Screening for Anxiety and Depressive Symptoms in Type 2 Diabetes Using Patient-Reported Outcome Measures: Comparative Performance of the EQ-5D-5L and SF-12v2

Fatima Al Sayah, Arto Ohinmaa, and Jeffrey A. Johnson

Abstract

Objective. To evaluate the performance of EQ-5D-5L (EuroQol Five-Dimension, Five-Level Questionnaire) and SF-12-v2 (12-item Medical Outcomes Health Survey–Short Form, Version 2) in screening for anxiety and depressive symptoms in adults with type 2 diabetes. **Methods.** Cross-sectional data from a population-based study of type 2 diabetes in Alberta, Canada, were used. Anxiety symptoms (using the 2-item Generalized Anxiety Disorder questionnaire) were categorized into absent (<3) versus present (≥3). Depressive symptoms (using the 8-item Patient Health Questionnaire) were categorized according to two severity cut-points: absent (<10) versus mild (≥10), and absent (<15) versus moderate-severe (≥15). The performance of the measures in screening for anxiety and depressive symptoms was evaluated using receiver operating curve (ROC) analysis. **Results.** Average age of participants ($N = 1,391$) was 66.8 years (SD 10.2), and 47% were female. Seventeen percent of participants screened positive for mild and 5.9% for moderate-severe depressive symptoms, and 11.3% for anxiety symptoms. For comorbid symptoms, 8.6% screened positive for anxiety and any depressive symptoms, and 4.6% for anxiety and moderate-severe depressive symptoms. The EQ-5D-5L anxiety/depression dimension and the SF-12 mental composite summary score had the best performance in screening for anxiety (area under ROC: 0.89, 0.89, respectively), depressive symptoms (any: 0.88, 0.92; moderate-severe: 0.90, 0.90), and comorbid anxiety and depressive symptoms (any: 0.92, 0.91; moderate-severe: 0.92, 0.90). These were followed by SF-12 feeling downhearted/depressed item (range = 0.83–0.85), while the lowest performance was for the EQ-5D-5L index score (0.80–0.84) and the SF-12 mental health domain (0.81–0.82). **Conclusion.** The EQ-5D-5L and the SF-12 are suitable tools for screening for anxiety and depressive symptoms in adults with type 2 diabetes. These tools present a unique opportunity for a standardized approach for routine mental health screening within the context of routine outcome measurement initiatives, where screening is recommended.

School of Public Health, University of Alberta, Edmonton, Alberta, Canada (FAS, AO, JAJ). The work was completed at the School of Public Health, University of Alberta, Edmonton, Alberta Canada. The work was presented at the ISPOR 22nd Annual International Meeting in May 20–24, 2017, in Boston, MA. All authors are members of the EuroQol Group. All authors have no other conflicts of interest to declare. This study was conducted at Alberta PROMs and EQ-5D Research and Support Unit (APERSU), which is funded by Alberta Health Services and Health Quality Council of Alberta.

Corresponding Author
Jeffrey A. Johnson, 2-040 Li Ka Shing Centre for Health Research Innovation, School of Public Health, University of Alberta, Edmonton, Canada AB T6G 2E1; telephone: 780-492-7317 (jeff.johnson@ualberta.ca).

Diabetes is a debilitating chronic condition that has considerable health and economic burden on public health systems around the world. The global prevalence of diabetes has doubled in the last 30 years reaching 8.5% in 2016.[1] In Canada, one in nine adults currently has diabetes, a prevalence that is estimated to increase by 44%

by the year 2025.[2] The burden of diabetes is primarily due to its complications and other comorbid chronic conditions. Mental health disorders, primarily anxiety and depression, are the most common conditions that coexist with diabetes, especially in older patients.[3–5] It is estimated that 14% of diabetes patients have anxiety symptoms,[6] and 19% have depressive symptoms.[7] However, despite the potential adverse effects of these mental health problems on diabetes outcomes and health care costs,[8–10] only one third of patients receive diagnosis and treatment.[3,11]

The Canadian Diabetes Association recommends that "all individuals with diabetes should be regularly screened for the presence of depressive and anxious symptoms."[12] Clinicians usually use common standardized measures of screening for these symptoms such as the Patient Health Questionnaire (PHQ) and the Generalized Anxiety Disorder (GAD) questionnaire; however, there is no uniform approach of screening for these symptoms in Canadian health care settings. Other patient-reported outcome measures (PROMs) of general health status have been increasingly used for routine outcome measurement in health systems in Canada and internationally. The EuroQol five dimensions questionnaire (EQ-5D) and the 12-item Medical Outcomes Health Survey–Short Form (SF-12) are the most commonly used generic measures in routine health outcome measurement initiatives. These measures, and other generic PROMs, assess mental health as a dimension of overall health using single or multiple items. The routine use of these tools in health care systems presents a unique opportunity for a standardized approach for routine mental health screening, if these tools are fit for this purpose.

There is a large body of evidence on the measurement properties of the EQ-5D and SF-12 measures, demonstrating their reliability and validity in various populations, including those with mental health conditions.[13–15] In recent years, there has been a considerable move from using these tools in research to using them in clinical settings.[16,17] This development has led to questions about their usefulness in individual patient management, particularly for those with chronic conditions. The high prevalence and costs of mental health problems in the context of diabetes, combined with evidence that behavioral factors are important for effective diabetes self-management, create an important opportunity to integrate mental health screening and treatment into multidisciplinary diabetes care, to improve patient and public health outcomes, and to help decrease health care expenditures. Given the crucial need for anxiety and depression screening in this population, we sought to examine whether these measures could be suitable as mental health screening tools in this patient population. Our objective was to evaluate the performance of the 5-level version of the EQ-5D (EQ-5D-5L) and version 2 of the SF-12 in screening for anxiety and depressive symptoms in a population-based sample of adults with type 2 diabetes.

Methods

Data Source

Cross-sectional data from an ongoing cohort study of adults with type 2 diabetes in Alberta, Canada, were used in this study. A detailed description of the cohort design and procedures was reported elsewhere.[18] Briefly, type 2 diabetes patients, over 18 years old and who were able to communicate in English, were eligible to participate in this study. Participants were recruited over a 2-year period (December 2011 to December 2013) using several approaches and strategies including invitations through primary care networks, diabetes clinics, and various forms of public advertisements. Data were collected between 2013 and 2015 via a postal self-administered survey that included several measures and sociodemographic questions. All participants provide a written informed consent. The sample was considered to be generally representative of the diabetes population in Alberta.[18] This study did not receive any external funding.

