

Identifying eating disorders in adolescents and adults with overweight or obesity: A systematic review of screening questionnaires

Eve T. House MND^{1,2}  | Natalie B. Lister PhD^{1,2}  | Anna L. Seidler PhD³  |
Haozhen Li MND⁴  | Wee Yee Ong MND⁴  | Caitlin M. McMaster PhD^{2,5}  |
Susan J. Paxton PhD⁶  | Hiba Jebeile PhD^{1,2} 

¹Institute of Endocrinology and Diabetes, The Children's Hospital at Westmead, Westmead, New South Wales, Australia

²Children's Hospital Westmead Clinical School, The University of Sydney, Westmead, New South Wales, Australia

³National Health and Medical Research Council Clinical Trials Centre, University of Sydney, Sydney, New South Wales, Australia

⁴Nutrition and Dietetics Group, School of Life and Environmental Sciences, Faculty of Science, The University of Sydney, Camperdown, New South Wales, Australia

⁵Boden Collaboration for Obesity, Nutrition, Exercise and Eating Disorders, Faculty of Medicine and Health, University of Sydney, Sydney, New South Wales, Australia

⁶School of Psychology and Public Health, La Trobe University, Melbourne, Victoria, Australia

Correspondence

Hiba Jebeile, Children's Hospital Westmead Clinical School, The University of Sydney, Westmead, NSW, Australia.
Email: hiba.jebeile@health.nsw.gov.au

Funding information

National Health and Medical Research Council, Grant/Award Numbers: GNT1128317, GNT1145748; University of Sydney

Action Editor: Kelly L. Klump

Abstract

Objective: This review aimed to examine the validity of self-report screening questionnaires for identifying eating disorder (ED) risk in adults and adolescents with overweight/obesity.

Method: Five databases were searched from inception to September 2020 for studies assessing validation of self-report ED screening questionnaires against diagnostic interviews in adolescents and adults with overweight/obesity. The review was registered with PROSPERO (https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=220013).

Results: Twenty-seven papers examining 15 questionnaires were included. Most studies validated questionnaires for adults (22 of 27 studies), and most questionnaires (12 of 15) screened for binge eating or binge-eating disorder (BED). The Eating Disorder Examination Questionnaire (sensitivity = .16–.88, specificity = .62–1.0) and Questionnaire on Eating and Weight Patterns (sensitivity = .07–1.0, specificity = .0–1.0) were most frequently validated (six studies each). Five studies of three questionnaires were in adolescents, with the Adolescent Binge-Eating Disorder Questionnaire having highest sensitivity (1.0) but lower specificity (.27). Questionnaires designed to screen for BED generally had higher diagnostic accuracy than those screening for EDs in general.

Discussion: Questionnaires have been well validated to identify BED in adults with overweight/obesity. Validated screening tools to identify other EDs in adults and any ED in adolescents with overweight/obesity are lacking. Thus, clinical assessment should inform the identification of patients with co-morbid EDs and overweight/obesity.

Public Significance: Individuals with overweight/obesity are at increased risk of EDs. This review highlights literature gaps regarding screening for ED risk in this vulnerable group. This work presents possibilities for improving care of individuals with

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *International Journal of Eating Disorders* published by Wiley Periodicals LLC.

overweight/obesity by reinventing ED screening tools to be better suited to diverse populations.

Abstract

Objetivo: Esta revisión tuvo como objetivo examinar la validez de los cuestionarios de detección de autorreporte para identificar el riesgo de trastorno de la conducta alimentaria (TCA) en adultos y adolescentes con sobrepeso/obesidad.

Método: Se realizaron búsquedas en cinco bases de datos desde su inicio hasta septiembre de 2020 para obtener estudios que evaluaran la validación de los cuestionarios de autorreporte de detección de TCA frente a entrevistas diagnósticas en adolescentes y adultos con sobrepeso/obesidad. La revisión se registró en PROSPERO (https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=220013).

Resultados: Se incluyeron veintisiete artículos que examinaron 15 cuestionarios. La mayoría de los estudios validaron cuestionarios para adultos (22 de 27 estudios) y la mayoría de los cuestionarios (12 de 15) detectaban atracones o trastorno por atracón (BED, en sus siglas en inglés). El Cuestionario de Examen de Trastornos alimentarios (Eating Disorder Examination Questionnaire) (sensibilidad = 0.16-0.88, especificidad = 0.62-1.0) y el Cuestionario sobre patrones de alimentación y peso (Questionnaire on Eating and Weight Patterns) (sensibilidad = 0.07-1.0, especificidad = 0.0-1.0) se validaron con mayor frecuencia (seis estudios cada uno). Cinco estudios de tres cuestionarios fueron en adolescentes, y el Cuestionario de Trastorno por Atracón en Adolescentes (Adolescent Binge-eating Disorder Questionnaire) tuvo la mayor sensibilidad (1,0) pero una menor especificidad (0,27). Los cuestionarios diseñados para detectar BED generalmente tuvieron una mayor precisión diagnóstica que los que detectaron TCA en general.

Discusión: Los cuestionarios han sido bien validados para identificar BED en adultos con sobrepeso/obesidad. Faltan herramientas de detección validadas para identificar otros TCA en adultos y cualquier tipo de trastorno de la conducta alimentaria en adolescentes con sobrepeso / obesidad. Por lo tanto, la evaluación clínica debe informar la identificación de pacientes con comorbilidad de TCA y sobrepeso/obesidad.

KEYWORDS

assessment, atypical anorexia nervosa, binge eating, binge-eating disorder, bulimia nervosa, diagnosis, disordered eating, Eating Disorder Examination Questionnaire, obesity, overweight

1 | INTRODUCTION

Obesity and eating disorders (EDs) are associated with significant morbidity and mortality (GBD Obesity Collaborators, 2017; van Hoeken & Hoek, 2020). Individuals with obesity are at increased risk of EDs possibly due to shared environmental (e.g., weight teasing, media use), cognitive (e.g., weight concern, body dissatisfaction), and behavioral (e.g., unhealthy weight control behaviors, loss of control eating) risk factors, with evidence of similar underlying genetic and biological mechanisms (Rancourt & McCullough, 2015).

Several epidemiological studies have found strong associations between EDs and overweight/obesity. Data from 14 countries were collected in the World Health Organisation World Mental Health Survey and examined to identify correlates of binge-eating disorder (BED) and bulimia nervosa (BN). It was found that there was a significantly larger proportion of people with obesity among those with a history of BN or BED than those without a clinical diagnosis of an ED (Kessler et al., 2013). Similarly, in a later study including over 12,000 US adults, individuals with obesity had higher lifetime prevalence of all EDs except anorexia nervosa (AN) compared with “normal/healthy” weight peers (Duncan et al., 2017). In another study of over

36,000 US adults, individuals with a history of BED were significantly more likely to have obesity compared to individuals without a history of ED (Udo & Grilo, 2018). Importantly, the prevalence of comorbid obesity and EDs is increasing faster than either condition alone (5.7-fold increase in obesity with comorbid recurrent binge eating and 8-fold increase in obesity with comorbid very strict dieting/fasting over a 20-year period) (Da Luz et al., 2017).

Individuals living with comorbid ED and obesity have poorer physiological and psychological health than those with obesity or an ED alone (Da Luz et al., 2018), presenting a challenge for the treatment of either condition. Early identification of EDs is essential to improve prognosis (Grange & Loeb, 2007). Obesity treatment guidelines highlight the need to screen for EDs as a routine part of care (Pfeiffle et al., 2019; Sherf Dagan et al., 2017), while ED treatment guidelines acknowledge the high prevalence of comorbid obesity and EDs (Hay et al., 2014). Thus, valid screening questionnaires that identify individuals at risk of ED in this population are imperative.

In the validation of screening questionnaires, sensitivity and specificity of tools are evaluated. Sensitivity refers to the probability that a person with an ED will be identified as having an ED by the questionnaire. Tools with high sensitivity will identify most individuals with EDs but may have a high false positive rate if they have lower specificity, for example, a tool with a sensitivity of .8 would correctly identify 80% of people with an ED but would have a 20% false negative rate. Specificity is the probability that a person without an ED will be correctly identified as not having an ED (Šimundić, 2009). The interpretation of diagnostic accuracy results and the desired level of sensitivity and specificity of tools is dependent on the context in which they will be used (Bossuyt et al., 2013). Highly sensitive screening questionnaires are essential in resource-limited clinical settings, to identify those individuals who may be at risk of an ED, such that appropriate clinical follow-up can be arranged to confirm or preclude the existence of a clinical ED (Warner, 2004).

Screening and assessment tools operationalize recognized ED risk factors (e.g., body dissatisfaction, dietary restraint, thin-ideal internalization) and clinical features (e.g., binge eating, overvaluation of weight and shape, purging behavior) to identify individuals who may be at risk of an ED. Historically, screening questionnaires to identify risk of any ED have been developed and validated in community or clinical samples with lower weights (Fairburn & Beglin, 1994; Morgan et al., 1999). Conceptually, it is possible that certain risk factors and clinical features incorporated into existing ED screening questionnaires may not have adequate specificity to identify patients with overweight/obesity who are at risk of EDs. For example, people with obesity have higher levels of body dissatisfaction than their lower weight peers (Moradi et al., 2022; Weinberger et al., 2016), thus, this criterion within an assessment tool needs to be able to adequately distinguish between individuals' perception of their weight and overvaluation of weight and shape. Similarly, dietary restraint occurs on a spectrum from flexible (gradual reduction, foods are limited in quantity rather than eliminated) to more rigid or extreme restraint (all-or-nothing mentality) (Schaumberg et al., 2016). In the context of obesity prevention or treatment, a reduction in grazing or snacking between

meals, energy-dense foods, non-hungry eating, and emotional eating, reflecting flexible restraint may be prescribed. Indeed a 2021 systematic review found that in the context of pediatric weight management, dietary restraint subscales tended to show a different direction of effect compared to other markers of ED risk, possibly indicating this was a measure of intervention adherence rather than eating pathology (House et al., 2021). Ideally, ED assessment tools would distinguish between flexible and rigid dietary restraint, particularly in the context of weight management (Ivezaj et al., 2021). In addition, differentiating between overeating (eating an unusually large amount of food) and binge eating (eating a large amount of food with loss of control) may be particularly important (Colles et al., 2008; Shomaker et al., 2010). Assessment of overeating within ED screening tools is usually in relation to peers and people with obesity, having higher energy requirements due to a larger body size, would be expected to eat a larger quantity of food to meet energy needs. Hence, it may be appropriate to focus on psychopathology related to eating patterns, such as loss of control or feelings of guilt and shame rather than quantity of food. Given that the development of some questionnaires may not have occurred in populations with overweight/obesity, it is unclear whether the way risk factors and clinical features are operationalized can adequately distinguish between individuals with obesity at risk of ED and those not at risk.

Previous systematic reviews examining the validity of ED screening questionnaires in community and clinical settings have included combined samples of individuals of any body size, with results not separated by body size (Burton et al., 2016; Kutz et al., 2020; Pursey et al., 2020). To date, no review has examined the validity of ED screening questionnaires for populations with overweight and obesity. It is important to identify which tools are most appropriate for people with overweight or obesity to inform appropriate use of these in clinical practice and research. Thus, this systematic review aimed to (1) identify validated self-report screening questionnaires for identifying ED in adults and adolescents with overweight or obesity and (2) to examine the validity of these questionnaires in this population.

2 | METHODS

Methodology was informed by Cochrane Diagnostic Test Accuracy (DTA) Reviews methodology (Deeks et al., 2013), reported according to Preferred Reported Items for Systematic Reviews and Meta-analyses (PRISMA) (Page et al., 2021). The review was registered with PROSPERO (https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=220013).

2.1 | Data sources and searches

A systematic search of five electronic databases—MEDLINE, EMBASE, CINAHL, PsycINFO, and PsycTests—was conducted from inception to September 2020. Search terms related to the Population (adolescent, adult), Index Test (survey, questionnaire), Reference Test

(clinical interview), and Diagnosis (EDs) (Table S1). Covidence online software (Veritas Health Innovation Ltd, Australia) was used to remove duplicates and screen results. Each paper was screened—first title and abstract, then full text—by two reviewers with conflicts resolved through discussion.

2.2 | Study selection

Studies that were conducted in adolescents (≥ 10 years) and adults with overweight or obesity, that is, body mass index (BMI) ≥ 25 kg/m² in adults; adult equivalent BMI ≥ 25 kg/m² (Cole et al., 2000), BMI z-score > 1 or BMI ≥ 85 th percentile in adolescents, in any setting were included. Studies with a subset of participants with overweight or obesity were included provided results were reported by weight status so that findings specific to these BMI groups could be identified. Studies that assessed the validation of a self-report ED screening questionnaire (index test) against a clinical diagnostic interview (reference standard) to identify risk of ED as defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM) including AN, BN, BED, and other specified feeding and eating disorder (OSFED), including night eating syndrome (NES), were included.

Articles published in English were included. Questionnaires and interviews conducted during or following an intervention, studies that used a reference standard that was not a clinical diagnostic interview, and studies that used any tool other than a self-report questionnaire as the index test were excluded.

2.3 | Data extraction and quality assessment

Data were extracted by one reviewer and cross-checked for accuracy. Data extracted included study setting, sample size, participant characteristics, screening questionnaires, reference standards, and validation outcomes (true positives [TP], false positives [FP], true negatives [TN], false negatives [FN], sensitivity [se], specificity [sp], negative predictive value [NPV], positive predictive value [PPV], area under the curve [AUC]). Where a two-by-two table was presented (including TP, FP, TN, and FN), se, sp, NPV, and PPV were calculated (Šimundić, 2009) (Table 1).

Study quality was assessed independently by two reviewers using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool (Reitsma et al., 2009) with discrepancies resolved through discussion. Nine of the 11 recommended items from the QUADAS (Reitsma et al., 2009) were included. Items 9 and 10—“Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?” and “Were uninterpretable/intermediate test results reported?” were removed as clinical data are not routinely incorporated in the interpretation of ED screening questionnaires and questionnaire scoring is unlikely to be uninterpretable. An additional item “Had test operators had appropriate training?” was included given its relevance to the reference standard (Reitsma et al., 2009). Thus, 10 items were included in the final quality assessment. Items were rated as “yes,” “no,” or “unclear” for each study.

