Differences in depressive symptom severity and architecture between men and women. A network approach.

Tarja Tuulikki Weckström

Master's Thesis

Psychology

Faculty of medicine

November 2019

Supervisor: Marko Elovainio

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildingsprogram – Degree Programme
Lääketieteellinen tiedekunta	Psykologian maisteriohjelma

Tekijä – Författare – Author

Tarja Weckström

Työn nimi – Arbetets titel – Title

Differences in depressive symptom severity and architecture between men and women. A network approach.

Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track

Psykologia

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Pro gradu-tutkielma	Marraskuu - 2019	47

Tiivistelmä – Referat – Abstract

Tavoitteet: Verkostoanalyysin teoreettisessa viitekehyksessä mielenterveysongelmia, kuten depressiota, käsitteellistetään yksilön eri oireiden yhteisvaikutuksen ilmentyminä. Depressio on yksi maailman vakavimmista terveyshaitoista ja oireverkostojen tarkasteleminen voi lisätä ymmärrystä depression etiologiasta. Tutkimuksen tavoitteena on lisätä ymmärrystä naisten tunnetusti suuremmasta taipumuksesta sairastua depressioon vertaamalla normaaliväestön naisten ja miesten depressio-oireverkostoja. Depression sukupuolieroja ollaan tutkittu laajasti, mutta tämä tutkimus tarkastelee ensimmäisenä oireverkostojen sukupuolieroja normaaliväestössä.
Menetelmät: Aineisto (miehet n=576; naiset n=866) on kerätty kansallisen LASERI-tutkimuksen yhteydessä. Tutkimuksessa vertaillaan miesten ja naisten Beckin Depressiokyselyn (BDI-II) muuttujien muodostamia osittaiskorrelaatio-oireverkostojen ominaisuuksia kahdessa aikapisteessä.
Tulokset: Oireverkostoissa esiintyi eroavaisuuksia sukupuolten välillä. Miehillä oli alhaisemmat oirekeskiarvot, mutta näkyvämpiä muutoksia oireiden yhteyksien välillä eri aikapisteissä, enemmän ja voimakkaampia negatiivia yhteyksiä sekä tiheämpi oireverkosto. Levottomuus näyttäytyi keskeisenä oireena ainoastaan miesten verkostoissa. Naisilla oireverkosto pysyi samankaltaisena ja oireet muodostivat kolme samaa oirerypästä molemmissa aikapisteissä, vaikka oirekeskiarvot poikkesivat tilastollisesti toisistaan eri aikapisteiden välillä. Väsymys näyttäytyi erityisen keskeisenä ainoastaan naisten verkostossa. Naiset myös raportoivat väsymystä tilastollisesti miehiä enemmän.
Johtopäätökset: Sukupuolten väliset erot oireiden välisissä yhteyksissä voivat selittää masennuksen esiintyvyyden sukupuolieroja. Miesten oireverkoston negatiiviset yhteydet ja naisten oireverkoston kolme oirerypästä ovat kohteita lisätutkimukselle. Väsymykseen kohdennetut toimet voivat olla oleellisia naisten masennuksen ennaltaehkäisyssä ja hoidossa.

Avainsanat – Nyckelord – Keywords

Verkostoanalyysi, masennus, sukupuolierot, masennusoireet, väsymys

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors

Marko Elovainio

Säilytyspaikka – Förvaringställe – Where deposited

Muita tietoja – Övriga uppgifter – Additional information

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildingsprogram – Degree Programme
Faculty of medicine	Psychology

Tekijä – Författare – Author

Tarja Weckström

Työn nimi – Arbetets titel – Title

Differences in depressive symptom severity and architecture between men and women. A network approach.

Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track

Psychology

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Master's thesis	November - 2019	47

Tiivistelmä – Referat – Abstract

Objectives: Depression presents one of the biggest global health concerns today. According to the network theory, mental disorders, such as depression, reflect co-occurring intercorrelating symptom effects. Thus, studying the properties of depressive symptom networks could enhance knowledge about the etiology of depression. In this study, network structures of men and women of the general population are compared, to enhance understanding of higher prevalence rates of depression in women. Although gender differences of depression are widely studied, this is the first study comparing the depressive symptom network structures of adult men and women in the general population.

Methods: The data (n = 567 men; n= 886 women) are from a national age cohort study (LASER). Partial correlation networks of BDI-II symptoms were compared in two time-points.

Results: Estimated networks had distinct gender differences. Men had lower mean scores, but more changes in the network structure across time, more negative edges, and higher network density. "Agitation" was highly central only in the men's networks. Women showed changes in mean sum scores, but network structures had few changes, and symptoms formed three distinct communities in both time-points. "Fatigue" was reported significantly more by women and was highly central only in the women's networks.

Implications: Differences in symptom networks between men and women may explain the gender-related differences in the prevalence of depression. Negative edges in the men's networks and the symptom communities in the women's networks are targets for more research. Fatigue could be a valuable target for preventing and treating women's depression.

Avainsanat – Nyckelord – Keywords

Network analysis, depression, gender difference, fatigue

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors

Marko Elovainio

Säilytyspaikka – Förvaringställe – Where deposited

Muita tietoja – Övriga uppgifter – Additional information

Abstract

Objectives: Depression presents one of the biggest global health concerns today. According to the network theory, mental disorders, such as depression, reflect co-occurring intercorrelating symptom effects. Thus, studying the properties of depressive symptom networks could enhance knowledge about the etiology of depression. In this study, network structures of men and women of the general population are compared, to enhance understanding of higher prevalence rates of depression in women. Although gender differences of depression are widely studied, this is the first study comparing the depressive symptom network structures of adult men and women in the general population.

Methods: The data (n = 567 men; n= 886 women) are from a national age cohort study (LASER). Partial correlation networks of BDI-II symptoms were compared in two time-points.

Results: Estimated networks had distinct gender differences. Men had lower mean scores, but more changes in the network structure across time, more negative edges, and higher network density. "Agitation" was highly central only in the men's networks. Women showed changes in mean sum scores, but network structures had few changes, and symptoms formed three distinct communities in both time-points. "Fatigue" was reported significantly more by women and was highly central only in the women's networks.

