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