Journal of Affective Disorders 189 (2016) 1–9

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research report

Predicting long-term depression outcome using a three-mode principal component model for depression heterogeneity

Rei Monden [a,*], Alwin Stegeman [b], Henk Jan Conradi [c], Peter de Jonge [a], Klaas J. Wardenaar [a]

[a] University of Groningen, University Medical Center Groningen, Interdisciplinary Center Psychopathology and Emotion regulation (ICPE), Department of Psychiatry, (CC-72), PO Box 30.001, 9700 Groningen, The Netherlands
[b] University of Groningen, Heijmans Institute of Psychological Research, Groningen, The Netherlands
[c] University of Amsterdam, Department of Clinical Psychology, Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 1 July 2015
Received in revised form
25 August 2015
Accepted 9 September 2015
Available online 12 September 2015

Keywords:
Major depressive disorder
Prognosis
Course
Three-mode Principal Component Analysis (3MPCA)
Beck Depression Inventory (BDI)

ABSTRACT

Background: Depression heterogeneity has hampered development of adequate prognostic models. Therefore, more homogeneous clinical entities (e.g. dimensions, subtypes) have been developed, but their differentiating potential is limited because neither captures all relevant variation across persons, symptoms and time. To address this, three-mode Principal Component Analysis (3MPCA) was previously applied to capture person-, symptom- and time-level variation in a single model (Monden et al., 2015). This study evaluated the added prognostic value of such an integrated model for longer-term depression outcomes.
Methods: The Beck Depression Inventory (BDI) was administered quarterly for two years to major depressive disorder outpatients participating in a randomized controlled trial. A previously developed 3MPCA model decomposed the data into 2 symptom-components ('somatic-affective', 'cognitive'), 2 time-components ('recovering', 'persisting') and 3 person-components ('severe non-persisting depression', 'somatic depression' and 'cognitive depression'). The predictive value of the 3MPCA model for BDI scores at 3-year ($n=136$) and 11-year follow-up ($n=145$) was compared with traditional latent variable models and traditional prognostic factors (e.g. baseline BDI component scores, personality).
Results: 3MPCA components predicted 41% and 36% of the BDI variance at 3- and 11-year follow-up, respectively. A latent class model, growth mixture model and other known prognostic variables predicted 4–32% and 3–24% of the BDI variance at 3- and 11-year follow-up, respectively.
Limitations: Only primary care patients were included. There was no independent validation sample.
Conclusion: Accounting for depression heterogeneity at the person-, symptom- and time-level improves longer-term predictions of depression severity, underlining the potential of this approach for developing better prognostic models.

1. Introduction

Although Major Depressive Disorder (MDD) is generally characterized by an episodic course (American Psychiatric Association, 2013), patients show considerable variation in their course (Kessler et al., 2005). Given the impact of depression on patients' lives (Alonso et al., 2004) and society (Kessler, 2012), predicting MDD patients' longer-term outcomes is of strong interest. Unfortunately, adequate prediction of depression outcomes in clinical practice has proven difficult. Prognostic research has identified several factors that are predictive of an unfavorable course of MDD, including alcohol use (Mueller et al., 1994), somatic problems (Huijbregts et al., 2010), high severity, long episode duration (Penninx et al., 2011), young age at onset (Karlsson et al., 2008), high neuroticism (Rhebergen et al., 2011), comorbidity (Patten et al., 2010) and increases on particular symptom dimensions (Wardenaar et al., 2012). However, these insights have not yet resulted in development of sufficiently accurate prediction models.

One reason for the current lack of specific prognostic models is the fact that depression is very heterogeneous. Depression symptomatology is broad and includes a range of affective, cognitive and somatic symptoms (e.g. Van Loo et al. (2012)). Consequently, patients with the same MDD diagnosis can have many different symptom patterns and course-trajectories (e.g. Goldberg (2011); Widiger and Clark (2000); Olbert et al. (2014)). Fried and Nesse (2015), for instance, observed 1030 unique symptom profiles in a sample of 3703 depressed patients, with the most

* Correspondence to: ICPE (CC-72), PO Box 30.001, 9700 RB Groningen. Fax: +31 50 36 19772.
E-mail address: r.tendeiro-monden@umcg.nl (R. Monden).

http://dx.doi.org/10.1016/j.jad.2015.09.018
0165-0327/

common profile occurring in only 1.8% of the patients. This diversity can be made more accessible for formal analysis by postulating heterogeneity within each of three *modes* of the depression construct: a '*symptom-*', '*person-*' and '*time-*' mode (Wardenaar and de Jonge, 2013; Monden et al., 2015). Within the symptom-mode, more homogeneous, different subdomains of depressive symptomatology can exist (e.g. van Loo et al. (2012), Shafer et al., 2006). Within the person-mode increasingly detailed subgroups, characterized by specific symptom-patterns can be discerned (e.g. Olbert et al. (2014); Fried and Nesse (2015)). Within the time-mode, many quantitatively (e.g. different baseline offset) and qualitatively (e.g. different course shapes) different course-trajectories can be discerned (e.g. Rhebergen et al. (2012), Wardenaar et al. (2014, 2015)). Different approaches have been used to identify the more homogeneous entities within each of these modes.

Data-driven studies using latent variable techniques, such as factor analysis (FA), latent class analysis (LCA) latent class growth analysis/growth mixture modeling (LCGA/GMM), and principal component analysis (PCA) have shown that relatively homogeneous symptom dimensions/classes can be identified, which improve differentiation between those with different prognoses. Studies using PCA, FA or related techniques, showed that different symptom-factors were associated with different long-term depression outcomes (e.g. Joiner and Lonigan (2000), Wardenaar et al. (2012)). Studies that used LCA to identify more homogeneous classes of patients, showed that these were associated with different long-term outcomes (e.g. Sullivan et al. (1998), Lamers et al. (2012)). Studies that used LCGA or GMM to model classes with different course-trajectories showed that class-membership (e.g. chronic vs. quick remission) was associated with depression outcomes (e.g. Wardenaar et al. (2014, 2015)).

