MULTIVARIATE BEHAVIORAL RESEARCH
2024, VOL. 59, NO. 4, 663–676

Methodological and Statistical Practices of Using Symptom Networks to Evaluate Mental Health Interventions: A Review and Reflections

Lea Schumacher[a], Julian Burger[b,c,d], Jette Echterhoff[a], and Levente Kriston[a]

[a]Department of Medical Psychology, University Medical Center Hamburg-Eppendorf; [b]Department of Psychological Methods, University of Amsterdam; [c]University Medical Center Groningen, University of Groningen; [d]Centre for Urban Mental Health, University of Amsterdam

ABSTRACT
The network approach to psychopathology, which assesses associations between individual symptoms, has recently been applied to evaluate treatments for mental disorders. While various options for conducting network analyses in intervention research exist, an overview and an evaluation of the various approaches are currently missing. Therefore, we conducted a review on network analyses in intervention research. Studies were included if they constructed a symptom network, analyzed data that were collected before, during or after treatment of a mental disorder, and yielded information about the treatment effect. The 56 included studies were reviewed regarding their methodological and analytic strategies. About half of the studies based on data from randomized trials conducted a network intervention analysis, while the other half compared networks between treatment groups. The majority of studies estimated cross-sectional networks, even when repeated measures were available. All but five studies investigated networks on the group level. This review highlights that current methodological practices limit the information that can be gained through network analyses in intervention research. We discuss the strength and limitations of certain methodological and analytic strategies and propose that further work is needed to use the full potential of the network approach in intervention research.

KEYWORDS
Symptom networks; network analysis; mental health interventions; psychopathology; review

The network approach to psychopathology has been gaining considerable popularity in the past years (Borsboom & Cramer, 2013; Robinaugh et al., 2020). This approach suggests that mental health problems develop and are sustained by symptoms mutually causing each other and describes mental disorders as networks of interacting symptoms (Borsboom, 2017). In addition to describing the symptomatology of a specific patient group with symptom networks, it was suggested to apply the network approach to plan and evaluate treatments for mental disorders (Blanchard & Heeren, 2022; McNally, 2016). In this framework, treatment effects are discussed in regards to the treatment's impact on symptom networks. More specifically, interventions may change the severity of specific symptoms, the interactions between symptoms, or impact symptom-triggering variables in the external field (Borsboom, 2017).

So far, interventions for mental disorders have mostly been evaluated by analyzing the presence/absence of a diagnosis or a composite score indicating the aggregated severity of several symptoms. Additional information could be gained with symptom networks, since these allow the analysis of the treatment effect on specific symptoms and symptom associations. Studying treatment effects on the symptom level seems promising due to several reasons. First, the effects of treatments for mental disorders might be symptom-specific, i.e., some symptoms are influenced while others are not (Bekhuis et al., 2018), and solely focusing on composite-scores or the presence of a diagnosis cannot reveal such symptom-specific effects (Kaiser et al., 2021). Second, large variations in symptom expressions have been observed for individuals with the same diagnosis, therefore, a diagnosis might not be a good description of the experienced problems of the target population (Fried & Nesse, 2015). Similarly, it was shown that individuals with similar symptom severity showed markedly different symptom associations (Ebrahimi et al., 2023). Fourth, when investigating symptom networks longitudinally throughout treatment, changes in symptom associations might offer some insights into the working mechanisms of the treatment (Hofmann et al., 2020). Thus, using symptom networks to evaluate mental health interventions could potentially broaden the knowledge on treatment effects by focusing on individual symptoms and their relations.

Statistical methods were developed to estimate symptom networks from empirical data (Bringmann et al., 2013; Epskamp & Fried, 2018). Such networks consist of nodes indicating observed symptoms and edges which show statistical relationships between the symptoms (Epskamp, Borsboom, et al., 2018). In the last years, network analysis has been frequently applied to investigate the symptomatology of specific patient groups (Robinaugh et al., 2020). Here, network analysis often assesses (a) the strength of edges, i.e., how strongly a symptom relates to another symptom, (b) the centrality of nodes, i.e., how strongly a symptom is associated with all other symptoms, and (c) overall connectivity, i.e., how strongly all symptoms are, at average, associated with each other. Additionally, researchers have started to use network analysis to evaluate treatments for mental disorders.

There are various conceptual, methodological and statistical options available for applying the network approach in intervention research. Therefore, several decisions need to be taken by applied researchers when evaluating treatment with network analyses. First, the study design impacts the extent to which a causal treatment effect can be evaluated. With data from (randomized) controlled study designs, the causal effect of a specific treatment on (1) specific symptoms or on (2) symptom associations can be evaluated. To investigate the direct and indirect symptom-specific effects of a given treatment, it was suggested to add a treatment node in network models, which indicates the allocation to a treatment group, a control group, or an alternative treatment (Blanken et al., 2019). Alternatively, when the research question concerns the treatment effect on symptom associations, networks can be compared between treatment groups.

Second, the quality and quantity of information that can be gained highly depend on what (kind of) variables are included as nodes in the network (Bringmann et al., 2022). Researchers need to decide which symptoms to consider and whether variables that are not symptoms (such as potential treatment effect modifiers) should be included. Possible interpretations are also highly impacted by whether nodes constitute an absolute or a change score for a symptom and if the included variables are measured in a reliable and valid way. Using change scores, e.g. the change between treatment initiations and termination, allows to investigate if treatment impacts change in specific symptoms and/or if change in one symptom is related to change in another symptom. Third, depending on the data at hand, cross-sectional or longitudinal networks with different underlying statistical models can be constructed. When only one or a few repeated observations per person are available for a large number of persons, cross-sectional networks can be estimated. It is often discussed that cross-sectional networks mostly display between-person associations, i.e. differences between individuals, which do not directly relate to intra-individual processes (Borsboom et al., 2021). This is, symptoms can vary between persons (between-person associations) and within a person (within-person associations), and these are not necessarily related to each other (Schuurman, 2023). Cross-sectional analyses cannot distinguish within- and between-person associations which needs to be considered when analyzing cross-sectional data (Schuurman, 2023). Cross-sectional networks can be estimated using Graphical Gaussian Models (GGMs) for continuous variables, Ising models for categorical variables, or Mixed Graphical Models (MGMs) for mixed variables (Epskamp & Fried, 2018; Finnemann et al., 2021; Haslbeck & Waldorp, 2020). When several repeated measures are available for each person, longitudinal networks can be constructed. Longitudinal networks indicate temporal associations between the included variables, mostly if a symptom at one time point is associated with itself and/or another symptom at the next time point. When repeated measures for several persons are available between-person and within-person associations can be separated (Epskamp, Waldorp, et al., 2018; Schuurman, 2023). This means it can be shown how symptoms relate within a person and between persons. Longitudinal networks can be estimated with a graphical vector autoregressive (GVAR) model with repeated observations of one individual, with multilevel GVAR models (mlGVAR) with repeated observations from many persons (Bringmann et al., 2013; Epskamp, Waldorp, et al., 2018) and with panelGVAR models with panel data (Epskamp, 2020a). Adaptions of these models or completely different models are also possible. Finally, there are various ways to describe the estimated networks. Different network parameters can be calculated describing node centrality or the network topology, different comparisons can be made (e.g., between treatment groups, at different time points or between treatment responders and non-responders), and different statistical analyses can be conducted.

