AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Psychopathology and Clinical Science

ISSN: 2769-7541

2024, Vol. 133, No. 2, 155–166
https://doi.org/10.1037/abn0000884

Depressive Symptoms as a Heterogeneous and Constantly Evolving Dynamical System: Idiographic Depressive Symptom Networks of Rapid Symptom Changes Among Persons With Major Depressive Disorder

Matthew D. Nemesure[1, 2], Amanda C. Collins[1, 3, 4], George D. Price[1, 2], Tess Z Griffin[1], Arvind Pillai[5], Subigya Nepal[5], Michael V. Heinz[1, 3], Damien Lekkas[1, 2], Andrew T. Campbell[1, 5], and Nicholas C. Jacobson[1, 2, 5, 6]

[1] Center for Technology and Behavioral Health, Geisel School of Medicine
[2] Quantitative Biomedical Sciences, Geisel School of Medicine
[3] Department of Psychiatry, Dartmouth-Hitchcock Medical Center
[4] Department of Psychology, Mississippi State University
[5] Department of Computer Science, Dartmouth College
[6] Department of Biomedical Data Science, Geisel School of Medicine

Major depressive disorder (MDD) is conceptualized by individual symptoms occurring most of the day for at least two weeks. Despite this operationalization, MDD is highly variable with persons showing greater variation within and across days. Moreover, MDD is highly heterogeneous, varying considerably across people in both function and form. Recent efforts have examined MDD heterogeneity byinvestigating how symptoms influence one another over time across individuals in a system; however, these efforts have assumed that symptom dynamics are static and do not dynamically change over time. Nevertheless, it is possible that individual MDD system dynamics change continuously across time. Participants (N = 105) completed ratings of MDD symptoms three times a day for 90 days, and we conducted time varying vector autoregressive models to investigate the idiographic symptom networks. We then illustrated this finding with a case series of five persons with MDD. Supporting prior research, results indicate there is high heterogeneity across persons as individual network composition is unique from person to person. In addition, for most persons, individual symptom networks change dramatically across the 90 days, as evidenced by 86% of individuals experiencing at least one change in their most influential symptom and the median number of shifts being 3 over the 90 days. Additionally, most individuals had at least one symptom that acted as both the most and least influential symptom at any given point over the 90-day period. Our findings offer further insight into short-term symptom dynamics, suggesting that MDD is heterogeneous both across and within persons over time.

Matthew D. Nemesure https://orcid.org/0000-0002-2369-600X

Matthew D. Nemesure is now at Digital Data Design Institute at Harvard Business School.

Matthew D. Nemesure and Amanda C. Collins contributed equally as cofirst authors.

Nicholas C. Jacobson has received a research grant from Boehringer-Ingelheim to study rapid symptom changes in schizophrenia. He receives no salary support or compensation for this work. This study was approved by the Dartmouth College Institutional Review Board (the Committee for the Protection of Human Subjects [STUDY00032081]). This study was not preregistered, and the data are not publicly available because of it being part of an ongoing data collection. The modified Patient Health Questionnaire-9 used in the current study is publicly available from the original methodological paper (Torous et al., 2015). Portions of this study and its results were presented at the 2022 Technology in Psychiatry Summit. A preprint of the first version of this article was published on PsyArXiv (https://psyarxiv.com/pf4kc/) and shared on Twitter. This work was supported by the National Institute on Drug Abuse (T32DA037202) and the National Institute of Mental Health and the National Institute of General Medical Sciences (R01MH123482). None of the other authors in this article have any conflicts of interest to disclose.

Matthew D. Nemesure served as lead for conceptualization, formal analysis, methodology, visualization, writing–original draft, and writing–review and editing and contributed equally to data curation and software. Amanda C. Collins contributed equally to formal analysis, methodology and visualization, writing–orginal draft, and writing–review and editing. George D. Price contributed equally to visualization, writing–original draft, writing–reviewing and editing. Tess Z Griffin served in a supporting role for data curation and contributed equally to project administration, writing–original draft, and writing–review and editing. Arvind Pillai contributed equally to data curation and software. Subigya Nepal contributed equally to data curation and software. Michael V. Heinz contribuetd equally to data curation and software. Damien Lekkas contributed equally to data curation and software. Andrew T. Campbell served as lead for suprevision and contributed equally to data curation and software. Nicholas C. Jacobson served as lead for data curation, investigation, and supervision and contributed equally to conceptualization and writing–reviewing and editing.

Correspondence concerning this article should be addressed to Matthew D. Nemesure, Digital Data Design Institute, Harvard Business School, Boston, MA, 02163, United States. Email: mnemesure@hbs.edu

General Scientific Summary

In the current study, we conducted a case series on five individuals with current major depressive disorder (MDD) who responded to a depression questionnaire three times a day and used a time-series methodology to determine, for each person, how each of the symptoms changed and influenced each other over the course of 90 days. Our findings indicate that MDD is not only heterogeneous in its manifestation between persons but can also be highly variable in presentation within an individual over time. As such, individuals with MDD who experience a more variable symptom presentation may benefit from personalized, time-sensitive clinical interventions.

Keywords: major depressive disorder, ecological momentary assessment, time series, network analysis, symptom dynamics

Supplemental materials: https://doi.org/10.1037/abn0000884.supp

Major depressive disorder (MDD) is one of the most common mental disorders and the leading cause of global disease burden (Smith, 2014). Unfortunately, current diagnostic conceptualizations are faced with several difficulties. First, while MDD is a complex, heterogeneous system with over 1,000 unique symptom presentations (Cramer et al., 2016; Fried & Nesse, 2015a), current diagnostic conceptualizations of MDD treat different MDD symptoms as interchangeable (e.g., meeting five of nine diagnostic criteria or using sum scores of individual symptoms of MDD occurring most of the day, nearly every day for at least 2 weeks to create a total score, reflecting overall depression severity). This conceptualization falls short of accurately capturing an individual's MDD given that presentations can vary substantially from person to person (Fried & Nesse, 2015a, 2015b), and viewing MDD as a sum score leaves out crucial information as to which symptoms are the most important and influential in a person's overall diagnostic presentation (Beard et al., 2016). A second problem with diagnostic conceptualizations is the assumption that within a major depressive episode, MDD is assumed to be chronic and unwavering (e.g., occurring "most of the day nearly every day for 2 weeks"). However, when persons with MDD are measured intensively within persons' daily lives, their symptoms are far from stable across days or weeks and in fact vary more substantially across hours within a day rather than across either weeks or months (Ebrahimi et al., 2021; Fried et al., 2022; Lorenz et al., 2020; Wichers et al., 2016, 2020).

Rather than using traditional diagnostic conceptualizations, MDD may be better conceptualized as a constantly evolving, complex system for individual persons. As such, MDD symptoms dynamically interact with each other such that one symptom can influence the development and maintenance of other MDD symptoms, thus contributing to the overall complex system of how MDD presents (Beard et al., 2016; Cramer et al., 2016). Moreover, symptom dynamics not only vary from person to person but can also vary widely within a single day for a given individual and on an hourly basis because of both internal (e.g., negative cognitions) and external factors (e.g., stressful life events; Bringmann et al., 2018). Thus, it is important to utilize methodological and statistical approaches to accurately capture the dynamic system of MDD.

Current Network Approaches to Symptom Variability

Ecological momentary assessments (EMAs) allow for researchers to collect several assessments throughout a single day for a longitudinal period and can be used to investigate short-term changes in symptom variability. Vector autoregressive (VAR) modeling is a statistical analysis that uses repeated measures regression to examine the temporal relationships between symptoms in psychopathology, including MDD (Jordan et al., 2020), and often includes data from EMAs. More recently, VAR modeling has been applied to network science to examine MDD as a symptom and investigate how symptoms change and influence each other over time (Epskamp et al., 2018; Jordan et al., 2020). When applied to network analysis, a statistical analysis used to examine interactive relationships between symptoms and how these relationships change over time (Borsboom & Cramer, 2013), VAR models are able to detect whether certain symptoms at one time point (e.g., insomnia) directly lead to increases or decreases of other symptoms (e.g., anhedonia) or the same symptom (e.g., insomnia) at the next time point.

