Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Validation of the SCOFF Questionnaire for Eating Disorders in a Multiethnic General Population Sample

Francesca Solmi, PhD1*
Stephani L. Hatch, PhD2
Matthew Hotopf, PhD2
Janet Treasure, PhD3
Nadia Micali, PhD1

ABSTRACT

Introduction: This study aimed to validate the SCOFF, an eating disorders (ED) screening questionnaire, in a multiethnic general population sample of adults.

Method: A two-stage design was employed using the South East London Community Health Study phases I and II data. A total of 1,669 participants were screened using the SCOFF in SELCoH-I, and 145 were administrated an ED clinical interview in SELCoH-II. We explored the diagnostic validity of the questionnaire restricting to the 145 individuals with the clinical questionnaire.

Results: Sensitivity and specificity of the SCOFF were 53.7 and 93.5%, respectively.

Conclusion: The SCOFF showed good levels of specificity but low sensitivity, resulting in a high percentage of false negatives. Given the low sensitivity found in our sample the SCOFF is likely to be a suboptimal measure for the identification of ED in the community.

Keywords: SCOFF; eating disorders; diagnostic validity

(Int J Eat Disord 2015; 48:312-316)

Introduction

Eating disorders (ED) and other specified feeding or eating disorders (OSFED) [the umbrella definition for subthreshold ED in the Diagnostic and Statistical Manual of mental disorders 5th ed (DSM-5)] have a lifetime prevalence of up to 9%1 and are associated with considerable physical and psychiatric comorbidity.2 However, ED are often undetected in the general population resulting in small proportions of individuals receiving treatment.3

The introduction of the SCOFF (an acronym describing five key screening questions for ED, which can be recalled through the mnemonic ‘Sick, Control, One stone, Fat, Food’)4 as a screening tool for ED in clinical settings has opened a window of opportunity for extending its use to routine screenings in the general population. The SCOFF has been validated in a number of primary-care based studies; in the UK4–7 and internationally;8–13 both in written and oral delivery;14 and compared to other instruments.15 Early studies have employed clinical ED cases and controls,4,5,7 used mainly female populations5,7,15–17 and few have included individuals older than 40 years of age.6,16,18 Overall, studies using clinical populations have yielded higher values of sensitivity (Se) and specificity (Sp)5,6,18 than community studies, the latter usually finding higher Sp than Se.11–13

However, most validation studies in the community have relied on young8,11 and homogeneous populations6,12 (i.e., females, limited ethnic representation), limiting the scope for the generalizability of these results. Therefore, in order to investigate the suitability of the SCOFF as both a screening tool at the community level and in general population surveys, we aimed to pilot a validation of the instrument in a multiethnic population-based sample of adults aged 16–90.

Method

Sample, Measures

This study employed data from the South East London Community Health Study (SELCoH) I and II, a two-phase...
general population survey of 1,669 individuals aged 16+ living in the London (UK) boroughs of Lambeth and Southwark. More details on the rationale, sampling, representativeness, and assessment of participants in SELCoHI is provided elsewhere. In SELCoHI, 1,669 participants completed the SCOFF questionnaire and underwent objective anthropometric measurements to calculate body mass index (BMI). In SELCoHI, participants who had given their consent to be re-contacted in SELCoHI and did not need an interpreter (N = 1,560) were eligible for inclusion in a clinical assessment of ED using the ED section of the Structured Clinical Interview for DSM-IV Axis I Disorders nonpatient edition (SCID-I-NP), if they: (i) had screened positive at the SCOFF in SELCoHI (N = 158); or (ii) had screened negative and had not screened positive for other mental health conditions (N = 599). SCOFF positive participants were gender matched with a randomly selected sample of eligible screen negatives.

