The Patient Health Questionnaire-9 as a Screening Tool for Depression in Individuals with Type 2 Diabetes Mellitus: The Maastricht Study

Citation for published version (APA):

Document status and date:
Published: 01/11/2016

DOI:
10.1111/jgs.14388

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.
Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:
repository@maastrichtuniversity.nl
providing details and we will investigate your claim.

Download date: 17 Sep. 2023
The Patient Health Questionnaire-9 as a Screening Tool for Depression in Individuals with Type 2 Diabetes Mellitus: The Maastricht Study

Eveline P. C. J. Janssen, MD,* † Sebastian Köhler, PhD, ‡ Coen D. A. Stehouwer, MD, PhD,§ Nicolaas C. Schaper, MD, PhD,¶ Pieter C. Dagnelie, MD, PhD,*** Simone J. S. Sep, PhD,*** Ronald M. A. Henry, MD, PhD, Carla J. H. van der Kallen, PhD,§ Frans R. Verhey, MD, PhD,* ‡ and Miranda T. Schram, PhD§

OBJECTIVES: To assess the psychometric properties and identify the best cutoff value of the Patient Health Questionnaire-9 (PHQ-9) for depression screening in individuals with type 2 diabetes mellitus (T2DM).

DESIGN: Observational population-based cohort study.

SETTING: The Maastricht Study.

PARTICIPANTS: Individuals with and without T2DM (mean age 58.6 ± 8.1, 44.6% male) according to an oral glucose tolerance test (N = 2,997).

MEASUREMENTS: Depressive disorder and depressive symptoms were measured using the Mini-International Neuropsychiatric Interview (MINI) as the reference and the PHQ-9. Cronbach alpha, Cohen’s kappa and receiver operating characteristic (ROC) analyses were used. Differences in factorial structure between participants with and without T2DM were tested using multigroup confirmatory factor analysis.

RESULTS: Based on the traditional PHQ-9 cutoff value, 133 (4.4%) participants had depressive symptoms (PHQ-9 score ≥ 10). Internal consistency of the PHQ-9 was good (Cronbach α = 0.87 with T2DM, 0.82 without T2DM), the kappa of agreement between the PHQ-9 and the MINI was moderate (0.40 with T2DM, 0.43 without T2DM). Area under the ROC curve for the PHQ-9 was 0.87 in participants with T2DM and 0.88 in those without. A PHQ-9 cutoff score of 5 provided the best sensitivity (92.3%), with acceptable specificity (70.4%), for T2DM, similar to sensitivity and specificity in individuals without T2DM. Factor analysis suggested a similar two-factor structure in both groups (affective and somatic symptoms).

CONCLUSION: Patient Health Questionnaire-9 performs well as a screening tool for depressive symptoms in individuals with and without T2DM based on the cutoff value of 5, indicating that the PHQ-9 can be used in two-stage screening in primary care to select individuals with T2DM for further psychological evaluation. J Am Geriatr Soc 64: e201–e206, 2016.

Key words: diabetes mellitus type 2; depressive disorder; Patient Health Questionnaire-9; Mini-International Neuropsychiatric Interview; screening

Type 2 diabetes mellitus (T2DM) is a chronic metabolic disorder, and psychological comorbidity such as depression is prevalent.1 Meta-analyses have demonstrated that a depressive disorder is nearly twice as prevalent in individuals with T2DM than in the general population, with a prevalence of 8% to 43.5%.2–5 Depression in individuals with T2DM is associated with poorer quality of life,6 poorer self-care, greater healthcare use, poorer glycemic control, and subsequently higher healthcare costs.6–10 The poorer glycemic control in individuals with T2DM with depressive symptoms may, at least in part, explain the excess risk of DM-related complications and mortality in these individuals.8–12 Therefore, early detection and treatment of depressive symptoms in individuals with T2DM may decrease complications and mortality and help to prevent a more-severe depressive disorder.13 The negative effects of depression in individuals with T2DM are not limited to those with severe depression but have also been observed in individuals with relatively mild or
subclinical depression.\textsuperscript{9,14} It is recommended that individuals with T2DM be systematically screened for depression,\textsuperscript{15} but this has not been implemented in primary care and referral diabetes mellitus healthcare settings in most countries, indicating that validated, standardized methods for depression screening are needed in guidelines for T2DM treatment.