Prior research using VAR models to investigate MDD symptom network dynamics has commonly made use of graphical ($N = 1$) or multilevel ($N > 1$) methods. Employment of these particular VAR models, however, may not provide an accurate picture of how symptoms dynamically fluctuate over time. First, the multilevel VAR model takes a group-level approach to examine network structures over time and cannot provide information as to how symptoms change at an individual level. Alternatively, graphical VAR takes an idiographic approach; however, relatively few studies have utilized this modeling for depression (de Vos et al., 2017; Kaiser & Laireiter, 2018; Wichers et al., 2016, 2020, 2021), and further research is needed to validate its ability to detect symptoms changes over time within a single individual. Second, both graphical and multilevel VAR models assume that the relationships between variables (e.g., insomnia and anhedonia) are static across time and that dynamic changes do not occur between variables across time (Bringmann et al., 2015; Haslbeck et al., 2021; Jordan et al., 2020; Lütkepohl, 2005), neglecting the possibility that MDD symptom networks themselves change over time. Thus, prior work has not yet been able to accurately capture the complex, dynamic system of MDD and has instead investigated MDD as a complex, static system.

Time-Varying VAR Modeling

In order to help address this inability to fully capture the dynamic nature of MDD symptoms across time, a time-varying auto- and cross-regressive modeling framework can be utilized. The major benefit of this approach is that it allows for both autoregressive (e.g., X at T1 impacts X at T2) and cross-regressive relationships

(e.g., X at T1 impacts Y at T2) to change with time (i.e., model nonstationary processes), and better reflects the reality of MDD in that it is still capable of capturing static symptom dynamics over time (which may be the case for some individuals) but can also model fluctuating dynamics as both internal and external factors arise.

Generalized additive models (GAMs) are uniquely poised to handle this modeling framework as they allow for nonlinear smooths to estimate changing coefficients over time (Hastie & Tibshirani, 1995). Specifically, GAMs do not have the assumption that all time points are equally spaced, allowing for the time in between each EMA to vary and not be fixed. When applied in a single-lag auto- and cross-regressive approach, this methodology allows us to estimate the changing relationship between each symptom of MDD, as measured by the Patient Health Questionnaire-9 (PHQ-9; Kroenke & Spitzer, 2002), and every other symptom (including the symptom of interest) at the following time point. For example, the relationship between anhedonia and itself (autoregressive) at the prior time point can be different on Day 2 than it is on Day 9, and there is a smoothed trajectory for this changing association over time. This same idea can also be applied to any cross-regressive association. For example, the association between concentration difficulties and psychomotor difficulties can be quite large at Time Point 15 and quite small at Time Point 30, and while this changing association may be because of changing external or internal factors, most importantly, this approach allows for capturing these changing dynamics.

Rationale

As previously noted, prior research has investigated MDD symptoms over time using a stationary approach (e.g., graphical and multilevel VAR), thus potentially leaving out important information about the heterogeneity of MDD symptoms and how they can dynamically change and influence each other over time. Moreover, relatively few studies have examined MDD symptoms with an idiographic network analysis approach.

There is one existing study that has examined the dynamic changes of MDD with a time-varying VAR approach (Siepe et al., 2022). Their findings indicate that MDD symptoms can substantially vary both across and within persons over time, providing further evidence highlighting the importance of investigating MDD as a nonstationary, dynamic system. While Siepe et al. (2022) uncovered important information about MDD symptom dynamics, they only assessed two depressive symptoms per day for 20 individuals. In addition, while they were able to examine variability across days, it is also important to examine variability within days given that symptoms can change over the course of hours (Ebrahimi et al., 2021; Fried et al., 2022; Lorenz et al., 2020; Wichers et al., 2016, 2020). In addition, using time-varying VAR modeling to examine the dynamics of all MDD symptoms may provide added insight into how symptoms influence each other over time. Siepe et al.'s (2022) analyses also utilized block bootstrapping methods to assess the stability of their estimates, which suggested that the parameter estimates were unstable; however, this method may not be suitable to use with time-varying coefficient models.

Thus, the purpose of the current case series is to examine the dynamics of MDD symptoms within individual persons using a time-varying VAR approach from EMA data. We first evaluate overall MDD symptom variability across the entire sample ($N = 105$), and then we provide a more detailed investigation of these dynamics in a smaller case series of selected participants ($n = 5$). Given that traditional bootstrapping methods, which are often used to validate network models, are not suitable for use with time-varying VAR, we have also included a qualitative analysis of the five individuals' daily diary entries to both better capture their lived experiences and to see how these entries map onto their MDD symptoms throughout the 90 days.

Method

Procedure

As a part of the Tracking Depression Study, an R01 study funded by the National Institute of Mental Health and the National Institute of General Medical Sciences, individuals over 18 years of age with current MDD were recruited remotely across the United States via Google Ads. Participants were required to use an Android-based phone as their primary mobile device. They were screened for current MDD and exclusion criteria through online surveys and virtual interviews, allowing for mental health assessments and for collection of demographic information. The mental health assessment included a Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition, which was administered by a board-certified psychiatrist.[1] Individuals were excluded from participation if, at any point during the screening process, they endorsed active suicidality, current or past psychotic symptoms or bipolar disorder, or if they did not meet criteria for a major depressive episode within the past 30 days. This study was approved by the Dartmouth College Institutional Review Board (the Committee for the Protection of Human Subjects [STUDY00032081]), and participants were asked to provide both written and verbal consent prior to taking part in study.

Following screening, qualifying participants were asked to install the smartphone application, MLife, on their Android device. MLife is a mobile sensing application developed to collect passive sensing and EMA data (R. Wang et al., 2014). Participants were instructed to keep MLife running throughout the 90-day study period and were prompted three times a day by the application to answer an EMA (i.e., a short survey), which included questions about depressive symptoms and a diary entry. EMA notifications were delivered starting 4 hr after participant self-reported wake time (morning EMA) and at 4-hr intervals thereafter (afternoon and evening EMAs), with a total of 270 EMA prompts per participant over the 90 days. Upon study completion, participants were compensated $1 per EMA completed.

PHQ-9

At each EMA, participants completed a modified version of the PHQ-9, a validated measure used to assess depression severity (Kroenke & Spitzer, 2002; Torous et al., 2015). In the current study, we utilized a modified PHQ-9 to make the EMA questions

[1] A board-certified psychiatrist completed the SCID-5 with each participant; however, inter-rater reliability was not assessed in the current study. Although the same individual administered the SCID-5 with all participants, which likely resulted in consistent clinical decisions, we were not able to assess whether these decisions would be consistent across independent clinicians.

more mobile-friendly (see the online supplemental materials). Participants were asked to use a sliding scale (rather than the original four-item Likert scale) to select a value ranging from 0 to 100 that best reflected how they felt, as done in prior work (Torous et al., 2015). Participants were asked to think of the sliding scale as ranging from a day in their past when the relevant question was not an issue at all (i.e., *0* or *not at all*) to a day in their past when the relevant question was the most applicable (i.e., *100* or *constantly*) and to assess how they had felt over the past 4 hr within this range.

Diary Entry Question

Participants were presented with an optional diary entry question at the end of each EMA, which prompted participants to describe how they had felt over the past 4 hr and why. Thus, participants were given the opportunity to provide diary entries up to 3 times per day, and they were provided the option of responding through either video, audio, or text (maximum of 2,500 characters). Participants were encouraged to use this space to either write about their emotions or provide any other relevant information, including changes in medication or therapy.

