Clinical versus patient-reported measures of depression in bariatric surgery

Sudarshan Srivatsan1 · Vinay Gudugunta1 · Kelly Z. Young1 · Aliasghar Arastu1 · Cameron R. Strong1 · Ruth Cassidy2 · Amir A. Ghaferi1,2,3,4

Received: 26 September 2017 / Accepted: 7 February 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract
Background Bariatric surgery patients with mental illness may experience worse surgical outcomes compared to those without. Depression is the most prevalent mental health diagnosis amongst Americans with obesity. Accurate diagnosis and treatment is of paramount importance to mitigate perioperative risk. Unfortunately, there is no standard method to screen patients for depression prior to surgery. Our goal was to understand the relationship between traditional clinical screening tools and a novel patient-reported depression screening survey, Patient Health Questionnaire 8 (PHQ-8), in the setting of the bariatric surgery preoperative assessment.

Methods The study included all adult bariatric surgery patients from January 2014 through June 2016. Patients who were not assessed using both the PHQ-8 and a traditional clinical depression screening were excluded from the study. There were a total of 4486 patients who met the eligibility criteria and were included in analysis. We used comparative statistics to examine the association between these screening tools and to test for contributing demographic, surgical, and socioeconomic factors.

Results The overall rate of clinically diagnosed depression in the study cohort was 45.6%. In comparison, 14.8% of all patients screened positive for depression using the PHQ-8. Of the patients without a traditional clinical diagnosis of depression, 10.2% screened positive for depression using the PHQ-8. This subset of undiagnosed patients was more likely to be non-white, employed, and had a higher BMI than their clinically diagnosed counterparts.

Conclusions and Relevance We found a higher rate of clinically diagnosed depression in our cohort compared to the general population. However, when using the validated PHQ-8 survey, the rate of depression more closely approximated the national incidence. Further, a significant proportion of patients were undiagnosed and/or misdiagnosed by current clinical assessments. Standardizing preoperative depression screening using validated patient-centered tools may prevent the consequences of untreated depression.

Keywords Bariatric surgery · Depression · Mental health

Depression is the most prevalent mental health diagnosis amongst the 15 million individuals with morbid obesity living in America [1–3]. Accurate diagnosis and treatment is of paramount importance to mitigate its potential sequelae. Depression is also associated with increased perioperative risk in surgical patients. With bariatric surgery as the most effective treatment for morbid obesity [4], depression diagnosis and treatment is an important component of the preoperative evaluation. About one-third of patients undergoing bariatric surgery have been diagnosed with depression [3]. While advances in surgical technique, perioperative care guidelines, and rigorous preoperative assessments have established bariatric surgery as one of the safest abdominal
procedures, there is no standardized method for preoperative mental health screening [5–7].

There is a wide range of depression screening practices and no consensus on best-practices or an optimal assessment tool [8]. In the context of bariatric surgery, traditional assessment methods include some combination of the following: clinical interviews, symptom inventories, personality/psychopathology test, tests of cognitive function, and projective personality tests [8]. Previous work raises the concern that current methods of screening do not fully capture all patients with depression, especially in certain demographics [8, 9]. For example, several studies have shown that the type of provider and method of evaluation can result in a wide range of diagnoses, especially amongst racial and ethnic minorities [9, 10]. This raises the problem of patients with depression who did not screen positive undergoing surgery without the appropriate treatment. Research across several other surgical subspecialties has also shown that undiagnosed and untreated depression results in increased morbidity and mortality following surgery [11]. Despite these compelling data, there is no standardized depression screening for bariatric surgery in the preoperative period, even with insurance and society mandates to perform such screening.

Therefore, a better understanding of how to accurately diagnose patients in the preoperative setting is paramount to improving bariatric surgery outcomes. In this study, we sought to understand the relationship between two diagnostic tools: traditional clinical diagnosis and a patient-reported depression screening tool, Personal Health Questionnaire-8 (PHQ-8) [12]. Specifically, we evaluated the demographics of patients who had a clinical diagnosis in conflict with their self-reported PHQ-8 result. The findings of this study will better inform efforts to standardize preoperative mental health screenings in bariatric surgery.

