Received: 11 January 2022 | Accepted: 11 July 2022

DOI: 10.1002/gps.5787

RESEARCH ARTICLE

International Journal of Geriatric Psychiatry WILEY

Network structure of time-varying depressive symptoms through dynamic time warp analysis in late-life depression

Denise C. R. van Zelst[1] | Eveline M. Veltman[1,2] | Didi Rhebergen[3] | Paul Naarding[4] | Almar A. L. Kok[5] | Nathaly Rius Ottenheim[2] | Erik J. Giltay[2,6,7]

[1]GGZ Rivierduinen, Leiden, The Netherlands

[2]Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands

[3]Mental Health Care Institute GGZ Centraal, Amersfoort, The Netherlands

[4]Department of Old-age Psychiatry, GGNet Apeldoorn/Zutphen, Zutphen, The Netherlands

[5]Department of Psychiatry, Amsterdam Public Health, Amsterdam University Medical Center, Vrije Universiteit, Amsterdam, The Netherlands

[6]Collaborative Antwerp Psychiatric Research Institute (CAPRI), Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium

[7]University Psychiatric Hospital Duffel, VZW Emmaüs, Duffel, Belgium

Correspondence
Erik J. Giltay, Department of Psychiatry, Leiden University Medical Center, Postal Box 9600, Postal Zone B1-P, 2300 RC Leiden, The Netherlands.
Email: e.j.giltay@lumc.nl

Abstract

Objectives: Late-life major depressive disorder (MDD) can be conceptualized as a complex dynamic system. However, it is not straightforward how to analyze the covarying depressive symptoms over time in case of sparse panel data. Dynamic time warping (DTW) analysis may yield symptom networks and dimensions both at the patient and group level.

Methods: In the Netherlands Study of Depression in Older People (NESDO) depressive symptoms were assessed every 6 months using the 30-item Inventory of Depressive Symptomatology (IDS) with up to 13 assessments per participant. Our sample consisted of 182 persons, aged ≥ 60 years, with an IDS total score of 26 or higher at baseline. Symptom networks dimensions, and centrality metrics were analyzed using DTW and Distatis analyses.

Results: The mean age was 69.8 years (SD 7.1), with 69.0% females, and a mean IDS score of 38.0 (SD = 8.7). DTW enabled visualization of an idiographic symptom network in a single NESDO participant. In the group-level nomothetic approach, four depressive symptom dimensions were identified: "core symptoms", "lethargy/somatic", "sleep", and "appetite/atypical". Items of the "internalizing symptoms" dimension had the highest centrality, whose symptom changes over time were most similar to those changes of other symptoms.

Conclusions: DTW revealed symptom networks and dimensions based on the within-person symptom changes in older MDD patients. Its centrality metrics signal the most influential symptoms, which may aid personalized care.

KEYWORDS
cluster analysis, dynamic time warp analysis, Inventory of Depressive Symptomatology, late-life depression, major depressive disorder, network analysis, time series

Key points

- Dynamic time warp (DTW) analysis yielded symptom networks and dimensions in patients with late-life depression, based on the tendency of symptom changes over time to covary.

- DTW and network analysis applied at individual patients may reveal symptoms with high centrality, which may aid personalized medicine.
- Four Inventory of Depressive Symptomatology symptom dimensions were identified at the group level: "Core symptoms" (14 items), "Lethargy/Somatic" (6 items), "Sleep" (3 items), and "Appetite/Atypical" (7 items).

1 | INTRODUCTION

Late-life major depressive disorder (MDD) is a common and often severe psychiatric disorder, with a worse long-term prognosis compared to younger adults.[1,2] Clinical presentation of late-life MDD differs from MDD in younger adults, with older adults often having more somatic symptoms and more loss of interest, and younger adults presenting more often with a lowered mood and guilt symptoms.[3–6] Research into late-life MDD is complicated by the high heterogeneity of symptoms across individual patients.[7] This heterogeneity has been difficult to address in research due to the use of dichotomous component criteria for MDD,[8] unweighted sum scores of rating scales as a measure of severity, and the idea of MDD symptoms as indicators of a single underlying "latent common cause".[9–11] However, there is increasing evidence that MDD symptoms are not psychometrically interchangeable.[12,13] From the complex dynamic systems perspective, MDD can be seen as an emergent property of the complex interaction of symptoms that may cause each other and may adapt in response to each other in a (causal) network structure.[7,14,15] Yet, to model such networks and dimensions from time series of symptom scores for the individual patients is challenging, especially in case of sparse panel data.[16,17]

Symptom networks and dimensions have been studied, using cross-sectional designs, but less so in prospective studies. In studies on the Inventory of Depressive Symptomatology (IDS) the most consistent symptom clusters of dimensions that were found using factor analysis were termed "Melancholic", "Core", and "Sleep".[18,19] It is however difficult to derive general conclusions for old-age psychiatry, because of inconsistent findings and because most studies were conducted in younger adults.[20–25] Moreover, diverse statistical methods (e.g., factor analysis, latent class analysis, hierarchical cluster analysis, vector autoregressive-based clustering) and different instruments to assess depressive symptoms (i.e., self-reported and observer-rated scales) were used.[17] Furthermore, most studies had cross-sectional designs, which do not consider the (dis)similarities of changes in depressive symptoms over time. Yet, some have studied panel data or time series of symptom severity by modeling their temporal relationships.[24,26,27] Their findings showed that not the trends of sum scores, but rather the idiosyncratic changes of symptoms could be considered the main data of interest. Some of these studies used analyses of lagged relationships, such as in vector-autoregressive (VAR) models that analyze dynamics typically one timepoint apart ($t - 1$). However, these models are not yet widely used in clinical practice, because it has been difficult to choose the proper time interval between assessments (e.g., hours, days, weeks, or months), because these complex models consume considerable computing time, and because results can be difficult to interpret.[28–30] Moreover, many severely depressed older patients are incapable to complete daily or even more frequent smartphone-based ecological momentary assessments (EMA) of depressive symptoms. New analytic approaches may be needed to analyze the covarying depressive symptoms over time in case of sparse panel data, especially in old-age psychiatry.

Our study aims to identify those symptoms which changes covary in older people with depression. When we can identify those symptoms that change together over time, we can determine networks and dimensions based on these similarities. Dynamic time warping (DTW) is a relatively new analytic technique in psychiatry, and not yet been previously been used in late-life MDD, but only in adult depressive inpatients.[24,31] The Netherlands Study on Depression in Older Persons (NESDO)[32] provides the opportunity to study the symptom networks of participants by providing 6-year longitudinal data with follow-up measurements every 6 months. We hypothesize that the DTW methodology will yield symptom networks and dimensions in late-life MDD that capture their within-patient dynamics. We apply the technique to the patient level (idiographic approach) as well as to the group level (nomothetic approach). We included both the idiographic and nomothetic analyses in support of the idea that these are complementary approaches, that can be combined for added strength.

2 | METHODS

2.1 | Sample

Data were obtained from the baseline and follow-up measurements of NESDO, a longitudinal multi-site naturalistic cohort study examining the course and consequences of depression in older people. From 2007 until 2010, persons aged 60–93 years were recruited from mental health care institutes and primary healthcare settings in five regions in the Netherlands. The NESDO cohort ($n = 510$) consisted of 378 depressed (diagnosed within the previous 6 months according to the DMS-IV-R criteria) and 132 non-depressed persons. Exclusion criteria were suspected dementia (with a Mini Mental State Examination score (MMSE) under 19), a primary diagnosis of dementia, and insufficient mastery of the Dutch language. The study design of NESDO has been described in detail previously.[32] For the current study, participants with a 6-month DSM-IV-R diagnosis of major depression were included, who had an IDS sum score of 26 or

higher (indicating moderate to severe depression), which yielded 182 (48.1% of 378 depressed) participants at baseline. To illustrate the idiographic approach we choose on participant from our NESDO sample. This DTW analysis in a single participant may yield reliable insight about the idiosyncratic dynamics of depressive symptoms over time, and the relative centrality of each of the depressive symptoms for this patient. It may help illustrate that a valid within-person process-based model could be developed.