Participants

At the time of this analysis, the Tracking Depression Study had a total of 105 participants who had completed the EMA portion of the larger trial and were included in the present analyses (86% women, 10% men, 2% nonbinary, and 2% other; 82% White, 5% Asian, 8% Black, 2% American Indian or Alaskan Native, and 3% other; 92% non-Hispanic and 8% Hispanic). The five participants included in the current analysis were selected based on their number of diary entries, and those with the greatest number of diary entries were chosen to allow for a more robust qualitative validation of model findings. The five illustrative participants met the criteria for MDD via the structured clinical interview for *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition and were between the ages of 20–40 years old, with the majority of participants identifying as women (80% women, $n = 4$) and non-Hispanic White (80%, $n = 4$; see Table 1 for individual demographic data).

Data Preparation

Data used for this analysis consisted of all PHQ-9 EMA data collected over the course of the 90-day study except for the question related to sleep difficulties, which we excluded based on it only being presented to participants once a day rather than three times per day. EMA entries were either fully completed or did not exist, and thus there were no EMAs that contained partial data. In this study, we evaluated symptom variability across all 105 participants as well as conducted a more nuanced investigation of five of these participants, selected based on their having contributed the greatest number of daily diary entries to allow for thorough qualitative validation of the modeling outputs. Across these five participants, they each answered 277, 266, 273, 216, and 168 EMAs, respectively,[2] and no participant went more than 3 days without answering a survey across the 90 days of the study. From this EMA data, eight data tables were generated per person, with each data table representing one MDD symptom as the outcome and all eight symptoms at a lag of one EMA as the predictors. These data tables had $(t-1)$ rows with one row for each EMA excluding the first because there are no lag one predictors at $t = 1$.

Modeling Approach

Following the data preparation, eight GAMs, representing each of the eight MDD symptoms as an outcome, were fit per person ($N = 5$) and included in the case series, resulting in 40 total GAMs. An example GAM formula looks as follows for the prediction of PHQ-9 Question 1 (Q1; having little interest or pleasure), where the remaining PHQ-9 questions Q1–Q8 (excluding the sleep question) at the time prior to the outcome $(t-1)$ are the predictors.

$$Q1_t = \beta_0 + f(t) * Q1_{t-1} + f(t) * Q2_{t-1} + \cdots + f(t) * Q8_{t-1} + \epsilon,\ \epsilon \sim N(0, \sigma^2). \quad (1)$$

In this formula, we evaluate the linear relationship between each MDD symptom as it predicts every other MDD symptom as a (non) linear smooth function (f) of time. Additionally, we add an L1 penalty term allowing predictors to be penalized to zero (Wood, 2017). This penalization was put in place to prevent spurious results from estimates with high variability. From each of these models, we obtained a coefficient for each lagged predictor at each time point (each EMA), and these changing coefficients represented the dynamic, directional relationships between the lagged predictor and the outcome for any given symptom.

Model Outputs and Evaluation

Given this modeling approach, the per-person outputs consisted of a coefficient for each symptom as it predicted every other symptom, including itself, at the next time point. These results were then made into an adjacency matrix of coefficients for each time point (EMA). From these adjacency matrices, directed networks could then be generated with nodes representing symptoms and edges representing the association for how well the starting node predicted the receiving node at the next time point. The primary outcome of interest was the changing outdegree for each node in the network. In this case, outdegree was essentially the sum of the absolute value of the coefficient for a given lag one predictor across each outcome. This value represented how influential a given symptom was on all other symptoms, including on itself at the following time point. With outdegree, we can evaluate the changing influence over time of any given symptom on all other PHQ-9 measured components of MDD.

This approach is similar to what would typically be seen in a Gaussian graphical model where nodes are represented by symptoms and edges are represented as partial correlations between symptoms across individuals (Epskamp et al., 2018; Yuan & Lin, 2007). As noted above, typically these models are validated via bootstrapping to assess whether the partial correlations persist over bootstrap iterations. Unfortunately, with the idiographic approach, there is no appropriate method to bootstrap over time points and thus a quantitative validation becomes implausible.

[2] Note that surveys beyond the 270 required were due to participants' entering surveys before the official start date or completing more than the required number of surveys in a given day. The internal consistency for these surveys across all participants was .765.

Table 1
Participant Demographics

Demographics	Overall sample	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5
Gender	86% Women	Transgender woman	Woman	Woman	Man	Woman
Race	82% White	White	Asian	White	White	White
Ethnicity	92% Non-Hispanic	Non-Hispanic	Non-Hispanic	Non-Hispanic	Hispanic	Non-Hispanic
Treatment	—	None	None	None[a]	Psychotherapy/antidepressant[b]	Antidepressant

Note. Overall sample includes 105 participants; psychotherapy indicates that the participant was seeing a mental health clinician for therapy (i.e., not medication management). Antidepressant indicates that the participant was taking a psychotropic medication for MDD (i.e., selective serotonin reuptake inhibitor or serotonin norepinephrine reuptake inhibitor). MDD = major depressive disorder.
[a] Participant 3 started an SSRI in April. [b] Participant 4 started esketamine in March.

To address this issue, we chose to perform a qualitative evaluation using written diary entries from the participants. All entries were aggregated into a single data file, which consisted of the participant number, date, and text from the diary entry. A trained predoctoral clinical psychology intern read through each participant's diary entries in this file and qualitatively evaluated how well they corresponded to the dynamic symptom fluctuations. Specifically, the intern searched for diary entries that were completed at the time in which an individual had a symptom shift (e.g., around September 20 for Participant 1). If relevant diary entries were not available at times in which symptoms shifted, the intern next selected diary entries that were relevant to when a given symptom demonstrated high outdegree (i.e., around March 29 for Participant 4). Entries that corresponded temporally to shifts in symptomatology or elevated symptom outdegree were noted and mapped back to the model outputs to qualify the modeling results. The intern selected an equal number of diary entries for each individual (i.e., three entries each); however, we included a fourth entry for Participant 1 given the number of times their most influential changed. All coauthors involved in the modeling agreed with the diary entry selections.

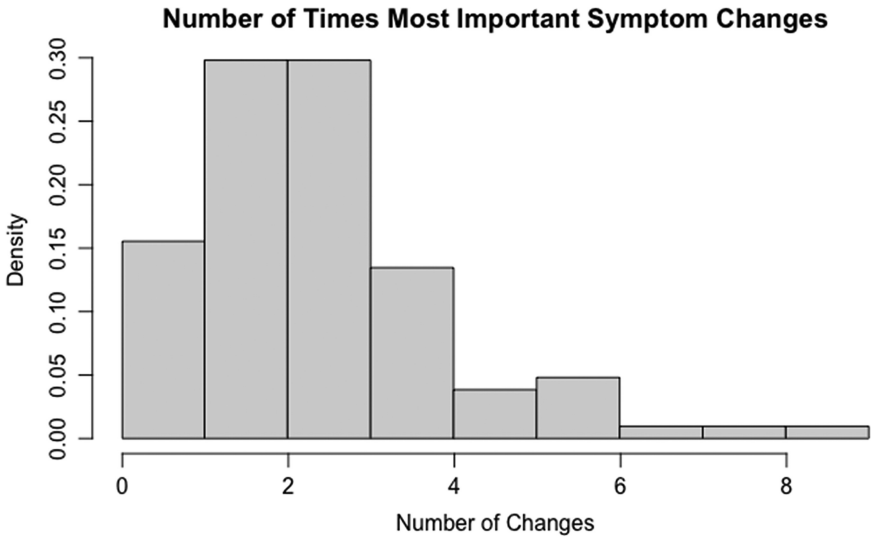
Transparency and Openness

The purpose of the current study was to examine how the dynamics of MDD symptoms within individual persons vary across time and was exploratory in nature; thus, we did not have any explicit hypotheses for the current study. This study was not preregistered, and the data is not publicly available because of it being part of an ongoing data collection. The modified PHQ-9 used in the current study is publicly available from the original methodological paper (Torous et al., 2015).

