

What does the beck depression inventory measure in myocardial infarction patients? a psychometric approach using item response theory and person-fit

KLAAS J. WARDENAAR,¹ ROB B.K. WANDERS,¹ ANNELIEKE M. ROEST,¹ ROB R. MEIJER² & PETER DE JONGE¹

¹ University Medical Center Groningen, Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), University of Groningen, Groningen, The Netherlands

² Department of Psychometrics and Statistics, University of Groningen, Groningen, The Netherlands

Key words

depression, myocardial infarction, item response theory, person-fit, Beck Depression Inventory (BDI)

Correspondence

Klaas J. Wardenaar, University Medical Center Groningen, Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE) (CC-72), University of Groningen; PO Box 30.001, 9700 RB Groningen, The Netherlands.
Telephone (+31) 50-36-14610
Fax (+31) 50-36-19722
Email: k.j.wardenaar@umcg.nl

Received 28 October 2014;
revised 4 February 2015;
accepted 17 March 2015

Abstract

Observed associations between depression following myocardial infarction (MI) and adverse cardiac outcomes could be overestimated due to patients' tendency to over report somatic depressive symptoms. This study was aimed to investigate this issue with modern psychometrics, using item response theory (IRT) and person-fit statistics to investigate if the Beck Depression Inventory (BDI) measures depression or something else among MI-patients.

An IRT-model was fit to BDI-data of 1135 MI patients. Patients' adherence to this IRT-model was investigated with person-fit statistics. Subgroups of "atypical" (low person-fit) and "prototypical" (high person-fit) responders were identified and compared in terms of item-response patterns, psychiatric diagnoses, socio-demographics and somatic factors. In the IRT model, somatic items had lower thresholds compared to depressive mood/cognition items. Empirically identified "atypical" responders ($n = 113$) had more depressive mood/cognitions, scored lower on somatic items and more often had a Comprehensive International Diagnostic Interview (CIDI) depressive diagnosis than "prototypical" responders ($n = 147$). Additionally, "atypical" responders were younger and more likely to smoke. In conclusion, the BDI measures somatic symptoms in most MI patients, but measures depression in a subgroup of patients with atypical response patterns. The presented approach to account for interpersonal differences in item responding could help improve the validity of depression assessments in somatic patients. *Copyright © 2015 John Wiley & Sons, Ltd.*

Introduction

Research findings suggests that there is a relationship between acute coronary syndromes (ACS), such as myocardial infarctions (MI), and depression. In community samples, depression has been found to be a risk factor for cardiovascular disease (CVD; van der Kooy *et al.*, 2007) and in ACS patients, depression has been found to increase the risk of new cardiac events (Meijer *et al.*, 2013) and mortality (Frasure-Smith *et al.*, 1993, 1995; Carney *et al.*, 2003; Bush *et al.*, 2001; Meijer *et al.*, 2013). Conversely, MI has been found to increase the risk of depression (Thombs *et al.*, 2006) and depression chronicity (Martens *et al.*, 2008). However, the supposed bidirectional link between depression and ACS is controversial. An often-heard criticism is that the association may be overestimated due to biased depression measurements (Koenig *et al.*, 1997; Sørensen *et al.*, 2005; Thombs *et al.*, 2008; Delisle *et al.*, 2012). Widely used questionnaires such as the Beck Depression Inventory (BDI) include items that assess somatic/functioning symptoms, which are common in depression but also in somatic illness, potentially leading to overestimated depression scores in patients with a somatic illness (Leentjens *et al.*, 2000; Moran and Mohr, 2005; Delisle *et al.*, 2012).

Despite the importance of issues with depression measurement in psychosomatic research, thorough investigations with modern psychometrics have been scarce. Previous factor analytical studies in somatic patients have shown that somatic items are part of a general depression severity scale but also constitute a specific domain within depression questionnaires (e.g. Thombs *et al.*, 2008; Chilcot *et al.*, 2011). This suggests that the endorsement of somatic items is not only explained by the presence of depression, but also by other sources, such as somatic problems and/or functional impairments. Although the earlier mentioned studies have shown the underlying structure of depression-related symptoms, they provide no insight into the endorsement probabilities of individual items. This makes factor analyses of limited use when trying to understand depressive symptom reporting in somatically ill patients. Investigation of the latter with item response theory (IRT, Embretson and Reise, 2000) could help to better understand how each symptom contributes to somatic patients' observed depression scores.

In an IRT model, each item in a questionnaire has a threshold on an underlying severity dimension that reflects its endorsement probability. Items that assess mild symptoms have low thresholds and are frequently reported; severe items have high thresholds and are less frequently reported. In an IRT model, the items' thresholds

and slopes are estimated based on raw item response data, providing insight into the measurement characteristics of the instrument. In somatic patients, somatic items of a depression scale (e.g. "energy loss") may be more frequently reported and thus have lower thresholds compared to items that cover depressive mood and cognitions (e.g. "sadness", "feeling guilty"; Wanders *et al.*, 2015). This would entail that depression scores in this group are more reflective of somatic problems/functional impairments than of the full breadth of depressive symptomatology.

However, within a sample of patients there is still considerable interpersonal variation in the level of adherence to the group-based IRT model that is estimated based on all subjects' data. In an MI sample, for example, the group-based IRT model may suggest that somatic problems are over-reported. In such a case, depressed patients that mainly report depressive mood/cognitions will not obey to the group-based IRT model and show a response pattern that is atypical for the group. This depressed subgroup is very interesting for psychosomatic research, but is easily overlooked when, based on the group-based IRT findings, all patients with increased BDI (or other questionnaire) scores are dismissed as merely over-reporting somatic depressive symptoms. A person-centered IRT approach can help to identify persons that do not conform to the group-based IRT model. In this approach, each person's level of adherence to the group-based IRT-model can be expressed by a *person-fit* statistic (Meijer and Sijtsma, 2001; Meijer, 2003). High person-fit indicates strong adherence to the group IRT-model (i.e. typical response patterns) and low person-fit indicates poor adherence (i.e. atypical response patterns). As such, person-fit can be used to investigate the interpersonal variations in item reporting and the factors that influence what a questionnaire measures across different persons. Eventually, this approach could help to better distinguish those patients, for whom increased depression scores reflect depression from those patients, for whom increased scores reflect only somatic illness.

The current study aimed to use a combination of group-level IRT and person-fit in a large MI patient sample ($n = 1135$) to identify and investigate patients with (somatically) biased BDI scores and patients with BDI scores reflecting true depression. The analyses were conducted in several steps. First, an IRT-model was fit to the complete BDI-data to investigate the item-characteristics. Second, each patient's person-fit was calculated and finally the association of person-fit with interview-based DSM-IV major depressive disorder (MDD) and other external variables was investigated.

Methods

Participants and procedures

Data came from two studies with similar inclusion criteria that were combined in previous studies as well (Zuidersma *et al.*, 2012; Bot *et al.*, 2012). The Myocardial Infarction and Depression Intervention Trial (MIND-IT) study (van den Brink *et al.* 2002; van Melle *et al.* 2007) was a multicenter randomized trial to investigate the effects of antidepressants in depressed MI patients. The Depression after Myocardial Infarction (DepreMI) study (Spijkerman *et al.*, 2006) was a naturalistic cohort study to investigate the effects of depression on cardiovascular outcome in MI patients.