2.2 | Subject characteristics

Socio-demographic variables, including age, gender, and years of education, were collected during the baseline interview. Clinical characteristics such as age of onset of the depressive disorder and comorbid anxiety disorders were assessed by the Composite International Diagnostic Interview (CIDI). Antidepressant and benzodiazepine use was assessed by interview and double-checked by chart review. Respondents were asked whether they were currently smoking (yes/no) and the number of alcoholic drinks per week. Body mass index (BMI) was calculated as kilograms divided by meter squared. Finally, the presence of cardiovascular disease (assessed by self-report supported by appropriate medication use (see Ref.[33] for detailed description), and the number of other self-reported chronic diseases were determined for which persons received treatment (including lung disease, osteoarthritis or rheumatic disease, cancer, ulcer, intestinal problem, liver disease, epilepsy and thyroid gland disease).[32]

2.3 | Measurements

Within the NESDO cohort, depressive symptoms were assessed every 6 months using the Dutch version of the 30-item Inventory of Depressive Symptomatology (IDS)[34] for a total duration of 6 years, with up to 13 waves. The IDS is a 28-item questionnaire that not only comprises all symptoms of depression as defined by the DSM-5, but also melancholic, atypical, and anxious symptoms.[34] The questionnaire has been shown to have adequate reliability, acceptable validity, good responsiveness and good discriminative ability.[34–36] Each item of the IDS is scored on a 0–3 scale. Persons were excluded if they had three or less IDS assessments. When we compared the 182 participating patients with moderate to severe depression with four or more assessments with the 52 excluded patients with moderate to severe depression with three or less assessments, we found that they had a similar gender distributions (69% vs. 65%, $p = 0.32$), mean age (of 69.9 vs. 70.4 years, $p = 0.58$), and mean number of years in education (10.4 vs. 9.7 years, $p = 0.18$).

Missing item scores were imputed based on the average score on other items within that individual (in only 0.85% of item scores). Opposite symptoms (i.e., increased or decreased weight/appetite, and increased or decreased sleep) were split, so each of these items became two separate items instead of one, yielding 30 IDS items. All 30 IDS items were group-level standardized into *z*-scores before DTW analysis.

2.4 | Statistical analysis

Sociodemographic and clinical variables were summarized as means with standard deviations (SD) or percentages, as appropriate.

DTW analysis was used to identify dimensions of depressive symptoms that tended to cluster dynamically over time, as described in more detail previously.[24,31] It combines both instantaneous (i.e., contemporaneous) relations as well as associations among symptom scores one time point away. Importantly, it can be applied to sparse time series data. When two symptom changes tended to occur at nearby timepoints, the resulting distance among this symptom pair will be smaller. The distances for each of the symptom pairs yields a distance matrix for each participant, from which symptom dimensions and symptom networks can be derived, both at the individual level (i.e., idiographic approach) and the group level (i.e., nomothetic approach). This algorithm can effectively cluster panel and time-series data,[37] as it calculates the strength of relationships (i.e., distance) between the changing symptom scores over time.[38,39]

Thus, within each patient, DTW calculated the "distance" between each pair of items, resulting in a 30 by 30 distance matrix for each individual. As we included 182 MDD patients, this resulted in 182 distance matrices. The time window was set at 1, meaning that changes between $t - 1$, t, and $t + 1$ were taken into account, using a programming approach with elastic stretching and compressing in order for the time-series of each pair of symptoms to become completely aligned through a warping path. In the supplements (Supplementary Figure 2), the need of using a time window is explained in more detail, as using no window can lead to the erroneous conclusion of a small distance, as item scores very far away in time are being aligned. Figure 1 explains the DTW analysis in detail for symptom scores of three items over time. It exemplifies the calculation of DTW distances for these three fictious symptom pairs, yielding a 3-by-3 distance matrix. S1 and S2 show a much more similar course over time than S3 (represented by the much smaller distance of 3 vs. 12 and 15).

The 182 distance matrices were subsequently analyzed on the group level (i.e., nomothetic approach), through a Distatis analysis, which is a generalization of principal component analysis (PCA). It is aimed to find the stable part of symptom dynamics among participants. For Distatis the raw data consists of the 182 distance matrices collected on the same 30 IDS symptoms. Distatis is a three-way extension of metric multidimensional scaling, in which the 182 distance matrices are transformed into 182 cross-product matrices, from which the best common representation of the observations (called the compromise) is calculated. The three compromise factors best describe the similarity structure of the 182 distance matrices.

These three factors are used as *x*, *y*, and *z*-axis values in the three-dimensional plot in which the symptoms are represented as points, such that the distances in the plot best reflect the similarities