Several questionnaires have been developed to assess depressive symptoms, of which the 9-item Patient Health Questionnaire (PHQ-9) is frequently used.\textsuperscript{16,17} The predictive value of the PHQ-9 has been studied in small samples of individuals with T2DM for major and in some studies also for minor depressive disorder.\textsuperscript{18–21} The questionnaire performed well as a screening instrument in comparison with the reference standards (clinical interview or the Mini-International Neuropsychiatric Interview (MINI))\textsuperscript{22}, but each study used a different cutoff value for the PHQ-9.\textsuperscript{18–20} Factor analytical studies in the general population or in individuals with depression suggest that the PHQ-9 consists of two correlated factors measuring affective and somatic symptoms of depression.\textsuperscript{23,24} Yet it is unknown whether the factor structure is similar for individuals with and without T2DM, in particular because somatic symptoms of depression might occur more often in individuals with T2DM.

Before the PHQ-9 can be implemented in daily clinical diabetes care, further psychometric evaluation of the questionnaire is required in a large group of individuals with T2DM. Furthermore, the best cutoff value of the PHQ-9 has to be identified, taking into account a prevalence of depression of less than 10% in the target population.\textsuperscript{25}

The aims of the present study were to identify the best cutoff value for the PHQ-9 questionnaire in individuals with and without T2DM using the MINI diagnostic interview as the reference standard and to test the psychometric properties of the PHQ-9 in individuals with and without T2DM.

METHODS

Study Population

Data from the Maastricht Study, an observational population-based cohort study, were used. The rationale and methodology have been described previously.\textsuperscript{26} In brief, the study focuses on the etiology, pathophysiology, complications, and comorbidities of T2DM and is characterized by an extensive phenotyping approach. All individuals aged 40 and 75 living in the southern part of the Netherlands were eligible for participation. Participants were recruited through mass media campaigns and from the municipal registries and the regional Diabetes Patient Registry using mailings. Recruitment was stratified according to known T2DM status for reasons of efficiency. The present report includes cross-sectional data from the first 3,451 participants, who completed the baseline survey between November 2010 and September 2013. Examinations of each participant were performed within a time window of 3 months. The institutional medical ethical committee (NL31329.068.10) and the Minister of Health, Welfare and Sports of the Netherlands approved the study on the basis of the Health Council’s opinion (Permit 131088–105234-PG). All participants provided written informed consent.

Assessment of T2DM Status

To determine glucose metabolism defined according to the World Health Organization 2006 criteria,\textsuperscript{27} all participants except those using insulin underwent an oral glucose tolerance test after an overnight fast.\textsuperscript{26} Participants using glucose-lowering medication were classified as having T2DM. Individuals with type 1 diabetes mellitus were excluded from the analyses.

Assessment of Depression

Major and minor depressive disorders according to Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria were assessed using the MINI, a short diagnostic structured interview.\textsuperscript{22,26} Major depressive disorder was diagnosed if participants had at least one core symptom (depressed mood or loss of interest) and at least four other symptoms of depression (significant weight change or change in appetite, insomnia or hypersomnia, psychomotor agitation or retardation, fatigue or loss of energy, guilt or worthlessness, diminished ability to think or concentrate or indecisiveness, and suicidal thoughts or plans). Persons with one core symptom and one to three other symptoms were classified as having a minor depressive disorder. The term “depression” in this study refers to participants with major or minor depressive disorder.

Depressive symptoms were assessed using a validated Dutch version of the PHQ-9, a self-administered questionnaire based on the DSM-IV criteria for a major depressive disorder.\textsuperscript{17} The PHQ-9 measured cognitive (thoughts about oneself and problems of the mind) and somatic symptoms of depression (various bodily sensations that a depressed individual perceives as unpleasant or worrisome). The questionnaire comprises nine items each rated on a 4-point scale ranging from 0 (not at all) to 3 (nearly every day). Response options are used to calculate a continuous total score ranging from 0 to 27; as a continuous variable, scores of 9 or less indicate no depression, 10 to 14 moderate depression, 15 to 19 moderately severe depression, and 20 to 27 severe depression. A predefined cutoff score of 10 can be used as a dichotomous scoring system for defining clinically relevant depressive symptoms.\textsuperscript{17}

Statistical Analysis

Statistical analyses were performed using SPSS version 20.0 (IBM Corp, Armonk, NY) and Mplus 7 (Muthén & Muthén, Los Angeles, CA). Dichotomous data were analyzed using the chi-square test and continuous data using an independent-samples t-test. The interitem consistency of the PHQ-9 was measured according to Cronbach alpha. Receiver operating characteristics (ROC) analyses were used to assess the sensitivity, specificity, and positive and negative predictive value and to test for overall accuracy and unweighted k for cutoff points on the PHQ-9. In this study on screening...
characteristics, higher sensitivity of the PHQ-9 is preferred over good specificity.