In MIND-IT, MI-patients were recruited from 11 hospitals in the Netherlands. Inclusion criteria were: age ≥ 18 and a documented increase in cardiac enzymes together with at least 20 minutes of chest pain or electrocardiographic (ECG) changes typical of an MI. Exclusion criteria were: the presence of disease influencing short-term life expectancy, inability to participate (e.g. communication problems, absence), receiving psychiatric treatment and/or participating in another trial. Of 2177 recruited patients, 331 met criteria of post-MI depression and were randomized to treatment (antidepressants and/or psychotherapy) or care-as-usual.

In DepreMI, 528 MI patients were recruited from four hospitals in the Netherlands. Patients were included if they had increased cardiac enzymes, chest pain for at least 20 minutes and pathological Q-waves in their ECG in at least two leads. Patients were excluded if they had a life expectancy < 1 year (due to non-cardiac illnesses), were in poor physical condition, had cognitive problems, spoke insufficient Dutch, and/or were scheduled to have their future check-ups in a non-participating hospital. Both MIND-IT and DepreMI were approved by the institutional ethical review boards of their respective participating institutions. All patients provided informed consent.

From both samples, baseline BDI data collected before the administration of any treatment were combined in a single dataset. Only patients with a complete BDI and psychiatric interview (Comprehensive International Diagnostic Interview [CIDI], WHO, 1990) were included in the analyses to enable a comparison between the BDI and interview-based depression diagnoses. In DepreMI, the CIDI was administered to most patients ($n = 442$) and to a subsample with a BDI ≥ 10 in MIND-IT ($n = 760$). Taken together, 1202 patients had the required data. Of these, 67 (5.6%) were excluded because they had missing responses on five or more BDI items. The study sample consisted of 1135 patients.

Measures

Depression

The BDI version 1 (Beck *et al.*, 1961, 1988) was administered to measure depression severity. The 21 items were scored on a four-point Likert scale (0–3). In DepreMI, the CIDI 1.1 (WHO, 1990) was administered and in MIND-IT, the CIDI 2.1 (WHO, 1997) was administered to evaluate whether the International Classification of Diseases – 10th revision (ICD-10) criteria (WHO, 1993) were met for depression after MI. The presence of anxiety disorders following MI (generalized anxiety disorder [GAD], panic disorder, social phobia, agoraphobia) was also assessed with the CIDI.

Demographic and clinical characteristics

All baseline clinical and demographic characteristics were assessed during hospitalization from the hospital charts. Left ventricular ejection fraction (LVEF) was assessed with radionuclide ventriculography, echocardiography, gated single photon emission computed tomography, angiography, magnetic resonance imaging, or clinical assessment. The following clinical variables were used in the current study: LVEF, Killip class, anterior site of MI, history of MI, history of cerebral vascular disease, history of peripheral vascular disease, family history of coronary artery disease, diabetes, hypertension, hypercholesterolemia, body mass index (BMI) and current smoking.

Missing data

One hundred and fifty-seven (13.8%) subjects had ≤ 3 missing responses, which were imputed with a k-nearest-neighbor (KNN) search (Troyanskaya *et al.*, 2001). KNN selects a number (k) of subjects that are most similar in terms of their response pattern to the subject, whose missing responses are to be imputed. Next, KNN imputes the weighted mean of the nearest neighbors' responses on the target item. Imputations were done with R-package "impute" (Hastie *et al.*, 2014) with default settings ($k = 10$). This method was chosen because item-reporting was hypothesized to vary across persons and imputation of scores calculated on the group level would not be in line with this.

Statistical analyses

Non-parametric item response theory (IRT)

Preliminary non-parametric IRT analyses (Meijer *et al.*, 2014) were done with R-package "KernSmoothIRT" (Mazza *et al.*, 2014) to evaluate the suitability of the data

for parametric IRT modeling by plotting the non-parametric probability curves (Ramsay, 1991) of each item's categories. The plots were inspected to check whether item-responses were meaningfully related to underlying severity (Meijer and Baneke, 2004; Sijtsma and Molenaar, 2002). Items that showed response behavior that was not in line with IRT assumptions were removed from subsequent analyses.

Factor analysis

An exploratory factor analysis (EFA) was conducted with the BDI data, using a weighted least squares (WLSMV) estimator for use with ordinal variables in Mplus (version 5; Muthén and Muthén, 1998–2007). The ratio between the Eigenvalues of the first and the respective Eigenvalues of the second, third and fourth factor were inspected to evaluate the extent of unidimensionality (Reise *et al.*, 2013).

Item response theory (IRT)

A graded response model (GRM, Samejima, 1969) was fit to the data with R-package “*ltm*” (Rizopoulos, 2006). Rather than choosing a model based on model fit, GRM was chosen a priori because it corresponded best with the categorical, ordered nature of the used data (Reise and Waller, 2003). Two parameters were fit for each item: the *slope* indicates the strength of the relationship between the item and underlying severity and the *thresholds* (0–1, 1–2, 2–3) indicate the severity of the symptom that is assessed by the item. The items' IRT parameters were inspected and items were ordered by their average thresholds.

Person-fit

The likelihood-based person-fit statistic l_z (Drasgow *et al.*, 1996) was used to quantify persons' adherence to the fitted IRT model. The l_z statistic reflects the likelihood of observing a response pattern, given the group-based IRT model. High l_z values (high person-fit) indicate strong consistency with the IRT-model whereas low l_z values indicate that a person's response pattern is atypical and adheres poorly to the group-based IRT-model (Meijer and Sijtsma, 2001; Meijer, 2003). To gain insight into its external correlates, person-fit was used as a continuous outcome variable in univariable linear regression analyses with external variables as determinants. Univariable analyses were conducted first, followed by a multivariable analysis including all independent variables with a univariate p -value < 0.10 . In another approach, external characteristics were compared between person-fit subgroups to gain insight into the more patient-specific characteristics of those with low versus high

person-fit. Because the person-fit statistic was not normally distributed, a Monte Carlo simulation procedure was used to simulate an empirical person-fit distribution with the same sample characteristics as the observed data, from which empirically-based unbiased person-fit cutoffs at a 10% significance level were derived (Seo and Weiss, 2013). The values corresponding to the 10% upper and lower part of the obtained empirical distribution were used as cutoffs to allocate patients in, respectively, prototypical and atypical subgroups based on their person-fit values. The other patients were allocated to a middle group. The atypical group ($n = 112$, 9.8%) had person-fit below the 10% person-fit cutoff ($l_z = -1.19$) and the prototypical group ($n = 148$, 13.0%) had person-fit above the 90% person-fit cutoff ($l_z = 1.08$). The “middle” person-fit subgroup consisted of 874 patients.

Results

Demographic and clinical characteristics

The descriptives of the study sample are shown in the left-most column of Table 1. Of the sample, 76.9% was male and the mean age was 60.6 years (standard deviation [SD] = 11.8). Of the patients, 27.8% had a low LVEF (< 45), 13.3% had a Killip class ≥ 2 and 33.9% had an anterior site of the index MI. Several patients had a history of a prior MI (15.7%), CVD (5.7%) and/or peripheral vascular disease (PVD) (9.0%). Health-related factors, such as current smoking (48.5%), hypercholesterolemia (61.1%), hypertension (32.2%), and family history of CVD (43.5%) were all common. The mean BDI score was 9.5 (SD = 6.7) and 40.6% of the sample had a CIDI diagnosis of post-MI depression. The most prevalent post-MI anxiety diagnoses were GAD (8.2%), social phobia (3.3%) and agoraphobia (2.4%).

