Published in final edited form as:

Int J Eat Disord. 2018 December ; 51(12): 1357–1360. doi:10.1002/eat.22972.

Identifying a male clinical cutoff on the Eating Disorder Examination-Questionnaire (EDE-Q)

Lauren M. Schaefer, Ph.D.^a, Kathryn E. Smith, Ph.D.^{a,b}, Rachel Leonard, Ph.D.^c, Chad Wetterneck, Ph.D.^c, Brad Smith, M.D.^c, Nicholas Farrell, Ph.D.^c, Bradley C. Riemann, Ph.D.^c, David A. Frederick, Ph.D.^d, Katherine Schaumberg, Ph.D.^e, Kelly L. Klump, Ph.D.^f, Drew A. Anderson, Ph.D.^g, and J. Kevin Thompson, Ph.D.^h

^aSanford Research, Fargo, North Dakota, USA

^bUniversity of North Dakota School of Medicine and Health Sciences, Department of Psychiatry and Behavioral Science, Fargo, North Dakota, USA

^cRogers Memorial Hospital, Oconomowoc, Wisconsin, USA

^dChapman University, Department of Psychology, Orange, California, USA

^eUniversity of Wisconsin-Madison, Department of Psychiatry, Madison, Wisconsin, USA

^fMichigan State University, Department of Psychology, East Lansing, Michigan, USA

^gUniversity at Albany-State University of New York, Department of Psychology, Albany, New York, USA

^hUniversity of South Florida, Department of Psychology, Tampa, Florida, USA

Abstract

Objective: Evidence suggests that eating disorders (EDs) may be under-detected in males. Commonly used measures of EDs such as the Eating Disorder Examination-Questionnaire (EDE-Q) were initially developed within female samples, raising concern regarding the extent to which these instruments may be appropriate for detecting EDs in males. The current study used receiver operating characteristic (ROC) curve analysis to (a) examine the accuracy of the EDE-Q global score in correctly classifying males with and without clinically significant ED pathology, and (b) establish the optimal EDE-Q global clinical cutoff for males.

Method: Participants were a clinical sample of 245 male ED patients and a control sample of 205 male undergraduates.

Results: EDE-Q global scores demonstrated moderate-high accuracy in predicting ED status (area under the curve = 0.85, 95% CI: 0.82–0.89). The optimal cutoff of 1.68 yielded a sensitivity of 0.77 and specificity of 0.77.

Discussion: Overall, results provide preliminary support for the discriminant validity of EDE-Q scores among males. However, concerns remain regarding the measure's ability to

comprehensively assess domains of disordered eating most relevant to males. Therefore, careful attention to the possibility for measurement bias and continued evaluation of the scale in males is encouraged.

Keywords

assessment; measurement; ROC curve analysis; discriminant validity; eating pathology; men

While males have often been regarded as a minority among individuals with eating disorders (EDs), evidence suggests that EDs are not uncommon in males, with some data indicating that ED behaviors are increasing more rapidly in males than in females (Murray et al., 2017). However, historically ED research has relied on primarily female samples, which questions the extent to which assessment instruments developed and validated with female samples are appropriate for use with males. The Eating Disorder Examination-Questionnaire (EDE-Q; Fairburn & Beglin, 1994) is one such measure that assesses a range of cognitive and behavioral ED symptoms, and is commonly used in clinical and research settings. Despite the widespread use of this measure and its practical advantages as a brief self-report assessment, relatively little is known regarding the validity of the EDE-Q in detecting clinically significant ED pathology among males.

The validity of an assessment instrument is in part determined by its ability to accurately classify cases (i.e., sensitivity) and non-cases (i.e., specificity). Several studies have empirically examined the EDE-Q global cutoff score that yields an optimal trade-off between sensitivity and specificity using receiver operating characteristic (ROC) curve analysis in female samples. Mond et al. (2004) found that an EDE-Q global score of 2.30 was associated with ideal sensitivity (.92) and specificity (.86) to detect EDs among a community sample of women, though this cutoff was found to be somewhat higher (2.80) in a study of women in primary care settings (Se = .80, Sp = .80; Mond et al., 2008). Similarly, Rø, Reas, and Stedal (2015) found that an EDE-Q global cutoff of 2.50 accurately discriminated between women with and without EDs (Se = .86, Sp = .86).

