Journal of Affective Disorders 190 (2016) 68–74

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research report

Diagnostic accuracy of the Patient Health Questionnaire-9 for assessment of depression in type II diabetes mellitus and/or coronary heart disease in primary care

G. Lennart van der Zwaan [a], Susan E.M. van Dijk [a], Marcel C. Adriaanse [a], Harm W.J. van Marwijk [b], Maurits W. van Tulder [a], Alide D. Pols [a,b], Judith E. Bosmans [a,*]

[a] Department of Health Sciences and the EMGO Institute for Health and Care Research, Faculty of Earth and Life Sciences, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands

[b] Department of General Practice and Elderly Care Medicine the EMGO Institute for Health and Care Research, VU University Medical Centre, Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 21 July 2015
Received in revised form
20 September 2015
Accepted 25 September 2015
Available online 9 October 2015

Keywords:
Depression
Diagnostic accuracy
Patient Health Questionnaire-9
Mini International Neuropsychiatric Interview
Type 2 diabetes
Coronary heart disease

ABSTRACT

Background: Depression is common among type 2 diabetes mellitus (DM2)/coronary heart disease (CHD) patients and is associated with adverse health effects. A promising strategy to reduce burden of disease is to identify patients at risk for depression in order to offer indicated prevention. This study aims to assess the diagnostic accuracy of the Patient Health Questionnaire-9 (PHQ-9) to be used as a tool to identify high risk patients.

Methods: In this cross-sectional study, 586 consecutive DM2/CHD patients aged >18 were recruited through 23 general practices. PHQ-9 outcomes were compared to the Mini International Neuropsychiatric Interview (MINI), which was considered the reference standard. Diagnostic accuracy was evaluated for minor and major depression, comparing both sum- and algorithm based PHQ-9 scores.

Results: For minor depression, the optimal cut-off score was 8 (sensitivity 71%, specificity 71% and an AUC of 0.74). For major depression, the optimal cut-off score was 10 resulting in a sensitivity of 84%, a specificity of 82%, and an AUC of 0.88. The positive predictive value of the PHQ-9 algorithm for diagnosing minor and major depression was 25% and 33%, respectively.

Limitations: Two main limitations apply. MINI Interviewers were not blinded for PHQ-9 scores and less than 10% of all invited patients could be included in the analyses. This could have resulted in biased outcomes.

Conclusions: The PHQ-9 sum score performs well in identifying patients at high risk of minor and major depression. However, the PHQ-9 showed suboptimal results for diagnostic purposes. Therefore, it is recommended to combine the use of the PHQ-9 with further diagnostics to identify depression.

1. Introduction

The presence of minor depression among patients with a chronic disease is high (Barth et al., 2004; Harter et al., 2007). In type 2 diabetes mellitus (DM2) and/or coronary heart disease (CHD) patients for example, the 12-month prevalence of minor depression ranges between 25% and 40% (Bot et al., 2010; Thombs et al., 2006). Approximately 40% of these DM2 and/or CHD patients with depressive symptoms will develop major depressive disorder within two years, indicating that the presence of depressive symptoms or minor depression is the most important predictor for major depression (Bot et al., 2010).

The yearly prevalence of major depression is comparable in DM2 patients and patients with CHD and adds up to about 10–20%, which is considerably high compared to the prevalence of 5% in the general Dutch population (Anderson et al., 2001; Backenstrass et al., 2006; Bot et al., 2010; de Graaf et al., 2012; Harter et al., 2007; Kroenke et al., 2001; Rodriguez et al., 2012; Rudisch and Nemeroff, 2003; Thombs et al., 2006). Major depression among DM2 and/or CHD patients is associated with lower quality of life, an increased risk of mortality, poor medication adherence and increased health care costs (Ciechanowski et al., 2000; Haddad et al., 2013; Lamers et al., 2008; Lin et al., 2009; Rodriguez et al., 2012; Simon, 1992; Stafford et al., 2007). Moreover, once patients are diagnosed with major depression, only roughly one third of the associated disease burden can be averted, even when optimal

* Corresponding author.
E-mail address: j.e.bosmans@vu.nl (J.E. Bosmans).

http://dx.doi.org/10.1016/j.jad.2015.09.045
0165-0327/

treatment is in place (Andrews et al., 2004).

These severe consequences of depression underline the importance of correctly identifying both minor and major depression in patients with DM2 or CHD, thereby enabling general practitioners (GP's) to provide suitable depression care. Previous studies have shown that collaborative and stepped care programs may successfully prevent the development of major depression in patients with minor depression (Cuijpers et al., 2008, Muñoz et al., 2010). However, minor depression often remains unrecognized, which hinders successful indicated prevention (Culpepper, 2012; Gilbody et al., 2008).

Although there is no evidence supporting the routine screening of all primary care patients for depression, general practitioners should be alert to the existence of depression in patients with chronic medical conditions such as DM2 and CHD (Thombs and Ziegelstein, 2014). Thus, GP's have a need for instruments that can be efficiently used to identify depression in their patients and to monitor the course of depression during treatment.