Measures

The EQ-5D-5L is a preference-based measure of general health status. Its classification system consists of five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), each with five levels of problems: no "1," slight "2," moderate "3," severe "4," and extreme or unable to perform a task "5."[19] Each health state is described as a vector of these five dimensions. With five levels for each dimension, the EQ-5D-5L describes 3,125 distinct health states, with 11111 being the best health state (full health) and 55555 the worst health state. The EQ-5D reference period is today. We used the recently developed preference-based index scoring for the Canadian population, which has a range from -0.148 for the worst (55555) to 0.949 for the best (11111) health states.[20]

The SF-12 (version 2) is a non–preference-based generic measure of health status. It includes 12 items (one specific to mental health, i.e., feeling downhearted or depressed), and yields 8 domain scores (physical functioning, role limitations due to physical problems, bodily

pain, general health, vitality, social functioning, role limitations due to emotional problems, and mental health [MH]), as well as physical and mental composite summary scores (PCS and MCS).[21] The SF-12 reference period is the last 4 weeks. The domain scores were transformed into T scores with a mean of 50 and standard deviation (SD) of 10.

Anxiety symptoms were assessed using the 2-item Generalized Anxiety Disorder (GAD2) questionnaire. GAD2 assesses the presence and severity of anxiety symptoms within the past 2 weeks.[22] It includes two items, each scored from 0 (not at all) to 3 (nearly every day), with higher scores indicating worse anxiety symptoms. A total score is computed as the sum of the scores of the two items (range = 0–6), with a cutoff point of ≥ 3 indicating the presence of anxiety symptoms.[22] As there is no established minimal important difference for GAD2, we applied the half standard deviation (1/2 SD) rule of thumb,[23] and the total score was categorized as follows: present (GAD2 ≥ 3) versus absent (GAD2 < 3) anxiety symptoms. The sensitivity and specificity of the GAD2 in detecting generalized anxiety disorder in primary care were reported to be 86% and 83%, respectively.[22]

Depressive symptoms were assessed using the 8-item Patient Health Questionnaire (PHQ8), a self-reported measure based on the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition) criteria for major depression.[24] PHQ8 assesses the presence of depressive symptoms within the past 2 weeks. Each of the PHQ8 items is scored from 0 (not at all) to 3 (nearly every day), with higher scores indicating worse depressive symptoms. A total score was computed as the sum of the scores of the 8 items (range = 0–24).[24] A 5-point difference or change in the PHQ8 score is considered clinically important.[25] Based on a meta-analysis on the use of PHQ in screening for major depressive disorder, the pooled sensitivity and specificity estimates were 78% and 87%, respectively.[26] The total PHQ8 score was categorized into two severity levels:

- Any depressive symptoms (PHQ8 ≥ 10) versus absent depressive symptoms (PHQ8 < 10)[25]
- Moderate-severe depressive symptoms (PHQ8 ≥ 15) versus absent moderate-severe depressive symptoms (PHQ8 < 15)[25]

Statistical Analysis

Descriptive statistics were computed for all variables in the overall sample. Differences in mean index scores, distribution of problems in the anxiety/depression dimension, and SF-12 MCS, MH, and distribution of feeling downhearted/depressed item were examined using t test and chi-square test as appropriate. Spearman correlation coefficient was used to examine the correlations between EQ-5D-5L and SF-12 components with GAD2 and PHQ9 scores. The performance of each of the components of the EQ-5D-5L (index score and anxiety/depression dimension) and the SF-12 (MCS, MH domain, and feeling downhearted/depressed item) in screening for anxiety and depressive symptoms was examined using AUROC (area under receiver operating curve) analysis. GAD2 and PHQ8 were used to define presence of anxiety and depressive symptoms, alone or in combination, as follows:

- Anxiety symptoms present (GAD2 ≥ 3) versus absent (GAD2 < 3)
- Any level of depressive symptoms (PHQ8 ≥ 10) versus absent (PHQ8 < 10)
- Moderate-severe depressive symptoms (PHQ8 ≥ 15) versus absent (PHQ8 < 15)
- Comorbid anxiety (GAD2 ≥ 3) and any level of depressive symptoms (PHQ8 ≥ 10) versus absent (GAD2 < 3 and PHQ8 < 10)
- Comorbid anxiety (GAD2 ≥ 3) and moderate-severe depressive symptoms (PHQ8 ≥ 15) versus absent (GAD2 < 3 and PHQ8 < 15)

The EQ-5D-5L index score was categorized into quintiles. The anxiety/depression dimension has five levels: 1 = no, 2 = mild, 3 = moderate, 4 = severe, and 5 = extreme. The SF-12 MCS score was also categorized into quintiles. The MH domain score was categorized into quartiles, while the "feeling downhearted/depressed" item has five levels: 1 = all the time, 2 = most of the time, 3 = some of the time, 4 = little of the time, and 5 = none of the time. Scores for the EQ-5D-5L index score quintiles, "feeling downhearted/depressed" item, MH domain quartiles, and MCS quintiles were reversed so that higher scores indicate more problems.

To identify the optimal threshold for each of the examined components, that is, the cut-point that maximizes sensitivity and specificity, we calculated sensitivity, specificity, and positive (LR +) and negative (LR −) likelihood ratios for presence of anxiety symptoms and the two severity levels of depressive symptoms at each cut-point (≥ 2, ≥ 3, ≥ 4, ≥ 5), and an overall AUROC with 95% confidence interval (CI). AUROC values were interpreted as follows: ≤ 0.5 noninformative; $0.5 < \text{AUROC} \leq 0.7$ less accurate; $0.7 < \text{AUROC} \leq 0.9$ moderately accurate; $0.9 < \text{AUROC} < 1.0$ highly

accurate; and AUROC = 1 perfect test.[27] All analyses were performed in STATA 13.1.[28]

Results

General Characteristics of Participants

Average age of participants (N = 1,391) was 66.8 years (SD 10.2); 47% were female, 49% had more than high school education, 24.5% had an annual household income of $\geq$\$80,000, and the majority (92.8%) were Caucasian. Mean GAD2 score was 0.95 (SD 1.4), with 11.3% screening positive for anxiety symptoms (GAD2 $\geq$3), and mean PHQ8 was 5.0 (SD 5.1), with 17.0% screening positive for any depressive symptoms (PHQ8 $\geq$10) and 5.9% for moderate-severe depressive symptoms (PHQ8 $\geq$15). Mean EQ-5D-5L index score was 0.79 (SD 0.17), with 53.4% reporting problems (levels 2–5) in mobility, 14.3% in self-care, 49% in usual activities, 74.6% in pain/discomfort, and 43.9% in anxiety/depression. Mean MCS score was 48.2 (SD 9.9), PCS score was 43.8 (SD 10.9), and MH score was 50.3 (SD 10.2), while 56.9% reported feeling downhearted/depressed.

