(30)       | 16           | 250         | Interview (12 months)             | **Internal consistency**: $\alpha = 0.90$ (factor: family insecurity); $\alpha = 0.82$ (individual insecurity) and $\alpha = 0.796$ (child hunger scales)  
**Structural validity**: Corroborated → factorial analysis via EFA. Solution with three factors. In most cases, factor loadings of the items were as expected and similar to those reported in previous studies  
**Construct validity**: Partially corroborated → individual insecurity and child hunger were inversely correlated with monthly per capita income, father’s education, mother’s education and father’s occupational status, and positively correlated with household size, as expected. However, household insecurity did not follow the same pattern |
| CCHIP            | USA, 1992(35)        | 8            | 377         | Interview (varied)                | **Internal consistency**: $\alpha = 0.86$  
**Face validity**: Corroborated  
**Structural validity**: Corroborated → factorial analysis via EFA. Solution with one factor, accounting for 52% of the variance  
**Construct validity**: Corroborated → FI was strongly associated in the expected direction with economic and socio-demographic variables, with reliance on strategies to cope with food shortages and with health problems of the children  
**Criterion validity**: Sens = 86%; Spec = 73% |
| USA, 1997(118)   | 8 (?)                | 193          | Interview (30 d) | -                                  | **Criterion validity**: Sens = 89%; Spec = 63% |
| R/CS and its linguistic and semantic variants | USA, 1992(37)        | 12           | 193         | Interview -                        | **Internal consistency**: $\alpha = 0.91$ (factor 1); $\alpha = 0.89$ (factor 2); $\alpha = 0.92$ (factor 3)  
**Structural validity**: Corroborated → factorial analysis via EFA. Solution with three factors accounting for 62% of the variance. Household factor alone accounted for 53% of the variance, indicating the strong presence of a single factor  
**Construct validity**: Corroborated → relationships between the R/CS and other scales (CCHIP and NHANES) occurred as expected  
**Criterion validity**: Sens = 89%; Spec = 63% |
| USA, 1997(118)   | 12 (?)               | 1993         | Interview - | -                                  | **Criterion validity**: Sens = 89%; Spec = 63% |
| HFIAS and its linguistic and semantic variants | Ethiopia, 2009(70)   | 9            | 99          | Interview (30 d)                  | **Internal consistency**: $\alpha = 0.85$ (1st application); $\alpha = 0.84$ (2nd) and $\alpha = 0.83$ (3rd)  
**Face validity**: Corroborated  
**Construct validity**: Corroborated → household per capita income was inversely associated with FI score. Dose–response trends between FI level and likelihood of previous-day consumption of various foods were observed for meats (not including fish), vegetables and fruits, with minor deviations between moderate and severe levels of FI  
**Criterion validity**: Sens = 89% |
| Tanzania, 2010(71) | 9                    | 237          | Interview (30 d) | -                                  | **Internal consistency**: $\alpha = 0.90$ (overall); $\alpha = 0.83$ (food quality subscale); $\alpha = 0.89$ (food quantity subscale)  
**Structural validity**: Corroborated → factorial analysis via EFA. Solution with two factors. Both factors explained 69% of the total variance  
**Construct validity**: Corroborated → FI was positively associated with maternal education, husband’s education, household wealth status, being of an agricultural rather than pastoral tribe and animal-source food consumption; it was negatively associated with maternal age and household size |
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Place and year</th>
<th>No. of items*</th>
<th>Sample size</th>
<th>Application method (recall period)</th>
<th>Psychometric results (type of study in italics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELCSA</td>
<td>Colombia, 2010</td>
<td>17</td>
<td>150</td>
<td>Interview (30 d)</td>
<td>• <em>Internal consistency</em>: $\alpha = 0.93$ (10 items for families with adults only); $\alpha = 0.96$ (17 items for families with adults, youths and children)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• <em>Structural validity</em>: Corroborated — factorial analysis via EFA. Solution with three factors, accounting for 75% of the variance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EbMHFI</td>
<td>Burkina Faso, 2006</td>
<td>9</td>
<td>126</td>
<td>Interview (varied)</td>
<td>• <em>Internal consistency</em>: $\alpha$ varied from 0.81 to 0.85 in the five applications of the scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• <em>Construct validity</em>: Corroborated — HFI was associated with predictor variables (dietary intake, nutritional status and economic status) as expected</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• <em>Face validity</em>: Corroborated</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• <em>Structural validity</em>: Corroborated — factorial analysis via EFA. Solution with one factor that explained 46% of the variation</td>
</tr>
<tr>
<td>MHFI</td>
<td>Costa Rica, 2008</td>
<td>14</td>
<td>213</td>
<td>Interview (12 months)</td>
<td>• <em>Internal consistency</em>: $\alpha = 0.89$ (three response categories) and $\alpha = 0.87$ (dichotomized responses)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• <em>Construct validity</em>: Corroborated — there was an association between the levels of FI and poverty line method, as well as unsatisfied basic needs method and integrated poverty classification</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFI</td>
<td>Kenya, 2011</td>
<td>4</td>
<td>6795</td>
<td>–</td>
<td>• <em>Internal consistency</em>: $\alpha = 0.73$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• <em>Structural validity</em>: Corroborated — factorial analysis via IRT. Likelihood ratio test rejected the one-parameter model in favour of the two-parameter one. Loewinger coefficient $H = 0.88$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• <em>Construct validity</em>: Corroborated — FI has a significant and negative association with household income level</td>
</tr>
<tr>
<td>TAPFI</td>
<td>USA, 2006</td>
<td>7</td>
<td>85</td>
<td>Interview –</td>
<td>• <em>Internal consistency</em>: $\alpha = 0.84$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• <em>Construct validity</em>: Corroborated — past FI was significant correlation with maternal education, crowding in the mother’s childhood household and past food insufficiency. Foreign-born Latino mothers reported significantly greater levels of past food insecurity than US-born mothers, demonstrating discriminant validity</td>
</tr>
<tr>
<td>FIE</td>
<td>USA, 2003</td>
<td>14</td>
<td></td>
<td>–</td>
<td>• <em>Criterion validity</em>: Sens $= 89$%; Spec $= 69$%</td>
</tr>
<tr>
<td>NHANES III_FIQ</td>
<td>USA, 1997</td>
<td>12(?)</td>
<td>193</td>
<td>Interview –</td>
<td>• <em>Criterion validity</em>: Sens $= 90$%; Spec $= 32$%</td>
</tr>
<tr>
<td>SAFSSM</td>
<td>USA, 2004</td>
<td>9</td>
<td>345</td>
<td>Self-administered (12 months)</td>
<td>• <em>Structural validity</em>: Partly corroborated — factorial analysis via the Rasch model. Poorer model fit. Response patterns differed between younger ($\leq 12$ years) and older ($&gt; 12$ years)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• <em>Criterion validity</em>: Simulation of precision of classification based on a population with ‘true’ FI distributed approximately as observed in the sample, gave the following results: on the threshold of FI (gross score $\geq 2$), Sens $= 0.89$, Spec $= 0.93$ and PPV $= 0.82$; on the threshold of FI with hunger (gross score $\geq 6$), Sens $= 0.77$, Spec $= 0.98$ and PPV $= 0.71$</td>
</tr>
</tbody>
</table>