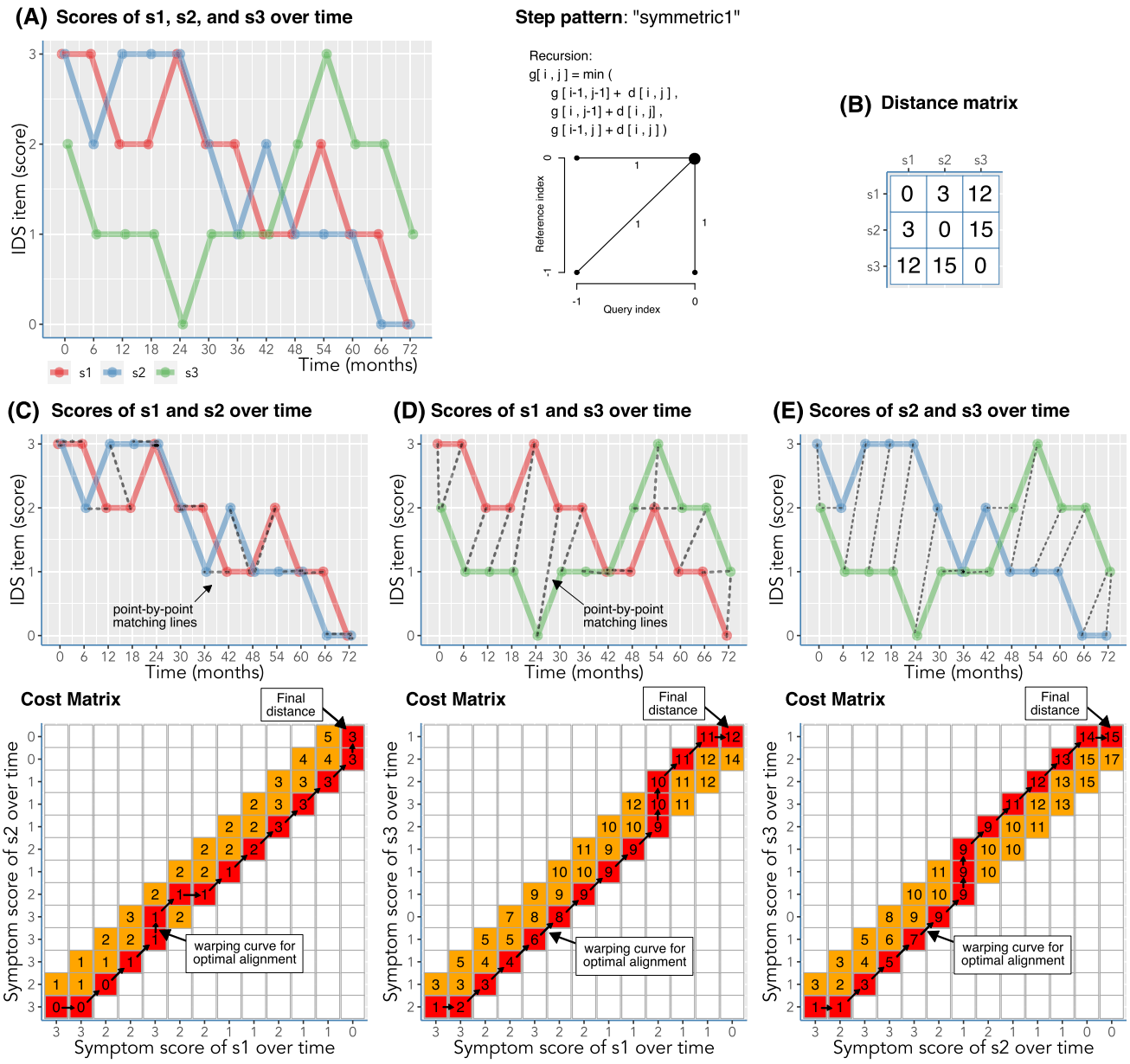

FIGURE 1 Explanation of the dynamic time warping (DTW) analysis. Say we have gathered three IDS symptom scores over time in a single fictitious patient, and (A) the (unstandardized) scores of these individual symptoms S1, S2, and S3 are given over time, with timepoint "0" being baseline. When we used the "symmetric1" step pattern recursion, (B) the following symmetric distance matrix was calculated. Parts (C)–(E) explain the calculations of DTW distances for the symptom pairs S1–S2, S1–S3, and S2–S3, respectively. The black lines illustrate the warped (i.e., elastic) modification of one item to get an optimal alignment, using the Sakoe-Chiba constraint (i.e., window type) of one time point before and after the current assessment.[59] The Cost Matrices show the optimal warping routes for each of these three calculations, yielding 3, 12, and 15 as for their respective distances. Thus, the distance between S1 and S2 was smallest, thus having the most similar course over time. As we had 30 IDS items, this calculation has been done 420 times [i.e., (30 × 30)/2, minus 30 of the diagonal] for each of the 182 individual patients, which yielded 182 (30-by-30) distance matrices

between the symptoms.[40] The first two compromise factors are also plotted into the *x*–*y* plane (with 95% confidence intervals based on 1000 bootstrap replicates of the factor scores), and all three compromise factors were also plotted in a 3-dimensional plot using the "plotly" r package. A hierarchical cluster analysis was applied according to "Ward.D2" clustering methods, which was visualized in a dendrogram. To estimate the optimal number of dimensions, a scree plot of dimension distances was created, with the elbow method being used to yield the number of clusters. This elbow can be observed as a sharp change in the slopes of adjacent line segments in the Scree plot, which location might indicate a good number of dimensions to retain. A network plot was created that was used to calculate the relative centrality values of each of the 30 IDS-SR items.

Using a random split with a subset of two groups of 91 patients, we compared the two separate DTW symptom networks, to study the stability of the network and dimensional structure. Node placement was done by using the Procrustes algorithm, to aid the visual comparison between the two networks and calculate their congruence coefficient.[41] The congruence coefficient (with the 95% CI) was estimated, through bootstrapping of 200 random splits of the 182 participants. A value below 0.85 is indicates poor similarity, a value in the range of 0.85–0.94 indicates fair similarity, and a value of 0.95 can be considered as being equal.[41]

The "dtw" (version 1.22.3), "parallelDist" (version 0.2.4), "qgraph" (version 1.6.9), "plotly" (version 4.10.0), "stats" (version 4.0.3), and "networktools" (version 1.2.3) packages for the R statistical software were used (R version 4.0.3; R Foundation for Statistical Computing, 2016. URL: https://www.R-project.org/, https://osf.io/s7jrc). In the supplements, a sample r script is given for DTW analyses of the IDS panel data from three participants.

3 | RESULTS

3.1 | Patient characteristics

Table 1 shows the sociodemographic and clinical characteristics of the included persons at baseline. The total sample consisted of 182 older participants who were severely depressed at baseline, of whom 69.0% were females with a mean age of 69.8 years (standard deviation [SD] = 7.1). The mean age of onset was 44.1 years (SD = 19.7) and the mean IDS total score was 38.0 (SD = 8.7). There were 75.8% who used an antidepressant.

3.2 | Idiographic approach

Figure 2 shows the DTW analyses and clustering of one NESDO participant. Figure 2A shows the individual IDS item scores over time. IDS items with more similar dynamics are clustered according to the dendrogram that is shown in Figure 2B. Figure 2C shows the symptom network based on the DTW analysis, and the standardized centrality statistics are plotted in Figure 2D, with item 8 "Responsiveness of mood" having the smallest distances to most other symptoms, resulting in the strongest connection strength. Items are colored from red (i.e., strongest centrality) to blue (i.e., weakest centrality). Thus, item 4 "Hypersomnia" showed the most independent course over time in this particular participant. Supplementary Figure 2 illustrates the importance of using a time window in the DTW analysis. Three fictitious item scores over time are analyzed and compared using a Sakoe-Chiba window-type of size one versus no window. To explain the steps towards the DTW distance, a local cost matrix and the path that minimizes the alignment between two item scores while aggregating the total distance is shown for each of the three symptom pairs.

TABLE 1 Sociodemographic and clinical characteristics of 182 depressed elderly patients

Variable	Mean (SD) or percentage
Sociodemographic characteristics	
Age, years	69.8 (7.1)
Female sex	68.7
Education, years	10.4 (3.6)
Clinical characteristics:	
Age of onset, years	44.1 (19.7)
IDS total score	38.0 (8.7)
Depression severity[†]	
Mild	0
Moderate severe	57.1
Severe	26.9
Very severe	15.4
Comorbid anxiety diagnosis, %	43.4
Physical health characteristics	
Current smoking	26.4
Alcohol, units per day, median (IQR)	0.03 (0.0–1.2)
Body mass index (BMI)	25.5 (7.6)
Somatic comorbidity	
Heart disease	22.0
Cardiovascular disease	19.2
Diabetes mellitus	11.0
# chronic diseases	2.2 (1.5)
Medication use	
Antidepressants	75.8
SSRIs	33.5
TCAs	18.7
Other AD	26.4
Antipsychotic	9.3
Benzodiazepines	48.4

Abbreviation: IDS, Inventory of Depressive Symptomatology.

[†]Depression severity, measured using the IDS sum scores at baseline: mild (0–25), moderate severe (26–38), severe (39–48), and very severe (≥49).

3.3 | Nomothetic approach

Figure 3 shows the nomothetic group-level analyses of the 182 NESDO patients with severe depression at baseline. Based on the scree plot analysis (Figure 3A), the most parsimonious number of dimensions was four. The DTW analysis and the subsequent Distatis analysis yielded three compromise factors, which are depicted in a three-dimensional plot (Online Figure 1, https://osf.io/4j5xg). The dendrogram based on the hierarchical cluster analyses (Figure 3A),