Participants were asked for ED symptoms occurred at the time of and since the SELCoHI assessment when answering SCID-I questions. Despite referring to DSM-IV diagnosis, the SCID-I also contains a section on binge eating disorder (BED), which means that all ED diagnoses were explored. Moreover, interviewers did not apply the ‘skip-rules’ of the SCID-I in order to avoid underestimating the prevalence of diagnoses and gathered information on type, frequency, and duration of ED behaviors in order to be able to subsequently derive DSM5 diagnoses. SELCoHI data was collected between June 2008 and December 2010 and SELCoHII data between August 2011 and March 2013. Consent was collected prior to participation to the study.

Analyses

Sample characteristics were described using cross tabulations and chi-square tests. Sensitivity, specificity, and positive and negative predictive values (PPV, NPV) were calculated on the sample interviewed in SELCoHII (N = 145) using sampling weights accounting for: (i) proportion of SCOFF positive and negative participants interviewed with the SCID over the whole sample who completed the SCOFF who agreed to take part to SELCoHII (N total: 1,560; screen positive: 158; screen negative: 1402); and (ii) proportion of ED diagnoses amongst screen positives and negatives, as previously recommended in two-phase epidemiological studies. These weights account for the real prevalence of the condition in the population when circumstances do not allow maintaining adequate sampling ratios, but do not account for characteristics associated with participation at follow-up (e.g., age, gender, and ethnicity). We weighted our data based on the sample of the 1,560 participants who agreed to be followed up, on the a priori knowledge that we could have not assessed the real ED status of the remaining 109 participants with the SCID-I and that limiting analyses to participants eligible for inclusion (i.e., without mental health comorbidities, common in individuals with ED) could have underestimated the number of false negatives and, thus overestimated the sensitivity. As a sensitivity analysis, we calculated sensitivity and specificity using the whole initial sample (N = 1,669) and the results did not change qualitatively (Se: 52.8; Sp: 93.7). Analyses were run in Stata12.

Results

Sample

Of the 322 participants who were eligible to take part to the ED module (158 screen positive and 164 screen negatives), 89 (56.3%) participants who had screened positive on the SCOFF and 88 (53.6%) of those who screened negative were lost to follow up. As seen in Supporting Information Table S1, no systematic differences existed between participants who took part in the study and those who were lost to follow-up with respect to age, marital status, ethnicity, education, and age. Among screen positive participants, however, more obese (67%) and underweight (100.0%) participants were lost to follow-up.

A total of 145 participants (76 (46.3%) SCOFF negative, 69 (43.7%) SCOFF positive) were assessed using the SCID-I interview. Of these, 31 (21.4%) received a threshold or subthreshold ED diagnosis. The majority of participants were female (75%), of White ethnicity (43%), between the ages of 25 and 34 years (28%), with a normal BMI (54%) and had at least a General Certificate of Secondary Education (GCSE) qualification (90%). No participant was underweight (Table 1).

Diagnostic Validity

DSM5 diagnosis was correctly predicted by the SCOFF for a total of 101 (69.7%; N = 73: no ED; N = 28: ED) participants; 3 (2.0%) participants were misclassified by the SCOFF as not having an ED and 41 (28.3%) as having an ED. Two of the three false negative participants had a diagnosis of binge eating disorder (BED) and one of OSFED presenting with excessive exercise (not in table). All false negatives were women, of White ethnicity, between the ages of 35 and 54 and ~60% were obese (Supporting Information Table S2).

Based on the established cutoff of ≥2 positive answers, the weighted sensitivity, specificity, PPV and NPV of the SCOFF were 53.7 (95% CI: 36.2–71.2), 93.5 (95% CI: 88.9–98.0), 40.6 (95% CI: 28.9–53.1), and 96.1 (95% CI: 88.9–99.2), respectively.
Discussion

This study aimed to assess the validity of the SCOFF as a screening tool for ED in a multiethnic general population sample of adults in London (UK).