CFSM/HFSSM, Core Food Security Measurement/Household Food Security Survey Module; SPHFSS, Self-Perceived Household Food Security Scale; HFSSM-6SF, HFSSM Six-Item Short Form; R/CSm_a, Modified Radimer/Cornell Scale (a); R/CSm_b, Modified Radimer/Cornell Scale (b); R/CSm_d, Modified Radimer/Cornell Scale (d); R/CSm_e, Modified Radimer/Cornell Scale (e); CCHIP, Community Childhood Hunger Identification Project; R/CS, Radimer/ Cornell Scale; HFAS, Household Food Insecurity Access Scale; ELCSA, Latin American Food Security Measurement Scale; EbMHFI, Experience-Based Measurement of Household Food Insecurity; MHFI, Measurement of Household Food Insecurity; IFI, Items of Food Insecurity; TAPFI, Tool to Assess Past Food Insecurity; FIE, Food Insecurity by Elders; NHANES III_FIQ, Food Insecurity Questions of NHANES III; SAFSSM, Self-Administered Food Security Survey Module for Children Aged 12 and Older; GOF, goodness of fit; FI, food insecurity; GLM, general linear model; IRT, Item Response Theory; 2PL, the two-parameter model; HFI, household food insecurity; DIF, differential item functioning; EFA, exploratory factor analysis; CHFSS, adapted Colombian Household Food Security Survey; SEM, structural equation model; Sens, sensitivity; Spec, specificity; PPV, positive predictive value. **‘?’ sign means that the study did not explicitly discriminate the number of items on the scale used. The same number of items as in the original version was thus considered. **
by the pioneering strategies proposed in the Community Childhood Hunger Identification Project\(^{(35)}\) and the Radiner research group at Cornell University\(^{(34,37)}\); and driven by the recognition of the Human Right to Adequate Food\(^{(42)}\) and the establishment of the Millennium Development Goals\(^{(3)}\), this movement reached different continents in the decades that followed. The large number of competing instruments is testimony to the commitment to this quest.