The potential of network analysis in intervention research is likely to depend strongly on such methodological and analytic choices. We realize that different analytic procedures are probably valuable for different research questions and contexts. Still, to our knowledge, there is no consensus on which methods are most suitable for the evaluation of treatment effects and there is very little guidance for applied researchers for choosing analytic strategies. An overview of which analytic choices have been previously made is also unavailable so far. Therefore, we systematically reviewed intervention studies that used network analysis to evaluate treatments for mental disorders. Through this review, we aimed to gain an overview of the methodological and analytic decisions that previous studies took and discuss the benefits and drawbacks of these. This can inform future studies using the network approach and, hopefully, increase the value of the network approach in intervention research.

Methods

Study search

We searched three bibliographic databases (PsycINFO, MEDLINE, and Web of Science) for intervention studies that utilized symptom network analyses. The title, abstract, keywords, and subject headings were searched by combining terms from three categories: (1) network analysis as the method of data analysis, (2) intervention study as the study design, and (3) individuals with mental health problems as the target population. The specific search terms can be found in the supplemental materials[1], Text S1. Additionally, we performed forward and backward reference search for the included studies and searched Google Scholar with the term "network intervention analysis". Finally, we checked the references of reviews on using the network approach in the field of mental health and psychopathology.

[1]All supplemental materials can be found here: https://osf.io/n4xp5

Inclusion criteria

The following inclusion criteria were applied: (1) the study conducted a network analysis which investigated the relation among psychological symptoms (and possibly other variables), (2) the study analyzed data which were collected before, after or during a treatment which was directed at psychological problems or mental disorders, (3) the analysis provides some information about the effect of the treatment, and (4) the study was published in a peer-reviewed journal. Network meta-analyses and network analyses in which nodes represented people or neural connections were excluded.

Study selection and data extraction

Titles and abstracts were screened with the above outlined inclusion criteria using the software Rayyan (Ouzzani et al., 2016); 20% of the titles and abstracts were double-screened by two independent raters. The full-texts of all studies that were found eligible in the first step were examined by two independent raters regarding the final decision to include the study in the review. Interrater agreement was quantified by calculating Cohen's κ. Disagreement between raters was resolved by discussion. Information on the sample characteristics, the intervention(s), the research design, the estimated networks, the statistical analysis, and the use of open science practices was extracted for all included studies. All variables are displayed in the supplemental materials, Table S1. Two independent raters extracted information on the main variables describing the network estimation and further statistical analyses. Disagreement was resolved by discussion.

This review was preregistered on the open science framework (https://osf.io/8txcy). The preregistered inclusion criteria were slightly adjusted, our rationale for this can be found in the supplemental materials, Table S2. Study materials and data (exact search terms, inclusion criteria, list of excluded studies, extracted data) are publicly available in the online supplemental material (https://osf.io/n4xp5/). This study is reported according to the extension of the Preferred Reporting Items for Systematic reviews and Meta-Analyses for scoping reviews (PRISMA-ScR, Tricco et al., 2018).

Results

Search results

The bibliographic database search in December 2021 yielded 4519 records, of which 4298 remained after deduplication. After title and abstract screening, the full-texts of 39 studies were screened, and 34 studies met all inclusion criteria. Interrater-reliability was $\kappa = 0.89$ and $\kappa = 0.86$ for the abstract screening and the full-text screening, respectively. The additional search (forward/backward reference search, Google scholar search, and search in network reviews) in April 2022 yielded the inclusion of additional 22 studies, leading to a total of 56 included studies. Of note, we considered such a high turn-out of the additional search due to the diverse network terminology and the substantial amount of ongoing studies. A detailed overview of the study selection procedure is displayed in Figure 1. A list of all included studies and an overview of all excluded studies can be found in the supplemental materials Text S2, and Table S3, respectively.

Study characteristics

Most studies were conducted in Europe ($n = 24$, 42.9%) or Northern America ($n = 16$, 28.6%) between 2015 and 2022, with the majority of studies being published between 2020 and 2022 ($n = 33$, 58.9%). Across all studies, the average age of the participants had a median of 41.4 years and the proportion of females had a median of 65.7%. The most often investigated patient group were persons with depressive symptoms or a diagnosis of a depressive disorder ($n = 29$, 51.8%). About a third of the included studies evaluated some form of cognitive behavioral therapy, and 21.4% evaluated antidepressants. The interventions had a mean length of 13.1 wk, ranging between 2 and 52 wk. About half of the studies classified as randomized controlled trials (RCTs, $n = 29$, 51.8%) and one study conducted an individual patient-data meta-analysis of RCTs. The other half of the studies was observational ($n = 26$, 46.4%)[2]. More than three-quarters of all included studies ($n = 44$) reported a secondary analysis of previously collected data. Study characteristics of all studies are displayed in Table S1 in the supplemental material.

Methodological and analytical decision points

Using network analysis to evaluate treatments for mental disorders entails various methodological and analytic decisions. In the following section, the identified studies are reviewed with regard to these decision points. The reviewed decision points roughly follow the general workflow used in network approaches to multivariate psychological data (Borsboom et al., 2021). As a vast majority of included studies reported a secondary data analysis of existing data, data collection is not separately reviewed. Instead the network structure estimation is reviewed for different types of underlying data. Of note, this overview provides an uncommented summary of the methodological and analytical decisions of previous studies. Recommendations regarding each of the decisions can be found in the discussion section.

Research question: the investigation of causal treatment effects

To investigate the causal effect of specific treatments, a (preferably randomized) controlled study design is necessary. In the current review, 30 studies (30/56) used data from a randomized controlled trial. Of these 30 studies, twelve included a treatment node in the network, often termed network intervention analysis (NIA) and allowing the investigation of symptom-specific treatment effects. Several studies estimated cross-sectional networks including a treatment node at different time points during the treatment ($n = 7/12$). Alternatively, NIA was conducted with changes scores that indicate change from the initiation to the termination of treatment ($n = 5/12$). Such networks can show if the treatment was associated with changes in the severity of specific symptoms. Most of the remaining analyses of RCTs that did not include a treatment node compared symptom networks between treatment groups ($n = 11/18$). The remaining seven of these 18 studies combined data from both treatment groups. More specifically, four studies investigated cross-sectional networks at different time points before, during, and after the treatment or cross-sectional networks of different response groups. One study estimated a longitudinal network to assess temporal symptom associations during treatment with combined data from both groups. Finally, two of the seven studies with combined data estimated symptom networks at baseline and evaluated if the network topology related to treatment outcomes.

