Evaluation of quality indicators for management of the National School Feeding Program in Brazil: a systematic review

Avaliação de indicadores de qualidade da gestão do Programa Nacional de Alimentação Escolar no Brasil: uma revisão sistemática


Palavras-chave Alimentação escolar, Avaliação de programas, Indicador

Abstract This article aims to identify studies that developed quality indicator for the management of the National School Feeding Program (PNAE, in Brazilian context) and to critically appraise the properties of their instruments. Systematic review using Scopus, Lilacs, Pubmed and Web of Science for data collection. The search was limited to studies between 2009 and 2019. The search strategy included search terms related to school feeding, program evaluation, and indicator. The indicators were evaluated using the Appraisal of Indicators through Research and Evaluation instrument. The search identified 1,355 studies, of which 14 were potentially relevant records and 10 met the inclusion criteria. Most studies used a literature review with consensus techniques in the development of the instrument and a framework format to evaluate the PNAE. None of them presented evidence of validity of the instrument. The highest level was achieved on the domain ‘Purpose, relevance and organizational context’, followed by ‘Stakeholder involvement’, ‘Additional evidence, formulation and usage’, and ‘Scientific evidence’. This review found gaps in the methodology of studies that had developed quality indicators for the management of PNAE. Future development of these instruments should include validity evidence.

Key words School feeding, Program evaluation, Indicator
Introduction

The Brazilian Constitution of 1988 recognized the right of students to a diet provided by the public network and ensured this universal service for students with a national feeding program. The Brazilian public intervention to provide meals to students through the school system started in 1954 when the National School Feeding Program (Programa Nacional de Alimentação Escolar - PNAE - in Brazilian Portuguese) were first implemented. This is a public policy of greater longevity in Brazil in the area of food and nutritional security and is considered one of the few programs in the world to be universal and free.

PNAE is set up by Federal Law nº 11.947 which regulates the provision of school meals. Since 2009, it has been required that 30.0% of the food budget of the PNAE be used to purchase foods directly from family farms. This measure was implemented with the aim of meeting the nutritional needs of students while at school, contributing to the growth, development, learning, and academic achievement of students, and promoting the formation of healthy eating habits.

The PNAE is a model from which other countries can draw important lessons. Multisectoral food and nutrition security strategy developed in Brazil prioritized the expansion of school feeding and brought significant changes in the design and implementation of this Program. Moreover, Brazil has requested to cooperate internationally, in partnership with the Food and Agriculture Organization of the United Nations and World Food Program, for development of others School Feeding Programs. In this sense, Brazil shared experiences and knowledge to other regions around the world, such as Latin America, Caribbean, and Africa. These opportunities provided changes in dietary habits through food and nutrition education actions and the incorporation of fresh and healthy food into schools. The Brazilian program is also exemplary for its reach, for instance, in 2018, it served 40.5 million public school students with a budget of 4 billion Brazilian real ($US 1 billion).

The existence of a strong legal framework with operational regulations supports consistent, high-quality service delivery. Successful management of the PNAE depends on a network of relationships involving professionals from different disciplines, such as education, the economic sector, family farming, civil society, and all levels of government (municipal, state, and federal). Decisions should be made through intersectoral collaboration and all actors must offer a local support network to allow efficient PNAE management.

Evaluating this program is the key to ensuring and improving the quality of its managers’ decision-making to optimize public health care resources. Quality in health services must permeate organizational policies and goals, based on the assumptions of safety and the satisfaction of users and professionals. In this sense, quality indicators can be used to assess quality improvement.

The term indicator is defined as ‘a quantitative measure that can be used to monitor and evaluate the quality of care provided to the user and the activities of the services’. The indicator is not a direct measure of quality, but rather a flag that identifies or directs attention to specific issues and needs periodic review. Indicators can be associated with the structure, process, and outcomes of healthcare. ‘Structure’ refers to the attributes of settings in which care occurs; ‘process’ expresses what is actually done in giving and receiving health care; and ‘outcome’ assesses the effects of care on the health status of the population.

