Comparing PHQ-9 Scores between Low-income Adults With and Without Self-Reported Type 2 Diabetes

Erin R. Brummett and Mary Kunes-Connell

Creighton University
Abstract

**Objective:** Depression and diabetes often co-exist and result in worsening health outcomes. Low-income and demographic variables have been shown to have an effect on both diabetes and depression. The presence of a relationship between low-income, demographic variables, and comorbid diabetes and depression is unclear. This pilot study was designed to compare depressive symptoms as measured by PHQ-9 scores between low-income adults with self-reported type 2 diabetes and those without in relation to gender, age, ethnicity, and education level. **Methods:** Low-income adults at a food pantry were given a demographic questionnaire and PHQ-9 to complete. Presence of diabetes was based on self-report. Wilcoxon signed-rank test was used to compare means and medians of the groups. **Results:** Each age, ethnicity, education, and gender cohort’s PHQ-9 scores were compared to the scores of those with diabetes in the same cohort. No statistical significance was found in any category. **Discussion:** Limitations included small sample size and reliance on self-report. Research confirms the link between diabetes and depression and confirms as well that low-income and demographic variables impact both diabetes and depression separately. More research needs to be done to examine the impact of low income and demographic variables on comorbid depression and diabetes.
Comparing PHQ-9 Scores between Low-income Adults With and Without Self-Reported Type 2 Diabetes

At any given time, 12-18% of individuals with type 2 diabetes meet criteria for a diagnosis of depression (Park, Katon, & Wolf, 2013). Conversely, individuals with depression have a significantly higher risk of developing diabetes type 2 (Ali, Stone, Peters, Davies, & Khunti, 2006; Park et al., 2013). It is hypothesized that diabetes and depression are associated with similar metabolic and chemical changes in the body (Ali et al., 2006). Comorbid diabetes and depression increase the risk for diabetic complications, mortality, and disability (Ali et al., 2006; Katon et al., 2010; Park et al., 2013). Katon et al. (2008) found the increase in mortality to be as high as 38%. Egede & Ellis (2010) and Simon et al. (2005) found that depressed persons with diabetes had higher health care costs, spent more time receiving healthcare, and spent more time in inpatient facilities.

It is hypothesized that there appears to be a multifactorial relationship between depression, diabetes, and situational and demographic variables. (Auchincloss et al., 2009; Kanjilal et al., 2006; Lysy et. al, 2013; Osborn et al., 2011; & Mezuk et al, 2013; & World Federation for Mental Health, 2012). However, it is unclear what role these variables might play in this relationship (Auchincloss et al., 2009; Kanjilal et al., 2006; Lysy et. al, 2013; & Mezuk et al., 2013). Lenz, Pugh, Milligan, Gift & Suppe’s 1997 “Theory Of Unpleasant Symptoms” was adapted to illustrate the synergistic effect of physiologic, psychological, and situational factors on symptom manifestation, functionality, and quality of life (figure 1).

Figure 1

Multifactorial and Synergistic Relationship of Depression, Diabetes, and Situational/Demographic Variables
Objective

The primary objective of this pilot study was to describe depressive symptoms as measured by the 9-item Patient Health Questionnaire (PHQ-9) between low-income adults with self-reported type 2 diabetes and those without self-reported type 2 diabetes. The research question for this pilot study was "Is there a difference in PHQ-9 scores in low-income adults
with and without self-reported type 2 diabetes in relation to gender, age, ethnicity, or education level?”

Prompt detection and treatment of depression in patients with type 2 diabetes can positively impact clinical, quality, and cost outcomes. (Mitchell et al., 2009). However, there are many barriers to depression screening in patients with diabetes including culture, communication, and physician attitude (Leng, Changrani, Tseng, & Gany, 2009; Sorkin et al., 2011; Swenson, Rose, Vittinghoff, Stewart, & Schillinger, 2008). Ethnic minorities, for example, receive less depression screening (Leng et al., 2009; Sorkin et al., 2011) despite a higher incidence of depression in these populations (Wagner, Tsimikas, Abbott, de Groot, & Heapy, 2007). Primary care providers rarely use formal depression screening tools when screening does occur (Baik et al., 2010; Osborn, Kozak, & Wagner, 2010).

