# Food Screeners for Adolescents: A Hybrid Review Focused on Their Psychometric Properties

by

Emma Roe

A thesis presented to the University of Ottawa in partial fulfillment of the requirements for the degree of Master of Science in Nutrition and Food Biosciences

School of Nutrition Sciences

Faculty of Health Sciences

University of Ottawa

Ontario, Canada, 2026

© Emma Roe, Ottawa, Canada, 2026

# Preface

This thesis is submitted in partial fulfillment of the requirements for the Master of Science in Nutrition and Food Biosciences at the University of Ottawa.

The research presented here is based on a hybrid scoping/systematic review conducted in accordance with established methodological guidelines for evidence synthesis.

Under the supervision of Dr. Claire Tugault-Lafleur, MSc candidate Emma Roe led the conceptualization of the study, the screening and selection of articles, quality appraisal, data synthesis, interpretation of findings, and preparation of the manuscript. Ms. Roe is the first author of this thesis, with Dr. Tugault-Lafleur serving as the senior author.

Portions of the writing were refined using Microsoft Copilot (Microsoft, 2026) and Grammarly (2026), which assisted with grammar, clarity, and organization. All conceptual decisions, interpretations, and final wording were made by the author.

# Abstract

Food screeners are a low-cost, low-burden option for nutrition surveillance, but their number and variability complicate tool selection for adolescents. This review aimed to describe food screeners used among adolescents in high-income countries and evaluate their validity and reliability. A hybrid approach was used: a scoping review to identify existing screeners and a systematic review to assess psychometric quality. Screeners with ≤30 items targeting healthy adolescents aged 10–17 years were eligible. Forty-seven screeners were identified through both searches. Instruments averaged 16 items with 22 including portion size questions, and 26 assessing alignment to dietary guidelines. Twenty-six screeners had psychometric evidence, and 18 met quality criteria. Construct (convergent) validity was most frequently examined, with most estimates falling in the poor–moderate range, whereas reliability was generally stronger, with most instruments demonstrating acceptable test–retest performance. Findings highlight substantial variability and key gaps in adolescent food screener evaluation.



**Keywords:** Food Intake; Screener; Validity; Reliability; Adolescents

## Acknowledgements

I would like to acknowledge the guidance and support of my supervisor, Dr. Claire Tugault-Lafleur, and thesis committee, Dr. Isabelle Giroux and Dr. Kathryn Walton, whose expertise and constructive feedback were essential to the development of this work. I would also like to thank Manel Khan and Jacob Cochrane for their assistance with the screening process, which played an important role in shaping the evidence base for this project. I extend my appreciation to my family and friends for their continued support during my graduate studies. Their patience and encouragement sustained me throughout this process.

# Table of Contents

**Preface** ..... **ii**

**Abstract** ..... **iii**

**Acknowledgements** ..... **iv**

**List of tables, figures and appendices** ..... **vii**

**List of abbreviations and symbols** ..... **vii**

**1 Introduction** ..... **1**

**2 Literature Review** ..... **2**

*2.1 Self-report dietary assessment methods* ..... *2*

2.1.1 Food records ..... 3

2.1.2 24-hour dietary recalls ..... 4

2.1.3 Food frequency questionnaires ..... 5

*2.2 Considerations when measuring diet in young people* ..... *7*

2.2.1 Measuring diet in children ..... 8

2.1.2 Measuring diet in adolescents ..... 9

*2.3 Psychometric properties of questionnaires* ..... *10*

2.3.1 Validity ..... 11

2.3.2 Reliability ..... 12

**3 Rationale and study objectives** ..... **13**

**4 Methodology** ..... **14**

*4.1 Study design* ..... *14*

*4.2 Search strategies* ..... *14*

*4.3 Eligibility criteria and study selection* ..... *15*

*4.4 Data extraction* ..... *16*

*4.5 Study quality assessment* ..... *17*

*4.6 Data synthesis and analysis* ..... *17*

**5 Results ..... 22**

*5.1 Study selection ..... 22*

*5.2 Study characteristics ..... 24*

*5.3 Food screener characteristics ..... 24*

*5.4 Content validity ..... 31*

*5.5 Construct and criterion validity ..... 32*

*5.6 Reliability ..... 38*

**6 Discussion ..... 42**

*6.1 Overview of the main findings ..... 42*

*6.2 Strengths and Limitations ..... 46*

*6.3 Future research directions ..... 48*

**7 Conclusion ..... 49**

**8 References ..... 50**

**9 Appendices ..... 61**

*9.1 Appendix 1: Search strategies ..... 61*

*9.2 Appendix 2: Inclusion and exclusion criteria ..... 63*

*9.3 Appendix 3: Table of quality assessment results determined from the COSMIN risk of bias checklist for food screeners designed for use with adolescents aged 10-17 years (n=26 screeners) 64*

*9.4 Appendix 4: Countries of origin of food screeners designed to measure food intake among adolescents aged 10-17 years (n = 47 screeners) ..... 68*

## List of tables, figures and appendices

**Figure 1.** Preferred reporting items for systematic reviews and meta-analyses flow diagram of the study selection process for the hybrid review
**Table 1.** Definitions of common indicators used in methodological studies evaluating the psychometric properties of food screeners, along with their threshold criteria as appropriate.
**Table 2.** Characteristics of self-administered food screeners developed for use among adolescents aged 10-17 years old (n = 47 screeners)
**Table 3.** Findings from methodological studies assessing the validity of food screeners designed for use with adolescents aged 10-17 years (n=13 studies assessing 13 screeners)
**Table 4.** Findings from methodological studies testing the reliability of food screeners designed for use among adolescents aged 10-17 years (n = 11 studies assessing 11 screeners)
**Appendix 1**. Search strategies
**Appendix 2.** Inclusion and exclusion criteria
**Appendix 3**. Table of quality assessment results determined from the COSMIN risk of bias checklist for food screeners designed for use with adolescents aged 10-17 years (n=26 screeners)
**Appendix 4.** Countries of origin of food screeners designed to measure food intake among adolescents aged 10-17 years (n = 47 screeners)

## List of abbreviations and symbols

AMPA: Automated Multiple Pass Method
ASA-24: (Automated Self-Administered 24-hour recall)
FFQ: Food Frequency Questionnaire
COSMIN: Consensus-based Standards for the selection of health Measurement Instruments
PROMs: Patient-Reported Outcome Measures

# 1 Introduction

Adolescence is a pivotal developmental period in which dietary behaviours are formed and often persist into adulthood, shaping long-term health and well-being. It is also characterized by rapid physical, cognitive, and psychosocial changes, which make adolescents particularly sensitive to environmental influences such as food marketing, peer norms, household food security, and school food environments. Consequently, the dietary patterns established during this stage can have far-reaching implications for chronic disease risk, mental health, and overall quality of life[1].

Accurately assessing adolescent dietary intake is critical for public health surveillance, program evaluation, and designing effective nutrition interventions. Traditional dietary assessment methods, such as 24-hour recalls and food records, are often burdensome, costly, and challenging for adolescents to complete. As a result, brief food frequency questionnaires (FFQ), or food screeners, have become a key tool in nutrition surveillance, offering a rapid, low-cost way to estimate usual dietary intake while minimizing participant burden[2].

Despite their value, the landscape of food screeners for adolescents is highly variable. Available tools differ substantially in purpose, structure, target dietary components, and psychometric quality. This variability poses challenges for researchers, who must identify the instrument that best fits their research objectives, population characteristics, and methodological requirements. Establishing and maintaining the validity and reliability of these tools is essential for generating robust evidence and accurately assessing the effectiveness of nutrition interventions.

These complexities underscore the need for a clear synthesis of available food screeners and their measurement properties to guide informed selection and application in adolescent nutrition research.

# 2 Literature Review

## 2.1 Self-report dietary assessment methods

Nutrient and food intake data are foundational to population health surveillance systems and inform the design of impactful health promotion initiatives [3]. However, measuring dietary intake is inherently complex due to the multifaceted nature of eating behaviours, the limitations of self-reporting tools and the variability in food composition databases; the influence of socio-cultural, environmental, and temporal factors also affect both individual and population-level dietary patterns[4].

Dietary assessment methods include self-reported tools such as 24-hour recalls, food frequency questionnaires, and food records, as well as objective approaches like direct observation and biomarker-based techniques [5]. Dietary assessment methods are subject to both random and systematic errors, which can significantly impact the accuracy and interpretation of dietary intake data [2]. Random errors arise from unpredictable fluctuations in reporting or day-to-day dietary intake variation that reduce precision without introducing biased results. Conversely, systematic error, also known as bias, occurs when measurements consistently deviate from the true value, typically due to factors like recall bias or social desirability, resulting in under- or over-reporting of intake[5,6]. While objective techniques such as doubly labelled water or urine tests can accurately estimate energy expenditure or specific nutrient intake, they do not capture the

types or contexts of foods consumed. Although observational designs like the duplicate diet method (where a researcher will collect an exact copy of their participants meal), can provide valuable context to objective measures, they are burdensome for participants and impose substantial costs on researchers [7]. Therefore, despite their inherent limitations, self-reported dietary instruments remain essential for understanding eating behaviours [8].

Dietary assessment instruments have both strengths and weaknesses, and the choice of instrument should align with the study's objectives. The next section summarizes frequently used self-reported dietary intake instruments with a focus on measurement error.

### 2.1.1 Food records

Food records require participants to document all food and beverage consumption in real time over a short period, typically two to three days. Food records can be weighed (where a participant is asked to weigh all the foods consumed) or estimated (where the portion of food consumed is estimated). Food records allow for detailed tracking of food intake and contextual factors, such as where food is purchased, prepared and eaten, and food characteristics (i.e. fresh, frozen, canned, etc.)[5].

One advantage of food records is that, because entries are made in real time, the cognitive burden is reduced compared to methods that involve recalling foods consumed the previous day or estimating usual intake over a longer period (e.g., fruit and vegetable intake over the past month), however, people can forget or neglect to record all of what they have consumed (e.g. night snack). As well, food records are designed to capture total food intake over a limited timeframe and may not reflect long-term dietary habits. They can also be burdensome to participants, and their accuracy can vary depending on factors

like participant motivation, capacities to estimate portion sizes and levels of health and food literacy [5]. Due to this high burden on participants, food records tend to have a comparatively lower quality of information provided compared to 24-hour recalls [5,9]. Finally, this instrument is prone to reactivity bias, which is when a person alters their own food intake because they know what they eat is being recorded[9].

### 2.1.2 24-hour dietary recalls

The 24-hour dietary recall method requires participants to remember and list all foods and beverages consumed the previous day (usually from midnight to midnight). A feature of the 24-hour dietary recall is the use of specific probes by an interviewer or an online program to capture specific details about the foods consumed [10]. Some 24-hour dietary recalls use the Automated Multiple-Pass Method (AMPM) to reduce bias in food intake estimates. This method includes a 5-step process where the participants are probed to ensure completeness of the recall. Participants will be prompted to provide clarification on any omitted foods, specify the timing of their meals, and offer additional details that may have been overlooked. Some example probes include "did you have anything to drink with this meal?" or "Did you put anything on your toast?" [11,12].

To get an accurate measure of an individual's usual intake, multiple days of 24-hour dietary recalls are necessary, and the number of days needed may vary widely depending on the food or nutrients of interest. For example, vitamin A intake varies significantly each day; assessing at least 30 days of intake is recommended to determine an individual's usual intake [9]. By conducting these recalls on different non-consecutive days, including both a weekend day and a weekday (or days for which the types of food and/or the eating pattern tend to differ), researchers can account for variations in eating

habits that might occur from day to day. This approach helps to get a more reliable and comprehensive picture of a person's typical diet [5]. There are several ways to administer 24-hour dietary recalls: phone interviews, in-person interviews or online (e.g. Automated Self-Administered 24-hour recall (ASA-24) [13]. Before technological advancements in web-based dietary recalls, this method was considered labour-intensive in large population-based research settings. However, technological advancements have made 24-hour dietary recalls a feasible tool. One study shows that online 24-hour dietary recalls are preferred by most participants and contribute to lower participant attrition than other self-report dietary assessments [14]. The ASA-24 was modeled to include the AMPM method to ensure it is performing as well as an in-person or phone interview [15].

Similar to food records, the 24-hour dietary recall can provide detailed information about foods consumed such as type, brand, amounts and preparation method (e.g. grilled vs. fried chicken). Additionally, it can offer a better estimate of the usual energy intake of subpopulations, such as children and adolescents, or cultural subgroups, compared to food records, because it can offer prompts that help people report their complete dietary intake [5]. However, 24-hour dietary recalls are susceptible to recall bias, often leading to under- or over-reporting of foods and nutrients. The demands associated with 24-hour dietary recalls may deter participation among certain subpopulations, particularly those with limited time, lower health literacy, or reduced digital access, thus introducing potential selection bias [2].

### 2.1.3 Food frequency questionnaires

FFQs are questionnaires designed to estimate the usual dietary intake, or intakes of a particular nutrient of interest, of individuals over a specific period of time (e.g. the

past week or the past month) [5]. FFQs are categorized based on their approach to portion size measurement: qualitative (no estimate on quantities of foods is given), semi-quantitative (portion sizes are chosen from a menu of options) and quantitative (participants are asked to recall not only frequency of foods but amounts consumed). FFQs are especially useful in epidemiological studies with larger samples, due to their low cost [5,9].

Compared to other self-report methods, FFQs place a lower burden on respondents (taking less time to complete than 24-hour dietary recalls) but are less precise. As such, FFQs are more often used to rank individuals according to specific nutrient intakes (e.g. low, medium, and high consumption of ultra-processed foods) and are not frequently used to measure energy intakes, as they give less accurate results[16]. FFQs can be limited by the inability to capture a wide range of foods consumed and need to be adapted to context (i.e. adapted to include popular cultural foods). They may also require a high level of cognitive ability, which can make these tools less accessible to certain populations [2].

Food screeners are short FFQs (e.g. include 10-30 items as opposed to long lists of foods). Like FFQs, food screeners are designed to capture an individual's usual intake, asking participants to estimate their intake over the past month or week. Like FFQs, food screeners can assess multiple components of a diet or focus on specific food groups or nutrients (e.g. dairy products, calcium). Food screeners offer a cost-effective alternative to more comprehensive methods like 24-hour dietary recalls, food records and longer FFQs [2]. Since food screeners are short, they are quick to complete, and are generally low burden for participants [9].