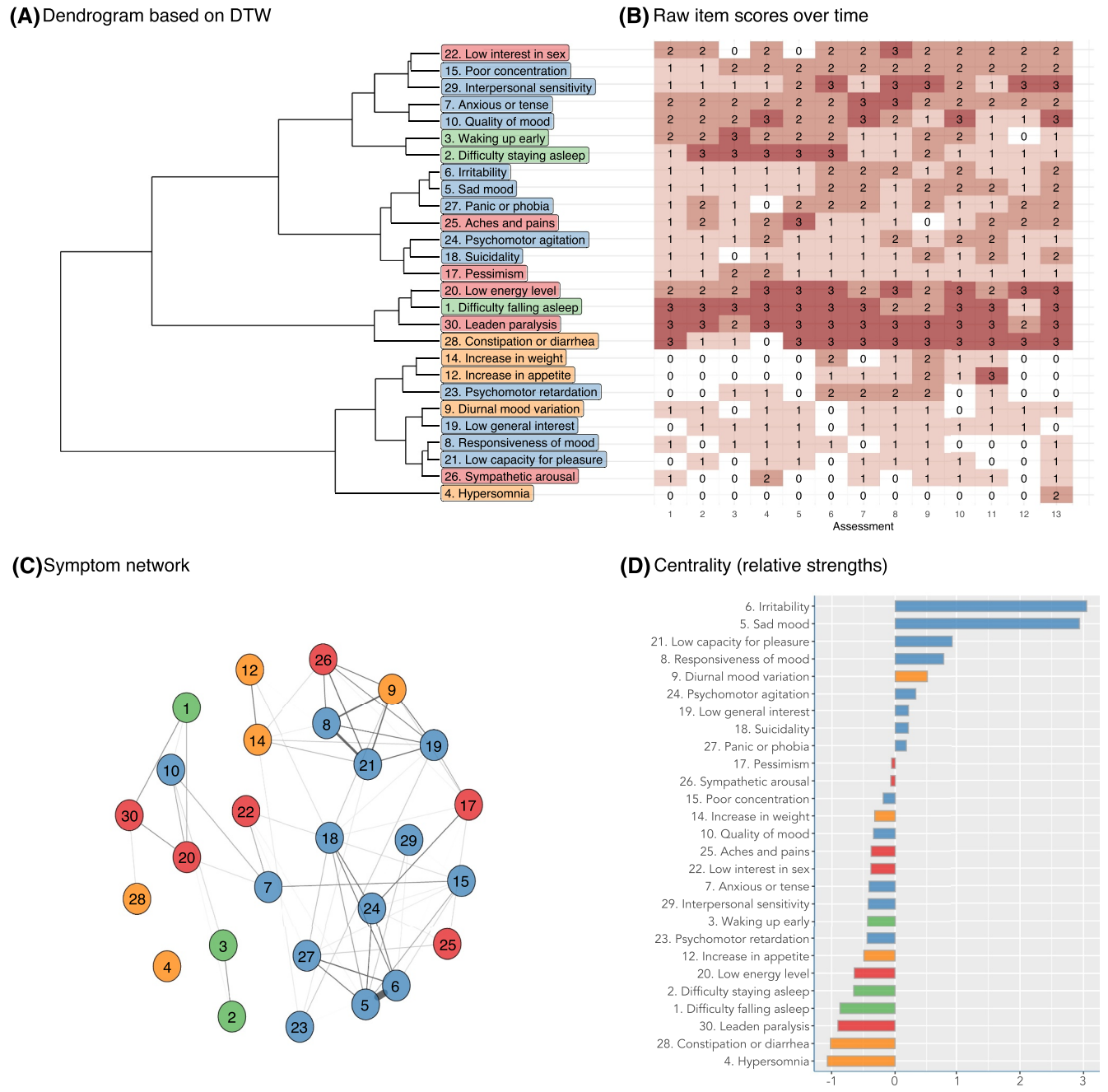

FIGURE 2 Dynamic time warping (DTW) analysis for a single Netherlands Study of Depression in Older People (NESDO) participant (idiographic approach). (A) The individual raw IDS-SR item scores over time are shown. (B) Using DTW analysis the distance matrix was assessed, which was entered in a hierarchical cluster analysis visualized in a dendrogram. (C) The distance matrix was also used to construct the symptom network graph (with one item that was constant being omitted). (D) The centrality statistics of this symptom network is shown, of which item 6 (Irritability) had the strongest strengths of connections with most other items.

the relative centrality (Figure 3B), and the compromise plots (Figure 3C and 3D) are given.

The four symptom dimensions thus consisted of items with similar course trajectories, namely: (1) Core symptoms (14 items: sad mood, irritability, anxious or tense, responsiveness of mood, quality of mood, poor concentration, low self-esteem, suicidality, low general interest, low capacity for pleasure, psychomotor retardation, psychomotor agitation, panic or phobia, and interpersonal sensitivity), (2) Lethargy/Somatic (6 items: pessimism, low energy level, low interest in sex, aches and pains, sympathetic arousal, and leaden paralysis), (3) Sleep (3 items: difficulty falling asleep, difficulty staying asleep, and waking up early) and (4) Appetite/atypical (7 items: hypersomnia, diurnal mood variation, decrease in appetite, increase in appetite, decrease in weight, increase in weight, and constipation or diarrhea).

Supplementary Figure 3 shows the four symptom dimensions when using three other hyperparameter settings (with different

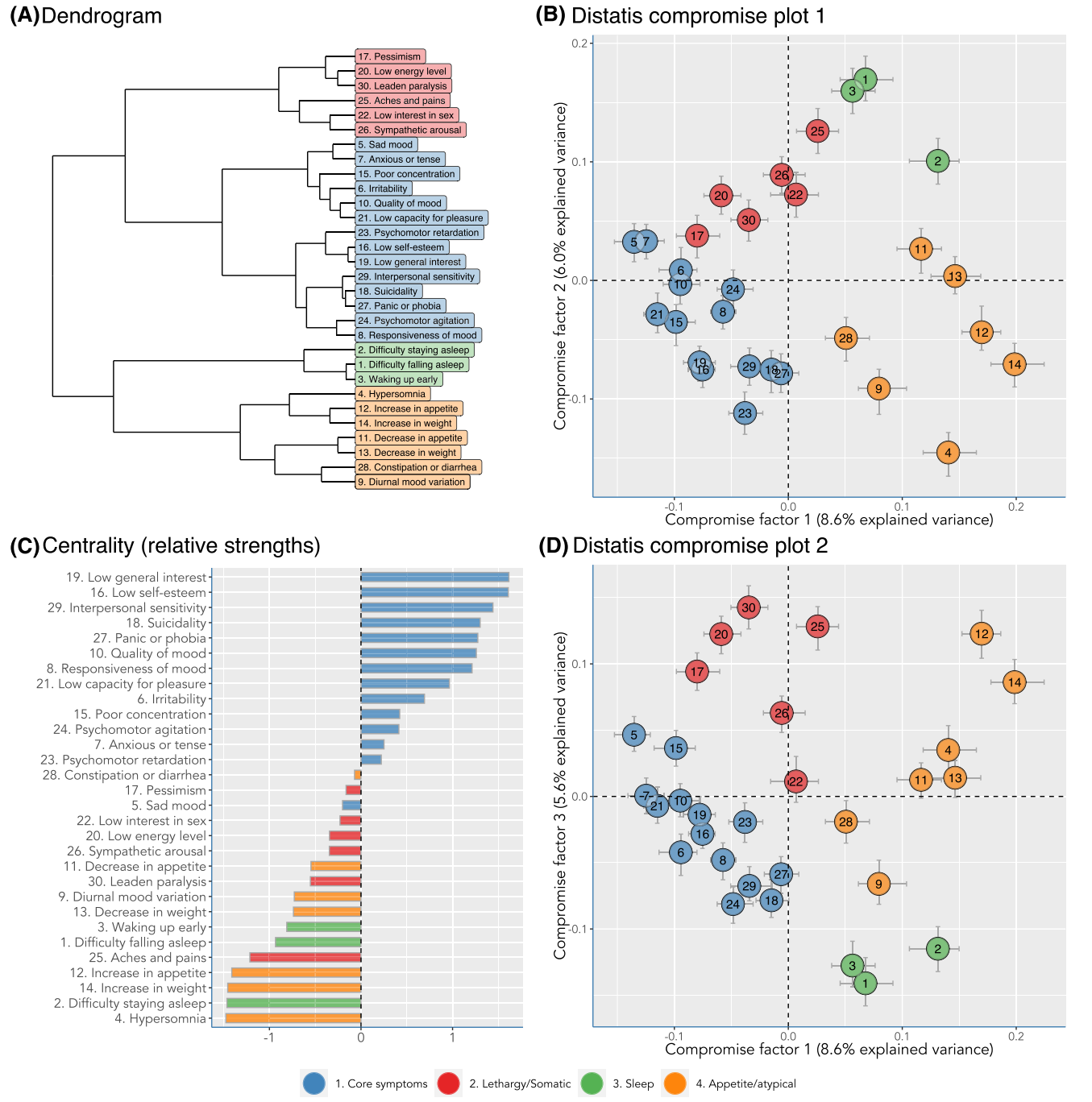

FIGURE 3 Nomothetic analyses based on all distance matrices from 182 participants. (A) A scree plot indicated four symptom dimensions. (B) A dendrogram was created, based on the Ward's (D2, i.e., general agglomerative hierarchical clustering procedure) clustering criterion on the compromise factors of the Distatis analysis of 182 distance matrices. (C) The Distatis analysis yielded three compromise factors. The position of each of the 30 IDS-SR items are shown in *x*–*y* scatter plot of the compromise space according to the first two of the three compromise factors. (D) Relative centrality of each of the 30 IDS-SR items based on the symptom network

window-types and window-sizes), and these different analyses results in largely similar symptom dimensions, meaning these are stable across the different settings. Moreover, Supplementary Figure 4 compares the DTW with the Euclidian distance approach and shows that the latter approach explains less variability. Like the DTW analysis, it does cluster the three sleep items together. Moreover, the increases and decreases in appetite and weight are clustering together. All the remaining 19 items are clustered more strongly together when using the Euclidian distance approach than when using DTW.