Few depression screening tools have been validated in patients with diabetes. Four screening tools appear most frequently in the literature for use with diabetes: the Beck Depression Inventory (BDI), the PHQ-9, the Hospital Anxiety and Depression Scale (HADS) and the CES-D (Roy, Lloyd, Pouwers, Holt, & Sartorius, 2011). The BDI, CES-D, and PHQ-9 all include questions on somatic complaints that may occur commonly in those with diabetes without depression. This may increase the false positive rate (Reddy, Philpot, Ford, & Dunbar, 2010; Roy et al., 2011; van der Feltz-Cornelis, 2011). Several studies examined ways to overcome this problem, including raising the cutoff values used for interpretation (Roy et al., 2011; van der Feltz-Cornelis, 2011; van Steenbergen-Weijenburg et al., 2010). Table 1 summarizes research studies examining the sensitivity and specificity of each test in regards to different cut-off values.
Roy et al.’s (2011) systematic review examined 234 studies on depression screening in persons with diabetes. They found the BDI, CES-D, HADS and PHQ-9 to have similar sensitivity and specificity. All four tools evidenced high rates of false positives and low rates of false negatives. Roy et al. recommend the BDI and the HADS, but note that these tools can be cost prohibitive. Though the CES-D had the highest specificity for depression in patients with type 2 diabetes, its length might be prohibitive in a primary care setting (McHale, Hendrikz, Dann, & Kenardy, 2008; Roy et al., 2011). Roy et al. also note that the PHQ-9 has similar sensitivity, reliability, and validity, though it is less specific than the CES-D.

Table 1

<table>
<thead>
<tr>
<th>Test</th>
<th>Sample Size</th>
<th>Type of Report</th>
<th>No. of</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>NPV (%)</th>
<th>PPV (%)</th>
<th>Cut Off Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHQ-9 Twist et al. (2013)</td>
<td>1500</td>
<td>Patient Self-report</td>
<td>9</td>
<td>95.24</td>
<td>70.42</td>
<td>99.3</td>
<td>24.2</td>
<td>≥ 10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>89.29</td>
<td>76.06</td>
<td>98.6</td>
<td>26.9</td>
<td>≥ 11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>86.9</td>
<td>80.28</td>
<td>98.4</td>
<td>30.4</td>
<td>≥ 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80.95</td>
<td>83.80</td>
<td>97.8</td>
<td>33.1</td>
<td>≥ 13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>77.38</td>
<td>87.32</td>
<td>97.5</td>
<td>37.6</td>
<td>≥ 14</td>
</tr>
<tr>
<td>van Steenbergen-Weijenburg et al. (2010)</td>
<td>197</td>
<td>Patient Self-report</td>
<td>9</td>
<td>91.9</td>
<td>59.4</td>
<td>34.3</td>
<td>96.9</td>
<td>≥ 8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>91.9</td>
<td>64.4</td>
<td>37.4</td>
<td>97.2</td>
<td>≥ 10</td>
</tr>
<tr>
<td>Test</td>
<td>Sample Size</td>
<td>Type of Report</td>
<td>No. of?</td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</td>
<td>NPV (%)</td>
<td>PPV (%)</td>
<td>Cut Off Score</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>----------------</td>
<td>---------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>---------</td>
<td>---------</td>
<td>---------------</td>
</tr>
<tr>
<td>CES-D</td>
<td></td>
<td>Patient Self-Report</td>
<td>20</td>
<td>70.5</td>
<td>71.4</td>
<td>85.2</td>
<td>50.8</td>
<td>≥ 13</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td></td>
<td></td>
<td>61.4</td>
<td>81.0</td>
<td>83.3</td>
<td>57.5</td>
<td>≥ 16</td>
</tr>
<tr>
<td>BDI-SF</td>
<td></td>
<td>Patient Self-Report</td>
<td>13</td>
<td>71 (CI 57-85)</td>
<td>81 (CI 76-86)</td>
<td>95 (CI 92-98)</td>
<td>38 (27-48)</td>
<td>≥ 8</td>
</tr>
<tr>
<td>Sultan, Luminet, &amp; Hartemann. (2011)</td>
<td>298</td>
<td></td>
<td></td>
<td>71 (CI 57-85)</td>
<td>81 (CI 76-86)</td>
<td>95 (CI 92-98)</td>
<td>38 (27-48)</td>
<td>≥ 8</td>
</tr>
<tr>
<td>HADS</td>
<td></td>
<td>Patient Self-Report</td>
<td>14</td>
<td>22.7</td>
<td>98.1</td>
<td>75.0</td>
<td>83.3</td>
<td>≥ 11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>79 (CI 66-91)</td>
<td>79 (CI 73-84)</td>
<td>96 (CI 93-98)</td>
<td>38 (27-48)</td>
<td>≥ 15</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td></td>
<td></td>
<td>79 (CI 66-91)</td>
<td>79 (CI 73-84)</td>
<td>96 (CI 93-98)</td>
<td>38 (27-48)</td>
<td>≥ 15</td>
</tr>
<tr>
<td></td>
<td>298</td>
<td></td>
<td></td>
<td>79 (CI 66-91)</td>
<td>79 (CI 73-84)</td>
<td>96 (CI 93-98)</td>
<td>38 (27-48)</td>
<td>≥ 15</td>
</tr>
</tbody>
</table>

Note. 95% confidence intervals in parentheses, NPV= negative predictive value, PPV= Positive predictive value, -------- indicates data not reported by authors.

