Screen for Disordered Eating: Improving the accuracy of eating disorder screening in primary care

Shira Maguen, Claire Hebenstreit, Yongmei Li, Julie V. Dinh, Rosemary Donalson, Sarah Dalton, Emma Rubin, Robin Masheb

San Francisco VA Medical Center, 4150 Clement Street, San Francisco, CA 94121, United States
University of California, San Francisco, 505 Parnassus Ave, San Francisco, CA 94143, United States
Department of Psychiatry, Yale School of Medicine, United States
VA Connecticut Healthcare System, United States

ARTICLE INFO

Keywords:
Eating disorders
Screening
Primary care
Binge eating

ABSTRACT

Objective: To develop a primary care eating disorder screen with greater accuracy and greater potential for generalizability, compared to existing screens.

Design: Cross-sectional survey to assess discriminative accuracy of a new screen, Screen for Disordered Eating (SDE), compared to Eating Disorders Screen for Primary Care (EDS-PC) and SCOFF screener, using prevalence rates of Binge Eating Disorder (BED), Bulimia Nervosa (BN), Anorexia Nervosa (AN), and Any Eating Disorder (AED), as measured by the Eating Disorder Examination Questionnaire (EDE-Q).

Results: The SDE correctly classified 87.2% (CI: 74.3%–95.2%) of BED cases, all cases of BN and AN, and 90.5% (CI: 80.4%–96.4%) of AED cases. Sensitivity estimates were higher than the SCOFF, which correctly identified 69.6% (CI: 54.2%–82.3%) of BED, 77.8% (CI: 40.0%–97.2%) of BN, 37.5% (CI: 8.52%–75.5%) of AN, and 66.1% (CI: 53%–77.7%) of AED. While the EDS-PC had slightly higher sensitivity than the SDE, the SDE had better specificity. The SDE outperformed the SCOFF in classifying true cases, the EDS-PC in classifying true non-cases, and the EDS-PC in distinguishing cases from non-cases.

Conclusions: The SDE is the first screen, inclusive of BED, valid for detecting eating disorders in primary care. Findings have broad implications to address eating disorder screening in primary care settings.

1. Introduction

Eating disorders are prevalent psychiatric disorders and a major public health burden in the United States (US). At least 30 million people of all ages and genders have an eating disorder in the US [1]. The 2013 publication of the Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (DSM-5) included revisions for Anorexia Nervosa (AN) and Bulimia Nervosa (BN), and added a new diagnostic category, Binge Eating Disorder (BED), that is more prevalent than both AN and BN combined [2]. To our knowledge existing eating disorder screens were validated prior to the publication of DSM-5. The generalizability of these screens for the full range of DSM-5 revised and new eating disorders is unknown.

Two standalone, brief measures that can be used in primary care include the Eating Disorder Screen for Primary Care (EDS-PC) [3] and the SCOFF (not an acronym) [4]. Both are 4–5 item measures that were validated in low to normal weight English medical clinics patients and/or university students [3–5]. While both screens can be helpful for screening in the populations for which they were developed, applicability for more diverse populations is unknown.

Detection of eating disorders is critical since it can prevent downstream medical and mental health complications associated with these diagnoses. Furthermore, timely detection can assist with early intervention efforts and better prognosis. Finally, detection can assist with better triage and appropriately honed treatment. Detection of eating disorders is especially important among veterans served by the Veterans Health Administration (VHA), given that a recent systematic review found prevalence estimates among veteran and service members to be comparable to or higher than prevalence estimates for the general population [6–7]. In fact, one study found that almost three-quarters of participants in VA’s weight loss program reported disordered eating [8]. Although screening for PTSD, depression, and alcohol use is conducted in primary care, the VA does not Screen for Disordered Eating, which is also true of most other healthcare systems.

While detection of eating disorders is key, having a streamlined screen is also critical for busy, fast-paced primary care and integrated...
care settings. Rather than rely on multiple screens that capture different eating disorders, reducing patient and provider burden through a unified screen can facilitate detection and triage to care.

Consequently, the goal of this study was to develop a primary care eating disorder screen with greater accuracy, compared to existing screens, and greater potential for generalizability by using a veteran sample. More specifically, we examined the discriminative accuracy of a brief primary care screen, as compared to two widely used eating disorder screens, which draws from several existing measures and includes BED.

