Behaviour Research and Therapy 122 (2019) 103440

Contents lists available at ScienceDirect

Behaviour Research and Therapy

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

Symptom-specific effectiveness of an internet-based intervention in the treatment of mild to moderate depressive symptomatology: The potential of network estimation techniques

Lynn Boschloo[a,*], Pim Cuijpers[a], Eirini Karyotaki[a,b], Thomas Berger[c], Steffen Moritz[d], Björn Meyer[e], Jan Philipp Klein[f]

[a] Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, VU University, Amsterdam, the Netherlands
[b] Department of Global Health and Social Medicine, Harvard Medical School, Harvard, USA
[c] Department of Clinical Psychology and Psychotherapy, Bern University, Bern, Switzerland
[d] Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[e] GAIA AG, Hamburg, Germany
[f] Department of Psychiatry and Psychotherapy, Lübeck University, Lübeck, Germany

ARTICLE INFO

Keywords:
Internet-based intervention
CBT
Depressive symptoms
Network estimation techniques
Precision psychiatry

ABSTRACT

The internet-based intervention Deprexis® has proven to be effective in improving overall depression severity. The current pragmatic randomized controlled trial included 1013 participants with mild to moderate symptomatology and aimed to identify the symptom-specific effects of the internet-based intervention Deprexis (intervention group) in comparison to care as usual (control group). All participants -in both conditions- were permitted to use any type of treatment. Of the nine considered symptoms (assessed with the Patient Health Questionnaire), seven showed larger improvements in the intervention condition relative to care as usual (effect sizes ranging from 0.15 to 0.31). No significant differences were found for the two other symptoms. In a next step, a network was estimated including treatment condition as well as changes in all nine symptoms. The resulting network suggests that four of the seven identified symptom-specific effects were direct, whereas the three other symptom-specific effects were indirect and could be explained by effects on other symptoms. Lastly, exploratory analyses showed that the intervention was more effective in improving overall depression severity for participants with higher scores on those four symptoms that were directly affected by the intervention; consequently, the network estimation techniques showed potential in precision psychiatry.

Funding

This work was supported by the German Federal Ministry of Health (grant number II A 5-2512 FSB 052) and the Netherlands Organization for Health Research and Development (Zon-MW; grant number 016-186-139).

1. Introduction

1.1. The importance of considering individual symptoms

Randomized controlled trials (RCTs) provide the best evidence about the general effectiveness of treatments. They have demonstrated, for example, that pharmacological (e.g., Mulrow et al., 2000), psychotherapeutic (e.g., Churchill et al., 2001) as well as internet-based (e.g., Karyotaki et al., 2017) interventions are effective in improving overall depression severity in various groups of participants. Although these findings are highly valuable, it is important to note that measures such as overall depression severity might not do justice to the diversity of depressive symptomatology.

An extensive review on this topic, for example, showed that individual depressive symptoms differ in various important dimensions, such as underlying biology and impact on impairment (Fried & Nesse, 2015a). In addition, the multifactorial structures of several commonly used depression scales are not stable over time and, consequently, scale scores may be inappropriate as outcome measures (Fried et al., 2016). Recent RCTs on the effectiveness of various depression treatments have also shown that individual depressive symptoms differ in their response

* Corresponding author. Department of Clinical, Neuro and Developmental Psychology, VU University, Van der Boechorststraat 7 (Room MF-A520), 1081, BT Amsterdam, the Netherlands.
E-mail address: l.boschloo@vu.nl (L. Boschloo).

https://doi.org/10.1016/j.brat.2019.103440
Received 30 November 2018; Received in revised form 1 July 2019; Accepted 16 July 2019
Available online 29 August 2019
0005-7967/

to treatment (Bekhuis et al., 2018; Blanken et al., 2019; Boschloo et al., 2019; Hieronymus, Emilsson, Nilsson, & Eriksson, 2016). It may, therefore, be valuable to consider individual symptoms as a more specific assessment of effectiveness, especially given the differences in clinical relevance across symptoms; for example, an effect on 'suicidal thoughts' or 'depressed mood' might be considered more relevant than an effect on 'concentration problems'.

1.2. The potential of network estimation techniques

Another advantage of a focus on individual symptoms is that it can help in identifying patterns in the diversity of symptomatology. An analysis of, for example, 3703 depressed patients revealed as many as 1030 unique symptom profiles based on combinations of depressive symptoms (Fried & Nesse, 2015b). Network estimation techniques specifically focus on individual symptoms and the complex patterns in which they typically co-occur; see, for example, Fried et al. (2017) for a review of empirical network studies and Epskamp, Borsboom, and Fried (2018) for a tutorial paper on network analyses.

Network estimation techniques can also be helpful in shedding light on the diversity of symptom-specific changes, for example, in response to treatment (Bekhuis et al., 2018; Blanken et al., 2019; Boschloo et al., 2019). Our study on the efficacy of antidepressant medication over cognitive behavioral therapy (Boschloo et al., 2019) showed that antidepressant medication was directly related to larger improvements in four symptoms, indicating that these larger improvements were independent of changes in other symptoms. The effect on one symptom was indirect, indicating that the larger improvement in that symptom was only present in patients who also experienced larger improvements in other symptoms. No significant differences between the two treatment conditions were found for another twelve symptoms.