Scoring systems are applied to food screeners to help researchers quantify the results from the screener, enabling statistical analysis. It takes the result from each question and combines them to produce a final score. Participant scores can be compared to evaluate how individuals perform relative to one another [2]. They can also help assess dietary alignment to healthy eating guidelines [17]. This can be done by deriving scores from dietary indexes (predetermined scoring systems derived from dietary guidelines), which may be applied directly to the food screener. Alternatively, results from a screener's own scoring system can be compared to scores from the diet index using 24-hour recalls or food records [18]. This is often used as a way to validate the results from a screener's scoring system, as such with the Canadian Food Intake Screener [17].

Food screeners are susceptible to systematic error, which is harder to avoid than random error. Unlike random error, systematic error cannot be reduced by repeated attempts as it is an error due to the measurement property, and is not person-specific [2]. However, it can be mitigated using the appropriate statistical methods when analyzing data, such as regression calibration, or by applying a scoring system to the screener that is based on a less biased instrument (such as a 24-hour dietary recall) [2]. Due to their brevity, food screeners may be unable to capture information about subtypes of food or other contextual factors (e.g., timing or food preparation), often limiting their assessment to only a portion of overall dietary intake.

## 2.2 Considerations when measuring diet in young people

While self-reported dietary instruments present several limitations, their use with children (<10 years old) and adolescents (≥10 years old, but <18 years old) can introduce an additional layer of complexity. Self-report instruments rely on the respondent's ability

to accurately recall and describe their food intake, which can lead to considerable measurement error when used with younger populations.

### 2.2.1 Measuring diet in children

Research suggests that adolescents aged 10 years and older generally have the cognitive ability to report their food intake without caregiver assistance [19]. In contrast, younger children may require their parents to report their intake for them (proxy reporting) or to assist with recalling foods consumed (parent-assisted dietary recalls). A common approach to balancing parent-proxies with self-reporting is to use parental assistance, allowing children to recall their diets for foods consumed outside the home, while parents contribute information about foods eaten at home. The appropriateness of this strategy depends on children's cognitive abilities and may not be suitable for younger groups (<8 years old) [19].

Children and youth vary considerably in their ability to engage with dietary assessment tools, so pilot testing with the intended population is recommended [13]. School-aged children are often interested in dietary reporting, but their limited cognitive abilities can impact the accuracy of their self-reports[8]. Factors such as short attention spans, limited food literacy, and a poor concept of time contribute to measurement errors. Additionally, since children are not typically involved in food preparation, they may be unaware of all the ingredients in their meals [4]. These limitations highlight the importance of choosing age-appropriate dietary assessment tools and considering context when interpreting data for younger populations.

Feasibility studies of online 24-hour dietary recalls show that children's ability to recall foods consumed varies by food type, with higher accuracy for main meals and

lower accuracy for side items like vegetables [10,20]. Children's total recall scores from diet indexes derived from 24-hour recalls are generally less accurate than those of older age groups, and their recall is affected by age and weight status, with younger and overweight children more prone to misreporting [21]. Common errors include under-reporting and reporting "phantom foods" (items not actually consumed) [19]. Children may also have difficulties accurately recalling portion sizes, but visual aids can help guide their assessment [22]. Self-reported portion sizes by children and youth can provide relative information on intake but may introduce errors when used as quantitative data. In this age group, food screeners face limitations similar to those found in 24-hour recalls, mainly due to the use of parental assistance, which can introduce several types of bias.

### 2.1.2 Measuring diet in adolescents

Unlike younger children, whose diets are typically structured by caregivers, adolescents experience greater autonomy over food choices, leading to more variable and less predictable eating patterns. Some adolescents may have lower motivation to engage in dietary reporting, which can lead to rushed or incomplete responses [1]. This lack of motivation can contribute to reporting errors, particularly in methods requiring detailed recall or sustained attention, such as 24-hour recalls or food records.

Adolescence is a developmental period marked by heightened body image concerns and increased prevalence of weight management behaviours, including dieting and disordered eating [23]. These factors can introduce systematic bias in dietary reporting, as adolescents may underreport energy-dense or socially stigmatized foods and overreport foods perceived as healthy. The tendency to misreport is further exacerbated by social desirability bias [23]. Adolescents may tailor their responses to align with

perceived norms or expectations, especially when dietary assessments are conducted in school settings or involve adult interviewers [1,2].

Social media has heightened these issues, as platforms showcasing idealized body types and curated eating patterns distort perceptions of what is considered "normal" or acceptable. Adolescents who encounter this content often feel pressured to conform, which can result in intentionally misreporting their diets or omitting foods that oppose these ideals. Misreporting not only compromises dietary data accuracy but also brings attention to the psychological toll dietary assessments may have on adolescents who may experience stress, guilt, or social desirability pressures when asked to recall and report their food intake[4].

On the other hand, since adolescents generally possess higher health knowledge and food literacy compared to younger children, they may also demonstrate a greater capacity to independently report their own dietary intake. This variability complicates both recall accuracy and the ability of standardized tools to capture habitual intake [1].

### 2.3 Psychometric properties of questionnaires

For a dietary assessment tool to be used by researchers, it must first undergo an assessment of its measurement properties to ensure the tool is both valid and reliable. Measurement properties refer to several characteristics of the instruments that ensure that the instrument is performing well, for its intended purpose, with its intended population[24]. When instruments are validated, the population of interest must be considered, as their validity only applies to the group in which it was tested.

### 2.3.1 Validity

Validation means ensuring that the tool is measuring what it is intended to measure in the population of interest[8]. Without prior validation, researchers cannot determine whether the data collected is an appropriate indicator of the concept or construct it seeks to measure. There are three different types of validity relevant to self-report dietary assessment tools [24]: content, convergent and construct validity.

Content validity refers to the degree to which an instrument adequately reflects the construct of interest for a specific purpose and target population. According to the COSMIN measurement frameworks[24], content validity has three key components: 1) Relevance (are the items relevant to the construct and population?); 2) Comprehensiveness (do the items adequately cover all important aspects of the construct?); and 3) Comprehensibility (are the items understood by respondents as intended?). To confirm content validity in a food screener, researchers often conduct expert panel evaluations to assess comprehensiveness while cognitive testing with the intended population enables researchers to measure comprehensibility [25].

Criterion validity refers to the extent to which an instrument aligns with a recognized gold standard. In dietary assessment, however, there are few true gold standards, and those that do exist are difficult and expensive to obtain[26]. This substantially limits the ability to establish 'true' criterion validity. However, criterion validity can still be evaluated by comparing diet scores from the food screeners to those generated using established gold-standard methods. These include doubly labelled water (for energy intake) and 24-hour urinary tests for protein, potassium, and sodium[26]. To substitute the established gold standards, which are difficult to obtain, researchers often compare scores obtained from the food screener with those from another dietary

assessment tool, such as a 24-hour recall, and determine the agreement between the instruments.

Construct validity indicates the degree to which an assessment instrument is consistent with specific hypotheses (such as differences between subgroups, or against other instruments) of what is being measured. This includes convergent and divergent validity, which demonstrates that a food screener should align with other similar instruments, like 24-hour recalls or food records, (convergent validity) and differ from dissimilar instruments (divergent validity)[27]. This differs from criterion validity in diet assessment since this approach directly compares food groups between instruments, instead of relying on final scores[24]. Structural validity, another subtype of construct validity, refers to the extent to which questionnaire scores accurately represent the dimensionality of the construct being measured. It is typically assessed using factor analysis, which examines how items in a food screener relate to one another (i.e., whether they form the expected clusters)[28].

### 2.3.2 Reliability

Reliability refers to the degree to which the measurement is free from measurement error, or the proportion of total variance in the measurement attributed to "true" differences within the sample [27]. There are several ways in which reliability can be tested, but the method most relevant to self-report dietary assessment tools is test-retest reliability [27]. When measuring test-retest reliability, the same group will complete the instrument at multiple time points. If the instrument provides similar results between different tests, then the test instrument is deemed to have good test-retest reliability. Internal consistency can also be assessed for each screener subscale (e.g., protein foods,

dairy foods). However, this is only meaningful once structural validity has been established, because internal consistency tests assume that the screener's structure is correct. It is typically evaluated using Cronbach's alpha, which quantifies the correlation among items within each subscale. The reliability of an instrument, however, does not guarantee its validity and accuracy (i.e. an instrument could be consistently providing false results). For a tool to be considered methodologically sound, it must be both reliable and valid [24].

## 3 Rationale and study objectives

Given the extensive range of food screeners available, researchers often face the challenge of choosing instruments that best align with the objectives and context of their studies. Although many studies have evaluated individual measurement properties, uncertainty remains regarding whether all relevant psychometric domains have been comprehensively assessed across existing tools. Therefore, this knowledge synthesis project aimed to:

(1) Describe the characteristics of food screeners designed for use among adolescents aged 10-17 years in high-income countries.

(2) Review and appraise the measurement properties of food screeners.

# 4 Methodology

## 4.1 Study design

A knowledge synthesis approach was used to evaluate the measurement properties of self-administered food intake screeners designed for adolescents. A hybrid review format was adopted to accommodate the distinct yet complementary aims of the review. This hybrid approach draws from scoping review methodology, while selectively integrating systematic review elements, such as a structured risk-of-bias appraisal and evidence grading, for components where rigorous synthesis is feasible [29,30]. By doing so, the review maintains flexibility in capturing diverse study designs and outcomes, while ensuring methodological transparency and evaluative depth where appropriate. This hybrid design has been previously used in the context of food safety and nutrition [31,32].

This hybrid review was registered in Open Science Framework and the systematic review sections were registered on Prospero in December 2025. The review followed the PRISMA-ScR methodology for scoping reviews for the scoping review sections [33], and the COnsensus-based Standards for the selection of heath Measurement Instruments (COSMIN) methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) for the systematic review sections [24,34].

## 4.2 Search strategies

Two searches were carried out in December 2025. For the scoping review, a single database was used to conduct the search. The search was restricted to papers published from high-income countries according to the World Bank 2024 definitions [35]. Boolean operators were used to combine common subject headings and keywords such as “adolescent”, “youth”, “teen”, “diet screener”, “food screener”, “diet questionnaire”,

"food questionnaire", "diet survey" and "food survey". Psychometric properties were intentionally excluded from the search strategy to allow screeners that had not yet been evaluated to be included.

For the systematic review, the search strategy was adapted from the scoping review, with the addition of key terms such as "validity", "reliability" and "psychometry" to limit the search to screeners that had been validated. Three database searches (Medline, Embase and CINHAL) were carried out to identify relevant publications and was supplemented with a search in the Register of Validated Short Dietary Assessment Instruments from the National Cancer Institute [36]. A research librarian was consulted when formatting and translating search strategies across databases. The complete search strategies can be found in **Appendix 1**.

### 4.3 Eligibility criteria and study selection

The inclusion and exclusion criteria for this review were developed using population, intervention, comparison, and outcome (PICO) terms. A full list of the inclusion and exclusion criteria can be found in **Appendix 2**. Both reviews used the same inclusion/exclusion criteria, except the systematic review screeners had to be previously validated, unlike the primary scoping review screeners. Inclusion criteria included: 1) Studies that included brief (30 items or fewer) food intake screeners administered to adolescents aged 10-17 years old; 2) Studies that used and/or evaluated screeners focused on total diet or at least one food group (i.e. screeners focused on single nutrients were excluded); and 3) Peer-reviewed studies conducted in high-income countries, as per the World Bank definition [35]. Studies were retained even when sample ages extended beyond 10–17 years, provided that most participants were within the target age range. Although

review articles were not part of the formal inclusion criteria, relevant reviews were used to conduct backward and forward citation searches. Any eligible primary studies identified through this process were included. Studies were screened using Covidence, an online platform designed to streamline the management of systematic reviews [37]. For each of the reviews, the first stage involved screening titles and abstracts, followed by a full-text review in the second stage. For the scoping review, the first author screened the articles. A calibration exercise was carried out where 5% of the articles were screened by a trained second reviewer to ensure at least 90% agreement between reviewers [38]. Both rounds were conducted independently by two reviewers for the systematic review. In cases of disagreement regarding study inclusion, reviewers met to discuss their rationale for including/excluding the article, and when consensus was not met, the senior author was consulted to resolve the discrepancy.

### 4.4 Data extraction

Data was extracted by the first author and checked for completeness and accuracy by two reviewers. Domains of interest were identified *a priori* based on widely accepted factors used to guide evidence synthesis and data extraction. These included the screener name, target population (such as age and participant demographics), number of items, survey format (paper or web-based), recall period, and whether the tool assessed dietary frequency, portion sizes, or alignment with dietary guidelines. Information on the availability of validity and reliability data was also extracted. To ensure comprehensive information on the screeners, the research team also searched for validation studies linked to primary articles that did not report psychometric evaluations of their instruments.

### 4.5 Study quality assessment

No methodological quality assessment was conducted for the scoping review. For the systematic review, the methodological quality of each study was evaluated by two independent reviewers using the COSMIN Risk of Bias checklist [24,39]. Following COSMIN's "worst score counts" principle [24], where the methodological property rating defaults to the worst rating given for each section, the overall methodological quality for each measurement property was rated as "very good", "adequate", "doubtful", or "inadequate". Then, a final overall quality grading of the study was given as "sufficient", "insufficient, or "inconsistent" following COSMIN's 75% principle[34] that states that at least 75% of the measurement properties need to be labelled as "adequate" or "very good" to be deemed "sufficient". Only studies rated as "sufficient" or "inconsistent" (and only areas of which were "adequate or "very good" in those cases) were included in the data synthesis. Subsequently, the methodological quality of evidence (which refers to "how confident that the reported findings are accurate") for the pooled narrative results of included studies were evaluated as part of the COSMIN risk of bias assessment[40]. This process resulted in the evidence rating of "high", "moderate", "low", or "very low"for the body of included studies.

### 4.6 Data synthesis and analysis

Data from the papers included in the scoping review were organized into a descriptive summary table. For the systematic review, findings for each instrument's psychometric properties were either narratively synthesized (e.g. content validity) or key quantitative indicators were curated into a summary table. Quantitative measures for various psychometric properties are shown in **Table 1.** Measures of association (e.g.