2. Methods

Participants were 407 female veterans who were recruited from one VHA medical center and its community outpatient clinics. Participants completed mailed surveys that included measures of disordered eating. Participants met the following inclusion/exclusion criteria: 1) were between the ages of 18 and 70; 2) had no lifetime history of any psychiatric disorder; and 3) had not attempted suicide in the past five years. Recruitment letters were sent to female veterans meeting these criteria, along with a pre-stamped postcard with the option to indicate whether the veteran would like to be contacted about a study of the relationship between stressors experienced in the military and eating behaviors. Questionnaires were mailed to interested participants to complete at home and return to the San Francisco VA Medical Center (SFVAMC) by mail. Participants received $30 for their participation. All procedures were approved by the University of California, San Francisco Institutional Review Board and the Human Research Protection Program at the SFVAMC.

2.1. Measures

Participants completed a questionnaire indicating their age, race/ethnicity, education, household income, marital status, and military history. Clinical variables were measured using written self-report measures designed for use in mental health and/or primary care settings, including the EDE-Q, EDS-PC and SCOFF, as well as a new five-item screen called Screen for Disordered Eating (SDE). The SDE is a self-report version of the Eating Disorders Examination (EDE) [9–10], an interviewer-administered evaluation considered the gold standard when assessing eating disorders. Two existing screening instruments were also used: the EDS-PC and SCOFF. The EDS-PC is a four-item measure (scored 0–4), and the SCOFF is a five-item measure (scored 0–5), both of which have cut scores of two or greater to indicate probable Anorexia Nervosa (AN) or Bulimia Nervosa (BN). Finally, we used the SDE, a five-item screening tool that was aimed at capturing traditional eating disorders as well as the addition of BED. Each of the five items of the SDE were derived based on face validity with eating disorder diagnoses (see Appendix A) and drawn from existing eating disorder measures, including one item each from the Dutch Eating Behaviors Questionnaire (DEBQ) [11], Patient Health Questionnaire (PHQ) [12], Minnesota Eating Behaviors Survey (MEBS) [13], Eating Attitudes Test (EAT-26) [14], and SCOFF [4]. The items from the DEBQ and the EAT-26 were dichotomized for consistency with the other items, with total scores ranging from 0 to 5.

The EDE-Q [10] was used as the benchmark against which we tested the other measures. BED, defined as an average of one or more objective binge episodes (OBE) per week without compensatory or purging behaviors [15] as well as BN and AN, were assessed using the EDE-Q.

2.2. Statistical analysis

In order to determine the cut score of the SDE, we compared the receiver operating characteristic (ROC) curve, sensitivity, and specificity yielded by several cutoffs and chose the optimal threshold that could maximize discriminative accuracy and obtain clinically meaningful classification of any eating disorder (AED).

In order to evaluate the discriminative accuracy of the SDE in the detection of BED, BN, AN, and AED, we used two screening instruments, the EDS-PC and SCOFF as comparative references.

To quantify the discriminative (or diagnostic) ability of the screening instruments, we assessed sensitivity, specificity, ROC curve, and positive and negative predictive values. Sensitivity refers to the proportion of those with eating disorders that are correctly classified by the screening instrument. Specificity refers to the proportion of those without eating disorders that are correctly classified by the screening instrument [16]. These two measures relate to the ability of the scale to identify those with and without a particular disorder and are critical for assessing the cutoff points. To test the effectiveness of the cutoff points, we supplemented our analysis with ROC curve. The area under the curve (AUC) is equal to the probability of correctly distinguishing cases from the non-cases. This global diagnostic accuracy measure has values ranging from 0.5 and 1, where 0.5 suggests non-informative discrimination and 1 indicates perfect discrimination [17]. In addition to the above measures, we assessed positive and negative predictive values (PPV and NPV). These measures depend on disorder prevalence and are useful clinical tools. They respectively represent the probability of having a disorder for someone with a positive result and the probability of being healthy for someone with a negative result.

3. Results

3.1. Characteristics of study sample

Sixty-three participants were diagnosed with any eating disorder (AED) based on EDE-Q assessment criteria. Specifically, 47 participants had BED, 9 had BN and 8 had AN. On average, participants with AED were 51.7 (± 12.5) years old, and participants with BN were younger (48.1 years) than those with BED (52.7 years) and AN (51.4 years). Those with eating disorders were predominantly Caucasian (50.8%), and those with BN and AN were composed of Caucasian, Asian and Other race, while the BED group additionally included African Americans and Hispanics. Other demographic and military variables by particular eating disorder can be found in Table 1.