Results

We first evaluated the distribution of symptom dynamic variability in the overall sample ($N = 105$). In this broader analysis, and as a means of simplifying a more complex set of results, symptom dynamic variability was defined as "the number of times the most influential feature changes" (see Figure 1 for symptom dynamic variability distribution). The average number of times the most predictive feature changed for an individual across the 90 days was 1.886

Figure 1
*Distribution of Symptom Dynamic Variability for Larger Study (*N = *105)*



Note. This histogram shows the distribution for the number of times the topmost predictive symptom changed over the course of the study for the first 105 individuals to complete the study.

times, with a median of 2 times. Furthermore, 89 out of the 105 individuals included in the study had their most influential symptom change at least once over the 90-day study period. To better understand if and how the outdegree coefficients appropriately represent the individuals' reported symptom severity, we calculated a correlation between the coefficient and reported symptoms for each individual (i.e., one correlation for each symptom). The z-scores of these correlations were then averaged across persons and converted back to Pearson's r. These results are listed in Table S1 and indicated a small, but positive, correlation ($r = .17–.38$) for each symptom, suggesting that the individuals' calculated metrics are positively related to their ratings of each symptom on the PHQ-9.

The primary idiographic results of this analysis are the changing symptom dynamics across the five participants selected for this case series. These are represented in Figures 2–6, which display the changing outdegree for each symptom as it predicts all other symptoms at the next time point. Through this, we can evaluate the changing influence of a given symptom on their depression profile as well as the overall variability of an individual's depression dynamics and presentation. Note, an even more nuanced representation of symptom dynamics is reported in the online supplemental materials. These animated gifs capture all symptom-to-symptom relationships over time instead of simply taking the sum of a symptom's influence.

As noted above, traditional bootstrapping methods are not suitable methods to validate models that use time-varying VAR. Thus, we used the empirical standard error to estimate the confidence intervals for the individual coefficient estimates for each symptom of the five participants. These figures are in the online supplemental materials and provide further information regarding the validity of our findings.

Participant 1

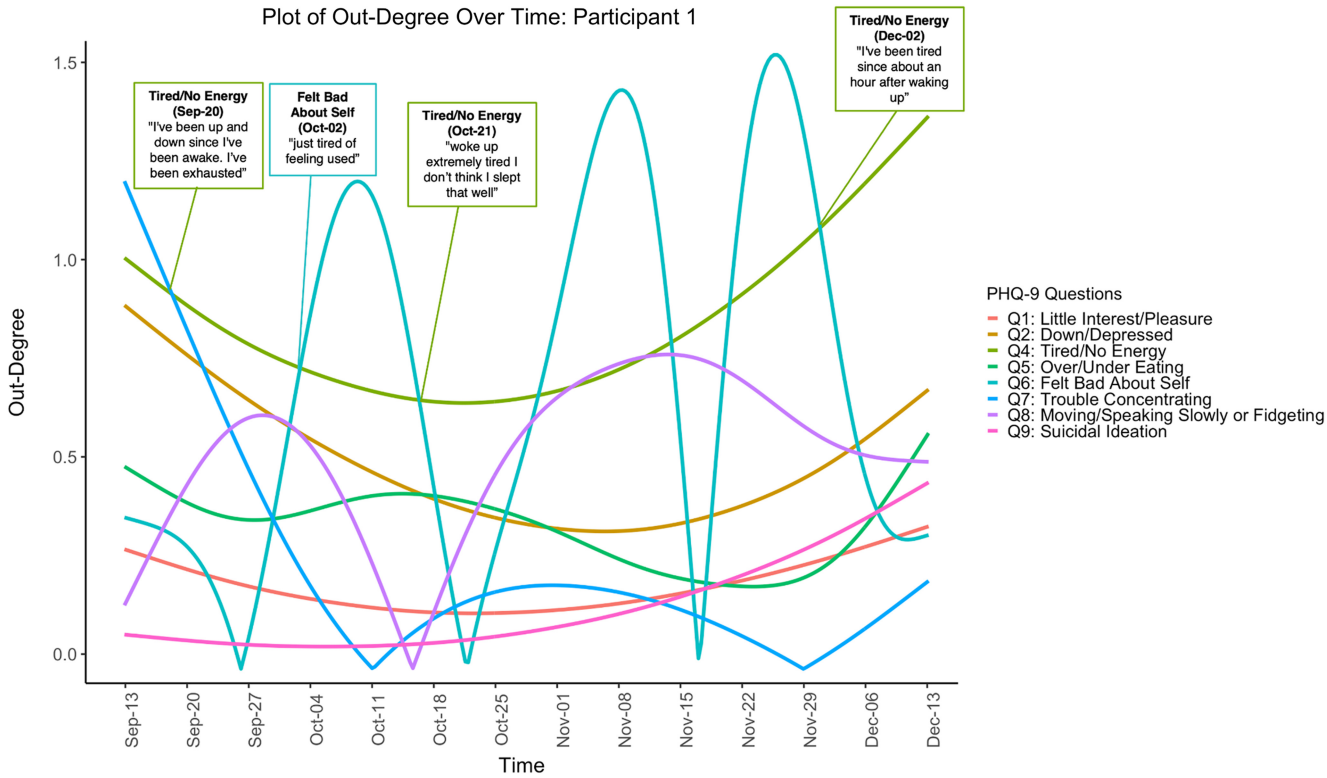
Participant 1 is a White, non-Hispanic transgender woman in the 20–40 age range. Of the five participants, Participant 1 had the most variable symptom profile in this case series as evidenced by their top influential symptom changing seven times (see Figure 2). Across the eight measured symptoms, this participant had three unique symptoms that were the most influential, each for a given period of time. Of note, Participant 1 seemed to experience a periodic effect for feeling bad about the self, where this symptom varied from high to low importance in an oscillatory manner. In addition, fatigue seemed to maintain relatively high importance, and it was the most influential symptom profile component when prior high-impact symptoms (e.g., feeling bad about the self) had dampened effects.

Participant 2

Participant 2 is an Asian, non-Hispanic woman in the 20–40 age range. The majority of this participant's symptom profile was influenced by anhedonia and depressed mood, and they demonstrated only one

Figure 2

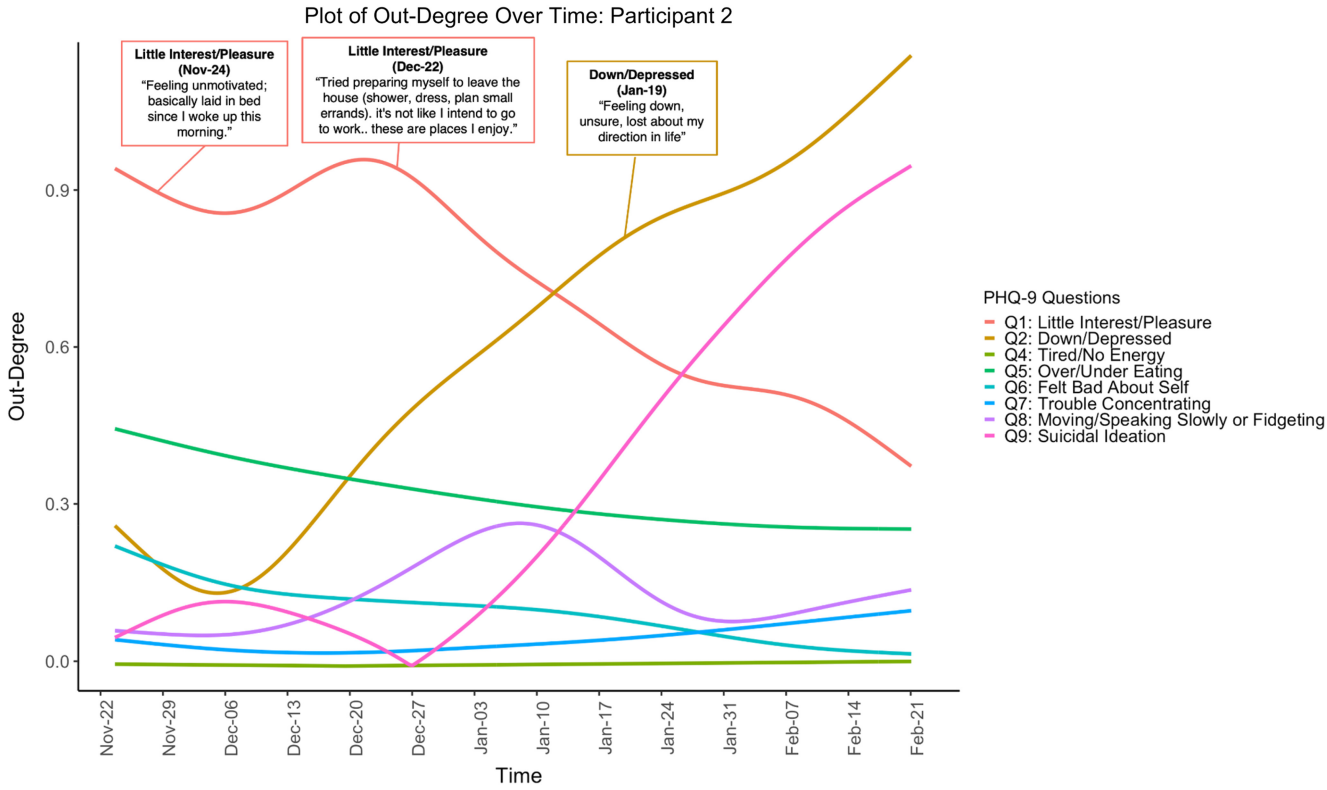
Changes in Symptom Dynamics for Participant 1