**Methods**

**Study population and data sources**

We utilized the Michigan Bariatric Surgery Collaborative (MBSC) clinical registry as our primary source of data. The MBSC aims to improve bariatric surgery outcomes and reduce variation in standard of care in the state of Michigan through rapid identification and implementation of best practices. Though the collaborative is voluntary, a hospital must perform a minimum of 25 bariatric procedures annually to be eligible. These procedures can include gastric bypass, laparoscopic adjustable gastric banding, biliopancreatic diversion (BPD) with or without duodenal switch (DS), and sleeve gastrectomy. Data from these procedures are collected pre-, peri-, and post-operatively and are clinically audited.

The study population included all adult bariatric surgery patients from January 2014 through June 2016 who completed a preoperative Patient Health Questionnaire Depression Scale (PHQ-8) and underwent traditional clinical depression screening per the discretion of their bariatric surgery program. Patients considered for this study included those who underwent the following operations: Roux-en-Y gastric bypass (RYGB), laparoscopic adjusted gastric band (LAGB), laparoscopic sleeve gastrectomy (LSG), and BPD/DS. Patients who were not assessed using both tools were excluded from the study.

This study was deemed exempt from review by the University of Michigan Institutional Research Board.

**Outcomes**

Data were collected on two measures of depression: traditional clinical depression diagnosis and the PHQ-8. The diagnosis of depression is made by a mental health provider and evaluated for study purposes by a trained data abstractor. Traditional depression screening methods used in the pre-op evaluation setting include clinical interview, symptom inventory, objective personality/psychopathology tests, tests of cognitive function, and projective personality tests [8]. The evaluation determines the adequacy and stability of treatment for existing mental health diagnoses, and is also utilized to diagnose and treat new mental health problems. The standardized MBSC definition of depression includes “clinical depression and depressive disorder, treated with medication, electroconvulsive therapy, and/or psychotherapy.”

The other tool, PHQ-8 (Fig. 1), is an established, validated, patient-reported depression specific questionnaire that has shown diagnostic and severity measure validity in large-scale clinical and epidemiological studies [12]. More recently, Kroenke et al. demonstrated that the PHQ-8 is a strong measure of depression in large population-based surveys [12]. The PHQ-8 is an eight-question survey, with a maximum score of 24 points. Previous studies have established that a PHQ-8 score > 10 is considered positive and diagnostic of clinical depression with a sensitivity of 88% and specificity of 88% [12]. Due to its relative brevity, the PHQ-8 is more convenient to use than most other traditional standard diagnostic depression algorithms and clinical screenings [12].

At the MBSC, the PHQ-8 questionnaire is built into the immediate preoperative patient intake, known as the “Baseline Questionnaire.” Patients receive either an automated email link containing the survey or a paper copy via mail if an email address was not provided. Both email and paper surveys are distributed by the MBSC Coordinating Center at the University of Michigan. 100% of patients in the study cohort completed the PHQ-8 survey.
Statistical analysis

The outcome of interest was the level of discordance between traditional clinical diagnosis of depression and the PHQ-8. In our first analytic step, we examined the association between patient-reported and clinically diagnosed depression, determining the proportion of patients with “undiagnosed” depression. This involved creating four distinct groups of patients (Fig. 2).

Next, we conducted a Chi-squared test using the variables age, race, gender, income level, type of insurance, and employment status to compare “undiagnosed” patients and all clinically diagnosed patients. All analyses were performed using statistical software (SAS) at the 5% significance level.