Implications: Differences in symptom networks between men and women may explain the gender-related differences in the prevalence of depression. Negative edges in the men's networks and the symptom communities in the women's networks are targets for more research. Fatigue could be a valuable target for preventing and treating women's depression.

TABLE OF CONTENTS

1 INTRODUCTION

Depression is one of the lead causes of disability in western societies, and over 300 million people are affected by it worldwide (WHO, 2018). It represents one of the most profound global health problems humans face today (Greden, 2001). Despite decades of research, depression remains challenging to treat, depressive episodes are often reoccurring, and residual symptoms after remission common (Holtzheimer & Nemeroff, 2006). Depression manifests in an array of cognitive, social, emotional, and physical symptoms, and the costs are high to the individuals affected by it, as well as to the society (Greden, 2001). Women are more likely to suffer from depression, showing around a 2:1 ratio of major depression diagnoses compared to men, and the gender difference in depression is well established and seen across nations (Salk, Hyde & Abramson, 2017).

Individuals with a depression diagnosis can express a multitude of symptom combinations. In a recent study, of 3703 clinically assessed depressed individuals, it was found that only 1.8% share the same combination of depressive symptoms (Fried & Nesse 2015a). Although different symptoms have different social, cognitive, and neurobiological origins, resulting in a heterogeneous clinical picture of depression, research does not traditionally focus on the symptom level. The network approach to mental disorders, is a new framework for understanding and studying mental disorders, which emphasizes the role of individual symptoms and relationships between symptoms (Borsboom & Cramer, 2013; Fried, 2015). This study aims to enhance understanding of the gender difference in depression by investigating the architecture and severity of depressive symptoms of men and women.

1.1 MAJOR DEPRESSIVE DISORDER

Depressive disorders, particularly the main diagnosis of ‘major depressive disorder' (MDD), belongs to the family of mood disorders, and is characterized by life disruptive symptoms, such as negative emotionality, sadness, hopelessness, and anhedonia. According to the current edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2015; DSM-V) an episode of major depression can be diagnosed when five out of 9 possible symptoms have persisted for at least two weeks and are causing the individual clinically significant distress and impairment. One of the symptoms must be 1) depressed mood or 2) loss of interest, which constitute higher-level symptoms of MDD diagnosis. The other seven symptoms are 3) significant weight loss or gain, or a decrease or increase in appetite, 4) insomnia or hypersomnia 5) an increase or a slowing

down /reduction of thought and of physical movement (psychomotor agitation or retardation), 6) fatigue or loss of energy, 7) feelings of worthlessness or excessive or inappropriate guilt, 8) diminished ability to think or concentrate, or indecisiveness, and 9) recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide. In a recent Epidemiology of Adult DSM-5 Major Depressive Disorder based on 36 309 US adult participants (Hasin et al., 2018) the estimated 12-month prevalence rate of MDD is 10.4% and the lifetime prevalence rate is 20.6% respectively. According to the same study, co-occurrence with other disorders is high, ranging from 36.4% comorbidity with anxiety disorders, 40.9% with personality disorders, and 45.3% with substance use. According to a World Health Organization estimate, the prevalence of depressive disorders in the European region ranges between 3.5-5% of all citizens (WHO, 2017).

The cost of depression can be devastating for the individuals affected by it, as it causes impairment in cognitive (Greden, 2001), emotional (Beck, 2002), social (Kessler, 2012) and physical functioning (Moussavi et al., 2007). The effects can be debilitating and be disruptive to crucial areas of life. Often episodes of depression coincide with transitional periods of development, which can lead to termination of education, marital dissatisfaction, divorce, negative parenting behavior, and reduced work performance (Kessler, 2012). In the WHO World Health Survey, based on observations from 245 404 participants from 60 countries, depression was seen to have the most significant decrement in health, when compared to other chronic diseases, such as angina, diabetes, arthritis or asthma (Moussavi et al., 2007). According to the same study, a physical illness with comorbid depression is also associated with worse health compared to just having one or the other. Depression can ultimately lead to suicide, and nearly 800 000 people die due to suicide annually (WHO, 2015). The cost of depression among other mental illnesses is significant also in monetary value. An OECD report estimated that the cost of mental illness amounts to more than 4% across EU countries, equivalent to over 600 billion euros per year. These costs include health service costs, social security programs, lost earnings and lower productivity (OECD/EU, 2018). Additional costs relating to informal care that is provided by family and friends, or other public service costs are harder to estimate (Center for Mental Health, 2010). Hence, developing more effective ways to prevent and treat mental illnesses, such as depression, remains a highly relevant topic in research.

1.2 GENDER DIFFERENCES IN DEPRESSION CHARACTERISTICS

Women are found more susceptible to depression across nations compared to men (Weissman, 1993; Angst et al., 2002; Van de Velde, Bracke & Levecque, 2010). The gender difference is seen in diagnoses of MDD, as well as in the severity of depression symptoms. In recent large meta-analyses, based on representative cross-national samples, females showed a higher ratio of depression diagnoses (OR= 1.95; d=0.37) and larger mean differences in depression symptoms (d=0.27) compared to men (Salk, Hyde & Abramson, 2017). There is a consistency in both analyses that the gender difference peaks in adolescence, then declines and remains stable until late mid-life (at around d=0.19-0.3 for both diagnoses and symptoms) (Salk, Hyde & Abramson, 2017).

Besides overall severity in symptoms sum scores, significant differences are reported in various symptoms between genders. Schuch et al. (2014) compared 30 depression symptoms in a nationally representative sample based on and found 11/30 significant differences in symptoms between men and women. The differences included men experiencing more "sad mood" ($p<.01$) and "loss of interest" ($p=.01$), and women experiencing more "increased weight" ($p<.01$), somatic (pain)complaints ($p=.01$), panic/phobic symptoms ($p=.02$), gastrointestinal complaints (diarrhea, constipation) ($p=.04$), increased interpersonal sensitivity ($p<.01$) and "mid-nocturnal insomnia" ($p=.01$). Women are seen to have more atypical symptoms of depression, such as hypersomnia and increase in weight/appetite (Angst et al., 2002) and present more comorbidity with anxiety-related symptoms and lower social functionality, whereas men consume more alcohol (Kornstein et al., 1995; Schuch et al., 2014).