Note. This figure shows the sum of the absolute value of outdegree for each symptom as it predicts every other symptom at the next time point. As an example, the value for Q1 (little interest/pleasure) for the last survey on September 13 is the sum of how predictive it is for all measured symptoms of the first survey on September 14. See the online article for the color version of this figure.

Figure 3

Changes in Symptom Dynamics for Participant 2



Note. This figure shows the sum of the absolute value of outdegree for each symptom as it predicts every other symptom at the next time point. As an example, the value for Q1 (little interest/pleasure) for the last survey on November 29 is the sum of how predictive it is for all measured symptoms of the first survey on November 30. See the online article for the color version of this figure.

change in their top influential symptom (see Figure 3). Toward the end of the 90-day period, suicidal ideation quickly became a larger influence on other symptoms despite starting as a symptom with a low amount of influence. The remaining PHQ-9 symptoms were relatively static with respect to their influence on this participant's MDD dynamics.

Participant 3

Participant 3 is a White, non-Hispanic woman in the 20–40 age range. Their top influential symptom changed twice (see Figure 4). For the first 2 months of the study, the symptom of suicidal ideation was the primary contributor to their symptom dynamics, increasing in influence for the first month and, while still dominant over the influence of other symptoms, slowly decreasing in importance for the second month. Throughout the 90 days, many symptoms dynamically changed in their influence, including fatigue, psychomotor difficulties, and feeling bad about the self. In the final month, psychomotor difficulties emerged as an important driver of other symptoms, and then suicidal ideation increased in its influence in the last week.

Participant 4

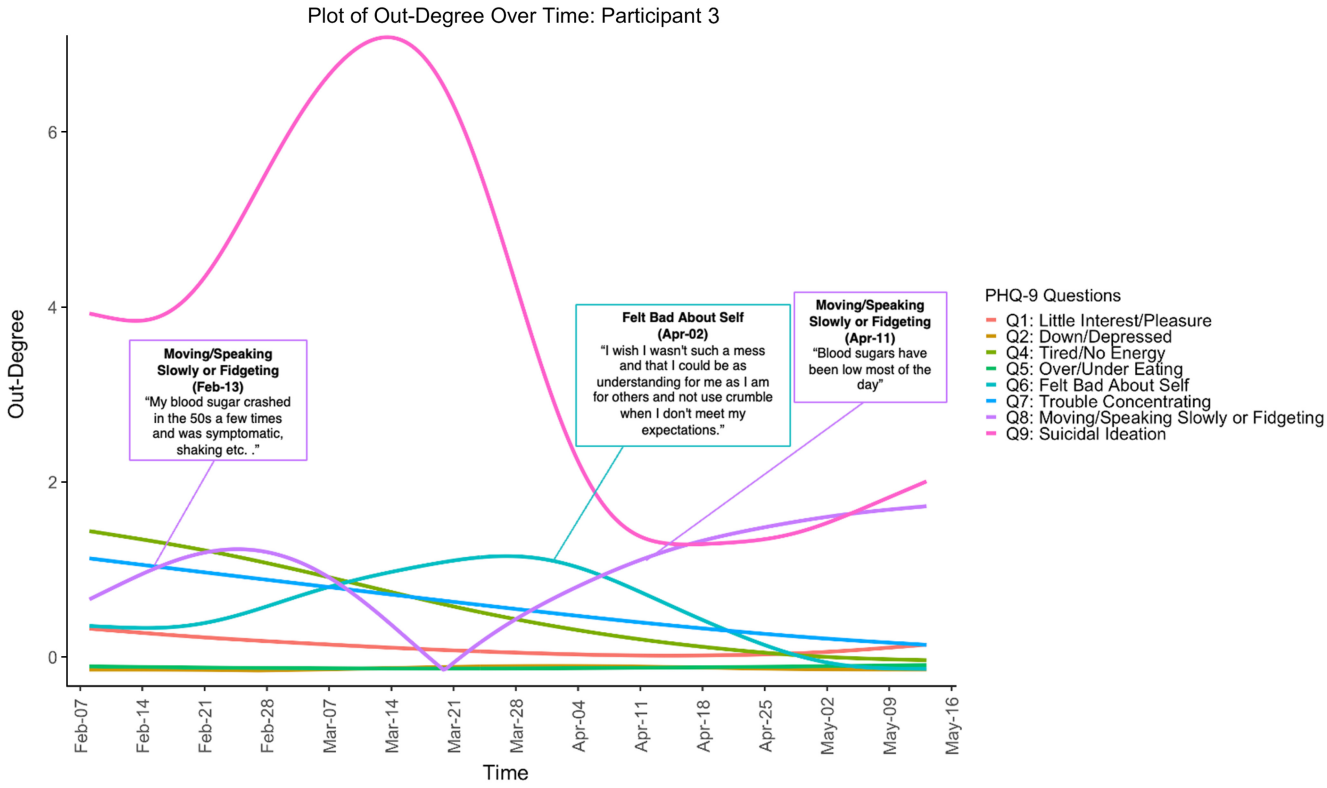
Participant 4 is a White, Hispanic man in the 20–40 age range. This participant had the least variable depression profile with no changes in their top influential symptom, depressed mood, which was maintained throughout the duration of the study (see Figure 5). Beyond this, however, all other symptoms also exhibited a relatively static level of predictiveness across the study. Anhedonia was initially an influential symptom, but it steadily decreased to a minimal outdegree, and then it increased slightly again. Feeling bad about the self also steadily increased throughout the study and remained the second most influential symptom throughout most of the 90 days. Participant 4's symptom variability profile exemplifies how MDD has been broadly defined in the past as a relatively static combination of symptoms. In context with the four other participants studied, this finding highlights the inherent flexibility of this methodology to capture not only dynamic MDD symptom fluctuations but also more consistent MDD experiences.

Participant 5

Participant 5 is a White, non-Hispanic woman in the 20–40 age range. Their top influential symptom changed three times (see Figure 6). At the outset of the study, the symptom of suicidal ideation seemed to drive their overall depression characterization. Over the first half of the study, however, this symptom influence decreased, followed by a sustained lack of impact starting midway into the study. Instead, feeling bad about the self and fatigue became the primary drivers of this participant's symptom profile across the second part of the study.

Figure 4

Changes in Symptom Dynamics for Participant 3



Note. This figure shows the sum of the absolute value of outdegree for each symptom as it predicts every other symptom at the next time point. As an example, the value for Q1 (little interest/pleasure) for the last survey on February 21 is the sum of how predictive it is for all measured symptoms of the first survey on February 22. See the online article for the color version of this figure.