The results of the network estimation techniques might also help in identifying participants who, based on their baseline symptomatology, would benefit more from a particular treatment relative to another. That is, participants primarily suffering from symptoms that are directly affected by the treatment would -by definition- benefit more from that treatment than participants primarily suffering from other symptoms. Network estimation techniques may therefore also show potential in "precision psychiatry".

1.3. Current study

This article focuses on the importance of considering individual symptoms as a more specific assessment of treatment effectiveness. In addition, we aim to explore the potential of network estimation techniques in shedding light on the complex patterns in which symptoms respond to treatment and in identifying participants who will benefit most from the intervention. We decided to focus on an internet-based intervention, as such treatments have great potential given their low-threshold accessibility, high anonymity, easy translatability and limited costs. The specific intervention, "Deprexis", has already proven to be effective for different outcomes and in different settings (e.g., Klein et al., 2016; Meyer et al., 2015; Schröder et al., 2014).

Data were derived from a large pragmatic RCT (N = 1013) considering the effectiveness of "Deprexis" relative to care as usual in participants with mild to moderate depressive symptomatology (Klein et al., 2016). This intervention has already proven to be successful in improving overall depression severity in this particular group (i.e., effect size within the small to moderate range) and we wanted to extend these findings by considering its effects on nine depressive symptoms as assessed with the Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer, & Williams, 2001). Secondly, we wanted to demonstrate the potential of network estimation techniques in shedding light on the diversity of symptom-specific changes and to examine direct and indirect symptom-specific effects. Lastly, we wanted to illustrate the potential of these network estimation techniques in identifying participants who, based on their baseline symptomatology, were likely to benefit more from the intervention relative to care as usual; that is, participants primarily suffering from symptoms that were directly affected by the internet-based intervention would -by definition- benefit more from the intervention than other participants.

2. Method

2.1. Trial design

This was a pragmatic multicenter trial, conducted at five sites in Germany, which was controlled, randomized and assessor-blinded. For the present study, we used data of the assessments conducted at baseline and at a 3-month post-assessment (the primary outcome of the trial); the monthly assessments following the post-assessment (i.e., 4–12 months of follow-up) were not used in the current study. The trial adheres to the guidelines of the CONSORT statement and its adaptation for internet interventions, CONSORTEHEALTH. It was approved by the Ethics Committee of the German Psychological Association (SM 04_2012) and registered in the ClinicalTrials.gov database (NCT01636752). The study protocol has been published (Klein et al., 2013) and no changes have been made to the protocol after study commencement. As extensive information on the trial design, participants and intervention has already been described in the study protocol (Klein et al., 2013) and the publication with the main findings (Klein et al., 2016), we only briefly describe these topics here.

2.2. Participants

This study included 1013 adults (18–65 years) with mild to moderate depressive symptoms, as assessed with the Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001). The PHQ-9 is a self-report questionnaire comprising nine items on the frequency of the nine symptoms of a DSM-IV major depressive disorder. All individual items are scored from 0 (never) to 3 (almost every day) and mild to moderate depressive symptomatology is defined as a total score between 5 and 14. Additional inclusion criteria were having internet-access and being able to communicate in German. Participants with acute suicidality (based on a structured assessment of current suicidal ideation and past suicide attempts; Hamilton, 1960; Sheehan et al., 1998; Posner et al., 2011) or a lifetime diagnosis of bipolar disorder or schizophrenia (based on the Mini International Neuropsychiatric Interview, MINI; Sheehan et al., 1998) were excluded. Neither a diagnosis of a depressive disorder nor current treatment was an inclusion criterion.

Participants were recruited via multiple settings, including inpatient and outpatient medical and psychological clinics, online forums for depression, health insurance companies and the media. Participants received extensive information about the aims and procedure of the study and were informed that they could withdraw at any time without having to disclose reasons. All participants provided (online) informed consent before the baseline assessment.

Of the 1013 included participants, 504 (49.8%) were randomly assigned to care as usual and 509 (50.2%) to the internet-based intervention (192 received the unguided and 317 received the guided version; stratification based on baseline severity of depressive symptomatology, see section 2.3). In total, 794 (78.4%) had complete post-assessment data on all individual depressive symptoms, with no difference between the care as usual versus intervention condition (79.2% versus 77.6%, p = 0.55; based on Chi-square test) and the unguided versus guided internet-intervention (80.2% versus 76.0%, p = 0.27; based on Chi-square test). Participants with missing post-assessment data did not differ from participants with complete data in overall depression severity, any of the nine individual depressive symptoms or any of the sociodemographics at baseline, except for age (41.1 versus 43.3 years, p = 0 .008; based on independent samples t-tests).

2.3. Treatment conditions

Participants were randomized equally (1:1) to either the internet-based intervention or care as usual (i.e., control) condition. Participants in both conditions were permitted to use any type of treatment (e.g., inpatient or outpatient psychiatric treatment), both before and during the study, and the investigators did not try to influence this. Participants in the intervention condition received access to the intervention in addition to their usual care. Participants in the care as usual condition received only care as usual and were offered access to the intervention after twelve months.