Checking data quality

The non-parametric IRT-plots are shown in Supplementary Material 1. The response behavior on most items did not show violations with respect to the expected form of the curves. However, for items 18 and 19, responses showed no meaningful relationship with underlying severity: higher categories did not become more likely to be endorsed as severity increased. Therefore, items 18 and 19 were removed from the subsequent analyses.

EFA was conducted next. The eigenvalues of the first four factors were respectively: 7.84, 1.89, 1.02 and 0.90. The ratio of the first to second eigenvalue was 4.14 and the ratios of the first to third and first to fourth eigenvalues were even larger. These results indicated that the first factor by far explained the most common variance and, thus, that

Table 1. Complete sample and subgroup characteristics in a group of patients with an acute coronary syndrome (N = 1135)

	Total sample	Person-fit subgroups			Test-statistic	p-Value
		"Atypical"	"Average"	"Prototypical"		
	N = 1135	n = 113 (10.0%)	n = 874 (77.0%)	n = 148 (13.0%)		
Person-fit (I_z), range	-4.77 to 1.82	-4.77 to -1.20	-1.18 to 1.08	1.08 to 1.82	—	
Age mean years (SD)	60.6 (11.8)	57.9 (11.9)	60.7 (11.9)	62.3 (11.0)	F=4.6	0.01
Female gender, n (%)	262 (23.1%)	24 (21.2%)	208 (23.8%)	30 (20.3%)	$\chi^2=1.12$	0.57
Cardiac severity	316 (27.8%)	32 (33.3%)	245 (34.5%)	39 (32.8%)	$\chi^2=0.17$	0.92
LVEF < 45, n (%)	151 (13.3%)	13 (11.5%)	118 (13.5%)	20 (13.7%)	$\chi^2=0.37$	0.83
Killip class ≥ 2 , n (%)	385 (33.9%)	42 (37.2%)	297 (34.0%)	46 (31.1%)	$\chi^2=1.07$	0.59
Anterior site of MI, n (%)	178 (15.7%)	18 (15.9%)	138 (15.8%)	22 (15.1%)	$\chi^2=0.06$	0.97
Cardiac vulnerability	66 (5.8%)	8 (7.1%)	55 (6.3%)	3 (2.1%)	$\chi^2=4.89^a$	0.08
History of cerebral vascular disease, n (%)	102 (9.0%)	7 (6.2%)	84 (9.6%)	11 (7.6%)	$\chi^2=1.85$	0.40
History of peripheral vascular disease, n (%)	494 (43.5%)	49 (44.1%)	390 (45.0%)	55 (37.4%)	$\chi^2=2.96$	0.23
Family history of coronary artery disease, n (%)	140 (12.3%)	16 (14.2%)	103 (11.8%)	21 (14.2%)	$\chi^2=1.05$	0.59
Somatic health	366 (32.2%)	31 (27.4%)	285 (32.7%)	50 (33.8%)	$\chi^2=1.43$	0.49
Diabetes, n (%)	694 (61.1%)	78 (69.0%)	523 (60.0%)	93 (62.8%)	$\chi^2=3.63$	0.16
Hypertension, n (%)	26.6 (4.0)	26.6 (4.4)	26.6 (3.9)	26.9 (4.7)	F=0.40	0.67
Hypercholesterolemia, n (%)	550 (48.5%)	71 (65.7%)	420 (49.6%)	59 (41.8%)	$\chi^2=14.5$	0.001
Body mass index (BMI), mean (SD)	461 (40.6%)	58 (51.3%)	347 (39.7%)	56 (37.8%)	$\chi^2=6.15$	0.046
Current smoking, n (%)	3.0 (2.8)	3.6 (3.0)	3.0 (2.8)	2.6 (2.7)	F=3.09	0.046
CIDI expression after MI	450 (40.4%)	55 (50.0%)	339 (39.5%)	56 (38.6%)	$\chi^2=4.7$	0.10
Post-MI CIDI depression, n (%)	371 (33.5%)	39 (35.1%)	285 (33.4%)	47 (32.6%)	$\chi^2=0.2$	0.91
Number of depressive symptoms, mean (SD)	503 (45.6%)	56 (51.4%)	388 (45.7%)	59 (40.7%)	$\chi^2=2.9$	0.24
Depressive mood, n (%) present ^b	63 (5.7%)	9 (8.4%)	43 (5.1%)	11 (7.5%)	$\chi^2=3.0$	0.22
Anhedonia/loss of interest, n (%) present ^b	529 (48.3%)	55 (52.4%)	411 (48.8%)	63 (42.9%)	$\chi^2=2.5$	0.29
Energy loss, n (%) present ^b	347 (31.2%)	41 (36.9%)	267 (31.2%)	39 (26.7%)	$\chi^2=3.1$	0.22
Appetite change, n (%) present ^b	229 (20.6%)	30 (27.8%)	179 (20.9%)	20 (13.8%)	$\chi^2=7.5$	0.02
Sleeping problems, n (%) present ^b	218 (19.6%)	33 (30.8%)	160 (18.6%)	25 (17.2%)	$\chi^2=9.6$	0.01
Psychomotor agitation/retardation, n (%) present ^b	455 (41.8%)	52 (48.1%)	350 (41.7%)	53 (37.3%)	$\chi^2=3.0$	0.23
Feeling worthless/guilty, n (%) present ^b	217 (19.7%)	20 (20.0%)	167 (19.8%)	28 (19.0%)	$\chi^2=0.1$	0.98
Loss of self-esteem, n (%) present ^b	0.63 (4.5)	1.4 (6.0)	0.55 (4.5)	0.50 (2.6)	F=2.0	0.14
Concentration problems, n (%) present ^b						
Preoccupations with death/suicide, n (%) present ^b						
Number of previous episodes, mean (SD)						

CIDI anxiety after MI	Generalized anxiety disorder (GAD), n (%)	94 (8.3%)	10 (8.8%)	71 (8.1%)	13 (8.8%)	$\chi^2=0.13$	0.94
Panic disorder (PD), n (%)	22 (1.9%)	1	21	0	—	—	—
Social phobia, n (%)	38 (3.3%)	10 (8.8%)	26 (3.0%)	2 (1.4%)	$\chi^2=10.1^a$	0.005	0.005
Agoraphobia without PD, n (%)	27 (2.4%)	6 (5.3%)	19 (2.2%)	2 (1.4%)	$\chi^2=5.01^*$	0.10	0.10
Agoraphobia with PD, n (%)	5 (0.4%)	0	3	0	—	—	—
Number of anxiety disorders, mean (SD)	0.16 (0.41)	0.24 (0.50)	0.16 (0.41)	0.12 (0.31)	F=2.95	0.053	0.053

Note: SD, standard deviation; MI, myocardial infarction; LVEF, left ventricular ejection fraction.