3.2. Threshold of the screen for disordered eating (SDE)

Total sum scores for the five SDE items ranged from 0 to 5. We examined using 1, 2, 3, 4, and 5 as the cut scores and consequently obtained five sets of ROC curve, sensitivity, and specificity measures (see Table 2). With a cutoff of 2, the SDE was most accurate in discriminating AED cases from non-cases (AUC = 0.74), higher than ROC curves associated with other cutoffs (AUC = 0.52–0.73). In addition, sensitivity and specificity with a cutoff of 2 had the best tradeoff; the SDE correctly classified 90.5% AED cases as positive and correctly classified 57.5% of non-cases of AED as negative. A cutoff of 2 had much higher sensitivity than 3–5 (4.8%–68.3%) and much higher specificity than 1 (26.0%) that picks up 74% of false positives (1–26%). Based on these comparisons, we chose 2 as the cutoff for SDE.

3.3. Classification by screeners

The classification performance of the SDE was determined by the proportion of participants endorsing each of the five items of the scale. A participant screened positive if she endorsed two or more items. As shown in Table 3, 205 participants screened positive on the SDE. Among participants with AED, 87.1% endorsed the item based on the DEBQ (“Do you often feel the desire to eat when you are emotionally upset or stressed?”), while only 56.2% of healthy participants endorsed this item. Of the 205 participants that screened positive on the SDE, 86.2% endorsed this item. The next most frequently endorsed items were endorsed by 75.8% (“Are you often preoccupied with a desire to...
be thinner?)" and 74.2% of participants with AED ("Do you often feel that you can't control what or how much you eat?). Half of those with any eating disorder endorsed the item from the SCOFF ("Do you believe yourself to be fat when others say you are thin?"). The item based on the MEBS was endorsed by only 12.9% of participants with eating disorders (Do you sometimes make yourself throw up to control your weight?), but by 66.7% of participants with BN.

The SDE captured 41 of 47 participants with BED (87.2%) and classified them as positive, and classified 195 of 355 participants without BED (54.9%) as negative. For BED, the SCOFF classified more healthy participants (78.2%) than those with eating disorders (69.6%); and the EDS-PC captured all those with eating disorders, but also categorized 60.8% of healthy individuals as having eating disorders.

The prevalence rates of BN and AN were low. Among the nine participants diagnosed with BN, all were classified as such by the EDS-PC and seven were classified as such by the SCOFF. Among the majority without BN, 290 (85.6%) were classified as negative by the SCOFF, 202 (51%) by the SDE and 135 (35.2%) by the EDS-PC. For AN, the SDE is better at capturing those with eating disorders (100%) than the SCOFF (51%) by the SDE and 135 (35.2%) by the EDS-PC. For BN, the SDE is better at capturing those with eating disorders (100%) than the SCOFF (92%) and the EDS-PC (91%).

The SDE captured 41 of 47 participants with BED (87.2%) and classified them as positive, and classified 195 of 355 participants without BED (54.9%) as negative. For BED, the SCOFF classified more healthy participants (78.2%) than those with eating disorders (69.6%); and the EDS-PC captured all those with eating disorders, but also categorized 60.8% of healthy individuals as having eating disorders.

The prevalence rates of BN and AN were low. Among the nine participants diagnosed with BN, all were classified as positive by the SDE, eight were classified as such by the EDS-PC and seven were classified as such by the SCOFF. Among the majority without BN, 290 participants (73.6%) were classified as negative by the SCOFF, 202 (51%) by the SDE and 135 (35.2%) by the EDS-PC. For AN, the SDE is better at capturing those with eating disorders (100%) than the SCOFF (92%) and the EDS-PC (91%).
Table 4
Percentage of participants who met criteria for each eating disorder by screening item.

<table>
<thead>
<tr>
<th>New eating screen items</th>
<th>BED</th>
<th>BN</th>
<th>AN</th>
<th>AED</th>
<th>NED</th>
<th>SDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Do you often feel the desire to eat when you are emotionally upset or stressed? (DEBQ)</td>
<td>87%</td>
<td>88.9%</td>
<td>75.0%</td>
<td>87.1%</td>
<td>56.2%</td>
<td>86.2%</td>
</tr>
<tr>
<td>2. Do you often feel that you can’t control what or how much you eat? (PHQ #6)</td>
<td>76.1%</td>
<td>88.9%</td>
<td>50.0%</td>
<td>74.2%</td>
<td>28.3%</td>
<td>66.5%</td>
</tr>
<tr>
<td>3. Do you sometimes make yourself throw up (vomit) to control your weight? (MEBS)</td>
<td>4.3%</td>
<td>66.7%</td>
<td>12.5%</td>
<td>12.9%</td>
<td>4.4%</td>
<td>11.3%</td>
</tr>
<tr>
<td>4. Are you often preoccupied with a desire to be thinner? (EAT-26 #11)</td>
<td>71.7%</td>
<td>88.9%</td>
<td>87.5%</td>
<td>75.8%</td>
<td>28.4%</td>
<td>67.0%</td>
</tr>
<tr>
<td>5. Do you believe yourself to be fat when others say you are thin? (SCOFF #4)</td>
<td>40.9%</td>
<td>77.8%</td>
<td>75.0%</td>
<td>50%</td>
<td>31.3%</td>
<td>56.7%</td>
</tr>
</tbody>
</table>

Note: BED = Binge Eating Disorder; BN = Bulimia Nervosa; AN = Anorexia Nervosa; AED = Any Eating Disorder; NED = No Eating Disorder; SDE = Screen for Disordered Eating.