Some authors have developed methodologies to evaluate the PNAE, but with specific proposals in few cities. Although the PNAE has previously been assessed for certain aspects in some Brazilian municipalities, no publications were found in the literature on the evaluation of the national management of this Program.

However, as the 2009 framework does not include guidance and indicators for results measurement, indicators for all existing processes and indicators of effectiveness additional researchers needed to be developed. Therefore, this systematic review aimed to identify quality indicator instruments to evaluate the management of PNAE and critically appraise their properties.

Methods

The protocol for this systematic review has been published in PROSPERO (the International Prospective Register of Systematic Reviews), and is available at: <http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42019111796>. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for reporting systematic reviews were used to undertake the present review.

Search strategy

A comprehensive literature search was performed in the Web of Science, PubMed/Medline, Scopus, and Lilacs databases to identify relevant
studies published between January 2009 and December 2019. The starting year 2009 was chosen because PNAE’s Law 11,947, which is on the food supply given to the students of public schools, was enacted that year. The search strategy included the use of the Health Sciences Descriptors (DeCS) and Medical Subject Headings (MeSH) for school feeding and quality indicator studies. The descriptors used were “school feeding” AND “program evaluation” OR “indicator”. The search strategy included other keywords about school feeding subjects: “school meal” OR “school food”. The full search strategy for all databases can be found in Chart 1. A grey literature search was conducted in Google Scholar (Google, Inc., Mountain View, CA, USA), using the search term “school feeding” AND “program evaluation”. Duplicate studies were eliminated.

Study selection

The present review was restricted to studies that (a) had been published in English, Portuguese, or Spanish; (b) developed an original instrument; (c) described a literature search to develop the quality indicator instrument; (d) were on instruments target the Brazilian National School Feeding Program, defined in Law no. 11,947 of 2009 (any food offered in the school environment, regardless of its origin, during the school term); (e) described only students of Brazilian public schools; and (f) described at least one type of quality indicator according to Donabedian’s conceptual framework of structure, process, and outcome.

Studies were excluded if they studied instruments targeted for school feeding programs from other countries or private school feeding services, did not define the type of services provided, or did not report the instrument.

All titles and abstracts were independently screened and selected by two authors (D.B. and T.S.S.S.). The full-text version of each article was obtained and reviewed to determine whether the article met the eligibility criteria. Disagreements were resolved by a third reviewer (T.M.L.). In addition, all the references cited in the included articles were reviewed to identify any studies that might have been missed.

Data extraction and analysis

For each included study, extracted information consisted of the year of publication, state, proposal of indicator, type of indicator according to Donabedian, format of the instrument, target public, the instrument domains, number of items of the instrument, instrument development, application of the instrument, and the instrument validation properties. Two reviewers (D.B. and T.S.S.S.) independently completed the data extraction, using a preformatted spreadsheet in Microsoft Excel version 2013. Disagreements were resolved by a third reviewer (T.M.L.).

Quality assessment

The quality assessment of the indicators was determined using the AIRE (Appraisal of Indicators through Research and Evaluation)
The AIRE instrument is a valid and reliable instrument specifically designed to appraise the quality of indicators\textsuperscript{18}. The instrument was derived from the AGREE (Appraisal of Guidelines Through Research and Evaluation) instrument\textsuperscript{19}, a widely used standard for assessing the methodological quality of practice guidelines. The AIRE instrument has been used previously in several systematic reviews on the quality of indicators\textsuperscript{20-23}, and in studies developing quality indicators\textsuperscript{24,25} in other fields of knowledge.

The AIRE instrument consists of 20 items addressing four domains: ‘Purpose, relevance, and organizational context’; ‘Stakeholder involvement’; ‘Scientific evidence’; and ‘Additional evidence, formulation, and usage’ (Chart 2). These four domains reflecting the methodological quality were used to address the research objectives. Each item presents a statement on the quality of indicators and is scored on a 4-point scale (1 ‘totally disagree or no information provided’ to 4 ‘strongly agree’). Two reviewers (D.B. and T.S.S.S.) independently evaluated each study applying this scale for all items of the AIRE instrument. At this time, disagreements in the evaluation are accepted and a third revision is not demand.