Although the above described research has provided valuable insights into the heterogeneity of depression and its role in depression prognosis, each of the used techniques (PCA, FA, LCA, LCGA) only allows for a partial explanation of all depression heterogeneity. This is due to the fact that each latent variable method assumes homogeneity within at least one mode of the depression data (Wardenaar and de Jonge, 2013; Monden et al., 2015). For example, PCA is a data-reduction technique to decompose scores on many variables into scores on a smaller number of components and FA describes variance shared among variables with one or more latent continuous variables (factors). When conducting PCA or FA, the resulting solution describes symptom heterogeneity, but no variation across persons. Conversely, LCA/LCGA/GMM models are based on the assumption that all heterogeneity across persons is captured by discrete class-membership, and that there is no residual symptom (co)variance within the classes (local independence), which is not in line with current dimensional views of psychopathology. Furthermore, PCA, FA and LCA are cross-sectional techniques that do not incorporate the time variations that are an important part of the clinical presentation of depression. Contrarily, LCGA and GMM describe inter-personal variations in course-trajectories, but do not take into account cross-sectional symptom-heterogeneity. Taken together, none of the traditionally used latent variable techniques capture all sources of inter-personal variation in a single model: neither captures variation across *persons* in how they vary in their *change over time* on *different symptom domains*. An integrated description of depression heterogeneity could provide more insight into these inter-personal variations, and tools to more specifically differentiate between patients.

To capture the three main sources of depression heterogeneity in a single model alternative statistical models are needed. When represented in a 'three-dimensional array' (or 'data cube'; Cattell (1966)) of various symptoms (symptom-mode) in a number of persons (person-mode) at different time points (time-mode), the heterogeneity of this multimodal data can be analyzed with Three-mode Principal Component Analysis (3MPCA; Kroonenberg and De Leeuw (1980), Tucker (1963, 1966), Kiers (2000)). 3MPCA is a multiway version of PCA to decompose three-dimensional data objects into a number of components. In the case of depression, 3MPCA can be used to summarize the heterogeneity of depression with a limited number of person-, symptom- and time-mode components, while accounting for the interactions between the different modes' components (Kroonenberg, 2008).

A previous application of 3MPCA in a sample of primary care depression patients, who were followed for two years (Monden et al., 2015) showed that the longitudinal depression data could be decomposed into two symptom-mode components ('cognitive' and 'somatic-affective'), two time-mode components ('improving' and 'persisting') and three person-mode components ('severe non-persisting depression', 'somatic depression' and 'cognitive depression'), providing an integrated and insightful description of the depression construct.

The aim of the present study was to evaluate if this 3MPCA model of depression heterogeneity showed added prognostic value compared to traditional cross-sectional prognostic factors (e.g. depression severity, personality), longitudinal prognostic factors (BDI change over time) and LCA and GMM class-solutions. As the 3MPCA model contained information about inter-personal variations in both depressive course and symptomatology, it was hypothesized to have superior prognostic value.

2. Methods

2.1. Participants and procedures

The data came from a randomized controlled trial to evaluate the efficacy of different combinations of treatment in primary care MDD patients, who were recruited from general practices. Detailed information on the inclusion and data collection procedure can be found elsewhere (Smit et al., 2005, 2006; Conradi et al., 2007, 2008) and is summarized below. Previous analyses showed no differences between the treatment groups in terms of remission on the BDI (Conradi et al., 2007).

Three-hundred-ninety-seven patients were referred by 49 GPs in the North of the Netherlands. Inclusion criteria were: having a history of a depressive episode, having no current life-threatening somatic disease, and receiving no current psychotherapy. Exclusion criteria were: presence of dementia, a bipolar/psychotic disorder, a primary diagnosis of substance abuse. These were confirmed by the Composite International Diagnostic Interview (CIDI: WHO (1997), Ter Smitten et al. (1998)). Of the initially referred 397 patients, 52 met exclusion criteria and 78 declined participation, resulting in a sample of 267 patients (67.3%). These patients were invited again to participate in the 3- and 11-year follow-up assessments. After 3-year follow-up, patients were free to use any necessary care. The study protocol was approved by the medical ethical committee of the University Medical Center Groningen. All participants signed informed consent.

For the 3MPCA analysis, patients were included if they provided BDI scores on at least 5 of 9 measurement-points (baseline, 3-, 6-, 9-, 12-, 15-, 18-, 21-, and 24-month) during the 2-year follow-up period. The resulting sample consisted of 219 patients (82.0%; Monden et al. (2015)). For the current analyses, only those with a 3- and/or 11-year follow-up assessment were included. Of the 267 patients, 141 (53%) provided 3-year follow-up data and 164 (61.4%) provided 11-year follow-up data. For 3-year follow-up analyses, 5 patients were excluded and for 11-year follow-up, 17 patients were excluded from prognostic analyses because they

either had missing BDI items at follow-up or they were not included in the 3MPCA model. The eventual samples included 136 patients with a 3-year follow-up and 145 patients with an 11-year follow-up.

2.2. Measures

2.2.1. Beck depression inventory

The BDI (Beck et al., 1961) is a 21-item self-report questionnaire, which was administered at baseline and at 3-, 6-, 9-, 12-, 15-, 18-, and 24-month follow-up. In addition, the BDI was administered at 3- and 11-year follow-up.

2.2.2. Other measures

Socio-demographic characteristics (i.e. age, gender, income, education level and working status) were assessed at baseline. In addition, the Symptoms Checklist-90 (SCL-90: Derogatis et al. (1973)), the Neuroticism-Extraversion-Openness-Five-Factor Inventory (NEO-FFI: Costa and McCrae (1989)) and the Medical Outcomes Study 36-item Short Form (MOS-SF-36: Ware and Sherbourne (1992)) were administered at baseline. At 11-year follow-up, medication use between 3- and 11-year follow-up (yes/no) was documented retrospectively.

2.3. Statistical analyses

2.3.1. Data imputation

Of all the BDI item-responses collected during the first two years, 7.8% was missing. These missing values were imputed 20 times (see Monden et al. (2015) for the full procedure) with the R-package 'Amelia ll' (Honaker and Blackwell, 2011). For the 3- and 11-year BDI measurements, imputation was not undertaken because these scores were used as primary outcomes.

2.3.2. Three-mode Principal Component Analysis (3MPCA)

3MPCA was previously applied to the complete 2-year data ($n=219$; Monden et al. (2015)) and decomposed the data into two symptom-mode components ('cognitive' and 'somatic-affective'), two time-mode components ('improving' and 'persisting') and three person-mode components ('severe non-persisting depression', 'somatic depression' and 'cognitive depression').

Because the number of excluded patients due to missing 3- and 11-year follow-up was considerable, the 3MPCA model could be different between the complete sample and the samples with 3- or 11-year follow-up. In that case, the predictive value of the 3MPCA could be affected not because of the model itself, but because of the change of the sample characteristics. Therefore, a 3MPCA model was also fitted in the subsamples ($n=136$ and $n=145$) and the component structures were compared with those of the complete data to evaluate the consistency of the models across the (sub)samples. If the 3MPCA models proved stable across (sub) samples, all prognostic analyses were conducted using the 3MPCA model from the complete sample ($n=219$). If the 3MPCA model-parameters were different across (sub)samples, prognostic analyses were conducted with subsample-specific 3MPCA model-components.