Individuals with anxiety and/or depressive symptoms had a considerably lower mean EQ-5D-5L index score compared with those without any symptoms (0.81): anxiety (0.59), any depressive symptoms (0.59), moderate-severe depressive symptoms (0.53), comorbid anxiety and any depressive symptoms (0.55), and comorbid anxiety and moderate-severe depressive symptoms (0.51) (Table 1). A similar pattern was observed for SF-12 MCS and MH domain, whereby scores were lower with the presence of anxiety and/or depressive symptoms. For the EQ-5D-5L anxiety/depression item and the SF-12 feeling downhearted/depressed item, the proportion of patients reporting more problems and frequency of symptoms were higher in those with anxiety, depressive symptoms, or both, compared with those without any symptoms.

Correlations of SF-12 components with GAD2 and PHQ8 scores were strong for MCS (−0.69 and −0.79, respectively), MH (−0.68 and −0.68, respectively), and feeling downhearted/depressed item (−0.64 and −0.63, respectively). These correlations for EQ-5D-5L ranged from moderate (GAD2: −0.48) to strong (PHQ8: −0.63) for the index score, and were strong for the anxiety/depression dimension (GAD2: 0.72; PHQ8: 0.69).

Performance of EQ-5D-5L and SF-12 in Screening for Anxiety Symptoms

All EQ-5D-5L and SF-12 components performed very well in screening for anxiety symptoms as assessed by

GAD2 (Table 2, Figure 1a), with the highest AUROC observed for EQ-5D-5L anxiety/depression dimension (0.89; 95% CI: 0.86, 0.91) and SF-12 MCS (0.89; 95% CI: 0.87, 0.91), followed by SF-12 feeling downhearted/depressed item (0.84; 95% CI: 0.81, 0.87) and MH domain (0.82, 95% CI: 0.80, 0.84), whereas the lowest performance was for the EQ-5D-5L index score (AUROC 0.79; 95% CI: 0.76, 0.82). Optimal performance was at cut-point $\geq$4 for the EQ-5D-5L index score (sensitivity 85.0%), cut-point $\geq$2 for the anxiety/depression dimension (sensitivity 96.8%), cut-point $\geq$5 for the MCS (sensitivity 97.2%), cut-point $\geq$4 for MH domain (sensitivity 91.5%), and cut-point $\geq$3 for feeling downhearted/depressed item (sensitivity 80.0%) (Table 2).

Performance of EQ-5D-5L and SF-12 in Screening for Depressive Symptoms

All EQ-5D-5L and SF-12 examined components performed very well at screening for depressive symptoms (both severity levels), as assessed by the PHQ8, and overall, the performance was slightly better than that for anxiety symptoms (Table 3, Figure 1b and 1c) with moderate-highly accurate AUROC values. For both levels of depressive symptoms, the best performance was for MCS (any: AUROC 0.92, 95% CI: 0.91, 0.93; moderate-severe: 0.90, 95% CI: 0.89, 0.92), followed by the EQ-5D-5L anxiety/depression dimension (any: 0.88, 95% CI: 0.86, 0.90; moderate-severe: 0.90, 95% CI: 0.87, 0.93), while the poorest performance was for the SF-12 MH domain (any: 0.82, 95% CI: 0.80, 0.85; moderate-severe: 0.81, 95% CI: 0.79, 0.83). For any depressive symptoms, the optimal performance was at cut-point $\geq$4 for EQ-5D-5L index score (sensitivity 84.6%), cut-point $\geq$2 for anxiety/depression dimension (sensitivity 95.3%), cut-point $\geq$4 for MCS (sensitivity 98.2%), cut-point $\geq$4 for MH domain (sensitivity 86.7%), and cut-point $\geq$3 for feeling downhearted/depressed item (74.0%). For moderate-severe depressive symptoms, the cut-point for optimal performance varied slightly, with generally lower sensitivity (Table 3).

Performance of EQ-5D-5L and SF-12 in Screening for Comorbid Anxiety and Depressive Symptoms

For comorbid anxiety and any depressive symptoms, the EQ-5D-5L anxiety/depression dimension had the best performance with a highly accurate AUROC of 0.92 (95% CI: 0.90, 0.94), followed by the SF-12 MCS (0.91; 95% CI: 0.89, 0.92), while the lowest performance was

Table 1 EQ-5D-5L and SF-12 Components in the Overall Sample and by Anxiety and Depressive Symptoms Subgroups[a]