Besides differences in prevalence rates of diagnoses and symptoms, neurobiological evidence indicates that processes on the molecular and neurobiological level, associated with MDD, vary depending on gender. Firstly, neurochemical differences are seen in growth factors, which are involved in regulating neural plasticity and seen in lower levels in depressed patients compared to healthy controls. The Brain-derived neurotrophic factors (BDNF) is seen associated with the duration of a depressive episode in women, but not in men, and in contrast the nerve growth factor (NGF) shows significant association with the duration of an MDE with men, but not in women respectively (Cardoso et al., 2014). Alterations in associations to MDD depending on gender are also seen in the functions of the endocrine system (e.g., the hypothalamic-pituitary–adrenal-axis (HPA-axis) and/or sex steroids), inflammation, hormones and metabolic functioning (e.g. leptin or glucocorticoids),

which all show clear associations with mood disorders (Schmidt, Shelton & Duman, 2011; Pariante, 2009; Pariante & Miller, 2001; Hryhorczuk, Sharma & Fulton, 2013).

For example, in line with the fact that depressed women show more symptoms of appetite changes, depressed women have a two-fold risk of having metabolic abnormalities compared to men (Hryhorczuk, Sharma & Fulton, 2013). There is also clear evidence that dysregulation of the HPA-axis is associated with mood disorders (Pariante, 2009; Pariante & Miller, 2001), and women seem to be more susceptible to the dysregulation of the HPA-axis (Weiss, Longhurst & Mazure, 1999). Inflammation molecules, on the contrary, are seen in increased levels in depression in men, but not in women, suggesting that men might be more prone to an inflammatory basis of depression (Elovainio et al., 2009; Ramsey et al., 2014). Sex steroids in men and gonadal hormones in women are linked to several neurobiological mechanisms and associations to MDD, including gender-related vulnerabilities to comorbidity of substance use in men and atypical symptoms in women (Walther et al., 2017; Halbreich & Kahn, 2007). The gender difference seen in the prevalence rates of depression are likely to be a mixture of biological, social and psychological factors. Investigating the gender difference on a symptom level may enhance our understanding of the developmental differences in depressive episodes and enable more specific hypothesis and targets for interventions and prevention.

1.3 PROBLEMS IN TRADITIONAL DIAGNOSTIC CLASSIFICATION

The current diagnostic system is an attempt to categorize and classify recognized symptom patterns that are seen in psychiatric patients. It is acknowledged that some symptoms, such as symptoms of MDD, are more likely to co-occur together compared to symptoms under another diagnostic label. Although having explicit diagnostic criteria have helped advance research and clinical practice, there are also concerns relating to the conceptual development of the current diagnostic system (Regier, Narrow, Kuhl & Kupfer, 2009). The biggest concern relates to high comorbidity between disorders, which implies that there is no clear separation between disorders. A problem, which follows from the overlapping of different disorders, is that individuals sharing the same diagnostic labels have many altering symptoms.

The problem of heterogeneity is also apparent in the context of MDD (Goldberg, 2011). Numerous different combinations can lead to a diagnosis of MDD according to the DSM-5, and two patients might only share one symptom in common. Some symptoms include their opposites, such as hypersomnia or insomnia, weight gain or weight loss, restlessness, or retardation of movements. To

classify the diverse clinical presentations of depression, the DSM-5 includes subtypes of major depression, with atypical, anxious, psychotic, or mixed features. Depressive disorders differ in their combinations of symptoms, onset or level of severity, and often, although more ambiguously, a diagnosis may include additional symptoms, unspecified labels, or comorbid diagnoses. As a result, patients with a depression diagnosis vary considerably in clinical presentations, genetics, neurobiology, clinical course, treatment responsiveness, and biological correlates (Rush, 2007; Fried & Nesse, 2015a).

So far, there are no biomarkers that are used to predict clinical outcomes, the development of effective treatments has been slow, and treating MDD is still very challenging (Holtzheimer, 2006). The effectiveness of a single treatment of anti-depressant medication in non-chronic patients with MDD is only at around 30% - 45% (Thase, Entsuah, & Rudolph, 2001). Moreover, approximately 20-30% of individuals with MDD have chronic treatment-resistant depression after sequential drug- and therapy interventions (Howland, 2014). Even for those who respond to treatment, relapse is prevalent, and the majority of individuals with MDD have residual symptoms (Fava & Rush, 2006). In clinical work, the heterogeneity of clinical presentations is recognized, as people rarely depict a clear cut "depression" or "anxiety," but more than often, they have multiple problems (Frank & Davidson, 2014). Many patients also experience problems, which are clinically significant but do not meet any existing diagnostic category (Löwe et al., 2008). In clinical work, clinicians have developed ways to get around complex issues and comorbidity by targeting common cognitive, emotional and behavioral mechanisms, that are seen across various disorders (Frank & Davidson, 2014).

The current diagnostic model assumes that the covariation between symptoms is due to a latent common cause, e.g. ‘major depression'. This traditional conceptualization of mental disorders represents an influential 20th-century single perspective explanatory model of psychopathology, which has been argued to resemble the disease model of medical science (Kendler, 2008; Borsboom & Cramer, 2013). According to the disease model, a distinct medical entity (such as a tumor) can cause symptoms, such as headaches, forgetfulness, and foggy eyesight (Borsboom & Cramer, 2013). However, mental disorders are increasingly seen to be influenced by multicausal psychological, social, and biological processes both within and outside the individual, which are difficult to explain to result from a common cause perspective. As a result, in the field of psychiatry and clinical psychology, the analogy is challenged. New explanations have become more pluralistic and multilevel (Kendler, 2008; Cramer 2013; Borsboom, Cramer & Kalis, 2018), that can better account for the high comorbidity rates and the psychological phenomena, that are seen across traditional

diagnostic categories. Examples of prominent new frameworks for studying psychopathology include higher-order externalizing and internalizing dimensions (Kessler, Petukhova & Zaslavsky, 2011), the RDoC project (NIMH, 2008), or the network approach to mental disorders which places symptoms in the focal point of psychopathology (Borsboom, Cramer & Kalis, 2018; Borsboom & Cramer, 2013).