The application of the 3MPCA consisted of the following five steps (details in: Monden et al. (2015)): (1) a fixed-effects three-way analysis of variance (ANOVA) was applied in each of the 20 imputed datasets after subtraction of the grand mean, to evaluate if a three-way interaction underlies the dataset (Kiers and Van Mechelen, 2001). (2) The generalized scree test (Kiers and Der Kinderen, 2003; Timmerman and Kiers, 2000) was used to select the number of components for each mode. (3) The stability of the solution was evaluated by inspection of the 3MPCA solutions' variation across the 20 imputed datasets and by using split-half procedures within each imputed dataset. (4) To get an interpretable 3MPCA solution, orthogonal Joint Orthomax rotation was used to obtain simple component structures for symptom-, time-mode, and their interactions were obtained. This rotation was executed with 'standard weights' but no weight on the person-mode (Kiers, 1998). (5) The average of the obtained 20 estimated solutions was calculated by a generalized Procrustes rotation (Kroonenberg, 2008; Kroonenberg and van Ginkel, 2012; Ten Berge, 1977). These analyses were conducted with the Tucker3.m program for Matlab (Kiers, 2000).

The symptom-components ('cognitive' and 'somatic-affective') were interpreted by inspecting loadings of the symptoms on each component. The time-components were interpreted by inspecting loadings of the 9 measurement points on the time-components. The first three (baseline, 3- and 6-months) loaded high on the first ('improving') component and the 9- to 24-month follow-ups loaded high on the second ('persisting') component. The person-mode components were interpreted by inspecting the interactions between symptom- and time-mode components for each person-mode component. For instance, scores on one person-component were associated with an interaction consisting of persisting somatic affective symptoms and decreasing cognitive symptoms, and was therefore interpreted as a 'somatic depression' component. A person's score on this component provides a continuous measure of the degree to which this phenotype applies to him/her. In contrast, when conducting a regular PCA on a cross-sectional assessment of depressive symptoms, the patients' scores on the resulting components would only provide information about baseline symptom-levels. Previous work also showed that the person-mode components were correlated with the SCL-90, NEO-FFI and MOS-SF-36, which was of additional help in interpreting each component's coverage. The 'severe non-persisting depression' person-mode component was associated with psychopathology ($r=0.60$) and negatively with quality of life ($r=-0.50$), the 'somatic depression' person-mode component was negatively correlated with physical functioning ($r=-0.45$), and a 'cognitive depression' person-mode component was positively correlated with neuroticism ($r=0.38$) and negatively with self-esteem ($r=-0.47$).

2.3.3. 3MPCA and missing outcomes

First, it was investigated whether drop-out at 3- or 11-year follow-up was associated with the 3MPCA person-mode components by conducting a multinomial logistic regression analysis using the total sample 3MPCA person-component scores as predictors and using either the presence of a 3-year follow-up or a 11-year follow-up BDI (1: absent/2:present) as outcome.

2.3.4. 3MPCA outcome prediction

To investigate the prognostic value of the 3MPCA, multivariate linear regression analyses were conducted, either using only the information from the person-mode components or the information from the whole 3MPCA model for outcome prediction. When the subjects' person-mode component scores were used as independent variables, one intercept and regression coefficients for each of the three person-mode components were estimated. To test the associations of the whole 3MPCA model with the outcomes, an intercept and coefficients for the two time-mode components were estimated.

2.3.5. Other known outcome predictors

Different sets of predictors were investigated and compared with the 3MPCA predictions. These sets were: (1) Latent trajectory classes from GMM applied to the BDI sum scores across the two-year period, (2) Latent classes from LCA applied to the baseline BDI, (3) Component scores from a traditional, cross-sectional PCA

of the baseline BDI, (4) The MOS-SF-36 scales at baseline, (5) The SCL-90 scales at baseline, (6) The NEO-FFI scales at baseline, (7) all independent variables in (3) to (7), (8) BDI item score differences between baseline and 24-month follow-up, and (9) BDI sum scores differences between baseline and 24-month follow-up.

To identify the optimal LCA-model describing the baseline cross-sectional heterogeneity in symptom-reporting and GMM-model describing heterogeneity in longitudinal course-trajectories, LCAs and GMMs were run in each imputed dataset. For LCA, the imputed item-scores were rounded to their closest discrete value (0, 1, 2 or 3) and a robust maximum likelihood estimation (MLR) was used to estimate models with increasing numbers of classes. The best-fitting model was identified by comparing the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) across models, with the lowest BIC/AIC indicating the best fit. After identification of the best-fitting model, patients' posterior class-probabilities for each class were averaged across imputed datasets and used as predictors. For GMM, the BDI sum scores from baseline to 24-month follow-up were calculated in each imputed dataset. GMMs were run in each dataset with freely estimated variances for the class-specific intercept and with variances of the slopes set to zero. Identification of the optimal model and class-allocation was done in the same way as the LCAs. Both models were run with Mplus (version 5) using multiple random starts to prevent identification of models at local maxima.

2.3.6. Outcome prediction analyses

For both 3- and 11-year follow-up, BDI sum scores were first used as outcomes. Second, sum scores on the two BDI symptom-domains that were identified with the 3MPCA (i.e., 'cognitive' and 'somatic-affective' domains; Monden et al. (2015)) were used as outcomes to investigate the domain-specific predictive ability of the model. The cognitive domain-score was calculated by summing the BDI item-scores on 'guilty feelings', 'past failure', 'self-criticism', 'self-dislike', 'body image', 'feeling punished', 'suicidal thoughts' and 'sadness'. The somatic-affective domain-score was calculated by summing the BDI item-scores on 'work difficulties', 'tiredness', 'loss of pleasure', 'indecisiveness', 'loss of interest in sex', 'loss of interest', 'agitation', 'changes in sleeping' and 'crying' (see Appendix 1).

To evaluate predictive value, both adjusted R^2 and residual plots were inspected. An adjusted R^2 indicates how well the model fits to the data. In addition, prediction precision was evaluated by inspection of residual plots, which provided insight in the congruence between predicted and observed values, the potential role of outliers and possible over- or underestimations. When the assumptions of multivariate linear regression were violated after transformation, robust regression with a bisquare weighted function was performed to evaluate the influence of these violations on estimated values.

3. Results

3.1. Descriptive information

Table 1 summarizes the descriptive information for the used samples. The 3-year follow-up group had a lower proportion of females and mean MOS-SF-36 social function scale score than the other samples. The 11-year follow-up group had a higher mean MOS-SF-36 pain scale score. There were no other differences. Mean baseline BDI sum scores (19.2–19.4 across samples) indicated moderate depression severity (BDI $\geq$ 19; Beck et al. (1988)).