Network structure estimation

Node selection and measurement. Across all studies, the median number of analyzed nodes was twelve, ranging between five and 47 nodes. While most of the studies ($n = 46/56$) included only symptoms in the networks, few studies also included additional variables such as schema beliefs, self-efficacy, or quality of life. More than two thirds of the studies ($n = 43/56$) did not report how they selected the nodes in the

[2]Data from five of these observational studies were originally from an RCT. However, data from just one treatment group was used in the reported analyses. Thus, we classified these studies as observational.

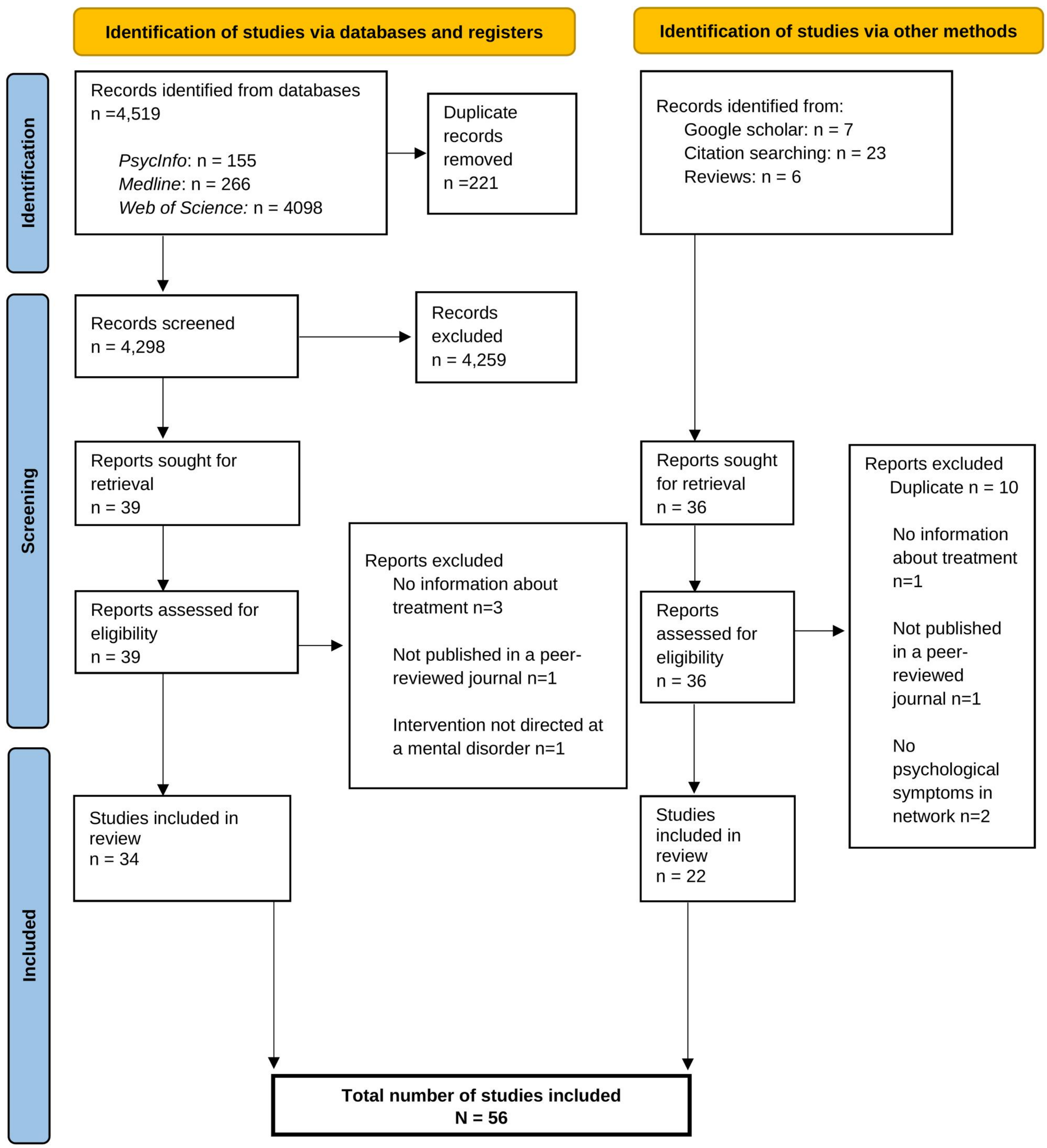

Figure 1. Flow chart for study selection.

networks. The remaining studies based their node selection on previous analyses ($n = 6/56$), on symptoms specified in diagnostic systems ($n = 3/56$), data availability ($n = 2/56$) and theoretical considerations ($n = 2/56$). The included variables were assessed by single items, mostly taken from symptom severity scales, in the majority of studies ($n = 40/56$), six studies (6/56) used composite scores, and ten studies (10/56) both. Nine out of the 56 studies calculated change scores, mostly between before and after treatment, and entered these as nodes in the network.

Data structure and the selection of statistical network models. The number of repeated measures per person that were used for the network analysis ranged from 1 to 120. Nearly half of the studies ($n = 25/56$) used one or two measurement points per person for their network analyses. Six studies estimated cross-sectional networks based on one data point before treatment and compared the networks of responders and non-responders or investigated if central symptoms at baseline could predict treatment outcome. All 19 studies that analyzed two measurement points estimated cross-sectional networks. Seven of these 19 studies calculated a change score between before and after treatment and entered these as nodes in the network analyses. Such networks allow investigating how changes in the severity of one symptom relate to changes in another symptom. The remaining twelve of the 19 studies compared networks before and after treatment and between responders and non-responders. The sample size for each of these cross-sectional network analyses ranged between 45 and 5614 with a median of 316.

Thirty studies (30/56) used data from more than two measurement points per person[3]. Half of these studies estimated cross-sectional networks ($n = 15/30$) at different time points and compared symptom networks across time before, during, and after treatment. The sample sizes for these cross-sectional networks ranged between 74 and 2862 with a median of 198 and they investigated between 3 and 14 different time points.

Longitudinal networks which indicate temporal associations between symptoms were modeled by 16 studies[4]. Different longitudinal network models were used when more than two measurement points per person were available. Most of the studies ($n = 12/16$) investigated longitudinal symptom networks with an mlGVAR model, which allows the estimation of time-lagged symptom associations. To handle a possible increase or decrease in the mean level of symptom severity in the investigated timeframe, i.e. a trend, five of these 12 studies detrended the data before analysis, four studies (4/12) included time as a predictor in the model, and one study (1/12) tested for a trend in the data. One study (1/12) added an interaction between time and the lagged symptoms to investigate if symptom associations change over time. Three quarters of the studies using mlGVAR models ($n = 9/12$), person-mean centered the data before estimating the models to disentangle within-person and between-person associations. A mlGVAR model allows the construction of three networks which show (1) temporal associations, (2) contemporaneous associations and (3) between-person associations (Borsboom et al., 2021; Epskamp, Waldorp, et al., 2018). Seven out of the twelve studies only reported the temporal network, four studies reported all three networks and one study reported the temporal and the contemporaneous network. While only half of these studies reported whether random intercepts and/or random slopes were included in the model ($n = 6/12$), all but one study investigated only the group-level fixed effects of these multilevel models. This means that average within-person symptom associations were evaluated. The studies using mlGVAR models analyzed data from a median of 20 repeated measures, ranging between four to 120 repeated observations. With sample sizes ranging between 5 and 1210 with a median of 44, the networks were based on a median number of data points of 438.5. One study used a panelGVAR model based on 3 repeated observations of 100 patients.