The scores for each of the four categories were calculated by summing the individual authors’ scores for the items in a category and standardizing this total as a percentage of the maximum possible score for that category. The maximum possible score for a category was calculated by multiplying the maximum score per item (score of 4) by the number of items in that category and the number of evaluators (two). The minimum possible score was calculated at the same time by using the minimum score per item (score of 1). A standardized domain score was calculated according to the instrument’s guidelines following the formula: \( \frac{\text{total obtained score} - \text{minimum possible score}}{\text{maximum possible score} - \text{minimum possible score}} \times 100\% \).

Source: Based on Koning et al.\textsuperscript{17}.
Example:
If 2 researchers give the following scores for Domain 1:

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>35</td>
</tr>
</tbody>
</table>

Maximum possible score = 4 (strongly agree) x 5 (items) x 2 (researchers) = 40
Minimum possible score = 1 (strongly disagree) x 5 (items) x 2 (researchers) = 10

The scaled domain score will be:
\[
\frac{\text{Total obtained score} - \text{Minimum possible score}}{\text{Maximum possible score} - \text{Minimum possible score}} \times 100\%
\]

\[
\frac{35 - 10 \times 100}{40 - 10} = \frac{25 \times 100}{30} = 83.33\%
\]

Figure 1. Example of the calculation procedure for Domain 1 of the Appraisal of Indicators through Research and Evaluation (AIRE) instrument.

Source: Author’s elaboration.

Results

Search results

A total of 1,355 unique records were identified from the databases. After reviewing the titles and abstracts, 14 articles were selected for full-text examination. Of these, eight studies met the inclusion criteria and were included in the present review. Reference tracking of the articles identified two additional eligible studies. As a result, a total of ten studies were included in the present review. A flowchart of the selection process of the literature search and reasons for exclusion is shown in Figure 2. A list of excluded studies is shown in Chart 3.

Characteristics of the studies

The characteristics of the ten included studies are shown in Chart 4. The studies were carried out in different Brazilian states: Bahia (four studies), Santa Catarina (three studies), São Paulo (two studies), and Rio de Janeiro (one study). All studies were published in Portuguese. The purposes of the indicators were to evaluate the effects of Brazil’s school feeding law on the quality of the menus (two studies), and to evaluate management of the PNAE (eight studies).

Most studies assessed only two types of indicators: structure and process (three studies), and process and outcome (three studies). The remaining four studies assessed all indicators proposed by Donabedian: structure, process, and outcome. Most of the studies used frameworks for the indicators (seven studies), and three studies used the questionnaire format.

The studies included actors from different segments: manager, nutritionist, school manager, school cook, student, and members of the Municipal School Nutrition Councils. Nevertheless, four studies included only the nutritionist, and another study included only the PNAE’s manager. One study did not describe the target population.

The number of instrument items ranged from 8 to 88. The number of domains ranged from 2 to 6. Most studies used similar domains such as: management, participation and social control, operationalization, promotion of healthy food, and material and human resources; however, each study used different items to measure their domains. Five studies used literature review combined with other techniques, such as an expert’s panel, workshops, and qualitative interviews with field visits to develop the instrument. However, five studies used only the literature review method. Almost all studies applied the instrument. Nevertheless, none of the studies presented evidence for the validity of the instrument. Although five studies discussed consensus techniques, the authors did not calculate any index of content validity.
Figure 2. Study selection flowchart through literature search.

Abbreviation: LILACS – Latin American and Caribbean Literature in Health Sciences

Chart 3. List of excluded studies.

<table>
<thead>
<tr>
<th>Reason for exclusion</th>
<th>Author, Year</th>
<th>Title</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not developed instrument</td>
<td>Gabriel et al., 2014</td>
<td>Evaluation of municipal management of the Brazilian National School Nutrition Program in the largest cities of Santa Catarina State, Brazil</td>
<td><em>Cad Saúde Pública</em> 2014; 30(9):2017-2023.</td>
</tr>
<tr>
<td>Other service</td>
<td>Constanty et al., 2016</td>
<td>National School Feeding Programme (PNAE) and Sustainability: The Case of Marechal Candido Rondon County</td>
<td><em>Desenvolv. Meio Ambiente</em> 2016; 36: 371-392.</td>
</tr>
</tbody>
</table>

Source: Author’s elaboration.
Chart 4. Descriptive data and characteristics of the included studies.