Comparisons Between Time-Varying Networks and Qualitative Data

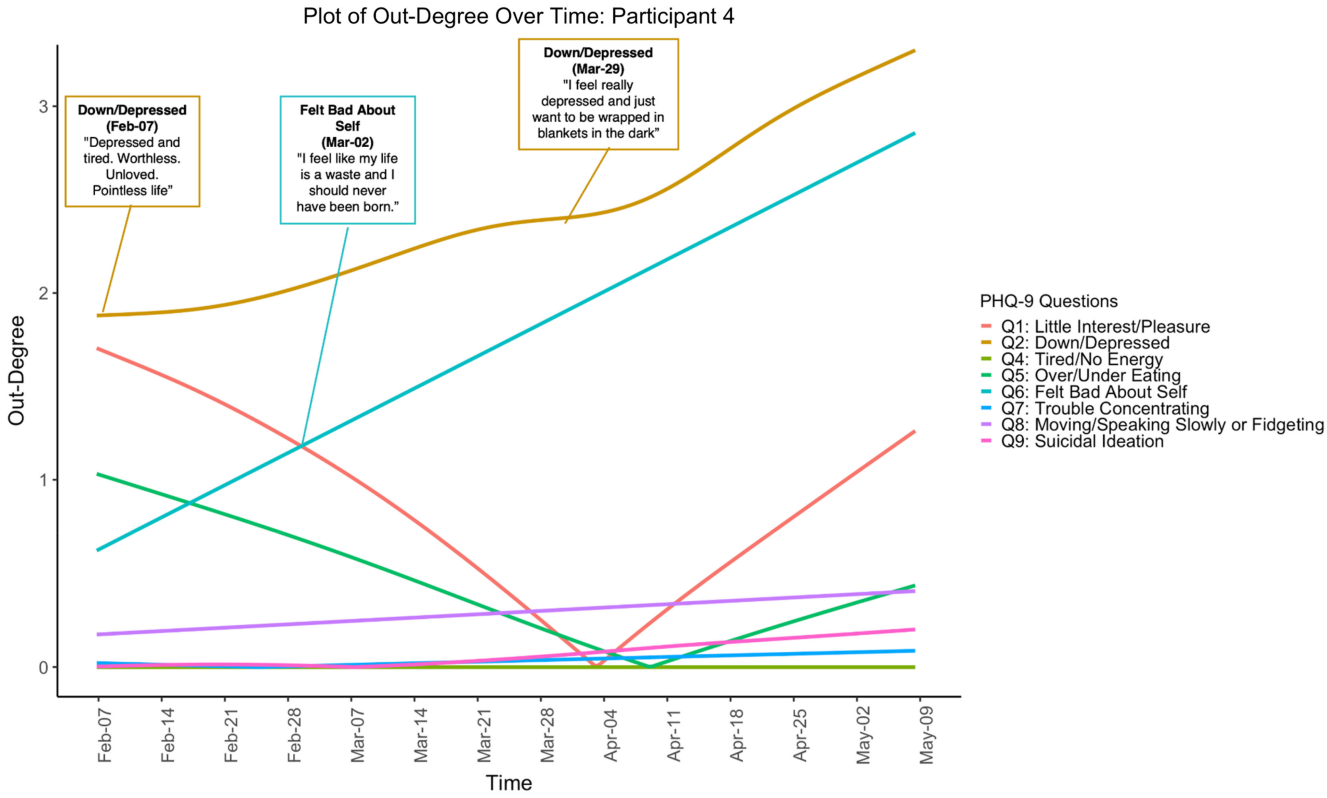
As noted above, we included selective diary entries to further validate the changes in symptom networks and investigate internal and external factors that may have influenced these changes. When examining this qualitative data (i.e., diary entries), a common theme emerged such that participants often wrote about symptoms, thoughts, or behaviors that were related to the most influential depressive symptom at the time instead of directly writing about the actual depressive symptom. Thus, it is likely that some of these depressive symptoms are capturing more than the symptom itself, including anxiety and somatic symptoms. For example, Participant 1 provided many diary entries about not getting enough sleep, feeling tired, or having a "busy day." Indeed, fatigue was one of their most influential symptoms, so it is likely that her fatigue is driven by lack of sleep and having busy days because of work and taking care of her child. Participant 3 provided more diary entries regarding her physical activity (e.g., exercising more) and medical problems (e.g., blood sugar decreasing) and increased anxiety throughout her 90 days. Participant 4 did not have any changes in outdegree in the network as depressed mood remained the most influential symptom in the network across the 90 days. This may have been exacerbated because of them ruminating daily on negative aspects of their life, including a recent breakup, hopelessness, and worthlessness.

Discussion

In the current study, we conducted a novel investigation of the dynamics of depressive symptoms within 105 participants over the course of 90 days using EMA data and a time-varying VAR approach and used five individuals as exemplars to illustrate this approach. In line with prior research, our results indicate that there is high heterogeneity across persons, such that the individual network composition is unique from person to person (de Vos et al., 2017; Kaiser & Laireiter, 2018; Siepe et al., 2022). Moreover, our results show that for most persons, individual depressive symptom networks can change dramatically in form across a 3-month period, as evidenced by some participants exhibiting significant variability within their symptom networks. Further investigation of symptom changes in the larger sample ($N = 105$) also revealed heterogeneity across persons, as evidenced by variability across the sample in the number of times that the most influential sample changed for a given individual (i.e., 0–8 times). Within the larger sample, 85% of individuals had their top symptom change at least once and 54% had this occur more than once. Furthermore, 46% of individuals had at least one symptom be both the most influential and least influential at some point over the course of the 90 days, and 70% of individuals had their most influential symptom fall into the bottom half. Taken together, our findings suggest that the dynamics of depressive symptom networks vary from person to person and are highly variable across time.

Figure 5

Changes in Symptom Dynamics for Participant 4



Note. This figure shows the sum of the absolute value of outdegree for each symptom as it predicts every other symptom at the next time point. As an example, the value for Q1 (little interest/pleasure) for the last survey on February 21 is the sum of how predictive it is for all measured symptoms of the first survey on February 22. See the online article for the color version of this figure.