TABLE 1 Diagnostic accuracy terms

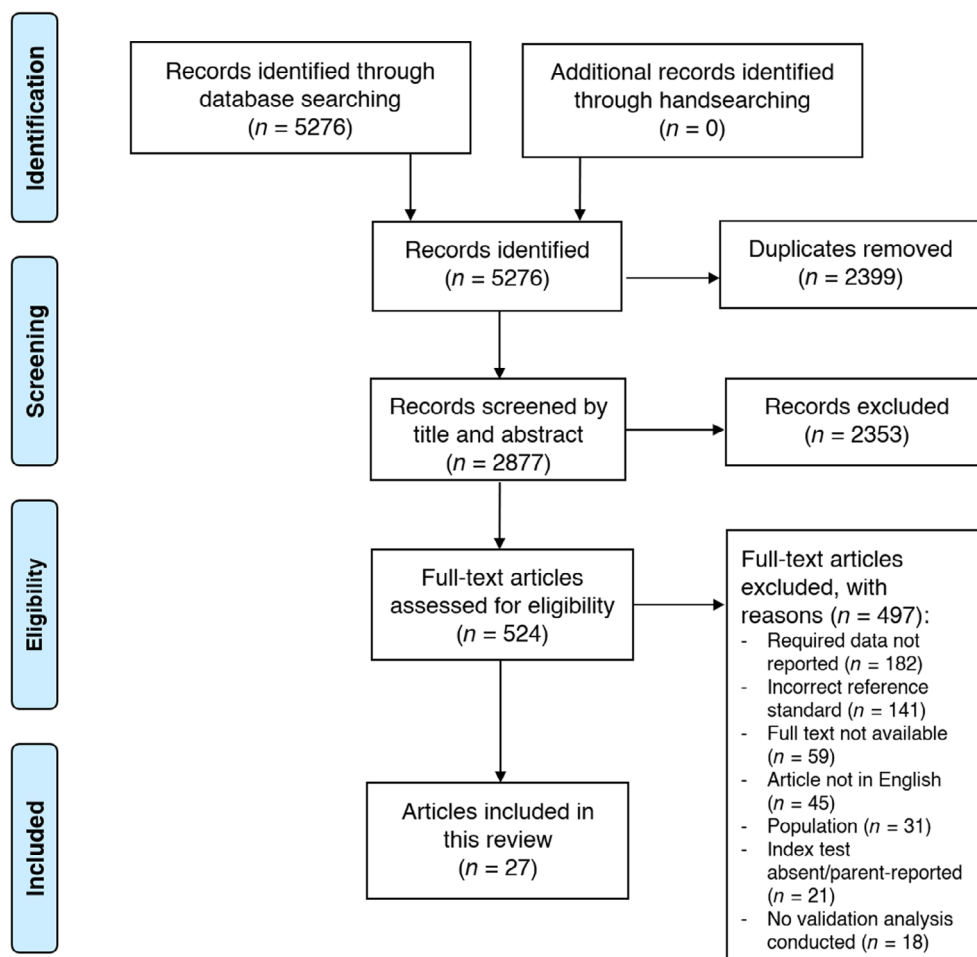
True positive (TP)	Individuals with the target condition who receive a positive screening questionnaire result
False positive (FP)	Individuals without the target condition who receive a positive screening questionnaire result
True negative (TN)	Individuals without the target condition who receive a negative screening questionnaire result
False negative (FN)	Individuals with the target condition who receive a negative screening questionnaire result
Sensitivity (se)	The “true positive rate,” that is, the probability that a person with the target condition (e.g., an eating disorder) will receive a positive screening questionnaire result. A sensitivity of 1.0 indicates that the screening questionnaire will identify all individuals with the target condition with no “false negative” results.
Specificity (sp)	The “true negative rate,” that is, probability that a person without the target condition will receive a negative screening questionnaire result. A specificity of 1.0 indicates that all individuals without the target condition will receive a negative screening questionnaire result with no “false positive” results.
Negative predictive value (NPV)	The proportion of people with a negative screening questionnaire result that do not have the target condition
Positive predictive value (PPV)	The proportion of people with a positive screening questionnaire result that have the target condition
Area under the curve (AUC)	AUC refers to the area under a receiver operating characteristics (ROC) curve. This provides a summary of the overall diagnostic accuracy of a test by combining sensitivity and specificity. An AUC of .5 indicates random chance that the test will correctly characterize patients, while an AUC of 1 indicates perfect diagnostic accuracy.

Source: Adapted from Šimundić (2009).

2.4 | Data synthesis and analysis

Due to the range of included tests and participant populations, a meta-analysis was not appropriate. Instead, a narrative synthesis of results was conducted. We adapted the Synthesis Without Meta-analysis (SWIM) guidelines to the reporting of diagnostic accuracy studies (Campbell et al., 2020). To account for heterogeneity between studies we reported findings by population—adolescents or adults—and index test examined. Findings were interpreted by assigning more weight to high-quality studies with large sample sizes, consistent with the SWIM guidelines (Item 4) studies considered to be at high risk of bias were assigned least weight (Campbell et al., 2020). Sensitivity

FIGURE 1 Preferred reporting items for systematic reviews and meta-analyses flow diagram of literature search and screening procedure



was prioritized over specificity, consistent with guidelines for screening tools in psychiatry (Warner, 2004). When screening for EDs minimizing “false negatives” was deemed to be important for accurately identifying individuals at higher risk of EDs who may then be flagged for further follow-up with a clinician to confirm or exclude an ED diagnosis.

3 | RESULTS

3.1 | Included studies

Of 5276 papers identified, 27 studies (Aardoom et al., 2012; Allison et al., 2008; Borges et al., 2005; Calugi et al., 2020; Chamay-Weber et al., 2017; De Man Lapidoth et al., 2007; De Zwaan et al., 1993; Decaluwé & Braet, 2004; Dymek-Valentine et al., 2004; Franklin et al., 2019; Freitas et al., 2006; Goldschmidt et al., 2007; Goossens & Braet, 2010; Grupski et al., 2013; Hartmann et al., 2014; Herman et al., 2016; Kalarchian et al., 2000; Liu et al., 2015; Mond et al., 2008; Orlandi et al., 2005; Parker et al., 2016; Quilliot et al., 2019; Ricca et al., 2000; Solmi et al., 2015; Vander Wal et al., 2005; Vander Wal et al., 2011; Wever et al., 2018) were included (Figure 1). Characteristics of studies conducted in adults and adolescents are summarized in

Tables 2 and 3 respectively, categorized by index test (Bohn et al., 2008; Chamay-Weber et al., 2017; De Man Lapidoth et al., 2007; Decaluwé & Braet, 2004; Fairburn & Beglin, 1994; Franklin et al., 2019; Ghaderi & Scott, 2002; Goldschmidt et al., 2007; Gormally et al., 1982; Henderson & Freeman, 1987; Herman et al., 2016; Morgan et al., 1999; Shapiro et al., 2007; Spitzer et al., 1992; Thelen et al., 1991; Vander Wal et al., 2005; Wadden & Foster, 2006; Wever et al., 2018). The diagnostic accuracy of index tests is summarized in Table 4. Overall, 12 index tests were validated in adults (Table 2) and 3 in adolescents (Table 3). Responses to individual quality assessment items for each study, categorized by index test can be found in Table S2. Performance against each quality assessment item is shown in Figure S1.

3.1.1 | Adults

Eating Disorder Examination Questionnaire

Six studies (range 41–433 participants) (Aardoom et al., 2012; Hartmann et al., 2014; Kalarchian et al., 2000; Mond et al., 2008; Parker et al., 2016; Vander Wal et al., 2011) validated the Eating Disorder Examination Questionnaire (EDE-Q) (Fairburn & Beglin, 1994). Four were conducted in the United States (Hartmann et al., 2014; Kalarchian

TABLE 2 Diagnostic accuracy of eating disorder screening questionnaires for adults with overweight and obesity, studies are presented in descending order of quality and then sample size

Author, year, country, setting, sample size (n), quality	Age (mean [SD] years); BMI (mean [SD] kg/m ²); sex (% female); race/ethnicity ^a ; SES	Index test version, cut-point, reference standard, diagnosis	Diagnostic accuracy		
			PPV/NPV	Sensitivity/specificity	Other measures of diagnostic accuracy
<i>Eating Disorder Examination Questionnaire (EDE-Q)</i>					
Questionnaire adapted from the EDE, assessing behavioral components of disordered eating across four domains: dietary restraint, eating concern, shape concern, and weight concern. Higher scores indicate greater psychopathology (Fairburn & Beglin, 1994).					
Hartmann et al., 2014, ^b USA, Clinical (hospital weight management center), n = 100 Quality: 9 of 10 positive ratings	Age 45.8 (12.0); BMI 41.9 (9.1); sex 72% female; race/ethnicity 81% White, 89% non-Hispanic/non-Latino; SES NR	EDE-Q version 6.0, score ≥4 on Q1, 2 or 3, Q10, Q11, 22 or 23, Q16, 17 or 18, and score = 0 on Q13 and 14, diagnostic interview for DSM-5 feeding and eating disorders, atypical AN	NR	NR	Prevalence of atypical AN: interview 0%, EDE-Q 15%
Kalarchian et al., 2000, USA, Clinical (gastric bypass, pre-operative), n = 98 Quality: 8 of 10 positive ratings	Age 38.4 (10.7); BMI 52.2 (10.1); sex 78% female; race/ethnicity 89% White, 9% Black, 2% Hispanic; SES NR	38-item EDE-Q, 2 binges per week, EDE (12th edition), BE	.57/.88	.59/.86	NR
Aardoom et al., 2012, The Netherlands, Clinical (outpatient obesity program and day stay patients), n = 433 Quality: 7 of 10 positive ratings	Group 1 (Obesity, no BED) (n = 321) Age 41.1 (12.4); BMI 41.9 (8.0); SES education level 38.9% low, 40.2% intermediate, 20.9% high Group 2 (BED) (n = 112) Age 36.7 (10.3); BMI 38.7 (7.8); SES education level 30.2% low, 43.4% intermediate, 26.4% high Sex 100% female; race/ethnicity 93% Caucasian, 7% not reported	36-item EDE-Q, cut-point not specified, standardized semi-structured interview based on DSM-IV, BED	NR	NR	AUC .72 (95% CI .67–.77)
Mond et al., 2008, ^b USA, Community, n = 147 (257 participants completed the EDE-Q and were invited to participate in an interview, 147 accepted) Quality: 7 of 10 positive ratings	Age 27.6 (6.5); BMI 28.1 (7.2); sex 100% female; race/ethnicity 88% Caucasian, 5% Native American, 3% African American, 2% Hispanic, 1% Asian; SES NR	22-item EDE-Q (excluded behavioral questions), 3.1, EDE, ED (AN, BN, BED, EDNOS—DSM-IV criteria)	.42/NR	.77/.73	AUC .84
Parker et al., 2016, ^{b,c} Australia, Clinical (bariatric surgery), n = 117 (405 participants were enrolled in the larger study, with 117 returning the completed EDE-Q and participating in the EDE) Quality: 7 of 10 positive ratings	Age 43.8 (11.6); BMI 42.5 (7.4); sex 79% female; race/ethnicity NR; SES NR	28-item EDE-Q, cut-point not specified, EDE (16.0), BE	NR	OBE .88/.62 SBE .16/.89 Subthreshold BE .20/.85 No disordered eating .46/.91	NR
Vander Wal et al., 2011, ^b USA, Community & University, n = 41	Age 52.0 (12.1); BMI 39.1 (9.3); sex 71% female; race/ethnicity 76%	39-item EDE-Q, restraint subscale ≥2.3–2.4 eating concern subscale ≥3.2	Restraint subscale .37/.64	Restraint subscale .40/.62	NR

TABLE 2 (Continued)

Author, year, country, setting, sample size (n), quality	Age (mean [SD] years); BMI (mean [SD] kg/m ²); sex (% female); race/ethnicity ^a ; SES	Index test version, cut-point, reference standard, diagnosis	Diagnostic accuracy		
			PPV/NPV	Sensitivity/specificity	Other measures of diagnostic accuracy
Quality: 6 of 10 positive ratings	Caucasian/White, 20% African American/Black, 7% Hispanic/Latino/Latina; SES NR	weight concern subscale ≥ 3.5 shape concern subscale ≥ 3.8 global score ≥ 3.2 –3.3, EDE (12.0D), BED	Eating concern subscale 1.0/.93 Weight concern subscale .68/.91 Shape concern subscale .59/.89 Global score .69/.84	Eating concern subscale .87/1.0 Weight concern subscale .87/.77 Shape concern subscale .87/.63 Global score .73/.81	
Questionnaire on Eating and Weight Patterns (QEWP/QEWP-R)					
28-item self-report questionnaire assessing binge eating and weight control behaviors (Spitzer et al., 1992). Positive screening result for BED if report ≥ 2 binge-eating episodes without compensatory behaviors per week and significant distress associated with binge eating (Spitzer et al., 1992).					
De Zwaan et al., 1993, USA/Austria, Clinical (not further specified), n = 100 Quality: 10 of 10 positive ratings	Age 39.2 (NR); BMI 35.9 (NR); sex 100% female; race/ethnicity NR; SES NR	QEWP (1992), cut-point not specified, SCID (DSM-III-R version with proposed DSM-IV criteria for BED), BED	.78/.80	.72/.84	NR
Borges et al., 2005, Brazil, Clinical (treatment seeking for weight loss/ binge eating), n = 89 Quality: 10 of 10 positive ratings	Participants with BED Age 35.2 (10.8); BMI 39.2 (11.8) Participants without BED Age 36.7 (11.8); BMI 37.9 (6.6) Sex 100% female; race/ethnicity 80% White; SES NR	QEWP-R (Portuguese version), cut-point not specified, SCID-I/P, BED and BE	BED .79/.56 BE .82/.73	BED .55/.80 BE .88/.63	NR
Hartmann et al., 2014, ^b USA, Clinical (hospital weight management center), n = 100 Quality: 9 of 10 positive ratings	Age 45.8 (12.0); BMI 41.9 (9.1); sex 72% female; race/ethnicity 81% White, 89% non-Hispanic/non-Latino; SES NR	28-item QEWP-R, BN = score = 1 on Q10, 11 and 12 with frequency ≥ 3 , score ≥ 3 on Q17, Score = 1 on Q18, 19, 20, 21, 22, or 23 with frequency ≥ 2 on Q18a, 19a, 20a, 21a, 22a, or 23a; BED = score = 1 on Q10, 11 and 12 with frequency ≥ 3 , score ≥ 3 on Q13, score ≥ 2 on Q15 or 16; subthreshold BN = criteria for BN with lower frequency (Q12 frequency of 1 or 2); subthreshold BED = criteria for BED with lower frequency (Q12 frequency of 1 or 2); purging disorder = score = 1 on Q18, 19 or 20 and frequency ≥ 2 on Q18a, 19a, 20a score ≥ 3 on Q17 and score = 0 on Q10 and	NR	BN/BED (DSM-IV) .21/.97 BN (DSM-IV) .50/.00 BED (DSM-IV) .71/.96 BN/BED (DSM-5) .5–.56/.98–1.00 Purging disorder (DSM-5) 1.0/.01 Subthreshold BED (DSM-5) NR/1.0	AUC BN/BED (DSM-IV) .59 BN (DSM-IV) .75 BED (DSM-IV) .84 <i>Prevalence of ED</i> BN (DSM-IV): 2% interview; 1% QEWP-R BED (DSM-IV): 9% interview; 7% QEWP-R BN (DSM-5): 2% interview; 1% QEWP-R BED (DSM-5): 9% interview; 7% QEWP-R Subthreshold BED (DSM-5): 5% interview; 1% QEWP-R Subthreshold BN (DSM-5):