<table>
<thead>
<tr>
<th>Author, Year, Country, State</th>
<th>Proposal of indicator</th>
<th>Type of indicator(^a)</th>
<th>Format; Target public</th>
<th>Domains evaluated</th>
<th>Number of items</th>
<th>Development</th>
<th>Application</th>
<th>Validation properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva, 2009(^b) Brazil, Bahia</td>
<td>Evaluate management of the PNAE</td>
<td>Process and Outcome</td>
<td>Framework; PNAE's manager, nutritionist, School manager, students, school cook, members of CAE</td>
<td>Management; Participation and social control, Promotion of healthy food</td>
<td>29</td>
<td>Literature review; Qualitative interview with field visit; Workshop; experts' panel consensus techniques</td>
<td>5 schools</td>
<td>Content validity; expert panel, without reported content validity index</td>
</tr>
<tr>
<td>Carvalho, 2012(^c) Brazil, Rio de Janeiro</td>
<td>Evaluate management of the PNAE</td>
<td>Structure, Process, and Outcome</td>
<td>Questionnaire; PNAE's manager, nutritionist, School manager, students, school cook, members of CAE</td>
<td>Management; Participation and social control, Promotion of healthy food</td>
<td>88</td>
<td>Literature review; Conceptual map; Workshop; Consensus techniques</td>
<td>14 schools</td>
<td>Content validity; workshop, without reported content validity index</td>
</tr>
<tr>
<td>Soares, 2012(^d) Brazil, Bahia</td>
<td>Evaluate management of the PNAE</td>
<td>Structure, Process, and Outcome</td>
<td>Questionnaire; PNAE's manager, nutritionist, School manager, students, school cook, members of CAE</td>
<td>Management; Operationalization</td>
<td>20</td>
<td>Literature review</td>
<td>1 school</td>
<td>Not reported</td>
</tr>
<tr>
<td>Calasans, 2013(^e) Brazil, Bahia</td>
<td>Evaluate management of the PNAE</td>
<td>Process and Outcome</td>
<td>Framework; Not reported</td>
<td>Management; Participation and social control, Promotion of healthy food</td>
<td>35</td>
<td>Literature review</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Gabriel, 2013(^f); 2014(^g) Brazil, Santa Catarina</td>
<td>Evaluate management of the PNAE</td>
<td>Structure and Process</td>
<td>Framework; PNAE's Nutricionist</td>
<td>Participation and social control; Operationalization; Promotion of healthy food; Material and human resources</td>
<td>22 (45 sub items)</td>
<td>Literature review; Logical theoretical model; 2-round Delphi consensus techniques</td>
<td>1 county's Pilot test and counties</td>
<td>Content validity; expert panel, without reported content validity index and field test</td>
</tr>
<tr>
<td>Calasans, 2014(^h) Brazil, Bahia</td>
<td>Evaluate management of the PNAE</td>
<td>Process and Outcome</td>
<td>Framework; PNAE's manager, nutritionist, School manager, students, school cook, members of CAE</td>
<td>Management; Participation and social control, Promotion of healthy food</td>
<td>35</td>
<td>Literature review</td>
<td>3 schools</td>
<td>Not reported</td>
</tr>
<tr>
<td>Goulart, 2015(^i) Brazil, Santa Catarina</td>
<td>Evaluate management of the PNAE in elementary school</td>
<td>Structure and Process</td>
<td>Questionnaire; PNAE's manager</td>
<td>Operationalization, Promotion of healthy food; Material and human resources</td>
<td>20 (59 sub items)</td>
<td>Literature review; 2-rounds experts' panel consensus techniques</td>
<td>1 School's pilot test and 11 elementary schools</td>
<td>Content validity; expert panel, without reported content validity index and field test</td>
</tr>
<tr>
<td>Bicalho, 2017(^j); 2019(^k) Brazil, São Paulo</td>
<td>Evaluate the effect of the PNAE's law on the quality of the menus</td>
<td>Structure, Process, and Outcome</td>
<td>Framework; PNAE's Nutricionist</td>
<td>Operationalization; Promotion of healthy food; Material and human resources</td>
<td>8</td>
<td>Literature review</td>
<td>38 schools</td>
<td>Not reported</td>
</tr>
</tbody>
</table>