Of the 63 participants with any eating disorder, 57 (90.5%) were classified as positive by the SDE, 41 (65.1%) by the SCOFF and 56 (88.9%) by the EDS-PC. Of the 339 patients without an eating disorder, 270 (79.6%) were classified as negative by the SCOFF, 195 (57.5%) by the SDE and 133 (39.2%) by the EDS-PC.

3.4. Sensitivity and specificity

Table 4 summarizes the classification information into discriminative accuracy measures. The SDE correctly classified 87.2% of true BED cases, all cases of BN and AN, and 90.5% of cases with AED. These sensitivity estimates are higher than those of the SCOFF, which correctly identified less than 80% of true cases for the four disorders (respectively 69.6% for BED, 77.8% for BN, 37.5% for AN and 66.1% for Any ED). Overall, the EDS-PC screening instrument performed better than the SCOFF and the new screening instrument, with sensitivity being 100% for BED and BN, 96.6% for AED, and 66.7% for AN.

The SDE was better than the EDS-PC in classifying non-cases correctly. More specifically, the SDE correctly identified 54.9% of non-cases of BED, while the EDS-PC correctly identified 39.2% of non-cases of BED. In other words, the SDE suggested 45.1% false positives (1–54.9%) of BED while the EDS-PC suggested 60.8% false positives of BED. The same pattern applied to BN, AN, and AED. On the other hand, the SCOFF was better than both other instruments in distinguishing non-cases correctly for all four disorders, namely, 78.2% of non-cases of BED, 73.6% of non-cases of BN, 72.5% of non-cases of AN and 79.9% of non-cases of AED.

3.5. ROC

When comparing ROC curve results across screening instruments, it is important to account for the results of sensitivity and specificity, because the ROC curve reflects the tradeoff between sensitivity and specificity. Overall, the SDE was better at distinguishing cases and non-cases than the EDS-PC. For instance, the probability of correctly classifying cases from non-cases of BN is 75.5% (AUC = 0.755) using the SDE, while the corresponding probability using the EDS-PC is 67.6%. This pattern applied to all other disorders. This comparison of ROC curve results between the two instruments is meaningful, because for both instruments, sensitivity is higher than specificity for all four disorders. Interpretation of the ROC curve results between the SDE and the SCOFF is not straightforward, because the SDE tends to have higher sensitivity than specificity and the SCOFF tends to have lower sensitivity than specificity.

3.6. Positive and negative predictive values

The SDE and the EDS-PC had comparable high negative predictive values. For the SDE, 100% of those screened negative for BN and AN did not have these disorders, and 97% of those screened negative for BED and AED did not have these disorders. For the EDS-PC, 100% of those screened negative for BED and BN did not have these disorders, and...
98.5% of those screened negative for AN and AED did not have these disorders. The SCOFF has a slightly lower negative predictive value compared with the other two scales, with the NPVs being 95.2% for BED, 99.3% for BN, 98.3% for AN and 92.8% for AED.

All positive predictive values for the three screening instruments were low (the highest being 37.6%), indicating relatively small proportions of those classified as positive truly did have the disorder. Overall, the SCOFF had a higher PPV compared with the other two instruments. Using the SCOFF, 37.6% of those screened as positive for AED had the disorder; using the SDE, 28.4% of those screened as positive for AED had the disorder; and using the EDS-PC, 22.1% of those screened positive for AED had the disorder. This pattern may be related to low prevalence of these disorders in our study population. The prevalence rates were 12.0% for BED, 2.2% for BN, 2.0% for AN, and 15.0% for AED. With higher prevalence, PPV tends to increase while NPV tends to decrease.

4. Discussion

We developed the first primary care eating disorder screener for the full range of DSM-5 eating disorders in order to capture patients who may require further intervention. We found that while the other standalone screening measures were either over- or under-capturing patients with eating disorders in our large sample, the SDE was able to strike a balance. This was the first known study to compare three known eating disorder screeners with a number of discriminative accuracy statistics, including ROC curve, and the first study to investigate eating disorder primary care screening for BED.