Pearson or Spearman correlation coefficients) and measures of agreement (e.g. Kappa coefficients, Bland Altman plots) were commonly used to assess associations between test scores and those from a reference instrument (convergent validity). For example, Nelson and Lytle used kappa coefficients to assess agreement between their beverage and fast-food screener and two 24-hour dietary recalls. Since their instrument was designed to classify individuals into high or low intake categories for each food item, the resulting scores were treated as categorical variables rather than continuous measures, making kappa coefficients an appropriate measure of agreement[41]. There are currently no established cutoffs classifying dietary instruments as having acceptable or poor validity and reliability. Therefore, the cutoffs presented in Table 1 align with the most commonly used cutoffs derived from a previous systematic review on the psychometric properties of screeners used in children [42], as well as methodological literature[43–51].

Table 1. Definitions of common indicators used in methodological studies evaluating the psychometric properties of food screeners, along with their threshold criteria as appropriate.

| Statistical indicator | Definition | Cut-offs for assessing validity and reliability |
|---|---|---|
| Correlation coefficients (r) (Pearson's, Spearman's or Tau's correlation coefficients) | A numerical value ranging from –1 to +1 that describes the direction (positive or negative) and strength (magnitude) of a linear association between two variables. Pearson's correlation is used for parametric data that meet assumptions of normality and linearity, whereas Spearman's correlation is used for non-parametric data that does not meet normality and linearity assumptions. Kendall Tau's correlation is also a non-parametric test but is preferred when working with ordinal variables. High correlation coefficients indicate a strong association between two variables, such as responses from a food screener and 24-hour recall data, or repeated administrations of the screener, suggesting that the measures are capturing similar patterns of intake [43]. | Poor or failing: ≤ \|0.50\|<br><br>Acceptable: >\|0.51\| |
| Kappa coefficient (k) | Kappa coefficients (which can range from 0 to 1.0) are used as measure of agreement. Kappa coefficients describe how consistently individuals are classified across instruments, beyond what would be expected by chance, in categorical data. A value closer to 1.0 suggest greater confidence that the screener reliably captures the intended dietary behavior [44]. | Poor or failing ≤ 0.40<br><br>Acceptable: > 0.41 |
| Bland-Altman analysis | Bland-Altman analysis is used to assess agreement between two methods of estimating intake. This approach evaluates the extent of systematic bias and random error by plotting the difference between methods against their average. The central line represents the mean difference, or bias, indicating whether the screener tends to systematically overestimate (positive) or underestimate intake (negative). The outer lines, known as the limits of agreement, capture the range within | Poor/failing:<br>Plot with dots falling outside the limits of agreement.<br><br>Acceptable: tight plot, with most points falling within the agreement lines. |

| | | |
|---|---|---|
| | which most differences fall, reflecting the degree of random error [45]. | |
| Cross-classification analysis | Cross-classification analysis evaluates the degree of agreement in categorical ranking between two instruments. Typically, intake estimates from both the screener and the reference method are divided into quantiles (e.g., tertiles or quartiles), and individuals are classified accordingly [48]. | Poor: placed in opposite categories or ≤ 50%<br><br>Acceptable: placed in adjacent or same category or > 50% |
| Paired t-test statistic, Wilcoxon signed rank test | Paired t-test is used to assess whether the mean intake estimates from the screener differ significantly from those obtained via a reference method, or from a different attempt. The Wilcoxon signed-rank test is a non-parametric alternative to the paired t-test, used when intake data are not normally distributed or when sample sizes are small A statistically significant value indicates systematic bias, suggesting that the screener consistently over- or underestimates intake relative to the reference [46]. | Poor or failing: p-value significant.<br><br>Acceptable: p-value not significant. |
| Specificity and Sensitivity | Sensitivity measures how effectively a test identifies true positives (i.e. how accurately the screener classifies subjects who have consumed the food). Specificity measures how effectively a test identifies true negatives (i.e. how accurately the screener classifies participants who have not consumed the food). Sensitivity and specificity are normally presented in percentages[51]. | Poor or failing: sensitivity or specificity is ≤ 70%<br><br>Acceptable: sensitivity and specificity is > 70% |
| Confirmatory/Exploratory Factor Analysis (CFA/EFA) | EFA is a descriptive technique that identifies the underlying structure among items in a tool. With EFA, researchers must then use additional statistical methods to determine whether their items have been grouped as expected. In contrast, CFA allows the researcher to specify a hypothesized structure in advance and then test whether the items fit that model. The performance of EFA and CFA is evaluated using goodness-of-fit statistics, including chi-square values, root mean square error of approximation, comparative fit index, and several others[49]. | Poor or failing: Goodness-of-fit statistics indicate substantial discrepancies between the proposed model and the actual grouping of items.<br><br>Acceptable: Goodness-of-fit statistics indicate reasonable agreement between the proposed model and the actual grouping of items. |

| | | |
|---|---|---|
| Intra-class correlation coefficient (ICC) | ICC is used to assess the test-retest reliability across repeated administrations of a food screener. ICC values range from 0 to +1, with higher values indicating stronger test-retest reliability. ICC evaluates absolute agreement by accounting for both between-subject and within-subject variability [47]. | Poor or failing: ≤ 0.50<br><br>Acceptable: >0.51 |
| Cronbach's alpha | Cronbach's alpha is a measure of internal consistency, with values ranging from 0 to 1. Internal consistency is defined as the extent to which all items measure the same construct. Higher values of alpha describe a closer agreement of items, whereas lower values describe lower consistency. Cronbach's alpha is also related to the number of items; the less items, the lower the alpha terms meaning the number of items needs to be taken into consideration when interpreting the alpha value [50]. | Poor or failing: ≤ 0.50<br><br>Acceptable:<br>>0.51 |

# 5 Results

## 5.1 Study selection

**Figure 1** illustrates the identification and selection process for the included articles. A total of 46 articles were identified, representing 47 unique food screeners, all of which were included in the scoping review. Of these, 26 screeners had documented psychometric evaluations and were therefore eligible for inclusion in the systematic review.

A comprehensive table summarizing the quality assessment results (only for the systematic review) is provided in **Appendix 3**. Among the 26 screeners evaluated, seven demonstrated sufficient methodology and reporting standards, and 11 showed inconsistent quality. Eight screeners were deemed insufficient and therefore not included in the analysis. Studies with inconsistent quality were considered only for domains that met criteria for sufficiency.

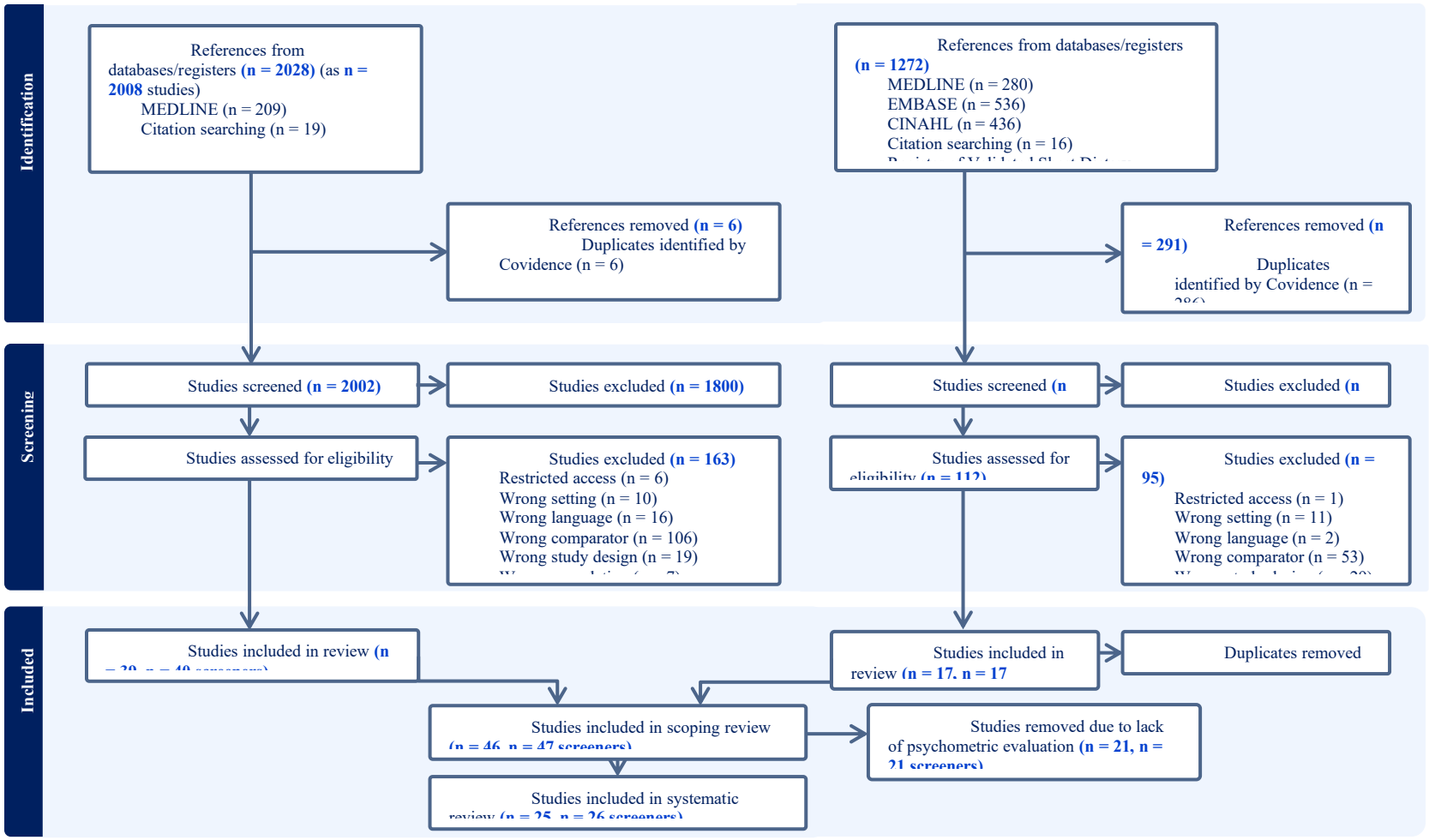



**Figure 1.** Preferred reporting items for systematic reviews and meta-analyses flow diagram of the study selection process for the hybrid review

## 5.2 Study characteristics

The screeners were developed across a wide range of countries, with the United States contributing the largest number (n=13), followed by Poland (n=6), Canada (n=6), Australia (n=5) and Spain (n=4) (see **Appendix 4**). Four screeners were developed collaboratively across multiple countries. Of the 26 screeners with psychometric evidence, 21 were validated in dedicated studies, and the remaining five were validated as parts of other studies. During the risk of bias appraisal, each measurement property received an individual grade, which was then synthesized into an overall rating for the study. Because only sufficient or inconsistent ratings (and only measurement properties graded adequate or very good when inconsistent) were included in the final determination, the study met the threshold for strong methodological quality[34]. Accordingly, the overall rating for the quality of evidence[24] in all measurement properties in the systematic review portion was deemed to be "high."

## 5.3 Food screener characteristics

**Table 2** presents the characteristics of the 47 food screeners from the scoping review. The number of items per screener varied considerably, with a median of 16 items. Three screeners did not report item counts. Paper-based formats remained the most common mode of administration (n=29). Fewer screeners were delivered online (n=6) or offered both paper- and web-based options (n=3), while eight studies did not report the mode of administration. Digital administration was more commonly used for screeners developed in Canada, Poland, South Korea, and Australia. Recall periods varied across screeners. A 7-day recall period was most frequently used, whether defined as a "usual week" (n=16) or a "past week" (n=12). The next most common approach was to assess diet over a 24-hour period, defined as either "usual day" (n=14) or "yesterday" (n=6). Longer recall windows such as "past month" (n=7) or "past year"

(n=6) were used less frequently. Four screeners did not specify a recall period, and five used more than one recall period for different items across their instrument. One study used "breakfast" as their recall period[52]. Most screeners assessed intake frequency (n=45) while portion size was included in only 12 instruments, typically for specific food groups such as beverages. Alignment with dietary guidelines was assessed in 22 screeners, most often in tools designed to evaluate diet quality or adherence to specific dietary patterns (e.g., Mediterranean diet indices or national dietary guidelines).

Table 2. Characteristics of self-administered food screeners developed for use among adolescents aged 10-17 years old (n=47 screeners)

| Screener name | Country | Age range | Sample population | Number of items | Survey format | Recall period | Assesses frequency? | Assesses portion size? | Assesses adherence to dietary guidelines? | Content validity | Construct/ criterion validity | Reliability |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **KIDMED 2.0** [53,54] | Spain | 10-17 | Southwestern | 15 | Paper | Usual week or usual day | | | ✓ | ✓ | ✓ | ✓ |
| **Modified beverage and snack questionnaire**[55] | U.S. | 5-17 | Cache County, Utah | 27 | Paper | Past Week | ✓ | | | | ✓ ** | |
| **NHANES dietary screener questionnaire**[56–58] | U.S. | 12-15 | Not reported | 26 | Paper or web-based | Past month | ✓ | | | | ✓** | |
| **KIDMED 2.0 Poland**[59] | Poland | 10-18 | Not reported | 26 | Web-based | Usual week or usual day | | | ✓ | ✓ * | ✓ | ✓ |
| **Food behaviour checklist modified for children**[60] | U.S. | 2-19 | Columbus, Ohio | 12 | Not reported | Usual day or past week | ✓ (only for fruits & vegetables) | | | | | ✓ |
| **Breakfast eating habits questionnaire** [61] | Denmark | 16-30 | Not reported | 13 | Paper | Breakfast during the past school week (Monday-Friday) | ✓ | | | ✓ ** | | |
| **Diet questionnaire**[52] | Australia | 14 | Not reported | 9 | Paper | Yesterday | ✓ | | | | | |
| **Youth risk behavior surveillance system questionnaire**[62] | U.S. | 14-18 | Not reported | 6 | Paper | Past week or yesterday or usual day | ✓ | ✓ | ✓ | | ✓ | |
| **Eating Behaviour Scale** [63] | U.S. | 12-18 | Low-income, Urban | 20 | Paper | Yesterday | ✓ | | ✓ | | ✓ | ✓ |
| **KIDMED**[64] | Estonia | 2-24 | Tartu | 16 | Paper | Usual day or usual week | | | ✓ | | | |