Five studies estimated person-specific longitudinal networks, i.e., modeled the network structure individually for each included patient. More specifically, two studies used dynamic time warp distance matrices to estimate person-specific longitudinal symptom networks based on 6 and 7 data points of 255 and 133 persons. One study estimated a VAR model for one patient based on 26 repeated observations. A unified structural equation model was estimated using 56 repeated observations for each participant in one study with a sample size of 1210 persons. Finally, one study estimated an mlGVAR model including only random effects to estimate patient level network parameters based on 13 repeated measures from 621 patients.

Network description

Forty-eight studies (48/56) calculated parameters to describe the role of each node, i.e. centrality indices, or the networks topology more globally. The three most common centrality indices were node strength/degree ($n = 34/48$), node closeness ($n = 21/48$), and node betweenness ($n = 19/48$). The overall connectedness was evaluated by 22 of the 48 studies and five studies investigated clusters in the network. A list of all network parameters can be found in the supplemental material, Table S4. Of all 36 studies comparing cross-sectional networks, 13 studies used the Network Comparison Test (van Borkulo et al., 2022). Three studies used permutation tests to statistically assess differences between longitudinal networks. Four studies used the Kolmogorow–Smirnov test, repeated measure ANOVA, *t*-test, or Wilcoson signed rank test to compare network parameters. Finally, five studies used the network parameters to predict some form of treatment outcome.

[3]One study did not report how many repeated measurement points were used.

[4]One study estimated a cross-sectional network and a longitudinal network.

Network stability analyses

Of all 40 cross-sectional network analyses, 23 reported conducting nonparametric bootstrap to estimate the stability of the edge weights. Of the 16 longitudinal network analyses, only one study reported that they compared their dataset to a previous simulation study to gauge the robustness of their analyses. Of the studies reporting network parameters, 17/48 performed a case-dropping bootstrap to evaluate the stability of these parameters. Twenty studies (20/56) indicated that there was no missing data or excluded participants with missing data. In four studies (4/56), imputation methods were used to handle missing data.

Software

Studies on cross-sectional networks reported to use the packages *qgraph* (Epskamp et al., 2012) 21 times, *mgm* (Haslbeck & Waldorp, 2020) eleven times, and *psychonetrics* (Epskamp, 2020b) twice in R (R Core Team, 2013) as well as the *NetworkX* package (Hagberg et al., 2008) in Python (Van Rossum & Drake, 2009) two times. For the analysis of longitudinal networks, the packages *mlVAR* (Epskamp et al., 2023) was reported five times, *qgraph* (Epskamp et al., 2012) seven times, *dtw* (Giorgino, 2009), *pheatmap* (Kolde, 2019), *parallelDist* (Eckert, 2022), *nmle* (Pinheiro et al., 2022), and *pompom* (Yang et al., 2021) in R each one time, as well as the command *mixed* in Stata (StataCorp., 2023) twice.

Reporting and open science practices

Information about several important analysis characteristics was not reported for a considerable number of studies (see supplemental material Table S5). 41 out of the 56 studies (73.2%) did not report how nodes were selected, 28 studies (50%) did not report handling of missing data, 10 studies (17.9%) did not specify which software package was used, and seven studies (12.5%) did not report the model that was estimated. Software code to reproduce the analysis was available for 11 studies (19.6%). Only one study shared their data, and three studies (5.4%) published correlation matrices to enable the reproduction of the networks. None of the network analyses were preregistered. The original trial that the data were taken from was registered in nine cases (16%).

Discussion

We conducted this systematic review to gain an overview on how network analysis has been applied to evaluate treatments for mental disorders and to assess which methodological and analytic decisions were taken in previous studies. The application of network analysis in intervention research has spurred in the last years, with the majority of studies being published since 2020. Most studies included patients with depressive symptoms and investigated the effect of CBT or antidepressants. In the following, we are reflecting on the previously reviewed methodological and analytical decision points, discuss important issues for each specific point, and give initial recommendations for these decisions.

Research question: the investigation of causal treatment effects

To investigate the causal effect of a specific treatment, a (randomized) controlled study design is needed. When this is given, researchers need to decide if treatment effects on specific symptoms, on symptom associations or on both should be evaluated. About half of all included studies used data from an RCT. Of these studies, 40% investigated symptom-specific treatment effects and 36.6% compared networks between treatment groups. To assess treatment effects on specific symptoms, NIA, in which a treatment node is added to the network (Blanken et al., 2019) seems to be a promising tool. NIA allows the investigation of direct and indirect symptom-specific effects of the treatment. Especially when only data at baseline and after treatment are available, additional information can be gained when NIA is conducted with change scores. These change score networks incorporate the information from two time points and indicate if the treatment is related to change in specific symptoms. Still, NIA for specific time points or with change scores does not disentangle within- from between-person associations. While this separation is possible with longitudinal analyses, NIA can often not be conducted using longitudinal models as the treatment does not vary over time for each person in most investigated RCTs. In the future, time-series designs with a large number of repeated measurements of multiple persons and within-person randomization to different treatments could allow NIA with longitudinal models and shed some light on within-person treatment effects on specific symptoms.

Second, if researchers are interested in the treatment effect on symptom interactions, symptom networks should be compared between treatment groups. In the set of studies included in this review, this was mostly done by estimating symptom networks separately for each treatment group. Here, networks should be compared statistically and not only visually, for example with the network comparison test (van Borkulo et al., 2022). A few reviewed studies combined data from both treatment groups for the network analyses. When it is assumed that different treatments change symptoms and their associations differently, network estimation with data combined from both treatment groups seems questionable. The treatment effect on symptom associations could also be modeled directly as done by Schumacher et al. (2023). With adequate parametrization, this allows a direct estimation of the size of the treatment effect.

Network structure estimation

Node selection and measurement. Most reviewed networks included only symptoms as nodes, which were often assessed by single items. Further, we found that just a small number of studies described how the variables that were included in the networks were selected. However, as most networks display conditional dependencies and the resulting network structure depends on which variables are included, node selection should be carefully considered (Bringmann et al., 2022). Not including influential (possibly confounding) variables could lead to spurious symptom associations and misleading interpretation. This emphasizes the need to develop general standards for symptoms inclusion in the field. Theoretical specifications need to be made regarding treatment effects on symptom and their associations, as these could also guide node selection. Until then, individual researchers should thoroughly consider which variables to include in the network and try to find clear criteria for the inclusion of variables for their specific study. This could, for example, be guided by symptoms included in the current classification systems like DSM 5 and ICD 11. At the same time, it should be noted that DSM criterion symptoms do not seem to be more central than other symptoms in depression (Fried et al., 2016).