The internet-based intervention is a 12-week program (Deprexis) comprising ten modules on topics that are broadly consistent with CBT and one additional summary module (see Meyer et al., 2009 and Berger, Hammerli, Gubser, Andersson, & Caspar, 2011 for full details of the intervention). These modules were based on strategies such as 1) Behavioral Activation, 2) Cognitive Modification, 3) Mindfulness and Acceptance, 4) Interpersonal Skills, 5) Relaxation, Physical Exercise and Lifestyle modification, 6) Problem Solving, 7) Childhood Experiences and Early Schemas, 8) Positive Psychology Intervention, 9) Dreamwork and Emotion-Focused Intervention, 10) Psychoeducation. The program is interactive by engaging the participant in exercises and continuously eliciting feedback in order to tailor subsequent content. It also contains audio recordings, worksheets, summary sheets and brief automatic daily messages, delivered either by e-mail or SMS. All participants who did not log in to the program within two weeks after randomization received an e-mail describing the login procedure once again.

The internet-based intervention can be used with or without guidance by a clinician (Berger et al., 2011). Participants with mild depressive symptomatology (PHQ-score between 5 and 9) at baseline received the unguided version of the intervention, whereas participants with moderate symptomatology (PHQ-score between 10 and 14) received the guided version of the intervention because of safety and efficacy considerations. Participants receiving the guided version of the intervention were actively contacted once a week by a trained e-mail supporter who provided brief feedback based on the participant's use of the program over the previous week. Participants could respond to these messages or contact the e-mail supporter themselves. The primary goal of the e-mail guidance was to motivate participants to engage with the program. More details on the training and supervision for e-mail supporters are provided in the study protocol (Klein et al., 2013).

2.4. Outcome measures

The nine individual depressive symptoms were assessed with separate items of the PHQ-9 (Kroenke et al., 2001), both before treatment and at a 3-month post-assessment. All items assess the frequency of symptoms during the past two weeks and are scored from 0 (never) to 3 (almost every day). We chose the PHQ-9 for the assessment of individual symptoms, as the sum score of this questionnaire was the primary outcome in the original publication on the general effectiveness of the intervention (Klein et al., 2016).

2.5. Data analysis

All non-network analyses were performed using SPSS (version 24). First, baseline characteristics were compared between participants in the intervention versus care as usual condition using X^2 statistics for categorical variables (i.e., gender) and independent samples t-tests for continuous variables (i.e., age, overall depression severity and individual depressive symptom scores). Paired t-tests were performed to compare the post-assessment symptom scores to baseline symptom scores for the two conditions separately. Independent samples t-tests were performed to determine whether change scores of individual symptoms differed between the two treatment conditions. Given the exploratory nature of the study, no adjustments for multiple comparisons were made and, instead, all individual p-values were reported. A p-value of < 0.05 was considered significant.

All network analyses were performed using statistical software R (version 3.3.3). First, we examined whether the internet-based intervention had direct or indirect effects on changes in individual symptoms. For this purpose, a network including treatment condition and changes in individual depressive symptoms was estimated with package *mgm* (Haslbeck & Waldorp, 2019) using a mixed graphical model. This package uses the *glmnet* package (Friedman, Hastie, & Tibshirani, 2010) to fit penalized generalized linear models to perform neighborhood selection (Meinshausen & Bühlmann, 2006). Package *qgraph* (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012) was used to visualize the network. The resulting network shows the direct and indirect connections between treatment condition and changes in all individual symptoms. As a sensitivity analysis, we tested whether the internet-based intervention affected the patterns in which changes in symptoms were related. A network including changes in all nine depressive symptoms was estimated with l1-regularized partial correlations for the care as usual condition and intervention condition separately; the package *NetworkComparisonTest* (Van Borkulo, 2019) was used to test whether the networks differed.

In addition, we wanted to illustrate the potential of network estimation techniques in identifying those participants who, based on their baseline symptomatology, were likely to benefit more from the intervention relative to care as usual. Participants primarily suffering from symptoms that were directly affected by the internet-based intervention would -by definition- benefit more from the intervention than other participants. To illustrate and test this, we calculated a severity measure based on the weighted sum of all nine symptoms, in which the weighting was based on the direct effects of the intervention on individual symptoms. We expected that the effect of the intervention on overall depression severity would be larger in participants with higher scores on this weighted severity measure; effect modification was explicitly tested using linear regression analyses and a p-value of < 0.10 was considered significant.

Lastly, a set of sensitivity analyses was performed to explore potential differences in results between the unguided versus guided version of the internet-based intervention. Participants with mild depressive symptoms received the unguided version and participants with moderate depressive symptoms received the guided version, and the analyses were repeated for the two groups separately.

3. Results

3.1. Sample characteristics

Of the 794 included participants who had complete post-assessment data on all individual symptoms (i.e., the sample for our analyses; see section 2.2), 399 received care as usual and 395 the intervention. No differences between conditions were found for any of the sociodemographics, any of the treatment types before and during the study, overall depression severity or any of the nine individual depressive symptoms at baseline (Table 1).

3.2. Symptom-specific effectiveness of the internet-based intervention

Although overall depression severity improved significantly in both conditions (both $p < 0.001$), this improvement was significantly larger for the intervention than for care as usual (Table 2; Cohen's $d = 0.38$). All individual depressive symptoms also improved significantly in both conditions, except for 'suicidal thoughts' in the care as usual condition (deterioration of 0.03 points, $p = 0.30$). Of the nine symptoms, seven showed significantly larger improvements in the intervention condition relative to care as usual (Table 2; Cohen's d ranging from 0.15 for 'suicidal thoughts' to 0.32 for 'concentration problems'). No significant differences were found for the other two symptoms ('poor appetite or

Table 1
Sample characteristics.