The common feature is that all instruments were developed from food insecurity and hunger concepts linked to a lack of access to food due to scarce financial resources. Questions (items) related to other aspects of the construct, such as food safety, have hardly been addressed so far\(^{(45)}\).

A feature found on some instruments relates to different forms of operationalization, which takes into account whether or not a child is present in the household (e.g. CCHIP, SPHFSS, CFSSM/HFSSM, HFSB\(_{b}\), Latin American Food Security Measurement Scale (ELCSA), Measurement of Household Food Insecurity (MHI), Items of Food Insecurity (IFI), Tool to Assess Past Food Insecurity (TAPFI) and R/CS). Adapting a version for families with children has been justified by the specific nutritional needs in this age group in terms of quantity, quality and regularity of food supply. The same rationale may also apply to the elder. However, despite research addressing food insecurity in this age group, there is only one particular measurement tool so far available in this respect\(^{(44-46)}\).

Maybe it would be timely to invest in the development of more food insecurity instruments specially tailored for this fast growing sub-population. Elderly persons often have distinct metabolic and nutritional requirements, and many find themselves in unwelcoming social contexts. Besides food-access problems due to inadequate resources to purchase food, there are also logistical problems, such as getting to food stores or even preparing meals.

There were discernible methods comprising the evaluation processes. Studies covered various countries, settings (household, health services and educational institutions) and population groups (women, adolescents and children, among others). They also used different forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (self-application, interview, telephone) and recall periods (1 to 12 months). This array of forms of application (satisfactory and positive scenario, not many instruments underwent strict testing according to the evaluated peer-reviewed literature. Of all twenty-four instruments identified, only one (CFSSM/HFSSM) has been evaluated through fifteen or more psychometric studies. Another instrument had seven studies (SPHFSS), two had five studies (HFSSM-6SF), yet four were tested only twice, and another eleven just once. In addition, most of these studies repeatedly explored features assessed in preceding ones (e.g. reliability/internal consistency), thus failing to shed new light on many other important psychometric features. Important properties, such as the sustainability of the theoretically proposed dimensional structure; item reliability and absence of measurement error correlations; factor-based convergent and discriminant validity; scalability; item positioning (difficulty/intensity) \(v.s.-v.t\) the latent trait; and the evaluation of measurement invariance (heterogeneity/item differential functioning), have not been extensively scrutinized\(^{(11,47-50)}\).

In fairness, there are studies addressing some dimensional aspects, but most are restricted to exploratory factor analyses and on the whole barely cover the confirmatory-type scrutiny as is the case of the mentioned properties. As understood from the evidence stemming from the literature found in peer-reviewed scientific journals, the exceptions are the CFSSM/HFSSM and to a lesser extent the SPHFSS and the HFSSM-6SF. Not only have these instruments been tested more frequently, but their history comprises several evaluations suggesting the tenability of the underlying properties.

In this regard the CFSSM/HFSSM stands out prominently, with several Rasch analyses disclosing the good performance of the component items, while also disclosing subtleties such as the difficulty in calibrating the instrument in families with and without children. In fact, the adequacy of the CFSSM/HFSSM may also be identified in the respective institutional and/or government literature. Several technical reports published by the USDA since 1995 consistently provide substantiation to the reliability and validity of the instrument\(^{(45,51-53)}\). For instance, a report by Hamilton \textit{et al.} (1997) found the food security scale to have good reliability, including good content validity and good construct validity\(^{(45)}\). Their report also showed expectedly high correlations between food security measured by the scales under scrutiny and weekly food expenditures per household, annual household income and income relative to the poverty line\(^{(45)}\). More recently, too, a study by Nord on the potential technical enhancements to the CFSSM/HFSSM using Rasch analysis showed favourable results regarding item severity parameters in both adult and child food security scales\(^{(54)}\). Valuable institutional information may also be found in regard to recent offshoots of the CFSSM/HFSSM; for instance, the ELCSA. Reports showed adequate internal consistency measured by Cronbach’s \(\alpha\) coefficient in the range of 0.91–0.96, and as anticipated, strong correlations between the scale and several sociodemographic variables like poverty, housing conditions and access to public services\(^{(55,56)}\).

Regarding the assessment of other aspects of validity, and more specifically in relation to criterion validity, a challenge concerns the lack of agreement as to the standard to use. Some studies turned to other food insecurity scales as reference, especially the CFSSM/HFSSM. This strategy may overestimate sensitivity and specificity because offshoot instruments usually have many items in common with the purported ‘gold standard’. This overestimation may have occurred in the studies of Connell \textit{et al.}\(^{(18)}\).