In addition to symptoms, it might be beneficial to also include hypothesized treatment processes as variables in the network in order to improve understanding of the treatments' working mechanisms (Hofmann et al., 2020). Johnson and Hoffart (2018) investigated, for example, the associations between symptoms and (meta-)cognitions for patients receiving metacognitive therapy or cognitive behavioral therapy. Including potentially relevant variables in the network next to symptoms takes advantage of the fact that networks can display associations between multiple potentially influential variables and therefore, might enrich the knowledge that can be gained about treatment effects from network analysis. However, broadening the scope of the modeled phenomenon beyond symptoms will also inevitably put more stress on the need for theoretical justification, highlighting an urgent need for the further development of the theory behind the network approach to psychopathology. Further, the integration of latent variables in network models (Epskamp, 2020a) might help to also consider broader constructs, which are difficult to measure with individual items.

This review also showed that symptoms were most commonly measured with a single item, which contrasts to established measurement practices (Allen et al., 2022). A reliable measurement of included variables is a pre-requisite to a reliable network analysis (Bringmann et al., 2022), therefore, future studies need to assess the validity of this approach (Allen et al., 2022). Fortunately, this topic seems to gain continuously increasing attention (Dejonckheere et al., 2022).

Data structure and the selection of statistical network models. About half of all reviewed studies analyzed one or two measurement points per person and estimated cross-sectional networks, often before, and after treatment or in responder and non-responder groups. Cross-sectional networks at different time points during treatment were also estimated in 15 studies with more than two observation points per persons. Such networks can show how symptom associations change throughout the course of treatment or may vary after two different kinds of treatments. However, as cross-sectional analyses do not disentangle within-person from between-person effects (Schuurman, 2023), it seems questionable how cross-sectional network analyses relate to within-person treatment effects (Bos et al., 2017). When only a limited number of data points are available for each person, it seems beneficial to estimate networks with change scores, as these incorporate within-person change at least to some extent. It has been argued that for relating treatment effects to the individual person, within-person instead of between-person treatment effects need to be assessed (Epskamp et al., 2018; Molenaar, 2004; Schuurman, 2023). Therefore, it needs to be considered to what extend results from between-person networks can generalize to treatment effects for the individual patient. This should be carefully reflected when interpreting cross-sectional symptom networks.

Nearly a third of all reviewed studies investigated longitudinal networks, which mostly show temporal associations between symptoms. Here, an mlGVAR model was most often used with a median of 20 repeated measurements. Using longitudinal networks to evaluate treatment effects has the benefit that between- and within-person processes can be investigated (Epskamp et al., 2018). While several studies only reported the temporal networks showing within-person associations, more information could be gained if between-person, contemporaneous and temporal networks would be reported. One major obstacle for longitudinal network models is that the popular VAR model, including panelGVAR and mlGVAR, assumes that symptom relations stay the same over time, i.e., are stationary (Bringmann et al., 2013; Epskamp, 2020a). When analyzing data which were collected during treatment, this assumption is likely to be violated. As seen in this review and a previous review of longitudinal networks (Blanchard et al., 2023), most studies handle the violation of stationarity by detrending the data or by including a trend in the model. However, as symptoms are hypothesized to change through treatments, their relations are also likely to change. Similarly, network theory proposes that treatment changes the symptom associations (Borsboom, 2017). Thus, by ignoring the change in symptom associations over time, we run the risk of missing a substantial part of the treatment effect. New methods to estimate time-varying VAR-models (Haslbeck et al., 2021; Schumacher et al., 2023) or dynamic time warp analysis (Booij et al., 2021; Hebbrecht et al., 2020) do not assume stationarity in symptom relations and, thus, have a large potential for using longitudinal networks in intervention research. Similarly, estimating longitudinal networks for a time period before and a time period after treatment as done by Kreiter et al. (2021) and Snippe et al. (2017) seems promising, as here the assumption of stationary symptom associations is more likely to be true. When using multilevel data, i.e., repeated measures from more than one person, person-mean centering the data before the analysis can be used to disentangle within-person associations from between-person associations. This is, for example, automatically done when using the *mlVAR* software (Epskamp et al., 2023). All but five studies assessed symptom associations on the group level, i.e., showed either between-person or averaged within-person symptom associations or a mix of both. This means that current network analyses in intervention research provided very little information on individual differences.

To assess changes in symptom associations, a time-variant longitudinal network model taking individual differences into account could be estimated for each treatment group from data provided by frequently repeated measurements in a randomized controlled trial (Schumacher et al., 2023). Then, within-person symptom associations and their change in response to treatment could be investigated and compared between treatment groups. Similarly, longitudinal symptom networks including a treatment node based on data from individuals receiving different treatment sequentially, e.g. in a cross-over trial, could show within-person symptom-specific treatment effects.

One frequently asked question relates to sample sizes needed to estimate longitudinal symptom networks. For a panelGVAR model that allows the investigation of averaged but not individual within-person symptom associations, estimations can be conducted with at least three repeated measures (Epskamp, 2020a). This model already allows separating within- from between-person associations. Using mlGVAR models, variations of each individual from the group-averaged within-person effects can also be estimated. However, here, more repeated measurements are needed. Our review showed that mlGVAR models were estimated by a median of 20 repeated observations per person in previous studies. However, the reviewed studies mostly focussed on the fixed effect, i.e., the group-averaged symptom associations. Therefore, it still seems unclear how many repeated measures need to be collected to reliably estimate person-specific variation of the group-level symptom associations. Here, recommendations from for general multilevel models (Bolin et al., 2019; Maas & Hox, 2005) might be applicable. To reliably estimate person-specific symptom networks based on data from one person, several hundred repeated measures are needed (Mansueto et al., 2022). Here, using prior knowledge and Bayesian estimation could be a promising path forward (Burger, Ralph-Nearman, et al., 2022). In general, Bayesian estimation might provide a more robust way to estimate networks with small sample sizes often encountered in intervention research, as uncertainty can be quantified (Huth et al., 2022; Schumacher et al., 2023). Additionally, advancements in the theoretical foundations of network analyses in intervention research might allow a focus on the most important variables. This would limit the number of analyzed nodes and therefore, enable estimations also with smaller sample sizes.

Network description

There are many possible parameters to describe the networks and the reviewed studies varied widely in how they described the network structure. When applying network parameters like centrality or density, researchers should be aware that the validity and clinical meaning of these parameters are still unclear and, therefore, need to be used with caution. Further, the evaluation of treatment effects often included comparisons between networks. Here, it is important that networks are compared not only visually but also statistically, for example using the network comparison test or permutation analyses (van Borkulo et al., 2022). This is also important for comparing symptom networks between individuals, as differences can be easily over interpreted (Hoekstra et al., 2023).