	Care as usual (N = 399)	Intervention (N = 395)	
	mean (sd)/%	mean (sd)/%	p
Sociodemographics			
Female gender	67.9%	69.9%	0.55
Age at baseline	43.0 (10.7)	43.7 (10.9)	0.32
Treatment in the 6 months prior to baseline			
Psychotherapy	34.3%	37.2%	0.40
Outpatient psychiatric treatment	32.1%	29.4%	0.41
Inpatient psychiatric treatment	7.8%	8.4%	0.76
Treatment between baseline and post-assessment			
Psychotherapy	35.2%	32.2%	0.38
Outpatient psychiatric treatment	27.1%	26.9%	0.94
Inpatient psychiatric treatment	2.3%	1.3%	0.29
Antidepressant medication	51.3%	48.9%	0.50
Baseline depressive symptomatology			
Overall depression severity	10.3 (2.5)	10.2 (2.4)	0.63
Individual symptoms			
Little interest/pleasure	1.4 (0.7)	1.4 (0.6)	0.93
Feeling depressed/hopeless	1.3 (0.6)	1.2 (0.6)	0.27
Sleeping problems	1.6 (0.9)	1.5 (0.9)	0.69
Tired or having little energy	1.7 (0.8)	1.7 (0.7)	0.85
Poor appetite or overeating	1.1 (0.9)	1.1 (0.8)	0.97
Feeling guilty/worthless	1.3 (0.8)	1.3 (0.8)	0.66
Concentration problems	1.2 (0.7)	1.2 (0.7)	0.22
Psychomotor agitation/ retardation	0.5 (0.7)	0.5 (0.6)	0.17
Suicidal thoughts	0.2 (0.4)	0.2 (0.4)	1.00

overeating': p = 0.06, 'psychomotor retardation/agitation': p = 0.19).

3.3. The potential of network estimation techniques

In both conditions, individual participants experienced substantially different changes in their precise symptomatology; we observed 367 (of 399; 92.0%) unique combinations of the nine symptom change scores in the care as usual condition and 361 (of 395; 91.4%) unique combinations in the intervention condition. Here, we demonstrate the potential of network estimation techniques in shedding light on the diversity of symptom-specific changes in response to treatment and in identifying participants who would likely benefit more from the intervention relative to care as usual.

To provide more information about the direct and indirect symptom-specific effects of the intervention over care as usual, a network was estimated including treatment condition and changes in individual symptoms (Fig. 1; see Supplemental Table S1 for the underlying matrix with connection strengths). The network suggests that four of the seven previously identified symptom-specific effects were, at least partly, direct; suggesting that the larger improvements in 'feeling guilty/worthless', 'concentration problems', 'sleeping problems' and 'tired or having little energy' in response to the intervention were independent of changes in other symptoms. The previously identified symptom-specific effects on 'little interest or pleasure', 'feeling depressed/hopeless' and 'suicidal thoughts' appear to be indirect, suggesting that the larger improvements in these symptoms were only present in participants who reported larger improvements in the four symptoms that were directly affected by the intervention. It is important to note that also the four effects on 'feeling guilty/worthless', 'concentration problems', 'sleeping problems' and 'tired or having little energy' could be partly indirect.

As a sensitivity analysis, we tested whether the internet-based intervention affected the patterns in which changes in symptoms were related. The two networks including changes in all nine depressive symptoms did not differ for the care as usual versus intervention condition (Supplemental Fig. S1; p = 0.76 for global connectivity, and Holm-Bonferroni corrected p-values were all $\geq$0.32 for individual connections).

Participants primarily suffering from symptoms that are directly affected by the treatment benefit -by definition- more from that treatment than participants primarily suffering from other symptoms. To illustrate this, a severity indicator was calculated in which scores on baseline symptoms were summed and weighted based on the magnitude of the direct effects of the intervention. As the internet-based intervention was only directly related to 'feeling guilty/worthless' (connection strength: 0.13), 'concentration problems' (connection strength: 0.12), 'tired or having little energy' (connection strength: 0.11) and 'sleeping problems' (connection strength: 0.02), the severity indicator was only based on these four symptoms. As expected, the effect of the intervention on the change in overall depression severity depended on this baseline severity indicator (effect-modification: p = 0.02); the intervention was more effective across participants with increasing levels on this baseline severity indicator (Fig. 2; effect sizes ranging from 0.15 for the first quintile to 0.50 for the last quintile). As a comparison, another severity measure was calculated based on the simple sum of the five symptoms that were not directly affected by the intervention; as expected, the effectiveness of the intervention did not depend on this baseline severity measure (effect-modification: p = 0.81).

3.4. Unguided and guided version of the internet-based intervention

To explore whether the reported effects for the overall internet-based intervention were driven by the effects of either the unguided or guided version of the intervention, a set of sensitivity analyses was

Table 2
Improvements in depressive symptomatology in the intervention compared to care as usual condition.