(Continues)

TABLE 2 (Continued)

Author, year, country, setting, sample size (n), quality	Age (mean [SD] years); BMI (mean [SD] kg/m ²); sex (% female); race/ethnicity ^a ; SES	Index test version, cut-point, reference standard, diagnosis	Diagnostic accuracy		Other measures of diagnostic accuracy
			PPV/NPV	Sensitivity/specificity	
		11, SCID-IV and diagnostic Interview for DSM-5 feeding and eating disorders, BN, BED and OSFED (subthreshold BED, subthreshold BN, purging disorder)			2% interview; 0% QEWP-R Purging disorder (DSM-5): 1% interview; 1% QEWP-R
Calugi et al., 2020, Italy, Clinical (rehabilitative residential treatment program), n = 604 Quality: 8 of 10 positive ratings	Age 53.1 (12.4); BMI 41.6 (7.5); sex 65% female; race/ethnicity NR; SES NR	QEWP-5 (Italian version), cut-point not specified, EDE (Italian version), BED and BE	BED .34/.96 OBE .48/.90 SBE .59/.87	BED .49/.93 OBE .69/.79 SBE .63/.86	NR
Dymek-Valentine et al., 2004, USA, Clinical (gastric bypass clinic, pre-operative), n = 168 Quality: 8 of 10 positive ratings	Age 39.5 (9.3); BMI 50.8 (9.2); sex 85% female; race/ethnicity 59% Caucasian, 26% African American, 11% Hispanic, 4% other; SES education level NR, no difference between education level of those with and without BED	28-item QEWP-R, cut-point not specified, ED-SCID (DSM-IV), BED and partial BED	BED .40/.94 partial BED .55/.88	BED .73/.80 partial BED .79/.70	NR
Parker et al., 2016, ^{a,b} Australia, Clinical (bariatric surgery), n = 101 (405 participants were enrolled in the larger study, with 101 returning the completed QEWP-R and participating in the EDE) Quality: 7 of 10 positive ratings NR	Age 43.8 (11.6); BMI 42.5 (7.4); sex 79% female; race/ethnicity NR; SES NR	28-item QEWP-R, cut-point not specified, EDE (16.0), BE	NR	OBE .71/.64 SBE .25/.86	Subthreshold BE .07/.88 No disordered eating .44/.79
<i>Binge-Eating Scale (BES)</i> 16-item self-report tool used to measure binge-eating behaviors, with higher scores indicating more severe binge-eating symptoms. A score of more than 17 indicates mild to moderate binge eating and score of 27 or above indicates severe binge eating (Gormally et al., 1982).					
Freitas et al., 2006, Brazil, Clinical (participants in clinical trial of obesity treatment), n = 178 Quality: 10 of 10 positive ratings	Age 36.4 (10.0); BMI 36.3 (3.8); sex 100% female; race/ethnicity NR; SES educational years 74.7% participants with >12 years schooling	BES (Portuguese version), 17, SCID-I/P (DSM-IV), BED	.67/.95	.98/.48	NR
Quilliot et al., 2019, France, Clinical (bariatric surgery, pre-operative), n = 340 (1484 patients were enrolled in this study and completed the clinical interview, the first 340 patients to present for bariatric surgery at the	Age 41.7 (11.2); BMI 46.3 (7.4); sex 78% female; race/ethnicity NR; SES NR	BES (French version), 17, SCID-I/P (DSM-IV), BED	.55/.71	.51/.75	NR

TABLE 2 (Continued)

Author, year, country, setting, sample size (n), quality	Age (mean [SD] years); BMI (mean [SD] kg/m ²); sex (% female); race/ethnicity ^a ; SES	Index test version, cut-point, reference standard, diagnosis	Diagnostic accuracy		
			PPV/NPV	Sensitivity/specificity	Other measures of diagnostic accuracy
recruitment site were asked to complete the BES) Quality: 9 of 10 positive ratings					
Grupski et al., 2013, USA, Clinical (gastric bypass, pre-operative), n = 473 Quality: 8 of 10 positive ratings	Age 41.7 (10.4); BMI 50.5 (9.2); sex 85% female; race/ethnicity NR; SES educational years –13.7 (2.5) (range 7–25 years)	BES (not further specified), 17 and 27, SCID (for BED), BED	Score >17 .37/.99 Score ≥27 .56/.91	Score >17 .94/.76 Score ≥27 .37/.96	NR
Ricca et al., 2000, ^b Italy, Clinical (outpatient obesity treatment), n = 344 Quality: 7 of 10 positive ratings	Age 43.5 (13.6); BMI 35.8 (6.1); sex 83% female; race/ethnicity NR; SES NR	16-item BES, 17 and 27, SCID (DSM-III-R, with DSM-IV BED criteria), BED	Score ≥17 .26/.98; score ≥27 .57/.96	Score ≥17 .85/.75; score ≥27 .61/.95	NR
Sick Control One Fat Food (SCOFF) Brief screening questionnaire consisting of five “yes/no” questions which can usually be completed in 30 s. A positive response (yes) to ≥2 questions is generally used to identify risk of ED (Mond et al., 2008; Morgan et al., 1999).					
Liu et al., 2015, Taiwan, Clinical (psychiatric clinic), n = 178 Quality: 10 of 10 positive ratings	Age 30.5 (7.8); BMI ≥27; sex 49% female; race/ethnicity NR; SES educational years—men 14.5 (2.5), women 14.3 (2.6)	M-SCOFF (Mandarin Chinese version), 2 (males) 3 (females), SCID-I/P (DSM-IV), ED (AN, BN, BED, EDNOS)	NR	Males .67/.65 females .96/.83	AUC Males .66 females .90
Mond et al., 2008, ^b USA, Community, n = 147 (257 participants completed the SCOFF and were invited to participate in an interview, 147 accepted) Quality: 7 of 10 positive ratings	Age 27.6 (6.5); BMI 28.1 (7.2); sex 100% female; race/ethnicity 88% Caucasian, 5% Native American, 3% African American, 2% Hispanic, 1% Asian; SES NR	SCOFF (US version), 2, EDE, ED (AN, BN, BED, EDNOS—DSM-IV criteria)	.3/NR	.69/.59	AUC .72
Solmi et al., 2015, UK, Community, n = 63 Quality: 7 of 10 positive ratings	Age NR for population w/overweight/obesity; BMI overweight (n = 31), obesity (n = 32); sex 75% female (including those in HWR); race/ethnicity 57% White, 29% Black, 3% Asian, 11% other; SES education level (including those in HWR)—no qualification n = 14, GCSE/A-level n = 70, degree level or above n = 61	SCOFF, 2, SCID-I/NP (DSM-IV), ED	Overweight .43/1.0 obesity .6/ .88 overweight and obesity .50/.95	Overweight 1.0/.64 obesity .82/.70 overweight and obesity .90/.66	NR

Night Eating Questionnaire (NEQ) and Night Eating Question

The NEQ is 14-item tool to assess the severity of night eating syndrome and its psychological and behavioral symptoms. Possible scores range from 0 to 52 with higher scores indicated greater NES symptoms (Allison et al., 2008). Vander Wal et al. (2005) used a single *Night Eating Question* to determine entry into the study followed by administration of the first nine items of the NEQ, as they appear in the Weight and Lifestyle Inventory (Wadden & Foster, 2006). The question “To what extent does snacking after dinner contribute to your weight problem,” was scored on a 5-point scale, participants with a score ≥4 were recruited (Vander Wal et al., 2005).

(Continues)

TABLE 2 (Continued)

Author, year, country, setting, sample size (n), quality	Age (mean [SD] years); BMI (mean [SD] kg/m ²); sex (% female); race/ethnicity ^a ; SES	Index test version, cut-point, reference standard, diagnosis	Diagnostic accuracy		
			PPV/NPV	Sensitivity/specificity	Other measures of diagnostic accuracy
Hartmann et al., 2014, ^b USA, Clinical (hospital weight management center), n = 100 Quality: 9 of 10 positive ratings	Age 45.8 (12.0); BMI 41.9 (9.1); sex 72% female; race/ethnicity 81% White, 89% non-Hispanic/non-Latino; SES NR	14-item NEQ, 30, diagnostic interview for DSM-5 feeding and eating disorders, NES	NR	.01/NR	Prevalence of NES: 6% interview, 4% NEQ
Vander Wal et al., 2005, ^d USA, Clinical (RCT participants), n = 59 Quality: 4 of 10 positive ratings	Age 50.7 (10.4); BMI 35.3 (7.4); sex 76% female; race/ethnicity NR; SES NR	Night Eating Question and 9-item NEQ (from the WALI); 4 for Night Eating Question 5–10 for NEQ, structured clinical interview for NES, NES	Night Eating Question .16–.98/ .25–.87 NEQ .23–.95/1.0	Night Eating Question .80–.98/ .11–.67 NEQ 1.0/ .00–.40	NR
Allison et al., 2008, ^c USA, Clinical (bariatric surgery candidates, pre-operative), n = 194 Quality: 3 of 10 positive ratings	Age 44.0 (10.7); BMI 50.4 (8.0); sex 83% female; race/ethnicity 69% Caucasian, 25% African American, 6% other; SES NR	14-item NEQ, 25 and 30, NESHI, NES	Score ≥25 .41/.95 score ≥30 .73/.94	NR	NR
Bulimic Investigatory Test, Edinburgh (BITE)					
33-item self-rating scale to assess symptoms and severity of binge eating and purging behavior (Henderson & Freeman, 1987). The possible score ranged from 0 to 30 (Ricca et al., 2000). A score of 10–19 indicates abnormal eating behavior and score of 20 or above indicates highly disordered eating patterns and presence of binge eating.					
Ricca et al., 2000, ^b Italy, Clinical (outpatient obesity treatment), n = 344 Quality: 7 of 10 positive ratings	Age 43.5 (13.6); BMI 35.8 (6.1); sex 83% female; race/ethnicity NR; SES NR	BITE (1987), 10 and 20, SCID (DSM-IV criteria for BED), BED	Score ≥10 .72/.98 score ≥20 .31/.93	Score ≥10 .91/.51 score ≥20 .33/.92	NR
Orlandi et al., 2005, Italy, Clinical (outpatient obesity treatment), n = 388 (this study recruited 710 patients with obesity, however, the EDE was only conducted at one of the two recruitment sites, thus n = 388 participants completed the interview) Quality: 6 of 10 positive ratings	Age 40.3 (13.2); BMI 35.2 (6.0); sex 81% female; race/ethnicity NR; SES NR	BITE (Italian version), 10 and 20, EDE-12.0D, BED	Score ≥10 .14/.99 score ≥20 .30/.95	Score ≥10 .93/.55 score ≥20 .41/.92	NR
Clinical Impairment Assessment (CIA)					
The CIA is a 16-item questionnaire assessing psychosocial impairment associated with eating disorders across three domains—personal, cognitive, and social (Bohn et al., 2008).					
Hartmann et al., 2014, ^b USA, Clinical (hospital weight management center), n = 100 Quality: 9 of 10 positive ratings	Age 45.8 (12.0); BMI 41.9 (9.1); sex 72% female; race/ethnicity 81% White, 89% non-Hispanic/non-Latino; SES NR	16-item CIA, 16, SCID-IV and diagnostic interview for DSM-5 feeding and eating disorders, any ED (BN, BED, and OSFED) according to DSM-IV and DSM-5 criteria	ED (DSM-IV) .20/.94 ED (DSM-5) .51/.79	ED (DSM-IV) .64/.69 ED (DSM-5) .59/.74	NR

Combination of the EDE-Q, QEWP-R and NEQ

These questionnaires have been described above.

TABLE 2 (Continued)

Author, year, country, setting, sample size (n), quality	Age (mean [SD] years); BMI (mean [SD] kg/m ²); sex (% female); race/ethnicity ^a ; SES	Index test version, cut-point, reference standard, diagnosis	Diagnostic accuracy		
			PPV/NPV	Sensitivity/specificity	Other measures of diagnostic accuracy
Hartmann et al., 2014, ^b USA, Clinical (hospital weight management center), n = 100 Quality: 9 of 10 positive ratings	Age 45.8 (12.0); BMI 41.9 (9.1); sex 72% female; race/ethnicity 81% White, 89% non-Hispanic/non-Latino; SES NR	28-item QEWP-R and EDE-Q version 6.0 and 14-item NEQ, cut-points specified above, diagnostic interview for DSM-5 feeding and eating disorders, any ED (BN, BED and OSFED)	NR	.47/.78	AUC .62
<i>Binge-Eating Disorder Screener (BEDS)</i>					
7-item self-report screening tool to identify risk of BED. To receive a positive BED result, participants must answer “yes” to the first two questions relating to overeating and distress related to this. They must answer “sometimes,” “often,” or “always” to Q3-6 and “sometimes” or “rarely/never” to Q7 (related to purging) (Herman et al., 2016).					
Herman et al., 2016, USA, Community, n = 80 Quality: 9 of 10 positive ratings	True positives (n = 16, 1 with normal weight) Age 42.4 (12.0); BMI 36.3 (9.4) False positives (n = 46, 9 with normal weight) Age 36.4 (10.6); BMI 32.4 (8.4) True negatives (n = 29, 4 with normal weight) Age 47.0 (11.4); BMI 31.4 (7.9) Sex 60% female; race/ethnicity 68% White, 14% Black, 7% Asian, 6% Hispanic or Latino; 4% other; SES NR	BEDS-7, cut-point not specified, modified SCID-I/NP, BED	BMI 25–29.9 .2/.85 BMI 30–39.9 .27/.91 BMI ≥40 .45/1.0 BMI ≥25 .29/.89	BMI 25–29.9 .60/.48 BMI 30–39.9 .88/ .34 BMI ≥40 1.00/ .40 BMI ≥25 .83/.40	NR
<i>Eating Disorders in Obesity questionnaire (EDO)</i>					
11-item questionnaire designed to screen for DSM-IV ED (De Man Lapidoth et al., 2007). Adapted from the Survey for Eating Disorders (SED) (Ghaderi & Scott, 2002) for use in weight management therapy settings. Eight of 11 items only relevant to patients reporting binge eating. “Binge eaters” (BE) reported out-of-control objective binge-eating episodes. Participants that met criteria for BN, BED, or EDNOS were classified as ED.					
De Man Lapidoth et al., 2007, Sweden, Clinical (surgical and behavioral obesity clinics), n = 97 Quality: 6 of 10 positive ratings	Age 41.1 (10.6); BMI 44.2 (7.7); sex 72% female; race/ethnicity NR; SES NR	11-item EDO, cut-point not specified, EDE (AN) questions excluded, BED described as per DSM-IV), ED and BE	ED .90/.97 BE .72/.90	ED .75/.99 BE .82/.83	NR
<i>Risk Factors for Binge-Eating Disorder in Overweight (REO)</i>					
30-item screening questionnaire designed to identify individuals with obesity who are at risk of BED. Each item is rated on a scale of 1–5 (1 = never; 5 = almost always) with a total score ranging from 30 to 150, with higher scores indicating more symptoms of BED (Wever et al., 2018).					
Wever et al., 2018, ^c The Netherlands, Clinical (psychology and nutrition practices), n = 50 Quality: 6 of 10 positive ratings	No BED (n = 27) Age 53.2 (12.5); BMI 35.5 (6.5); sex 78% female; race/ethnicity NR; SES NR BED (n = 23) Age 37.0 (12.6); BMI 36.8 (6.7); sex 83% female; race/ethnicity NR; SES NR	30-item REO, 83.5, EDE not further specified, BED	.07/.03	.95/.82	AUC .89

Bulimia Test-Revised (BULIT-R)

36-item questionnaire (with 28 scored questions), originally designed to identify bulimic symptoms according to the DSM-III. Scores range from 28 to 140, with higher scores indicating more bulimic symptomatology (Thelen et al., 1991).