Results

The study cohort included 4486 eligible patients from the MBSC (Table 1). Patients underwent the following operations: LSG—82.7%, RYGB—16%, LAGB—0.8%, and

---

Over the **last 2 weeks**, how often have you been bothered by any of the following problems? (*circle one number on each line*)

<table>
<thead>
<tr>
<th>How often during the past 2 weeks were you bothered by...</th>
<th>Not at all</th>
<th>Several days</th>
<th>More than half the days</th>
<th>Nearly every day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Little interest or pleasure in doing things...</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2. Feeling down, depressed, or hopeless...</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3. Trouble falling or staying asleep, or sleeping too much...</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4. Feeling tired or having little energy...</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5. Poor appetite or overeating...</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6. Feeling bad about yourself, or that you are a failure, or have let yourself or your family down...</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7. Trouble concentrating on things, such as reading the newspaper or watching television...</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8. Moving or speaking so slowly that other people could have noticed. Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual...</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

---

**Fig. 1** PHQ-8 Questionnaire [12]

**Fig. 2** The “undiagnosed” patient population included patients who were not previously clinically diagnosed with depression but tested positive for depression for PHQ8
BPS/DS—0.4%. Of the eligible patients, 21.5% were male, the mean age was 45.8 years (SD: 11.8), the average BMI was 47.8 (SD: 8.3), 83.6% had 3 or more comorbidities, 60.3% were married or living with a significant other, and 59.9% had an income equal to or greater than $45,000. A total of 75.0% were White/Caucasian, while the remaining 25.0% were Black, Hispanic, Asian, Native Hawaiian/Pacific Islander, American Indian/Alaskan Native, Other, or Multi-Racial. A total of 77.5% of patients had some college/technical education and 65.4% worked part or full time (Table 1).

More than half of all patients screened negative for clinical depression. (Fig. 3) Of those patients, 10.2% screened positive for depression on the PHQ-8, representing an “undiagnosed” population. Amongst the 46% of patients who screened positive for clinical depression, 20.2% screened positive for depression on the PHQ-8 while 79.8% had a negative PHQ-8 screening. (Figs. 2, 3).

Compared to patients with a positive clinical diagnosis for depression, the “undiagnosed” patients, those patients with a negative clinical screening for depression and a positive PHQ-8 depression screening, were more likely to be male (27.0 vs. 14.4%, \( p < .0001 \)), of higher BMI (49.5 ± 9.2 vs. 47.5 ± 8.5, \( p = .0002 \)), of minority race (34.0 vs. 20.5%, \( p < .001 \)), and employed (65.3 vs. 48.7%, \( p < .001 \)) (Fig. 3).

### Table 1 Demographics of patient population

<table>
<thead>
<tr>
<th>Procedure type</th>
<th>All MBSC</th>
<th>Positive clinical</th>
<th>Positive PHQ8</th>
</tr>
</thead>
<tbody>
<tr>
<td>RYGB (%)</td>
<td>16.0</td>
<td>17.1</td>
<td>15.0</td>
</tr>
<tr>
<td>LAGB (%)</td>
<td>0.8</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>LSG (%)</td>
<td>82.7</td>
<td>82.1</td>
<td>84.0</td>
</tr>
<tr>
<td>BPD/DS (%)</td>
<td>0.4</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>Age at procedure: mean (± SD)</td>
<td>45.8 (± 11.8)</td>
<td>46.2 (± 11.5)</td>
<td>45.5 (± 11.6)</td>
</tr>
<tr>
<td>Male (%)</td>
<td>21.5</td>
<td>14.4</td>
<td>20.9</td>
</tr>
<tr>
<td>BMI: mean (± SD)</td>
<td>47.8 (± 8.3)</td>
<td>47.5 (± 8.5)</td>
<td>48.4 (± 9.0)</td>
</tr>
<tr>
<td>Race and ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White or Caucasian, Non-Hispanic (%)</td>
<td>75.0</td>
<td>79.5</td>
<td>73.6</td>
</tr>
<tr>
<td>Other (Black, Hispanic, Asian, Native Hawaiian/Pacific Islander, American Indian/Alaskan Native, Other, or Multi-Racial) (%)</td>
<td>25.0</td>
<td>21.0</td>
<td>26.5</td>
</tr>
<tr>
<td>Income ≥ $45,000 per year (%)</td>
<td>51.9</td>
<td>47.7</td>
<td>42.9</td>
</tr>
<tr>
<td>Married or living with SO (%)</td>
<td>60.3</td>
<td>58.0</td>
<td>55.9</td>
</tr>
<tr>
<td>≥ 3 Total comorbidities (%)</td>
<td>83.6</td>
<td>92.8</td>
<td>87.2</td>
</tr>
<tr>
<td>Some college/technical school or greater (%)</td>
<td>77.5</td>
<td>77.5</td>
<td>70.8</td>
</tr>
<tr>
<td>Working part- or full-time (%)</td>
<td>65.4</td>
<td>57.7</td>
<td>54.9</td>
</tr>
</tbody>
</table>