	Care as usual (N = 399)	Intervention (N = 395)	p	Cohen's d
	mean (sd)	mean (sd)		
Overall depression severity	1.22 (4.31)	2.83 (4.10)	< 0.001	0.38
Individual symptoms				
Little interest/pleasure	0.21 (0.84)	0.36 (0.81)	0.01	0.18
Feeling depressed/hopeless	0.19 (0.81)	0.33 (0.80)	0.01	0.17
Sleeping problems	0.11 (1.08)	0.32 (0.98)	0.004	0.20
Tired or having little energy	0.16 (0.96)	0.45 (0.95)	< 0.001	0.30
Poor appetite or overeating	0.10 (0.96)	0.23 (0.93)	0.06	0.14
Feeling guilty/worthless	0.27 (0.87)	0.54 (0.86)	< 0.001	0.31
Concentration problems	0.07 (0.88)	0.34 (0.83)	< 0.001	0.32
Psychomotor agitation/retardation	0.14 (0.78)	0.21 (0.70)	0.19	0.09
Suicidal thoughts	−0.03 (0.53)	0.05 (0.53)	0.03	0.15

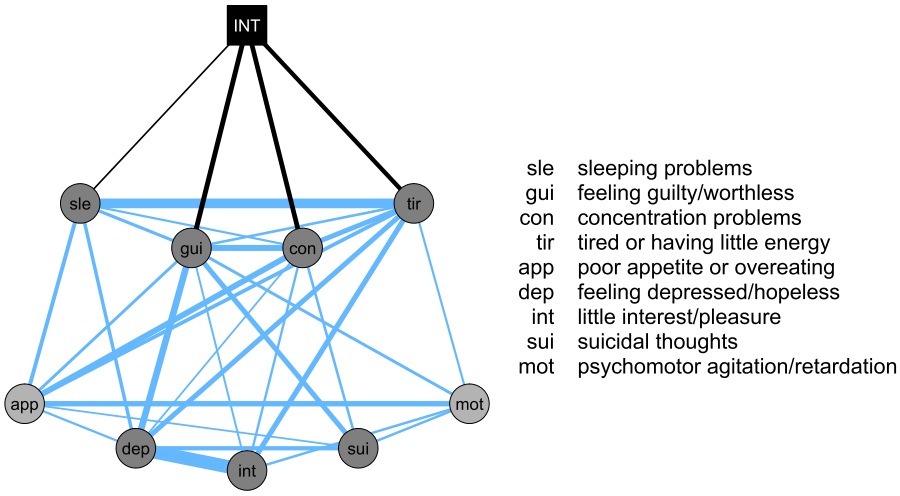

Fig. 1. Network structure including both treatment condition and improvements in individual depressive symptoms.
Black lines indicate direct connections between treatment condition and improvements in individual symptoms (all effects of the intervention are positive). Blue lines represent positive connections between improvements in symptoms, red lines represent negative connections between improvements in symptoms (not present in this network). Thicker lines represent stronger connections. Symptoms that showed larger improvements in the intervention relative to care as usual in the general analysis (see Table 2) are colored dark grey, while symptoms showing similar responses to the conditions are colored light grey. The network is presented at $\gamma = 0.25$. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

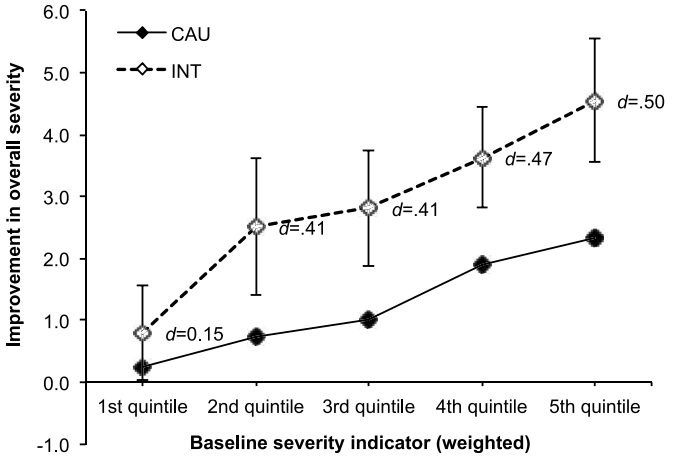

Fig. 2. Improvements in overall depression severity in the care as usual (CAU) versus intervention (INT) condition, stratified for participants with increasing scores on a baseline severity indicator (weighted).

performed in which results were stratified and compared for patients with mild depressive symptomatology (receiving the unguided version of the intervention or care as usual) or moderate depressive symptomatology (receiving the guided version of the intervention or care as usual). First, we determined the symptom-specific effectiveness of the two versions and found that the guided version had more significant effects than the unguided version (six versus three symptom-specific effects; see Supplemental Tables S2 and S3); effect sizes were, however, largely comparable for the two versions of the intervention. Second, we estimated a network including the treatment condition and changes in individual symptoms for patients with mild and moderate symptomatology separately and found that the networks did not differ ($p = 0.45$ for global connectivity, and Holm-Bonferroni corrected p-values were all ≥ 0.72 for individual connections). Lastly, we explored whether the weighted severity indicator based on the four symptoms that were directly affected by the intervention (in the primary analyses) was helpful in identifying those participants who would benefit more from either the unguided or the guided version of the intervention than from care as usual; both versions were more effective for participants with increasing levels on this severity indicator (effect modification: both were $p = 0.09$). The effectiveness of both the unguided and the guided version of the intervention were independent of the baseline severity indicator based on the five symptoms that were not directly affected by the intervention (effect-modification: $p = 0.69$ for the unguided and $p = 0.53$ for the guided version).