^aFor χ^2 tests with cell-counts < 5, Fisher's exact test is reported.

^bIndividual item-level CIDI symptoms were available for 1112 participants (98.0%).

additional dimensions did not explain much additional variance. These results are in line with bifactor modeling results by Brouwer *et al.* (2013), showing that the total score of the updated BDI (the BDI-II) adequately reflected overall depression severity and that only limited variance was explained by additional domain-specific factors.

Item response theory (IRT)

The IRT parameters are shown in Table 2. All items at the lowest end of the (theta) severity dimension (i.e. with the lowest mean thresholds) covered somatic symptoms and/or functional impairments. “Fatigability” (item 17), “work inhibition” (item 15), “sleep problems” (item 16), “loss of libido” (item 21), “indecisiveness” (item 13), “somatic preoccupation” (item 20) and “lack of satisfaction” (item 4) had mean thresholds ranging from 1.00 to 1.84. Items that covered depressive mood and/or depressive cognitions (e.g. “depressed mood” [item 1], “guilty feelings” [item 5], “social withdrawal” [item 12] and “suicidal thoughts” [item 9]) were all located higher on the severity dimension with average thresholds ranging from 2.29 to 3.42. Similar ordering of mean thresholds was seen when the IRT model was fitted in the DepreMI and MIND-IT subsamples separately (see Supplementary Material 2). These results indicated that somatic/functional BDI items were reported at lower severity levels and that mood/cognitive BDI items were only reported at higher severity levels.

Person-fit

Person-fit distribution plots are shown in Figure 1. The frequency distribution of person-fit showed some negative skewness with a tail extending into the lower person-fit range. The distribution was quite uniform across increasing levels of severity (theta), although very low person-fit was more common at higher severity levels. Prior to regression analyses, person-fit was transformed as follows: $-1 \times (\ln(I_z^{\max} - I_z))$. The results of the regression analyses with person-fit as dependent variable and the external variables as independent variables are shown in Table 3. Age and hypertension were positively associated and smoking was negatively associated with person-fit. In addition, person-fit was negatively associated with the number of reported CIDI depressive symptoms (beta = -0.111), the presence of several individual CIDI symptoms (e.g. “feeling worthless/guilty” [beta = -0.110], “depressive mood” [beta = -0.089] and “psychomotor problems” [beta = -0.081]), CIDI social phobia, agoraphobia, and the total number of present anxiety disorders. In the multivariable analyses, only smoking, “feeling worthless/guilty”, “depressive mood” and “psychomotor problems” were

Table 2. Item response theory (IRT) item-parameters for the beck depression inventory (BDI) in a sample of patients with acute cardiac syndromes ($n = 1135$) ordered by mean item-threshold-value

BDI item	Slope	Threshold 1 ^a	Threshold 2 ^a	Threshold 3 ^a	Mean threshold
17-Fatigability	1.13	-1.97	1.62	3.35	1.00
15-Work inhibition	1.53	-0.78	1.30	2.57	1.03
16-Sleep problems	0.95	-0.51	1.59	2.84	1.31
21-Loss of libido	0.80	-0.10	1.74	2.88	1.51
13-Indecisiveness	1.65	0.20	1.42	3.33	1.65
20-Somatic preoccupation	1.56	-0.14	1.83	3.59	1.76
2-Pessimism	2.52	1.17	1.88	2.43	1.83
4-Lack of satisfaction	2.08	-0.01	2.46	3.08	1.84
11-Irritability	1.06	0.32	2.54	2.84	1.90
1-Mood	1.89	0.78	2.20	3.90	2.29
6-Sense of punishment	1.09	1.99	2.49	2.52	2.33
10-Crying	1.12	0.89	3.06	3.15	2.37
3-Sense of failure	1.88	1.58	2.53	3.68	2.60
5-Guilty feelings	1.86	1.51	2.73	3.59	2.61
8-Self accusation	1.56	1.27	3.08	4.04	2.80
14-Body image	1.41	1.95	2.95	3.80	2.90
7-Self dislike	1.89	1.28	3.62	4.07	2.99
12-Social withdrawal	1.49	1.37	3.20	5.06	3.21
9-Suicidal thoughts	1.68	2.06	3.76	4.42	3.41

Parameters based on a graded response IRT model.

^aThe category thresholds represent the level of depression (theta) necessary to report the category or higher.

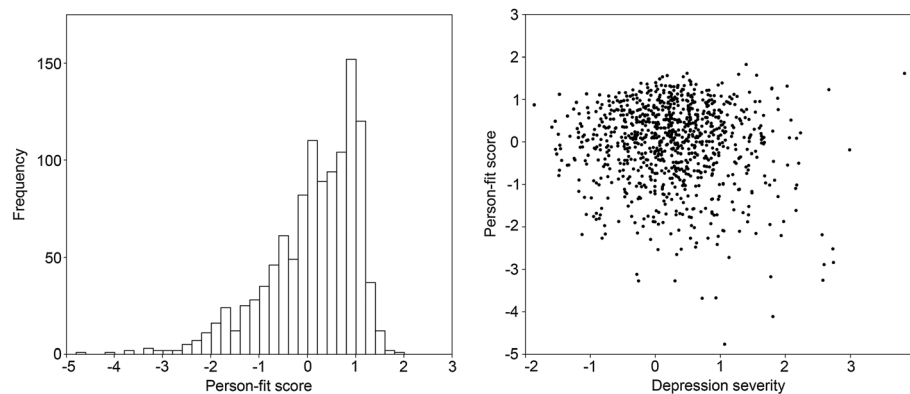


Figure 1. The person-fit frequency distribution (left) and the individual person-fit values (y -axis) plotted against severity (θ) according to the item response model (right).

negatively associated with person-fit. These results indicated that lower adherence to the group-based IRT model was associated with more features of clinical depression.

Person-fit subgroups

The mean item-scores for the atypical and prototypical subgroups are shown in the upper panel of Figure 2. Most mean item scores were higher in the atypical group, in line

with the finding that person-fit was inversely related with the number of depressive symptoms and the observation that the atypical group reported comparatively higher severity (θ). When these quantitative severity differences were adjusted by centering the mean scores in the atypical and prototypical groups on the middle group (Figure 2, lower panel), the groups' qualitative response pattern differences became more clearly visible. Compared to the prototypical group, the atypical group showed relatively

Table 3. Regression analyses to investigate the associations of externally determined variables with person-fit (outcome) in a group of patients with an acute coronary syndrome ($N = 1135$)