| Screener name | Country | Age range | Sample population | Number of items | Survey format | Recall period | Assesses frequency? | Assesses portion size? | Assesses adherence to dietary guidelines? | Content validity | Construct/ criterion validity | Reliability |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Short FFQ**[65] | New Caledonia | 11-16 | Not reported | 26 | Paper | Usual day | Not reported | | ✓ | | | |
| **Questionnaire to asses fruit and vegetables** [66,67] | Denmark, Norway, Iceland, Belgium, Portugal, Spain | 11-12 | Not reported | 12 | Paper | Usual week | ✓ | ✓ | ✓ | ✓*** | ✓*** | ✓*** |
| **Questionnaire to assess fatty food intake** [68] | U.S., Hungary, Netherlands | 14-19 | Not reported | 11 | Paper | Usual day | ✓ | | | | | |
| **Short FFQ** [69] | Australia/Torres Straight Islands | 10-12 | Aboriginal and non-Indigenous rural | 30 | Paper | Usual day, usual week | ✓ | ✓ | | | ✓ | ✓ |
| **Beverage intake questionnaire**[70,71] | U.S. | 6-18 | New River Valley, Virginia | 15 | Paper | Past month | ✓ | ✓ | ✓ | | ✓ | ✓ |
| **Fruit and vegetable questionnaire** [72,73] | Norway | 10-12 | Kongsberg and Hønefoss | 16 | Paper | Usual week | ✓ | | | | ✓*** | ✓*** |
| **Fruit and vegetable questionnaire** [74] | U.K. | 15-18 | London | 2 | Paper | Usual day | ✓ | ✓ | ✓ | | ✓** | |
| **Questions to assess fruit and vegetable consumption**[75] | Poland | 13 | Not reported | 3 | Paper | Usual week | ✓ | | | | | |
| **Diet questionnaire** [76] | Canada | 12-17 | Toronto and Niagara region, Ontario | 9 | Web-based | Usual day | ✓ | | ✓ | | | |
| **Block screening questionnaire for fat intake**[77] | Poland | 13-21 | Female only | 15 | Paper | Past year | ✓ | | | | | |
| **Short form FFQ 4 polish children** [78] | Poland | 6-15 | Not reported | 9 | Paper | Past year | ✓ | | | | | ✓ |
| **Block screening questionnaire for** | Poland | 13-18 | Not reported | 9 | Paper | Past year | ✓ | | | | | |

| Screener name | Country | Age range | Sample population | Number of items | Survey format | Recall period | Assesses frequency? | Assesses portion size? | Assesses adherence to dietary guidelines? | Content validity | Construct/ criterion validity | Reliability |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **fruit/vegetable/fiber intake** [79] | | | | | | | | | | | | |
| **OPIC dietary intake questionnaire**[80] | New Zealand | 9-13 | South Auckland | Not reported | Not reported | Not reported | not reported | Not reported | Not reported | | | ✓ ** |
| **COMPASS (Cq) questionnaire**[81] | Canada | 14-15 | Southwestern Ontario | 4 | Web-based | Yesterday | ✓ | | ✓ | | ✓ | ✓ |
| **Short food frequency questionnaire for Norwegian children** [82] | Norway | 9, 13 | Not reported | 23 | Paper | Past year | ✓ | | | | ✓ | |
| **Dietary adherence tool** [83] | South Korea | 12-18 | Not reported | 13 | Web-based | Usual week, usual day | ✓ | ✓ (only for water) | ✓ | ✓** | ✓ | |
| **Questionnaire to assess snack consumption behaviours** [84] | Poland | 10-15 | Not reported | 20 | Not reported | Usual day | ✓ | ✓ | ✓ | ✓** | | ✓** |
| **Targeted nutrient questionnaire** [85] | U.S. | 13-17 | Iowa | 22 | Not reported | Past week | ✓ | ✓ | ✓ | | | |
| **Beverage consumption questionnaire** [86] | Ireland | 12-18 | Not reported | 23 | Paper | Usual week | ✓ | ✓ | ✓ | | | |
| **1995 NNS**[87] | Australia | 12-17 | Not reported | 6 | Paper | Usual day | ✓ | | | | | |
| **Beverage and snack questionnaire** [88] | U.S. | 12-13 | Seattle, Washington | 19 | Paper | Past week | ✓ | | ✓ | ✓*** | ✓*** | ✓*** |
| **Short food frequency questionnaire adapted from INTERHEART FFQ** [89] | Canada | 14-15 | Survey, British Columbia and Brampton, Ontario with South Asian descent | 25 | Paper | Usual day, usual week, usual month | ✓ | | | | | |

| Screener name | Country | Age range | Sample population | Number of items | Survey format | Recall period | Assesses frequency? | Assesses portion size? | Assesses adherence to dietary guidelines? | Content validity | Construct/ criterion validity | Reliability |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Diet history questionnaire** [90] | U.S. | 15-17 | Pittsburgh, Pennsylvania | 9 | Not reported | Usual week | ✓ | | | | | |
| **ISCOLE FFQ**[91–94] | Finland, U.S and Columbia | 11-17 | Helsinki, Espoo, Vantaa, Baton Rouge and Bogotá | 23 | Not reported | Usual week | ✓ | | | | ✓ | ✓ |
| **CoCu FFQ**[95] | Germany | 10-19 | Leipzip, | 14 | Paper | Usual day or usual week | ✓ | ✓ | ✓ | | ✓ ** | ✓ ** |
| **SPAN diet questions**[96] | U.S. | 10-18 | Texas | 25 | Not reported | Yesterday | ✓ | | | | | |
| **HBSC questionnaire**[97–99] | Austria, Belgium, France, Germany Greece, Hungary, Italy, Spain and Sweden | 11,13,15 | Vienna, Gent, Lille, Dortmund, Athens, Heraklion, Pécs, Rome, Zaragoza and Stockholm | 15 | Paper or web-based | Past week | ✓ | | | | ✓ | ✓*** |
| **Brief Mediterranean diet screener**[100] | Spain | 3-80 | Girona | 15 | Paper | Past year | ✓ | ✓ | ✓ | | | |
| **Short diet quality screener**[100] | Spain | 3-80 | Girona | 18 | Paper | Past year | ✓ | ✓ | | | | |
| **NDNS short FFQ**[101] | U.K. | 11-18 | Not reported | 20 | Paper or web-based | Past month | ✓ | ✓ | | ✓** | | |
| **PrimeScreen**[102] | U.S. | 13-18 | Massachusetts | 18 | Not reported | Past month | ✓ | | | | | |
| **Foodstuffs questionnaire**[103] | Poland and Czech Republic | 10-12 | Rogów, Wodzislaw, Śląski Bohumín and Šilheřovice | 13 | Not reported | Not reported | ✓ | | ✓** | ✓* | | |
| **Canadian Food Intake Screener for Adolescents**[104] | Canada | 10-17 | Not reported | 10 | Paper | Past week | ✓ | | ✓ | ✓ | | |

| Screener name | Country | Age range | Sample population | Number of items | Survey format | Recall period | Assesses frequency? | Assesses portion size? | Assesses adherence to dietary guidelines? | Content validity | Construct/ criterion validity | Reliability |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **French adaptation of the Beverage intake questionnaire[105]** | Canada | 14-17 | Québec | 15 | Web-based | Past month | ✓ | ✓ | ✓ | ✓* | ✓ | ✓ |
| **Krece plus food questionnaire[106]** | Spain | 9-12 | Reus | 16 | Paper | Not reported | not reported | | ✓ | | | |
| **Jeddah FFQ[107]** | Saudi Arabia | 13-18 | Jeddah city | 17 | Not reported | Past week | ✓ | | ✓** | ✓** | | |
| **Dietary habits questionnaire[108]** | New Zealand | 15-18 | Not reported | 17 | Paper | Past month | ✓ | ✓ | ✓ | ✓* | ✓** | |

*Content validity only partially addressed (for example, the content of the screener was reviewed by experts but not cognitively tested in the target population).
** Psychometric property did not pass quality assessment per the COSMIN risk of bias requirements[24].
*** Evaluated in a separate study published prior to 2010.

## 5.4 Content validity

Among the 47 tools examined, 12 had been assessed for content validity. However, two studies[67,88] were evaluated for content validity prior to 2010, and five screeners[83,84,101,103,107] did not pass the COSMIN risk of bias quality assessment[83,84,101,103,107] and were therefore excluded from the analysis. Within the five studies that met inclusion met quality assessment, content validity was assessed using expert panel reviews[54,59,104,105,108] with only two studies [54,104] conducting cognitive interviews with the target population. Expert panels varied widely in size, ranging from groups of seven[105] or eight[54] experts to larger panels of up to twenty[104] members. Panels typically included researchers, dietitians, public health professionals and pediatricians whose role was to verify alignment with dietary guidelines, ensure comprehensive coverage of all diet constructs, and confirm that items reflected the intended constructs. In two studies[59,105], experts also contributed to cultural adaptation processes, ensuring that items were relevant and acceptable for specific populations. Two studies[54,104] quantified expert agreement, reporting levels of 75% and 100%, indicating strong consensus on item relevance. Cognitive interviewing assessing comprehensibility of the screener items were usually conducted in a small sample of participants (n=6-11) by asking participants to read each item with specific probes to ensure participants understood the question in the way it was intended to. Only one study[104] reported detailed findings related to item comprehension such as unclear phrasing, readability issues, or keyword confusion. These insights led to revisions such as rewording items, adjusting response options, or modifying sequencing to improve flow and comprehension.

## 5.5 Construct and criterion validity

**Table 3** summarizes the findings for the 13 methodological studies[54,59,60,62,63,69,71,81–83,94,99,105] (n=13 screeners) that assess criterion or construct validity. Most studies (n=9/13) evaluated construct (convergent) validity[54,62,63,69,71,81,82,94,105] while two studies measured only criterion validity[59,99]. No studies measured both construct and criterion validity. Among the nine studies that assessed convergent validity by statistically comparing screener food groups with those from a reference instrument, five used 24-hour dietary recalls and four used food records as the reference method. All nine studies used measures of association, such as Pearson's or Spearman's correlation coefficients to compare the screener with its reference instrument. In addition, three of the screeners also employed measures of agreement, such as Cohen's kappa.

Across studies assessing convergent validity, correlation coefficients were frequently low, often falling below 0.30 for most food groups. Correlation coefficients also varied across food categories. For example, vegetables, grains, and fried foods tended to show the weakest correlations, whereas fruits, dairy, and sugar-sweetened beverages generally demonstrated comparatively higher values. Studies that also validated their screeners using measures of agreement such as Cohen's kappa coefficients[54,63,81] reported slightly higher values, typically around 0.40–0.50, indicating acceptable validity. Lowest agreement values were seen for olive oil, meat and grain products, whereas agreement was higher for breakfast foods, milk and fruits. Two studies[82,94] used other statistical indicators, including calculating sensitivity and

specificity for correctly classified foods, as well as examining levels of food misclassification. Both studies reported acceptable validity.

Two studies assessed criterion validity by comparing their final diet scores derived from the food screener to either weight status[59] , which showed acceptable validity, or blood biomarkers[99] , which showed poor validity.

Two screening tools[60,83] were tested for construct (structural) validity using confirmatory or exploratory factor analysis. Only one screener[94] was validated through linear regression methods for hypothesis-testing. All three screeners were deemed to provide acceptable construct validity.

Table 3. Findings from methodological studies assessing the validity of food screeners designed for use among adolescents aged 10-17 years (n = 13 studies assessing 13 screeners)

| Screener name [ref] | Population, sample size | Reference instrument | Methods and statistics used to assess validity | Findings | Classification based on cut-off* |
|---|---|---|---|---|---|
| **KIDMED 2.0**[54] | 12-18 years old, n = 419 | Seven-day food record | Agreement between the screener and the reference instrument was assessed using kappa statistics for 16 food groups. | Cohen's kappa (range): -0.14-0.47. Lowest kappa for olive oil; highest for breakfast foods, industrial pastries, and fruit. | Poor – acceptable |
| **KIDMED 2.0 Poland**[59] | 10-18 years old, n = 102 | Anthropometric (weight and height) | Correlation coefficients assessed association between total diet scores derived from 16 food groups (obtained from KIDMED index) and BMI percentiles. | Spearman's correlation: -0.854. | Acceptable |
| **Food behaviour checklist modified for children**[60] | 8-10 years old, n = 97 | N/A | Confirmatory factor analysis for model confirmation. | CFI = 0.962, RMSEA = 0.064, insignificant chi square test of model fit ($p < 0.11$). | Acceptable |
| **Youth risk behavior surveillance system questionnaire**[62] | 14-18 years old, n = 610 | Three 24-hour dietary recalls | Mean intakes of 14 food groups (split into three sets of questions) were compared between test and reference instrument using paired t-tests and correlation coefficients. | Mean intake differences (range): 0.27-0.98 & 0.36-0.99 servings/day (based on the past 7 days and yesterday, respectively).<br><br>Pearson coefficients (range): 0.05-0.28 & 0.14-0.29 (based on the past 7 days and yesterday, respectively). Lowest r for vegetables; highest for fruits and fruit juice. | Poor - acceptable<br><br>Poor |

| | | | | | |
|---|---|---|---|---|---|
| **Eating behaviour scale**[63] | 12-18 years old, n = 387 | One 24-hour dietary recall (group 1 completed the test instrument first while group 2 completed the reference instrument first). | Mean intakes of 10 food groups compared using correlation analyses, separately for group 1 and group 2. | Spearman's correlations (range): group 1: 0.723 - 0.961 & group 2: 0.821 - 0.983. Lowest r for fried meats and beans; highest for dairy and non-nutritious foods such as chips, ice cream, doughnuts, brownies and chocolate candies. | Acceptable |
| | | | Agreement between instruments was tested using kappa coefficients. | Cohen's kappa (range): group 1: 0.42-0.87; group 2: 0.65-0.83. Group 2 showed a slightly higher agreement than group 1. | Acceptable |
| **Short FFQ**[69] | 10-12 years old, n = 205 | Three 24-hour dietary recalls | Mean intake (weight/volume) and frequency for the 21 food groups were compared across survey response categories using Kruskal–Wallis tests, and the Wilcoxon rank-sum test was applied to assess whether there was a statistically significant increasing or decreasing trend across the ordered response categories. | 16/21 food groups showed a significant ($p<0.05$) positive trend (higher screener responses corresponded to higher mean intake). | Acceptable |
| | | | Kendall Tau correlations coefficients were used to assess the strength of the association of the two rankings from the test and the reference instrument. | Kendall Tau Correlations (range): -0.1 - 0.3. No correlation found for bread, red meat, meat products, chicken and sweet foods; highest for fruit juice and breakfast cereals. | Poor |
| **Beverage intake questionnaire**[71] | 12-18 years old n= 326 | Four 24-hour dietary recalls | Mean intakes of 15 beverage groups compared using paired t-tests. | Mean intake differences (range): 0 to 4 (fl oz). Lowest mean intake difference for energy drinks, sweet tea and water; highest | Poor - acceptable |