As is shown in Figure 4 the network did not significantly change when splitting up the participants into two random groups of 91

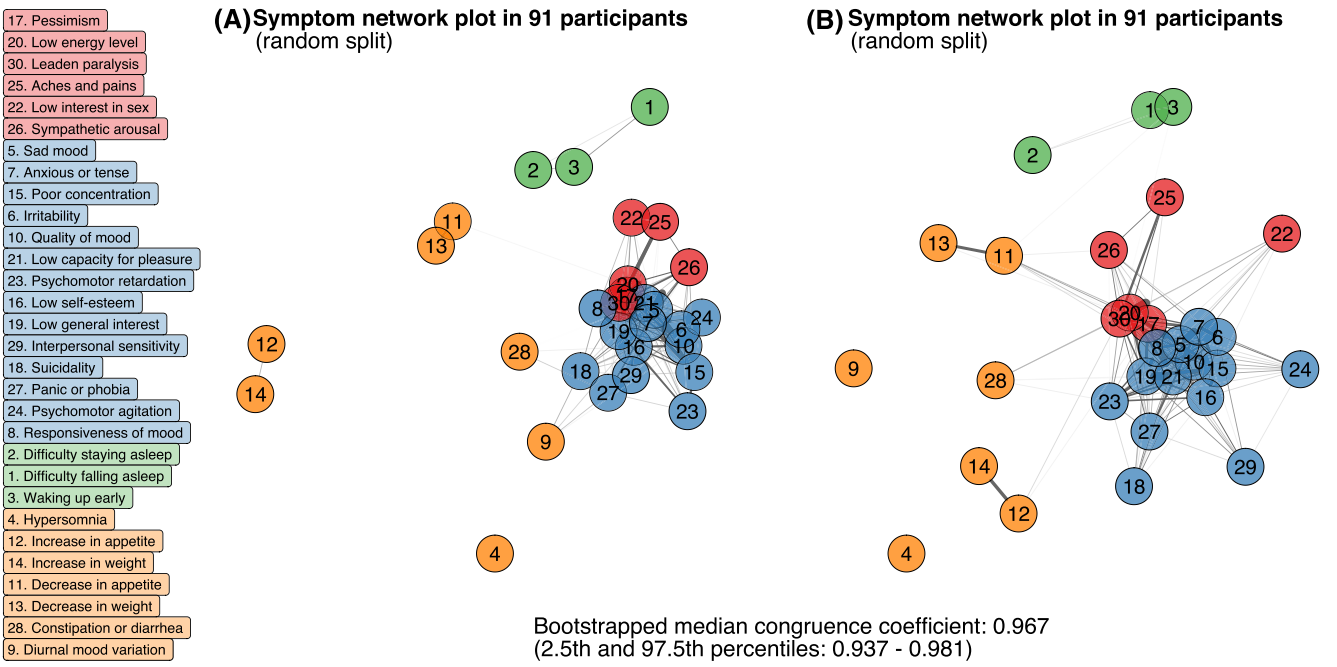

FIGURE 4 Network plots of two subsamples (A, B) of the 182 patients. We used an automated split with a subset of 91 patients each, in which we conducted separate dynamic time warping (DTW) analyses. Node placement was done by using the Procrustes algorithm (from the R Package "networktools"), to aid the visual comparison between the two networks. The congruence coefficient was high at 0.967 (95%CI: 0.937–0.981)

participants each. The median congruence coefficient was very high at 0.977, when we derived the network plots from 200 random splits of the data set.

As is shown in Figure 5, there is much overlap in the symptom dimensions in men and women, as supported by the high congruence coefficient (i.e., 0.945) when stratifying for gender. The "Lethargy/Somatic" and "Sleep" symptom dimensions were similar across the genders. However, the symptoms "23. Psychomotor retardation" and "27. Panic or phobia" clustered within the "Appetite/atypical" symptom dimension in men, whereas "28. Constipation or diarrhea" clustered within the "Core symptoms" dimension in women. Moreover, sleep problems tended to cluster more closely with decreases in weight and appetite in women than in men, whereas increases in weight and appetite tended to cluster as a separate cluster in women but not in men.

4 | DISCUSSION

The present study is the first to analyze the similarity of symptom changes over time in older depressed adults using DTW analyses. In the group-level nomothetic approach, four depressive symptom clusters and their course trajectories were identified. The very high congruence factor showed that these four dimensions were highly stable between participants. Analysis of the symptom network of one participant illustrates that DTW could also be a fruitful idiographic approach to identify dimensions and the most influential symptoms within a single participant. Symptom clusters were mostly consistent among men and women, with some subtle differences. Item 28, "Constipation and diarrhea" clustered more strongly with the core depressive symptoms in women than in men. Other cross-sectional studies that compared IDS items among men and women found that this symptom was more prevalent in women than in men.[42,43] Moreover, the atypical items 12 and 14 of increases in appetite and weight strongly clustered separately in women but not in men, and these symptoms were also more prevalent in women than in men in these previous studies.[42,43] Although our findings do not support distinct female symptom dimensions, it does suggest some somatic symptoms of depression may develop differently in time among the sexes.

Studies that focused on symptom clusters in late-life depression often used latent class analyses,[44–47] and factor analyses.[18,19] Their most consistent findings were symptoms dimensions termed "Melancholic", "Core", and "Sleep". Except for the consistent "Sleep" dimension, our dimensions differed substantially from the symptom factors from previous cross-sectional studies. The most likely explanation for these inconsistencies may be the differences in analysis techniques used. Other potential factors involve the use of different depression scales, inclusion criteria, and age ranges. A previous Exploratory Factor Analysis (EFA) on the IDS data in the current NESDO cohort revealed a three-factor solution.[4] They included 229 participants with MDD in the previous month, and found evidence for three symptom factors ("Somatic", "Mood", and "Motivation"), while six items loaded on more than one of these factors. However, these three factors differed substantially from our four dynamic symptom dimensions, for examples the "Somatic" factor combined symptoms from three different dimensions that we found (i.e., 1. Core symptoms, 3. Sleep, and 4.