(Continues)

TABLE 2 (Continued)

Author, year, country, setting, sample size (n), quality	Age (mean [SD] years); BMI (mean [SD] kg/m ²); sex (% female); race/ethnicity ^a ; SES	Index test version, cut-point, reference standard, diagnosis	Diagnostic accuracy		
			PPV/NPV	Sensitivity/specificity	Other measures of diagnostic accuracy
Vander Wal et al., 2011, ^b USA, Community and University, n = 41 Quality: 6 of 10 positive ratings	Age 52.0 (12.1); BMI 39.1 (9.3); sex 71% female; race/ethnicity 20% African American/Black, 76% Caucasian/White, 7% Hispanic/Latino/Latina; SES education level—graduated high school/ GED n = 4, partial college n = 8, graduated 2-year college n = 7, graduated 4-year college n = 8, partial graduate/professional school n = 3, completed graduate/professional school n = 11	36-item BULIT-R, 80, EDE (12.0D), BED	.94/1.0	1.0/.96	NR
<i>The Binge-Eating Disorder Test (BEDT)</i> 23-item questionnaire that is derived from the items related to binge-eating disorder symptoms of the BULIT-R (Thelen et al., 1991).					
Vander Wal et al., 2011, ^b USA, Community and University, n = 41 Quality: 6 of 10 positive ratings	Age 52.0 (12.1); BMI 39.1 (9.3); sex 71% female; race/ethnicity 20% African American/Black, 76% Caucasian/White, 7% Hispanic/Latino/Latina; SES education level—graduated high school/ GED n = 4, partial college n = 8, graduated 2-year college n = 7, graduated 4-year college n = 8, partial graduate/professional school n = 3, completed graduate/professional school n = 11	23-item BEDT, 75, EDE (12.0D), BED	1.0/1.0	1.0/1.0	NR

Abbreviations: AN, anorexia nervosa; AUC, area under the curve; BE, binge eating; BED, binge-eating disorder; BEDS, Binge-Eating Disorder Screener; BEDS-7, 7-item Binge-Eating Disorder Screener; BEDT, Binge-Eating Disorder Test; BES, Binge-Eating Scale; BITE, Bulimic Investigatory Test, Edinburgh; BMI, body mass index; BN, bulimia nervosa; BULIT-R, Bulimia Test-Revised; CIA, Clinical Impairment Assessment; DSM-III, Diagnostic and Statistical Manual of Mental Disorders, 3rd edition; DSM-III-R, Diagnostic and Statistical Manual of Mental Disorders-Revised 3rd edition; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition; DSM-5, Diagnostic and Statistical Manual of Mental Disorders, 5th edition; ED, eating disorder; EDE, Eating Disorder Examination; EDE-12.0D, Eating Disorder Examination 12th Edition; EDE-Q, Eating Disorder Examination Questionnaire; EDO, Eating Disorders in Obesity Questionnaire; ED-SCID, eating disorder portion of the Structured Clinical Interview based on the DSM (SCID); EDNOS, eating disorder not otherwise specified; FN, false negative; FP, false positive; HWR, healthy weight range; M-SCOFF, Mandarin Chinese version of the Sick Control One Fat Food questionnaire; NES, night eating syndrome; NEQ, Night Eating Questionnaire; NESHI, Night Eating Syndrome History and Inventory; NPV, negative predictive value; NR, not reported; OBE, objective binge eating; OSFED, other specified feeding and eating disorder; PPV, positive predictive value; Q, question; QEWP, Questionnaire on Eating and Weight Patterns; QEWP-R, Questionnaire on Eating and Weight Patterns-Revised; RCT, randomized controlled trial; REO, Risk Factors for Binge-Eating Disorder in Overweight; SED, Survey for Eating Disorders; SBE, subjective binge eating; SCID, Structured Clinical Interview based on the DSM; SCID-I/NP, Structured Clinical Interview based on the DSM-non-patient edition; SCID-I/P, Structured Clinical Interview based on the DSM (patient edition); SCOFF, Sick Control One Fat Food questionnaire; SD, standard deviation; SES, socioeconomic status; TP, true positive; UK, United Kingdom; USA, United States of America; WALL, Weight and Lifestyle Inventory.

^aTerms used to describe race/ethnicity are those used by the authors in the original papers.

^bThese studies utilized multiple questionnaires, thus will be presented more than once in the table.

^cSome diagnostic accuracy results in these studies were converted from percentages to decimals for ease of comparison with other outcomes.

^dDiagnostic accuracy and cut-off values presented as a range as this varied dependent on which of six definitions of the Night Eating Syndrome were used: (a) eating more in the evening than any other time of day; (b) eating at least 50% of one's daily caloric intake after 7 p.m.; (c) eating more in the evening than any other time of day and no morning appetite; (d) eating at least 50% of one's daily caloric intake after 7 p.m. and no morning appetite; (e) eating more in the evening than any other time of day, no morning appetite, and a sleep disturbance; (f) eating at least 50% of one's daily caloric intake after 7 p.m., no morning appetite, and sleep disturbance.

et al., 2000; Mond et al., 2008; Vander Wal et al., 2011) and one each in the Netherlands (Aardoom et al., 2012) and Australia (Parker et al., 2016). Four studies were in clinical settings (Aardoom et al., 2012; Hartmann et al., 2014; Kalarchian et al., 2000; Parker et al., 2016), including two in bariatric surgery patients (Kalarchian et al., 2000; Parker et al., 2016), and two in community settings (Mond et al., 2008; Vander Wal et al., 2011). Mean age of participants ranged from 27.62 to 52 years, and BMI from 28.1 to 52.2 kg/m². Studies recruited samples with 72%–100% females and five studies reported ethnicity, with samples being predominantly White populations (Aardoom et al., 2012; Hartmann et al., 2014; Kalarchian et al., 2000; Mond et al., 2008; Vander Wal et al., 2011). One study received nine positive ratings (Hartmann et al., 2014), one study received eight positive ratings (Kalarchian et al., 2000), three studies received seven positive ratings (Aardoom et al., 2012; Mond et al., 2008; Parker et al., 2016) and one study received six positive ratings (Vander Wal et al., 2011), of 10 quality assessment items (Reitsma et al., 2009).

Diagnostic accuracy. Four (Kalarchian et al., 2000; Mond et al., 2008; Parker et al., 2016; Vander Wal et al., 2011) studies validated the EDE-Q against the Eating Disorder Examination Diagnostic Interview (EDE) (Fairburn & Cooper, 1993), one study used a standardized semi-structured interview based on DSM-IV (Aardoom et al., 2012) and one study used a diagnostic interview for DSM-5 feeding and EDs (Hartmann et al., 2014). The largest study ($n = 433$) by Aardoom et al. (quality rating of seven) (Aardoom et al., 2012) found an AUC of .72 when identifying patients with obesity with and without BED in outpatient weight management programs. Hartmann et al. (2014) received the highest quality rating and used the EDE-Q to identify atypical AN. In this study, the EDE-Q identified 15% of participants as having atypical AN, while no participants with atypical AN were identified with the diagnostic interview. Kalarchian et al. (2000) also received a high quality rating and examined the diagnostic accuracy of the EDE-Q to identify BED in 98 bariatric surgery candidates, finding moderate sensitivity ($se = .59$) and higher specificity ($sp = .86$). Results varied among remaining studies with sensitivity ranging from .16 when identifying subjective binge eating to .88 when identifying objective binge eating, and specificity ranging from .62 when identifying objective binge eating to .89 when identifying subjective binge eating.

Questionnaire on Eating and Weight Patterns

Six studies (89–604 participants) (Borges et al., 2005; Calugi et al., 2020; De Zwaan et al., 1993; Dymek-Valentine et al., 2004; Hartmann et al., 2014; Parker et al., 2016) validated the Questionnaire on Eating and Weight Patterns (QEWP) or QEWP-revised version (QEWP-R) (Spitzer et al., 1992), three in the USA (De Zwaan et al., 1993; Dymek-Valentine et al., 2004; Hartmann et al., 2014), one with a study site in Austria (De Zwaan et al., 1993), and one study each in Brazil (Borges et al., 2005), Italy (Calugi et al., 2020), and Australia (Parker et al., 2016). All studies were in clinical settings, two in bariatric surgery patients (Dymek-Valentine et al., 2004; Parker et al., 2016). Mean age ranged from 35.2 to 53.1 years, mean BMI

from 35.9 to 50.8 kg/m², all studies had $\geq 65\%$ female participants and the three studies reporting ethnicity included 59% (Dymek-Valentine et al., 2004), 80% (Borges et al., 2005), and 81% White participants (Hartmann et al., 2014). Studies received 10 (Borges et al., 2005; De Zwaan et al., 1993), 9 (Hartmann et al., 2014), 8 (Calugi et al., 2020; Dymek-Valentine et al., 2004), and 7 (Parker et al., 2016) positive ratings from 10 quality assessment items (Reitsma et al., 2009).

Diagnostic accuracy. Four studies (Borges et al., 2005; De Zwaan et al., 1993; Dymek-Valentine et al., 2004; Hartmann et al., 2014) validated the QEWP against the Structured Clinical Interview for DSM disorders (First & Gibbon, 2004), two (Calugi et al., 2020; Parker et al., 2016) against the EDE (Fairburn & Cooper, 1993), and one against a diagnostic interview for DSM-5 feeding and EDs (Hartmann et al., 2014) to identify BED and/or binge-eating behaviors. Hartmann et al. (2014) also validated the QEWP-R to identify BN against DSM-IV and DSM-5 criteria, and subthreshold BN, subthreshold BED, and purging disorder against DSM-5 diagnoses. The largest study, by Calugi et al. (2020), found the QEWP had higher specificity ($sp = .93$) when identifying BED than when used to identify objective and subjective binge-eating behaviors. Conversely, the QEWP had higher sensitivity ($se = .63$) when identifying objective and subjective binge-eating behaviors compared to when identifying BED (Calugi et al., 2020). The two highest quality studies (Borges et al., 2005; De Zwaan et al., 1993) found the QEWP had specificity $>.8$ and sensitivity $>.5$ to identify BED. In the remaining studies, sensitivity ($se = .21$ – $.73$) and specificity ($sp = .8$ – 1.00) varied by study quality and outcome (binge-eating behavior, BED, or BN) (Dymek-Valentine et al., 2004; Hartmann et al., 2014; Parker et al., 2016).

Binge Eating Scale

Four studies (178–473 participants) (Freitas et al., 2006; Grupski et al., 2013; Quilliot et al., 2019; Ricca et al., 2000) examined the Binge Eating Scale (BES) (Gormally et al., 1982), one each in Brazil (Freitas et al., 2006), United States (Grupski et al., 2013), France (Quilliot et al., 2019), and Italy (Ricca et al., 2000). All studies were in clinical settings, two in bariatric surgery candidates (Grupski et al., 2013; Quilliot et al., 2019). Mean age ranged from 36.4 to 43.5 years, BMI from 35.8 to 50.5 kg/m², $\geq 78\%$ females. One study each received 10 (Freitas et al., 2006), 9 (Quilliot et al., 2019), 8 (Grupski et al., 2013), and 7 (Ricca et al., 2000) positive ratings of 10 quality assessment items (Reitsma et al., 2009).

Diagnostic accuracy. All studies validated the BES against the Structured Clinical Interview for DSM disorders for diagnosis of BED (First & Gibbon, 2004). The two largest and high-quality studies ($n = 473$ and $n = 340$) (Grupski et al., 2013; Quilliot et al., 2019) compared the diagnostic accuracy of the BES at two cut-points, finding a cut-point of 17 yielded a higher sensitivity ($se \geq .85$) and a cut-point of 27 yielded a higher specificity ($sp \geq .95$). Across studies, at a cut-point of 17, sensitivity ranged from .51 to .98 and specificity from .48 to .76.

3.1.1.4 Sick, Control, One, Fat, Food questionnaire

Three studies (63–178 participants) (Liu et al., 2015; Mond et al., 2008; Solmi et al., 2015) examined the Sick, Control, One, Fat, Food (SCOFF) questionnaire (Morgan et al., 1999) to identify any ED, one study each in Taiwan (Liu et al., 2015), United States (Mond et al., 2008), and United Kingdom (Solmi et al., 2015). Two studies were in community settings (Mond et al., 2008; Solmi et al., 2015) and one in a psychiatric clinic (Liu et al., 2015). Mean age was 27.6 (Mond et al., 2008) and 30.5 years (Liu et al., 2015) in two studies and BMI 28.1 kg/m² in one study (Mond et al., 2008). Studies included 49%–100% females (Liu et al., 2015; Mond et al., 2008; Solmi et al., 2015) with two studies reporting predominantly White ethnicity (Mond et al., 2008; Solmi et al., 2015). One study received 10 positive ratings (Liu et al., 2015) and two (Mond et al., 2008; Solmi et al., 2015) received 7 positive ratings of 10 quality assessment items (Reitsma et al., 2009).