Compared to the CES-D, BDI and HADS, the PHQ-9 has the largest body of evidentiary support when comparing sample sizes for use in this population. The tool is also a more reasonable length for use in screening. The aggregated data collected by Twist et al. (2013) and van Steenbergen-Weijenburg et al. (2010) suggest that a cut-off point of ≥ 10, the recommended cut-off in patients without diabetes, is not sensitive enough in the diabetic population. Both studies found that a cut-off point of ≥ 12 demonstrated sensitivity and specificity. None of these
studies were conducted among low-income populations; therefore this data may not be generalizable to low-income populations.

Methods

Design

The study was non-experimental and descriptive. Following approval by the institutional review board, participants were recruited from a low-income food pantry located in an underserved area of an urban, Midwestern city during a weekday morning. A flyer was posted for recruitment. Inclusion criteria included adults, ages 19 or above, with the ability to read and write in English. Informed consent was obtained from eligible adults who expressed interest. Participants were asked to anonymously complete a demographic survey and the PHQ-9 tool. If participants had difficulty reading the PHQ-9, they could request that the researcher read the tool aloud to them. Only one participant made this request. Those who completed the study received a food gift worth less than two dollars. A pamphlet listing free and reduced cost community mental health resources was offered to all participants.

Outcome Measures

Depressive symptoms were measured using the PHQ-9. The PHQ-9 scores were grouped, as suggested on the tool, with a score of 1-4 indicating minimal depression; 5-9 representing mild depression; 10-14 representing moderate depression; 15-19 representing moderate-severe depression; and 20-27 representing severe depression. The final question on the PHQ-9 addresses the impact of symptoms on daily life. This was not included in analysis or discussion for this study.

Demographic data were collected with regard to the presence or absence of self-reported diabetes, age, education level, and ethnicity. If participants self-reported diabetes they were
PHQ-9 SCORES IN LOW INCOME DIABETIC ADULTS

asked whether they were diagnosed as a child or adult. This was done to avoid confusion from using the terms type 1 and type 2 diabetes. Ethnicity was grouped using the categories from the United States Census. Figure 2 lists the demographic survey questions.

Figure 2

Demographic Survey

1. Have you ever been told you had diabetes?
   Are you first told you had diabetes when you were a child or an adult?
   □ 1 Child □ 2 Adult

2. What is your age? __________

3. What is your gender?
   □ 1 Male □ 2 Female

4. Are you Hispanic or Latino?
   □ 1 No □ 2 Yes

5. What is your race? (check one box)
   □ 1 American Indian or Alaska Native
   □ 2 Asian
   □ 3 Black or African American
   □ 4 Native Hawaiian or Other Pacific Islander
   □ 5 White
   □ 6 Other: ____________________________

6. How much schooling have you had? (Years of formal schooling completed)
   (check one box)
   □ 1 Did not complete high school
   □ 2 Completed high school
   □ 3 Received education beyond high school (includes attending any post high school classes)

Data Analysis

Statistical analysis was completed using the discrete score produced by the PHQ-9. The findings did not follow a normal distribution, therefore non-parametric analysis was completed using the Wilcoxon signed rank test to compare differences in the means, medians, and standard deviations (S.D.) between the group with self-reported diabetes and the group without in relation
to age, gender, ethnicity and education level. The score for each variable in the population with diabetes was compared to the score of the same variable in the population without diabetes.

Results

Fifty-five datasets were completed; seven were incomplete and excluded. Subsequently, 48 datasets were analyzed. Of the sample, 12 participants self-identified as having diabetes, with two of 12 self-reporting being diagnosed as children. The study was originally intended to focus on those with diabetes type 2. Due to the difficulty in obtaining a reliable self-report and the small number of those who reported child onset, all participants reporting diabetes were considered together in this study.