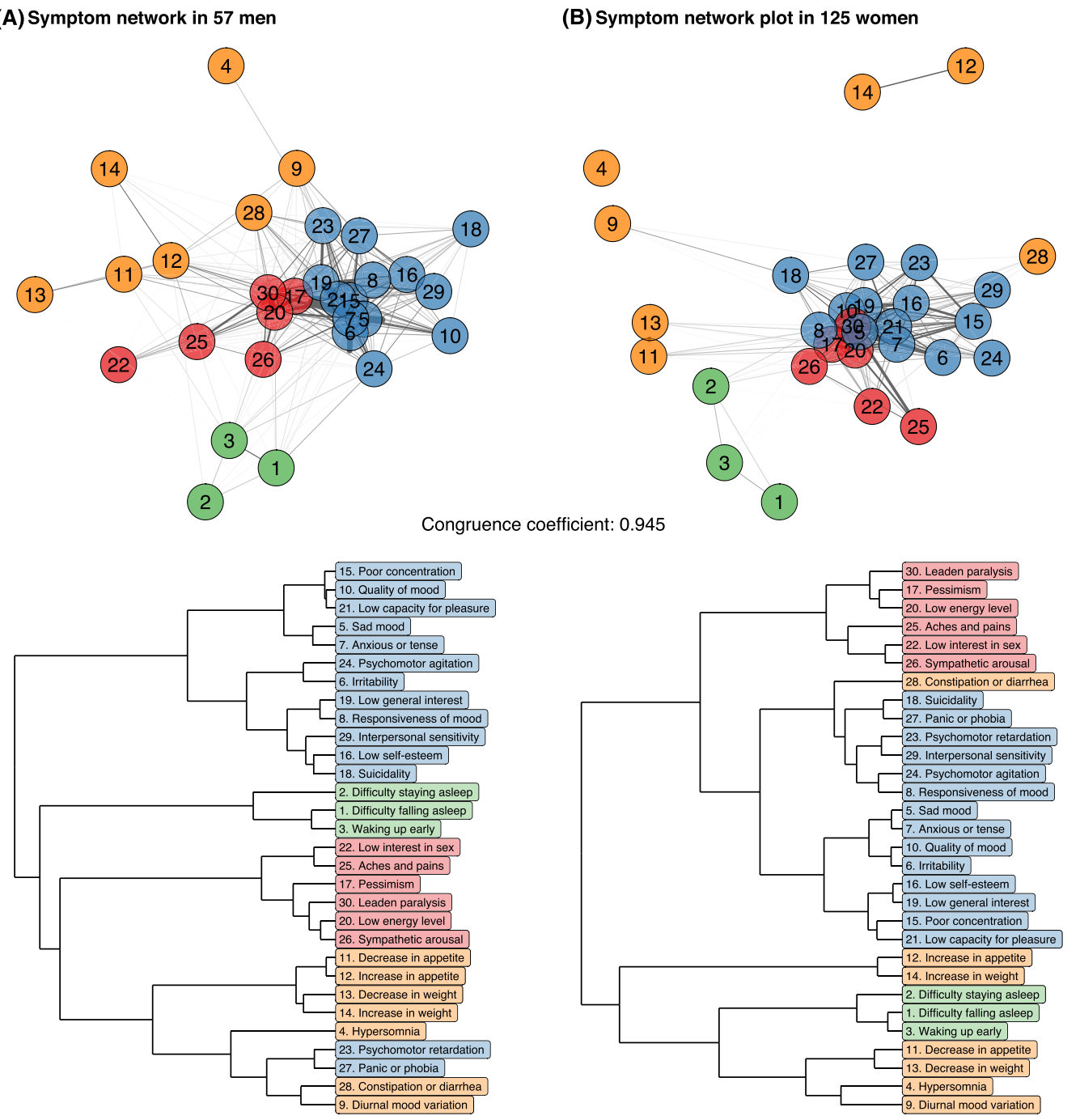

FIGURE 5 Network plots stratified for gender of the total of 182 patients. (A) The symptom network in 57 men and the dendrogram. (B) The symptom network in 125 women and the dendrogram. The congruence coefficient was large (0.945)

Appetite/atypical). Other factor analyses on IDS items were all in younger adults rather than older participants.[7,16,21,22,25,35,48–51] However, the clinical presentation of MDD in younger adults differs from that in older adults.[3,6] Using EFA, confirmatory factor analysis, Rasch analysis, principal component analysis, or network analysis, these studies found different groupings of IDS items, with little overlap among each other or with our finding, but some consistencies. Sleep symptoms were often closely associated,[22,48] and also appetite and weight symptoms,[48,49,51] as well as sleep symptoms in combination with appetite and weight symptoms.[4,35] These latter findings support the idea of the subtype of depressive disorder with atypical symptoms, within both adult and older patients. The many discrepancies with previous studies could likely be ascribed to the fact that we were using panel data (i.e., time series with sparse data) that first analyzed change profiles within patients, after which these data were aggregated. This raises the idea that using sum scores of the four symptom dimensions that we found would be more helpful in monitoring treatment effects, rather than using those factors found in cross-sectional analyses. The change profiles are strongly associated within each of the four dimensions that we found.

The field of research on symptom dynamics over time is growing. Idiographic and nomothetic symptom networks have been studied with different analytical techniques.[21,24,25,52] A recent comparable DTW analysis on younger depressed adults using Hamilton Rating Scale for Depression (HRSD) instead of IDS data,[24] had findings that partly overlapped with regard to the sleep and core-symptom dimensions. These previous and our findings strengthen concerns on the loss of clinically relevant information when using only sum scores for assessing treatment progress in daily practice,[7,20] since much information that represents the dynamical symptom complexity gets lost. Moreover, specific symptom dimensions were found to have a distinct etiology in old-age MDD.[53]

DTW may offer some opportunities for personalized medicine in psychiatry. In our single participant, DTW analysis of IDS data showed that item 6 (Irritability) was a central symptom with the highest connection strength, followed by several other core depressive symptoms. Thus, irritability most strongly covaried with her other symptoms, and whether such a central symptom is causally linked to other symptoms should be explored further using techniques employing directed networks.[54] The connection strengths are likely to be different for each individual participant or depressive patient. It could be hypothesized that targeting treatment on such central symptoms early in therapy may lead to a more rapid resolution of closely connected depressive symptoms.[27,48] However, as discussed by McNally,[55] selectively deactivating a symptom is more easily said than done, as it is difficult to target a single symptom without simultaneously affecting others. For example, in our individual participant this would imply that (psycho)therapy should be focusing on her core depressive symptoms, rather than focusing on somatic symptoms like sleep and appetite that were much more loosely connected. DTW may thus reveal individual symptom patterns, which might lead to person-tailored pharmaco- and psychotherapy.

Strengths of the current study are the use of the innovative DTW clustering method to study panel data of individual symptoms scores, the comprehensive set of depressive symptoms, the long-term follow-up, different health-care settings, and the broad range of old age. As we group-standardized symptoms scores before the DTW analyses, the clustering of symptoms in dimensions was based on change over time.[17] The DTW algorithm differs from the analytical tools traditionally used in psychiatry research, first as it is based on non-linear relationships rather than linear regression, second focusses on change profiles rather than absolute levels (as symptom scores were group-level standardized before the DTW analyses), third compares temporal dynamics with those that occur both at the same time points or close by in time, and fourth models individual patients (i.e., idiographic approach) first, after which findings are aggregated to search for commonalities.[56,57] There are also some limitations that need to be discussed. First, the time interval between assessments was 6 months, so the symptom dimensions are based on relapses and recovery of symptoms over longer time periods. Therefore, symptom changes that fluctuated over much shorter time periods could not be taken into account. For an idiographic analysis to be of clinical importance during the treatment phase of depression, more assessments are preferable. As this study consisted of participants with up to 13 assessments, further studies are necessary to explore the effects of different number of assessments and different time intervals between assessments on the DTW findings. Symptom dynamics also likely cluster differently when using different scales, because items differ among them.[58,59] Second, we could not consider treatment effects and the effects of recovery from MDD. Distance matrices will likely fluctuate over time, as they are dynamic, multi-causal, partly random, and manifest idiosyncratically,[60] because of many complex interactions with environmental factors, including treatments received. The goal of treatment is to reduce symptoms, resulting in less variances and means for symptom severity over the course of therapy. Future studies should ideally compare symptom networks among randomized groups receiving different treatments. Third, items that do not change over time (e.g., scoring zero throughout follow-up) tend to cluster together, which should be considered when conducting DTW.

In conclusion, DTW enabled the estimation and visualization of symptom networks in late-life depression. In the group-level nomothetic approach four symptom dimensions were identified, namely: core symptoms/lethargy, internalizing symptoms, sleep, and appetite/atypical. DTW analysis is a new way to capture the underlying structure of the dynamics of depressive symptoms over time both in individuals as well as in groups of participants. DTW has a promising potential for clinical practice and could be a continuation of already available evidence of the value of measurement-based care in psychiatry.[61] Inflammatory and vascular risk markers and other somatic measurements could also be included in the analyses which may help to uncover specific etiology.