	Overall (N = 1,391)	No Anxiety or Depressive Symptoms[b] (N = 1,215)	Anxiety Present[c] (N = 157)	Any Depressive Symptoms[d] (N = 237)	Moderate-Severe Depressive Symptoms[e] (N = 82)	Anxiety and Any Depressive Symptoms[f] (N = 114)	Anxiety and Moderate-Severe Depressive Symptoms[g] (N = 61)
EQ-5D-5L							
Index score (mean ± SD)	0.79 ± 0.17	0.81 ± 0.15	0.59 ± 0.22	0.59 ± 0.22	0.53 ± 0.23	0.55 ± 0.22	0.51 ± 0.24
Anxiety/depression, n (%)							
Level 1: None	768 (55.7)	738 (61.2)	5 (3.3)	11 (4.7)	0 (0)	0 (0)	0 (0)
Level 2: Mild	400 (29.0)	353 (29.3)	45 (29.2)	90 (38.5)	23 (28.4)	27 (24.1)	13 (21.7)
Level 3: Moderate	177 (12.8)	111 (9.2)	73 (47.4)	102 (43.6)	35 (43.2)	56 (50.0)	24 (40.0)
Level 4: Severe	27 (2.0)	3 (0.3)	25 (16.2)	25 (10.7)	18 (22.2)	23 (20.5)	18 (30.0)
Level 5: Extreme	6 (0.4)	0 (0)	6 (3.9)	6 (2.6)	5 (6.2)	6 (5.4)	5 (8.3)
SF-12							
MCS (mean ± SD)	48.2 ± 9.9	49.8 ± 8.5	33.2 ± 8.0	34.6 ± 7.2	30.3 ±7.8	31.4 ± 7.4	28.5 ± 8.1
MH domain (mean ± SD)	50.3 ± 10.2	51.9 ± 8.8	36.0 ± 9.3	38.4 ± 9.3	34.5 ± 10.1	34.4 ± 9.3	32.1 ± 10.3
Feeling downhearted or depressed, n (%)							
None of the time	586 (42.5)	562 (46.7)	7 (4.5)	12 (5.1)	3 (3.7)	3 (2.7)	3 (5.0)
Little of the time	385 (27.9)	349 (29.0)	24 (15.5)	49 (20.9)	11 (13.6)	17 (15.2)	7 (11.7)
Some of the time	309 (22.4)	251 (20.9)	70 (45.2)	109 (46.4)	31 (38.3)	44 (39.3)	18 (30.0)
Most of the time	77 (5.6)	34 (2.8)	42 (27.1)	54 (23.0)	26 (32.1)	37 (33.0)	22 (36.7)
All of the time	21 (1.5)	8 (0.7)	12 (7.7)	11 (4.7)	10 (12.4)	11 (9.8)	10 (16.7)

EQ-5D-5L, EuroQol Five-Dimension, Five-Level Questionnaire; GAD2, 2-item Generalized Anxiety Disorder questionnaire; MCS, mental composite summary; MH, mental health; PHQ8, 8-item Patient Health Questionnaire; SD, standard deviation; SF-12, 12-item Medical Outcomes Health Survey–Short Form.

[a]Differences in mean EQ-5D-5L index score, proportion of problems in the anxiety/depression dimension, mean SF-12 MCS and MH domain, and proportion of frequency in the downhearted/depressed item, between those with and without symptoms in all examined groups ([b-g]) were statistically significant at $P < 0.05$ based on t tests and chi-square tests.

[b]GAD2 <3 and PHQ8 <10.
[c]GAD2 ≥3 and PHQ8 <10.
[d]GAD2 <3 and PHQ8 ≥10.
[e]GAD2 <3 and PHQ8 ≥15.
[f]GAD2 ≥3 and PHQ8 ≥10.
[g]GAD2 ≥3 and PHQ8 ≥15.

Table 2 Performance of EQ-5D-5L and SF-12 Components in Screening for Anxiety Symptoms (GAD2 ≥3)[a]

Cut-Point	Sensitivity	Specificity	LR+	LR−	AUROC (95% CI)
EQ-5D-5L index score quintiles					
Overall					0.79 (0.76, 0.82)
≥2	98.7%	22.4%	1.27	0.06	
≥3	95.4%	40.6%	1.61	0.11	
≥4	**85.0%**	**62.8%**	**2.28**	**0.24**	
≥5	54.9%	84.3%	3.50	0.53	
EQ-5D-5L Anxiety/depression dimension					
Overall					0.89 (0.86, 0.91)
≥2	**96.8%**	**62.3%**	**2.56**	**0.05**	
≥3	67.5%	91.4%	7.82	0.36	
≥4	20.1%	99.8%	121.18	0.80	
≥5	3.9%	100.0%	—	0.96	
SF-12 MCS quintiles					
Overall					0.89 (0.87, 0.91)
≥2	99.3%	22.5%	1.28	0.03	
≥3	98.6%	44.8%	1.79	0.03	
≥4	**97.2%**	**67.1%**	**2.95**	**0.04**	
≥5	78.9%	87.1%	6.12	0.24	
SF-12 MH quartiles					
Overall					0.82 (0.80, 0.84)
≥2	100.0%	15.1%	1.18	0.00	
≥3	98.7%	45.4%	1.81	0.03	
≥4	**91.5%**	**69.3%**	**2.98**	**0.12**	
SF-12 Feeling downhearted/depressed item					
Overall					0.84 (0.81, 0.87)
≥2	95.5%	47.6%	1.82	0.10	
≥3	**80.0%**	**77.1%**	**3.50**	**0.26**	
≥4	34.8%	96.6%	10.22	0.67	
≥5	7.7%	99.3%	10.39	0.93	

AUROC, area under receiver operating curve; CI, confidence interval; LR−, negative likelihood ratio; LR +, positive likelihood ratio; MCS, mental composite summary; MH, mental health; SF-12, 12-item Medical Outcomes Health Survey–Short Form.
[a]EQ-5D-5L index score: 2 = quintile 2, 3 = quintile 3, 4 = quintile 4, 5 = quintile 5. EQ-5D-5L anxiety/depression dimension: 2 = mild problems, 3 = moderate problems, 4 = severe problems, 5 = extreme problems/unable to perform. SF-12 MCS quintiles: 2 = quintile 2, 3 = quintile 3, 4 = quintile 4, 5 = quintile 5. SF-12 MH quartiles: 2 = quartile 2, 3 = quartile 3, 4 = quartile 4. SF-12 feeling downhearted/depressed item: 2 = little of the time, 3 = some of the time, 4 = most of the time, 5 = all of the time. Bolded values indicate the cutoff that maximizes sensitivity and specificity for each component.

for the SF-12 MH domain (0.82; 95% CI: 0.80, 0.84), which was moderately accurate (Table 4, Figure 1d). Similar results were observed for comorbid anxiety and moderate-severe depressive symptoms (Table 4, Figure 1e), where the EQ-5D-5L anxiety/depression dimension had the highest AUROC (0.92; 95% CI: 0.89, 0.94), and the SF-12 MH domain had the lowest AUROC (0.81; 95% CI: 0.79, 0.83). For comorbid anxiety and any depressive symptoms, the optimal performance was at cut-point ≥4 for EQ-5D-5L index score (sensitivity 89.2%), cut-point ≥3 for anxiety/depression dimension (sensitivity 75.9%), cut-point ≥5 for MCS (sensitivity 86.5%), cut-point ≥4 for MH domain (sensitivity 93.7%), and cut-point ≥3 for feeling downhearted/depressed item (sensitivity 82.1%). The cut-points for optimal performance were similar for comorbid anxiety and moderate-severe depressive symptoms with varying degrees of sensitivity and specificity.