Abbreviation: PNAE (National School Feeding Program); CAE (Municipal School Nutrition Councils). \(^a\)According to Donabedian conceptual framework; 
Source: Author's elaboration.
Quality assessment results

The results of the assessment quality of the indicators using the AIRE instrument are shown in Table 1. The methodological quality of indicators presented in these studies varied considerably.

Most of sets of indicators presented in these studies obtained the highest scores for the items ‘Purpose, relevance and organizational context’ (range 73% - 93%) and the lowest scores for the item ‘Stakeholder involvement’ (range 0% - 72%), followed by ‘Additional evidence, formulation, and usage’ (range 22% - 46%), and ‘Scientific evidence’ (range 22% - 44%). No study achieved high assessment quality scores in all four categories. Carvalho presented a set of indicators with high scores on domains ‘Purpose, relevance and organizational context’ (93%) and ‘Stakeholder involvement’ (72%). The study of Soares had the lowest scores for the set of indicators on domains ‘Scientific evidence’ (22%) and ‘Additional evidence, formulation, usage’ (22%). The majority items that scored poorly were ‘the supporting evidence has been critically appraised’, ‘a strategy for risk adjustment has been considered and described’, ‘the indicator measures what it is intended to measure (validity)’, ‘the indicator measures accurately and consistently (reliability)’, and ‘the indicator has sufficient discriminative power’.

Discussion

Summary of evidence

To our knowledge, this is the first review to identify and assess studies that have developed quality indicators for the management of the PNAE. Ten studies were found that developed instruments to measure and evaluate the PNAE. The present review has provided a comprehensive critical analysis of the study characteristics and the measurement properties of the studies.

We systematically searched the literature in five electronic reference databases and thoroughly reviewed and evaluated a vast number of articles. The selection of articles, data extraction, and quality assessment were independently conducted by two reviewers, which increases the reliability of the results. Therefore, we can be confident that the present review provides a comprehensive overview of the available indicators.

The present review highlighted relevant gaps in the quality of the instruments. No study achieved high assessment quality scores in all four domains of the AIRE instrument. Moreover, none of them developed an instrument with evidence of validity, and this limited the psychometric quality of the instruments. Likewise, no study included goals and frequency for the indicators to support decision-making of the stakeholders. Consistent with Donabedian’s framework, further research should include the development and validation of the indicators underlying the

<table>
<thead>
<tr>
<th>Studies</th>
<th>Purpose, relevance and organizational context (%)</th>
<th>Stakeholder involvement (%)</th>
<th>Scientific evidence (%)</th>
<th>Additional evidence, formulation, usage (%)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva, 2009[27]</td>
<td>87</td>
<td>61</td>
<td>44</td>
<td>44</td>
<td>44-87</td>
</tr>
<tr>
<td>Carvalho, 201[23]</td>
<td>93</td>
<td>72</td>
<td>39</td>
<td>46</td>
<td>39-93</td>
</tr>
<tr>
<td>Soares, 201[24]</td>
<td>77</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22-77</td>
</tr>
<tr>
<td>Calasans, 201[28]</td>
<td>73</td>
<td>0</td>
<td>33</td>
<td>33</td>
<td>0-73</td>
</tr>
<tr>
<td>Gabriel, 2013[30], 2014[31]</td>
<td>87</td>
<td>67</td>
<td>33</td>
<td>35</td>
<td>33-87</td>
</tr>
<tr>
<td>Calasans, 2014[32]</td>
<td>73</td>
<td>0</td>
<td>33</td>
<td>46</td>
<td>0-73</td>
</tr>
<tr>
<td>Goulart, 2015[33]</td>
<td>90</td>
<td>56</td>
<td>44</td>
<td>37</td>
<td>37-90</td>
</tr>
<tr>
<td>Bicalho, 2016[30], 2019[34]</td>
<td>77</td>
<td>22</td>
<td>28</td>
<td>39</td>
<td>22-77</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td><strong>73-93</strong></td>
<td><strong>0-72</strong></td>
<td><strong>22-44</strong></td>
<td><strong>22-46</strong></td>
<td><strong>0-93</strong></td>
</tr>
</tbody>
</table>

Source: Author’s elaboration.
structures, processes, and outcomes. This approach may provide a comprehensive evaluation of the quality of the management of the PNAE.