**RYGB** Roux-en-y gastric bypass, **LAGB** laparoscopic adjusted gastric band, **LSG** laparoscopic sleeve gastrectomy, **BPD/DS** biliopancreatic diversion/duodenal switch

**Fig. 3** Stratification of patients based on clinical diagnosis of depression and the PHQ-8, a patient-reported depression screening tool. 10.2% of patients who were not clinically diagnosed with depression screened positive for depression via the patient-reported PHQ-8 (indicated in black).
However, patients with 3 or more comorbidities (78.2 vs. 92.8%, \( p < .0001 \)) were less likely to be “undiagnosed” (Fig. 4).

**Discussion**

This study demonstrates discordance between a traditional clinical diagnosis of depression and a previously validated, patient-reported method of diagnosing depression among Michigan bariatric surgery patients. Of the patients who did not clinically screen positive for depression, 10.2% tested positive for depression based on their PHQ-8 scores, potentially representing a missed or “undiagnosed” group of high risk patients. Additionally, patients who were minorities, males, and employed were more likely to be missed by traditional methods of preoperative clinical screening.

Prior studies have demonstrated that mental illness is prevalent amongst patients with obesity [3, 13]. Though weight loss generally has a positive impact on mental health [13, 14], the findings after weight loss due to bariatric surgery are less definitive [15–17]. In a study following bariatric surgery patients from 2006 to 2011, the rate of self-harm emergencies after surgery increased by 50%, with nearly all of the self-harm emergencies happening in patients with a previous mental health diagnosis [18]. Specifically, the authors of this study found that patients with a history of major depression accounted for almost all of the self-harm events. To mitigate these risks, over the last two decades, most bariatric surgery candidates in North America have been required to undergo preoperative evaluation by a mental health professional [19–22]. These evaluations are conducted via clinical interviews and symptom inventories. However, clinical interview practices can vary widely [8]. In fact, the American Society for Metabolic and Bariatric Surgery acknowledges that there is no consensus on the structure and content of these interviews despite their own recommendation for preoperative psychological evaluation [23]. This lack of standardization may lead to under-diagnosis of depression [24]. Previous studies have also noted that bariatric surgery candidates tend to present themselves favorably during psychological assessment which may also account for potential under-diagnosis of depression [25]. Our findings corroborate this by highlighting an undiagnosed group of patients undergoing bariatric surgery who may be susceptible to increased risk of postoperative self-harm.

We also identified a subset of patients who were clinically diagnosed with depression, while screening negative on the PHQ-8. This may represent a population who have been successfully treated and thus, no longer report symptoms consistent with depression. This is consistent with prior studies that demonstrate improved scores on patient-reported depression screening tools after appropriate treatment [26]. In a trial of depression treatment, PHQ-9 (a survey similar to the PHQ-8) responsiveness to treatment was validated over a 6-month span [26]. The study found that the PHQ-9 accurately reflects positive changes in depression after treatment. In the current study, the PHQ-8 focuses on patient-reported symptoms within the last month while a clinical diagnosis of depression is the result of any documentation of depression in the patient’s overall medical history. We hypothesize that the 37% of patients who clinically screened positive but screened negative using the PHQ8 represent an adequately treated patient population. In addition, we hypothesize that

![Fig. 4](image-url) Comparative demographics for positive clinical screening versus positive PHQ-8, negative clinical diagnosis. All results \( p < .05 \)
the lower rate of depression detected by the PHQ-8 in our study can also be attributed to the clinically diagnosed subset receiving treatment. This may further emphasize that accurate preoperative diagnosis may lead to appropriate treatment in the perioperative period. However, additional evaluation of these patients is required in order to completely understand this association.