4. Discussion

4.1. Principal findings

This article focused on the importance of considering individual symptoms as a more specific assessment of treatment effectiveness as well as the potential of network estimation techniques in shedding light on the symptom-specific changes in response to treatment and in identifying participants who will benefit more from the one treatment relative to the other. We found that seven of the nine symptoms showed larger improvements in the internet-based intervention relative to the care as usual condition, and one was borderline significant. The network estimation techniques further suggested that the internet-based intervention had a direct effect on four symptoms and the effects on the other three symptoms were, thus, indirect. In addition, we illustrated the potential of these network estimation techniques in identifying participants who were likely to benefit more from the intervention relative to care as usual; as expected, the intervention was more effective for participants suffering from those symptoms that were directly affected by the intervention.

4.2. Symptom-specific effectiveness of the internet-based intervention

Previous studies have already demonstrated that the internet-based intervention Deprexis is superior to care as usual in improving overall depression severity after three and six months (Klein et al., 2016) as well as over the course of one year (Klein et al., 2017). Klein et al. (2016), for example, showed that the internet-based intervention was effective in improving overall depression severity after three months, with an effect size within the small to moderate range. The current study is an important extension, as it demonstrated that the intervention had an effect on seven depressive symptoms after three months, including symptoms that are highly clinically relevant such as 'feeling depressed/hopeless' and 'suicidal thoughts'. It is important to note that the seven symptoms differed in their response to the intervention, with effect sizes ranging from 0.15 for 'suicidal thoughts' to 0.32 for 'concentration problems'. No effects were found for 'poor appetite or overeating' (borderline significant) and 'psychomotor agitation/retardation' (not significant). As studies on other treatment types have also observed substantial differences across symptom-specific effects, we believe that considering individual symptoms as a more detailed assessment of effectiveness would be an important step forward; especially, since it could inform participants more precisely about the preferred treatment option.

4.3. The potential of network estimation techniques

Fried and Nesse (2015b) previously showed that patients with the same diagnosis of major depressive disorder differed substantially in their exact symptomatology. The response of individual symptoms to the internet-based intervention as well as care as usual was also diverse, as we found that more than 90% of participants in each condition were unique in their combination of changes in individual symptoms. Information about this diversity is largely lost when symptoms are combined into scale scores of overall severity.

Network estimation techniques showed that the internet-based intervention was, at least partly, directly related to larger improvements in 'feeling guilty/worthless', 'concentration problems', 'sleeping problems' and 'tired or having little energy', indicating that these larger improvements were independent of changes in other symptoms. This could be explained by the content of the intervention; the modules on Cognitive Modification as well as Mindfulness and Acceptance may have resulted in decreases in 'feeling guilty/worthless', whereas the module on Relaxation, Physical Exercise and Lifestyle modification may have been beneficial for 'sleeping problems', 'tired or having little energy' and 'concentration problems'. The effects on 'little interest or pleasure' and 'feeling depressed/hopeless' and 'suicidal thoughts' were indirect, suggesting that these larger improvements were only present in participants who also reported larger improvements in the four symptoms that were directly affected by the intervention.

The results from these network estimation techniques have potential in generating hypotheses regarding the working mechanisms of treatment, as they may suggest a causal chain in which changes in specific symptoms result in changes in specific other symptoms. For example, we found that an improvement in 'feeling guilty/worthless', which was directly affected by the internet-based intervention, was related to an improvement in 'feeling depressed/hopeless' and 'suicidal thoughts', suggesting that improved self-worth may have resulted in a decrease in depressed mood and suicidality. Other potential causal chains are that, for example, an improvement in 'sleeping problems' may cause an improvement in 'being tired or having little energy', and an improvement in 'being tired or having little energy' may cause an improvement in 'little interest or pleasure' as well as 'feeling depressed/hopeless'.

It is, however, important to note that changes in symptoms were assessed simultaneously and that the temporal relationships between them remain unknown. In addition, the network estimation techniques are not intended to formally test for mediation. For example, the regularization techniques set weak connections to zero and, consequently, conservatively identify the most relevant connections in the network. A weak direct connection between the internet-based intervention and, for example, an improvement in 'feeling depressed/hopeless' or 'suicidal thoughts' could be set to zero and thus, in reality, this effect would not be fully indirect. To examine whether the hypothesized causal chains are really present and, more fundamentally, to determine whether a network model or another type of model (e.g. a common cause or hybrid model) best describes the dynamics of symptoms, it would be more appropriate to use data from the experience sampling method including large numbers of assessments with short time intervals. In addition, it would be interesting to include comorbid symptomatology (e.g. anxiety symptoms), other outcomes (e.g., quality of life) or other factors that are believed to play a role in the working mechanisms of treatment (e.g., social support) (see paragraph 4.4 for a more thorough discussion of this topic).