Network stability analysis

Network stability was only assessed in half of the reviewed studies and power was mostly discussed post-hoc as a possible limitation. The power and robustness of psychopathology networks is a debated topic and (initial) recommendations depend on the model type, number of nodes, and expected effect sizes (Bringmann et al., 2022; Lafit et al., 2021). For cross-sectional networks, the stability of edges and network parameter should be assessed with bootstrapping methods to gauge the robustness of the network (Epskamp, Borsboom, et al., 2018). For longitudinal networks, robustness has been only evaluated with cross-validation (Bringmann et al., 2022). It is highly important that individual studies conduct robustness analyses of the estimated networks. Here, Bayesian approaches to network analysis could be promising as these can quantify the certainty of specific network structures and provide credibility intervals for all network parameters (Huth et al., 2022; Schumacher et al., 2023).

Reporting and open science practices

As network analysis (in intervention research) is rather new, no reporting standards have been established yet. When it is unclear what to report, the use of open science practices becomes also more difficult. In this review, information on the investigated interventions and the analysis was missing for a considerable number of studies. Burger, Isvoranu, et al. (2022) recently proposed reporting standards, which hopefully will lead to more consistent reporting of network analyses. Further, no study preregistered their analysis and only a few shared their code and data, making it more difficult to reproduce the analyses. While many studies were exploratory, it still seems recommendable to preregister a priori analytic decisions to limit the impact of (post-hoc) analytic decisions. We realize that within clinical contexts, data is often more sensitive and difficult to share while protecting patients' anonymity. Still, especially when the code is openly available, the analyses are more easily reproducible and researchers can learn from each other. Precise reporting and good reproducibility are especially important under light of the very large researcher degrees of freedom of network analysis in intervention research.

Limitations of the current review

Not all studies that were included in the review aimed to evaluate a treatment with the network approach, possibly inflating the variability of the methods found. These studies were still included because they analyzed data that were collected within the realm of treatment and conducted some analyses that allowed some inference in regards to the treatment effect. Additionally, this review included all kinds of interventions that were directed at mental health problems. Different network methods might be more suitable for different kind of interventions, and this could not be assessed in this study. We took this broad approach because we wanted to provide a comprehensive overview of network analyses that were used in mental health intervention research.

Overall evaluation of the current methodological state and outlook

This review showed that the previously used methodological and statistical practices have several limitations. In the identified studies, the importance of variable selection, variable measurement, and the assessment of the robustness of the analyses have not been comprehensively addressed so far. Given that the median sample size across all cross-sectional analyses was rather low, the number of repeated measures for longitudinal analyses also largely varied and the stability of the estimations were not regularly investigated, the degree of replicability of current symptom networks in intervention research is largely unknown. In general, the reliability and replicability of symptom networks are strongly debated topics (Neal et al., 2022). Thus, it is not yet clear if the findings of the individual studies can be expected to replicate. Similarly, the clinical validity of the network theory has hardly been explicitly tested. So far, there is limited evidence for the clinical importance of central symptoms (Rodebaugh et al., 2018; Spiller et al., 2020) and evidence for the claims of the network theory in regards to treatment effects is only emerging (Schumacher et al., 2023). Since these aspects are indispensable to gain valuable knowledge from network analysis, the field needs to move forward addressing these issues.

Even more importantly, it became apparent that currently available and used statistical network models can only indirectly and to a limited extend provide the information that was hoped to be gained through the network approach. This is, it was suggested that network analyses can inform on symptom-specific treatment effects (Blanken et al., 2019), provide information on how (causal) symptom associations change through treatment and, thereby offer insights into the working mechanisms of investigated treatments (Hofmann et al., 2020). As cross-sectional networks, which were most often used, display (partial) correlations, no causal interactions between symptoms can be inferred. Therefore, inference about possible working mechanisms seems difficult from these analyses. Longitudinal network models, in contrast to cross-sectional network models, can separate within from between-person associations and can indicate temporal associations. Therefore, inference about within-person effects and possible causal mechanisms might be more appropriate here but is still far from being unproblematic (Epskamp, van Borkulo, et al., 2018). Furthermore, as most longitudinal models assume stationarity, i.e., no change in symptom associations, no change due to treatment could be directly investigated. It should be noted that network theory suggests that individuals differ in their symptom networks (Borsboom and Cramer, 2013; Borsboom, 2017). As the majority of studies investigated group-level symptom networks, individual differences for treatment effects on symptom networks are also largely unknown. In sum, current methodological and statistical practices provide limited information on the causal, dynamic (i.e., changing) and possibly person-specific interactions among symptoms and how these are influenced by treatment.

Additionally, as a reviewer of the present manuscript pointed out, a considerable discrepancy or even a paradox exists between the current state of theory and the existing methods of network science for mental health. On the one hand, effects and working mechanisms of existing psychological (and to some extent also pharmacological and somatic) treatments are difficult to differentiate and threaten to merge into rather unspecific conceptual conjectures, a phenomenon frequently termed as “Dodo effect” or “Dodo Bird Verdict” (Cuijpers et al., 2019; Wampold et al., 1997). On the other hand, the network science methods model detailed relationships of a large number of psychological variables, being very specific with regard to their content. As a result of this conflict, some call for more strongly formalized (i.e., specific) theories (Haslbeck et al., 2019) while empirical researchers try to adjust their findings on the rather vague theory by focusing on global rather than specific network parameters (e.g., node-wise centrality or network connectivity instead of individual symptom associations). Although some research on the effects of mental health treatments on symptom networks exists that brings theory and observations explicitly together (Schumacher et al., 2023), considerable theoretical, methodological, and empirical efforts will be needed to close this gap.

We think that in such a new, emerging field, it can be expected that the originally used methodological practices need to be continuously developed. There are ongoing methodological developments, e.g., Bayesian approaches (Burger, Ralph-Nearman, et al., 2022; Huth et al., 2022), time-varying approaches (Haslbeck et al., 2021; Schumacher et al., 2023), and new estimation approaches like Group Iterative Multiple Model Estimation (Sanford et al., 2022), which are likely to advance the field further. Similarly, further development of the network theory of psychopathology could improve this line of research, as more clearly specified network-theoretical claims can be more easily tested. In our opinion, further methodological and statistical developments would be highly valuable, as the network approach to psychopathology offers a new perspective on the treatment of mental disorders and has the potential provide new insights about treatments and their effects. Network analyses allow the exploration of associations among various specific symptoms, the display of a complex picture of a multitude of variables, and an assessment of symptom-specific treatment effects. These kind of analyses are likely to provide more detailed information than analyses focussing on the composite score of various different symptoms or the presence of a diagnosis. With methodological advances, network analysis can be a tool for the investigation of treatment processes and personalization of treatments (Burger, Ralph-Nearman, et al., 2022), both increasingly important topics in intervention research.

We are aware that much progress is being made currently, and several newer approaches already address some of the mentioned weaknesses. We hope that the current review can offer initial guidance for applied researchers with regard to available methods and issues that deserve particular attention. Future studies need to address the scope of applicability of different methodological and analytic options. With a better understanding of which options are suitable for which questions, network analyses in intervention research can hopefully help us learn more about treatment effects on symptom networks and with this improve our understanding of treatments for mental health problems.