Discussion

We found that the EQ-5D-5L (index score and anxiety/depression dimension) and SF-12 (MCS, MH domain, and "feeling downhearted/depressed" item) generally performed very well in screening for anxiety and/or depressive symptoms in this sample of adults with type 2 diabetes. In particular, the anxiety/depression dimension of the EQ-5D-5L and the MCS score of the SF-12 had similarly strong performance in screening for these symptoms compared to the reference measures (GAD2 and PHQ8) with highly accurate AUROC values. These findings are consistent with previous studies that found the SF-12 and SF-36 to be suitable tools for screening for anxiety and/or depression in the general population,[29,30] as well as in patients with coronary heart disease,[31] and those with chronic spinal pain.[32] Although there is an abundance of evidence on the use of EQ-5D instruments for measuring health status in individuals with mental health problems,[13–15] we are aware of only one study that suggested the EQ-5D-3L anxiety/depression dimension as a useful measure to differentiate among individuals with major depressive episode and anxiety and their counterparts,[33] and thus our results are consistent with those from that study.

The World Health Organization screening criteria, also known as the Wilson and Jungner criteria,[34] include 10 items that should be met in order to endorse screening for a particular disease. Having a "suitable test or examination" is one of these criteria that, on its own, is not sufficient to recommend screening. In this study, we found the EQ-5D-5L and SF-12 to be suitable tools to screen for anxiety and depressive symptoms in patients with diabetes in primary care. Nonetheless, other criteria need to be examined and addressed; such exploration is not
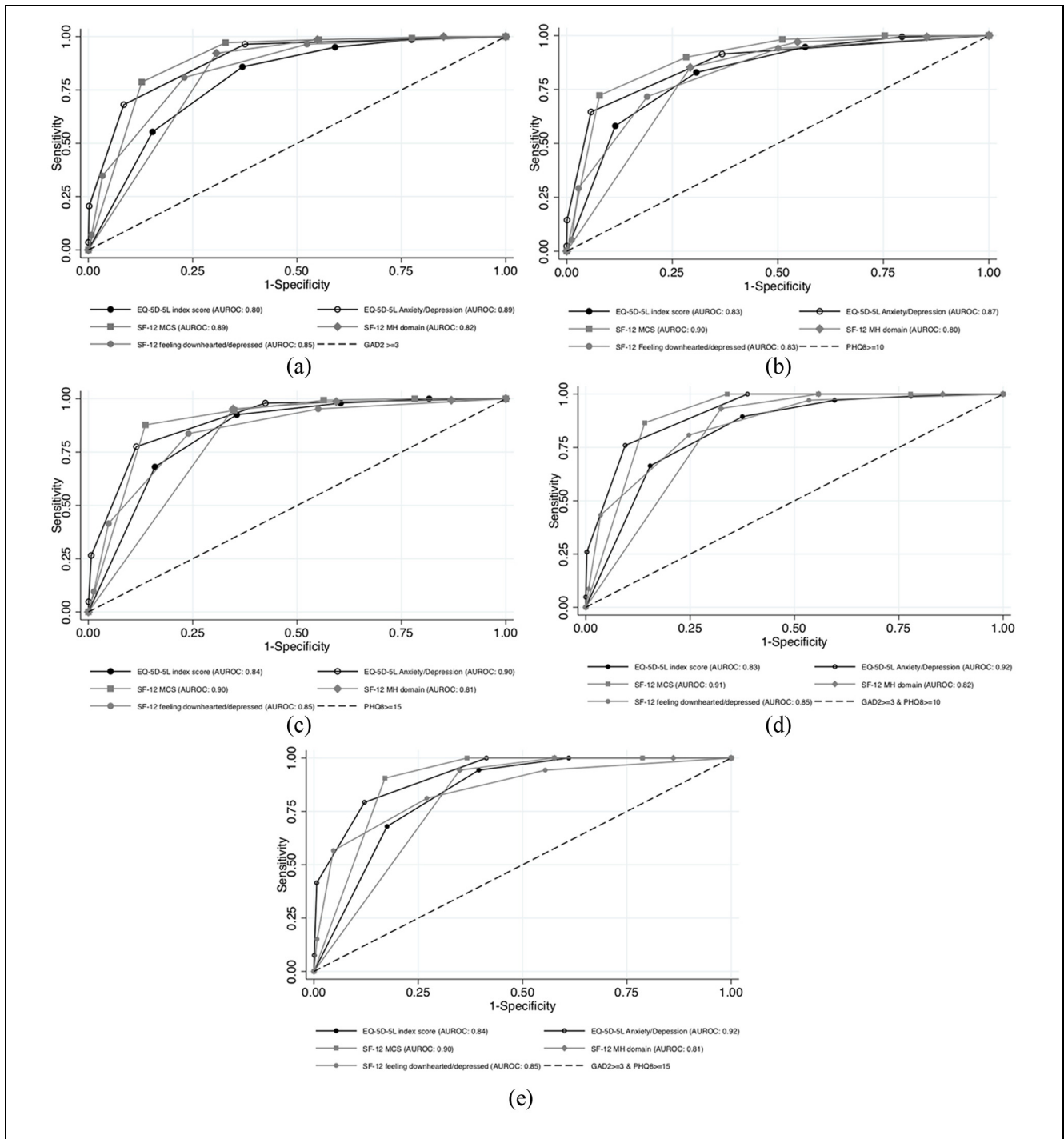
Figure 1 Overall AUROC for the performance of EQ-5D-5L and SF-12 components in screening for anxiety and depressive symptoms.

within the scope of this article. Although depression is the leading cause of disability worldwide and a major contributor to the overall global burden of disease,[35] mental health screening is not regularly done and guidelines differ between agencies despite the availability of suitable screening tests. In Canada, while the Canadian

Table 3 Performance of EQ-5D-5L and SF-12 in Screening for Any (PHQ8 ≥10) and Moderate-Severe (PHQ8 ≥15) Depressive Symptoms[a]