3.2. Three-Mode Principal Component Analysis

The results of 3MPCA in the complete data (n=219), in the subgroup with a 3-year follow-up (n=136), and in the subgroup with an 11-year follow-up (n=145) are shown in Appendix 1. Because high congruence ($\geq$0.97) was observed between the 3MPCA model in the original sample and 3MPCA models fitted in the subsamples, the component scores from the originally fitted 3MPCA solution were used in all the 3-year and 11-year follow-up analyses. This was done to keep in line with previous work and to facilitate comparability across the subsample-specific results. Missing either 3-year or 11-year follow-up data was not associated with any of the person-mode components. This indicated that

Table 1
Baseline characteristics of the study groups.

Baseline variable	Complete sample	Sample with Complete 3-year follow-up	Sample with complete11-year follow-up
N	219	136	145
Median follow-up period in months (IQR)	–	37.9 (37.4–38.3)	140.4 (129.1–151.7)
Female (%)	144 (65.8)	79 (58.1)	96 (66.2)
Baseline age: mean years (SD)	43.3 (11.1)	43 (10.8)	42.4 (10.6)
Baseline age: range	17–69	17–69	21–64
Mean BDI sum score (SD)	19.4 (9.1)	19.7 (8.9)	19.2 (9.0)
Psychiatric characteristics (SCL-90)			
Mean sum score of depression scale (SD)	42.5 (12.5)	43.3 (12.8)	42.8 (12.9)
Mean sum score of anxiety scale (SD)	21.8 (7.8)	21.9 (7.5)	21.5 (7.7)
Mean sum score of psycho neuroticism scale (SD)	195 (54.5)	194 (52.6)	194 (54.7)
Personality traits (NEO-FFI)			
Mean score of neuroticism scale (SD)	42.3 (6.5)	42.1 (6.4)	42.3 (6.8)
Mean score of extraversion scale (SD)	32.7 (6.8)	32.9 (6.7)	32.2 (7.3)
Quality of life (MOS-SF-36)			
Mean score of social functions scale (SD)	45.9 (21.5)	43.2 (21.8)	45.7 (22.1)
Mean score of mental health scale (SD)	40.5 (16.5)	39.7 (17.4)	40.6 (16.5)
Mean score of pain scale (SD)	65.7 (26.5)	66.9 (25.9)	69.2 (26.3)

SD=standard deviation, BDI=Beck Depression Inventory, SCL-90=Symptom Checklist-90, NEO=Neuroticism-Extraversion-Openness-Five-Factor Inventory, MOS-SF-36=Medical Outcomes Study 36-item Short Form; IQR=Interquartile range; patients with missing data on the baseline variables were excluded from the calculations (between 1 and 26).

missing follow-up data was not associated with the 3MPCA component scores (Appendix 2).

3.3. Prediction of follow-up BDI sum scores

In the prediction models, the normality assumption of multivariate regression analysis was violated with right-skewed BDI item-data. Because transformation did not solve this problem, robust regression analysis was performed alongside linear regression (see Appendix 3). However, estimated model coefficients and R^2–values were comparable between robust and regular techniques. Therefore, regular regression results are presented below.

Comparison of BIC-values across LCA models with increasing classes (see Appendix 4) showed a 2-class LCA model to fit best to the data in all 20 imputed datasets. The model had one class showing low scores on the BDI-items and another class showing relatively higher scores. For the GMM, a 2-class model was also found to fit the data most consistently across the imputed datasets (adding a third class either led to an increase, or a minimal decrease of the BIC; see Appendix 4). The GMM had one class characterized by persistently high BDI scores over time and one class characterized by decreasing BDI scores over time. A regular PCA was also run on the baseline BDI-data and a 2-component model was selected based on a Scree-plot.

The results of multivariate linear regression analyses to predict the BDI sum scores at 3- and 11-year follow-up are presented in Table 2. Several interesting observations were made in these results. First, the person-mode components showed the highest explained variance in follow-up BDI scores of all tested predictors, followed by the GMM solution. Second, the R^2 of the 3MPCA model was more than two times higher than that of the baseline PCA model. Third, when the time aspect was incorporated in the traditional prediction models by using GMM or BDI item- or sum-score differences between baseline and 2-year follow-up, this still yielded R^2-values (range: 0.04–0.32) that were lower than those for the 3MPCA model. Fourth, the predictive value of traditional baseline predictors (MOS-SF-36, SCL-90 and NEO-FFI) were found to be limited (maximum $R^2=0.10$). Finally, 3MPCA predictions of 3- and 11-year follow-up severity were comparable, which was not the case for the GMM. In addition, predictions were stable across the 20 imputed datasets (standard deviations for R^2 across datasets were all ≤ 0.01, both when using the 3- and 11-year BDI score as outcome). Additional multivariate analyses including both the 3MPCA model and other predictors, showed that 82.3–83.0% of the 3-year follow-up R^2 and 73.7–75.8% of the 11-year follow-up R^2 was uniquely explained by the 3MPCA model.

The associations of the person-mode components with 3- and 11-year follow-up (Table 3) indicated that the 'somatic depression' person-component was significantly associated with BDI sum scores at 3- and 11-year follow-up, but that the 'cognitive depression' person-component was only associated with BDI scores at 11-year follow-up. This implies that the 'somatic depression' person-component is the most important predictor for both 3- and 11-year BDI sum scores and that the 'cognitive depression' person-mode component is associated with long-term depression outcome. The 'severe non-persisting depression' component was not associated with any follow-up score, indicating that this component is not related to chronicity or relapse of depression in the long run. The associations of the complete 3MPCA model with 3- and 11-year follow-up (Table 4) indicated that only the 'persisting' time component was associated with follow-up scores.

Additional analyses adjusting the predictions of the 11-year BDI for medication-use between 3- and 11-year follow-up showed no change in the R^2 statistics for the 3MPCA model.

Table 2
Explained variance of different predictors in the Beck Depression Inventory sum scores or two symptom domains' scores at 3- and 11-year follow-up.