Outcome = $-1 \times (\ln(I_z^{\max} - I_z))^a$		Univariable ^b		Multivariable ^c	
		beta	p-Value	beta	p-Value
	Female gender	-0.011	0.71	—	—
	Age	0.114	<0.001	0.046	0.21
	<i>Cardiac severity factors</i>				
	LVEF < 45	0.001	0.97	—	—
	Killip class ≥ 2	-0.005	0.86	—	—
	Anterior site of MI	-0.038	0.20	—	—
	<i>Cardiac vulnerability factors</i>				
	History of MI	0.016	0.59	—	—
	History of cerebral vascular disease	-0.030	0.32	—	—
	History of peripheral vascular disease	-0.005	0.88	—	—
	Family history of coronary artery disease	-0.044	0.14	—	—
	<i>Somatic health characteristics</i>				
	Diabetes	-0.004	0.89	—	—
	Hypertension	0.061	0.04	0.036	0.28
	Hypercholesterolemia	-0.021	0.48	—	—
	Body mass index, mean (SD)	-0.008	0.80	—	—
	Current smoking	-0.115	<0.001	-0.078	0.03
	<i>CIDI depression after MI</i>				
	Depression present after MI	-0.052	0.08	-0.025	0.60
	Number of CIDI depressive symptoms	-0.111	0.001	-0.181	0.14
	<i>Individual symptoms after MI</i>				
	Depressive mood	-0.089	0.003	0.108	0.04
	Anhedonia/loss of interest	-0.030	0.32	—	—
	Energy loss	-0.048	0.11	—	—
	Appetite loss	-0.014	0.65	—	—
	Sleeping problems	-0.074	0.01	-0.048	0.33
	Psychomotor agitation/retardation	-0.081	0.007	-0.095	0.04
	Feeling Worthless/Guilty	-0.110	<0.001	-0.101	0.03
	Loss of Self-esteem	-0.052	0.08	0.017	0.70
	Concentration problems	-0.058	0.06	-0.062	0.22
	Preoccupations with death/suicide	-0.050	0.10	—	—
	#of previous depressive episodes	-0.052	0.08	-0.043	0.20
	<i>CIDI anxiety after MI</i>				
	Generalized anxiety disorder	0.020	0.51	—	—
	Panic disorder (PD)	-0.056	0.06	-0.071	0.05
	Social phobia	-0.067	0.03	-0.071	0.07
	Agoraphobia	-0.059	0.048	-0.062	0.11
	Agoraphobia with PD	-0.040	0.18	—	—
	Number of anxiety disorders present	-0.063	0.03	0.058	0.24

MI, myocardial infarction; LVEF, left ventricular ejection fraction; PTCA, percutaneous transluminal coronary angioplasty; CABG, coronary artery bypass graft (surgery).

^a I_z was transformed in this way because it was negatively skewed. The transformed outcome was multiplied by -1 to make sure that positive coefficients reflect positive associations and negative coefficients reflect negative associations.

^bTransformed outcome mean = 0.34 (SD = 0.55; range: -1.57 to 1.89).

^cConducted in a subsample of participants that missed none of the covariates ($n = 924$; 81.3%); transformed outcome mean = 0.38 (SD = 0.55; range = -1.57 to 1.78).

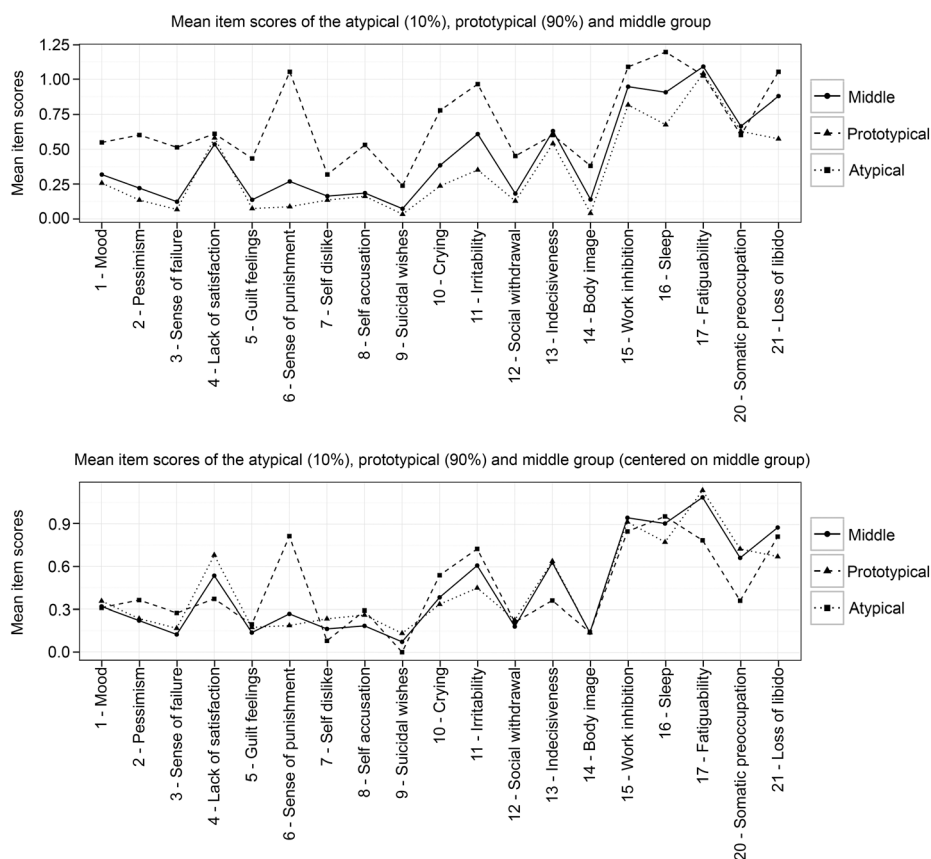


Figure 2. Mean item scores for the atypical, prototypical and middle person-fit groups. The upper panel displays the raw means, the lower panel displays the means after centering on the middle group (to eliminate quantitative severity differences and allow for better qualitative comparison of response patterns). Patients were allocated to the subgroups based on the 10% lowest (atypical) and 10% highest (prototypical) person-fit cutoff values. Because the observed person-fit was not normally distributed, a Monte Carlo simulation procedure was used to simulate an empirical person-fit distribution with the same sample characteristics as the observed data, from which unbiased person-fit cutoffs were derived.

higher scores on “pessimism” (item 2), “sense of failure” (item 3), “sense of punishment” (item 5), “crying” (item 10) and “irritability” (item 11) and relatively lower scores on “lack of satisfaction” (item 3), “indecisiveness” (item 13), “fatigability” (item 17) and “somatic preoccupation” (item 20). These results indicated that person-fit was related both to overall depression severity and specific patterns of item-endorsement.

Person-fit subgroups and external variables

The characteristics and comparisons of the subgroups are shown in Table 1. The atypical group was younger (mean age: 57.9 years) than the other groups, and the prototypical group was the oldest (mean age: 62.3 years). Of all health-related factors, only current smoking was more common in the atypical group. In addition, there were

more patients with an actual CIDI MDD diagnosis in the atypical group and they more often reported “feeling worthless/guilty” and “loss of self-esteem”. Other CIDI symptom ratings did not differ across subgroups. Finally, there were more patients with a CIDI diagnosis of agoraphobia in the atypical group than in the other groups.