However, recent evidence has raised concern regarding whether these cutoffs would yield appropriate sensitivity and specificity for males. For example, several studies have indicated that males with EDs evidence lower EDE-Q scores compared to female clinical samples (Dahlgren, Stedal, & Rø, 2017; Jennings & Phillips, 2017; Mantilla & Birgegård, 2016; Smith et al., 2017). At present, the reasons for these findings are not entirely clear. While females with EDs may exhibit more severe ED pathology, the observed differences could also reflect measurement biases, as the EDE-Q does not fully assess all relevant domains for males with EDs (e.g., muscularity concerns; Murray et al., 2017). Nevertheless, the finding that males with EDs consistently evidence significantly lower EDE-Q scores than their female counterparts questions the application of previously established EDE-Q cutoffs with males. That is, it may be that a lower cutoff provides a better balance between sensitivity and specificity among males.

Therefore, the aims of the present study were to apply ROC curve analysis to (a) examine the accuracy of the EDE-Q global score in classifying males with and without clinically significant ED pathology, and (b) establish the optimal EDE-Q global clinical cutoff for

males. Doing so may address previously identified problems related to the under-detection of EDs in males and ultimately improve their access to specialized ED treatment (Murray et al., 2017).

Method

Participants

The clinical sample was comprised of 245 male patients receiving treatment for an ED within Rogers Memorial Hospital's residential (79%) and partial hospital (21%) programs between 2002 and 2015. Anorexia nervosa restricting subtype was the most common primary diagnosis (30.6%), followed by bulimia nervosa (30.2%), eating disorder not otherwise specified (22.0%), anorexia nervosa binge-purge subtype (16.3%) and binge-eating disorder (0.8%). Twenty-nine percent of the sample was diagnosed with a comorbid psychiatric disorder. The most common secondary diagnoses were major depressive disorder (9.2%) and obsessive-compulsive disorder (3.9%). The mean EDE-Q global score in the clinical sample was 3.06 ($SD = 1.60$), which is similar to scores reported in other male clinical samples (Jennings & Phillips, 2017), but lower than those reported in female clinical samples (Welch, Birgegård, Parling, & Ghaderi, 2011).

Participants for the control sample were drawn from a larger dataset of 256 undergraduate males attending one of four geographically diverse universities in the United States. The mean EDE-Q global score in the full college sample was 1.15 ($SD = 1.06$), which is comparable to scores reported in similar college male samples (Darcy, Hardy, Crosby, Lock, & Peebles, 2013; Lavender, De Young, & Anderson, 2010). Behavioral frequency items from the EDE-Q were used to remove individuals with symptoms of an eating disorder based on DSM-5 criteria (i.e., body mass index $< 18.5 \text{ kg/m}^2$, binge eating or purging 4 times in the past month) from the college database. The final control sample was composed of 205 males with a mean EDE-Q global score of 1.04 ($SD = 0.94$). Table 1 contains additional demographic information for each of the samples.

Measures

Demographic Information.—Demographic information for the clinical sample was obtained from patients' medical records. Control participants self-reported their age, race/ethnicity, height, and weight. Self-reported height and weight were used to calculate body mass index (BMI).