1.4 DEPRESSION AS A DYNAMIC SYSTEM

The network approach to mental disorders is a new framework for studying and conceptualizing psychopathology as dynamic symptom networks. Symptoms are seen to have direct relationships, which causes covariation between symptoms. In other words, mental disorders are conceptualized, as the emergent property of co-occurring and intercorrelating (and causal) symptom effects, instead of latent variables (Borsboom, 2017; Fried, 2015). Symptoms can be activated by other symptoms and/or a multitude of internal or external factors (e.g., loss of a spouse, catastrophic thinking or neurobiological dysfunction), making the connections highly multifactorial (Borsboom, Cramer & Kalis, 2019). For example, an external event of stress may lead to concentration problems, which may lead to sleep problems, which may lead to fatigue and depressed mood. The displays of current disorders are explained as endless variations of complex networks of symptoms, where the focus of research is placed on the connections between items in the network.

The relationships between symptoms are the basis for understanding mental health and illness in the network framework. When symptoms connect strongly, the activation of one symptom (e.g., worry) can lead to the activation of other symptoms more easily (e.g. loss of sleep, anxiety, concentration problems) (Cramer et al., 2016). When enough symptoms are activated, causal relationships between them can create dynamic networks, feedback loops and self-sustaining structures, which can remain active, even after the external stressor has passed (Borsboom, 2017). A recovery from an episode happens when connections between symptoms dissolve or symptoms deactivate. In remission, the network returns to a state of equilibrium, where the self-sustaining parameters of symptom activity rest below a threshold value (Borsboom, 2017).

In this framework, comorbidity and heterogeneity are natural consequences of the network structure, as symptoms may act as bridges to other symptoms across traditional diagnostic categories (Borsboom & Cramer, 2013). Furthermore, each person has their unique symptom networks, causal pathways and connections, vulnerabilities, and risk factors in developing episodes of mental distress

(Borsboom & Cramer, 2013). The ontological insights of the network approach have laid a basis for a growing number of methodological tools and techniques for visualizing and comparing patterns of network interactions and properties. The application of network analysis is rapidly growing, and it has been used to study the interactions of symptoms for many diagnostic categories and comorbid disorders, such as depression and anxiety (reviews of the application of network analysis: Conteras et al., 2019; Fried et al., 2017). The new insights network studies reveal about symptom interactions may help form new hypotheses, guide the development of novel interventions, and increase understanding of depressive illnesses (McNally, 2016).

1.5 NETWORK ANALYSIS OF DEPRESSION

The most commonly used network with psychological data is a weighted partial correlation network, which consists of nodes (variables) and edges (connections between nodes) (Eskamp, Borsboom, & Fried, 2018). A weighted network means that the connections in the network have altering magnitudes (Constantini et al., 2015). Partial correlations are the pairwise values derived from an inverse of a covariance matrix. The values reflect direct associations between two items after controlling information from all other items. This means, that partial correlations are especially suitable for revealing interactions between symptoms. With the use of typical correlations, a connection between two items could be due to shared connections to other items. However, in a partial correlation network, an edge between two nodes, depicts an association between the nodes after controlling information from all other nodes (Constantini et al., 2015). There are already several studies, which have utilized the network approach to reveal symptom associations, risk and protective factors, the course of illness and characteristics of depression, and the number of network studies is rapidly increasing (Fried et al., 2017).

1.5.1 The concept of centrality

Symptom networks can be assessed on both local and global levels (Constantini et al., 2015). On the local level, centrality measures can be calculated for individual nodes, which indicate the magnitude of connections a node has to other nodes. A symptom which has high centrality is likely to influence or be influenced by other symptoms more easily, compared to symptoms with weak connections. Many studies use different variables, so a comparison between centralities in different studies is difficult (Fried et al., 2017). Similarly, different network analysis techniques, or factors such as

whether the study was cross-sectional or temporal, hamper comparisons. However, there is some consistency regarding the centrality of depressive symptoms.

Several studies have shown similar results, that the two core items of DSM-5 "depressed mood" and "loss of interest" and the two non-core DSM-items of "concentration problems" and "fatigue/energy loss," are central symptoms in depression symptom networks among both depressed individuals (Fried et al., 2016a; Boschloo, van Borkulo, Borsboom, & Schoevers, 2016). Furthermore, centrality measures are associated with the course of illness in depression. For example, “depressed mood, "loss of interest", "concentration problems" and "fatigue" were found to predict the onset of a later diagnosis of major depression, in a study based on 501 participants, with no lifetime depressive or anxiety diagnosis in the baseline assessment of the Netherlands Study of Depression and Anxiety (Boschloo, van Borkulo, Borsboom, & Schoevers, 2016). The researchers found that “depressed mood, "loss of interest," "concentration problems" and "fatigue" were the strongest predictors for the onset of a later MDD, compared to other symptoms ($r=.87$, $p< .001$).

A study by van Borkulo et al. (2015) looked at how centrality was associated with clinical outcomes in clinically depressed patients. They found that "feelings of guilt," "fatigue," and "loss of energy" were significantly more central in the baseline assessment of those patients with a persistent depression after two years, compared to patients with depression in remission after two years (difference in node strength: "feelings of guilt": Cohen d=1,18; "fatigue/loss of energy": Cohen d=1,13). "Fatigue" has been associated with worse treatment outcomes also with adolescents. In a prospective study, that compared the network structure of adolescents undergoing treatment for depression symptoms (McElroy, Napoleone, Wolpert & Patalay, 2019), "fatigue" was the most central symptom at baseline for those who did not improve. Interestingly, "low mood" was the most central symptom at baseline for those who responded positively to treatment.

It is possible that when depression is accompanied by "feelings of guilt," "fatigue," and "loss of energy," depression may be more treatment-resistant. "Low mood" as the main problem, might instead yield better results in treatment. There is evidence that "depressed mood" (the equivalent of sadness), is associated with lower baseline severity (Ahmed, 2019) and that men are more prone to report symptoms of "sadness" (Schuch et al., 2014). In this light, "depressed mood" might be more central to men, whereas ‘feelings of guilt," "fatigue," and "loss of energy" might be more central to women.