Discussion

This study was aimed to use a combination of group- and person-centered IRT approaches to investigate (1) whether in a sample of MI patients the BDI was biased towards measuring predominantly somatic/functional problems rather than actual depressive symptoms, and (2) whether it is possible to distinguish true depressive patients using person-fit. The group-based IRT results showed that somatic/functional items all had thresholds

at the lower end of the severity spectrum, and were thus likely to be endorsed by most patients. Items covering depressive mood/cognitions all had higher thresholds and were thus less likely to be endorsed by most patients. Thus, at lower severity levels the BDI predominantly measures somatic symptoms, whereas at higher levels it measures depressive mood and cognitions. Variation in the adherence to this group-based IRT-model was investigated by calculating person-fit statistics. Low person-fit indicated that depressive mood/cognitions symptoms were endorsed, whereas somatic symptoms were not (or scarcely) endorsed (low adherence to the group-based IRT model) and regression analyses showed that this was positively associated with more CIDI depressive symptomatology and CIDI anxiety diagnoses. Comparison of person-fit subgroups (e.g. atypical versus prototypical responders) showed that those with atypical response patterns were younger, smoked more often, were more likely to have a CIDI diagnosis of MDD or social phobia and reported more depressive symptoms on the CIDI. There were no differences in the rates of CIDI somatic symptoms. These results suggested that for most MI patients, BDI scores reflect the presence of somatic symptoms, whereas for atypical patients BDI scores more often reflect symptoms indicative of depression according to clinical classifications.

These results have several interesting implications. First, results indicate that the BDI predominantly measures somatic/functional symptoms in MI patients with low BDI scores, in line with previous suggestions (e.g. Leentjens *et al.*, 2000; Moran and Mohr, 2005; Delisle *et al.*, 2012). Although the current study looked at the BDI, which is just one of several broadly used questionnaires, the identified measurement properties are unlikely to be unique to the BDI. Other broadly used questionnaires such as the Patient Health Questionnaire (PHQ-9; Kroenke *et al.*, 2001) and BDI-II (Beck *et al.*, 1996) also contain items that are likely to be over endorsed by somatic patients. However, an investigation of the BDI-II in MI patients showed that on this instrument, somatic symptoms were less over-reported by MI patients, although, in line with the currently presented group-based IRT model, somatic symptoms did account for a majority (73.9%) of low (BDI-II < 4) scores and this percentage decreased with increasing BDI-II scores (35.5% for those with BDI-II > 12; Thombs *et al.*, 2010).

The findings suggest that in most MI patients, BDI scores reflect predominantly somatic problems. This is especially the case for those with relatively low BDI scores, since mainly somatic items were endorsed in the low severity range of the scale, in contrast to the mood/cognitive

items. This relation between response-behavior and severity is important because previous studies on the association between depression and MI/ACS have often included “depressed” patients based on relatively low scale cutoffs (e.g. BDI ≥ 10 for mild-moderate depression, Frasurre-Smith *et al.* 1995, 1999), which are lower than most clinical cutoffs (e.g. Beck *et al.*, 1988: BDI ≥ 10 for mild and BDI ≥ 19 for moderate depression). The current results (Table 2) suggest that a patient could meet such a cutoff by reporting almost exclusively somatic symptoms. Also, when using the BDI (or other instrument) as a continuous determinant, the range of the BDI scores is likely to influence the extent to which scores are biased due to over-reporting, with more somatic bias in the low severity range.

However, to state that the BDI always measures predominantly somatic symptoms in MI patients (and other somatic patient populations) would be too simplistic given the second implication of the results. The person-fit analyses showed that, despite the group-level IRT model, not all individual patients reported strictly somatic symptomatology. Indeed, lower person-fit was associated with increased key-features of depression and more comorbid anxiety, more indicative of the presence of psychopathology according to clinical classifications. If all BDI scores would be dismissed as being reflective of patients’ somatic symptomatology and useless to detect depressive severity, these patients would be overlooked. This would be a shame because in any population of somatic patients there can be truly depressed patients that will respond accordingly on a depression questionnaire. Person-fit could help to distinguish such relevant cases from the majority of non-depressed MI-patients. If developed further, this could be of scientific and clinical interest, as it could help to improve the currently suboptimal specificity of questionnaires to detect true depressive cases.

The earlier finding suggests that person-fit could be helpful in clinical practice to identify potentially relevant cases. Once a group-based IRT-model has been established in a norm-population, computation of a person-fit statistics for individual patients is straightforward. An individualized person-fit statistic can then be used to evaluate whether a patient responds in a way that is atypical for the population he/she belongs to and could indicate the need for closer scrutiny of the individual item responses. Especially when questionnaires are administered digitally, such a procedure could be implemented quite easily.

The IRT findings are of conceptual interest in the light of previously developed dimensional approaches to distinguish between the somatic and mood/cognitive aspects of depression in psychosomatic research (e.g. Ormel and de

Jonge, 2011). Some studies have used factor analyses to extract distinct factors and have used these in psychosomatic research (e.g. de Jonge *et al.*, 2006). There is an ongoing discussion about whether factor models of depression should take the form of a factor model with co-existing (correlated) factors (e.g. de Jonge *et al.*, 2006) or that one should use a hierarchical, bifactor approach (Thombs *et al.*, 2008). The current IRT study did not look deeply into the matter of (multi)dimensionality, but did indicate that there may exist yet another distinction between somatic versus mood/cognitive items, with one cluster of items (somatic) positioned along the lower range of the severity spectrum and a second cluster (mood/cognitive) along the upper range. This indicates that in MI patients, item-clustering is related to the severity of the reported BDI score. Presumably, in populations without somatic illnesses, item-thresholds could be more evenly distributed along the severity dimension, making such clustering effects less likely.

Interestingly, we found no differences between the prototypical and atypical subgroup on measures of cardiac vulnerability such as LVEF, Killip class, and history of MI. These findings appear to be in contrast with a previous study on these data that showed stronger associations between cardiac vulnerability measures and the somatic depressive symptoms dimension compared to the cognitive/affective symptom dimension (de Jonge *et al.*, 2006). However, another report on data from the MIND-IT study showed that a lower LVEF was not only associated to higher baseline BDI scores, but also to increased rates of a depression diagnosis in the year following MI (van Melle *et al.*, 2005), suggesting that a reduced LVEF is not only characteristic for patients exhibiting somatic/functional symptoms (as seen in the prototypical subgroup), but also for those with a formal depression diagnosis (as seen in the atypical subgroup).

The present results could explain previous findings from studies on the association between depression and MI/ACS outcome. For instance, one study showed that when adjusting for BDI score, previously observed associations between clinical depression and poor cardiac outcomes diminished, whereas BDI score remained predictive of cardiac outcomes (Zuidersma *et al.*, 2013). Another study among ACS patients found that an overall depression severity measure (Hospital Anxiety and Depression Scale, HADS-D) was associated with mortality, whereas the BDI Fast-Screen, which includes only the mood/cognitive items of depression, was not. Again, the full severity instrument could reflect a large amount of somatic symptom severity. Both studies' findings could be explained by the current finding that the BDI score is a

somatic severity indicator in most patients, and thus, a good predictor of poor cardiac outcome. In line with this, the observation that even minimally increased BDI scores are associated with increased mortality among MI-patients (Bush *et al.*, 2001) could be explained by the current finding that low BDI scores reflect predominantly somatic symptomatology.