Diagnostic accuracy. Liu et al. (2015) conducted the largest and highest rated study, validating the SCOFF against the Structured Clinical Interview for DSM disorders (First & Gibbon, 2004) to identify any ED. They found the usual cut-point of two achieved the greatest diagnostic accuracy for males (AUC .66), while a cut-point of three was optimal for females (AUC .90) (Liu et al., 2015). Across studies, at the usual cut-point of two, sensitivity ranged from .67 to 1.0 and specificity from .59 to .7.

Night Eating Questionnaire

The Night Eating Questionnaire (NEQ) (Vander Wal et al., 2005) was validated in three studies (Allison et al., 2008; Hartmann et al., 2014; Vander Wal et al., 2005), with 194 (Allison et al., 2008), 100 (Hartmann et al., 2014), and 59 (Vander Wal et al., 2005) participants, conducted in the United States in clinical settings, one in bariatric surgery candidates (Allison et al., 2008). Studies were in predominantly female (Allison et al., 2008; Hartmann et al., 2014; Vander Wal et al., 2005) and White (Allison et al., 2008; Hartmann et al., 2014) samples. Studies had positive ratings on 9 (Hartmann et al., 2014), 4 (Vander Wal et al., 2005), and 3 (Allison et al., 2008) of 10 quality assessment items (Reitsma et al., 2009).

Diagnostic accuracy. Hartmann et al. (2014) examined the concordance between the NEQ and a diagnostic interview for DSM-5 feeding and EDs. They found that the NEQ identified four participants as having NES, however, all were diagnosed with different EDs using the diagnostic interview. Conversely, seven participants were diagnosed with NES using a diagnostic interview, none of which were identified using the NEQ.

Bulimia Investigatory Test, Edinburgh

The Bulimia Investigatory Test, Edinburgh (BITE) (Henderson & Freeman, 1987) was validated in two studies (Orlandi et al., 2005; Ricca et al., 2000) with 388 (Orlandi et al., 2005) and 344 (Ricca et al., 2000) predominantly female participants, conducted in Italy in outpatient weight management clinics. The studies received seven

(Ricca et al., 2000) and six (Orlandi et al., 2005) positive ratings of 10 quality assessment items (Reitsma et al., 2009).

Diagnostic accuracy. Studies examined the diagnostic accuracy of the BITE to identify BED at a cut-point of 10 and 20, against the EDE (Fairburn & Cooper, 1993; Orlandi et al., 2005) or the Structured Clinical Interview for DSM disorders (First & Gibbon, 2004; Ricca et al., 2000). Studies found sensitivity was better at a cut-point of 10 ($se = .91$ and $.93$) and specificity at a cut-point of 20 ($sp = .92$) (Orlandi et al., 2005; Ricca et al., 2000).

Screening questionnaires assessed in a single study

There were six questionnaires that were only validated in single studies (Clinical Impairment Assessment [CIA] [Bohn et al., 2008; Hartmann et al., 2014]; Binge-Eating Disorder Screener [BEDS] [Herman et al., 2016]; Eating Disorders in Obesity Questionnaire [EDO] [De Man Lapidoth et al., 2007]; Risk Factors for Binge-Eating Disorder in Overweight [REO] [Wever et al., 2018]; Bulimia Test-Revised [BULIT-R]; Binge-Eating Disorder Test [BEDT] [Thelen et al., 1991; Vander Wal et al., 2011]). One study (Hartmann et al., 2014) reported the validity of the CIA (Bohn et al., 2008) as an ED initial screening tool. The CIA was developed as a measure of psychological impairment that occurs due to ED psychopathology; however, it was included in this review as this study used the CIA (with a cut-off of 16) as an initial screening tool to identify participants with any ED. This study also reports the validity of the combination of the EDE-Q, QEWP-R, and NEQ to identify any ED (BN, BED, or OSFED). Three studies were conducted in the United States (Hartmann et al., 2014; Herman et al., 2016; Vander Wal et al., 2011) and one each in Sweden (De Man Lapidoth et al., 2007) and the Netherlands (Wever et al., 2018), in clinical (De Man Lapidoth et al., 2007; Hartmann et al., 2014; Wever et al., 2018) (including bariatric surgery patients for the EDO; De Man Lapidoth et al., 2007), community settings (Herman et al., 2016; Vander Wal et al., 2011), or university settings (Vander Wal et al., 2011). The sample size ranged from 41 to 100 predominantly female participants ($\geq 60\%$ female) and three studies reported a majority of White participants (Hartmann et al., 2014; Herman et al., 2016; Vander Wal et al., 2011). The CIA, combined questionnaires, and BEDS were validated in studies receiving nine positive ratings (Hartmann et al., 2014; Herman et al., 2016), and the EDO, REO, BULIT-R, and BEDT in studies receiving six positive ratings (De Man Lapidoth et al., 2007; Vander Wal et al., 2011; Wever et al., 2018) of 10 quality assessment items (Reitsma et al., 2009).

Diagnostic accuracy. The largest and highest quality study (Hartmann et al., 2014) examined the validity of the CIA as an ED screening tool. This study also reports the validity of the combination of the EDE-Q, QEWP-R, and NEQ to identify any ED using DSM-IV and DSM-5 criteria validated against the Structured Clinical Interview for DSM-IV disorders (First & Gibbon, 2004) and a diagnostic interview for DSM-5 feeding and EDs (Hartmann et al., 2014). They found that the CIA (Bohn et al., 2008) had moderate sensitivity ($se = .59$ – $.64$) and specificity ($sp = .69$ – $.74$) when used as an initial screening instrument for

TABLE 3 Diagnostic accuracy of eating disorder screening questionnaires for adolescents with overweight and obesity, studies are presented in descending order of quality and then sample size

Author, year, country, setting, sample size (n), quality	Age (mean [SD] years); BMI (mean [SD] kg/m ²); sex (% female); race/ethnicity ^a ; SES	Index test version, cut-point, reference standard, diagnosis	Diagnostic accuracy		
			PPV/NPV	Sensitivity/specificity	Other measures of diagnostic accuracy
<i>Children's Eating Disorder Examination Questionnaire and Youth Eating Disorder Examination Questionnaire (ChEDE-Q and YEDE-Q)</i>					
ChEDE-Q and YEDE-Q are adaptations of the EDE-Q for use in children and adolescents. Modifications include simplification of descriptive terms and addition of images and vignettes to help young people understand behaviors described in the questionnaire. Higher scores indicate greater psychopathology (Decaluwé & Braet, 2004; Goldschmidt et al., 2007).					
Goldschmidt et al., 2007, USA, Online, n = 35 Quality: 10 of 10 positive ratings	Age 13.8 (1.6); BMI 35.3 (7.1); sex 71% female; race/ethnicity 51% Caucasian, 40% African American, 3% Hispanic, 6% other; SES NR	YEDE-Q (adapted from EDE-Q 5.2), no cut-point specified, ChEDE 12.0, BE	1.0/.9	.57/1.0	NR
Decaluwé & Braet, 2004, Belgium, Clinical (inpatient obesity treatment), n = 139 Quality: 9 of 10 positive ratings	Age 12.8 (1.8); adjusted BMI 172.3 (25.6)%; sex 58% female; race/ethnicity NR; SES NR	ChEDE-Q (Dutch version), cut-point not specified, ChEDE, BE	.22/.97	.79/.68	NR
Goossens & Braet, 2010, Belgium, Clinical (inpatient weight management) and community (non-treatment seeking), n = 235 (a total of 429 participants were recruited, 235 completed the diagnostic interview) Quality: 5 of 10 positive ratings	Age 14.3 (1.5); adjusted BMI 160.6 (29.6)%; sex 56% female; race/ethnicity NR; SES NR	ChEDE-Q (Dutch version), cut-point not specified, ChEDE, disordered eating behaviors	NR	NR	% concordant/non-concordant cases Objective overeating 52.56/47.44 OBE 68.51/31.49 SBE 66.38/33.62 self-induced vomiting 85.96/14.04 laxative/diuretic use 84.68/15.32 appetite depressant 84.26/15.74 excessive exercise 68.94/31.06
<i>Adolescent Binge-Eating Disorder Questionnaire (ADO-BED)</i>					
French language, 10-item, questionnaire to assess binge-eating behavior in adolescents with obesity. Q1–3 are yes/no questions relating to binge-eating behavior, Q4 and 5 related to frequency of the behaviors and Q6 is a yes/no question relating to purging behavior (Chamay-Weber et al., 2017).					
Chamay-Weber et al., 2017, Switzerland, Clinical (pediatric obesity care center), n = 94 Quality: 10 of 10 positive ratings	Age median (range) = 14 (11–18); BMI all participants >97th percentile for age and sex; sex 60% female; race/ethnicity NR; SES NR	ADO-BED (French version), cut-point not specified (diagnostic accuracy reported by question), SCID (BED portion of French version), BED	Q1 .31/1.0 Q2 .36/.93 or 2 .28/1.0 Q3a .43/.75 Q3b .47/.75 Q3c .35/.78 Q3d .36/1.0 Q3e .44/.87 Q4 .65/.80 Q5 .33/.73 Q6 .29/.69	Q1 1.0/.36 Q2 .86/.56 or 2 1.0/.27 Q3a .43/.75 Q3b .38/.81 Q3c .76/.38 Q3d 1.0/.21 Q3e .81/.54 Q4 .52/.87 Q5 .81/.24 Q6 .33/.65	NR

(Continues)

TABLE 3 (Continued)

Author, year, country, setting, sample size (n), quality	Age (mean [SD] years); BMI (mean [SD] kg/m ²); sex (% female); race/ethnicity ^a ; SES	Index test version, cut-point, reference standard, diagnosis	Diagnostic accuracy		
			PPV/NPV	Sensitivity/specificity	Other measures of diagnostic accuracy
<i>Children's Brief Binge-Eating Questionnaire (CBBEQ)</i>					
7-item tool based on the Children's Binge-Eating Disorder Scale (C-BEDS) (Shapiro et al., 2007). The first six items consist of "yes/no" questions and the seventh question asks about the onset of behavior. A cut-off score ≥ 8 was used to identify individuals at risk of BED (Franklin et al., 2019).					
Franklin et al., 2019, USA, Clinical (obesity clinic), n = 70 Quality: 6 of 10 positive ratings	Age 13.6 (2.7); BMI 36.9 (8.5); sex 57% female; race/ethnicity 64% Caucasian, 27% African American, 8% other; 63% Hispanic race/ethnicity; SES insurance type 67% Medicaid, 19% private insurance, 14% Children's Health Insurance Program	CBBEQ (adapted from C-BEDS), 8, EDA (for DSM-5 BED), BED	NR	1.0/.93	PLR .33 NLR 1.0

Abbreviations: ADO-BED, Adolescent Binge-Eating Disorder Questionnaire; BE, binge eating; BED, binge-eating disorder; BMI, body mass index; CBBEQ, Children's Brief Binge-Eating Questionnaire; C-BEDS, Children's Binge Eating Disorder Scale; ChEDE, Children's Eating Disorder Examination; ChEDE-Q, Children's Eating Disorder Examination-Questionnaire; DSM-5, Diagnostic and Statistical Manual of Mental Disorders, 5th edition; EDA, Eating Disorder Assessment; EDE-Q, Eating Disorder Examination-Questionnaire; NLR, negative likelihood ratio; NPV, negative predictive value; NR, not reported; OBE, objective binge eating; PLR, positive likelihood ratio; PPV, positive predictive value; Q, question; SBE, subjective binge eating; SCID, Structured Clinical Interview based on the Diagnostic and Statistical Manual of Mental Disorders; SD, standard deviation; SES, socioeconomic status; USA, United States of America; YEDE-Q, Youth Eating Disorder Examination-Questionnaire.

^aTerms used to describe race/ethnicity are those used by the authors in the original papers.

EDs. The combination of questionnaires had lower sensitivity ($se = .47$) than specificity ($sp = .78$) when used to identify any ED (Hartmann et al., 2014). The high-quality study by Herman et al. (2016) examined the validity of the 7-item BEDS in participants with overweight and obesity against a modified version of the Structured Clinical Interview for DSM disorders (First & Gibbon, 2004). They found that the questionnaire yielded sensitivity and NPV of $\geq .83$ for individuals with BMI above 30 (Herman et al., 2016).

3.1.2 | Adolescents

Children's and Youth Eating Disorder Examination Questionnaires

Three studies (range 35–235 participants) validated adaptations of the EDE-Q—the Children's and Youth Eating Disorder Examination Questionnaires (ChEDE-Q and YEDE-Q) (Decaluwé & Braet, 2004; Goldschmidt et al., 2007; Goossens & Braet, 2010). While the ChEDE-Q and YEDE-Q appear to have been developed independently they were analyzed together as both are adaptations of the EDE-Q which were informed by the development of the CHEDE and retain the subscales used in the EDE-Q (Decaluwé & Braet, 2004; Goldschmidt et al., 2007; Goossens & Braet, 2010). Two studies were in Belgium (Decaluwé & Braet, 2004; Goossens & Braet, 2010) and one in United States (Goldschmidt et al., 2007), one in a clinical setting (Decaluwé & Braet, 2004), one in a combination of clinical (treatment-seeking) and community (non-treatment seeking) settings (Goossens & Braet, 2010), and one was online (Goldschmidt et al., 2007). Mean age of participants ranged from 12.8 to 14.3 years

and mean BMI categorized participants as having severe obesity. Studies included 56%–71% females and one study reported ethnicity of participants as 51% White and 40% African American (Goldschmidt et al., 2007). One study received 10 (Goldschmidt et al., 2007), one 9 (Decaluwé & Braet, 2004), and one 5 positive ratings (Goossens & Braet, 2010) of 10 quality assessment criteria (Reitsma et al., 2009).

Diagnostic accuracy. Studies validated the ChEDE-Q/YEDE-Q against the Children's Eating Disorder Examination interview (ChEDE) (Bryant-Waugh et al., 1996). The two highest rated studies found the ChEDE-Q yielded higher sensitivity ($se = .79$) and NPV (NPV = .97) (Decaluwé & Braet, 2004) than the YEDE-Q which yielded higher specificity ($sp = 1.0$) and PPV (PPV = 1.0) (Goldschmidt et al., 2007). The largest study (Goossens & Braet, 2010) reported disagreement between ChEDE-Q and ChEDE finding the questionnaire identified a higher proportion of objective overeating and a lower proportion of self-induced vomiting, laxative/diuretic misuse, and appetite depressant misuse. The two studies examining the ChEDE-Q reported different outcome measures (see Table 3).