**General view of the studies**

All studies developed the instrument in Portuguese and were carried out in Brazil, but only to states in the northeast, south, and southeast of the country. Considering that PNAE is a national program, it was not necessary to have a specific instrument for each state. But would be important to test characteristics of reliability and validity or adapt the instrument in other regions. Adapting existing instruments for each setting is necessary to guarantee the instruments’ linguistic and cultural appropriateness.

Quality indicators can be categorized according to structure, process, and outcome. However, most studies assessed two types of indicators: structure and process or process and outcome. There is a consensus that the three types of indicators complement each other and can assist in obtaining service with better quality. An indicator that evaluates structure can assist, under favourable or unfavourable conditions, in the achievement of the objectives of the PNAE in the other dimensions. Likewise, the process indicator assesses what the provider did for the PNAE and how well it was done. However, structure is not a necessary condition for the processes to occur. The structure of the program will fulfill its purposes if the applied processes are appropriate. In addition, both types of indicators (structure and process) will only reach their ultimate goals with the achievement of good outcomes.

The number of indicators in the instruments ranged from 8 to 88. Three studies divided indicators in sub-item’s sets. For example, the indicator ‘Adequacy of school cooks team’ is considered to assess the sub-items ‘number of school meals/school cooks ratio’ and ‘Extra tasks for school cooks’. However, these subitems should not be considered indicators. According to Tanaka et al., the indicator is a numeric variable that can be an absolute number, a two-events ratio, or a quality event. One study classified some items of the instrument, such as ‘Has the school kitchen a Standard Operating Procedures?’ and ‘Were the school cooks trained to use the Standard Operating Procedures?’ as process indicators. Nevertheless, according to Santos et al., based on Donabedian’s conceptual framework, these items are considered structure indicators as they refer to the actual PNAE law and the human resources training. Another important finding of the present review is that only two studies developed a short set including eight indicators. The literature recommends the choice of three to five indicators for their importance, synthesis capability, and ease of data collection because too many indicators may cause operational difficulty.

One important aspect of the development of quality indicators is the enrollment of stakeholders with different perspectives on quality management. The combination of PNAE’s legal and institutional mechanisms for the participation of civil society and the partnership of different government sectors set the conditions for the promotion of intersectorality. It is recommended to include the perspectives of all potential end users including the service recipient, their families, health professionals, and managers. However, the included studies mainly considered the points of view of the manager and nutritionists of the PNAE. Some studies included the perspectives of the students, school managers, school cooks, or members of Municipal School Nutrition Councils in the development of the quality indicators. Therefore, specific challenges in measuring results are related to one of the major strengths of the programme: its integrated and multisectoral approach. The management of the PNAE depends on a network of relationships involving different areas: education, the economic sector, family farming, civil society, and all levels of government.

Quality indicators can be developed using non-systematic or systematic evidence, combined or not with expert opinion methods. Five studies used the literature review method combined with consensus techniques. Consensus techniques are group facilitation techniques that explore the level of consensus among a group of experts while synthesizing opinions. Group judgements are preferable to individual judgments, which are prone to personal bias.

Five studies used only the literature review method to develop the instrument. Many areas of healthcare that have a limited or methodologically weak evidence base, especially within the evaluation of public policy, require other evidence, including expert opinion. Systematic research methods that also involve consensus are the best methods for developing quality indicators in many areas of health care where the scientific evidence base is limited. In the development of indicators, the use of expert opinion is necessary in order to obtain more validity evidence. Therefore, the instruments from the five studies that did not perform a literature review are not suitable for further application.
Finally, three studies\textsuperscript{16,29-31} in the present review were from dissertations. We included these studies because they report important additional information regarding the development process of indicators.