As individual depressive symptoms differed in their response to the internet-based intervention, we expected that participants primarily suffering from symptoms that were directly affected by the intervention would benefit more from the intervention relative to care as usual. This was also illustrated by our findings; participants who scored higher on the symptoms 'feeling guilty/worthless', 'concentration problems', 'sleeping problems' and 'tired or having little energy' benefitted most from the intervention (effect sizes ranging from 0.41 to 0.50). In contrast, no significant difference in the effectiveness of the intervention was found for participants scoring the lowest on these four symptoms. These findings are an important step forward in "precision psychiatry", as they can inform participants more precisely about the effectiveness of the intervention based on their baseline symptomatology.

4.4. Strengths and weaknesses of the study and suggestions for further research

Strength of the current study is that we considered individual depressive symptoms as a more detailed assessment of the effectiveness of the internet-based intervention relative to care as usual. It is, however, important to note that much is still unclear about the reliability and validity of assessing individual symptoms. An advantage of the PHQ-9 is that all items have the same answering categories and that the sensitivity to detect changes in symptom frequency is thus, at least in theory, comparable across symptoms. Previous studies on the symptom-specific efficacy of other treatment types often used the Hamilton Depression Rating Scale with its varying response categories and, in general, larger effects were observed for symptoms with more response categories (Boschloo et al., 2019; Hieronymus et al., 2016).

A weakness of the PHQ-9 is that opposite symptoms are aggregated into one item; for example, poor appetite and overeating are combined into a single item and the same is true for psychomotor agitation and retardation as well as hypersomnia and insomnia. Aggregation of these symptoms may have distorted the results regarding the symptom-specific effects of the intervention and could provide an alternative explanation for not finding significant treatment effects on 'poor appetite and overeating' and 'psychomotor agitation and retardation'. It may also have impacted the estimated networks. Previous network studies showed that disaggregated symptoms were connected to different symptoms (Boschloo et al., 2015); for example, psychomotor agitation was connected to insomnia but not hypersomnia, whereas psychomotor retardation was connected to hypersomnia and not insomnia. Therefore, we would like to advice other researchers to consider disaggregated symptoms in their research on treatment effectiveness.

Another weakness of the PHQ-9 is that it focuses exclusively on the nine depressive symptoms and we believe that it would be valuable to adopt a transdiagnostic approach by also considering comorbid symptomatology (Boschloo, 2018). Previous studies on the symptom-specific efficacy of other depression treatments observed, for example, additional effects on specific anxiety and somatic symptoms (Bekhuis et al., 2018; Boschloo et al., 2019; Hieronymus et al., 2016), and also other symptomatology, such as alcohol problems, might be valuable to consider. In addition, it would be interesting to include other clinically relevant outcomes (e.g., aspects of daily functioning or quality of life), as all this information combined may help in making a deliberate decision about starting a particular treatment or not. In future research on the potential working mechanisms of treatment, it would be interesting to not only focus on symptomatology or other outcomes but also to consider factors that have been hypothesized to play a role in the working mechanisms of treatment (e.g., locus of control or social support).

A limitation of this study is that the analyses are based on the participants with complete data at baseline and 3-month post-assessment and not the intention-to-treat sample. It is, however, unlikely that this has substantially affected the results, as the observed effectiveness of the intervention in improving overall depression severity is highly comparable for the two sample (i.e., effect size of 0.38 in our sample versus 0.39 in the intention-to-treat sample as published by Klein et al., 2016). In addition, none of the nine individual depressive symptoms at baseline were a significant predictor of missing data at the post-assessment.

Lastly, it is important to emphasize that our findings are based on one randomized controlled trial. Replication studies in similar as well as

different settings are needed to prove the validity of our findings. In addition, it is important to emphasize that the current study is based on a *pragmatic* RCT aiming to examine the effectiveness of the internet-based intervention in a real-world situation. Such a design is not appropriate for determining the working mechanisms of treatment, as it does not adequately control for potential confounding factors such as using other types of treatment before or during the study or differences between the unguided and guided version of the intervention.

5. Conclusion

The current study demonstrated the importance of considering individual symptoms as a more specific assessment of treatment effectiveness. We found that individual depressive symptoms differed considerably in their response to the internet-based intervention Deprexis relative to care as usual; seven symptoms showed larger improvements in the internet-based intervention (effect sizes ranging from 0.15 to 0.31), whereas no significant differences were found for the two other symptoms. We also demonstrated the potential of network estimation techniques in shedding light on the symptom-specific changes in response to treatment and in identifying participants who are likely to benefit more from the intervention than from care as usual. Four of the seven symptom-specific effects were, at least partly, direct, whereas the effects on the other three symptoms were, thus, indirect. As expected, participants with higher scores on those symptoms that were directly affected by the intervention improved most from the internet-based intervention relative to care as usual (effect sizes as high as 0.50), whereas no difference in effectiveness was found for participants scoring the lowest on these four symptoms.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2019.103440.

References

Bekhuis, E., Schoevers, R. A., De Boer, M. K., Peen, J., Dekker, J., Van, H., et al. (2018). Symptom-specific effects of psychotherapy versus combined therapy in the treatment of mild to moderate depression: A network approach. *Psychotherapy and Psychosomatics, 87*, 121–123.