1.5.2 The concept of density

Similarly, as the strength of a single node suggests that the node is more pathogenic, a network having strong density (overall connectivity) is proposedly more pathogenic compared to a loosely connected network. In a dense network, symptom activation may spread easier and create “negative spirals” (Cramer, 2017). Density is calculated as the sum of absolute values of all edges. There is contradicting evidence whether denser networks are more pathogenic, or whether denser networks at baseline predict worse treatment outcomes of depression. According to network theory, in the same study as referred above, van Borkulo et al. (2015) found that patients who had a persistent depression had a more densely connected network at baseline compared to those who remained in remission. However, a study aiming to replicate the finding with a clinical sample of adolescents found no difference in network connectivity at baseline (Schweren et al., 2015). The direction of the association was the same none-the-less (Schweren et al., 2015). More recently, McElroy, Napoleone, Wolpert, and Patalay (2019) looked at routine psychiatric data from 3017 adolescents undergoing at least three sessions in publicly funded services. They found that looser baseline density was associated with better recovery. However, the group who had a weaker network density (profiting from treatment), also showed the largest increase in density during the intervention. The effect of increasing network density is reported with decreasing symptom severity also in other antidepressant-interventions studies of clinically depressed adults (Bos et al, 2018; Madhoo & Levine, 2016). These findings suggest that where less densely connected networks can be beneficial initially, stronger interaction between symptoms may support recovery from depression, indicating "positive spirals" in symptom interactions (McElroy, Napoleone, Wolpert, & Patalay, 2019).

The role of density is an exciting topic of investigation, and there are still open questions regarding it (Fried et al., 2016b). For example, density has sometimes been reported higher among the general population compared to patient groups. An example of this was seen in a study with cancer patients (n=4020) who showed lower network density compared to controls (n=4020), although individuals with cancer experience more frequent and severe depressive symptoms (Hartung et al., 2019). Additionally, the lower network density in cancer patients was accompanied by lower symptom predictability compared to the general population. Symptom predictability depicts how much of the variance of a given symptom is explained by other symptoms in the network (Haslbeck and Fried, 2017). Lower density and lower predictability suggest that the occurrence of depressive symptoms in cancer patients is affected by many other variables outside the network which was estimated in the study. Factors that could trigger depressive symptoms, could relate to the illness, be due to pain,

associate to cancer treatments or with neurological damage (Hartung et al., 2019). Reflecting on the results above, network density should be interpreted carefully, as higher density can be paired with lower mean values and lower severity of symptoms. Higher network density can be a sign of recovery and a "better working" network, on the contrary, what one may assume according to the current formulation of network theory (Cramer, 2017). Thus, the association between gender and density is difficult to assess a priori, as the implications of high/low density are unclear. These results also imply that although it is possible to graph networks, the context in which it interpreted is essential. Analyzing symptom predictability may increase the interpretability of the results.

1.5.3 The concept of clustering

The global structure of a symptom network can also be analyzed by looking at patterns, in which symptoms cluster together (Constantini et al., 2015). In network analysis clusters are often described as communities. A few studies report community structures of depressive symptoms. One study based on a clinical sample found two distinct communities which reflect atypical and melancholic subtypes of depression (Bringmann et al., 2015). In another study, three communities were found initially, and just as network density was seen to increase, communities saw an increase from three to five during a clinical intervention (Bos et al., 2018). As only a few studies have reported community structures of depressive symptoms among clinical samples, looking at the community structure of depressive symptoms networks of men and women, in a general population sample, is purely explorative.

1.5.4 Comparing networks of men and women

To my knowledge, two previous studies have examined the gender differences of depression from a network perspective. Firstly, van Borkulo et al. (2017) selected adult men (N=351) and women (N=709) with a past year major depressive order, from the Netherlands Study of Depression and Anxiety. They compared male and female cross-sectional symptom networks using the Network Comparison Test (NCT; van Borkulo et al., 2017), which was introduced in the same article. The NCT is a permutation technique that compares two independent networks to a network of subsamples drawn from the two datasets. The permutation technique can be used to compare the overall structure, network density (global strength) or differences between specific edge values or node centrality indices. It estimates networks using Pearson's correlations by default. There was no significant difference in the network structure between the two genders, resulting that no further investigation of density or specific edge differences was made.

A second study (Mullarkey, Marchetti & Beevers, 2018) used the NCT to compare adolescent boys (N=646) and girls (N= 744) in a diverse community sample, who were assessed using the Children's Depression Inventory (CDI; Kovacs, 1985). They found a significant difference in the network structures (maximum difference = 1.16; p= .01), but not in network densities. A closer look at specific edge weights, using a Bonferroni-Holm correction, revealed that the only edge to significantly differ between the groups was the association between "body image" and "self-hatred," which was stronger in adolescent girls compared to boys (1.50 -0.34; p< .001). "Self-hatred" was one of the most central symptoms for both genders, along with "loneliness," "sadness" and "pessimism."

The current investigation contributes to the ongoing investigation of gender differences in depression. This study complements the previous studies, which were based on an adult clinical sample and a community sample of adolescents, by looking at depression symptom networks of adult men and women in a large general population sample. A network analysis of depressive symptoms in the general population will give insight into whether there is a difference in symptom interactions at lower "baseline" levels between genders in adulthood. Replication studies, with large representative samples of individuals with recurrent depression, have shown that the network structure of depression symptoms show strong similarity in network structures across different levels of genetic and environmental risk factors (vanLoo et al., 2018). In this light, individuals who have been clinically diagnosed with depression may depict a more homogeneous group, compared to the general population, as they have all already been susceptible for developing major depression. An investigation of depressive symptom networks among the general population can shed light on why women are more vulnerable to develop depression in the first place. Comparing the networks in both local (centrality of specific nodes) and global levels (density, structure and clustering), will enable a sensitive analysis of gender-dependent dynamics of symptom interaction, which may have implications for the prognosis and prevention and intervention of MDD.

2 RESEARCH QUESTIONS

1. Are there differences in self-reported depressive symptom severity between men and women?
2. Are there differences in the centrality of symptoms, community structure and/or overall connectedness in the depression symptom networks between men and women?

3. Do the network structures change correspondingly between genders, assessed by network properties in two time-points.

3 METHOD

3.1 PARTICIPANTS

The participants (female N=882, male N=576) are part of an ongoing Young Finns age cohort study that started in 1980. The first sample of 4320 children and adolescents aged 3, 6, 9, 12, 15, 18 was randomly chosen from the Finnish national registry. This study will look at data from 2007 and 2012 from participants who finished all questions relating to depressive symptoms in both time points. The participants' age ranged from 30 to 45 years old in the 2007 sample, and 35 to 50 years in the 2012 sample respectively. Only those participants who filled in the whole questionnaire at both time points were included in the study.