Prediction of BDI sum score	**Adjusted R^2**	
	3 Years	**11 Years**
Independent variables		
Person-mode component	0.41	0.36
3MPCA solution*	0.41	0.35
GMM with 2 classes	0.32	0.24
LCA with 2 classes	−0.01	−0.01
Baseline PCA with 2 components	0.15	0.13
SCL-90 scores at baseline	0.04	0.07
NEO-FFI scores at baseline	0.10	0.03
MOS-SF-36 scores at baseline	0.09	0.07
PCA Comp2, SCL, NEO, MOS	0.24	0.09
Baseline – 2 years: BDI item scores	0.22	0.14
Baseline – 2 years: BDI sum scores	0.04	0.04
Prediction of cognitive domain-scores		
Independent variables		
Person-mode component	0.31	0.37
3MPCA solution*	0.26	0.23
GMM with 2 classes	0.20	0.12
LCA with 2 classes	0.00	0.00
Baseline PCA with 2 components	0.19	0.11
SCL-90 scores at baseline	0.10	0.12
NEO-FFI scores at baseline	0.16	0.05
MOS-SF-36 scores at baseline	0.06	0.05
PCA Comp2, SCL, NEO, MOS	0.24	0.13
Baseline – 2 years: BDI sum scores	0.00	0.02
Prediction of somatic-affective domain-scores		
Independent variables		
Person-mode component	0.47	0.31
3MPCA solution*	0.36	0.32
GMM with 2 classes	0.27	0.24
LCA with 2 classes	0.00	0.00
Baseline PCA with 2 components	0.08	0.09
SCL-90 scores at baseline	0.05	0.03
NEO-FFI scores at baseline	0.03	0.03
MOS-SF-36 scores at baseline	0.06	0.06
PCA Comp2, SCL, NEO, MOS	0.15	0.06
Baseline – 2 years: BDI sum scores	0.07	0.04
LCA with 2 classes	0.00	0.00

* Person-, symptom-, time-component and core array.

3.4. Prediction of specific symptom-domains

Predictions of cognitive or somatic-affective symptom scores at 3- and 11-year follow-up are shown in Table 2. The 3MPCA components showed the highest R^2-statistics for both outcomes. Interestingly, baseline PCA together with baseline SCL-90, NEO-FFI and MOS-SF-36 showed the third-best predictive value for 'cognitive' domain-scores, while the GMM showed the third-best predictive value for 'somatic-affective' domain-scores. Comparison of the R^2-values between the two outcome domains showed that the 3MPCA components explained more variance in the somatic-affective domain than in the cognitive domain.

The estimated associations of the person-mode components (Table 3) showed that the 'somatic depression' person-component was associated most strongly with the somatic-affective BDI domain score at both follow-ups, whereas the 'cognitive depression' person component showed the strongest associations with the cognitive BDI domain score at 3- and 11-year follow-up. Only the 'persisting' time-component was associated with the domain-scores at 3- and 11-year follow-up (see Table 4). This indicated that two of the three person-mode components had symptom-specific predictive value. In addition, the findings indicated that persistence of any symptomatology was predictive of all symptom-

Table 3
Associations of the three person-mode components with the Beck Depression Inventory (BDI) total score, cognitive domain score and somatic-affective domain score at 3- and 11-year follow-up.

Predictors		Outcomes					
		BDI at 3-year follow-up			BDI at 11-year follow-up		
		Sum score Coef. (*p*-val)	Cognitive Coef. (*p*-val)	Somatic-affective Coef. (*p*-val)	Sum score Coef. (*p*-val)	Cognitive Coef. (*p*-val)	Somatic-affective Coef. (*p*-val)
Person-mode component	Severenon-persisting depression	2.15(0.75)	3.06(0.28)	0.50(0.90)	−4.24(0.61)	−1.35(0.68)	−2.47(0.61)
	SomaticDepression	69.33(< 0.001)	16.42(< 0.001)	45.92(< 0.001)	59.78(< 0.001)	14.89(< 0.001)	37.77(< 0.001)
	Cognitivedepression	12.22(0.07)	45.92(< 0.001)	−7.04(0.07)	36.88(< 0.001)	23.01(< 0.001)	9.48(0.03)

domains in the long run.

Visual inspection of the residual plots (Fig. 1) indicated that predictions of 3-year follow-up scores were more accurate than predictions of 11-year follow-up scores. In addition, the cognitive domain scores were predicted more accurately (smaller residuals) than scores on the somatic-affective domain at both follow-ups.

4. Discussion

This study investigated the predictive value of a 3MPCA model of depression for long-term depression-outcome and compared its predictive performance to traditionally used prognostic factors. A model containing only the person-mode component scores of the previously identified 3MPCA model explained most variance in BDI sum scores and domain-specific scores at both follow-ups. Interestingly, more traditional latent variable models and prognostic factors (e.g. LCA, GMM, baseline PCA, NEO-FFI, $\Delta_{\text{Baseline BDI-24month BDI}}$) were much less predictive. The somatic-affective BDI domain at follow-up was most strongly associated with the 'somatic-affective depression' person-component and the BDI cognitive domain showed the strongest associations with the 'cognitive depression' person-mode component. Interestingly, the 'severe non-persisting' component scores showed no associations with long-term BDI scores. A model containing the whole 3MPCA model attained the second highest explained variance in the follow-up BDI scores. As expected, the model coefficients showed that BDI follow-up scores were associated with the 'persisting' time-component of the 3MPCA model. Interestingly, both in the model with person-mode components and the model with the complete 3MPCA model, somatic-affective domain-scores were predicted with higher explained variance. However, plotting residual scores revealed that cognitive domain-scores were predicted with higher accuracy than somatic-affective domain-scores. These results may seem contradictory, but R^2 is defined by a correlation between observed and estimated scores, and therefore sensitive to the effects of outliers. As a result, high R^2 does not necessarily reflect higher accuracy, which seems to be the case here. Taken together, these findings supported the hypothesis that 3MPCA components are better predictors of long-term depression outcomes than a range of traditional predictors and latent variable models.

The presented results provide a proof-of-principle for the use of 3MPCA in prognostic research. The results illustrate that integrating three important sources of depression heterogeneity yields a comprehensive set of person-component scores with good predictive value. Compared to known predictors, 3MPCA explained much more variance in the depression outcomes. This could be due to the fact that, unlike traditional prognostic factors, person-components describe two important aspects of patients' clinical picture at a time: they capture inter-personal differences in symptomatology (cognitive vs. somatic), which have been found to be associated with long-term depression prognosis (e.g. Patten et al. (2010), Riihimäki et al. (2011), Wardenaar et al. (2012)) and they reflect the dynamic of symptomatology over time, which has previously been shown to be predictive of long-term depression severity (e.g. Rhebergen et al. (2011), Wardenaar et al. (2014, 2015)). The importance of including this temporal aspect was further exemplified by the fact that the components of a regular baseline PCA or the classes of a baseline LCA explained much less variance. These findings emphasize the importance of accounting for time-heterogeneity when trying to predict depression outcomes. The other way around, the importance to account for symptom heterogeneity was illustrated by the observation that temporal difference-scores on the BDI scale and the GMM, which do not account for symptom-mode heterogeneity, explained less variance than the person-mode components.