Article information

Conflict of interest disclosures: We have no conflict of interest to disclose.

Ethical Principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Role of the Funders/Sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

References

Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science. European Journal of Psychological Assessment, *38*(1), 1–5. https://doi.org/10.1027/1015-5759/a000699

Bekhuis, E., Schoevers, R., de Boer, M., Peen, J., Dekker, J., Van, H., & Boschloo, L. (2018). Symptom-specific effects of psychotherapy versus combined therapy in the treatment of mild to moderate depression: A network approach. *Psychotherapy and Psychosomatics*, *87*(2), 121–123. https://doi.org/10.1159/000486793

Blanchard, M. A., & Heeren, A. (2022). Ongoing and future challenges of the network approach to psychopathology: From theoretical conjectures to clinical translations. In *Comprehensive clinical psychology*. (pp. 32–46) Elsevier. https://doi.org/10.1016/B978-0-12-818697-8.00044-3

Blanchard, M. A., Contreras, A., Kalkan, R. B., & Heeren, A. (2023). Auditing the research practices and statistical analyses of the group-level temporal network approach to psychological constructs: A systematic scoping review. *Behavior Research Methods*, *55*(2), 767–787. https://doi.org/10.3758/s13428-022-01839-y

Blanken, T. F., Van Der Zweerde, T., Van Straten, A., Van Someren, E. J. W., Borsboom, D., & Lancee, J. (2019). Introducing Network Intervention Analysis to Investigate Sequential, Symptom-Specific Treatment Effects: A Demonstration in Co-Occurring Insomnia and Depression. *Psychotherapy and Psychosomatics*, *88*(1), 52–54. https://doi.org/10.1159/000495045

Bolin, J. H., Finch, W. H., & Stenger, R. (2019). Estimation of random coefficient multilevel models in the context of small numbers of level 2 clusters. *Educational and Psychological Measurement*, *79*(2), 217–248. https://doi.org/10.1177/0013164418773494

Booij, M. M., van Noorden, M. S., van Vliet, I. M., Ottenheim, N. R., van der Wee, N. J. A., Van Hemert, A. M., & Giltay, E. J. (2021). Dynamic time warp analysis of individual symptom trajectories in depressed patients treated with electroconvulsive therapy. *Journal of Affective Disorders*, *293*, 435–443. https://doi.org/10.1016/j.jad.2021.06.068

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

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

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 58. https://doi.org/10.1038/s43586-021-00055-w

Bos, F. M., Snippe, E., de Vos, S., Hartmann, J. A., Simons, C. J. P., van der Krieke, L., de Jonge, P., & Wichers, M. (2017). Can we jump from cross-sectional to dynamic interpretations of networks implications for the network perspective in psychiatry. *Psychotherapy and Psychosomatics*, *86*(3), 175–177. https://doi.org/10.1159/000453583

Bringmann, L. F., Albers, C., Bockting, C., Borsboom, D., Ceulemans, E., Cramer, A., Epskamp, S., Eronen, M. I., Hamaker, E., Kuppens, P., Lutz, W., McNally, R. J., Molenaar, P., Tio, P., Voelkle, M. C., & Wichers, M. (2022). Psychopathological networks: Theory, methods and practice. *Behaviour Research and Therapy*, *149*, 104011. https://doi.org/10.1016/j.brat.2021.104011

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLOS One*, *8*(4), e60188. https://doi.org/10.1371/journal.pone.0060188

Burger, J., Isvoranu, A.-M., Lunansky, G., Haslbeck, J. M. B., Epskamp, S., Hoekstra, R. H. A., Fried, E. I., Borsboom, D., & Blanken, T. F. (2022). Reporting standards for psychological network analyses in cross-sectional data. *Psychological Methods*, *28*(4), 806–824. https://doi.org/10.1037/met0000471

Burger, J., Ralph-Nearman, C., & Levinson, C. A. (2022). Integrating clinician and patient case conceptualization with momentary assessment data to construct idiographic networks: Moving toward personalized treatment for eating disorders. *Behaviour Research and Therapy*, *159*, 104221. https://doi.org/10.1016/j.brat.2022.104221

Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The role of common factors in psychotherapy outcomes. *Annual Review of Clinical Psychology*, *15*(1), 207–231. https://doi.org/10.1146/annurev-clinpsy-050718-095424

Dejonckheere, E., Demeyer, F., Geusens, B., Piot, M., Tuerlinckx, F., Verdonck, S., & Mestdagh, M. (2022). Assessing the reliability of single-item momentary affective measurements in experience sampling. *Psychological Assessment*, *34*(12), 1138–1154. https://doi.org/10.1037/pas0001178

Ebrahimi, O. V., Borsboom, D., Hoekstra, R. H. A., Epskamp, S., Ostinelli, E. G., Bastiaansen, J. A., & Cipriani, A. (2023). Towards precision in the diagnostic profiling of patients: Leveraging symptom dynamics in the assessment of major depressive disorder. *PsyArXiv*. https://doi.org/10.31234/osf.io/wh6cf

Eckert, A. (2022). *parallelDist*: Parallel Distance Matrix Computation using Multiple Threads. R package version 0.2.6, https://CRAN.R-project.org/package=parallelDist.

Epskamp, S. (2020a). Psychometric network models from time-series and panel data. *Psychometrika*, *85*(1), 206–231. https://doi.org/10.1007/s11336-020-09697-3

Epskamp, S. (2020b). *psychonetrics: Structural equation modeling and confirmatory network analysis* (0.8) [Computer software]. https://CRAN.R-project.org/package=psychonetrics

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4), 1–18. https://doi.org/10.18637/jss.v048.i04

Epskamp, S., Deserno, M. K., Bringmann, L. F., Veenman, M. (2023). *mlVAR: Multi-level vector autoregression* (0.5.1) [Computer software]. https://cran.r-project.org/web/packages/mlVAR/index.html

Epskamp, S., van Borkulo, C. D., van der Veen, D. C., Servaas, M. N., Isvoranu, A.-M., Riese, H., & Cramer, A. O. J. (2018). Personalized network modeling in psychopathology: The importance of contemporaneous and temporal connections. *Clinical Psychological Science*, *6*(3), 416–427. https://doi.org/10.1177/2167702617744325

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

Finnemann, A., Borsboom, D., Epskamp, S., & van der Maas, H. L. J. (2021). The theoretical and statistical ising model: A practical guide in R. *Psych*, *3*(4), 594–618. https://doi.org/10.3390/psych3040039

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

Fried, E. I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016). What are 'good' depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders*, *189*, 314–320. https://doi.org/10.1016/j.jad.2015.09.005

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, *31*(7), 1–24. https://doi.org/10.18637/jss.v031.i07

Hagberg, A., Swart, P., & S Chult, D. (2008). Exploring network structure, dynamics, and function using NetworkX. Los Alamos National Lab (LANL).