To assess whether PHQ-9 specific items (e.g., somatic or affective symptoms) are related to different latent dimensions of depression, confirmatory factor analysis was conducted, following a strategy described previously.23 The PHQ-9 consists of Likert-type items that were treated as ordinal variables by specifying a mean and variance-corrected weighted least squares estimator. The model with the best fit was chosen based on the root mean square error of approximation (RMSEA, <.05 indicates close fit), Comparative Fit Index, and Tucker-Lewis Index (values >0.95 indicate close fit for both). Chi-square is also reported (P >.05 desirable), although it is sensitive to sample size and is often significant in even only moderately large samples. Finally, measurement invariance was tested to evaluate whether the factor model is similar for participants with and without T2DM (multigroup confirmatory factor analysis). In the case of strong factorial invariance, groups have the same factor pattern (same number of factors and same items correlating with the same factors), intercepts (or item thresholds in the case of ordinal items), and factor loadings.28

RESULTS

Demographical and Clinical Characteristics

Of the 3,451 participants, 454 were excluded from the analyses because they had Type 1 or other type of diabetes (n = 41) or missing MINI (n = 141) or PHQ-9 (n = 294) values (Figure S1). Of the remaining 2,997 participants, 793 (26.5%) had T2DM. Depressive disorder, according to the MINI, was present in 61 (7.7%) participants with T2DM and in 88 (4.0%) of those without.

Demographical and clinical characteristics according to T2DM and depression status are presented in Table 1. Significant differences were found in all demographic characteristics, lifestyle factors, clinical characteristics, and use of medication, except for sleep medication, between participants with and without T2DM.

In the T2DM group, individuals with current depression had higher glycosylated hemoglobin levels and were significantly more likely to report lifetime depression; to be current smokers; and to use antidepressants, anxiolytics, and antipsychotics than those who did not have depression.

Individuals with T2DM with depression were significantly more often male; were older; had a higher body mass index; and more often reported prior cardiovascular disorders, hypertension, and use of glucose-lowering medication, statins, and antihypertensives than those without T2DM.

Agreement Between PHQ-9 and MINI Diagnostic Interview

Based on the traditional PHQ-9 cutoff value of 10, 133 (4.4%) participants, 57 with and 76 without T2DM, were found to have depressive symptoms. The internal consistency of the PHQ-9 was good, and Cronbach alpha was 0.87 for individuals with T2DM and 0.82 for those without. The kappa measure of agreement between PHQ-9 with a cutoff score of 10 and the MINI was moderate (0.40 with T2DM, 0.43 without).

In the group with T2DM, overlap between a MINI diagnosis of depression and presence of a PHQ-9 score of 10 or greater was small (Table 2). Sensitivity was 42.6%, and specificity was 95.7%, with a positive predictive value of 45.6% and a negative predictive value of 95.8% (Table 3). Similar poor results were found in the group without T2DM and the total population (Table 3).

ROC Analyses

Receiver operating characteristic analyses were performed to define a potential better cut-off value for the PHQ-9. These analyses showed an area under the ROC curve (AUC) of 0.87 in individuals with T2DM and 0.88 in those without (both good). When the groups were combined, an AUC of 0.88 was found (Figure S2). Table 3 illustrates that a PHQ-9 cutoff score of 5 provided the best sensitivity (92.3%), with moderate but still acceptable specificity (70.4%) for individuals with T2DM. Sensitivity for individuals without T2DM was 92.9%, and specificity was 73.9%.

Factor Structure

Confirmatory factor analysis was performed following a previously reported strategy,23 four models were tested (a single-factor model and three-two-factor models that compared different item-factor loading: Table S1), and the best fit was found for a two-factor model consisting of four nonsomatic (anhedonia, depressed mood, feeling of worthlessness, thoughts of death) and five somatic (fatigue, appetite changes, sleep difficulties, concentration difficulties, psychomotor agitation and retardation) items. This model showed a close fit (χ² = 196.9, degrees of freedom = 26, P < .001; RMSEA = 0.047, 90% confidence interval = 0.04–0.05, Comparative Fit Index = 0.991, Tucker-Lewis Index = 0.987) (Table S2). Testing the model in individuals with and without T2DM separately showed good fit in both groups (data not shown). Formal testing for measurement invariance across both groups showed support for strong (scalar) factorial invariance.