Clinical Implications

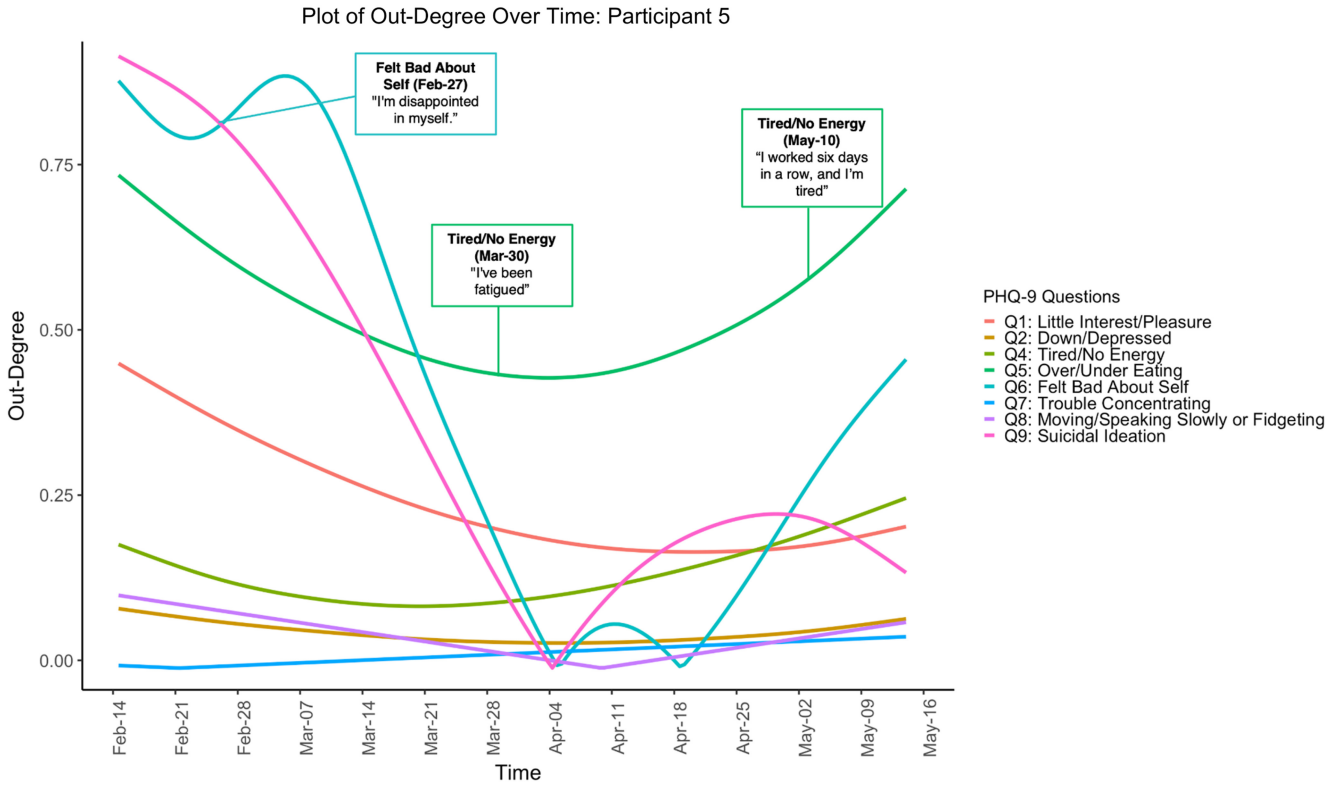
Our findings hold important clinical implications for treatment as well. The field of network science has thus far provided important information about the development and maintenance of depressive symptoms. If reflecting causal relationships, centrality measures (e.g., outdegree) can give us information about which depressive symptoms are the most influential over others in a network and potentially suggest which symptoms can serve as important targets for clinical interventions. For example, for individuals where anhedonia emerges as the most impactful symptom, interventions targeting this symptom (e.g., positive affect treatments) may be more beneficial than other treatments (Wichers et al., 2021). As currently explored with graphical and multilevel VAR models, the symptom that emerges as the most impactful may indicate that this symptom is a risk factor and an important intervention target overall, but these models do not assess time-sensitive changes in symptom dynamics and intervention needs. As evidenced by our findings, MDD is better represented as a heterogeneous, dynamic system, given that, for some individuals, symptoms and symptom dynamics change dramatically across time. Moreover, symptoms also dynamically change during treatment, often a result of direct therapeutic change. Thus, investigating depressive symptoms with a dynamic, time-varying approach may provide better information as to how symptom dynamics change over time in response to psychological and pharmacological therapies (Bringmann, 2021; Bringmann et al., 2017) and may also help bridge the gap between network science and clinical practice for providing personalized therapeutic care based on person-specific networks.

Based on the differences between the five case studies presented here and prior research (Fisher, 2015; Jacobson & Nemesure, 2021), individuals likely benefit from different treatments depending on their initial presentation. For example, anhedonic depressed individuals may benefit more from positive affect treatments, and primarily depressed individuals may benefit more from cognitive-behavioral therapy. Thus, taking a "one-size-fits-all" treatment approach across individuals can be potentially problematic and ineffective. Additionally, our findings indicate that the dynamic nature of depressive symptoms may be better suited for interventions that are more time-sensitive and fluid rather than traditional, weekly in-person interventions. Thus, a one-size-fits-all treatment approach within an individual may also be potentially problematic as a patient's therapeutic needs will most likely fluctuate over time in response to treatment or other internal (e.g., negative cognitions) or external factors (e.g., stressful life events).

Fortunately, digital interventions represent a growing field in the literature, with several interventions currently in use for MDD (Moshe et al., 2021). Digital interventions offer an advantage over

Figure 6

Changes in Symptom Dynamics for Participant 5



Note. This figure shows the sum of the absolute value of outdegree for each symptom as it predicts every other symptom at the next time point. As an example, the value for Q1 (little interest/pleasure) for the last survey on February 21 is the sum of how predictive it is for all measured symptoms of the first survey on February 22. See the online article for the color version of this figure.

traditional in-person interventions as they are often cheaper, less time consuming, and available in the moment to individuals (Wilhelm et al., 2020). Given the range in variability of symptom changes from person to person, those who experience greater fluctuations in symptoms may benefit more from digital interventions that can be used in the moment than weekly in-person interventions. Just-in-time, adaptive interventions in particular can be utilized for those individuals whose symptoms tend to change dynamically over the course of hours or days (Teepe et al., 2021; L. Wang & Miller, 2020). Thus, being able to monitor individuals' symptom dynamics over time and implement just-in-time, adaptive interventions in response to specific symptom changes may help advance personalized treatment.

Limitations

Although our findings provide important, novel information as to how depressive symptom networks vary on an idiographic level, there are several limitations of the current study. First, because of space constraints, we were unable to include all participants in the current study and consequently only selected the five individuals with the most written diary entries to include in our qualitative analyses and illustration. While these individuals were more inclined to write diary entries and may not have been representative of the broader population, we picked them specifically so that we could validate whether the modeling approach was accurately detecting symptom changes. Moreover, given the time-series nature of the data, a quantitative validation would not have worked as bootstrapping is not suitable with an idiographic approach; thus, we estimated the confidence intervals for our estimates based on the empirical standard error. However, despite the selection process as a potential limitation, the symptom variability for these five individuals proved representative of the range of variability for all participants in the sample (i.e., Participant 1 had significant variability in symptoms over time, and Participant 4 had no variability).

Second, given the nature of the time-varying VAR model, we were unable to include the symptom related to sleep difficulties as this symptom was only measured once per day (compared to three times per day for all other symptoms). Thus, it is possible that excluding this item impacted the variability of symptoms overall for some individuals. For example, sleep difficulties could indeed be the most influential for some individuals; however, we were unable to capture this phenomenon with the current sampling framework and inherent missingness (Bringmann, 2021).

Third, we recruited participants online via Google Ads, allowing us to sample participants more representative of the general population within the United States than had we used a community or

clinical sample (e.g., from a local hospital). Given that we did not recruit from patients in a hospital or outpatient clinic, it is unclear whether our sample extends to a more specific clinical, treatment-seeking sample of depressed individuals. However, three participants endorsed receiving treatment for MDD (i.e., psychotherapy and/or psychotropic medication) at some point during their 90 days in the study; thus, it is possible that we would see similar results if investigated within a clinical setting.

Fourth, only a lag of one was assessed in the present study, and our time frame between lags was consistently 4 hr. We chose to assess this lag because of its interpretability and a lack of methodology to determine other optimal lags in this framework. Future work may wish to investigate different ways of identifying lags (and different time frames) and incorporating them into the model to best determine the optimal lag for the symptoms of interest. Moreover, we only investigated outdegree as a metric for dynamic changes across the 90 days. Recent work has raised concerns about the suitability of this metric to investigate heterogeneity across time (both within and between persons; Hoekstra et al., 2023). Thus, although our findings provide important insights into how individuals' symptom presentations change over time, it is important to use these types of metrics (e.g., outdegree) along with other data to investigate heterogeneity and clinically meaningful changes over time.

Finally, the MLife app utilized for the current study was developed for use on Android devices. Thus, participants were required to own and use an Android phone as their primary device, resulting in exclusion of participants who used smartphones other than an Android (e.g., iPhones). Given that Android devices constitute 44% of smartphone usage in the United States (statcounter, 2022), our sample does not accurately reflect the larger U.S. population with regard to smartphone usage.

Conclusions

In the current study, we conducted the first case series investigating the symptom dynamics of MDD using time-varying VAR models. Our findings support prior research that MDD is a dynamic, constantly evolving system and suggest that the dynamics of depressive symptoms are person-specific and can dramatically change over time in response to both internal and external factors. Moreover, our findings suggest that digital interventions may be promising toward providing personalized, in-the-moment treatment for depressed individuals. Thus, monitoring depressive symptoms with intensive, longitudinal data may allow for better detection of symptom changes and for implementation of time-sensitive interventions.