Screening questionnaires assessed in a single study

The Adolescent Binge-Eating Disorder Questionnaire (ADO-BED) and Children's Brief Binge-Eating Questionnaire (CBBEQ) were validated in single studies, both studies were conducted in clinical settings with majority female participants (Chamay-Weber et al., 2017; Franklin et al., 2019), one study reported ethnicity (64% White participants) (Franklin et al., 2019). The study validating the ADO-BED received positive ratings on 10 (Chamay-Weber et al., 2017) and the study of

TABLE 4 A summary of the validation of eating disorder screening questionnaires in adolescents and adults with overweight/obesity separated by diagnosis

Questionnaire	Number of studies	Sample size (range of n)	Quality (range/10)	Diagnoses	Sensitivity/specificity by diagnosis
Adults					
Eating Disorder Examination Questionnaires (EDE-Q)	6	n = 41–433	6–9	BE, BED, ED, AAN	BE—.16–.88/.62–.89 BED—.40–.87/.62–1.0 ED—.77/.73 AAN—sensitivity and specificity NR
Questionnaires on Eating and Weight Patterns (QEWP/QEWP-revised)	6	n = 89–604	7–10	BE/subthreshold BED, BED, BN, purging disorder	BE/subthreshold BED—.07–.88/.63–1.0 BED—.49–.73/.80–.93 BN—.5/.0 Purging disorder—1.0/.01
Binge-Eating Scale (BES)	4	n = 178–473	7–10	BED	Cut-point of 17—.51–.98/.48–.76 Cut-point of 27—.37–.61/.95–.96
Sick Control One Fat Food (SCOFF)	3	n = 63–178	7–10	ED	.67–1.0/.59–.83
Night Eating Questionnaire (NEQ)	3	n = 59–194	3–9	NES	.01–1.0/.0–.4
Bulimic Investigatory Test, Edinburgh (BITE)	2	n = 344–388	6–7	BED	Cut-point of 10—.91–.93/.51–.55 Cut-point of 20—.33–.41/.92
CIA	1	n = 100	9	ED	.59/.74
Combination of QEWP-R + EDE-Q + NEQ	1	n = 100	9	ED	.47/.78
Binge-Eating Disorder Screener (BEDS)	1	n = 80	9	BED	.6–1.0/.34–.48 (dependent on BMI range)
Eating Disorders in Obesity questionnaire (EDO)	1	n = 97	6	ED, BE	ED—.75/.99 BE—.82/.83
Risk Factors for Binge-Eating Disorder in Overweight (REO)	1	n = 50	6	BED	.95/.82
Bulimia Test-Revised (BULIT-R)	1	n = 41	6	BED	1.0/.96
The Binge-Eating Disorder Test (BEDT)	1	n = 41	6	BED	1.0/1.0
Adolescents					
Children's/Youth Eating Disorder Examination Questionnaire (ChEDE-Q/YEDE-Q)	3	n = 35–235	5–10	BE, DEBs	BE—.57–.79/.68–1.0 DEBs—sensitivity and specificity NR
Adolescent Binge-Eating Disorder Questionnaire (ADO-BED)	1	n = 94	10	BED	1.0/.27 (if using first two questions for screening)
Children's Brief Binge-Eating Questionnaire (CBBEQ)	1	n = 70	6	BED	1.0/.93

Abbreviations: AAN, atypical anorexia nervosa; BE, binge eating; BED, binge-eating disorder; BMI, body mass index; BN, bulimia nervosa; DEBs, disordered eating behaviors; ED, eating disorder; NES, night eating syndrome; NR, not reported; OSFED, other specified feeding and eating disorder.

the CBBEQ on 6 (Franklin et al., 2019) of the 10 quality assessment items (Reitsma et al., 2009).

Diagnostic accuracy. The larger and higher quality study validated the ADO-BED (Chamay-Weber et al., 2017) against the Structured Clinical Interview for DSM disorders (First & Gibbon, 2004) to identify BED (Chamay-Weber et al., 2017). In this study, Questions 1 and

2 ask about cravings in the absence of hunger and loss of control overeating and were used as an initial screen for all participants. Questions 3–6 were completed only by participants with a positive answer to at least one of the first two questions. They found that questions 1 and 2 had high sensitivity ($se > .85$) but lower specificity ($sp \leq .56$) for the identification of adolescents at risk of BED (Chamay-Weber et al., 2017).

4 | DISCUSSION

This is the first systematic review to examine the validation of ED screening questionnaires in adolescents and adults with overweight and obesity. Results indicate that a wide range of index tests—15 different questionnaires—have been assessed in this population. The most extensively examined were the EDE-Q and QEWP in adults and only three questionnaires (adaptations of the EDE-Q, the ADO-BED, and CBBEQ) were tested in adolescents. Diagnostic accuracy varied widely between studies and ED diagnoses. Importantly, studies were primarily conducted in samples of White females highlighting a major evidence gap for assessing ED risk in males with obesity and culturally diverse communities, which may be disproportionately affected by obesity (Anekwe et al., 2020). Emerging research on questionnaires designed specifically for use in individuals with overweight or obesity and/or specifically designed to identify BED, suggests a higher degree of diagnostic accuracy compared to questionnaires without these foci and warrants further research.

4.1 | Adults

Due to inconsistency of evidence, it is not possible to recommend a single tool for ED screening in adults with obesity. The SCOFF was most extensively validated for the identification of any ED, showing moderate to high sensitivity (.67–1.00) at the usual cut-point of two (Liu et al., 2015; Mond et al., 2008; Solmi et al., 2015) and good accuracy in females at a cut-point of three (Liu et al., 2015). The sensitivity of the SCOFF in this population appears to be consistent with that found in a review of the diagnostic accuracy of the SCOFF in the general population where it was found to have a pooled sensitivity of .86 (95% CI .78–.91), while the specificity appeared to be slightly lower in this population with a range of .59–.83 compared to pooled specificity of .83 (95% CI .77–.88) (Kutz et al., 2020). However, in the previous review, sensitivity of the SCOFF was lower in locations with higher rates of obesity. It was noted that the SCOFF had much higher pooled sensitivity and specificity when used in case–control studies identifying AN and BN ($se = .96$, $sp = .89$) and that included studies did not reflect racial and weight diversity seen outside of AN and BN, suggesting the SCOFF may not perform as well when used to identify BED and OSFED (Kutz et al., 2020). The EDO, a tool designed as a simple screener to identify EDs in a weight management setting, had a sensitivity of .75 in a single study (De Man Lapidoth et al., 2007) and may warrant further research updating to meet DSM-5 criteria.

The EDE-Q, the most commonly recommended screening tool in primary care (Mond et al., 2008), only had moderate accuracy in adults with obesity. Only one of the included studies examined the diagnostic accuracy of the EDE-Q to identify any ED and reported results by BMI category (Mond et al., 2008). In this study, the EDE-Q had higher sensitivity in people with “normal weight” compared to those with “overweight” ($se = .83$ compared with .77) with comparable specificity between these groups ($sp = .72$ compared with .73). The EDE-Q was also found to have poor concordance with the DSM-5 interview

in identifying atypical AN in adults undergoing weight management (Hartmann et al., 2014). It has been suggested that a higher cut-point on the EDE-Q be used for individuals with a high BMI and to identify OSFED (Mond et al., 2008; Rø et al., 2015) and that an additional question designed to capture significant weight loss be added to distinguish between a healthful desire to lose weight and disordered eating pathology (Hartmann et al., 2014). Of note, studies validating the commonly used Eating Attitudes Test (EAT-26) (Galmiche et al., 2019; Garner et al., 1982; Maïano et al., 2013) were not included in this review. We identified two studies validating the EAT-26 in samples with a mean BMI >30; however, they included participants within a broad BMI range and did not report data by BMI category (Orbitello et al., 2006; Siervo et al., 2005). These studies found the diagnostic accuracy of the EAT-26 varied based on the cut-point used (Orbitello et al., 2006) and that a lower cut-point than is generally used is required to optimize the balance between sensitivity and specificity, an important implication of these findings is that prevalence estimates of ED in adults with obesity may be inaccurate if determined based on a screening questionnaire alone.

Screening questionnaires have been more extensively validated to identify BED (12 of 15 questionnaires), with the BES having a high sensitivity overall at a cut-point of 17 (three of four studies finding sensitivity $\geq .85$) (Freitas et al., 2006; Grupski et al., 2013; Quilliot et al., 2019; Ricca et al., 2000) and the QEWP-R having moderate to high sensitivity and specificity in identifying BED, appearing to be preferable to the EDE-Q. Several newer tools designed to identify BED (e.g., REO; Wever et al., 2018), had promising results, warranting further research. The focus on BED may be influenced by the high prevalence of overweight or obesity among populations with this diagnosis (Duncan et al., 2017; Hay et al., 2015). However, it also highlights a lack of evidence for screening for BN and OSFED such as atypical AN and NES. This is particularly important as OSFED has a higher lifetime prevalence than any other ED (Micali et al., 2017), and individuals with obesity are more likely to experience not only BED but also BN and OSFED (Hay et al., 2015). Validation of screening tools to identify BN and OSFED in populations with overweight and obesity is an important area for future research, given the significant physical and psychological impacts when these go undiagnosed (Grange & Loeb, 2007).

4.2 | Adolescents

Studies validating ED screening questionnaires in adolescents with obesity are limited, with only five eligible studies identified. Importantly, a validated screening tool to identify any ED is not currently available. While adolescent or youth adaptations of the EDE-Q are generally recommended to screen for any ED, to our knowledge, variations in cut-points that may be required have not been tested. This highlights a significant gap in the literature considering the potential shortfalls with the EDE-Q identified with adult populations and is concerning given that adolescence is a period of increased risk for the development of ED (Hudson et al., 2007). We identified two tools

(ADO-BED and CBBEQ), tested in one study each, designed to identify BED in youth with obesity, which showed high sensitivity (Chamay-Weber et al., 2017; Franklin et al., 2019). Based on these findings, identification of risk of any ED using questionnaires alone may not be appropriate for adolescents with obesity. Furthermore, there are reports detailing cases of atypical AN among adolescents with a history of obesity (Matthews et al., 2019; Wolter et al., 2009). Often atypical AN has a longer treatment delay than other EDs due to “normal” or high weight status (Matthews et al., 2019; Sawyer et al., 2016). However, the psychological and physical symptoms of those presenting to treatment are no less severe than AN (Matthews et al., 2019; Sawyer et al., 2016). Considering the high-risk life stage of adolescence, the development of screening tools is required to identify a range of EDs.

4.3 | Recommendations for future research

Overall, our findings indicate that questionnaires that are not tailored to individuals with obesity may include questions that are less relevant to this group. Consistent with the 2021 systematic review which found dietary restraint may not be a useful measure of ED risk among adolescents engaged in weight management (House et al., 2021), within this review, the study by Vander Wal et al. (2011) found that the dietary restraint subscale of the EDE-Q had a lower sensitivity and specificity than the other subscales in a community sample with overweight/obesity. Further, another study identified in our search but not included in our review reported higher scores on the dietary restraint and eating concern subscales of the EDE-Q compared to the EDE interview in patients' post-bariatric surgery (de Zwaan et al., 2004). Interestingly, a new screening tool aiming to identify loss of control and binge eating specifically (Manasse et al., 2021), found that items included in the screening tool based on the DSM-5 criteria for BED did not have good accuracy at distinguishing binge eating from overeating. This highlights that further consideration of the individual subscales of such tools may be warranted to determine the optimal cut-point to identify individuals with overweight or obesity at risk of an ED.

The results of this review, suggest that further development of screening tools tailored for adolescents and adults with overweight/obesity is needed. The input of individuals with lived experience of obesity and EDs would be extremely valuable to gain insight into the constructs that may be measured as well as the sensitive administration of screening tools (Musić et al., 2022). Four studies included in this review examined such tools and provide insight into the steps required to develop these targeted screening questionnaires (Chamay-Weber et al., 2017; De Man Lapidoth et al., 2007; Franklin et al., 2019; Wever et al., 2018). In the development of questionnaires targeted toward people with overweight or obesity, studies included in this review recommend removing items that may introduce ambiguity, with particular attention required to ensure items examining binge eating may not also capture instances of overeating (De Man Lapidoth et al., 2007), similarly items examining dietary restriction need to have

adequate specificity to distinguish rigid and flexible dietary restraint in tools designed for use throughout weight management therapy (House et al., 2021). In the development of the REO, Wever et al. (2018) identified several clinical characteristics that may be used to differentiate patients with obesity with and without BED, eight such characteristics were incorporated into the development of the REO—weight and shape preoccupation; binge eating and disturbed eating; emotional eating; frequent previous dieting attempts; disturbance in psychosocial functioning; depressive symptoms; self-control; and impulsivity.

Furthermore, additional considerations are required in the development of questionnaires for adolescents with obesity. Foremost, questionnaires must be developmentally appropriate and at a reading level that can be interpreted by the target audience, in the development of the CBBEQ readability was tested to ensure the tool was at a reading level suitable for children from 7 years of age (Franklin et al., 2019). Similarly, Chamay-Weber et al. (2017) involved adolescents in the development of the ADO-BED to ensure that their interpretation of items was consistent with the intended constructs being measured. It has also been suggested that examining loss of control eating among adolescents may be a more useful measure of ED risk than BE, given that adolescents may exhibit less control over the amount of food they consume and thus may experience LOC eating when consuming a small, moderate, or large amount of food dependent on the portions provided to them (Tanofsky-Kraff et al., 2020).

4.4 | Implications for clinical practice

When interpreting the results of this review, selection of screening tools should be informed by the context in which they will be used. In interpreting the main findings of this review, we prioritized sensitivity over specificity as it was deemed important to reduce the probability of FN results in clinical practice. However, when estimating prevalence at a population level, tools require higher specificity as well to reduce the probability of FP results, improving accuracy of these estimates. Furthermore, this review incorporates different types of screening questionnaires, there are pure screening tools, such as the SCOFF and BEDS, as well as longer form questionnaires, such as the EDE-Q and EDO which assess psychopathology and may differentiate ED diagnoses in addition to serving as a screener. If the use of a screener is solely to identify patients who may require clinical follow-up a pure screener may be appropriate. Where clinicians or researchers wish to assess the change in ED symptoms/behaviors over time a long-form questionnaire may be more appropriate.

4.5 | Strengths and Limitations

This review has several strengths. It was conducted in accordance with the Cochrane DTA guidelines and provides a comprehensive review of validation studies for screening questionnaires in

adolescents and adults with overweight and obesity. It addresses an important gap in the literature, given the high prevalence of ED in populations with overweight and obesity, which necessitates well-validated tools to identify at-risk individuals. In conducting this review, all validation studies of ED screening questionnaires that were compared to a diagnostic interview were screened at full text which ensured that relevant studies were not missed. However, there were also some limitations of this review. As this review focussed on the diagnostic accuracy of screening questionnaires, other components of validity—including construct validity of the measured risk factors—were not considered. Included studies were required to report diagnostic accuracy data for individuals with overweight or obesity separately which led to some studies being excluded when data were not reported as required. Several included studies also did not report the demographic characteristics of study populations, particularly ethnicity and socioeconomic status (SES), which makes it difficult to determine the applicability of the results and may prevent future meta-analyses examining the validity of these screening tools for certain sub-groups of the population.