Berger, T., Hammerli, K., Gubser, N., Andersson, G., & Caspar, F. (2011). Internet-based treatment of depression: A randomized controlled trial comparing guided with unguided self-help. *Cognitive Behaviour Therapy, 40*, 251–266.

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*, 52–54.

Boschloo, L. (2018). Zooming in and zooming out: A network perspective on the comorbidity of depression and anxiety. *Journal of the American Academy of Child & Adolescent Psychiatry, 57*, 914–915.

Boschloo, L., Bekhuis, E., Weitz, E. S., Reijnders, M., DeRubeis, R. J., Dimidjian, S., et al. (2019). The symptom-specific efficacy of antidepressant medication vs. cognitive behavioral therapy in the treatment of depression: Results from an individual patient meta-analysis. *World Psychiatry, 18*, 183–191.

Boschloo, L., Van Borkulo, C. D., Rhemtulla, M., Keyes, K. M., Borsboom, D., & Schoevers, R. A. (2015). The network structure of symptoms of the diagnostic and statistical manual of mental disorders. *PLoS One, 10*, e0137621.

Churchill, R., Hunot, V., Corney, R., Knapp, M., McGuire, H., Tylee, A., et al. (2001). A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technology Assessment, 5* 1-173.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods, 50*, 195.

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*, 1–18.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*, 1–22.

Fried, E. I., & Nesse, R. M. (2015a). Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. *BMC Medicine, 13* 72.

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

Fried, E. I., Van Borkulo, C. D., Cramer, A. O., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology, 52*, 1–10.

Fried, E. I., Van Borkulo, C. D., Epskamp, S., Schoevers, R. A., Tuerlinckx, F., & Borsboom, D. (2016). Measuring depression over time… or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychological Assessment, 28*, 1354–1367.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry, 23*, 56–62.

Haslbeck, J. M. B., & Waldorp, L. J. (2019). mgm: estimating time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software,* 1–49. Retrieved from https://arxiv.org/pdf/1510.06871.pdf.

Hieronymus, F., Emilsson, J. F., Nilsson, S., & Eriksson, E. (2016). Consistent superiority of selective serotonin reuptake inhibitors over placebo in reducing depressed mood in patients with major depression. *Molecular Psychiatry, 21*, 523–530.

Karyotaki, E., Riper, H., Twisk, J., Hoogendoorn, A., Kleiboer, A., Mira, A., et al. (2017). Efficacy of self-guided internet-based cognitive bevavioral therapy in the treatment of depressive symptoms: A meta-analysis of individual participant data. *JAMA Psychiatry, 74*, 351–359.

Klein, J. P., Berger, T., Schröder, J., Späth, C., Meyer, B., & Caspar, F. (2016). Effects of a psychological internet intervention in the treatment of mild to moderate symptoms: Results of the EVIDENT Study, a randomized controlled trial. *Psychotherapy and Psychosomatics, 85*, 218–228.

Klein, J. P., Berger, T., Schröder, J., Späth, C., Meyer, B., Lutz, W., et al. (2013). The EVIDENT-trial: Protocol and rationale of a multicenter randomized controlled trial testing the effectiveness of an online-based psychological intervention. *BMC Psychiatry, 13*, 239.

Klein, J. P., Späth, C., Schröder, J., Meyer, B., Greiner, W., Hautzinger, M., et al. (2017). Time to remission from mild to moderate depressive symptoms: One year results from the EVIDENT-study, an RCT of an internet intervention for depression. *Behavioral Research and Therapy, 97*, 154–162.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*, 606–613.

Meinshausen, N., & Bühlmann, P. (2006). High-dimensional graphs and variable selection with the Lasso. *Annals of Statistics, 34*, 1436–1462.

Meyer, B., Berger, T., Caspar, F., Beevers, C. G., Andersson, G., & Weiss, M. (2009). Effectiveness of a novel integrative online treatment for depression (Deprexis): Randomized controlled trial. *Journal of Medical Internet Research, 11* e15–e15.

Meyer, B., Bierbrodt, J., Schröder, J., Berger, T., Beevers, C. G., Weiss, M., et al. (2015). Effects of an internet intervention (Deprexis) on severe depression symptoms: Randomized controlled trial. *Internet Interventions, 2*, 48–59.

Mulrow, C. D., Williams, J. W., Jr., Chiquette, E., Aguilar, C., Hitchcock-Noel, P., Lee, S., et al. (2000). Efficacy of newer medications for treating depression in primary care patients. *The American Journal of Medicine, 108*, 54–64.

Posner, K., Brown, G. K., Stanley, B., Brent, D. A., Yershova, K. V., Oquendo, M. A., et al. (2011). The columbia suicide severity rating scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. *American Journal of Psychiatry, 168*, 1266–1277.

Schröder, J., Brückner, K., Fischer, A., Lindenau, M., Köther, U., Vettorazzi, E., et al. (2014). Efficacy of a psychological online intervention for depression in people with epilepsy: A randomized controlled trial. *Epilepsia, 55*, 2069–2076.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weillerm, E., et al. (1998). The mini-international neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*, 22–33.

Van Borkulo, C. D. (2019). Statistical Comparison of Two Networks Based on Three Invariance Measures. *R package.* https://cran.r-project.org/web/packages/NetworkComparisonTest/NetworkComparisonTest.pdf.