The PHQ-9 scores suggested some degree of depressive symptoms in all participants. Fifteen participants had PHQ-9 scores categorizing them as minimally depressed; 13 participants as mildly depressed; seven participants as moderately depressed; nine participants as moderate-severely depressed, and four participants as severely depressed. The 12 participants with self-reported diabetes had a PHQ-9 score mean of 11.9, median of 11.5, and S.D. of 2.31. The 36 participants without self-reported diabetes had a mean PHQ-9 score of 8.3, median of 7.0, and S.D. of 1.05. No statistically significant difference was found between the groups.

The participants ranged in age from 19-70. The largest group, aged 50-59, had a sample size of 18, with 4 people reporting diabetes. As seen in Table 2, the 30 to 39 age group had a much higher mean and median PHQ-9 score for those with self-reported diabetes compared to their counterparts without self-reported diabetes. The differences were not, however, statistically significant. This difference was not observed in other age groups. No statistically significant differences in PHQ-9 scores for those who self-reported diabetes compared to those without self-reported diabetes were seen in any age group.
Thirty-five females, eight with self-reported diabetes, and 13 males, four with self-reported diabetes, participated. Men with self-reported diabetes had a mean PHQ-9 score of 14.3, median score of 12.5, and a S.D. of 5.31. Men without self-reported diabetes had a mean PHQ-9 score of 5.6, median score of 3.0 and a S.D. of 1.98. Females with self-reported diabetes had a mean PHQ-9 score of 10.8, median score of 11.5, and a S.D. of 2.438. Females without self-reported diabetes had a mean PHQ-9 score of 9.2, a median score of 8.0, and a S.D. of 1.20. No statistically significant difference was found in regards to gender for those with and those without self-reported diabetes.

Sixteen participants identified themselves as Caucasian and thirty-two as non-Caucasian with 29 of those being Black, two as American Indian/Alaskan Native, and one as other. For
analysis all ethnicities were grouped into Caucasian (n=16) and non-Caucasian (n=32). Seven non-Caucasian participants self-reported diabetes; their PHQ-9 score was 11.1, median was 11.0, and S.D. was 2.81. Non-Caucasians without self-reported diabetes had a mean of 7.1, median of 5.0, and a S.D. of 1.13. Five Caucasian participants self-reported diabetes in this group, their PHQ-9 score mean was 13.0, median was 12.0, and S.D. was 4.25. Caucasians without self-reported diabetes had a mean of 11.1, median of 11.0, and a S.D. of 2.2. No statistically significant difference was found in regards to ethnicity for those with and those without self-reported diabetes.

Level of education was grouped into three categories: no high school completion, high school completion, and beyond high school completion. Nine participants identified themselves as not completing high school. Two of these nine self-reported diabetes and had a PHQ-9 score mean of 8.5, median of 8.5, and S.D. of 3.50. For those without self-reported diabetes who did not complete high school their score mean was 9.3 and median was 12.0. Nineteen participants completed high school. Three of these 19 self-reported diabetes; their PHQ-9 score mean was 15.0, median was 19.0, and S.D. was 4.51. The 16 participants without self-reported diabetes who completed high school had a mean of 7.6, median of 4.5, and a S.D. of 1.89. Twenty participants had education beyond high school. Seven of these participants self-reported diabetes. Their PHQ-9 score mean was 11.6, median was 11.0, and S.D. was 3.46. Participants without diabetes who completed beyond high school had a mean of 8.7, median of 8.0, and S.D. of 1.45. No statistically significant differences were found between those with and those without self-reported diabetes in regards to education level.

Discussion
This study was conducted to describe differences in depressive symptoms as evidenced by PHQ-9 scores in low-income adults with self-reported type 2 diabetes and those without self-reported type 2 diabetes in relation to demographic variables. Though the literature supports that rates of depression may likely be higher in this population, no statistically significant differences were found in any of the groupings. However, the small sample size may preclude such findings. Kanjilal et al. (2007), Mezuk et al. (2013), and Osborn et al.’s (2011) larger studies found significant correlations between depression and diabetes in low-income populations.

Kanjilal et al. (2007) examined four national, cross-sectional surveys, administered in the United States from 1971-2002. Mezuk et al.’s (2013), study of 336,340 participants in the United Kingdom, living in high deprivation neighborhoods, found significantly higher rates of both depression and diabetes than in affluent neighborhoods. Not only does there appear to be a significant relationship between low income and incidence of diabetes, the disparity between people with high versus low income developing diabetes appears to be increasing, having tripled in the United States from 1971 to 2002 (Kanjilal et al., 2007)). Osborn et al.’s (2011) study of 69,000 low-income adults found that 42% of the population had symptoms of depression. The rate of depression was significantly higher in those with diabetes.