The current study had several strengths, including the modern psychometric analyses, large sample size, and the fact that the psychometric results could be linked to clinically relevant external factors. However, several study limitations should also be kept in mind when interpreting the results. First, patients with a BDI < 10 in the MIND-IT sample were excluded because they had no CIDI assessment, potentially leading to selection bias. However, IRT-analyses in all patients with and without a CIDI assessment ($n = 2469$) showed a similar ordering of low item thresholds for somatic items and higher thresholds for depressive mood/cognition items (results not shown), indicating that the effect of sample selection on the IRT result was limited. Second, the number of assessed external variables is limited and other relevant determinants of person-fit could exist. Third, the results apply to the BDI and, although similar effects would be expected, the generalizability of the results to other, more recently developed instruments (e.g. BDI-II, Beck *et al.*, 1996; Inventory of Depressive Symptomatology, Rush *et al.*, 1996) should be investigated. Future research could focus on such replication efforts but also on further model development to better differentiate between groups of patients based on their response tendencies. Also, to find out whether the clustering of somatic symptoms at the lower end of the severity depression spectrum purely reflects somatic disease or also the presence of a latent depression subtype, future research should repeat the current analyses in a somatically healthy sample. The current (atypical/prototypical) subgroups were shown to be different in terms of psychopathology characteristics (e.g. depression diagnoses), but differentiation between distinct types of patients based on their item reporting could be further improved by use of a data-driven mixture approach to person-fit or mixture IRT (Cohen and Bolt, 2005). Eventually, such an approach could help to even better distinguish those that report increased depression scores as an expression of mental health problems from the patients that report increased depression scores as an expression of their somatic illness.

In conclusion, the results indicate that the BDI measures predominantly somatic symptoms in the majority of MI patients with low BDI scores. However, person-fit could be used to identify a subgroup of patients with atypical response patterns and for whom BDI scores were

indicative of clinical depression, which showed the potential usefulness of person-fit statistics for clinical purposes. In addition, the overall findings illustrate the potential usefulness of a person-centered IRT approach to gain more insight in symptom-specific link between depression measurements and cardiac outcomes.

Acknowledgments

The current study was supported by a VICI grant (no. 91812607) received by Peter de Jonge from the Netherlands

Organization for Scientific Research (NWO-ZonMW). MIND-IT was sponsored by the Netherlands Heart Foundation (no. 97.016). MIND-IT received educational grants from Organon (The Netherlands) and Lundbeck (Denmark). DepreMI was sponsored by a grant from the Netherlands Organization for Scientific Research (ZonMW; grant no. 904-57-106). The authors report no conflicts of interest.

Declaration of interest statement

The authors have no competing interests.

References

- Beck A.T., Ward C., Mendelson M. (1961) Beck Depression Inventory (BDI). *Archives of General Psychiatry*, **4**(6), 561–571.
- Beck A.T., Steer R.A., Garbin M.G. (1988) Psychometric properties of the Beck Depression Inventory: twenty-five years later. *Clinical Psychology Review*, **8**(1), 77–100.
- Beck A.T., Steer R.A., Brown G.K. (1996) *BDI-II Manual*, San Antonio, TX, The Psychological Corporation.
- Bot M., Pouwer F., Zuidersma M., van Melle J.P., de Jonge P. (2012) Association of coexisting diabetes and depression with mortality after myocardial infarction. *Diabetes Care*, **35**(3), 503–509.
- Brouwer D., Meijer R.R., Zevalkink J. (2013) On the factor structure of the Beck Depression Inventory-II: G is the key. *Psychological Assessment*, **25**(1), 136–145.
- Bush D.E., Ziegelstein R.C., Tayback M., Richter D., Stevens S., Zahalsky H., Fauerbach J.A. (2001) Even minimal symptoms of depression increase mortality risk after acute myocardial infarction. *American Journal of Cardiology*, **88**(4), 337–341.
- Carney R.M., Blumenthal J.A., Catellier D., Freedland K.E., Berkman L.F., Watkins L.L., Czajkowski S.M., Hayano J., Jaffe A.S. (2003) Depression as a risk factor for mortality after acute myocardial infarction. *American Journal of Cardiology*, **92**(11), 1277–1281.
- Chilcot J., Norton S., Wellsted D., Almond M., Davenport A., Farrington K. (2011) A confirmatory factor analysis of the Beck Depression Inventory-II in end-stage renal disease patients. *Journal of Psychosomatic Research*, **71**(3), 148–153.
- Cohen A.S., Bolt D.M. (2005) A mixture model analysis of differential item functioning. *Journal of Educational Measurement*, **42**(2), 133–148.
- de Jonge P., Ormel J., van den Brink R.H., van Melle J.P., Spijkerman T.A., Kuijper A., van Veldhuisen D.J., van den Berg M.P., Honig A., Crijns H.J., Schene A.H. (2006) Symptom dimensions of depression following myocardial infarction and their relationship with somatic health status and cardiovascular prognosis. *American Journal of Psychiatry*, **163**(1), 138–144.
- Delisle V.C., Abbey S.E., Beck A.T., Dobson K.S., Dozois D.J., Grace S.L., Stewart D.E., Ziegelstein R.C., Thombs B.D. (2012) The influence of somatic symptoms on Beck Depression Inventory scores in hospitalized postmyocardial infarction patients. *Canadian Journal of Psychiatry*, **57**(12), 752–758.
- Drasgow F., Levine M.V., Zickar M.J. (1996) Optimal identification of mismeasured individuals. *Applied Measurement in Education*, **19**(1), 47–64.
- Embretson SE, Reise SP. (2000) *Item Response Theory for Psychologists*, Hillsdale, NJ, Erlbaum.
- Frasure-Smith N., Lespérance F., Talajic M. (1993) Depression following myocardial infarction: impact on 6-month survival. *JAMA*, **270**(15), 1819–1825.
- Frasure-Smith N., Lesperance F., Talajic M. (1995) Depression and 18-month prognosis after myocardial infarction. *Circulation*, **91**(4), 999–1005.
- Frasure-Smith N., Lespérance F., Juneau M., Talajic M., Bourassa M.G. (1999) Gender, depression, and one-year prognosis after myocardial infarction. *Psychosomatic Medicine*, **61**(1), 26–37.
- Hastie T., Tibshirani R., Narasimhan B., Chu G. (2014). impute: Imputation for Microarray Data. <http://www.bioconductor.org/packages/release/bioc/manuals/impute/man/impute.pdf> [3 June 2014].
- Koenig H.G., George L.K., Peterson B.L., Pieper C. F. (1997) Depression in medically ill hospitalized older adults: prevalence, characteristics, and course of symptoms according to six diagnostic schemes. *American Journal of Psychiatry*, **154**(10), 1376–1383.
- Kroenke K., Spitzer R.L., Williams J.B. (2001) The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, **16**(9), 606–613.
- Leentjens A.F., Verhey F.R., Luijckx G.J., Troost J. (2000) The validity of the Beck Depression Inventory as a screening and diagnostic instrument for depression in patients with Parkinson's disease. *Movement Disorders*, **15**(6), 1221–1224.
- Martens E.J., Smith O.R., Winter J., Denollet J., Pedersen S.S. (2008) Cardiac history, prior depression and personality predict course of depressive symptoms after myocardial infarction. *Psychological Medicine*, **38**(2), 257–264.
- Mazza A., Punzo A., McGuire B. (2014). KernSmoothIRT: nonparametric item response theory. <http://cran.r-project.org/web/packages/KernSmoothIRT/index.html> [3 June 2014].
- Meijer R.R. (2003) Diagnosing item score patterns on a test using item response theory based person-fit statistics. *Psychological Methods*, **8**(1), 72–87.
- Meijer R.R., Baneke J.J. (2004) Analyzing psychopathology items: a case for nonparametric item response modeling. *Psychological Methods*, **9**(3), 354–368.
- Meijer R.R., Sijtsma K. (2001) Methodology review: evaluating person fit. *Applied Psychological Measurement*, **25**(2), 107–135.
- Meijer A., Conradi H.J., Bos E.H., Anselmino M., Carney R.M., Denollet J., Doyle F., Freedland K.E., Grace S.L., Hosseini S.H., Lane D.A., Pilote L., Parakh K., Rafanelli C., Sato H., Steeds R.P., Welin C., de Jonge P. (2013). Adjusted prognostic association of depression following myocardial infarction with mortality and cardiovascular events: individual patient