| | | | | | |
|---|---|---|---|---|---|
| | | | | reduced-fat milk, fat-free milk, and fruit juice). | |
| | | | Pearson correlations were used to assess agreement between intakes from the test and reference instrument for 15 beverage groups. | Pearson coefficients (range): 0.26 – 0.72. | Poor - acceptable |
| **COMPASS (Cq) questionnaire** [81] | 14-15 years old, n = 178 | Web-based food record (the EaTracker tool) and weight status | Mean intakes for four food groups (fruit and vegetable, grain products, meat & alternatives and milk and alternatives) were assessed using correlational analyses (Spearman's coefficients for weight status and kappa coefficients for dietary intake). | Kappa coefficient (range): 0.33-0.48 Lower kappa for meat and grain products; highest for milk and fruit and vegetables. | Poor - acceptable |
| | | | | Spearman correlation coefficient with weight status (males and females combined): 0.86. | Acceptable |
| **Short food frequency questionnaire for Norwegian children** [82] | 9 and 13 years old, n = 1642 | Four-day food records | Mean intakes for 23 food groups were compared using Spearman correlation coefficients. | Spearman's coefficients (range): 0.14-0.63. Lowest r for pizza and potato chip; highest for fruits and vegetables. | Poor - acceptable |
| | | | Sensitivity (true consumers) and specificity (true non-consumers) of food intake compared to food records was calculated. | Sensitivity (range): 76-84% correctly identified. | Acceptable |
| | | | | Specificity (range): 66-83% correctly identified. | Acceptable |
| **Dietary adherence tool**[83] | 12-18 years old, n = 723 | N/A | Exploratory factor analysis for model confirmation. | KMO = 0.781, Bartlett's test of sphericity $\chi 2 = 1043.15$ ($p < 0.001$). "Good" model fit. | Acceptable |
| **ISCOLE FFQ** [94] | 9-11 year old, n = 282 | Three-day food records | Mean intakes for 23 food groups were compared using Spearman's correlation coefficients. | Spearman's coefficient (range): 0.00-0.64. Lowest coefficients for fruit, deep fried food and grains; highest for milk and sugar sweetened beverages. | Poor – acceptable |

| | | | | | |
|---|---|---|---|---|---|
| | | | Cross-classifications percentages (%) used to assess misclassification of foods. | Cross-classifications (range): 0-1% for 19/23 food groups. Whole milk (4.7%), dark-green vegetables (2.2%), orange vegetables (1.3%) and diet sodas (1.6%) had the highest misclassification rates. | Acceptable |
| | | | Construct (hypothesis-testing) validity was assessed using logistic regression to test differences in diet scores based on sociodemographic characteristics. | P <0.05 for country differences and for parental education level which was deemed "consistent among previous research". | N/A |
| **HBSC FFQ questionnaire**[99] | Not reported, adolescents n = 390 | Blood biomarkers | Correlation analysis used to compare blood biomarkers and diet scores. | Pearson's correlation (range): 0.066-0.315. | Poor |
| | | | | Spearman's correlation (range): 0.058-0.289. | Poor |
| **French adaptation of the beverage intake questionnaire**[105] | 14-17 years old n = 60 | Two 24-hour dietary recalls | Mean intake differences for 17 beverages were compared using correlation analyses. | Mean differences (range): -1.2 to 463.9 ml/day. Lowest mean differences for sweetened teas, hot chocolate and sports drinks; highest for total water. | Poor – acceptable |
| | | | | Spearman correlation (range): 0.13 to –0.67. | Poor – acceptable |

*Refer to Table 1 for convergent validity cut-offs.

## 5.6 Reliability

**Table 4** summarizes the reliability evidence for the 11 methodological studies reviewed[54,59,60,63,69,71,78,81,84,94,105] (n=11 screeners). Nine screener[53,58,63,69,71,78,81,94,105] were tested for test-retest reliability and two screeners[60,84] were measured for internal consistency. Most screeners (n=7/9) were classified as providing acceptable test-retest reliability[54,60,69,71,78,84,105]. However, measures of agreement were generally lower for meat, grains, and certain beverages, including sports drinks and sweetened tea. In contrast, fruits, vegetables, dairy products, sugar-sweetened beverages, and some snack foods demonstrated higher agreement, as evidenced by stronger correlations, kappa values, or ICCs. Although test-retest reliability was not uniformly strong across all measures, adolescent food screeners tended to perform consistently over time, and the majority met accepted reliability thresholds. Internal consistency was assessed with Cronbach's alpha values ranging from 0.43 to 0.80 across food groups. Both studies[60,84] demonstrated acceptable values for most items.

Table 4. Findings from methodological studies testing the reliability of food screeners designed for use among adolescents aged 10-17 years (n = 11 studies assessing 11 screeners)

| Screener name [ref] | Population, sample size | Methods and statistics used to assess reliability | Reliability results | Classification based on cut-off* |
|---|---|---|---|---|
| **Kidmed 2.0** [54] | 12-18 years old<br>n = 419 | Agreement between attempts was assessed using kappa statistics over a two-week interval. | Cohen's kappa (range): 0.29-0.72 for 16 food groups. Lower kappa for protein consumption; higher for fruit and vegetables, pastries and sweets. | Poor-acceptable |
| **Kidmed 2.0 PL** [59] | 10-18 years old<br>n = 102 | Grouped correlation analysis between the two attempts was measured over a two-week interval. | Spearman's correlation= 0.876. | Acceptable |
| **Food behaviour checklist modified for children** [60] | 8-10 years old<br>n = 97 | Internal consistency measured using Cronbach's alpha for each food factor. | Cronbach's alpha (range): 0.43-0.67. Lower alpha for milk consumption and healthy eating factors; higher for fruit and vegetable consumption factor. | Poor-acceptable |
| **Eating behaviour scale**[63] | 12-18 years old<br>n = 387 | Correlation analysis between the two attempts was measured over a five-hour interval. Correlation analysis and kappa statistics was used to assess their agreement. | Spearman's correlation (range): 0.744-0.985 for 10 food groups. Lower r for beans and vegetables; higher for meat, non-nutritious foods and drinks such as such as chips, ice cream, doughnuts, brownies, chocolate candies and soda.<br><br>Cohen's kappa (range): 0.49-0.79 for 10 food groups. | Acceptable<br><br>Acceptable |
| **Short FFQ**[69] | 10-12 years old<br>n = 205 | Quadratic weighted and simple kappa statistics were used to assess agreement between attempts over a two-week interval. | Cohen's kappa (range): 0.30-0.82 for 21 food groups. Lower kappa for bread, milk, pasta and meat; higher for butter, types of bread, types of milk and salty snacks. | Acceptable |

| | | | | |
|---|---|---|---|---|
| **Beverage intake questionnaire** [71] | 6-18 years old<br>n= 326 | Mean intakes of beverage groups for each attempt compared using paired t-tests and correlation analysis over a two-three-week interval. | Mean intake differences: 0 –12 (fl oz). Statistically significant for ($p<0.005$) for 3/15 groups. | Acceptable |
| | | | Pearson's coefficients (range): 0.313 – 0.78. Lower r for energy drinks and sweetened juice; higher for 100% fruit juice and fat-free milk. | Acceptable |
| **Short form FFQ 4 polish children** [78] | 11-15 years old<br>n = 630 | Agreement between the attempts was tested using kappa statistics and cross-classifications over a two-week interval | Cohen's kappa (range): 0.30-0.54 for 13 food groups. Lowest kappa for fruit, vegetables and sweets; highest for breakfast foods and school meals. | Poor – acceptable |
| | | Bland Altman plots were used to assess agreement between attempts by comparing diet indices applied to each attempt | Bland Altman plots: very scattered plot both above and below the agreement lines. | Acceptable |
| | | | Cross-classifications (range): 42.9 – 80.1% agreement for 13 food groups. | Poor – acceptable |
| **COMPASS (Cq) questionnaire** [81] | 14-15 years old<br>n = 178 | Mean intakes were compared between attempts using intra-class correlations over a one-week interval. | ICC (range): 0.41-0.69 for 8 food groups. Lower ICC for meat and grain products; higher for milk and fruit and vegetables. | Poor – acceptable |
| **Questionnaire to assess snack consumption behaviours** [84] | 10-15 years old<br>n= 50 | Internal consistency was measured using Cronbach's alpha. | Cronbach's alpha: 0.80. | Acceptable |
| **ISCOLE FFQ** [94] | 9-11 years old<br>n = 321 | Mean intakes were compared between attempts using intra-class correlations and Wilcoxon's signed rank test over a five-week interval. | ICC (range): 0.37-0.78 for 23 food groups. Lower ICC for dark-green and orange vegetables; higher for low-fat milk and sports drinks. | Poor - acceptable |

| | | | | |
|---|---|---|---|---|
| | | Cross-classifications were used to assess the percentage (%) of misclassifications between attempts. | Misclassification rate (range): 0.0-4.7%. | Acceptable |
| | | | Wilcoxon's signed rank test (range): <0.001-0.504 (significant differences for 16/23 groups). | Poor |
| **French adaptation of the beverage intake questionnaire** [105] | 14-17 years old<br>n = 59 | Mean differences were compared using correlation analysis for 17 sweetened beverage groups over a two-week interval | ICC (range): 0.09-0.93. Lowest r for hot chocolate and diet sodas; highest for sweetened coffee, vitamin water and flavored milk. | Acceptable |

*Refer to Table 1 for reliability cut-offs.

# 6 Discussion

## 6.1 Overview of the main findings

In this review, we set out to describe existing food screeners used with adolescents, and to synthesize their measurement properties, including validity and reliability. We identified 47 screeners used in high-income countries and analyzed 26 for their psychometric properties. The screeners identified in the scoping review originated from a wide range of countries and varied substantially in their structure, purpose, and administration. When selecting an appropriate screener, researchers are therefore required to align their screener of choice with the specific aims of their study, ensuring it was developed with the target population in mind[24]. By consolidating key information on each screener's characteristics and evaluation status, this review provides a foundation to support researchers in making informed decisions about which screener best fits their needs.

The systematic review revealed that only three screeners, including the KIDMED 2.0[54], the KIDMED 2.0 PL[59] and the French adaptation of the Beverage Intake Screener[105], were comprehensively assessed for content, construct or criterion validity, and reliability. Two other screeners[67,88] were identified as being assessed for all psychometric properties, but since their evaluation was conducted prior to 2010, they were not included in the systematic analysis. These screeners showed that a more rigorous approach is feasible, in contrast to most screeners, which were evaluated on only one or two properties, leaving important aspects of validity unexamined. The Eating Behaviour Scale[63], while not comprehensively assessed, was reported to have the highest

correlation coefficients across all food groups. This may result from the screener's shorter reference period, which allows less room for random error.

Despite the importance of examining the content validity of diet screeners, only five out of 47 food screeners were properly assessed for this measurement property. Most methodological studies used measures of association to evaluate the construct or criterion validity of their tools. Content validity is crucial because it shows whether a screener's items adequately represent the behaviours it is meant to measure, forming the basis for all subsequent measures of validity. Without strong content validity, statistical analyses cannot reliably demonstrate that the screener measures adolescent dietary behaviours as intended[109]. Further, several studies relied solely on pilot testing, which, while useful for feasibility and clarity, cannot show whether a screener's content represents the behaviours it aims to measure[24,25]. This overreliance could suggest a broader misunderstanding of content validity, which highlights the need for structured approaches such as expert review and cognitive interviewing. Without the proper approaches, a screener may not truly reflect adolescent dietary behaviours, therefore limiting their interpretability and usefulness.

While the quality of the assessments varied greatly, construct (convergent) validity was the most frequently assessed form of validity. Most studies used correlation analysis to compare data obtained from a screener with data from a reference instrument, such as 24-hour recalls or food records, ensuring that the reference instrument matched the recall period of the food screener. For example, the Eating Behaviour Scale[63] appropriately justified using a single 24-hour recall because the screener was designed to measure intake the day prior, allowing for direct comparison. In contrast, other screeners

aimed to capture usual monthly or yearly intake and needed several 24-hour recalls or food records. To capture usual intake, most studies used food intake data collected over either three[62,69,94] or four[71,81] days. Due to this difference in their intended recall period, it is not appropriate to expect comparable statistical associations across tools; instruments designed to capture intake "yesterday," will naturally show stronger alignment with single-day reference measures than screeners aiming to estimate usual weekly or monthly intake, which are inherently more prone to recall error.

Most screeners demonstrated poor to acceptable construct or criterion validity, with few showing consistently strong validity across food groups. Studies that assessed convergent validity using agreement-based metrics tended to report more favourable results than those relying on correlation coefficients. Agreement statistics often appear higher because they only assess whether two tools place individuals into the same general intake categories, even when the actual amounts reported differ. Correlation coefficients, however, capture those differences in magnitude, so they tend to be lower when one tool systematically over- or underestimates intake[45]. As a result, stronger agreement values may reflect the choice of statistic rather than genuinely better validity. For tools aiming to compare direct intake values (i.e., those using "yesterday" as the recall period), correlation statistics can be effective because the magnitudes align. However, screeners that assess usual intake should rely on agreement measures, since their magnitudes are not expected to match directly.

Performance of construct or criterion validity often varied depending on the specific dietary component assessed. Consistent with past literature[1,2,23], foods that were easier for adolescents to recall tended to be those they were both more willing and more

able to report. Grains and proteins were often underreported[69,81,94], which aligns with prior findings that show omissions may vary based on food groups among adolescents. In addition, as adolescents gain greater autonomy over their food intake (especially snacks which they may be responsible for preparing), they may be more likely to remember and accurately report these items[1]. Indeed, we found that snacks and fruits were recalled more consistently[54,62,69,71,81,82], while less noticeable foods such as vegetables incorporated into mixed dishes or cooking oils, were frequently omitted[54,62].

One of the key aspects of construct validity is structural validity. Structural validity was not commonly measured in the food screeners reviewed. A screener with high structural validity would show whether items cluster into meaningful dietary constructs, for example, whether fruit and vegetable items load onto a shared "healthy eating" factor or snack-related items form a distinct behavioural domain. This is particularly important for screeners adapted to specific cultural contexts or dietary guidelines[110]. Without evidence of high structural validity, it is unclear whether screeners measure coherent constructs or just sets of loosely related items.