4.6 | Conclusions

Tools designed to measure BED have been extensively validated in adults with overweight and obesity, while existing screening tools designed to identify the broad spectrum of EDs have not been sufficiently validated in this population, though tools designed specifically for adults with obesity show promise. Validated screening questionnaires to identify ED risk in adolescents with obesity are lacking. Thus, questionnaires may have sufficient sensitivity to identify risk of BED in adults with obesity, while additional clinical information (e.g., biochemistry, weight history, and behavioral indicators) should be used in clinical decision-making to identify adult and adolescent patients at risk of other EDs.

AUTHOR CONTRIBUTIONS

Eve T. House: Conceptualization; data curation; formal analysis; methodology; project administration; validation; writing – original draft. **Natalie B. Lister:** Conceptualization; methodology; supervision; writing – reviewing and editing. **Anna L. Seidler:** methodology; writing – review and editing. **Haozhen Li:** Data curation; formal analysis; writing – original draft. **Wee Yee Ong:** Data curation; formal analysis; writing – original draft. **Caitlin M. McMaster:** Writing – review and editing. **Susan J. Paxton:** Conceptualization; methodology; writing – review and editing. **Hiba Jebile:** Conceptualization; data curation; methodology; supervision; writing – original draft.

ACKNOWLEDGMENTS

ETH is supported on the NHMRC funded clinical trial GNT1128317. NBL is supported by an NHMRC Peter Doherty Early Career Fellowship GNT1145748. CMM is supported by a Research Training Stipend (The University of Sydney). HJ is supported by the Sydney Medical

School Foundation (The University of Sydney). The authors acknowledge the technical assistance provided by the Sydney Informatics Hub, a Core Research Facility of the University of Sydney. Open access publishing facilitated by The University of Sydney, as part of the Wiley - The University of Sydney agreement via the Council of Australian University Librarians.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Eve T. House  <https://orcid.org/0000-0002-6953-4786>

Natalie B. Lister  <https://orcid.org/0000-0002-9148-8632>

Anna L. Seidler  <https://orcid.org/0000-0002-0027-1623>

Haozhen Li  <https://orcid.org/0000-0001-9811-7423>

Wee Yee Ong  <https://orcid.org/0000-0002-5110-5115>

Caitlin M. McMaster  <https://orcid.org/0000-0003-2951-0330>

Susan J. Paxton  <https://orcid.org/0000-0002-8486-1996>

Hiba Jebile  <https://orcid.org/0000-0002-8623-6174>

REFERENCES

- Aardoom, J. J., Dingemans, A. E., Slof Op't Landt, M. C., & Van Furth, E. F. (2012). Norms and discriminative validity of the Eating Disorder Examination Questionnaire (EDE-Q). *Eating Behaviors, 13*(4), 305–309. <https://doi.org/10.1016/j.eatbeh.2012.09.002>
- Allison, K. C., Lundgren, J. D., O'Reardon, J. P., Martino, N. S., Sarwer, D. B., Wadden, T. A., Crosby, R. D., Engel, S. G., & Stunkard, A. J. (2008). The Night Eating Questionnaire (NEQ): Psychometric properties of a measure of severity of the night eating syndrome. *Eating Behaviors, 9*(1), 62–72. <https://doi.org/10.1016/j.eatbeh.2007.03.007>
- Anekwe, C. V., Jarrell, A. R., Townsend, M. J., Gaudier, G. I., Hiserodt, J. M., & Stanford, F. C. (2020). Socioeconomics of obesity. *Current Obesity Reports, 9*, 272–279. <https://doi.org/10.1007/s13679-020-00398-7>
- Bohn, K., Doll, H. A., Cooper, Z., O'Connor, M., Palmer, R. L., & Fairburn, C. G. (2008). The measurement of impairment due to eating disorder psychopathology. *Behaviour Research and Therapy, 46*(10), 1105–1110. <https://doi.org/10.1016/j.brat.2008.06.012>
- Borges, M. B. F., Morgan, C. M., Claudino, A. M., & da Silveira, D. X. (2005). Validation of the Portuguese version of the Questionnaire on Eating and Weight Patterns: Revised (QEWPR) for the screening of binge eating disorder. *Brazilian Journal of Psychiatry, 27*(4), 319–322. <https://doi.org/10.1590/s1516-44462005000400012>
- Bossuyt, P., Davenport, C., Deeks, J., Hyde, C., Leeflang, M., & Scholten, R. (2013). Chapter 11: Interpreting results and drawing conclusions. In J. Deeks, P. Bossuyt, & C. Gatsonis (Eds.), *Cochrane handbook for systematic reviews of diagnostic test accuracy version 0.9*. The Cochrane Collaboration.
- Bryant-Waugh, R. J., Cooper, P. J., Taylor, C. L., & Lask, B. D. (1996). The use of the eating disorder examination with children: A pilot study. *International Journal of Eating Disorders, 19*(4), 391–397. [https://doi.org/10.1002/\(SICI\)1098-108X\(199605\)19:4<391::AID-EAT6>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1098-108X(199605)19:4<391::AID-EAT6>3.0.CO;2-G)
- Burton, A. L., Abbott, M. J., Modini, M., & Touyz, S. (2016). Psychometric evaluation of self-report measures of binge-eating symptoms and

- related psychopathology: A systematic review of the literature. *International Journal of Eating Disorders*, 49(2), 123–140. <https://doi.org/10.1002/eat.22453>
- Calugi, S., Pace, C. S., Muzi, S., Fasoli, D., Travagnin, F., & Dalle Grave, R. (2020). Psychometric properties of the Italian version of the questionnaire on eating and weight patterns (QEWP-5) and its accuracy in screening for binge-eating disorder in patients seeking treatment for obesity. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 25(6), 1739–1745. <https://doi.org/10.1007/s40519-019-00818-1>
- Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline. *BMJ*, 368, l6890. <https://doi.org/10.1136/bmj.l6890>
- Chamay-Weber, C., Combesure, C., Lanza, L., Carrard, I., & Haller, D. M. (2017). Screening obese adolescents for binge eating disorder in primary care: The adolescent binge eating scale. *The Journal of Pediatrics*, 185, 68–72.e1. <https://doi.org/10.1016/j.jpeds.2017.02.038>
- Cole, T. J., Bellizzi, M. C., Flegal, K. M., & Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: International survey. *BMJ*, 320(7244), 1240–1243. <https://doi.org/10.1136/bmj.320.7244.1240>
- Colles, S. L., Dixon, J. B., & O'Brien, P. E. (2008). Loss of control is central to psychological disturbance associated with binge eating disorder. *Obesity (Silver Spring)*, 16(3), 608–614. <https://doi.org/10.1038/oby.2007.99>
- Da Luz, F. Q., Hay, P., Touyz, S., & Sainsbury, A. (2018). Obesity with comorbid eating disorders: Associated health risks and treatment approaches. *Nutrients*, 10(7), 829. <https://doi.org/10.3390/nu10070829>
- Da Luz, F. Q., Sainsbury, A., Mannan, H., Touyz, S., Mitchison, D., & Hay, P. (2017). Prevalence of obesity and comorbid eating disorder behaviors in South Australia from 1995 to 2015. *International Journal of Obesity*, 41(7), 1148–1153. <https://doi.org/10.1038/ijo.2017.79>
- De Man Lapidoth, J., Ghaderi, A., Halvarsson-Edlund, K., & Norring, C. (2007). Psychometric properties of the Eating Disorders in Obesity questionnaire: Validating against the Eating Disorder Examination interview. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 12(4), 168–175. <https://doi.org/10.1007/BF03327594>
- De Zwaan, M., Mitchell, J. E., Specker, S. M., Pyle, R. L., Mussell, M. P., & Seim, H. C. (1993). Diagnosing binge eating disorder: Level of agreement between self-report and expert-rating. *International Journal of Eating Disorders*, 14(3), 289–295. [https://doi.org/10.1002/1098-108x\(199311\)14:3<289::aid-eat2260140307>3.0.co;2-0](https://doi.org/10.1002/1098-108x(199311)14:3<289::aid-eat2260140307>3.0.co;2-0)
- de Zwaan, M., Mitchell, J. E., Swan-Kremeier, L., McGregor, T., Howell, M. L., Roerig, J. L., & Crosby, R. D. (2004). A comparison of different methods of assessing the features of eating disorders in post-gastric bypass patients: A pilot study. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association*, 12(6), 380–386. <https://doi.org/10.1002/erv.602>
- Decaluwé, V., & Braet, C. (2004). Assessment of eating disorder psychopathology in obese children and adolescents: Interview versus self-report questionnaire. *Behaviour Research and Therapy*, 42(7), 799–811.
- Deeks, J., Bossuyt, P., & Gatsonis, C. (Eds.). (2013). *Cochrane handbook for systematic reviews of diagnostic test accuracy*. methods.cochrane.org/sdt/handbook-dta-reviews. The Cochrane Collaboration.
- Duncan, A. E., Ziobrowski, H. N., & Nicol, G. (2017). The prevalence of past 12-month and lifetime DSM-IV eating disorders by BMI category in US men and women. *European Eating Disorders Review*, 25(3), 165–171. <https://doi.org/10.1002/erv.2503>
- Dymek-Valentine, M., Rienecke-Hoste, R., & Alverdy, J. (2004). Assessment of binge eating disorder in morbidly obese patients evaluated for gastric bypass: SCID versus QEWP-R. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 9(3), 211–216. <https://doi.org/10.1007/BF03325069>
- Fairburn, C., & Cooper, Z. (1993). The eating disorder examination. In C. Fairburn & G. Wilson (Eds.), *Binge eating: Nature, assessment and treatment* (12th ed., pp. 317–360). Guildford Press.
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16(4), 363–370.
- First, M. B., & Gibbon, M. (2004). The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II). In M. Hilsenroth & D. Segal (Eds.), *Comprehensive handbook of psychological assessment, personality assessment* (Vol. 2). John Wiley & Sons, Inc.
- Franklin, E. V., Simpson, V., Berthet-Miron, M., Gupta, O. T., & Barlow, S. E. (2019). A pilot study evaluating a binge-eating screener in children: Development of the Children's Brief Binge-Eating Questionnaire in a pediatric obesity clinic. *Clinical Pediatrics*, 58(10), 1063–1071. <https://doi.org/10.1177/0009922819863664>
- Freitas, S. R., Lopes, C. S., Appolinario, J. C., & Coutinho, W. (2006). The assessment of binge eating disorder in obese women: A comparison of the binge eating scale with the structured clinical interview for the DSM-IV. *Eating Behaviors*, 7(3), 282–289. <https://doi.org/10.1016/j.eatbeh.2005.09.002>
- Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating disorders over the 2000–2018 period: A systematic literature review. *The American Journal of Clinical Nutrition*, 109(5), 1402–1413. <https://doi.org/10.1093/ajcn/nqy342>
- Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The eating attitudes test: Psychometric features and clinical correlates. *Psychological Medicine*, 12(4), 871–878.
- GBD Obesity Collaborators. (2017). Health effects of overweight and obesity in 195 countries over 25 years. *New England Journal of Medicine*, 377(1), 13–27. <https://doi.org/10.1056/NEJMoa1614362>
- Ghaderi, A., & Scott, B. (2002). The preliminary reliability and validity of the Survey for Eating Disorders (SEDs): A self-report questionnaire for diagnosing eating disorders. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association*, 10(1), 61–76. <https://doi.org/10.1002/erv.425>
- Goldschmidt, A. B., Doyle, A. C., & Wilfley, D. E. (2007). Assessment of binge eating in overweight youth using a questionnaire version of the child eating disorder examination with instructions. *International Journal of Eating Disorders*, 40(5), 460–467.
- Goossens, L., & Braet, C. (2010). Screening for eating pathology in the pediatric field. *International Journal of Pediatric Obesity*, 5(6), 483–490. <https://doi.org/10.3109/17477160903571995>
- Gormally, J., Black, S., Daston, S., & Rardin, D. (1982). The assessment of binge eating severity among obese persons. *Addictive Behaviors*, 7(1), 47–55. [https://doi.org/10.1016/0306-4603\(82\)90024-7](https://doi.org/10.1016/0306-4603(82)90024-7)
- Grange, D. I., & Loeb, K. L. (2007). Early identification and treatment of eating disorders: Prodrome to syndrome. *Early Intervention in Psychiatry*, 1(1), 27–39. <https://doi.org/10.1111/j.1751-7893.2007.00007.x>
- Grupski, A. E., Hood, M. M., Hall, B. J., Azarbad, L., Fitzpatrick, S. L., & Corsica, J. A. (2013). Examining the Binge Eating Scale in screening for binge eating disorder in bariatric surgery candidates. *Obesity Surgery*, 23(1), 1–6. <https://doi.org/10.1007/s11695-011-0537-4>
- Hartmann, A. S., Gorman, M. J., Sogg, S., Lamont, E. M., Eddy, K. T., Becker, A. E., & Thomas, J. J. (2014). Screening for DSM-5 other specified feeding or eating disorder in a weight-loss treatment-seeking obese sample. *The Primary Care Companion for CNS Disorders*, 16(5). <https://doi.org/10.4088/PCC.14m01665>
- Hay, P., Chinn, D., Forbes, D., Madden, S., Newton, R., Sugenor, L., Touyz, S., & Ward, W. (2014). Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the treatment of eating disorders. *Australian and New Zealand Journal of Psychiatry*, 48(11), 977–1008. <https://doi.org/10.1177/0004867414555814>
- Hay, P., Girosi, F., & Mond, J. (2015). Prevalence and sociodemographic correlates of DSM-5 eating disorders in the Australian population.