References

Beard, C., Millner, A. J., Forgeard, M. J. C., Fried, E. I., Hsu, K. J., Treadway, M. T., Leonard, C. V., Kertz, S. J., & Björgvinsson, T. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine*, *46*(16), 3359–3369. https://doi.org/10.1017/S0033291716002300

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Bringmann, L. F. (2021). Person-specific networks in psychopathology: Past, present, and future. *Current Opinion in Psychology*, *41*, 59–64. https://doi.org/10.1016/j.copsyc.2021.03.004

Bringmann, L. F., Ferrer, E., Hamaker, E. L., Borsboom, D., & Tuerlinckx, F. (2018). Modeling nonstationary emotion dynamics in dyads using a time-varying vector-autoregressive model. *Multivariate Behavioral Research*, *53*(3), 293–314. https://doi.org/10.1080/00273171.2018.1439722

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological Methods*, *22*(3), 409–425. https://doi.org/10.1037/met0000085

Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Medicine*, *45*(4), 747–757. https://doi.org/10.1017/S0033291714001809

Cramer, A. O. J., van Borkulo, C. D., Giltay, E. J., van der Maas, H. L. J., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PLoS ONE*, *11*(12), Article e0167490. https://doi.org/10.1371/journal.pone.0167490

de Vos, S., Wardenaar, K. J., Bos, E. H., Wit, E. C., Bouwmans, M. E. J., & de Jonge, P. (2017). An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. *PLoS ONE*, *12*(6), Article e0178586. https://doi.org/10.1371/journal.pone.0178586

Ebrahimi, O. V., Burger, J., Hoffart, A., & Johnson, S. U. (2021). Within- and across-day patterns of interplay between depressive symptoms and related psychopathological processes: A dynamic network approach during the COVID-19 pandemic. *BMC Medicine*, *19*(1), Article 317. https://doi.org/10.1186/s12916-021-02179-y

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, *53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of Consulting and Clinical Psychology*, *83*(4), 825–836. https://doi.org/10.1037/ccp0000026

Fried, E. I., Flake, J. K., & Robinaugh, D. J. (2022). Revisiting the theoretical and methodological foundations of depression measurement. *Nature Reviews Psychology*, *1*(6), 358–368. https://doi.org/10.1038/s44159-022-00050-2

Fried, E. I., & Nesse, R. M. (2015a). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders*, *172*, 96–102. https://doi.org/10.1016/j.jad.2014.10.010

Fried, E. I., & Nesse, R. M. (2015b). Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. *BMC Medicine*, *13*(1), Article 72. https://doi.org/10.1186/s12916-015-0325-4

Haslbeck, J. M. B., Bringmann, L. F., & Waldorp, L. J. (2021). A tutorial on estimating time-varying vector autoregressive models. *Multivariate Behavioral Research*, *56*(1), 120–149. https://doi.org/10.1080/00273171.2020.1743630

Hastie, T., & Tibshirani, R. (1995). Generalized additive models for medical research. *Statistical Methods in Medical Research*, *4*(3), 187–196. https://doi.org/10.1177/096228029500400302

Hoekstra, R. H. A., Epskamp, S., & Borsboom, D. (2023). Heterogeneity in individual network analysis: Reality or illusion? *Multivariate Behavioral Research*, *58*(4), 762–786. https://doi.org/10.1080/00273171.2022.2128020

Jacobson, N. C., & Nemesure, M. D. (2021). Using artificial intelligence to predict change in depression and anxiety symptoms in a digital intervention: Evidence from a transdiagnostic randomized controlled trial. *Psychiatry Research*, *295*, Article 113618. https://doi.org/10.1016/j.psychres.2020.113618

Jordan, D. G., Winer, E. S., & Salem, T. (2020). The current status of temporal network analysis for clinical science: Considerations as the paradigm shifts? *Journal of Clinical Psychology*, *76*(9), 1591–1612. https://doi.org/10.1002/jclp.22957

Kaiser, T., & Laireiter, A.-R. (2018). Process-symptom-bridges in psychotherapy: An idiographic network approach. *Journal for Person-Oriented Research*, *4*(2), 49–62. https://doi.org/10.17505/jpor.2018.06

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, *32*(9), 509–515. https://doi.org/10.3928/0048-5713-20020901-06

Lorenz, N., Sander, C., Ivanova, G., & Hegerl, U. (2020). Temporal associations of daily changes in sleep and depression core symptoms in patients suffering from major depressive disorder: Idiographic time-series analysis. *JMIR Mental Health*, *7*(4), Article e17071. https://doi.org/10.2196/17071

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

Moshe, I., Terhorst, Y., Philippi, P., Domhardt, M., Cuijpers, P., Cristea, I., Pulkki-Råback, L., Baumeister, H., & Sander, L. B. (2021). Digital interventions for the treatment of depression: A meta-analytic review. *Psychological Bulletin*, *147*(8), 749–786. https://doi.org/10.1037/bul0000334

Siepe, B. S., Sander, C., Schultze, M., Kliem, A., Ludwig, S., Hegerl, U., & Reich, H. (2022). Temporal dynamics of depressive symptomatology: An idiographic time series analysis applying network models to patients with depressive disorders. *PsyArXiv*. https://doi.org/10.31234/osf.io/hnw69

Smith, K. (2014). Mental health: A world of depression. *Nature*, *515*(7526), 180–181. https://doi.org/10.1038/515180a

statcounter. (2022). *Mobile operating system market share United States of America*. https://gs.statcounter.com/os-market-share/mobile/united-states-of-america

Teepe, G. W., Da Fonseca, A., Kleim, B., Jacobson, N. C., Salamanca Sanabria, A., Tudor Car, L., Fleisch, E., & Kowatsch, T. (2021). Just-in-time adaptive mechanisms of popular mobile apps for individuals with depression: Systematic app search and literature review. *Journal of Medical Internet Research*, *23*(9), Article e29412. https://doi.org/10.2196/29412

Torous, J., Staples, P., Shanahan, M., Lin, C., Peck, P., Keshavan, M., & Onnela, J.-P. (2015). Utilizing a personal smartphone custom app to assess the Patient Health Questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. *JMIR Mental Health*, *2*(1), Article e3889. https://doi.org/10.2196/mental.3889

Wang, L., & Miller, L. C. (2020). Just-in-the-moment adaptive interventions (JITAI): A meta-analytical review. *Health Communication*, *35*(12), 1531–1544. https://doi.org/10.1080/10410236.2019.1652388

Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., & Campbell, A. T. (2014). *StudentLife: Assessing mental health, academic performance and behavioral trends of college students using smartphones*. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 3–14). https://doi.org/10.1145/2632048.2632054

Wichers, M., Groot, P. C., & Psychosystems, ESM Group, EWS Group. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, *85*(2), 114–116. https://doi.org/10.1159/000441458

Wichers, M., Riese, H., Hodges, T. M., Snippe, E., & Bos, F. M. (2021). A narrative review of network studies in depression: What different methodological approaches tell us about depression. *Frontiers in Psychiatry*, *12*, Article 719490. https://doi.org/10.3389/fpsyt.2021.719490

Wichers, M., Smit, A. C., & Snippe, E. (2020). Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: A confirmatory single-subject time-series study. *Journal for Person-Oriented Research*, *6*(1), 1–15. https://doi.org/10.17505/jpor.2020.22042

Wilhelm, S., Weingarden, H., Ladis, I., Braddick, V., Shin, J., & Jacobson, N. C. (2020). Cognitive-behavioral therapy in the digital age: Presidential address. *Behavior Therapy*, *51*(1), 1–14. https://doi.org/10.1016/j.beth.2019.08.001

Wood, S. N. (2017). *Mixed GAM computation vehicle with automatic smoothness estimation*. R package version 1.8–12.

Yuan, M., & Lin, Y. (2007). Model selection and estimation in the Gaussian graphical model. *Biometrika*, *94*(1), 19–35. https://doi.org/10.1093/biomet/asm018

Received December 2, 2022
Revision received August 31, 2023
Accepted September 13, 2023 ■