Haslbeck, J. M. B., & Waldorp, L. J. (2020). mgm: Estimating time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software*, *93*(8), 1–48. https://doi.org/10.18637/jss.v093.i08

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

Haslbeck, J. M. B., Ryan, O., Robinaugh, D., Waldorp, L., & Borsboom, D. (2019). *Modeling psychopathology: From data models to formal theories* [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/jgm7f

Hebbrecht, K., Stuivenga, M., Birkenhäger, T., Morrens, M., Fried, E. I., Sabbe, B., & Giltay, E. J. (2020). Understanding personalized dynamics to inform precision medicine: A dynamic time warp analysis of 255 depressed inpatients. *BMC Medicine*, *18*(1), 400. https://doi.org/10.1186/s12916-020-01867-5

Hoekstra, R. H. A., Epskamp, S., Nierenberg, A. A., Borsboom, D., & McNally, R. J. (2023). Testing similarity in longitudinal networks: The Individual Network Invariance Test (INIT). *PsyArXiv*. https://doi.org/10.31234/osf.io/ugs2r

Hofmann, S. G., Curtiss, J. E., & Hayes, S. C. (2020). Beyond linear mediation: Toward a dynamic network approach to study treatment processes. *Clinical Psychology Review*, *76*, 101824. https://doi.org/10.1016/j.cpr.2020.101824

Huth, K. B. S., de Ron, J., Mohammadi, R., Luigjes, J., Goudriaan, A. E., van Holst, R. J., Marsman, M. (2022). *Bayesian analysis of cross-sectional network psychometrics: A conceptual introduction and tutorial in R and JASP*. Open Science Framework. https://osf.io/https://osf.io/pxk6q?view_only=b3079e44c4a04f9ca4c4bc40dce87703

Johnson, S. U., & Hoffart, A. (2018). Metacognitive therapy versus cognitive behavioral therapy: A network approach.

Frontiers in Psychology, *9*, 2382. https://doi.org/10.3389/fpsyg.2018.02382

Kaiser, T., Boschloo, L., Berger, T., Meyer, B., Späth-Nellissen, C., Schröder, J., Hohagen, F., Moritz, S., & Klein, J. P. (2021). Maintaining outcomes of internet-delivered cognitive-behavioral therapy for depression: A network analysis of follow-up effects. *Frontiers in Psychiatry*, *12*, 598317. https://doi.org/10.3389/fpsyt.2021.598317

Kolde, R. (2019). *Package 'pheatmap'*. https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf

Kreiter, D., Drukker, M., Mujagic, Z., Vork, L., Rutten, B. P. F., van Os, J., Masclee, A. A. M., Kruimel, J. W., & Leue, C. (2021). Symptom-network dynamics in irritable bowel syndrome with comorbid panic disorder using electronic momentary assessment: A randomized controlled trial of escitalopram vs. placebo. *Journal of Psychosomatic Research*, *141*, 110351. https://doi.org/10.1016/j.jpsychores.2020.110351

Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., & Ceulemans, E. (2021). Selection of the number of participants in intensive longitudinal studies: A user-friendly shiny app and tutorial for performing power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science*, *4*(1), 251524592097873. https://doi.org/10.1177/2515245920978738

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86

Mansueto, A. C., Wiers, R. W., van Weert, J. C. M., Schouten, B. C., & Epskamp, S. (2022). Investigating the feasibility of idiographic network models. *Psychological Methods*, *28* (5), 1052–1068. https://doi.org/10.1037/met0000466

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy*, *86*, 95–104. https://doi.org/10.1016/j.brat.2016.06.006

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Neal, Z. P., Forbes, M. K., Neal, J. W., Brusco, M. J., Krueger, R., Markon, K., Steinley, D., Wasserman, S., & Wright, A. G. (2022). Critiques of network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *2*(1), 90. https://doi.org/10.1038/s43586-022-00177-9

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Pinheiro, J., Douglas, B., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., Ranke, J. (2022). *Package 'nlme'*. https://cran.r-project.org/web/packages/nlme/nlme.pdf

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Robinaugh, D. J., Hoekstra, R. H. A., Toner, E. R., & Borsboom, D. (2020). The network approach to psychopathology: A review of the literature 2008–2018 and an agenda for future research. *Psychological Medicine*, *50*(3), 353–366. https://doi.org/10.1017/S0033291719003404

Rodebaugh, T. L., Tonge, N. A., Piccirillo, M. L., Fried, E., Horenstein, A., Morrison, A. S., Goldin, P., Gross, J. J., Lim, M. H., Fernandez, K. C., Blanco, C., Schneier, F. R., Bogdan, R., Thompson, R. J., & Heimberg, R. G. (2018). Does centrality in a cross-sectional network suggest intervention targets for social anxiety disorder? *Journal of Consulting and Clinical Psychology*, *86*(10), 831–844. https://doi.org/10.1037/ccp0000336

Sanford, B. T., Ciarrochi, J., Hofmann, S. G., Chin, F., Gates, K. M., & Hayes, S. C. (2022). Toward empirical process-based case conceptualization: An idionomic network examination of the process-based assessment tool. *Journal of Contextual Behavioral Science*, *25*, 10–25. https://doi.org/10.1016/j.jcbs.2022.05.006

Schumacher, L., Klein, J. P., Elsaesser, M., Härter, M., Hautzinger, M., Schramm, E., & Kriston, L. (2023). Implications of the network theory for the treatment of mental disorders: A secondary analysis of a randomized clinical trial. *JAMA Psychiatry*, *80*(11), 1160–1168. https://doi.org/10.1001/jamapsychiatry.2023.2823

Schuurman, N. K. (2023). A 'within/between problem' primer: About (Not) separating within-person variance and between-person variance in psychology. [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/7zgkx

Snippe, E., Viechtbauer, W., Geschwind, N., Klippel, A., de Jonge, P., & Wichers, M. (2017). The impact of treatments for depression on the dynamic network structure of mental states: Two randomized controlled trials. *Scientific Reports*, *7*(1), 46523. https://doi.org/10.1038/srep46523

Spiller, T. R., Levi, O., Neria, Y., Suarez-Jimenez, B., Bar-Haim, Y., & Lazarov, A. (2020). On the validity of the centrality hypothesis in cross-sectional between-subject networks of psychopathology. *BMC Medicine*, *18*(1), 297. https://doi.org/10.1186/s12916-020-01740-5

StataCorp. (2023). Stata 18 Multilevel Mixed-Effects Reference Manual. College Station, TX: Stata Press.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., & Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

van Borkulo, C. D., van Bork, R., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2022). Comparing network structures on three aspects: A permutation test. *Psychological Methods*, *28*(6), 1273–1285. https://doi.org/10.1037/met0000476

Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*. CreateSpace.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empiricially, 'all must have prizes. *Psychological Bulletin*, *122*(3), 203–215. https://doi.org/10.1037/0033-2909.122.3.203

Yang, X., Ram, N., & Molenaar, P. (2021). *Person-oriented method and perturbation on the model* [Computer software].