Although most studies assert that the psychometric properties have been confirmed (corroborated), this
should be regarded with some caution. As pointed out above, some detailed research is still needed before a clear-cut consensus may be reached. A solid choice for a particular instrument demands guidance from in-depth sequential analyses; something that, as seen, is yet to be achieved concerning most of the measurement tools addressing food insecurity. Although there are a few clear frontrunners in this dispute, perhaps much could be gained from more and better evidence regarding the other instruments, as well. The greater the choice, the better the decisions that may follow.

The interpretation of the present study’s results requires perspective in the light of its limitations. Despite our effort to meet with rigour the inclusion criteria established for the review, missing out some articles may not be ruled out since the literature is bound to hold publications with positive results. However, in an attempt to minimize such a publication bias, we sought to scrutinize carefully the reference sections of all psychometric and review articles in order to locate those studies missed out by the employed algorithm-based search strategy. Likewise, language bias may not be ruled out since studies with positive and interesting results are more likely to be published in English, while those with negative or non-significant results tend to be published in other languages. In order to lessen this possible language bias, we opted to broaden the linguistic scope beyond English by also including studies in Spanish and Portuguese.

The methodological option to restrict the review to peer-reviewed scientific journals published in PUBMED, LILACS and/or SciELO – and thus disregarding research reports, dissertations, theses and other sources not published in indexed journals – certainly narrowed the scope of the study. An ancillary search identified several government sources addressing some of the scales covered in this review (e.g. Brazil, Canada, Colombia, USA and Iran). Surely, a step forward would be to expand the approach by accounting for information arising from a wider institutional literature, not only by assessing hard evidence found in published media, but also through personal contacts with researchers and institutions proposing the scales. On a positive note, however, the current synthesis presents a favourable picture not withstanding the remaining gaps. As evolved up to now, some measurement tools (e.g. the CFSM/HFSSM and its variants and offshoots) are already apt in assisting researchers and decision makers in evaluating food insecurity.

Conclusion

The present study sought to provide detailed information about the extant experience-based HFI instruments used in most studies worldwide. From the twenty-four measurement tools identified, the CFSM/HFSSM, HFSSM-6SF and SPHFSS were the most frequently used and evaluated, holding the largest number of psychometric and applied studies conducted in different socio-economic and cultural contexts. Still, according to the peer-reviewed literature used here, these instruments and the others above of all would gain from further psychometric evaluations so that an array of not yet explored properties may be addressed. Despite these shortcomings, overall, one may conclude that current knowledge about the quality of the tools available to measure food insecurity is moving forward.

To reiterate, according to the state-of-the-art shown in the current evaluation, the CFSM/HFSSM and its linguistic variants may be recommended without much hesitation. However, as also pointed out earlier, few psychometric studies were found for most of the other scales. Initiatives intended to raise awareness about the need for psychometric assessments of existing scales would thus be encouraging. For instance, an interesting endeavour would be to establish a database repository encompassing all results from psychometric and utilization studies so that interested researchers could carry out their own psychometric analyses and/or use the scales for different purposes. Also of help would be to have any evidence so far disclosed in institutional and/or government reports, additionally published in peer-reviewed journals so that the findings and discussions on the different methods of food insecurity classification could be fully accessed and used by researchers in the field worldwide.

In concluding, we hope the present paper may serve as an incentive for further studies, given the importance of using reliable and valid instruments, not only in epidemiological research, but also as screening and decision-making tools to guide actions to address food insecurity at local and national levels.

Acknowledgements

Financial support: E.S.M. was partially supported by the Brazilian National Research Council (CNPq; grant number 162030/2012-6). C.L.d.M. was partially supported by the CNPq (grant number 302851/2008-9). M.E.R. was partially supported by the CNPq (grant numbers 306909/2006-5 and 301221/2009-0). The CNPq had no role in the design, analysis or writing of this article. Conflict of interest: None. Authorship: E.S.M. collaborated in designing the study, undertook the literature review, assisted in analysing the data and wrote the final draft of the manuscript. C.L.d.M. and M.E.R. designed the study, supervised the data collection process, undertook the analysis and collaborated in writing the final draft of the manuscript. M.M.L.A. collaborated in designing the study, undertook the literature review and assisted in analysing the data. R.S.-C. collaborated in designing the study and supervised the data collection process. Ethics of human subject participation: Ethical approval was not required for this study.
References