We found that the SDE outperformed the SCOFF in classifying true cases, outperformed the EDS-PC in classifying true non-cases, and outperformed the EDS-PC in distinguishing cases from non-cases of BED, BN, AN and AED. The SDE was developed for use in primary care in order to capture those with traditional eating disorders, such as AN or BN, as well as those with more newly conceptualized cases of eating disorders, such as BED. An important goal was to ensure that those with BED were captured by the screen, given that it was included in the DSM-5 after both the EDS-PC and SCOFF were developed. To address this gap, a new screening measure for binge eating in primary care has been validated in a veteran sample, however, it does not screen for other eating disorders [18]. Furthermore, because veterans may evidence disordered eating at the same or greater levels than civilians, we wanted to ensure a tool that could be used with veterans. For example, we know that about three-quarters of veterans within a weight loss program evidenced disordered eating, so it is critical to include BED as well as ensuring a tool that can capture the highest-risk veterans.

Being able to identify individuals with disordered eating through screening is important not only because eating disorders can cause death and disability, but also because providers may be missing opportunities to provide optimal care for overweight or obese patients. For example, eating disorder symptoms and unhealthy eating habits may be overlooked, referrals for eating disorder treatment not made, and weight loss efforts may be hindered by behaviors such as binge eating and purging. This could lead to frustration for patients and providers, particularly for overweight patients who try to lose weight without success. There are also lost opportunities for patients to disclose their eating symptoms for fear of embarrassment. Screening can help open up and normalize the conversation about this topic. Screening can not only help identify those who need more comprehensive evaluation and tailored care, but also serve a preventative measure. More specifically, we found that 75% of veterans that are returning from deployments to the Middle East and seeking VA care are overweight or obese [19]. If a certain proportion of those veterans have subthreshold or undiagnosed eating disorders, screening can help identify them early before symptoms become calcified and are more difficult to treat. Veterans may have unique struggles with weight and eating due to military requirements such as “making weight” and being monitored regularly for eating and weight [7]. Consequently, when separating from the military, unique factors may place veterans at higher risk for disordered eating, such as poor eating habits during service (e.g., eating quickly due to demands of war and limited access to healthy food while deployed) and military-related stress and trauma. Furthermore, many veterans may be injured during their service and can no longer exercise to ward off the impact of poor eating or high caloric diets. For all of these reasons, as well as studies demonstrating accelerated weight gain after separation [19–20], screening for disordered eating is especially important in this population.

There are several limitations to this study. First, our study sample only included women veterans. While prior studies have also used predominantly or entirely females when determining accuracy of screening instruments, it is important to include men, especially given the higher rates of males with BED, compared to the rates of males with AN and BN, and recent studies suggesting high rates of BED among male veterans [8,21]. Second, while the EDE-Q is an excellent self-report instrument to help identify probable eating disorders, we were not able to employ clinician-administered diagnostic instruments due to our data collection methods (i.e., mailed surveys). The EDE-Q is a self-report measure that is based on the gold-standard EDE clinical interview, but is not itself the gold standard. In the future, we hope to validate the SDE with a clinician-administered interview to verify our findings. Finally, given that the number of individuals with some eating disorders were small, it is difficult to know how well these findings generalize to other individuals with these eating disorders. Although this is generally a significant challenge with eating disorder studies due to the prevalence of these disorders, we hope to address this issue in future studies by having even larger sample sizes.

A screening tool that can be used to capture the full range of eating disorders can facilitate early identification and treatment. Given that veterans evidence similar or higher rates of eating disorders than the general population, an instrument that can be used in primary care settings to identify veterans for further evaluation is important. We found that the SDE, a five-item tool, had good discriminative accuracy and could easily be integrated into existing systems of care, particularly with integrated care providers embedded in primary care, to facilitate the transition to mental health. Furthermore, the VA is rolling out nationwide provider training for evidence-based therapy for eating disorders, so use of a tailored veteran screen is optimal. Future research should continue to examine accuracy and validity of this new tool, as well as how to best facilitate its implementation within VA. Finally, validation of this screen in non-VA primary care settings has broad implications given that existing research has not addressed screening for eating disorder patients with BED in primary care.

Conflicts of interest

None of the authors have any conflicts of interest to report.

Acknowledgements

This research was supported by the Veterans Health and Integration Program (VHIP; PI: Maguen). Writing of this manuscript was supported by the Department of Veterans Affairs Office of Academic Affiliations, Advanced Fellowship Program in Women's Health (Hebenstreit). We would also like to thank Rebecca Gloria for her thoughtful review of our manuscript. These data were presented at the Society for Behavioral Medicine Conference on March 30, 2017.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.genhosppsych.2017.09.004.
References