The 'severe non-persisting' person-component was previously found to be related to severe baseline psychopathology, low quality of life and low self-esteem (Monden et al., 2015), but was hardly associated with long-term depression scores in the current study. This may due to the fact that this component reflected mainly baseline severity and was associated with a large severity decrease during the 'recovering' time-phase, whereas the current analyses showed the importance of the 'persisting' time-component for the prediction of 3- and 11-year BDI scores. These findings fit in with the observation that other baseline severity indicators (e.g. SCL-90, NEO-FFI, MOS-SF36) also showed limited predictive

Table 4
Associations of the three 3MPCA model with the Beck Depression Inventory (BDI) total score, cognitive domain score and somatic-affective domain score at 3- and 11-year follow-up.

Predictors		Outcomes					
		BDI at 3-year follow-up			BDI at 11-year follow-up		
		Sum score Coef. (*p*-val)	Cognitive Coef. (*p*-val)	Somatic-affective Coef. (*p*-val)	Sum score Coef. (*p*-val)	Cognitive Coef. (*p*-val)	Somatic-affective Coef. (*p*-val)
Time-mode component	Improving	−0.02(0.76)	0.03(0.17)	−0.04(0.29)	−0.03(0.65)	0.01(0.79)	−0.03 (0.37)
	Persisting	0.23(< 0.001)	0.06(< 0.001)	0.14(< 0.001)	0.24(< 0.001)	0.08(< 0.001)	0.14(< 0.001)

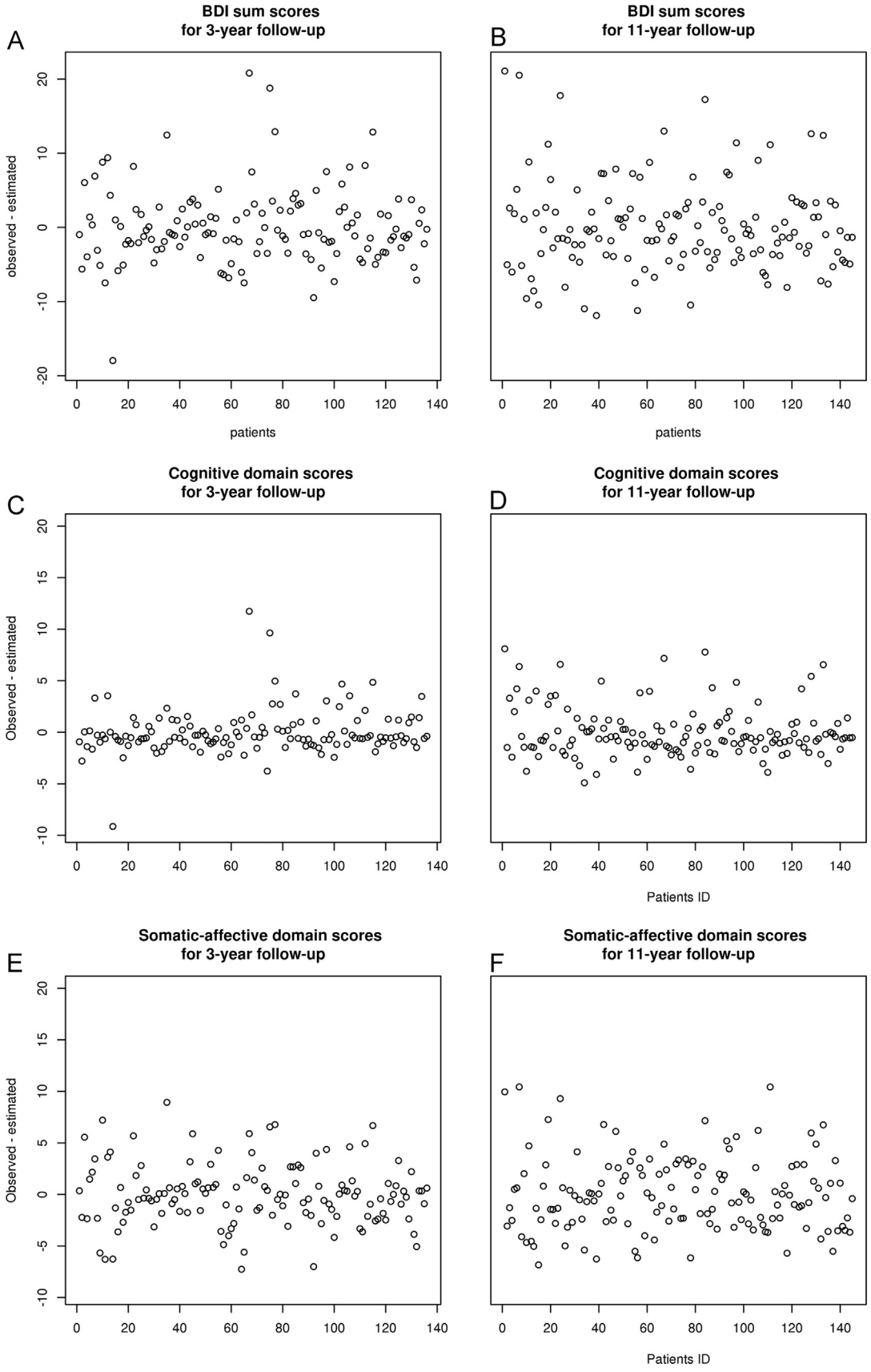

Fig. 1. Residual plots of BDI sum scores and symptom-domain scores.

value. Together, these findings suggest that the patterns of persistence of symptomatology over time were more important for long-term prediction than the baseline severity levels.

Although the current study took a more integrative approach to explaining heterogeneity, the results fit in with previous work using more traditional latent variable techniques. The many studies attempting to decrease symptom-heterogeneity with FA and/or PCA have found highly mixed results, both in terms of the number and the content of identified factors (e.g. van Loo et al. (2012), Shafer et al. (2006)). This variation can be due to the use of different instruments, samples and/or analytical choices, but could also stem from the fact that essential information about person-level and/or time-level heterogeneity is not included in these analyses. Although the current analysis did consider these sources of heterogeneity, the results were seemingly in line with previous studies that found mood/cognitive and somatic/vegetative factors (e.g. Joiner and Lonigan, 2000, Wardenaar et al. (2012)). Comparisons of the current 3MPCA results with previous LCA and LCGA/GMM studies is quite hard due to the conceptual differences between the approaches. Studies using LCA have found subtypes of depression with different prognoses (e.g Lamers et al. (2012)). Furthermore, studies that used LCGA/GMM to identify subgroups of patients with similar course-trajectories have also shown that these are associated with different long-term outcomes (e.g. Wardenaar et al. (2014, 2015)).