Food screeners remain useful for capturing dietary intakes quickly and at low cost, making them appealing for researchers trying to assess food consumption patterns in large groups or populations. Their modest convergent validity for several food groups, together with generally strong test–retest reliability[54,60,69,71,78,84,105], suggests they are appropriate for classifying levels of intake and can measure food intake reliably, even if they cannot estimate precise quantities[9]. Most studies that assessed test–retest reliability reported acceptable stability over short intervals, indicating that adolescents tend to provide consistent responses when the screener is administered more than once. For

monitoring population-level behaviours, tracking adherence to recommendations, or classifying adolescents into broad dietary categories, screeners provide meaningful, low-burden insights.

### 6.2 Strengths and Limitations

The hybrid design strengthened this review by supporting both objectives, beginning with broad coverage for the scoping review. Because not every screener required a full quality assessment, we were able to include a wider range of tools than a strictly systematic design would have permitted. This inclusivity was particularly important for mapping the landscape of food screeners and capturing the diversity of formats, target populations, and measurement aims.

The decision to cap screeners at a maximum of 30 items enhanced feasibility for this objective; however, it also resulted in the exclusion of longer instruments that may possess stronger psychometric properties. For instance, the Block Kids Food Screener[111], a widely used and well-regarded tool for assessing dietary intake among adolescents, contains 41 items and was therefore excluded despite its established utility and robust measurement properties. This decision was made to ensure that the included tools remained brief. Because no established cut-off exists to distinguish a food screener from a full food frequency questionnaire, a threshold of 30 items was applied. This threshold was selected because it captures most instruments commonly classified as screeners and approximates a completion time of 10 minutes or less based on data survey tools[112,113], thereby maintaining the intended brevity of the instruments.

In the systematic review, the flexibility of the hybrid approach allowed for targeted, rigorous synthesis where appropriate; however, the broad scope introduced

several limitations that constrained the depth and comparability of the validity evaluation. Screeners were validated with diverse reference methods and statistical approaches, limiting direct comparison across screeners. Variation in the constructs assessed by each screener further limited comparability; a more focused scope (e.g., screeners targeting only fruit and vegetable intake) would have allowed clearer cross-screener comparisons. Inconsistent reporting quality across validation studies created additional barriers, as essential methodological details were often missing or inadequately described.

In the systematic review, the COSMIN Risk of Bias[114] appraisal was used to assess the quality of validation studies for the screeners. A comprehensive citation search was conducted to ensure all the relevant validation data was obtained. Screeners were considered to have passed the quality assessment if more than 75% of their measurement properties were rated at least "adequate." This rigorous approach strengthens the report by ensuring that only studies meeting established quality standards were included. The COSMIN guidelines were selected because they provide the most comprehensive and widely accepted framework for evaluating the methodological quality of studies assessing measurement properties. Since food screeners function as measurement instruments, COSMIN offered an appropriate and rigorous structure for assessing content, construct, and criterion validity, as well as reliability, in a transparent and standardized manner.

Finally, although COSMIN provides defined criteria for evaluating measurement properties[114], these thresholds are not tailored to dietary assessment tools, which often measure multiple dietary components. This means that it was necessary to define our own cut-offs from methodological literature. Within a food screener, some

food items scored in the acceptable range while others scored in the lower/unacceptable range, making it difficult to provide an overall 'grade' for the tool.

### 6.3 Future research directions

Future work should prioritize improving the transparency, consistency, and methodological rigour of food screener validation. Authors should use clear, standardized validation terminology, as several studies in this review did include validation evidence but were only identified through the scoping search, not the systematic search, because their terminology did not include the expected validation keywords. Explicitly indexing validation-related concepts would improve discoverability and reduce the risk of overlooking relevant screeners. Greater transparency is also needed; many screeners are described as fully validated despite only reporting some aspects of validity (often convergent validity). Some validation procedures do not follow best practices as outlined by the COSMIN guidelines[40]. In studies that were excluded from this review, methods used to assess validity were either incomplete, insufficiently described, or not aligned with COSMIN's recommended standards for evaluating measurement properties. Clear reporting on which aspects of validity have and have not been established would help researchers judge a tool's readiness for use.

Existing frameworks, such as the COSMIN guidelines[25], provide a strong foundation for improving the validation of dietary screening tools. Developing or adopting food screener-specific reporting guidelines could promote more appropriate validation methods, more consistent reporting, and clearer expectations for methodological quality, especially given the improper or incomplete approaches used in some studies. Future research should also define cutoffs for commonly used validation

statistics, as the lack of agreed-upon thresholds makes it difficult to judge whether screener performance is adequate for practice. Consensus-building efforts, such as expert panels or methodological reviews, could further standardize expectations and improve comparability across studies. Together, these steps would strengthen the evidence base and support the development of more robust food screeners for this population.

# 7 Conclusion

This review mapped the landscape of food screeners available for adolescents and found that, despite the wide range of screeners available, evidence supporting their validity remains uneven, inconsistently reported, and often difficult to interpret. By bringing together existing screeners and identifying key methodological gaps, the review clarifies the current state of the field and highlights shortcomings in validation practices. These findings underscore the need for more coherent, transparent, and methodologically rigorous approaches to developing and evaluating brief, practical, and scientifically robust screeners for adolescent dietary assessment.

# 8 References


1. Livingstone MBE. Issues in dietary intake assessment of children and adolescents. *Br J Nutr*. 2022;127(9):1426-1427. doi:10.1017/S0007114522000770

2. Krebs-Smith S, Subar AF, Thompson F, et al. Dietary Assessment Primer. Accessed March 21, 2025. https://dietassessmentprimer.cancer.gov/

3. Health Canada. Food and Nutrition Surveillance. July 26, 2004. Accessed October 9, 2025. https://www.canada.ca/en/health-canada/services/food-nutrition/food-nutrition-surveillance.html

4. Kirkpatrick S, Raffoul A. Measures Registry User Guide: Individual Diet. Published online January 2017.

5. Lee R, Nieman D. Measuring Diet. In: *Nutritonal Assessment*. 5th ed. McGraw-Hill Science/Engineering/Math; 2009:608.

6. Ravelli MN, Schoeller DA. Traditional Self-Reported Dietary Instruments Are Prone to Inaccuracies and New Approaches Are Needed. *Front Nutr*. 2020;7. doi:10.3389/fnut.2020.00090

7. Shim JS, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. *Epidemiol Health*. 2014;36:e2014009. doi:10.4178/epih/e2014009

8. Kirkpatrick S, Baranowski T, Subar AF, Tooze JA, Frongillo EA. Best Practices for Conducting and Interpreting Studies to Validate Self-Report Dietary Assessment Methods. *J Acad Nutr Diet*. 2019;119(11):1801-1816. doi:10.1016/j.jand.2019.06.010

9. Bailey RL. Overview of Dietary Assessment Methods for Measuring Intakes of Foods, Beverages, and Dietary Supplements in Research Studies. *Curr Opin Biotechnol*. 2021;70:91-96. doi:10.1016/j.copbio.2021.02.007

10. Baranowski T, Islam N, Baranowski J, et al. Comparison of a Web-based versus traditional diet recall among children. *J Acad Nutr Diet*. 2012;112(4):527-532. doi:10.1016/j.jada.2011.10.002

11. Conway JM, Ingwersen LA, Vinyard BT, Moshfegh AJ. Effectiveness of the US Department of Agriculture 5-step multiple-pass method in assessing food intake in obese and nonobese women12. *Am J Clin Nutr*. 2003;77(5):1171-1178. doi:10.1093/ajcn/77.5.1171

12. Moshfegh AJ, Rhodes DG, Baer DJ, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes1. *Am J Clin Nutr*. 2008;88(2):324-332. doi:10.1093/ajcn/88.2.324

13. National Cancer Institute. ASA24® Dietary Assessment Tool | EGRP/DCCPS/NCI/NIH. December 2, 2024. Accessed March 12, 2025. https://epi.grants.cancer.gov/asa24/

14. Thompson FE, Dixit-Joshi S, Potischman N, et al. Comparison of Interviewer-Administered and Automated Self-Administered 24-Hour Dietary Recalls in 3 Diverse Integrated Health Systems. *Am J Epidemiol*. 2015;181(12):970-978. doi:10.1093/aje/kwu467

15. Subar AF, Kirkpatrick SI, Mittl B, et al. The Automated Self-Administered 24-Hour Dietary Recall (ASA24): A Resource for Researchers, Clinicians and Educators from the National Cancer Institute. *J Acad Nutr Diet*. 2012;112(8):1134-1137. doi:10.1016/j.jand.2012.04.016

16. Freedman LS, Commins JM, Moler JE, et al. Pooled Results From 5 Validation Studies of Dietary Self-Report Instruments Using Recovery Biomarkers for Energy and Protein Intake. *Am J Epidemiol*. 2014;180(2):172-188. doi:10.1093/aje/kwu116

17. Hutchinson JM, Dodd KW, Guenther PM, et al. The Canadian Food Intake Screener for assessing alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations: scoring system and construct validity. *Appl Physiol Nutr Metab Physiol Appl Nutr Metab*. 2023;48(8):620-633. doi:10.1139/apnm-2023-0018

18. Bell LK, Golley RK, Magarey AM. A short food-group-based dietary questionnaire is reliable and valid for assessing toddlers' dietary risk in relatively advantaged samples. *Br J Nutr*. 2014;112(4):627-637. doi:10.1017/S0007114514001184

19. Livingstone MBE, Robson PJ, Wallace JMW. Issues in dietary intake assessment of children and adolescents. *Br J Nutr*. 2004;92 Suppl 2:S213-222. doi:10.1079/bjn20041169

20. Raffoul A, Hobin EP, Sacco JE, et al. School-Age Children Can Recall Some Foods and Beverages Consumed the Prior Day Using the Automated Self-Administered 24-Hour Dietary Assessment Tool (ASA24) without Assistance. *J Nutr*. 2019;149(6):1019-1026. doi:10.1093/jn/nxz013

21. Livingstone MBE, Robson PJ. Measurement of dietary intake in children. *Proc Nutr Soc*. 2000;59(2):279-293. doi:10.1017/S0029665100000318

22. de Vlieger NM, Weltert M, Molenaar A, et al. A systematic review of recall errors associated with portion size estimation aids in children. *Appetite*. 2020;147:104522. doi:10.1016/j.appet.2019.104522

23. Bodega P, de Cos-Gandoy A, Fernández-Alvira JM, Fernández-Jiménez R, Moreno LA, Santos-Beneit G. Body image and dietary habits in adolescents: a systematic review. *Nutr Rev*. 2023;82(1):104-127. doi:10.1093/nutrit/nuad044

24. Mokkink LB, Elsman EBM, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Qual Life Res*. 2024;33(11):2929-2939. doi:10.1007/s11136-024-03761-6

25. Hutchinson JM, Williams TE, Westaway AM, et al. Development of the Canadian Food Intake Screener to assess alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations. *Appl Physiol Nutr Metab*. 2023;48(8):603-619. doi:10.1139/apnm-2023-0019

26. Bingham S. The use of 24-h urine samples and energy expenditure to validate dietary assessments. *Am J Clin Nutr*. 1994;59(1):227S-231S. doi:10.1093/ajcn/59.1.227S

27. Gleason PM, Harris J, Sheean PM, Boushey CJ, Bruemmer B. Publishing Nutrition Research: Validity, Reliability, and Diagnostic Test Assessment in Nutrition-Related Research. *J Am Diet Assoc*. 2010;110(3):409-419. doi:10.1016/j.jada.2009.11.022

28. Marsh HW, Morin AJS, Parker PD, Kaur G. Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis. *Annu Rev Clin Psychol*. 2014;10(1):85-110. doi:10.1146/annurev-clinpsy-032813-153700

29. Christou E, Parmaxi A, Zaphiris P. A systematic exploration of scoping and mapping literature reviews. *Univers Access Inf Soc*. 2025;24(1):941-951. doi:10.1007/s10209-024-01120-3

30. Cumpston MS, McKenzie JE, Welch VA, Brennan SE. Strengthening systematic reviews in public health: guidance in the Cochrane Handbook for Systematic Reviews of Interventions, 2nd edition. *J Public Health*. 2022;44(4):e588-e592. doi:10.1093/pubmed/fdac036

31. Arero Dassie G, Dhugassa Girsha W, Kalu Beggi S. A hybrid scoping and systematic review of food safety and quality control in selected low- and middle-income countries: challenges, policies, and effective interventions in Africa, Asia, and Latin America. *Front Food Sci Technol*. 2025;5. doi:10.3389/frfst.2025.1570588

32. Little M, Rosa E, Heasley C, Asif A, Dodd W, Richter A. Promoting Healthy Food Access and Nutrition in Primary Care: A Systematic Scoping Review of Food Prescription Programs. *Am J Health Promot AJHP*. 2022;36(3):518-536. doi:10.1177/08901171211056584

33. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169(7):467-473. doi:10.7326/M18-0850

34. Prinsen C a. C, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil*. 2018;27(5):1147-1157. doi:10.1007/s11136-018-1798-3

35. World Bank. WDI - The World by Income and Region. 2025. Accessed October 20, 2025. https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html

36. National Cancer Institute. Search the Register of Validated Short Dietary Assessment Instruments. Published online February 9, 2023. https://epi.grants.cancer.gov/diet/shortreg/register.php

37. Veritas Health Innovation. Covidence systematic review software. Covidence. 2025. Accessed October 20, 2025. https://www.covidence.org/

38. Mak S, Thomas A. Steps for Conducting a Scoping Review. *J Grad Med Educ*. 2022;14(5):565-567. doi:10.4300/JGME-D-22-00621.1

39. Ma LL, Wang YY, Yang ZH, Huang D, Weng H, Zeng XT. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? *Mil Med Res*. 2020;7(1):7. doi:10.1186/s40779-020-00238-8

40. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res*. 2018;27(5):1171-1179. doi:10.1007/s11136-017-1765-4

41. Nelson MC, Lytle LA. Development and evaluation of a brief screener to estimate fast-food and beverage consumption among adolescents. *J Am Diet Assoc*. 2009;109(4):730-734. doi:10.1016/j.jada.2008.12.027

42. Bell LK, Golley RK, Magarey AM. Short tools to assess young children's dietary intake: a systematic review focusing on application to dietary index research. *J Obes*. 2013;2013:709626. doi:10.1155/2013/709626

43. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth Analg*. 2018;126(5):1763. doi:10.1213/ANE.0000000000002864

44. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Medica*. 2012;22(3):276-282.

45. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet Lond Engl*. 1986;1(8476):307-310.

46. Lombard MJ, Steyn NP, Charlton KE, Senekal M. Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. *Nutr J*. 2015;14:40. doi:10.1186/s12937-015-0027-y

47. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012

48. Masson LF, MCNeill G, Tomany JO, et al. Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. *Public Health Nutr*. 2003;6(3):313-321. doi:10.1079/PHN2002429

49. Santor DA, Haggerty JL, Lévesque JF, et al. An Overview of Confirmatory Factor Analysis and Item Response Analysis Applied to Instruments to Evaluate Primary Healthcare. *Healthc Policy*. 2011;7(Spec Issue):79-92.

50. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53-55. doi:10.5116/ijme.4dfb.8dfd

51. Monaghan TF, Rahman SN, Agudelo CW, et al. Foundational Statistical Principles in Medical Research: Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value. *Medicina (Mex)*. 2021;57(5):503. doi:10.3390/medicina57050503

52. Dumuid D, Mellow ML, Stanford TE, et al. Many different roads lead to Rome: equivalence of time-use for activity, sedentary and sleep behaviours and dietary intake profiles among adolescents. *J Act Sedentary Sleep Behav*. 2022;1:6. doi:10.1186/s44167-022-00005-1

53. Agostinis-Sobrinho C, Santos R, Rosario R, et al. Optimal Adherence to a Mediterranean Diet May Not Overcome the Deleterious Effects of Low Physical Fitness on Cardiovascular Disease Risk in Adolescents: A Cross-Sectional Pooled Analysis. *Nutrients*. 2018;10(7). doi:10.3390/nu10070815

54. López-Gajardo MA, Leo FM, Sánchez-Miguel PA, López-Gajardo D, Soulas C, Tapia-Serrano MA. KIDMED 2.0, An update of the KIDMED questionnaire: Evaluation of the psychometric properties in youth. *Front Nutr*. 2022;9:945721. doi:10.3389/fnut.2022.945721

55. Aguilar SS, Wengreen HJ, Lefevre M, Madden GJ, Gast J. Skin Carotenoids: A Biomarker of Fruit and Vegetable Intake in Children. *J Acad Nutr Diet*. 2014;114(8):1174-1180. doi:10.1016/j.jand.2014.04.026

56. Au LE, Gurzo K, Gosliner W, Webb KL, Crawford PB, Ritchie LD. Eating School Meals Daily Is Associated with Healthier Dietary Intakes: The Healthy Communities Study. *J Acad Nutr Diet*. 2018;118(8):1474-1481.e1. doi:10.1016/j.jand.2018.01.010

57. Ritchie LD, Wakimoto P, Woodward-Lopez G, et al. The Healthy Communities Study Nutrition Assessments: Child Diet and the School Nutrition Environment. *Am J Prev Med*. 2015;49(4):647-652. doi:10.1016/j.amepre.2015.06.016

58. Thompson FE, Midthune D, Kahle L, Dodd KW. Development and Evaluation of the National Cancer Institute's Dietary Screener Questionnaire Scoring Algorithms123. *J Nutr*. 2017;147(6):1226-1233. doi:10.3945/jn.116.246058

59. Bober J, Gaszyńska E. Validation of KIDMED 2.0 PL—Mediterranean Diet Quality Index for Polish Children and Adolescents. *Nutrients*. 2025;17(16):2636. doi:10.3390/nu17162636

60. Branscum P, Sharma M, Kaye G, Succop P. An evaluation of the validity and reliability of a food behavior checklist modified for children. *J Nutr Educ Behav*. 2010;42(5):349-352. doi:10.1016/j.jneb.2009.12.005

61. Christensen CB, Mikkelsen BE, Toft U. The effect of introducing a free breakfast club on eating habits among students at vocational schools. *BMC Public Health*. 2019;19(1):369. doi:10.1186/s12889-019-6701-9

62. Eaton DK, Olsen EO, Brener ND, et al. A Comparison of Fruit and Vegetable Intake Estimates from Three Survey Question Sets to Estimates from 24-Hour Dietary Recall Interviews. *J Acad Nutr Diet*. 2013;113(9):1165-1174. doi:10.1016/j.jand.2013.05.013

63. Fahlman M, McCaughtry N, Martin J, Garn AC, Shen B. Assessing the Eating Behaviors of Low-income, Urban Adolescents. *Am J Health Educ*. 2012;43(3):165-171.

64. Galan-Lopez P, Dominguez R, Pihu M, Gisladottir T, Sanchez-Oliver AJ, Ries F. Evaluation of Physical Fitness, Body Composition, and Adherence to Mediterranean Diet in Adolescents from Estonia: The AdolesHealth Study. *Int J Environ Res Public Health*. 2019;16(22). doi:10.3390/ijerph16224479

65. Galy O, Paufique E, Nedjar-Guerre A, et al. Living in Rural and Urban Areas of New Caledonia: Impact on Food Consumption, Sleep Duration and Anthropometric Parameters Among Melanesian Adolescents. *Nutrients*. 2020;12(7). doi:10.3390/nu12072047

66. Gebremariam MK, Lien N, Torheim LE, et al. Perceived rules and accessibility: measurement and mediating role in the association between parental education and vegetable and soft drink intake. *Nutr J*. 2016;15(1):76. doi:10.1186/s12937-016-0196-3

67. Haraldsdóttir J, Thórsdóttir I, de Almeida MDV, et al. Validity and Reproducibility of a Precoded Questionnaire to Assess Fruit and Vegetable Intake in European 11- to 12-Year-Old Schoolchildren. *Ann Nutr Metab*. 2005;49(4):221-227. doi:10.1159/000087276

68. Gerrits JH, O'Hara RE, Piko BF, et al. Self-control, diet concerns and eater prototypes influence fatty foods consumption of adolescents in three countries. *Health Educ Res*. 2010;25(6):1031-1041. doi:10.1093/her/cyq055

69. Gwynn JD, Flood VM, D'Este CA, et al. The reliability and validity of a short FFQ among Australian Aboriginal and Torres Strait Islander and non-Indigenous rural

children. *Public Health Nutr*. 2011;14(3):388-401. doi:10.1017/S1368980010001928

70. Hedrick VE, Savla J, Comber DL, et al. Development of a Brief Questionnaire to Assess Habitual Beverage Intake (BEVQ-15): Sugar-Sweetened Beverages and Total Beverage Energy Intake. *J Acad Nutr Diet*. 2012;112(6):840-849. doi:10.1016/j.jand.2012.01.023

71. Hill CE, MacDougall CR, Riebl SK, Savla J, Hedrick VE, Davy BM. Evaluation of the Relative Validity and Test-Retest Reliability of a 15-Item Beverage Intake Questionnaire in Children and Adolescents. *J Acad Nutr Diet*. 2017;117(11):1757-1766.e5. doi:10.1016/j.jand.2017.05.011

72. Hilsen M, van Stralen MM, Klepp KI, Bere E. Changes in 10-12 year old's fruit and vegetable intake in Norway from 2001 to 2008 in relation to gender and socioeconomic status - a comparison of two cross-sectional groups. *Int J Behav Nutr Phys Act*. 2011;8(101217089):108. doi:10.1186/1479-5868-8-108

73. Andersen LF, Bere E, Kolbjornsen N, Klepp KI. Validity and reproducibility of self-reported intake of fruit and vegetable among 6th graders. *Eur J Clin Nutr*. 2004;58(5):771-777. doi:10.1038/sj.ejcn.1601875

74. Huang P, O'Keeffe M, Elia C, et al. Fruit and vegetable consumption and mental health across adolescence: evidence from a diverse urban British cohort study. *Int J Behav Nutr Phys Act*. 2019;16(1):19. doi:10.1186/s12966-019-0780-y

75. Jodkowska M, Oblacinska A, Tabak I. How well do Polish teenagers meet health behaviour guidelines?. *Przegl Epidemiol*. 2014;68(1):65-51.

76. Korczak DJ, Perruzza S, Chandrapalan M, et al. The association of diet and depression: an analysis of dietary measures in depressed, non-depressed, and healthy youth. *Nutr Neurosci*. 2022;25(9):1948-1955. doi:10.1080/1028415X.2021.1918981

77. Kowalkowska J, Lonnie M, Wadolowska L, Czarnocinska J, Jezewska-Zychowicz M, Babicz-Zielinska E. Health- and Taste-Related Attitudes Associated with Dietary Patterns in a Representative Sample of Polish Girls and Young Women: A Cross-Sectional Study (GEBaHealth Project). *Nutrients*. 2018;10(2). doi:10.3390/nu10020254

78. Kowalkowska J, Wadolowska L, Hamulka J, et al. Reproducibility of a Short-Form, Multicomponent Dietary Questionnaire to Assess Food Frequency Consumption, Nutrition Knowledge, and Lifestyle (SF-FFQ4PolishChildren) in Polish Children and Adolescents. *Nutrients*. 2019;11(12):2929. doi:10.3390/nu11122929

79. Krusinska B, Kowalkowska J, Wadolowska L, Wuenstel JW, Slowinska MA, Niedzwiedzka E. Fibre-Related Dietary Patterns: Socioeconomic Barriers to

Adequate Fibre Intake in Polish Adolescents. A Short Report. *Nutrients*. 2017;9(6). doi:10.3390/nu9060590

80. Kulkarni AA, Swinburn BA, Utter J. Associations between diet quality and mental health in socially disadvantaged New Zealand adolescents. *Eur J Clin Nutr*. 2015;69(1):79-83. doi:10.1038/ejcn.2014.130

81. Leatherdale ST, Laxer RE. Reliability and validity of the weight status and dietary intake measures in the COMPASS questionnaire: are the self-reported measures of body mass index (BMI) and Canada's food guide servings robust? *Int J Behav Nutr Phys Act*. 2013;10(1):42. doi:10.1186/1479-5868-10-42

82. Lillegaard ITL, Øverby NC, Andersen LF. Evaluation of a short food frequency questionnaire used among Norwegian children. *Food Nutr Res*. Published online January 16, 2012. doi:10.3402/fnr.v56i0.6399

83. Lim J, Lee S, Hwang JY, Oh J. Empowering Healthy Adolescents: A Dietary Adherence Tool Incorporating Environmental Factors Based on Korean Guidelines. *Nutrients*. 2025;17(7):1102. doi:10.3390/nu17071102

84. Malczyk E, Walkiewicz KW, Muc-Wierzgon M, Dziegielewska-Gesiak S. Estimation of Intake of Fat, Saturated Fatty Acids, and Trans Fatty Acids from Sweet and Salty Snacks Among Children and Adolescents. *Nutrients*. 2025;17(9). doi:10.3390/nu17091572

85. Marshall TA, Van Buren JM, Warren JJ, Cavanaugh JE, Levy SM. Beverage Consumption Patterns at Age 13 to 17 Years Are Associated with Weight, Height, and Body Mass Index at Age 17 Years. *J Acad Nutr Diet*. 2017;117(5):698-706. doi:10.1016/j.jand.2017.01.010

86. Millar S, O'Donoghue M, McNulty B, Kirwan L, McKevitt A. A cross-sectional observation on habitual non-alcoholic beverage consumption among adolescents from four Irish post-primary schools. *Public Health Nutr*. 2017;20(3):404-412. doi:10.1017/S1368980016002627

87. Morley BC, Scully ML, Niven PH, et al. What factors are associated with excess body weight in Australian secondary school students?. *Med J Aust*. 2012;196(3):189-192. doi:10.5694/mja11.11184

88. Neuhouser ML, Lilley S, Lund A, Johnson DB. Development and Validation of a Beverage and Snack Questionnaire for Use in Evaluation of School Nutrition Policies. *J Am Diet Assoc*. 2009;109(9):1587-1592. doi:10.1016/j.jada.2009.06.365

89. Noor S, Dehghan M, Lear SA, et al. Relationship between diet and acculturation among South Asian children living in Canada. *Appetite*. 2020;147(6jw, 8006808):104524. doi:10.1016/j.appet.2019.104524

90. Park Y, Dodd KW, Kipnis V, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. *Am J Clin Nutr*. 2018;107(1):80-93. doi:10.1093/ajcn/nqx002

91. Pereira S, Katzmarzyk PT, Hedeker D, Maia J. Change and Stability in Sibling Resemblance in Obesity Markers: The Portuguese Sibling Study on Growth, Fitness, Lifestyle, and Health. *J Obes*. 2019;2019(101526295):2432131. doi:10.1155/2019/2432131

92. Katzmarzyk PT, Chaput JP, Fogelholm M, et al. International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE): Contributions to Understanding the Global Obesity Epidemic. *Nutrients*. 2019;11(4):848. doi:10.3390/nu11040848

93. Mikkilä V, Vepsäläinen H, Saloheimo T, et al. An international comparison of dietary patterns in 9–11-year-old children. *Int J Obes Suppl*. 2015;5(2):S17-S21. doi:10.1038/ijosup.2015.14

94. Saloheimo T, González SA, Erkkola M, et al. The reliability and validity of a short food frequency questionnaire among 9–11-year olds: a multinational study on three middle-income and high-income countries. *Int J Obes Suppl*. 2015;5(2):S22-S28. doi:10.1038/ijosup.2015.15

95. Poulain T, Spielau U, Vogel M, Körner A, Kiess W. CoCu: A new short questionnaire to evaluate diet composition and culture of eating in children and adolescents. *Clin Nutr*. 2019;38(6):2858-2865. doi:10.1016/j.clnu.2018.12.020

96. Ranjit N, Evans MH, Byrd-Williams C, Evans AE, Hoelscher DM. Dietary and activity correlates of sugar-sweetened beverage consumption among adolescents. *Pediatrics*. 2010;126(4):e754-61. doi:10.1542/peds.2010-1229

97. Rasmussen M, Pedersen TP, Johnsen NF, Krolner RF, Holstein BE. Persistent social inequality in low intake of vegetables among adolescents, 2002-2014. *Public Health Nutr*. 2018;21(9):1649-1653. doi:10.1017/S136898001800040X

98. Vereecken CA, Maes L. A Belgian study on the reliability and relative validity of the Health Behaviour in School-Aged Children food-frequency questionnaire. *Public Health Nutr*. 2003;6(6):581-588. doi:10.1079/PHN2003466

99. Vandevijvere S, Geelen A, Gonzalez-Gross M, et al. Evaluation of food and nutrient intake assessment using concentration biomarkers in European adolescents from the Healthy Lifestyle in Europe by Nutrition in Adolescence study. *Br J Nutr*. 2013;109(4):736-747. doi:10.1017/S0007114512002012

100. Schroder H, Benitez Arciniega A, Soler C, Covas MI, Baena-Diez JM, Marrugat J. Validity of two short screeners for diet quality in time-limited settings. *Public Health Nutr*. 2012;15(4):618-626. doi:10.1017/S1368980011001923

101. Shaw S, Crozier S, Strommer S, Inskip H, Barker M, Vogel C. Development of a short food frequency questionnaire to assess diet quality in UK adolescents using the National Diet and Nutrition Survey. *Nutr J*. 2021;20(1):5. doi:10.1186/s12937-020-00658-1

102. Switkowski KM, Kronsteiner-Gicevic S, Rifas-Shiman SL, Lightdale JR, Oken E. Evaluation of the Prime Diet Quality Score from Early Childhood Through Mid-Adolescence. *J Nutr*. 2024;154(6):1890-1906. doi:10.1016/j.tjnut.2024.04.014

103. Szczepańska E, Deka M, Calyniuk B. Studies to determine nutrition behaviour amongst middle school pupils living in the border areas of Poland and the Czech Republic. *Rocz Panstw Zakl Hig*. 2013;64(3):191-196.