Cut-Point	Any Depressive Symptoms (PHQ8 ≥10)					Moderate-Severe Depressive Symptoms (PHQ8 ≥15)				
	Sensitivity	Specificity	LR+	LR−	AUROC (95% CI)	Sensitivity	Specificity	LR+	LR−	AUROC (95% CI)
EQ-5D-5L index score quintiles										
Overall					0.83 (0.80, 0.86)					0.83 (0.79, 0.86)
≥2	98.7	24.3	1.30	0.05		98.8	21.4	1.26	0.06	
≥3	95.7	43.9	1.71	0.10		98.8	39.2	1.62	0.03	
≥4	**84.6**	**67.2**	**2.57**	**0.23**		**92.5**	**61.3**	**2.39**	**0.12**	
≥5	59.7	88.7	5.26	0.46		65.0	83.0	3.83	0.42	
EQ-5D-5L Anxiety/depression dimension										
Overall					0.88 (0.86, 0.90)					0.90 (0.87, 0.92)
≥2	**95.3**	**67.0**	**2.89**	**0.07**		100.0	59.6	2.48	0.00	
≥3	56.8	93.8	9.22	0.46		**71.6**	**88.6**	**6.25**	**0.32**	
≥4	13.3	99.9	144.01	0.87		28.4	99.3	39.12	0.72	
≥5	2.6	100.0	—	0.97		6.2	99.9	76.55	0.94	
SF-12 MCS quintiles										
Overall					0.92 (0.91, 0.93)					0.90 (0.89, 0.92)
≥2	100.0	24.7	1.33	0.00		100.0	21.6	1.27	0.00	
≥3	100.0	49.3	1.97	0.00		100.0	43.1	1.76	0.00	
≥4	**98.2**	**73.2**	**3.67**	**0.02**		100.0	64.4	2.80	0.00	
≥5	74.9	91.6	8.87	0.27		**90.4**	**84.1**	**5.70**	**0.11**	
SF-12 MH quartiles										
Overall					0.82 (0.80, 0.85)					0.81 (0.79, 0.83)
≥2	100.0	16.3	1.20	0.00		100.0	14.3	1.17	0.00	
≥3	97.4	48.8	1.90	0.05		100.0	43.2	1.76	0.00	
≥4	**86.7**	**73.2**	**3.24**	**0.18**		**93.8**	**66.3**	**2.78**	**0.09**	
SF-12 Feeling downhearted/depressed item										
Overall					0.84 (0.81, 0.86)					0.85 (0.81, 0.89)
≥2	94.9	51.1	1.94	0.10		96.3	45.4	1.77	0.08	
≥3	**74.0**	**80.5**	**3.79**	**0.32**		**82.7**	**74.3**	**3.21**	**0.23**	
≥4	27.7	97.6	11.54	0.74		44.4	95.6	10.01	0.58	
≥5	4.7	99.3	6.35	0.96		12.4	99.3	17.00	0.88	

AUROC, area under receiver operating curve; EQ-5D-5L, EuroQol Five-Dimension, Five-Level Questionnaire; LR−, negative likelihood ratio; LR+, positive likelihood ratio; MCS, mental composite summary; MH, mental health; PHQ8, 8-item Patient Health Questionnaire; SF-12, 12-item Medical Outcomes Health Survey–Short Form.

[a] EQ-5D-5L index score: 2 = quintile 2, 3 = quintile 3, 4 = quintile 4, 5 = quintile 5. EQ-5D-5L anxiety/depression dimension: 2 = mild problems, 3 = moderate problems, 4 = severe problems, 5 = extreme problems/unable to perform. SF-12 MCS quintiles: 2 = quintile 2, 3 = quintile 3, 4 = quintile 4, 5 = quintile 5. SF-12 MH quartiles: 2 = quartile 2, 3 = quartile 3, 4 = quartile 4. SF-12 feeling downhearted/depressed item: 2 = little of the time, 3 = some of the time, 4 = most of the time, 5 = all of the time. Bolded values indicate the cutoff that maximizes sensitivity and specificity for each component.

Diabetes Association clinical practice guidelines recommend screening for depressive symptoms,[12] the Canadian Task Force for Preventive Health Care advises against routine screening for depression in primary care settings, even among those who may be at increased risk of depression.[36] In the United Kingdom, the National Institutes for Health and Clinical Excellence recommends targeted case identification rather than general population screening,[37] while the US Preventive Services Task Force recommends universal screening where adequate systems are in place to ensure accurate diagnosis,

effective treatment, and appropriate follow-up.[38] These guidelines and others have considered the presence of supports for treatment and follow-up as a key determinant for screening, particularly while considering the harms of mental health screening such as labelling and stigma, false diagnosis, and lack of treatment and follow-up.

In using any of these instruments for the purpose of screening for anxiety and/or depressive symptoms, several factors need to be considered. For the EQ-5D-5L, the index score, which is a preference-based summary

Table 4 Performance of EQ-5D-5L and SF-12 Components in Screening for Comorbid Anxiety (GAD2 ≥3) With Any Depressive Symptoms (PHQ8 ≥10), and Comorbid Anxiety (GAD2 ≥3) With Moderate-Severe Depressive Symptoms (PHQ8 ≥15)[a]