In line with previous general population studies, we found that specificity of the instrument was higher than its sensitivity, and that the latter was lower than what previously found in some studies,¹¹,¹² but not others.¹³ All of these studies employed a younger population than ours. PPV was low, which is common for low prevalence conditions, but NPV was high. Low sensitivity suggests that high proportions of individuals with an ED are not identified by the SCOFF.

Several factors could account for this finding. It is possible that, in the absence of follow-up questions (i.e., such as probing questions contained in the SCID) the ego-syntonicity of ED could lead to negative answers. Moreover, the focus on the fat/thin dichotomy in assessing body dissatisfaction could introduce gender biases not accounting for different cognitions in men²³ or in individuals who are overweight or obese. Although the former were not represented amongst the participants who were false negatives in our sample, two of the three of false negatives were obese. More research using larger mixed gender sample is warranted to test the validity of the SCOFF in men. The limited numbers of questions contained in the SCOFF could also mean that important behaviors (e.g., laxative use, excessive exercise) central to the diagnosis of bulimia nervosa or OSFED such as purging disorder, are missed, as in the case of the participant presenting with excessive exercise. Similarly, although we could not provide evidence of this from our sample, it can be speculated that without a measure of BMI, cases of AN where the individual is stable on a low, unhealthy BMI, could be missed as question 3 of the SCOFF only enquires about recent weight loss. Finally, the SCOFF was developed when BED was not yet a recognized diagnosis and it is possible that its questions are limited in identifying the condition, which might be more prevalent in older adults, as was the case amongst false negatives in our sample. This could explain the lower sensitivity we found compared to other community studies employing younger participants.

This study has several strengths. It employed a large representative and ethnically diverse general population sample, suggesting that findings are generalizable to similarly diverse populations. Although the validation was conducted on a subsample of individuals (less than 10% of the full study population), sampling weights were employed to account for differential sampling across screen positives and negatives to ensure that estimates of sensitivity and specificity reflected the prevalence of ED in the whole sample. However, some limitations should also be accounted for. The SCID interview was conducted 2–3 years following the administration of the SCOFF; recall bias and regression to the mean could thus have occurred to some extent. Since ED are chronic conditions and no incident cases (i.e., new ED onsets between the administration of the SCOFF and the SCID) were found in interviews, we suggest that the diagnoses were not greatly under- or over-estimated and that false negatives could not be attributed to ED cases with onset occurring after the administration of the SCOFF. It is possible, however, that different recall patterns could have occurred. We tried to minimize this by asking about present and past ED behaviors in general and then for their duration in order to identify overlaps with the time of the SCOFF interview. Interviewers were also blinded as to the screening status of participants in order to avoid observer bias. Substantial losses to follow-up occurred and whilst the sample of screen

---

### TABLE 1. Socio-demographic characteristics of the interviewed sample in SELCoHII

<table>
<thead>
<tr>
<th>Socio-Demographic Characteristics</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>145 (100)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>36 (24.8)</td>
</tr>
<tr>
<td>Female</td>
<td>109 (75.2)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>83 (57.2)</td>
</tr>
<tr>
<td>Black</td>
<td>42 (29)</td>
</tr>
<tr>
<td>Asian</td>
<td>4 (2.8)</td>
</tr>
<tr>
<td>Other</td>
<td>16 (11)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>No qualification</td>
<td>14 (9.7)</td>
</tr>
<tr>
<td>GCSE/A-level</td>
<td>70 (48.3)</td>
</tr>
<tr>
<td>Degree level or above</td>
<td>61 (42)</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>75 (54.3)</td>
</tr>
<tr>
<td>Overweight</td>
<td>31 (22.5)</td>
</tr>
<tr>
<td>Obese</td>
<td>32 (23.2)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>16–24</td>
<td>38 (26.2)</td>
</tr>
<tr>
<td>25–34</td>
<td>41 (28.3)</td>
</tr>
<tr>
<td>35–44</td>
<td>23 (15.9)</td>
</tr>
<tr>
<td>45–54</td>
<td>26 (17.9)</td>
</tr>
<tr>
<td>55–64</td>
<td>9 (6.2)</td>
</tr>
<tr>
<td>65+</td>
<td>8 (5.5)</td>
</tr>
</tbody>
</table>