Eating Disorder Examination – Questionnaire (EDE-Q).—The EDE-Q assesses disordered eating attitudes and behaviors experienced in the last 28 days. Twenty-two items are rated on a 7-point scale ranging from 0 (*no days/not at all*) to 6 (*everyday/markedly*), yielding four subscale scores assessing Restraint, Shape Concern, Weight Concern, and Eating Concern. The global score represents an average of the four subscale scores, with higher scores indicating greater eating pathology. The measure also includes six free-response items assessing the frequency of eating disorder behaviors (e.g., bingeing, purging). In the current study, participants in the clinical sample were administered the 36-item EDE-Q 4 (Fairburn & Beglin, 1994), while participants in the control sample were administered

the 28-item EDE-Q 6 (Fairburn & Beglin, 2008). Despite differences in the overall length of the scales, the 22 items contributing to the calculation of the global score on each version are highly comparable across measures. Internal consistency for the global score was .96 in the clinical sample and .91 in the control sample.

Procedure

Upon admission, participants for the clinical sample completed the EDE-Q. ED diagnoses were made by board-certified psychiatrists with extensive experience in the assessment and treatment of eating disorders, and were based on clinical interview and DSM-IV-TR (American Psychiatric Association, 2000). The Human Subjects Committee at Rogers Memorial Hospital reviewed and approved the procedure for recruitment and data collection. Participants for the control sample were recruited through each university's undergraduate research pool. Informed consent was provided electronically and measures were completed online via secure survey software. Participants received extra course credit upon completion. All study procedures were approved by the recruiting university's Institutional Review Board.

Data Analyses

ROC curve analysis was used to test the performance of the EDE-Q global in predicting disordered eating status and identify a cutoff score. The ROC curve is a graphical representation of a measure's sensitivity plotted against its false positive rate (i.e., 1-specificity). The area under the curve (AUC) summarizes a test's overall accuracy, or ability to distinguish cases from non-cases, based on the average value of sensitivity for all possible values of specificity. AUC values are defined as non-informative (< 0.50), less accurate (0.51 to 0.70), moderately accurate (0.71 to 0.90), highly accurate (0.91 to 0.99), and perfect (1.0; Swets, 1988). Consistent with previous work identifying EDE-Q cutoffs in females (e.g., Mond et al., 2008), the optimal clinical cutoff was defined as the EDE-Q global score that maximized both sensitivity (i.e., ability to detect individuals with significant eating pathology) and specificity (i.e., ability to screen out individuals with nonclinical eating pathology). All analyses were conducted using SPSS Statistics version 25.0.

Results

As shown in Figure 1, EDE-Q global scores demonstrated moderate-high accuracy in predicting eating disorder status, $AUC = 0.85$, 95% CI [0.82–0.89]. In other words, there is a 85% probability that a randomly selected individual from the clinical sample would obtain a higher EDE-Q global score than a randomly selected individual from the control sample. A cutoff score of 0.02 provided the optimal sensitivity (0.99), however, this score was associated with poor specificity (0.09). A cutoff score of 3.95 provided the optimal specificity (1.00), however, this score was associated with poor sensitivity (0.36). The cutoff that provided the optimal balance of sensitivity and specificity was a mean EDE-Q global score of 1.68, which had a sensitivity of 0.77 and a specificity of 0.77¹.

¹Notably, the clinical and control samples differed on important demographic characteristics (See Table 1). As a measure's accuracy and optimal cutoff may be impacted by sample demographics (Rø et al., 2015), the discriminant validity and optimal EDE-Q global cutoff was also examined by age (i.e., 18–20, 21–30) and race (i.e., white, non-white). BMI was not examined as the dataset did not

Discussion

The current study fills an important gap in the literature by (a) examining the performance of a widely-used measure of eating pathology, the EDE-Q, in discriminating between males with and without significant eating pathology, and (b) identifying an optimal clinical cutoff score on the EDE-Q global among males. Results indicate that the scale is able to distinguish between cases and controls with moderate-high accuracy. Examination of sensitivity and specificity values suggest that an EDE-Q global cutoff score of 1.68 best differentiates individuals with and without clinically significant eating pathology, which is considerably lower than the cutoffs identified in female samples (e.g., Mond et al., 2004).