It is important to note that the primary aim of 3MPCA and related techniques is data-reduction and that 3MPCA was used here to explore the optimal way to capture depression heterogeneity and to evaluate the prognostic added value of this approach. Although results of this sort can be suggestive of the existence of certain 'true' or 'causal' dimensions and/or subtypes of depression, such conclusions should not be drawn based on a single study. 3MPCAs conducted in different datasets could show regularities in how interactions between symptom-, person- and time-level heterogeneity of depression are explained. Insight in such consistencies could help formulate an empirical, integrative model of depression heterogeneity, which could be used to formulate working-definitions for clinical subtypes of depression. Such subtypes could take the form of classifications that can help clinicians differentiate between patients with different patterns of symptom-specific persistence (e.g. somatic vs. cognitive depression), each with specific treatment and/or prevention indications.

The present study had several strengths, including the number of longitudinal assessments, the availability of long-term follow-up data and the possibility to compare predictions of multiple models and prognostic factors. However, there were also limitations. First, the results apply to a general practice sample and generalizability to other populations should be evaluated. Second, although high R^2–statistics were observed, much variance in long-term BDI scores remained unexplained, indicating that other factors/error also play an important role. Third, the results could not be evaluated in an independent sample. Fourth, the present results apply to the BDI and using other instruments could have led to different results. Finally, patients without 3- or 10-year follow-up scores were not included in the study, which could have led to bias due to selective attrition.

Further research could apply 3MPCA to data collected with other instruments and in other target-populations. Moreover, studies could explore the associations of 3MPCA models with clinical factors and biomarkers. For instance, person-mode components can be correlated with external variables. Alternatively, 3MPCA could be run with symptom data and biological/clinical data to get insight into their mutual associations.

In conclusion, a 3MPCA model that captures variations in course and symptomatology across patients showed better predictive ability for long-term depression severity than other predictive factors. These findings suggests that in order to optimize outcome prediction in depression, different sources of variation among patients should ideally be captured in the predictor-variables.

Acknowledgements

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). The trial from which the data were sourced was financially supported by grants from the Dutch Organization for Scientific Research (NWO), the Medical Sciences Program and Chronic Diseases Program, Research Foundations of the Health Insurance Company 'Het Groene Land', the Regional Health Insurance Company (RZG), the Netherlands Foundation for Mental Health (NFMH), and the University Hospital Groningen to J. Ormel and H.J. Conradi.

Appendix A. Suplementary information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jad.2015.09.018.

References

Alonso, J., Angermeyer, M.C., Bernert, S., Bruffaerts, R., Brugha, T.S., Bryson, H., Girolamo, G., de Graaf, R., Demyttenaere, K., Gasquet, I., Haro, J.M., Katz, S.J., Kessler, R.C., Kovess, V., Lepine, J.P., Ormel, J., Polidori, G., Russo, L.J., Vilagut, G., Almansa, J., Arbabzadeh-Bouchez, S., Autonell, J., Bernal, M., Buist-Bouwman, M.A., Codony, M., Domingo-Salvany, A., Ferrer, M., Joo, S.S., Martinez-Alonso, M., Matschinger, H., Mazzi, F., Morgan, Z., Morosini, P., Palacin, C., Romera, B., Taub, N., Vollebergh, W.A., 2004. Disability and quality of life impact of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr. Scand. 109 (Suppl s420), 38–46.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders. APA, Washington DC, Fifth edition.

Beck, A.T., Ward, C.H., Mendelson, M., 1961. An inventory for measuring depression. Archiv. Gen. Psychiatry 4, 561–571.

Beck, A.T., Steer, R.A., Garbin, M.G., 1988. Psychometric properties of the beck depression inventory: twenty-five years later. Clin. Psychol. Rev. 8, 77–100.

Cattell, R.B., 1966. The data box: its ordering of total resources in terms of possible relational system. In: Cattell, R.B. (Ed.), Handbook of Multivariate Experimental Psychology. Rand-McNally, Chicago, pp. 67–128.

Conradi, H.J., De Jonge, P., Kluiter, H., Smit, A., Van der Meer, K., Jenner, J.A., Van Os, T.W.D.P., Emmelkamp, P.M.G., Ormel, J., 2007. Enhanced treatment for in primary care: long term outcomes of a psycho-educational prevention program alone and enriched with psychiatric consultation or cognitive behavioral therapy. Psychol. Med. 37 (6), 849–862.

Conradi, H.J., De Jonge, P., Ormel, J., 2008. Prediction of the three-year course of recurrent depression in primary care patients: different risk factors for different outcomes. J. Affect. Disord. 105 (1–3), 267–271.

Costa Jr., P.T., McCrae, R.R., 1989. NEO-PI/FFI Manual Supplement. Psychological Assessment Resources, Odessa.

Derogatis, L.R., Lipman, R.S., Covi, L., 1973. SCL-90: an outpatient psychiatric rating scale – preliminary report. Psychopharmacol. Bull. 9, 13–28.

Fried, E.I., Nesse, R.M., 2015. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR*D study. J. Affect. Disord. 172, 96–102.

Goldberg, D., 2011. The heterogeneity of "major depression". World Psychiatry 10, 226–228.

Honaker, J. King, G. Blackwell, M., 2011. Amelia II: a program for missing data. J. Stat. Softw. 45, 1–47.

Huijbregts, K.M., van der Feltz-Cornelis, C.M., van Marwijk, H.W., de Jong, F.J., van der Windt, D.A., Beekman, A.T., 2010. Negative association of concomitant physical symptoms with the course of major depressive disorder: a systematic review. J. Psychosom. Res. 68, 511–519.

Joiner, T.E., Lonigan, C.J., 2000. Tripartite model of depression and anxiety in youth psychiatric inpatients: relations with diagnostics status and future symptoms. J. Clin. Child Psychol. 29 (3), 372–382.

Karlsson, L., Kiviruusu, O., Miettunen, J., Heila, H., Holi, M., Ruuttu, T., Tuisku, V., Pelkonen, M., Marttunen, M., 2008. One-year course and predictors of outcome of adolescent depression, a case – control study in Finland. J. Clin. Psychiatry 69, 844–853.

Kessler, R.C., Chiu, W.T., Demler, O., Merikangas, K.R., Walters, E.E., 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Archiv. Gen. Psychiatry 62, 617–627.

Kessler, R.C., 2012. The costs of depression. Psychiatr. Clin. N. Am. 35, 1–14.