- data meta-analysis. *British Journal of Psychiatry*, **203**(2), 90–102.
- Meijer R.R., Tendeiro J., Wanders, R. (2014). The use of nonparametric item response theory to explore data quality. In Reise S.P., Revicki D. (eds) *Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment*, pp. 85–110, London, Routledge.
- Moran P.J., Mohr D.C. (2005) The validity of Beck Depression Inventory and Hamilton Rating Scale for Depression Items in the assessment of depression among patients with multiple sclerosis. *Journal of Behavioral Medicine*, **28**(1), 35–41.
- Muthén L.K., Muthén B.O. (1998–2007) *Mplus User's Guide*, 5th edition, Los Angeles, CA, Muthén & Muthén.
- Ormel J., de Jonge P. (2011) Unipolar depression and the progression of coronary artery disease: toward an integrative model. *Psychotherapy and Psychosomatics*, **80**(5), 264–274.
- Ramsay J.O. (1991) Kernel smoothing approaches to nonparametric item characteristic curve estimation. *Psychometrika*, **56**(4), 611–630.
- Reise S.P., Waller N.G. (2003) How many IRT parameters does it take to model psychopathology items? *Psychological Methods*, **8**(2), 164–184.
- Reise S.P., Scheines R., Widaman K.F., Haviland M.G. (2013) Multidimensionality and structural coefficient bias in structural equation modeling. A bifactor perspective. *Educational and Psychological Measurement*, **73**(1), 5–26.
- Rizopoulos D. (2006) ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software*, **17**(5), 1–25.
- Rush A.J., Gullion C.M., Basco M.R., Jarrett R.B., Trivedi M.H. (1996) The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychological Medicine*, **26**(3), 477–486.
- Samejima F. (1969) Estimation of Latent Trait Ability using a Response Pattern of Graded Scores (*Psychometric Monograph No. 17*), Richmond, VA, Psychometric Society.
- Seo D.G., Weiss D.J. (2013) lz Person-fit index to identify misfit students with achievement test data. *Educational and Psychological Measurement*, **73**(7), 994–1016.
- Sijsma K., Molenaar I. (2002) *Introduction to Non-parametric Item Response Theory*, Thousand Oaks, CA, Sage.
- Sorensen C., Friis-Hasché E., Haghfelt T., Bech P. (2005) Postmyocardial infarction mortality in relation to depression: a systematic critical review. *Psychotherapy and Psychosomatics*, **74**(2), 69–80.
- Spijkerman T.A., van den Brink R.H., May J.F., Winter J.B., van Melle J.P., de Jonge P., Crijns H.J., Ormel J. (2006) Decreased impact of post-myocardial infarction depression on cardiac prognosis? *Journal of Psychosomatic Research*, **61**(4), 493–499.
- Thombs B.D., Bass E.B., Ford D.E., Stewart K.J., Tsilidis K.K., Patel U., Fauerbach J.A., Bush D.E., Ziegelstein R.C. (2006) Prevalence of depression in survivors of acute myocardial infarction. *Journal of General Internal Medicine*, **21**(1), 30–38.
- Thombs B.D., Ziegelstein R.C., Beck C.A., Pilote L. A (2008) General factor model for the Beck Depression Inventory-II: validation in a sample of patients hospitalized with acute myocardial infarction. *Journal of Psychosomatic Research*, **65**(2), 115–121.
- Thombs B.D., Ziegelstein R.C., Pilote L., Dozois D. J., Beck A.T., Dobson K.S., Fuss S., de Jonge P., Grace S.L., Stewart D.E., Ormel J., Abbey S.E. (2010) Somatic symptom overlap in Beck Depression Inventory-II scores following myocardial infarction. *British Journal of Psychiatry*, **197**(6), 61–66.
- Troyanskaya O., Cantor M., Sherlock G., Brown P., Hastie T., Tibshirani R., Botstein D., Altman R. B. (2001) Missing value estimation methods for DNA microarrays. *Bioinformatics*, **17**(6), 520–525.
- van den Brink R.H., van Melle J.P., Honig A., Schene A.H., Crijns H.J., Lambert F.P., Ormel J. (2002) Treatment of depression after myocardial infarction and the effects on cardiac prognosis and quality of life: rationale and outline of the Myocardial Infarction and Depression-Intervention Trial (MIND-IT). *American Heart Journal*, **144**(2), 219–225.
- van der Kooy K., van Hout H., Marwijk H., Marten H., Stehouwer C., Beekman A. (2007) Depression and the risk for cardiovascular diseases: systematic review and meta-analysis. *International Journal of Geriatric Psychiatry*, **22**(7), 613–626.
- van Melle J.P., de Jonge P., Ormel J., Crijns H.J., van Veldhuisen D.J., Honig A., Schene A.H., van den Berg M.P., MIND-IT investigators. (2005) Relationship between left ventricular dysfunction and depression following myocardial infarction: data from the MIND-IT. *European Heart Journal*, **26**(24), 2650–2656.
- van Melle J.P., de Jonge P., Honig A., Schene A.H., Kuyper A.M., Crijns H.J., Schins A., Tulner D., van den Berg M.P., Ormel J. (2007) Effects of antidepressant treatment following myocardial infarction. *British Journal of Psychiatry*, **190**(6), 460–466.
- Wanders R.B., Wardenaar K.J., Kessler R.C., Penninx B.W., Meijer R.R., de Jonge P. (2015) Differential reporting of depressive symptoms across distinct clinical subpopulations: what difference does it make? *Journal of Psychosomatic Research*, **78**(2), 130–136.
- World Health Organization (WHO). (1990) *Composite International Diagnostic Interview (CIDI)*, Geneva, WHO.
- World Health Organization (WHO). (1993) *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*, Geneva, WHO.
- World Health Organization (WHO). (1997) *Composite International Diagnostic Interview (CIDI, Version 2.1. World Health Organization)*, Geneva, WHO.
- Zuidersma M., Ormel J., Conradi H.J., de Jonge P. (2012) An increase in depressive symptoms after myocardial infarction predicts new cardiac events irrespective of depressive symptoms before myocardial infarction. *Psychological Medicine*, **42**(4), 683–693.
- Zuidersma M., Conradi H.J., van Melle J.P., Ormel J., de Jonge P. (2013). Self-reported depressive symptoms, diagnosed clinical depression and cardiac morbidity and mortality after myocardial infarction. *International Journal of Cardiology*, **167**(6), 2775–2780.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.