Although these results suggest that the EDE-Q may be useful for identifying males with elevated ED symptoms, important caveats must be noted. It is possible that the lower EDE-Q cutoff observed among males may reflect actual gender differences in levels of eating pathology that signal clinical concern among males and females. However, it has also been suggested that the EDE-Q (and many similar measures) may better assess ED symptomatology common among female presentations of the disorder (i.e., thinness-oriented attitudes and behaviors) than symptomatology more common among male presentations (i.e., muscularity-oriented attitudes and behaviors; Murray et al., 2017). For example, although calculation of the EDE-Q global score weighs subscales assessing weight concern and shape concern equally, evidence indicates that compared with females, males with EDs are less concerned with weight but equally concerned with shape (Strober et al., 2006). Thus, males with EDs may not demonstrate high scores on weight-related items despite engagement in clear disordered eating patterns. In contrast, behaviors that are more common among males with EDs such as a pathological pursuit of muscularity are not indexed within the EDE-Q, which may result in the under-detection of these behaviors and a concomitant deflation of global scores. Therefore, apparent sex differences in EDE-Q scores should be interpreted with caution.

Limitations of the current study include the use a self-report measure to identify the control sample and non-structured clinical interviews to identify the clinical sample. Therefore, the use of a structured clinical interview to identify cases and non-cases is an important avenue for future work in this area. Further, although items contributing to the EDE-Q 4 and EDE-Q 6 global score are highly similar, it is possible that slight differences in item wording may contribute to differences in scores observed between the clinical and control groups. Finally, although subgroup analyses indicated a limited impact of age and race/ethnicity on measure accuracy and clinical cutoffs, additional work is needed to examine the generalizability of results to diverse populations.

In sum, the current study suggests that the EDE-Q global score may be useful in identifying EDs in men, though results support this use of a lower cutoff than those reported for female samples. Application of this cutoff may serve to improve ED screening procedures in males,

contain an adequate number of non-cases for some weight categories (e.g., underweight). Overall, the EDE-Q global demonstrated moderate to high accuracy in discriminating between cases and non-cases. Although slightly higher cutoffs were observed for non-white respondents (1.81) and those aged 18–20 (1.71), compared to white respondents (1.50) and those aged 21–30 (1.62), the impact of age and race appeared to be relatively small. Results from these subgroup analyses are available as online Supplemental Materials.

and in turn, identify those in need of specialized ED treatment. However, concerns remain regarding the measure's ability to comprehensively assess all domains of disordered eating relevant to males, particularly muscularity-oriented concerns. Therefore, careful attention to the possibility for measurement bias and continued evaluation of the scale in males is encouraged.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institute of Mental Health (grant number T32 MH082761).