### TABLE 2. Summary of weighted diagnostic validity measures for the SCOFF questionnaire

<table>
<thead>
<tr>
<th>Diagnostic Measures</th>
<th>Values (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>53.7 (36.2–71.2)</td>
</tr>
<tr>
<td>Specificity</td>
<td>93.5 (88.9–98.0)</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>40.6 (28.9–53.1)</td>
</tr>
<tr>
<td>Negative predicted value</td>
<td>96.1 (88.9–99.2)</td>
</tr>
</tbody>
</table>

SOLMI ET AL.
negative seems to be representative of the overall sample (Supporting Information Table S1). 4 (100%) underweight participants amongst the screen positives lost to follow-up could index missed ED cases (especially AN) (Supporting Information Table S1). Moreover, it appears that losses to follow-up in the screen positive group could have occurred with respect to participants of Asian and other ethnic backgrounds (Supporting Information Table S1). This could introduce some degree of selection bias, which our sampling weights could not account for. However, given the small proportion and the lower weight assigned to screen positive participants interviewed compared to screen negatives, this is unlikely to bias the overall sensitivity and specificity estimates.11,12 Only participants with no mental health comorbidities were eligible to be assessed in SELCoHII and were therefore interviewed. Although we weighted Se and Sp to represent the whole sample that agreed to be followed up, we could be overestimating or underestimating Se. On the one hand, as ED are comorbid with a number of psychiatric conditions24 and more false negatives could have occurred and been missed in the noneligible sample. On the other hand, underestimation could also have occurred, as the false negatives ratio was based on three individuals only and, although weights were applied, uncertainty around the estimate (reflected in the wide 95% CI) exists. Future studies should aim at conducting both interviews simultaneously and on the whole sample to improve accuracy of findings.

Recent studies found high levels of psychiatric comorbidity in individuals who screened positive on the SCOFF.25–27 Given its high specificity, this is likely to reflect the high levels of psychopathology in individuals with ED. However, although good at ruling out an ED (i.e., high specificity) the SCOFF should be used with caution as an ED screening tool, as the low sensitivity in our sample indicates that a substantial number of individuals with ED might be missed. This is particularly relevant to population-based samples and suggestions of using the SCOFF as a screening tool in the community.13 Based on our findings, more research is needed to assess whether rephrasing some of its questions could improve the diagnostic validity of the SCOFF without compromising its brevity. Moreover, research exploring the diagnostic validity of the SCOFF across different population subgroups (e.g., overweight and obese, males, ethnic minorities) is warranted to improve the measure.

SLH and MH receive salary support from the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, King’s College London. This research was supported by the Bio-medical Research Nucleus data management and informatics facility at South London and Maudsley NHS Foundation Trust, which is funded by the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London and a joint infrastructure grant from Guy’s and St Thomas’ Charity and the Maudsley Charity. This research was also funded by a National Institute of Health Research (NIHR) clinician scientist award to Dr N Micali and by a grant received by the British Academy. These funders had no involvement in study design, data collection, analysis or the decision to submit for publication. The authors have no financial involvement (including employment, fees, share ownership) or affiliation with any organisation whose financial interests may be affected by material in the manuscript, or which might potentially bias it. This publication is the work of the authors and Nadia Micali will serve as guarantors for the contents of this paper. The authors are grateful for the support received the Open Access Funding Team at the University College London Library Services towards publishing this manuscript as an open access publication.

References


17. Siervo M, Boschi V, Papa A, Bellini O, Falconi C. Application of the SCOFF, Eating Attitude Test 26 (EAT 26) and Eating Inventory (TFEQ) Questionnaires in young women seeking diet-therapy. Eat Weight Disord 2005;10:76–82.