104. Tugault-lafleur C, Desgreniers V, Besesette G, Al Kazzi R, Haines J. Development and content validity of a food intake screener for Canadian youth based on Canad'a Food Guide 2019 healthy eating recommendations. *Submitt Appl Physiol Nutr Metab*. Published online April 2025.

105. Vézina-Im LA, Beaulieu D, Turcotte S, et al. Validation d'un questionnaire mesurant la consommation de boissons sucrées, jus et eau auprès d'adolescents francophones. *Can J Diet Pract Res*. 2023;84(3):141-148. doi:10.3148/cjdpr-2023-002

106. Voltas N, Arija V, Aparicio E, Canals J. Longitudinal study of psychopathological, anthropometric and sociodemographic factors related to the level of Mediterranean diet adherence in a community sample of Spanish adolescents. *Public Health Nutr*. 2016;19(10):1812-1822. doi:10.1017/S1368980015003560

107. Washi SA, Ageib MB. Poor diet quality and food habits are related to impaired nutritional status in 13- to 18-year-old adolescents in Jeddah. *Nutr Res N Y N*. 2010;30(8):527-534. doi:10.1016/j.nutres.2010.07.002

108. Wong JE, Skidmore PML, Williams SM, Parnell WR. Healthy Dietary Habits Score as an Indicator of Diet Quality in New Zealand Adolescents. *J Nutr*. 2014;144(6):937-942. doi:10.3945/jn.113.188375

109. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. 2018;27(5):1159-1170. doi:10.1007/s11136-018-1829-0

110. Indu P, Vidhukumar K, Chacko D, Menon V, Grover S, Gupta S. Criterion validity, construct validity, and factor analysis: An introductory overview. *Indian J Psychiatry*. 2025;67(9):916-921. doi:10.4103/indianjpsychiatry_911_25

111. Hunsberger M, O'Malley J, Block T, Norris JC. Relative validation of Block Kids Food Screener for dietary assessment in children and adolescents. *Matern Child Nutr*. 2015;11(2):260-270. doi:10.1111/j.1740-8709.2012.00446.x

112. Sharma H. How short or long should be a questionnaire for any research? Researchers dilemma in deciding the appropriate questionnaire length. *Saudi J Anaesth*. 2022;16(1):65-68. doi:10.4103/sja.sja_163_21

113. Chudoba B. How Long Should A Survey Be? Insights And Best Practices. SurveyMonkey. Accessed April 22, 2026. https://www.surveymonkey.com/curiosity/survey_completion_times/

114. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol*. 2010;10:22. doi:10.1186/1471-2288-10-22

Portions of the writing were refined using Microsoft Copilot (Microsoft, 2026) and Grammarly (2026), which assisted with grammar, clarity, and organization. All conceptual decisions, interpretations, and final wording were made by the author.

# 9 Appendices

## 9.1 Appendix 1: Search strategies

**Medline**

Primary Search Conducted: Dec 1st, 2025

Link: Proposed Search Strategy

| # | Searches | Results |
|---|---|---|
| 1 | exp adolescent/ | 2368617 |
| 2 | (adolescen* or teen* or youth*).ti,ab,kf. | 528448 |
| 3 | 1 or 2 | 2517563 |
| 4 | "diet surveys"/ | 10204 |
| 5 | (diet* adj1 (screen* or questionnaire* or survey)).ti,ab,kf. | 3598 |
| 6 | (food* adj1 (screen* or questionnaire* or survey )).ti,ab,kf. | 1197 |
| 7 | 4 or 5 or 6 | 14102 |
| 8 | 3 AND 7 | 3590 |
| 9 **Scoping Review** | 8 AND 2010:2025.(sa_year) | **2050** |
| 10 | ("28608728" or "40944152").ui. | 2 |
| 11 | 9 AND 10 | 2 |
| 12 | exp "reproducibility of results"/ or exp "sensitivity and specificity"/ | 1047705 |
| 13 | (validit* or reliab* or psychomet* ).ti,ab,kf. | 912815 |
| 14 | 12 or 13 | 1735155 |
| 15 **Systematic Review** | 9 AND 14 | **280** |
| 16 | (“28673606” or “37739731”).ui. | 2 |
| 17 | 15 AND 16 | 2 |

**Embase**

Primary Search Conducted: Dec 1st, 2025

Link: Proposed Search Strategy

| # | Searches | Results |
|---|---|---|
| 1 | exp adolescent/ | 2180018 |
| 2 | (adolescen* or teen* or youth*).ti,ab,kf. | 698538 |
| 3 | 1 or 2 | 2370774 |
| 4 | "diet surveys"/ | 327589 |
| 5 | (diet* adj1 (screen* or questionnaire* or survey)).ti,ab,kf. | 5274 |
| 6 | (food* adj1 (screen* or questionnaire* or survey )).ti,ab,kf. | 1387 |
| 7 | 4 or 5 or 6 | 332040 |
| 8 | exp "reproducibility of results"/ or exp "sensitivity and specificity"/ | 824270 |
| 9 | (validit* or reliab* or psychomet* ).ti,ab,kf. | 1186624 |
| 10 | 8 or 9 | 1860244 |
| 11 | 3 AND 7 AND 10 AND 2010:2025.(sa_year) | **536** |

**CINAHL**

Primary Search Conducted: Dec 9th, 2025

| # | Searches | Results |
|---|---|---|
| S1 | (MH "Adolescent Nutrition") OR (MH "Adolescent Health") OR (MH "Adolescence+") | 655010 |
| S2 | XB (adolescen* or teen* or youth) | 252292 |
| S3 | S1 or S2 | 726343 |
| S4 | ((MH "Questionnaires+") OR (MH "Surveys+")) AND (MH "Diet+")) | 29917 |
| S5 | XB (diet* N1 (screen* or questionnaire* or survey)) | 3457 |
| S6 | XB (food* N1 (screen* or questionnaire* or survey)) | 10938 |
| S7 | S4 or S5 or S6 | 36537 |
| S8 | (MH "Reproducibility of Results") OR (MH "Sensitivity and Specificity") | 157868 |
| S9 | XB (validit* or reliab* or psychomet*) | 211107 |
| S10 | S8 or S9 | 330964 |
| S11 | S3 AND S7 AND S10 AND (DT20100101-20251231) | **436** |

## 9.2 Appendix 2: Inclusion and exclusion criteria

| CATEGORY | INCLUSION CRITERIA | EXCLUSION CRITERIA |
|---|---|---|
| **PARTICIPANTS** | Studies involving children and youth ≥10 to ≤17 years.<br><br>Studies with healthy participants. | Children <10 years or adults ≥18 years.<br><br>Studies that enroll participants diagnosed with a disease (including obesity) or hospitalized with an illness or injury. |
| **INTERVENTION/ EXPOSURE** | Screener can be completed independently (self-administered), with or without parental assistance.<br><br>Multifactorial screeners assessing more than a single food or nutrient.<br><br>Screener includes 30 items or less. | Screener is filled out by parents (proxy reporting).<br><br>Screener assess a single food (e.g. milk) or nutrient (e.g. fat intake).<br><br>>30 items. |
| **COMPARATOR (IF APPLICABLE)** | Twenty-four-hour dietary recalls, food frequency questionnaires, food records, or nutritional biomarkers (i.e., doubly labelled water, carotenoid) can be included as comparators to assess convergent validity | N/A |
| **OUTCOMES (SYSTEMATIC REVIEW ONLY)** | Screener must have been previously tested for validity or reliability | Screener has not yet been tested for validity or reliability |
| **TIMING** | Studies with a publication date from January 2010 to present date. | Studies with a publication date prior to January 2010. |
| **STUDY DESIGNS** | Randomized controlled trials (RCTs)<br><br>Non-randomized studies (e.g., quasi-RCTs)<br><br>Quasi experimental studies: controlled<br><br>before/after (CBA), interrupted-time series (ITS)<br><br>Observational studies: prospective and retrospective cohorts, case-control studies, cross-sectional studies based on survey data | Systematic reviews (meta-analyses)<br><br>Narrative reviews/Scoping reviews<br><br>Abstracts, conference proceedings<br><br>Pre-prints<br><br>Letters to the editor<br><br>Commentaries<br><br>Case reports<br><br>Uncontrolled trials or before/after studies<br><br>Qualitative studies |
| **SETTING** | Studies done in high-income countries as defined by the World Bank Organisation [35] | Studies done in middle or low-income countries |
| **OTHER** | Studies published in peer-review journals<br>Studies published in English or French | Screener does not include assessment of food intake/food choices (i.e., assessing food-related practices or eating behaviours like skipping breakfast, food literacy).<br>Articles that were not peer-reviewed<br>Studies published in another language other than English or French. |

## 9.3 Appendix 3: Table of quality assessment results determined from the COSMIN risk of bias checklist for food screeners designed for use with adolescents aged 10-17 years (n=26 screeners)

| Screener name | PROM development | Content validity | Structural Validity | Cross-cultural validity | Convergent validity | Construct (discriminative) validity | Internal consistency | Test re-test reliability | Overall rating |
|---|---|---|---|---|---|---|---|---|---|
| KIDMED 2.0[54] | Very good | Very good | N/A | N/A | Very good | N/A | N/A | Very good | Sufficient |
| Modified beverage and snack questionnaire [55] | Inadequate | N/A | N/A | N/A | Inadequate | N/A | N/A | N/A | Insufficient |
| NHANES dietary screener questionnaire[57] | Inadequate | N/A | N/A | N/A | Inadequate | N/A | N/A | N/A | Insufficient |
| KIDMED 2.0 Poland[59] | Very good | Adequate | N/A | N/A | N/A | Adequate | N/A | Adequate | Sufficient |
| Food behaviour checklist modified for children[60] | Inadequate | N/A | Very good | N/A | N/A | N/A | Very good | N/A | Inconsistent |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Breakfast eating habits questionnaire [61] | Doubtful | Doubtful | N/A | N/A | N/A | N/A | N/A | N/A | Insufficient |
| Youth risk behavior surveillance system questionnaire[62] | Inadequate | N/A | N/A | N/A | Very good | N/A | N/A | N/A | Inconsistent |
| Eating behaviour scale[63] | Inadequate | N/A | N/A | N/A | Very good | N/A | N/A | Very good | Inconsistent |
| Short FFQ[69] | Inadequate | N/A | N/A | N/A | Very good | N/A | N/A | Very good | Inconsistent |
| Beverage intake questionnaire[71] | Inadequate | N/A | N/A | Very good | Very good | N/A | N/A | Adequate | Sufficient |
| Fruit and vegetable questionnaire[74] | Inadequate | N/A | N/A | N/A | Inadequate | N/A | N/A | N/A | Insufficient |
| Short form FFQ 4 polish children[78] | Inadequate | N/A | N/A | N/A | N/A | N/A | N/A | Very good | Inconsistent |
| OPIC dietary intake questionnaire [80] | Inadequate | N/A | N/A | N/A | N/A | N/A | Doubtful | N/A | Insufficient |
| COMPASS (Cq) questionnaire[81] | Inadequate | N/A | N/A | N/A | Very good | N/A | N/A | Very good | Inconsistent |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Short food frequency questionnaire for Norwegian children [82] | Inadequate | N/A | N/A | N/A | Very good | Very good | N/A | N/A | Inconsistent |
| Dietary adherence tool[83] | Inadequate | Doubtful | Very good | N/A | N/A | N/A | N/A | N/A | Inconsistent |
| Questionnaire to assess snack consumption behaviours [84] | Inadequate | Doubtful | N/A | N/A | N/A | N/A | Adequate | Doubtful | Insufficient |
| ISCOLE FFQ[91–94] | Inadequate | N/A | N/A | Very good | Very good | N/A | N/A | Very good | Sufficient |
| HBSC questionnaire[97–99] | Inadequate | N/A | N/A | N/A | Very good | N/A | N/A | N/A | Inconsistent |
| CoCu FFQ[95] | Inadequate | N/A | N/A | N/A | Inadequate | N/A | N/A | Doubtful | Insufficient |
| NDNS short FFQ[101] | Adequate | Doubtful | N/A | N/A | Very good | N/A | N/A | N/A | Inconsistent |
| Foodstuffs questionnaire [103] | Very good | Adequate | N/A | N/A | N/A | N/A | N/A | N/A | Sufficient |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Canadian Food Intake Screener for Adolescents[104] | Very good | Very good | N/A | N/A | N/A | N/A | N/A | N/A | Sufficient |
| French adaptation of the Beverage intake questionnaire [105] | Very good | Very good | N/A | N/A | Very good | N/A | N/A | Very good | Sufficient |
| Jeddah FFQ [107] | Doubtful | Doubtful | N/A | N/A | N/A | N/A | N/A | N/A | Insufficient |
| Dietary habits questionnaire[108] | Adequate | Adequate | N/A | N/A | Inadequate | N/A | N/A | N/A | Inconsistent |

## 9.4 Appendix 4: Countries of origin of food screeners designed to measure food intake among adolescents aged 10-17 years (n = 47 screeners)