Since DTW is a relatively new analytic technique in psychiatry there is lot of details left to explore (like the minimum number of assessments that are required, the window-type, and the window-size to yield the most stable outcomes). The DTW approach should be developed further, for example, testing the symptom and environmental factors in a 2 or 3 week period using EMA or ESM prior to or during the first weeks of treatment.[56,58,60] Such an idiographic analysis might lead to actionable insights for improved treatment and prognosis, which may ultimately aid personalized medicine in psychiatry.

AUTHOR CONTRIBUTIONS

Giltay, Veltman, Rhebergen and van Zelst contributed to developing the idea and analysis plan for the study. Giltay performed statistical analyses and produced the figures. Van Zelst and Giltay wrote the initial draft of the manuscript, which was commented on and edited by all co-authors.

ACKNOWLEDGMENTS

The infrastructure for the NESDO study (https://nesdo.onderzoek.io/) is funded through the Fonds NutsOhra [project0701-065]; Stichting tot Steun VCVGZ, NARSAD The Brain and Behaviour Research Fund [grant ID 41080]; and the participating universities and mental health

care organizations (VU University Medical Centre, Leiden University Medical Centre, University Medical Centre Groningen, UMC St Radboud, GGZ inGeest, GGNet, GGZ Nijmegen, GGZ Rivierduinen, Lentis, and Parnassia). This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Denise C. R. van Zelst https://orcid.org/0000-0001-5267-8452

REFERENCES

1. Mitchell AJ, Subramaniam H. Prognosis of depression in old age compared to middle age: a systematic review of comparative studies. *Am J Psychiatry*. 2005;162(9):1588-1601. https://doi.org/10.1176/appi.ajp.162.9.1588
2. Schaakxs R, Comijs HC, Lamers F, Kok RM, Beekman ATF, Penninx BWJH. Associations between age and the course of major depressive disorder: a 2-year longitudinal cohort study. *Lancet Psychiatry*. 2018;5(7):581-590. https://doi.org/10.1016/s2215-0366(18)30166-4
3. Hegeman JM, de Waal M, Comijs H, Kok R, van der Mast R. Depression in later life: a more somatic presentation? *J Affect Disord*. 2015;170:196-202. https://doi.org/10.1016/j.jad.2014.08.032
4. Hegeman JM, Wardenaar K, Comijs H, de Waal M, Kok R, van der Mast R. The subscale structure of the Inventory of Depressive Symptomatology Self Report (IDS-SR) in older persons. *J Psychiatr Res*. 2012;46(10):1383-1388. https://doi.org/10.1016/j.jpsychires.2012.07.008
5. Krishnan KR, Hays JC, Tupler LA, George LK, Blazer DG. Clinical and phenomenological comparisons of late-onset and early-onset depression. *Am J Psychiatry*. 1995;152(5):785-788.
6. Schaakxs R, Comijs HC, Lamers F, Beekman ATF, Penninx BWJH. Age-related variability in the presentation of symptoms of major depressive disorder. *Psychol Med*. 2017;47(3):543-552. https://doi.org/10.1017/s0033291716002579
7. Fried EI. Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. *Front Psychol*. 2015;6:309. https://doi.org/10.3389/fpsyg.2015.00309
8. Association AP. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub; 2013.
9. Borsboom D. Psychometric perspectives on diagnostic systems. *J Clin Psychol*. 2008;64(9):1089-1108. https://doi.org/10.1002/jclp.20503
10. Cramer AO, Waldorp LJ, van der Maas HLJ, Borsboom D. Comorbidity: a network perspective. *Behav Brain Sci*. 2010;33(2-3):137-150. discussion 150-93. https://doi.org/10.1017/s0140525x09991567
11. Kendler KS. The phenomenology of major depression and the representativeness and nature of DSM criteria. *Am J Psychiatry*. 2016;173(8):771-780. https://doi.org/10.1176/appi.ajp.2016.15121509
12. Fried EI, Nesse RM. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR*D study. *J Affect Disord*. 2015;172:96-102. https://doi.org/10.1016/j.jad.2014.10.010
13. van Eeden WA, Hemert AM, Carlier IVE, Penninx BW, Giltay EJ. Severity, course trajectory, and within-person variability of individual symptoms in patients with major depressive disorder. *Acta Psychiatr Scand*. 2019;139(2):194-205. https://doi.org/10.1111/acps.12987
14. Cramer AO, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS One*. 2016;11(12):e0167490. https://doi.org/10.1371/journal.pone.0167490
15. Robinaugh DJ, Hoekstra RHA, Toner ER, Borsboom D. The network approach to psychopathology: a review of the literature 2008–2018 and an agenda for future research. *Psychol Med*. 2020;50(3):353-366. https://doi.org/10.1017/s0033291719003404
16. van Borkulo C, Boschloo L, Borsboom D, Penninx BWJH, Waldorp LJ, Schoevers RA. Association of symptom network structure with the course of [corrected] depression. *JAMA Psychiatry*. 2015;72(12):1219-1226. https://doi.org/10.1001/jamapsychiatry.2015.2079
17. Borsboom D, Deserno MK, Rhemtulla M, et al. Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*. 2021;1(1):58. https://doi.org/10.1038/s43586-021-00055-w
18. Parker RD, Flint EP, Bosworth HB, Pieper CF, Steffens DC. A three-factor analytic model of the MADRS in geriatric depression. *Int J Geriatr Psychiatry*. 2003;18(1):73-77. https://doi.org/10.1002/gps.776
19. Veltman EM, van Hulten S, Twisk J, et al. Differences in speed of response of depressive symptom dimensions in older persons during electroconvulsive therapy. *J ECT*. 2019;35(1):35-39. https://doi.org/10.1097/yct.0000000000000506
20. Fried EI, Nesse RM, Zivin K, Guille C, Sen S. Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. *Psychol Med*. 2014;44(10):2067-2076. https://doi.org/10.1017/s0033291713002900
21. van Loo HM, de Jonge P, Romeijn JW, Kessler RC, Schoevers RA. Data-driven subtypes of major depressive disorder: a systematic review. *BMC Med*. 2012;10(1):156. https://doi.org/10.1186/1741-7015-10-156
22. Wardenaar KJ, van Veen T, Giltay EJ, den Hollander-Gijsman M, Penninx BW, Zitman FG. The structure and dimensionality of the Inventory of Depressive Symptomatology Self Report (IDS-SR) in patients with depressive disorders and healthy controls. *J Affect Disord*. 2010;125(1-3):146-154. https://doi.org/10.1016/j.jad.2009.12.020
23. Boschloo L, van Borkulo CD, Rhemtulla M, Keyes KM, Borsboom D, Schoevers RA. The network structure of symptoms of the diagnostic and statistical manual of mental disorders. *PLoS One*. 2015;10(9):e0137621. https://doi.org/10.1371/journal.pone.0137621
24. Hebbrecht K, Stuivenga M, Birkenhager T, et al. Understanding personalized dynamics to inform precision medicine: a dynamic time warp analysis of 255 depressed inpatients. *BMC Med*. 2020;18(1):400. https://doi.org/10.1186/s12916-020-01867-5
25. Beijers L, Wardenaar KJ, van Loo HM, Schoevers RA. Data-driven biological subtypes of depression: systematic review of biological approaches to depression subtyping. *Mol Psychiatry*. 2019;24(6):888-900. https://doi.org/10.1038/s41380-019-0385-5
26. Bringmann LF, Vissers N, Wichers M, et al. A network approach to psychopathology: new insights into clinical longitudinal data. *PLoS One*. 2013;8(4):e60188. https://doi.org/10.1371/journal.pone.0060188
27. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: a tutorial paper. *Behav Res Methods*. 2018;50(1):195-212. https://doi.org/10.3758/s13428-017-0862-1
28. Haslbeck JMB, Bringmann LF, Waldorp LJ. A tutorial on estimating time-varying vector autoregressive models. *Multivar Behav Res*. 2021;56(1):120-149. https://doi.org/10.1080/00273171.2020.1743630
29. Bulteel K, Tuerlinckx F, Brose A, Ceulemans E. Clustering vector autoregressive models: capturing qualitative differences in within-person dynamics. *Front Psychol*. 2016;7:1540. https://doi.org/10.3389/fpsyg.2016.01540
30. van de Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proc Natl Acad Sci USA*. 2014;111(1):87-92. https://doi.org/10.1073/pnas.1312114110