References

- American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington, DC: Author.
- Dahlgren CL, Stedal K, & Rø Ø (2017). Eating Disorder Examination Questionnaire (EDE-Q) and Clinical Impairment Assessment (CIA): Clinical norms and functional impairment in male and female adults with eating disorders. *Nordic Journal of Psychiatry*, 71(4), 1–9. [PubMed: 27598719]
- Darcy AM, Hardy KK, Crosby RD, Lock J & Peebles R (2013). Factor structure of the Eating Disorder Examination Questionnaire (EDE-Q) in male and female college athletes. *Body Image*, 10(3), 399–405. doi: 10.1016/j.bodyim.2013.01.008 [PubMed: 23453695]
- Fairburn CG, & Beglin SJ (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16(4), 363–370. [PubMed: 7866415]
- Fairburn CG, & Beglin SJ (2008). Eating Disorder Examination Questionnaire (EDE-Q 6.0) In Fairburn CG (Ed.), *Cognitive behavior therapy and eating disorders* (pp. 309–314). New York, NY: Guilford Press.
- Jennings KM, & Phillips KE, (2017). Eating Disorder Examination Questionnaire (EDE-Q): Norms for a clinical sample of males. *Archives of Psychiatric Nursing*, 31(1), 73–76. doi: 10.1016/j.apnu.2016.08.004 [PubMed: 28104062]
- Lavender JM, De Young KP, & Anderson DA (2010). Eating Disorder Examination Questionnaire (EDE-Q): Norms for undergraduate men. *Eating Behaviors*, 11(2), 119–121. doi:10.1016/j.eatbeh.2009.09.005 [PubMed: 20188296]
- Mantilla EF, & Birgegård A (2016). Eating Disorder Examination Questionnaire: Norms and clinical reference data from adolescent boys and girls in Sweden. *Psychiatry Research*, 239, 164–162. doi: 10.1016/j.psychres.2016.03.022
- Mond JM, Hay PJ, Rodgers B, Owen C, & Beumont PV (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, 42(5), 551–567. doi:10.1016/S0005-7967(03)00161-X [PubMed: 15033501]
- Mond JM, Myers TC, Crosby RD, Hay PJ, Rodgers B, Morgan JF, & ... Mitchell JE (2008). Screening for eating disorders in primary care: EDE-Q versus SCOFF. *Behaviour Research and Therapy*, 46(5), 612–622. doi:10.1016/j.brat.2008.02.003 [PubMed: 18359005]
- Murray SB, Nagata JM, Griffiths S, Calzo JP, Brown TA, Mitchison D, & ... Mond JM (2017). The enigma of male eating disorders: A critical review and synthesis. *Clinical Psychology Review*, 57, 1–11. doi:10.1016/j.cpr.2017.08.001 [PubMed: 28800416]
- Rø Ø, Reas DL, & 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, 408–412. doi: 10.1002/erv.2372 [PubMed: 26094887]
- Smith KE, Mason TB, Murray SB, Griffiths S, Leonard RC, Wetterneck CT, & ... Lavender JM (2017). Male clinical norms and sex differences on the Eating Disorder Inventory (EDI) and Eating

- Disorder Examination Questionnaire (EDE-Q). *International Journal of Eating Disorders*, 50(7), 769–775. doi:10.1002/eat.22716 [PubMed: 28436086]
- Strober M, Freeman R, Lampert C, Diamond J, Teplinsky C, & DeAntonio M (2006). Are there gender differences in core symptoms, temperament, and short-term prospective outcome in anorexia nervosa? *International Journal of Eating Disorders*, 39, 570–575. doi: 10.1002/eat.20293 [PubMed: 16937384]
- Strother E, Lemberg R, Stanford SC, & Turberville D (2012). Eating disorders in men: Underdiagnosed, undertreated, and misunderstood. *Eating Disorders: The Journal of Treatment & Prevention*, 20(5), 346–355. doi:10.1080/10640266.2012.715512
- Swets JA (1988). Measuring the accuracy of diagnostic systems. *Science*, 240(4857), 1285–1293. doi: 10.1126/science.3287615 [PubMed: 3287615]
- Welch E, Birgegård A, Parling T, & Ghaderi A (2011). Eating Disorder Examination Questionnaire and clinical impairment assessment questionnaire: General population and clinical norms for young adult women in Sweden. *Behaviour Research and Therapy*, 49(2), 85–91. doi:10.1016/j.brat.2010.10.010 [PubMed: 21185552]

Figure 1.
ROC Curve Analysis Showing the Area Under the Curve for the Prediction of Disordered Eating Status Using the EDE-Q Global Score.
Note. Area under the ROC curve = 0.85; 95% Confidence Interval: 0.82–0.89.

Table 1

Demographic Information for Clinical and Control Samples

	Clinical Sample (N = 245)	Control Sample (N = 205)	Significance
Age <i>M</i> (<i>SD</i>)	21.75 (3.19)	20.34 (1.79)	$t(511) = 6.29^*$
BMI <i>M</i> (<i>SD</i>)	21.92 (6.92)	24.38 (3.98)	$t(503) = 4.99^*$
Race/Ethnicity %			$\chi^2 (4, N = 519) = 166.71^*$
White or Caucasian	81.3%	39.7%	
Black or African American	0.8%	10.3%	
Hispanic or Latina	1.6%	12.7%	
Asian	0.8%	27.9%	
Multiracial or other	15.5%	9.3%	

*
 $p < .001$