Cut-Point	Anxiety (GAD2 ≥3) and Any Depressive Symptoms (PHQ8 ≥10)					Anxiety (GAD2 ≥3) and Moderate-Severe Depressive Symptoms (PHQ8 ≥15)				
	Sensitivity	Specificity	LR +	LR −	AUROC (95% CI)	Sensitivity	Specificity	LR +	LR −	AUROC (95% CI)
EQ-5D-5L index score quintiles										
Overall					0.83 (0.79, 0.86)					0.83 (0.79, 0.86)
≥2	99.1%	22.1%	1.30	0.04		100.0%	21.2%	1.27	0.00	
≥3	97.3%	40.1%	1.60	0.06		100.0%	38.7%	1.63	0.00	
≥4	**89.2%**	**62.3%**	**2.40**	**0.20**		**93.2%**	**60.4%**	**2.35**	**0.11**	
≥5	65.8%	84.4%	4.20	0.40		66.1%	82.3%	3.73	0.41	
EQ-5D-5L Anxiety/depression dimension										
Overall					0.92 (0.90, 0.94)					0.92 (0.89, 0.94)
≥2	100.0%	61.2%	2.58	0.00		100.0%	58.7%	2.42	0.00	
≥3	**75.9%**	**90.5%**	**8.00**	**0.26**		**78.3%**	**87.9%**	**6.48**	**0.25**	
≥4	25.9%	99.8%	104.00	0.74		38.3%	99.3%	53.50	0.62	
≥5	5.4%	100.0%	—	0.95		8.3%	99.9%	104.75	0.92	
SF-12 MCS quintiles										
Overall					0.91 (0.89, 0.92)					0.90 (0.88, 0.92)
≥2	100.0%	22.2%	1.28	0.00		100.0%	21.3%	1.27	0.00	
≥3	100.0%	44.3%	1.80	0.00		100.0%	42.4%	1.74	0.00	
≥4	100.0%	66.1%	2.90	0.00		100.0%	63.3%	2.72	0.00	
≥5	**86.5%**	**85.8%**	**6.10**	**0.15**		**90.6%**	**83.0%**	**5.32**	**0.11**	
SF-12 MH quartiles										
Overall					0.82 (0.80, 0.84)					0.81 (0.79, 0.83)
≥2	100.0%	14.7%	1.17	0.00		100.0%	14.1%	1.16	0.00	
≥3	100.0%	44.4%	1.80	0.00		100.0%	42.5%	1.74	0.00	
≥4	**93.7%**	**67.8%**	**2.91**	**0.09**		**94.9%**	**65.3%**	**2.70**	**0.08**	
SF-12 Feeling downhearted/depressed item										
Overall					0.85 (0.83, 0.90)					0.85 (0.81, 0.91)
≥2	97.3%	46.7%	1.83	0.06		95.0%	44.8%	1.72	0.11	
≥3	**82.1%**	**75.7%**	**3.38**	**0.24**		**83.3%**	**73.3%**	**3.12**	**0.23**	
≥4	42.9%	96.5%	12.29	0.59		53.3%	95.4%	11.55	0.49	
≥5	9.8%	99.3%	14.80	0.91		16.7%	99.3%	23.30	0.84	

AUROC, area under receiver operating curve; CI, confidence interval; EQ-5D-5L, EuroQol Five-Dimension, Five-Level Questionnaire; GAD2, 2-item Generalized Anxiety Disorder questionnaire; LR−, negative likelihood ratio; LR +, positive likelihood ratio; MCS, mental composite summary; MH, mental health; PHQ8, 8-item Patient Health Questionnaire; SF-12, 12-item Medical Outcomes Health Survey–Short Form.
[a]EQ-5D-5L index score: 2 = quintile 2, 3 = quintile 3, 4 = quintile 4, 5 = quintile 5. EQ-5D-5L anxiety/depression dimension: 2 = mild problems, 3 = moderate problems, 4 = severe problems, 5 = extreme problems/unable to perform. SF-12 MCS quintiles: 2 = quintile 2, 3 = quintile 3, 4 = quintile 4, 5 = quintile 5. SF-12 MH quartiles: 2 = quartile 2, 3 = quartile 3, 4 = quartile 4. SF-12 feeling downhearted/depressed item: 2 = little of the time, 3 = some of the time, 4 = most of the time, 5 = all of the time. Bolded values indicate the cutoff that maximizes sensitivity and specificity for each component.

score of the five dimensions measured by the instrument, and the anxiety/depression dimension, both appeared suitable to screen for anxiety/depressive symptoms in this population of type 2 diabetes. However, the reference period of this instrument is "today," and each dimension asks about the severity of problems, and not their frequency or impact on life; these characteristics may influence the application of the tool. The key advantage of the EQ-5D is that it is very short and easy to complete even in clinical settings, and screening for anxiety and depressive symptoms could be done using only one item, "anxiety/depression dimension." As for the SF-12 that appeared to have a similar performance to that of the EQ-5D-5L, the reference period is 4 weeks, and the items assess the frequency of symptoms and the impact of such symptoms on one's life. To use the SF-12 as a screener for anxiety or depressive symptoms, one needs to ask all of the 12 items, and apply the scoring algorithm to obtain

the MH and MCS scores. The choice between EQ-5D-5L and SF-12 to screen for anxiety and/or depressive symptoms depends not only on their suitability as a screening test but also on the clinical context and the purpose of screening, as well as the implications of such screening.

The advantage of using these instruments is in the context of routine outcome measurement, as generic PROMs, which have multiple purposes. Preference-based measures like the EQ-5D may be used in economic evaluations. However, PROMs have been increasingly used in noneconomic applications including quality improvement within health care systems, surveillance of population health, and patient management in clinical practice. Routine mental health screening using these generic PROMs could be implemented within the context of these large-scale applications of these measurements. For example, in Canada, the provincial health ministry in Alberta has mandated that all chronic disease patients within primary care settings complete a health status measure, such as the EQ-5D-5L, on an annual basis for routine monitoring of outcomes and evaluation of health care services. Our results support the use of this measure to screen for anxiety and depressive symptoms in adults with type 2 diabetes in primary care; a population that reportedly suffers from high levels of mental health disorders that is often neither diagnosed nor treated.[39]

A few limitations should be considered in the interpretation of the results of this study. First, anxiety and depressive symptoms assessment was based on self-report. Although GAD2 and PHQ8 are widely used and established screening measures of anxiety and depressive symptoms, respectively, comparing the performance of the EQ-5D-5L and SF-12 to clinical assessments of these symptoms would provide a more robust conclusion about their suitability for this purpose. Second, some of the examined groups, namely, those with moderate-severe depressive symptoms ($N = 82$) and those with comorbid anxiety and moderate-severe depressive symptoms ($N = 61$) had small samples, which limits the analysis. Finally, this study included predominantly Caucasian adults with type 2 diabetes in Alberta, Canada; the generalizability of the results of this study is limited to similar populations, and therefore exploration in other populations is warranted.

In summary, we found the EQ-5D-5L and the SF-12 to be suitable tools for screening for anxiety and depressive symptoms in adults with type 2 diabetes. These tools present a unique opportunity for a standardized approach for routine mental health screening in primary care settings within the context of routine outcome measurement initiatives, particularly for patients with chronic diseases. Examining the suitability and usefulness of these tools and other PROMs in mental health screening in other chronic disease populations and clinical settings is crucial.