3.2 BDI-II

Beck's depression inventory II (BDI-II: Beck, Steer & Brown, 1996) is a multiple-choice self-report inventory for assessing depressive symptoms of adults and children over 12. In Beck's theory, a wholesome negative way for evaluating one's self, experiences, and future provides reason for depression symptoms to arise. BDI-II is based on Aaron T. Beck and his colleague's original version of the inventory in 1961 and has been modified to reflect the depression symptomology of DSM-IV (Beck, Steer & Brown, 1996). The inventory consists of 21 multiple choice question on a four-measure scale (0-3). The sum score of the items reflect severity, which ranges from no depression (0-13), to mild depression (14-19), to moderate depression (20-28) to severe depression (29-63). However, the BDI-II is not a diagnostic tool, and a high score does not necessarily reflect severe depression. The internal consistency measured by Cronbach's Alpha, was very good in all samples, ranging between .96-.97.

Bootstrapping resampling method (Efron, 1982) was used to compare mean sum scores between samples, as the method is especially valid when data are skewed, have unequal sample sizes or outliers (Howell, 2009). In comparison, outliers easily influence the Wilcoxon-test for non-parametric data

and, and the t-test assumes normally distributed data. Thus, means were compared by drawing 10,000 subsamples from the original data and computing the mean each time. The process creates a distribution of means for each sample. The means of the two distributions can then be compared with a t-test (Howell, 2009). Bootstrapping was also used to obtain confidence intervals (CIs) for symptom mean scores. Each symptom was resampled 10,000 times and the mean of the subsample was recalculated each time. Means were reorganized in numerical order and 2,5% was cut from both ends of the resampled mean distribution to obtain 95% CIs. This procedure was repeated for each symptom distribution across all samples.

3.3 NETWORK ANALYSIS

The symptom networks in this study are partial correlation networks, which are based on a Gaussian graphical model (GGM; Laurizen, 1996) and most commonly used with psychological data (Epskamp, Borsboom, & Fried, 2018). The GGM can be used with ordinal data, as is the case with this data. The partial correlations, in this study, are based on threshold functions and polychoric correlations, which is common with non-normal ordinal data (Epskamp, Borsboon, & Fried, 2017; Muthén, 1984). As explained in more detail by Muthén (1984), a latent score threshold is calculated for each answering category (0, 1, 2, 3) along the Gaussian curve. Correlations between latent variables can then be estimated pairwise, without further transformations, using the threshold estimates.

3.3.1 Network estimation

Graphical LASSO regularization (glasso: (Friedman, Hastie, & Tibshirani, 2008) was used, which is specifically aimed at partial correlations (Epskamp, Borsboom, & Fried, 2018). Regularization cleans off very weak connections in the network, by limiting the sum of absolute partial correlation coefficients. As a result, all estimates shrink and some edges are reduced to zero (meaning the absence of an edge). By limiting the number of edges, the network becomes more interpretable. Regularization creates a pool of possible networks and the best-regularized network was selected by using EBIC-model selection. The EBIC selection minimizes the Extended Bayesian Information Criterion (BIC) (Chen & Chen, 2008), and has been found both sensitive and especially good in retrieving the true network with a moderately large n and glasso regularization (Foygel & Drton, 2010). The EBIC hyperparameter λ was set to 0.5 as suggested by Foygel and Drton (2010). Symptom networks were

estimated using the R-package bootnet function ‘estimateNetwork' and plotted using the R-package qgraph function ‘plot'.

Network comparisons were conducted by visual assessment. The NCT, which has been used in previous studies cannot be applied to this data, as it is based on Pearson’s correlations or binary data. Enabling comparisons, all networks were standardized by setting the maximum value to correspond with the largest maximum value of 0.62. Furthermore, to ease comparison, the layouts of the networks were averaged with the qgraph function ‘averageLayout'. Analyses were executed using the statistical computing software R 3.4.4 (R-Core team, 2014). A full R-script of the analyses is found in the Appendix, as is encouraged in the network literature (Epskamp, Borsboom, & Fried, 2018).

3.3.2 Centrality, predictability and community structure

Strength (Opsahl, Agneessens, & Skvoretz, 2010) and Expected Influence (Robinaugh, Millner & McNally, 2016) centralities were estimated using the bootnet-package in R (Epskamp, Borsboom, & Fried 2017). Strength is calculated as the sum of absolute edge values in a node and it signifies the magnitude of connections a node has to other nodes. When node strength is high, the node can influence many other nodes directly, and be influenced by them (Constantini et al., 2015). ExpectedInfluence is a relatively new centrality metrics which, like strength centrality, measures the overall connectivity of a node (Robinaugh, Millner & McNally, 2016). The difference to strength is that ExpectedInfluece accounts for whether the connection is negative or positive. Hence, ExpectedInfluence is especially informative in networks with negative edges.

Following the guidelines of Epskamp, Borsboom and Fried (2018) centrality indices were estimated for stability, by using bootstrapping with the bootnet-package. 1,000 subsets were drawn from the original data and each time a different number of cases was dropped and the order of the indices recorded. Correlations for the order of the centrality indices were calculated between each subset and the original sample. The correlation stability coefficient (CS-coefficient; Epskamp, Borsboom & Fried, 2018) tells the maximum percentage of cases that can be dropped from the original sample to retain a correlation of 0.7 in at least 95% of the samples. I.e. a very low CS-coefficient means that the order of the indices is not stable or accurate even when a small drop in the sample is made. The CS-coefficient should be at least over 0.25 and preferably over 0.5 (Epskamp, Borsboom & Fried, 2018). CS-coefficients were calculated for strength and ExpectedInfluence.

Predictability (Haslbeck & Waldorp, 2018) of nodes was estimated using the mgm-package in R. Predictability of an item (Haslbeck & Fried, 2017) tells how much of the variance of a single item is being explained by other items in the network. Predictability estimation ranges from 0 (node variance is not explained by other nodes) to 1 (node variance is explained perfectly by other nodes). However, predictability cannot be estimated using polychoric correlations. Instead the mgm- function uses a Mixed Graphical Model (MGM) to estimate networks. MGM treats variables as continuous, and network estimation is based on the neighborhood regression approach. In order to interpret predictability estimates meaningfully, the MGM network should have the same structure to the GGM network.