Diagnostic Algorithm of the PHQ-9

A diagnostic algorithm can be used as an alternative sum-score for the PHQ-9; the core symptoms of depressive disorder (depressed mood, anhedonia) have to be present along with five or more other items for more than half of the days.29 This algorithm had worse consistency with MINI depressive disorder than the continuous total score (data not shown).

Sensitivity Analyses

Because the use of psychoactive drugs (Table 1) could affect the scoring of the PHQ-9, a sensitivity analyses was performed excluding participants who used anxiolytics, antidepressants, or antipsychotics. This did not materially change the results (data not shown).
DISCUSSION

The PHQ-9 performed similarly in individuals with and without T2DM in the current study. In individuals with and without T2DM, a PHQ-9 cutoff of 5 was found to distinguish far better between individuals with and without depression as diagnosed using the MINI structured interview than the traditional cutoff 10. Further psychometric evaluation demonstrated a similar factor structure of the PHQ-9 for individuals with and without T2DM, suggesting two correlated factors (affective and somatic symptoms).

The application of a cutoff score of 5 for the PHQ-9 demonstrated sensitivity of 92.3% and specificity of 70.4% in individuals with T2DM and comparable results in individuals without T2DM. The sensitivity of the traditional PHQ-9 cutoff score of 10 was below acceptable values, in contrast to the majority of earlier studies. To the knowledge of the authors of the current study, this is the first study to investigate a different ratio of sensitivity and specificity.
preventing depression-mediated effects on diabetic complications.

To the knowledge of the authors, this is the first study to conduct confirmatory factor analyses in individuals with and without T2DM. The factor structure of somatic and nonsomatic items of the PHQ-9 showed no difference in individuals with and without T2DM, despite the presence of a chronic somatic disease in those with T2DM, which could have led to higher scores for the somatic items. Thus, the full questionnaire is useful regardless of T2DM status.

This study is based on the Maastricht Study, a large, observational, population-based cohort study, which has several strengths. Most notably, its sample size allows the psychometric properties of the PHQ-9 to be studied properly given the low prevalence of major depression in the population. In addition, state-of-the-art and validated measures were used to assess DM and depression status.

Certain limitations of the study should be acknowledged. The study population was mainly Caucasian, which limits extrapolation to non-Caucasian populations. Furthermore, the recruitment strategy could have led to selection bias in the sense that individuals with more-severe depressive disorder or DM or with greater comorbidity were less likely to participate in the Maastricht Study. The prevalence of depression might therefore be underestimated and may be higher in an unselected population with T2DM, although it is comparable with previously published data. The lower PHQ-9 cutoff used in the final analyses would detect these more severely depressed individuals.

**CONCLUSION**

The PHQ-9 performs well as a depression screening tool in individuals with and without T2DM individuals based on the cutoff value of 5 and can consequently be used in a two-stage screening approach in primary DM care. Possible early detection of depression in individuals with T2DM may improve outcomes in terms of mental and physical health.

**ACKNOWLEDGMENTS**

This study is supported by the European Regional Development Fund as part of OP-ZUID; the Province of Limburg; the department of Economic Affairs of the Netherlands (grant 310.041); Stichting the Weijerhorst, the Pearl String Initiative Diabetes; the Cardiovascular Center Maastricht Cardiovascular Research Institute Maastricht; School of Nutrition, Toxicology and Metabolism; Stichting Anndal; Health Foundation Limburg; and unrestricted grants from Janssen, Novo Nordisk, and Sanofi.

**Conflict of Interest:** None.

**Author Contributions:** Janssen: study concept and design, analysis, interpretation, writing of manuscript. Köhler, Schram: data acquisition, conceptualization, study concept and design, analysis, interpretation, writing of manuscript. Stehouwer, Schaper, Dagnelie, Sep, Henry, van der Kallen, Verhey: data acquisition, conceptualization, data interpretation. All authors: creation and revision of manuscript.
Sponsor’s Role: There was no sponsor involvement in data collection, analysis, interpretation, or manuscript preparation.

REFERENCES


SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Factor Models.
Table S2. Comparison-of-Fit Statistics for the Four Patient Health Questionnaire-9 Factor Analytical Models in the Total Group.
Figure S1. Flowchart.
Figure S2. Receiver operating characteristic curve in total group.

Please note: Wiley-Blackwell is not responsible for the content, accuracy, errors, or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.