**Assessment of quality indicators**

The set of indicators presented in the studies varied in the methodological quality and the information available about development process. Some studies described set of indicators in detail, with a clear definition of numerators, denominators, and/or performance standards as well as the development process, whereas other studies presented set of indicators without more detailed information about methodology process.

The sets of indicators presented by Carvalho\textsuperscript{33} study had the highest methodological scores according the AIRE instrument. The development process for these sets was described more precisely and elaborately. Moreover, no study obtained high assessment quality scores in all four categories.

Overall, in terms of assessment quality, most of the studies\textsuperscript{26,27,31-33} reached a high-quality level on the domains and ‘Purpose, relevance, and organizational context’ and ‘Stakeholder involvement’. However, the studies did not describe satisfactorily the domains ‘Scientific evidence’ and ‘Additional evidence, formulation, and usage’. Information about formal endorsement of the indicators was barely available in the studies. They may have put less emphasis on this type of information, resulting in lower quality scores on these aspects. We have tried to resolve this by incorporating as much information as possible about the indicator sets when evaluating their quality.

The characteristics of the quality indicators of the studies in the present review varied widely. The addressed content, the organizational context, and the criteria for interpretation were described in detail in all studies. On the other hand, no information was available about reliability and validity in the studies. Two studies\textsuperscript{31,32} discussed some aspect of validity (e.g. the cut-off of the experts consensus) but did not describe a process of validating the instrument. In addition, no studies considered a sufficient discriminative power; no strategy for risk adjustment was considered and described. Characteristics of reliability and validity are very important to develop or adapt research-measuring instruments\textsuperscript{44-46}.

Indicator sets without a robust development process (i.e. those sets scoring poorly in a methodological assessment) can still be considered as potential quality indicators. They can be used in other quality assessment initiatives, on the condition that they will be further studied\textsuperscript{30}. In the literature concerning quality indicators, there are some disagreements on the types of indicators that are most suitable for the assessment of quality. Therefore, publication of the methodological characteristics of quality indicator sets, including an extensive description of the development process, is recommended\textsuperscript{37}.

**Limitations**

Although this systematic review makes a significant contribution to the quality of public-school feeding policies literature, some limitations must be acknowledged. As demonstrated in this comprehensive review, few studies developed an instrument to assess the quality of local management of the PNAE. It is possible that some studies were missed because they were not indexed in the databases searched or were published by institutions, foundations, or societies. This setting was minimized to track down relevant grey literature by manually checking the reference lists during the full-text screening and using Google’s internet search.

Finally, the assessment quality of the indicator sets of the studies included in the present review might have been underestimated in some aspects. Following the instructions from the AIRE instrument, the lowest score was assigned to an item when no information was provided in the article or dissertation.

**Conclusions**

The PNAE monitoring and evaluation mechanisms in the studies presented in this review focused on implementation. Few studies were identified proposing quality indicator instruments to evaluate the management of the PNAE. The literature review combined with consensus techniques was used for instrument development in some studies although the authors did not perform the reliability and validity of the indicators. The highest level of the quality assessment according AIRE instrument was achieved on the domains ‘Purpose, relevance and organizational context’ and ‘Stakeholder involvement’, however the studies did not describe satisfactorily the domains and ‘Additional evidence, formulation, and usage’ and ‘Scientific evidence’.

This study was carried out by understanding that an evaluation is considered positive when it is possible to contribute for identifying problems
and proposing solutions. The quality indicators could contribute to reducing the gap in the area of evaluation and improvement of the implementation of this Program. However, one challenge identified in this study is inadequate research to support the school feeding policy. Therefore, the flaws observed in this study showed that further research concerning the development of quality indicators with rigorous methods of evidence of validity are necessary to evaluate the management of the PNAE.

Collaborations

D Bicalho, TM Lima and B Slater contributed to the planning and design of the research. Bicalho, TM Lima and TSS Santos conducted the search in the database and the screening procedure. TM Lima resolved the disagreements. Bicalho wrote the manuscript. All authors commented the drafts and read and approved the version of the final manuscript forwarded.

References