A community analysis was made using the R-package EGAnet (Golino & Christensen, 2019) using the function 'EGA.' The Exploratory Graph Analysis technique (EGA: Golino & Epskamp, 2017) estimates the number of clusters of symptoms that interconnect more densely to each other, compared to the rest of the network. It also shows which items belong to each cluster. EGA is specifically designed for estimating the number of dimensions in a network structure, based on partial correlations with graphical Lasso regularization and EBIC selection. It uses the Walktrap algorithm for weighted undirected networks, which in short, uses random walks to find densely connected symptom communities (Pons & Latapy, 2005). It reveals groupings by getting trapped in them, rather than being able to move freely in the network from one community to another.

4 RESULTS

4.1 DESCRIPTIVE STATISTICS OF DEPRESSIVE SYMPTOMS

Figure 1 shows the distributions and means of the four samples. Women showed significantly higher mean sum scores compared to men in both years (2007: $p > .000$; 2012: $p > .033$- $.003$). Women also showed a significant difference in the mean sum scores between 2007 and 2012 ($p > .0142$). There was no significant difference in the mean sum scores in men between the two measuring points ($p > .483$). Figure 2 shows mean scores of each symptom and bootstrapped CIs in each sample. 12/21 symptoms had significant differences between women and men in at least one possible comparison based on the bootstrapped CIs. In 6 symptoms, women had significantly higher means only in the 2007 sample compared to men ("sadness," "self-criticalness," "agitation," "indecisiveness," "changes

in sleep," "changes in appetite"). In 4 symptoms, women had significantly higher means compared to either one or both men's samples, but the differences were not significant in all comparisons ("guilty feelings," "crying," "loss of energy," "irritability"). In 2/21 symptoms women had significantly higher means in all comparisons ("fatigue" and "loss of interest in sex").

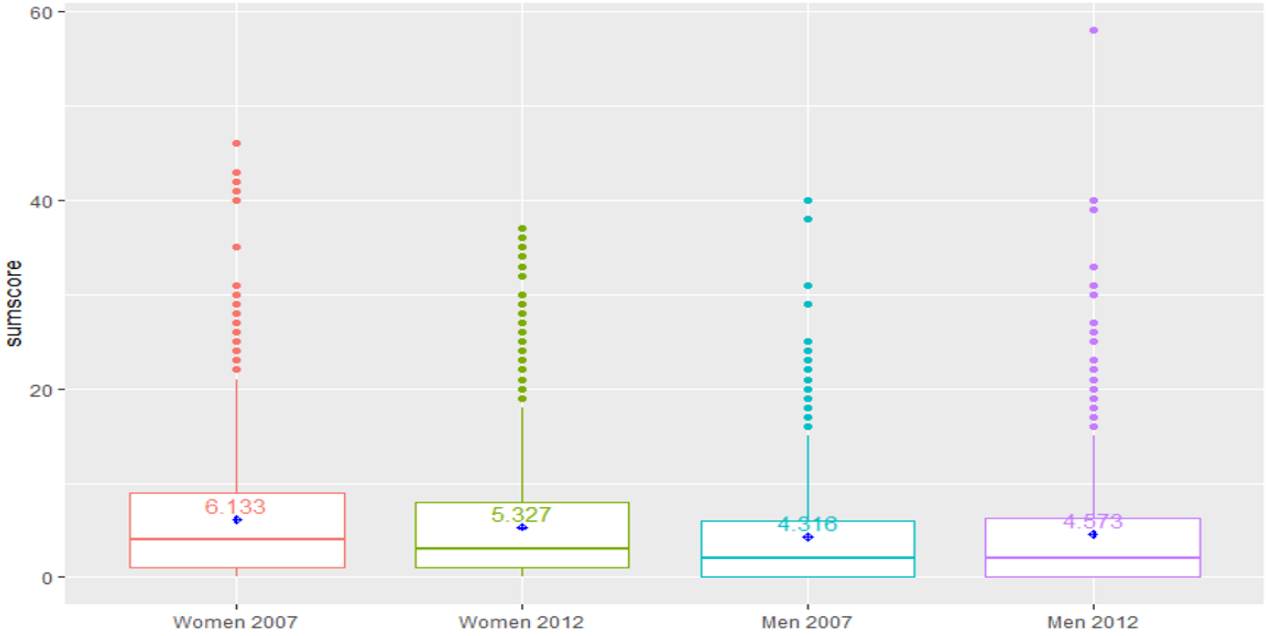

Figure 1: Box-plot graph of depressive symptoms of men and women in two time-points. The horizontal lines depict median values. Means are shown as blue dots and presented in numbers. Dots outside the boxes represent outliers.