Demographic variables have been shown to be involved in this relationship. Kanjilal et al. (2007) found that education and literacy level played a factor. A chart review examining 88,886 patients concluded that the relationship between low income and incidence of diabetes was stronger in females (Lysy et al., 2013). Mezuk et al. (2013) found that low-income women, Caucasians, the less educated, and the un-employed tended to be at increased risk for both depression and diabetes.
Though the PHQ-9 shows both validity and reliability in the population with diabetes, there is evidence suggesting that the standard cut-off score of $\geq 10$ used for interpretation results in false positives. This may be due, in part, to the overlap of somatic symptoms of depression and diabetes (Reddy et al., 2010; Roy et al., 2011; van der Feltz-Cornelis, 2011).

The PHQ-9 recommends using a score of $\geq 10$ to indicate the presence of depression in patients without diabetes. Various studies have examined the specificity and sensitivity of the PHQ-9 in the persons with diabetes using different cut-off scores to overcome high rates of false positives using the cut-off score of $\geq 10$. Van Steenbergen-Weijenburg et al. (2010) conducted a study comparing 197 diabetic patients’ PHQ-9 scores. The PHQ-9 was less sensitive, but more specific if the cut-off score was raised to $\geq 12$, rather than $\geq 10$. The results of the study by van Steenbergen-Weijenburg et al. were sufficient for van der Feltz-Cornelis (2011) to recommend use of the PHQ-9 in patients with diabetes using the cutoff score of $\geq 12$. Twist et al.’s (2013) large study ($n=1500$), examined different cut-off points to screen for depression in patients with type 2 diabetes using the PHQ-9. They concluded that $\geq 12$ was the ideal balance of sensitivity and specificity.

Despite greater than 30 participants in this study being classified by their PHQ-9 score as moderately depressed, only 5 participants accepted the mental health resources pamphlet. Several factors could account for this low rate. The PHQ-9 may have a high rate of false positives in this sample secondary to cultural, socioeconomic, or health literacy factors. Somatization and the denial of depression play a part in recognizing depression in minority populations (Leng et al., 2009). Hispanic ethnicity, less education, and lower income were all predictors of self-reporting more than one reason to deny depression symptoms to a physician. Wagner, White Perkins, Piette, Lipton, and Aikens (2009) reported that African-American
patients with diabetes were 5.9 times less likely than Caucasian patients to discuss their symptoms of depression with anyone \( (p = 0.008) \). The reasons reported included believing they should be able to handle it themselves, fear of being placed on medication, and fear of being referred to a counselor (Bell et al., 2011).

**Limitations**

A major limitation of this pilot study was small sample size. The number of participants was too small to obtain any significance, even if such significance does exist in the population. Prior to data collection a power analysis was completed and a sample size of \( >350 \) was required to obtain an 80\% power. This was determined to not be feasible with the available resources. Future research could be done using the same method, over an increased number of days, to increase the sample size. A larger multi-site sample, comparing low-income to higher income populations, could elucidate significance that is unable to be seen in this study.

Another limitation was in the self-reporting process. Confusion of the terms ‘type 1’ or type ‘2’ resulted in using the terms ‘child onset’ or ‘adult onset’. This prohibited differentiating the types as was originally planned. Confusion over presence or absence of diabetes was not anticipated but was observed. At least 3 respondents asked the researcher if having “pre-diabetes” or “high blood sugar” qualified as having diabetes. To prevent bias respondents were asked to use their best judgment in answering the question. Similar studies identified diabetes through a medical chart review rather than self-report. Inaccurate self-reporting may skew this study’s results. Health literacy potentially impacted participants’ results. Health literacy was a factor that commonly occurred in the literature review and further research should consider this in the context of self-report questionnaires.

**Conclusion**
Many studies have suggested improvement in disease and quality of life outcomes when comorbid depression is treated those with diabetes, including an improvement in hemoglobin A1C and a lower cost of health care. (Echeverry Duran, Bonds, Lee, & Davidson, 2009; Lamers Jonkers, Bosma, Knottnerus, & van Eijk, 2011; Lin et al., 2012; Katon et al., 2010; Simon et al., 2007). Both depression and diabetes appear to be influenced by both low-income and demographic variables, though the relationship is unclear. The small sample size of this study prevented finding significance if any does exist; a larger sample size would help elucidate the relationship between demographic variables in a low-income population with comorbid diabetes and depression. Regardless, based on a general review of existing evidence, depression should be considered in patients with type 2 diabetes, particularly those with a low-income. Screening, using a low-cost, user-friendly tool, in a primary care setting, may result in earlier identification, timely treatment, and more positive outcomes for this population.
References