Identifying patient populations at increased risk for missed depression diagnoses would be an important first step in improving the accuracy of diagnosis and treatment. Certain demographics are at risk for missed diagnoses when evaluated by clinical interview. Several studies demonstrate that despite the similar rates of depression seen in different ethnicities [27–29], the population-adjusted rate of office-based diagnoses of depression is significantly lower for Hispanic and Black patients compared to White patients [10]. Additionally, males are only half as likely as females to be diagnosed with depression [30, 31]. In a self-reported depressive symptom study, males were less likely to be diagnosed with depression than females with similar symptoms [32]. This difference may result from depressed males experiencing different symptoms than those symptoms currently included as diagnostic criteria [33]. Depression screening tools are reliant on traditional symptoms of depression, such as stress, irritability, problems with sleep, and loss of interest in things that are normally enjoyed. In one study, when authors included what they called more externalized symptoms, such as anger attacks/aggression, substance abuse, and risk-taking behavior in addition to those traditional symptoms, they found no gender difference in rates of depression [33]. This difference may also be explained by gender differences in health utilization of health care. Studies have shown that women utilize health care resources more frequently earlier in life [34, 35]. Another frequently undiagnosed group of patients are those who are employed [24]. In a study comparing depression diagnosis results between structured diagnostic interview and clinical interview, the undiagnosed group was entirely employed, demonstrating a similar trend to our data [24].

There are limitations with our study. First, our findings may not be generalizable to patients outside of the state of Michigan. However, our patient population (Table 1) is similar to those of other bariatric surgery centers participating in the nationwide Longitudinal Assessment of Bariatric Surgery (LABS) program [36]. Second, we do not have granular data on how the original clinical diagnosis of depression was made. For example, we are unsure whether some providers used clinical interviews, symptom inventories, objective personality/psychopathology tests, or one of many other methods used to diagnose depression [8]. However, we have no reason to believe that current preoperative traditional screening practices varied systematically across Michigan providers. Lastly, while this study demonstrates the value of the PHQ-8 screening tool in capturing a significant number of undiagnosed patients, we cannot claim the PHQ-8 to be the gold standard screening tool. There may be other tools with varying sensitivity. It is also important to note that the PHQ-8 is a screening test; however, as was noted in the Methods section, these screening tests have been validated for detecting depression.

The variation in screening for depression prior to bariatric surgery may result in an undiagnosed, untreated patient population. Currently, there is no consensus on best practice for diagnosing depression. Identifying best practices for preoperative mental health assessment may reduce the number of undiagnosed patients, leading to improvements in postoperative outcomes and lowering risk of self-harm. Additionally, focusing on populations at risk for being undiagnosed may improve the adequacy of diagnosis and treatment. We believe that a supplementary standardized screening modality should be added to the pre-op screening for bariatric surgery candidates to potentially improve surgical outcomes. This study can be used to motivate further investigation into methods of standardization of mental health evaluations for bariatric surgery patients.

Acknowledgements Sudarshan Srivatsan had full access to all data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Support for the MBSC is provided by Blue Cross Blue Shield of Michigan; however, the opinion beliefs, and viewpoints expressed by the author do not necessarily reflect those of BCBSM or any of its employees.

Compliance with ethical standards

Disclosures Dr. Ghaferi receives research funding from the Agency for Healthcare Research and Quality (K08 HS023621, P30 HS024403), the National Institute of Aging (R01 AG042340), and the Patient Centered Outcomes Research Institute (CE-1304-6596). He also receives salary support from Blue Cross Blue Shield of Michigan as the Director of the Michigan Bariatric Surgery Collaborative. Sudarshan Srivatsan, Vinay Guduguntla, Kelly Z. Young, Aliashgar Arastu, Cameron R. Strong, Ruth Cassidy have no conflicts of interest or financial ties to disclose.

References