Several instruments have been developed over the past years to identify and monitor depression (Lowe et al., 2004) and many of them are tested in patients with chronic medical illnesses (Meader et al., 2011). One of these instruments is the Patient Health Questionnaire-9 (PHQ-9) which is an extensively applied self-reported questionnaire comprising the nine depression items from the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) (Kroenke and Spitzer, 2002). A recent meta-analysis reported good diagnostic accuracy of the PHQ-9 for the case identification of major depression in patients with chronic medical illness. This meta-analysis identified 6 studies that all used the Diagnosis and Statistical Manual of Mental Disorders (DSM) or the International Classification of Diseases (ICD) as reference standard. The pooled sensitivity and specificity were 84% and 88%, respectively, and the pooled positive and negative likelihood ratios were 6.77 and 0.19, respectively (Meader et al., 2011). Another study, not included in this systematic review, compared the performance of the PHQ-9 to a diagnosis of major depression according to the DSM-IV criteria in clinical outpatients with DM2. Specificity was found to be good (80–87%), but sensitivity was relatively low (58–76%) (van Steenbergen-Weijenburg et al., 2010). Additionally, Haddad et al. (2013) found similar results in 2013 in primary care patients with CHD (sensitivity 59–94%, specificity 84–95%).

However, studies included in the aforementioned meta-analysis were conducted in a very heterogeneous group of patients with different chronic physical health problems. Furthermore, to our knowledge, no information is available for the performance of the PHQ-9 to detect minor depression in DM2 and CHD patients in primary care, which is important because these patients visit the GP often and are at high risk for major depression. Therefore, this study aims to assess the diagnostic accuracy of the PHQ-9 to identify minor and major depression in DM2 and CHD patients without known depression in primary care.

2. Methods

2.1. General design

This study was conducted alongside the StepDep study; a cluster randomized controlled trial to evaluate the cost effectiveness of a stepped-care intervention to prevent major depression in DM2 and/or CHD patients in primary care. Detailed information on this project can be found elsewhere (van Dijk et al., 2013). Data for purposes of the current study were collected using a cross-sectional design in a convenience sample from the StepDep study.

2.2. Ethical approval

The study-protocol was approved by the Medical Ethics Committee of the VU University Medical Centre. This study was conducted corresponding the principles of the Declaration of Helsinki and the Dutch Medical Research Involving Human Subjects Act.

2.3. Study population

The study population was recruited through 23 general practices with 58 GP's and 128,980 enlisted patients in the vicinity of Amsterdam and Twente in the Netherlands. GP's constructed an initial list of all patients with DM2 and or CHD (angina pectoris, acute myocardial infarction, other/chronic ischemic heart disease, coronary sclerosis, previous myocardial infarction) according to the ICPC ($n=7534$) (International Classification of Primary Care, see supplementary material 1) and were requested to exclude patients fulfilling the pre-set exclusion criteria. In total, 1184 patients (16%) were excluded from participation by the GP, because they had type 1 diabetes mellitus, dementia/Alzheimer, major psychiatric conditions (schizophrenia, bipolar depression, affective psychosis, borderline, suicidal attempts), mental retardation, visual impairment, illiteracy, recent loss of a significant other, pregnancy, the use of anti-depressants, or not mastering the Dutch language (supplementary material 1). All remaining individuals ($n=6350$) were invited by mail, through a letter from their GP, to complete the written version of the PHQ-9, which was attached to the invitation letter. Demographical information such as gender and zip code was collected concurrently with the PHQ-9.

A total of 2876 patients (45%) returned the PHQ-9 by mail and of those 1076 (37%) gave informed consent to participate. These patients were approached by telephone for a MINI interview. The interviews took place within two weeks after receiving the PHQ-9 and were administered by trained interviewers. Patients who responded but did not give informed consent ($n=1800$) were excluded from the study. Fig. 1 represents a flowchart of the recruitment of respondents. Another 448 (41%) patients were excluded because they could not be contacted within two weeks. Additionally, 33 were excluded of whom five recently lost a significant other, 14 used anti-depressants at time of the inclusion, 11 did not have DM2 or CHD, and 15 had incomplete/missing data. Valid scores for the final (complete-case) analysis were obtained from 583 patients.

2.4. Measurements

2.4.1. PHQ-9

The PHQ-9 is a brief instrument for screening and diagnosing depressive symptoms that was developed by Kroenke et al. (2001). The continuous sum score can be used for screening purposes while the dichotomous algorithm score can be used for diagnostic purposes. The PHQ-9 comprises the following nine items which evaluate the presence of one of the nine symptoms of depression based on the DSM-IV criteria: (a) depressed mood, (b) anhedonia, (c) trouble sleeping, (d) feeling tired, (e) change in appetite or weight, (f) guilt or worthlessness, (g) trouble concentrating, (h) feeling slowed down or restless, (i) suicidal thoughts. These nine items have the following answer categories: “not at all” (0), “various days” (1), “more than half the days” (2) and “almost every day” (3), resulting in a sum score of 0 to 27. The Dutch version of the PHQ-9, which was translated and validated by Zuithof et al., was used for the current study (Zuithof et al., 2010).

Earlier research reported scores in the 5–9 range are considered an indication for minor depression while scores of 10 or higher were considered an indication for major depression (Kroenke et al., 2001; Kroenke & Spitzer 2002). Since little is