Kiers, H.A.L., 1998. Joint orthomax rotation of the core and component matrices resulting from three-mode principal components analysis. J. Classif. 15,

245–263.
Kiers, H.A.L., 2000. Displaying results from three-way methods. J. Chemom. 14, 151–170.
Kiers, H.A.L., Van Mechelen, I., 2001. Three-way component analysis: principles and illustrative application. Psychol. Methods 6 (1), 84–110.
Kiers, H.A.L., Der Kinderen, A., 2003. A fast method for choosing the numbers of components in Tucker 3 analysis. Br. J. Math. Stat. Psychol. 56, 119–125.
Kroonenberg, P.M., De Leeuw, J., 1980. Principal component analysis of three-mode data by means of alternating least squares algorithms. Psychometrika 45, 69–97.
Kroonenberg, P.M., 2008. Applied Multiway Data Analysis. Wiley Series Probab. Stat. John Wiley and Sons, Hoboken NJ.
Kroonenberg, P.M., van Ginkel, J.R., 2012. Combination rules for multiple imputation in three-way analysis illustrated with chromatography data. Curr. Anal. Chem. 8, 224–235.
Lamers, F., Rhebergen, D., Merikangas, K.R., De Jonge, P., Beekman, A.T.F., Penninx, B. W.J.H., 2012. Stability and transitions of depressive subtypes over a 2-year follow-up. Psychol. Med. 42 (10), 2083–2093.
Monden, R., Wardenaar, K.J., Stegeman, A., Conradi, H.J., de Jonge, P., 2015. Simultaneous decomposition of depression heterogeneity on the Person-, Symptom- and Time-level: the use of the three-mode principal component analysis. PLoS One 10 (7), e0132765.
Mueller, T.I., Lavori, P.W., Keller, M.B., Swartz, A., Warshaw, M., Hasin, D., Coryell, W., Endicott, J., Rice, J., Akiskal, H., 1994. Prognostic effect of the variable course of alcoholism on the 10-year course of depression. Am. J. Psychiatry 151, 701–706.
Olbert, C.M., Gala, G.J., Tupler, L.A., 2014. Quantifying heterogeneity attributable to polythetic diagnostic criteria: theoretical framework and empirical application. J. Abnorm. Psychol. 123 (2), 452–462.
Patten, S.B., Wang, J.L., Williams, J.V., Lavorato, D.H., Khaled, S.M., Bulloch, A.G., 2010. Predictors of the longitudinal course of major depression in a Canadian population sample. Can. J. Psychiatry 55, 669–676.
Penninx, B.W., Nolen, W.A., Lamers, F., Zitman, F.G., Smit, J.H., Spinhoven, P., Cuijpers, P., de Jong, P.J., van Marwijk, H.W., van der Meer, K., Verhaak, P., Laurant, M.G., de Graaf, R., Hoogendijk, W.J., van der Wee, N.V., Ormel, J., van Dyck, R., Beekman, A.T., 2011. Two-year course of depressive and anxiety disorders, results from the Netherlands Study of Depression and Anxiety (NESDA). J. Affect. Disord. 133, 76–85.
Rhebergen, D., Batelaan, N.M., De Graaf, R., Nolen, W.A., Spijker, J., Beekman, A.T., Penninx, B.W., 2011. The 7-year course of depression and anxiety in the general population. Acta Psychiatr. Scand. 123 (4), 297–306.
Riihimäki, K.A., Vuorilehto, M.S., Melartin, T.K., Isometsä, E.T., 2011. Five-year outcome of major depressive disorder in primary health care. Psychol. Med. 16, 1–11.
Shafer, A.B., 2006. Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. J Clin Psychol 62 (1), 123–146.
Smit, A., Tiemens, B.G., Ormel, J., Kluiter, H., Jenner, J.A., van der Meer, K., Van Os, T. W., Conradi, H.J., 2005. Enhanced treatment for depression in primary care: first year results on compliance, self-efficacy, the use of antidepressants and contacts with the primary care physician. Prim. Care Commun. Psychiatry 10, 39–49.
Smit, A., Kluiter, H., Conradi, H.J., van der Meer, K., Tiemens, G., Jenner, J.A., Van Os, T.W., Ormel, J., 2006. Short-term effects of enhanced treatment for depression in primary care: results from a randomized controlled trial. Psychol. Med. 36, 15–26.
Sullivan, P.F., Kessler, R.C., Kendler, K.S., 1998. Latent class analysis of lifetime depressive symptoms in the national comorbidity survey. Am. J. Psychiatry 155 (10), 1398–1406.
Timmerman, M.E., Kiers, H.A.L., 2000. Three-mode principal component analysis: choosing the numbers of components and sensitivity to local optima. Br. J. Math. Stat. Psychol. 53, 1–16.
Ten Berge, J.M.F., 1977. Orthogonal Procrustes rotation for two or more matrices. Psychometrika 42, 267–276.
Ter Smitten, M.H., Smeets, R.M.W., Van der Brink, W., 1998. Composite International Diagnostic Interview (CIDI). Basis version 2.1. Lifetime manual. WHO, Amsterdam.
Tucker, L.R., 1963. Implications of factor analysis of three-way matrices for measurement of change. In: Harris, C.W. (Ed.), Problems in Measuring Change. University of Wisconsin Press, Madison, pp. 122–137.
Tucker, L.R., 1966. Some mathematical notes on three-mode factor analysis. Psychometrika 31, 279–311.
van Loo, H.M., de Jonge, P., Romeijn, J.W., Kessler, R.C., Schoevers, R.A., 2012. Data-driven subtypes of major depressive disorder: a systematic review. BMC Med. 10, 156.
Ware, J.E., Sherbourne, C.D., 1992. The MOS 36-item short form health survey (SF-36). 1. Conceptual Framework and item selection. Med. Care 30 (6), 473–483.
Wardenaar, K.J., Giltay, E.J., van Veen, T., Zitman, F.G., Penninx, B.W., 2012. Symptom dimensions as predictors of the two-year course of depressive and anxiety disorders. J. Affect. Disord. 136, 1198–1203.
Wardenaar, K.J., de Jonge, P., 2013. Diagnostic heterogeneity in psychiatry: towards an empirical solution. BMC Med. 11, 201.
Wardenaar, K.J., Conradi, H.J., de Jonge, P., 2014. Data-driven course trajectories in primary care patients with major depressive disorder. Depression Anxiety 31, 778–786.
Wardenaar, K.J., Monden, R., Conradi, H.J., de Jonge, P., 2015. Symptom-specific course trajectories and their determinants in primary care patients with major depressive disorder: Evidence for two etiologically distinct prototypes. J. Affect. Disord. 179, 38–46.
WHO, 1997. The Composite International Diagnostic Interview (CIDI). World Health Organization, Geneva.
Widiger, T.A., Clark, L.A., 2000. Toward DSM-V and the classification of Psychopathology. Psychol. Bull. 126 (6), 946–963.