- Journal of Eating Disorders*, 3(1), 1–7. <https://doi.org/10.1186/s40337-015-0056-0>
- Henderson, M., & Freeman, C. (1987). A self-rating scale for bulimia the 'BITE'. *The British Journal of Psychiatry*, 150(1), 18–24. <https://doi.org/10.1192/bjp.150.1.18>
- Herman, B. K., Deal, L. S., DiBenedetti, D. B., Nelson, L., Fehnel, S. E., & Brown, T. M. (2016). Development of the 7-item binge-eating disorder screener (BEDS-7). *The Primary Care Companion for CNS Disorders*, 18(2). <https://doi.org/10.4088/PCC.15m01896>
- House, E. T., Gow, M. L., Lister, N. B., Baur, L. A., Garnett, S. P., Paxton, S. J., & Jebelle, H. (2021). Pediatric weight management, dietary restraint, dieting, and eating disorder risk: A systematic review. *Nutrition Reviews*, 79(10), 1114–1133. <https://doi.org/10.1093/nutrit/nuaa127>
- Hudson, J. I., Hiripi, E., Pope, H. G., Jr., & Kessler, R. C. (2007). The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biological Psychiatry*, 61(3), 348–358. <https://doi.org/10.1016/j.biopsych.2006.03.040>
- Ivezaj, V., Carr, M. M., Brode, C., Devlin, M., Heinberg, L. J., Kalarchian, M. A., Sysko, R., Williams-Kerver, G., & Mitchell, J. E. (2021). Disordered eating following bariatric surgery: A review of measurement and conceptual considerations. *Surgery for Obesity and Related Diseases*, 17(8), 1510–1520. <https://doi.org/10.1016/j.soard.2021.03.008>
- Kalarchian, M. A., Wilson, G. T., Brolin, R. E., & Bradley, L. (2000). Assessment of eating disorders in bariatric surgery candidates: Self-report questionnaire versus interview. *International Journal of Eating Disorders*, 28(4), 465–469. [https://doi.org/10.1002/1098-108x\(200012\)28:4<465::aid-eat17>3.0.co;2-2](https://doi.org/10.1002/1098-108x(200012)28:4<465::aid-eat17>3.0.co;2-2)
- Kessler, R. C., Berglund, P. A., Chiu, W. T., Deitz, A. C., Hudson, J. I., Shahly, V., Aguilar-Gaxiola, S., Alonso, J., Angermeyer, M. C., Benjet, C., Bruffaerts, R., de Girolamo, G., de Graaf, R., Haro, J. M., Kovess-Masfety, V., O'Neill, S., Posada-Villa, J., Sasu, C., Scott, K., ... Xavier, M. (2013). The prevalence and correlates of binge eating disorder in the World Health Organization World Mental Health Surveys. *Biological Psychiatry*, 73(9), 904–914. <https://doi.org/10.1016/j.biopsych.2012.11.020>
- Kutz, A. M., Marsh, A. G., Gunderson, C. G., Maguen, S., & Masheb, R. M. (2020). Eating disorder screening: A systematic review and meta-analysis of diagnostic test characteristics of the SCOFF. *Journal of General Internal Medicine*, 35, 885–893. <https://doi.org/10.1007/s11606-019-05478-6>
- Liu, C.-Y., Tseng, M.-C. M., Chen, K.-Y., Chang, C.-H., Liao, S.-C., & Chen, H.-C. (2015). Sex difference in using the SCOFF questionnaire to identify eating disorder patients at a psychiatric outpatient clinic. *Comprehensive Psychiatry*, 57, 160–166. <https://doi.org/10.1016/j.comppsy.2014.11.014>
- Mañano, C., Morin, A. J., Lanfranchi, M.-C., & Therme, P. (2013). The Eating Attitudes Test-26 revisited using exploratory structural equation modeling. *Journal of Abnormal Child Psychology*, 41(5), 775–788. <https://doi.org/10.1007/s10802-013-9718-z>
- Manasse, S. M., Michael, M. L., Lin, M., Gillikin, L., Zhang, F., Forman, E. M., & Juarascio, A. (2021). Can a short screening tool discriminate between overeating and binge eating in treatment-seeking individuals with obesity? *Obesity*, 29(4), 706–712. <https://doi.org/10.1002/oby.23128>
- Matthews, A., Peterson, C. M., & Mitan, L. (2019). Adolescent males with atypical anorexia nervosa and premorbid obesity: Three case reports. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 24(5), 963–967. <https://doi.org/10.1007/s40519-019-00702-y>
- Micali, N., Martini, M. G., Thomas, J. J., Eddy, K. T., Kothari, R., Russell, E., Bulik, C. M., & Treasure, J. (2017). Lifetime and 12-month prevalence of eating disorders amongst women in mid-life: A population-based study of diagnoses and risk factors. *BMC Medicine*, 15(1), 12. <https://doi.org/10.1186/s12916-016-0766-4>
- Mond, J. M., Myers, T. C., Crosby, R. D., Hay, P. J., Rodgers, B., Morgan, J. F., Lacey, J. H., & Mitchell, J. E. (2008). Screening for eating disorders in primary care: EDE-Q versus SCOFF. *Behaviour Research and Therapy*, 46(5), 612–622. <https://doi.org/10.1016/j.brat.2008.02.003>
- Moradi, M., Mozaffari, H., Askari, M., & Azadbakht, L. (2022). Association between overweight/obesity with depression, anxiety, low self-esteem, and body dissatisfaction in children and adolescents: A systematic review and meta-analysis of observational studies. *Critical Reviews in Food Science and Nutrition*, 62(2), 555–570. <https://doi.org/10.1080/10408398.2020.1823813>
- Morgan, J. F., Reid, F., & Lacey, J. H. (1999). The SCOFF questionnaire: Assessment of a new screening tool for eating disorders. *BMJ*, 319(7223), 1467–1468. <https://doi.org/10.1136/bmj.319.7223.1467>
- Musić, S., Elwyn, R., Fountas, G., Gnat, I., Jenkins, Z. M., Malcolm, A., Miles, S., Neill, E., Simpson, T., Yolland, C. O., & Phillipou, A. (2022). Valuing the voice of lived experience of eating disorders in the research process: Benefits and considerations. *Australian and New Zealand Journal of Psychiatry*, 56(3), 216–218. <https://doi.org/10.1177/0004867421998794>
- Orbitello, B., Ciano, R., Corsaro, M., Rocco, P., Taboga, C., Tonutti, L., Armellini, M., & Balestrieri, M. (2006). The EAT-26 as screening instrument for clinical nutrition unit attenders. *International Journal of Obesity*, 30(6), 977–981. <https://doi.org/10.1038/sj.ijo.0803238>
- Orlandi, E., Mannucci, E., & Cuzzolaro, M. (2005). Bulimic Investigatory Test, Edinburgh (BITE). A validation study of the Italian version. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 10(1), e14–e20. <https://doi.org/10.1007/BF03354662>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Whiting, P. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Parker, K., Mitchell, S., O'Brien, P., & Brennan, L. (2016). Psychometric evaluation of disordered eating measures in bariatric surgery candidates. *Obesity Surgery*, 26(3), 563–575. <https://doi.org/10.1007/s11695-015-1780-x>
- Pfeiffli, S., Pellegrino, F., Kruseman, M., Pijollet, C., Volery, M., Soguel, L., & Bucher Della Torre, S. (2019). Current recommendations for nutritional management of overweight and obesity in children and adolescents: A structured framework. *Nutrients*, 11(2), 362. <https://doi.org/10.3390/nu11020362>
- Pursey, K. M., Hart, M., Jenkins, L., McEvoy, M., & Smart, C. E. (2020). Screening and identification of disordered eating in people with type 1 diabetes: A systematic review. *Journal of Diabetes and Its Complications*, 34(4), 107522. <https://doi.org/10.1016/j.jdiacomp.2020.107522>
- Quilliot, D., Brunaud, L., Mathieu, J., Quenot, C., Sirveaux, M.-A., Kahn, J.-P., Ziegler, O., & Witkowski, P. (2019). Links between traumatic experiences in childhood or early adulthood and lifetime binge eating disorder. *Psychiatry Research*, 276, 134–141. <https://doi.org/10.1016/j.psychres.2019.05.008>
- Rancourt, D., & McCullough, M. B. (2015). Overlap in eating disorders and obesity in adolescence. *Current Diabetes Reports*, 15(10), 78. <https://doi.org/10.1007/s11892-015-0645-y>
- Reitsma, J., Rutjes, A., Whiting, P., Vlassov, V., & Leeflang, M. (2009). Assessing methodological quality. In J. Deeks, P. Bossuyt, & C. Gatsonis (Eds.), *Cochrane handbook for systematic reviews of diagnostic test accuracy version 1*. The Cochrane Collaboration. <http://srdta.cochrane.org/>
- Ricca, V., Mannucci, E., Moretti, S., Di Bernardo, M., Zucchi, T., Cabras, P., & Rotella, C. (2000). Screening for binge eating disorder in

- obese outpatients. *Comprehensive Psychiatry*, 41(2), 111–115. [https://doi.org/10.1016/s0010-440x\(00\)90143-3](https://doi.org/10.1016/s0010-440x(00)90143-3)
- Rø, Ø., Reas, D. L., & Stedal, K. (2015). Eating disorder examination questionnaire (EDE-Q) in Norwegian adults: Discrimination between female controls and eating disorder patients. *European Eating Disorders Review*, 23(5), 408–412.
- Sawyer, S. M., Whitelaw, M., Le Grange, D., Yeo, M., & Hughes, E. K. (2016). Physical and psychological morbidity in adolescents with atypical anorexia nervosa. *Pediatrics*, 137(4), e20154080. <https://doi.org/10.1542/peds.2015-4080>
- Schaumberg, K., Anderson, D., Anderson, L., Reilly, E., & Gorrell, S. (2016). Dietary restraint: what's the harm? A review of the relationship between dietary restraint, weight trajectory and the development of eating pathology. *Clinical Obesity*, 6(2), 89–100. <https://doi.org/10.1111/cob.12134>
- Shapiro, J. R., Woolson, S. L., Hamer, R. M., Kalarchian, M. A., Marcus, M. D., & Bulik, C. M. (2007). Evaluating binge eating disorder in children: Development of the children's binge eating disorder scale (C-BEDS). *International Journal of Eating Disorders*, 40(1), 82–89. <https://doi.org/10.1002/eat.20318>
- Sherf Dagan, S., Goldenshluger, A., Globus, I., Schweiger, C., Kessler, Y., Kowen Sandbank, G., Ben-Porat, T., & Sinai, T. (2017). Nutritional recommendations for adult bariatric surgery patients: Clinical practice. *Advances in Nutrition*, 8(2), 382–394. <https://doi.org/10.3945/an.116.014258>
- Shomaker, L. B., Tanofsky-Kraff, M., Elliott, C., Wolkoff, L. E., Columbo, K. M., Ranzhofer, L. M., Roza, C. A., Yanovski, S. Z., & Yanovski, J. A. (2010). Salience of loss of control for pediatric binge episodes: Does size really matter? *The International Journal of Eating Disorders*, 43(8), 707–716. <https://doi.org/10.1002/eat.20767>
- Siervo, M., Boschi, V., Papa, A., Bellini, O., & Falconi, C. (2005). Application of the SCOFF, Eating Attitude Test 26 (EAT 26) and Eating Inventory (TFEQ) Questionnaires in young women seeking diet-therapy. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 10(2), 76–82. <https://doi.org/10.1007/BF03327528>
- Šimundić, A.-M. (2009). Measures of diagnostic accuracy: Basic definitions. *The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine*, 19(4), 203–211.
- Solmi, F., Hatch, S. L., Hotopf, M., Treasure, J., & Micali, N. (2015). Validation of the SCOFF questionnaire for eating disorders in a multiethnic general population sample. *International Journal of Eating Disorders*, 48(3), 312–316. <https://doi.org/10.1002/eat.22373>
- Spitzer, R. L., Devlin, M., Walsh, B. T., Hasin, D., Wing, R., Marcus, M., Stunkard, A., Wadden, T., Yanovski, S., Agras, S., & Mitchell, J. (1992). Binge eating disorder: A multisite field trial of the diagnostic criteria. *International Journal of Eating Disorders*, 11(3), 191–203. [https://doi.org/10.1002/1098-108X\(199204\)11:3<191::AID-EAT2260110302>3.0.CO;2-S](https://doi.org/10.1002/1098-108X(199204)11:3<191::AID-EAT2260110302>3.0.CO;2-S)
- Tanofsky-Kraff, M., Schvey, N. A., & Grilo, C. M. (2020). A developmental framework of binge-eating disorder based on pediatric loss of control eating. *American Psychologist*, 75(2), 189–203. <https://doi.org/10.1037/amp0000592>
- Thelen, M. H., Farmer, J., Wonderlich, S., & Smith, M. (1991). A revision of the bulimia test: The BULIT—R. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3(1), 119–124. <https://doi.org/10.1037/1040-3590.3.1.119>
- Udo, T., & Grilo, C. M. (2018). Prevalence and correlates of DSM-5–defined eating disorders in a nationally representative sample of US adults. *Biological Psychiatry*, 84(5), 345–354. <https://doi.org/10.1016/j.biopsych.2018.03.014>
- van Hoeken, D., & Hoek, H. W. (2020). Review of the burden of eating disorders: Mortality, disability, costs, quality of life, and family burden. *Current Opinion in Psychiatry*, 33(6), 521–527. <https://doi.org/10.1097/YCO.0000000000000641>
- Vander Wal, J. S., Stein, R. I., & Blashill, A. J. (2011). The EDE-Q, BULIT-R, and BEDT as self-report measures of binge eating disorder. *Eating Behaviors*, 12(4), 267–271. <https://doi.org/10.1016/j.eatbeh.2011.07.006>
- Vander Wal, J. S., Waller, S. M., Klurfeld, D. M., McBurney, M. I., & Dhurandhar, N. V. (2005). Night eating syndrome: Evaluation of two screening instruments. *Eating Behaviors*, 6(1), 63–73. <https://doi.org/10.1016/j.eatbeh.2004.04.004>
- Wadden, T. A., & Foster, G. D. (2006). Weight and lifestyle inventory (WALI). *Obesity*, 14(S3), 99S–118S. <https://doi.org/10.1038/oby.2006.289>
- Warner, J. (2004). Clinicians' guide to evaluating diagnostic and screening tests in psychiatry. *Advances in Psychiatric Treatment*, 10(6), 446–454. <https://doi.org/10.1192/apt.10.6.446>
- Weinberger, N. A., Kersting, A., Riedel-Heller, S. G., & Luck-Sikorski, C. (2016). Body dissatisfaction in individuals with obesity compared to normal-weight individuals: A systematic review and meta-analysis. *Obesity Facts*, 9(6), 424–441. <https://doi.org/10.1159/000454837>
- Wever, M. C., Dingemans, A. E., Geerets, T., & Danner, U. N. (2018). Screening for binge eating disorder in people with obesity. *Obesity Research & Clinical Practice*, 12(3), 299–306. <https://doi.org/10.1016/j.orcp.2018.02.002>
- Wolter, H., Schneider, N., Pfeiffer, E., & Lehmkuhl, U. (2009). Diagnostic crossover from obesity to atypical anorexia nervosa – A case report. *Obesity Facts*, 2(1), 52–53. <https://doi.org/10.1159/000194641>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: House, E. T., Lister, N. B., Seidler, A. L., Li, H., Ong, W. Y., McMaster, C. M., Paxton, S. J., & Jebeile, H. (2022). Identifying eating disorders in adolescents and adults with overweight or obesity: A systematic review of screening questionnaires. *International Journal of Eating Disorders*, 55(9), 1171–1193. <https://doi.org/10.1002/eat.23769>