References

1. World Health Organization. *Global Report on Diabetes*. Geneva: World Health Organization; 2016.
2. Diabetes Canada. Diabetes statistics in Canada [cited June 23, 2017]. Available from: http://www.diabetes.ca/how-you-can-help/advocate/why-federal-leadership-is-essential/diabetes-statistics-in-canada#_ftn1
3. Ducat L, Philipson LH, Anderson BJ. The mental health comorbidities of diabetes. *JAMA*. 2014;312(7):691–2.
4. Dziemidok P, Makara-Studzińska M, Jarosz MJ. Diabetes and depression: a combination of civilization and life-style diseases is more than simple problem adding—literature review. *Ann Agric Environ Med*. 2011;18(2):318–22.
5. Sears C, Schmitz N. The relationship between diabetes and mental health conditions in an aging population. *Can J Diabetes*. 2016;40(1):4–5.
6. Grigsby AB, Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. Prevalence of anxiety in adults with diabetes: a systematic review. *J Psychosom Res*. 2002;53(6):1053–60.
7. Roy T, Lloyd CE. Epidemiology of depression and diabetes: a systematic review. *J Affect Disord*. 2012;142 (Suppl.):S8–S21.
8. Black S, Markides K, Ray L. Depression predicts increased incidence of adverse health outcomes in older Mexican Americans with type 2 diabetes. *Diabetes Care*. 2003;26(10):2822–8.
9. Ciechanowski P, Katon WJ, Russo J. Depression and diabetes: impact of depressive symptoms on adherence, function, and costs. *Arch Intern Med*. 2000;160(21):3278–85.
10. Lustman PJ, Anderson RJ, Freedland KE, deGroot M, Carney RM, Clouse RE. Depression and glycemic control: a meta-analysis review of the literature. *Diabetes Care*. 2000;23(7):934–42.
11. Johnson JA, Al Sayah F, Wozniak L, et al. Collaborative care versus screening and follow-up for patients with diabetes and depressive symptoms: results of a primary care-based comparative effectiveness trial. *Diabetes Care*. 2014;37(12):3220–6.
12. Diabetes Canada. Diabetes and mental health [cited June 23, 2017]. Available from: http://guidelines.diabetes.ca/Browse/Chapter18
13. Brazier J, Connell J, Papaioannou D, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. *Health Technol Assess*. 2014;18(34):1–188.
14. König HH, Born A, Günther O, et al. Validity and responsiveness of the EQ-5D in assessing and valuing health

status in patients with anxiety disorders. *Health Qual Life Outcomes*. 2010;8:47.

15. Sapin C, Fantino B, Nowicki ML, Kind P. Usefulness of EQ-5D in assessing health status in primary care patients with major depressive disorder. *Health Qual Life Outcomes*. 2004;2:20.

16. Appleby J, Devlin N, Parkin D. *Using Patient Reported Outcomes to Improve Healthcare*. 1st ed. West Sussex: John Wiley; 2016.

17. Devlin NJ, Appleby J. *Getting the Most Out of PROMs: Putting Health Outcomes at the Heart of NHS Decision-Making*. London: The King's Fund; 2010.

18. Al Sayah F, Majumdar SR, Soprovich A, et al. The Alberta's Caring for Diabetes (ABCD) study: rationale, design and baseline characteristics of a prospective cohort of adults with type 2 diabetes. *Can J Diabetes*. 2015;39(Suppl. 3):S113–S119.

19. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–36.

20. Xie F, Pullenayegum E, Gaebel K, et al. A time trade-off-derived value set of the EQ-5D-5L for Canada. *Med Care*. 2016;54(1):98–105.

21. Fleishman JA, Selim AJ, Kazis LE. Deriving SF-12v2 physical and mental health summary scores: a comparison of different scoring algorithms. *Qual Life Res*. 2010;19(2):231–41.

22. Skapinakis P. The 2-item Generalized Anxiety Disorder scale had high sensitivity and specificity for detecting GAD in primary care. *Evid Based Med*. 2007;12(5):149.

23. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41(5):582–92.

24. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord*. 2009;114(1–3):163–73.

25. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *CMAJ*. 2012;184(3):E191–E196.

26. Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): a meta-analysis. *Gen Hosp Psychiatry*. 2015;37(6):567–76.

27. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988;240(4857):1285–93.

28. StataCorp. Stata Statistical Software: Release 13.1 [Computer software]. College Station (TX): StataCorp; 2013.

29. Vilagut G, Forero CG, Pinto-Meza A, et al. The mental component of the Short-Form 12 Health Survey (SF-12) as a measure of depressive disorders in the general population: results with three alternative scoring methods. *Value Health*. 2013;16(4):564–73.

30. Gill SC, Butterworth P, Rodgers B, Mackinnon A. Validity of the mental health component scale of the 12-item Short-Form Health Survey (MCS-12) as measure of common mental disorders in the general population. *Psychiatry Res*. 2007;152(1):63–71.

31. Tavella R, Air T, Tucker G, Adams R, Beltrame JF, Schrader G. Using the Short Form-36 mental summary score as an indicator of depressive symptoms in patients with coronary heart disease. *Qual Life Res*. 2010;19(8):1105–13.

32. Walsh TL, Homa K, Hanscom B, Lurie J, Sepulveda MG, Abdu W. Screening for depressive symptoms in patients with chronic spinal pain using the SF-36 Health Survey. *Spine J*. 2006;6(3):316–20.

33. Supina AL, Johnson JA, Patten SB, Williams JV, Maxwell CJ. The usefulness of the EQ-5D in differentiating among persons with major depressive episode and anxiety. *Qual Life Res*. 2007;16(5):749–54.

34. Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. Geneva: World Health Organization; 1968.

35. World Health Organization. Depression [cited June 26, 2017]. Available from: http://www.who.int/mediacentre/factsheets/fs369/en/

36. Joffres M, Jaramillo A, Dickinson J, et al. Recommendations on screening for depression in adults. *CMAJ*. 2013;185(9):775–82.

37. UK National Collaborating Centre for Mental Health. *Depression: The Treatment and Management of Depression in Adults (Updated Edition)*. Leicester: British Psychological Society; 2010.

38. Siu AL; US Preventive Services Task Force (USPSTF); Bibbins-Domingo K, et al. Screening for depression in adults: US Preventive Services Task Force recommendation statement. *JAMA*. 2016;315(4):380–7.

39. Clarke DM, Currie KC. Depression, anxiety and their relationship with chronic diseases: a review of the epidemiology, risk and treatment evidence. *Med J Aust*. 2009;190(7 Suppl.):S54–S60.