31. Booij MM, van Noorden MS, van Vliet IM, et al. Dynamic time warp analysis of individual symptom trajectories in depressed patients treated with electroconvulsive therapy. *J Affect Disord*. 2021;293:435-443. https://doi.org/10.1016/j.jad.2021.06.068
32. Comijs HC, van Marwijk HW, van der Mast RC, et al. The Netherlands study of depression in older persons (NESDO); a prospective cohort study. *BMC Res Notes*. 2011;4(1):524. https://doi.org/10.1186/1756-0500-4-524
33. Vogelzangs N, Kritchevsky SB, Beekman ATF, et al. Obesity and onset of significant depressive symptoms: results from a prospective community-based cohort study of older men and women. *J Clin Psychiatry*. 2010;71(4):391-399. https://doi.org/10.4088/jcp.08m04743blu
34. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med*. 1996;26(3):477-486. https://doi.org/10.1017/s0033291700035558
35. Corruble E, Legrand J, Duret C, Charles G, Guelfi J. IDS-C and IDS-sr: psychometric properties in depressed in-patients. *J Affect Disord*. 1999;56(2-3):95-101. https://doi.org/10.1016/s0165-0327(99)00055-5
36. Trivedi MH, Rush AJ, Ibrahim HM, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med*. 2004;34(1):73-82. https://doi.org/10.1017/s0033291703001107
37. Ding H, Trajcevski G, Scheuermann P, Wang X, Keogh E. Querying and mining of time series data: experimental comparison of representations and distance measures. *Proc VLDB Endow*. 2008;1(2):1542-1552. https://doi.org/10.14778/1454159.1454226
38. Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans Acoust Speech Signal Process*. 1978;26(1):159-165. https://doi.org/10.1109/tassp.1978.1163055
39. Giorgino T. Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package. *Journal of Stat*. 2009;31(7):1-24.
40. Abdi H, Williams LJ, Valentin D, Bennani-Dosse M. STATIS and DISTATIS: optimum multitable principal component analysis and three way metric multidimensional scaling. *WIREs Comput Stat*. 2012;4(2):124-167. https://doi.org/10.1002/wics.198
41. Lorenzo-Seva U, Berge JMFT. Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*. 2006;2(2):57-64. https://doi.org/10.1027/1614-2241.2.2.57
42. Schuch JJ, Roest AM, Nolen WA, Penninx BW, de Jonge P. Gender differences in major depressive disorder: results from the Netherlands study of depression and anxiety. *J Affect Disord*. 2014;156:156-163. https://doi.org/10.1016/j.jad.2013.12.011
43. Marcus SM, Young EA, Kerber KB, et al. Gender differences in depression: findings from the STAR*D study. *J Affect Disord*. 2005;87(2-3):141-150. https://doi.org/10.1016/j.jad.2004.09.008
44. Lee CT, Leoutsakos JM, Lyketsos CG, Steffens DC, Breitner JCS, Norton MC. Latent class-derived subgroups of depressive symptoms in a community sample of older adults: the Cache County Study. *Int J Geriatr Psychiatry*. 2012;27(10):1061-1069. https://doi.org/10.1002/gps.2824
45. Mezuk B, Kendler KS. Examining variation in depressive symptoms over the life course: a latent class analysis. *Psychol Med*. 2012;42(10):2037-2046. https://doi.org/10.1017/s003329171200027x
46. Hybels CF, Landerman LR, Blazer DG. Latent subtypes of depression in a community sample of older adults: can depression clusters predict future depression trajectories? *J Psychiatr Res*. 2013;47(10):1288-1297. https://doi.org/10.1016/j.jpsychires.2013.05.033
47. Veltman EM, Lamers F, Comijs H, et al. Depressive subtypes in an elderly cohort identified using latent class analysis. *J Affect Disord*. 2017;218:123-130. https://doi.org/10.1016/j.jad.2017.04.059
48. Fried EI, Epskamp S, Nesse RM, Tuerlinckx F, Borsboom D. What are 'good' depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *J Affect Disord*. 2016;189:314-320. https://doi.org/10.1016/j.jad.2015.09.005
49. Gullion CM, Rush AJ. Toward a generalizable model of symptoms in major depressive disorder. *Biol Psychiatry*. 1998;44(10):959-972. https://doi.org/10.1016/s0006-3223(98)00235-2
50. Hegeman JM, Kok RM, van der Mast RC, Giltay EJ. Phenomenology of depression in older compared with younger adults: meta-analysis. *Br J Psychiatry*. 2012;200(4):275-281. https://doi.org/10.1192/bjp.bp.111.095950
51. Ulbricht CM, Chrysanthopoulou SA, Levin L, Lapane KL. The use of latent class analysis for identifying subtypes of depression: a systematic review. *Psychiatry Res*. 2018;266:228-246. https://doi.org/10.1016/j.psychres.2018.03.003
52. Fisher AJ, Reeves JW, Lawyer G, Medaglia JD, Rubel JA. Exploring the idiographic dynamics of mood and anxiety via network analysis. *J Abnorm Psychol*. 2017;126(8):1044-1056. https://doi.org/10.1037/abn0000311
53. Naarding P, Schoevers R, Janzing J, Jonker C, Koudstaal P, Beekman A. A study on symptom profiles of late-life depression: the influence of vascular, degenerative and inflammatory risk-indicators. *J Affect Disord*. 2005;88(2):155-162. https://doi.org/10.1016/j.jad.2005.07.002
54. Borsboom D, Cramer AO. Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol*. 2013;9(1):91-121. https://doi.org/10.1146/annurev-clinpsy-050212-185608
55. McNally RJ. Network analysis of psychopathology: controversies and challenges. *Annu Rev Clin Psychol*. 2021;17(1):31-53. https://doi.org/10.1146/annurev-clinpsy-081219-092850
56. Heino MTJ, Knittle K, Noone C, Hasselman F, Hankonen N. Studying behaviour change mechanisms under complexity. *Behav Sci*. 2021;11(5):77. https://doi.org/10.3390/bs11050077
57. Wright AGC, Woods WC. Personalized models of psychopathology. *Annu Rev Clin Psychol*. 2020;16(1):49-74. https://doi.org/10.1146/annurev-clinpsy-102419-125032
58. Myin-Germeys I, Kasanova Z, Vaessen T, et al. Experience sampling methodology in mental health research: new insights and technical developments. *World Psychiatry*. 2018;17(2):123-132. https://doi.org/10.1002/wps.20513
59. Hopwood CJ, Bleidorn W, Wright AGC. Connecting theory to methods in longitudinal research. *Perspect Psychol Sci*; 2021;17(3):884-894.
60. Hekler EB, Klasnja P, Chevance G, Golaszewski NM, Lewis D, Sim I. Why we need a small data paradigm. *BMC Med*. 2019;17(1):133. https://doi.org/10.1186/s12916-019-1366-x
61. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry*. 2006;163(1):28-40. https://doi.org/10.1176/appi.ajp.163.1.28

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: van Zelst DCR, Veltman EM, Rhebergen D, et al. Network structure of time-varying depressive symptoms through dynamic time warp analysis in late-life depression. *Int J Geriatr Psychiatry*. 2022;1-12. https://doi.org/10.1002/gps.5787