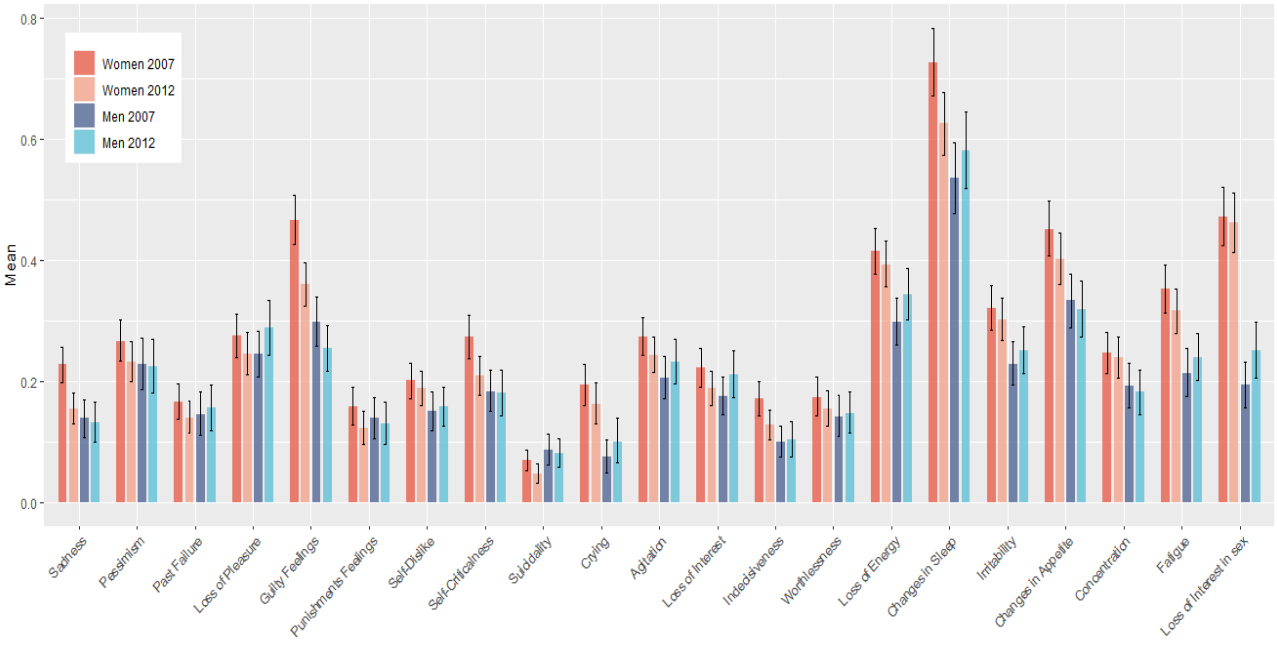

Figure 2: Mean scores of BDI-II symptoms for men and women in two time-points. The black bars indicate bootstrapped 95% confidence intervals

4.2 NETWORK ANALYSIS

Figure 3 shows the estimated partial correlation networks of BDI-II symptoms for men and women in two time points. Centrality analysis results are shown in Figure 4 for both genders in two time points. CS-coefficients, which tell the stability of centrality were at good levels for ExpectedInfluence (.205-.439), but Strength centrality was under the recommended cut-off point of .25 (.128-.205), so only ExpectedInfluence centrality was assessed. Predictability of symptoms in the networks could not be meaningfully estimated, as there were visible differences in the network structure based on neighborhood regression (MGM) compared to the networks estimated using polychoric correlations (GGM). Estimated networks based on the EGA-community analysis are shown in figure 5. Table 1 shows descriptive statistics for each network, with means and standard deviations.

.

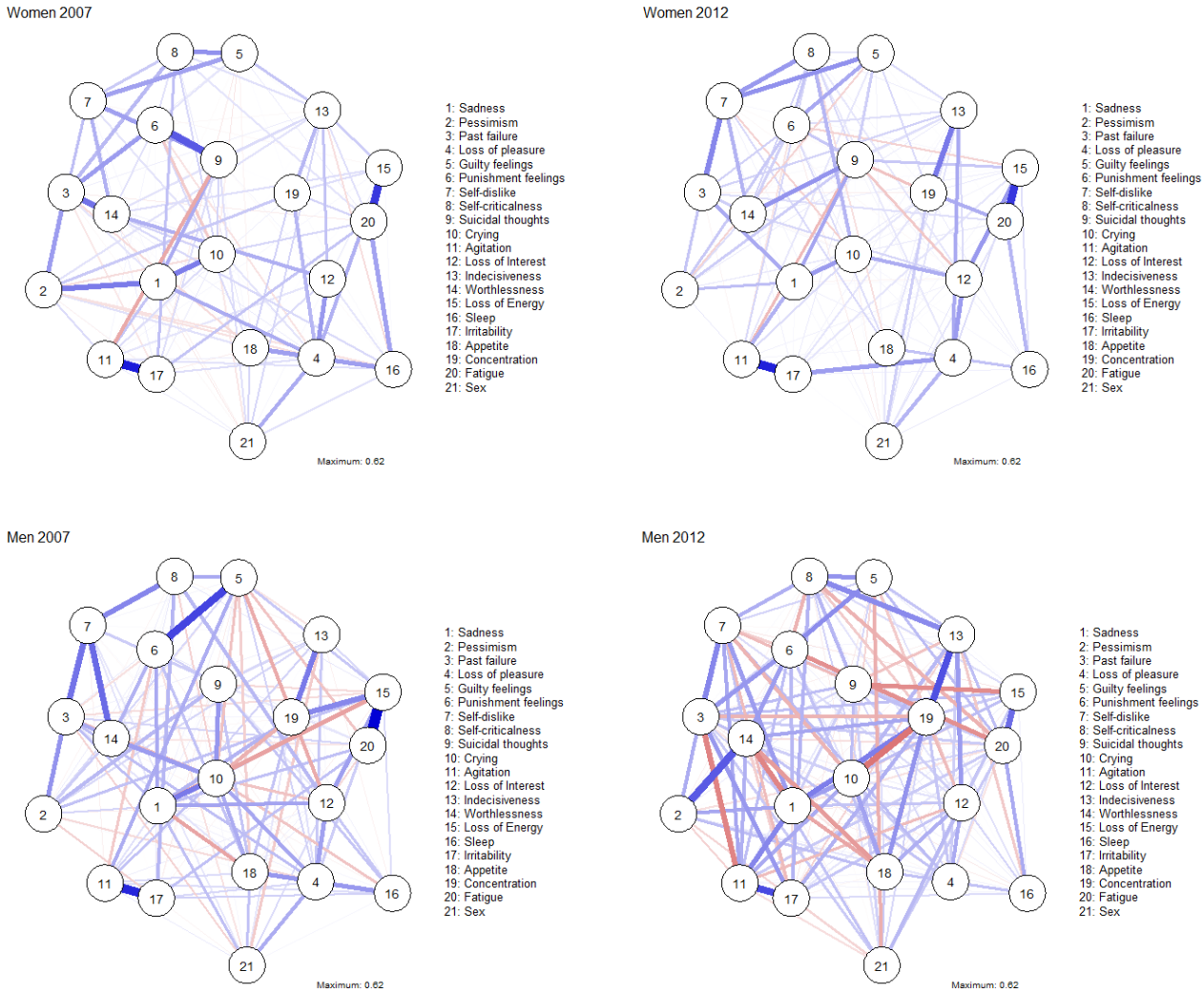

Figure 3. Networks of BDI-II symptoms for men and women in two time points. The Fruchterman-Reingold algorithm makes stronger edges wider, and usually, more connected nodes are placed closer together. Positive edges are painted blue and negative edges red.

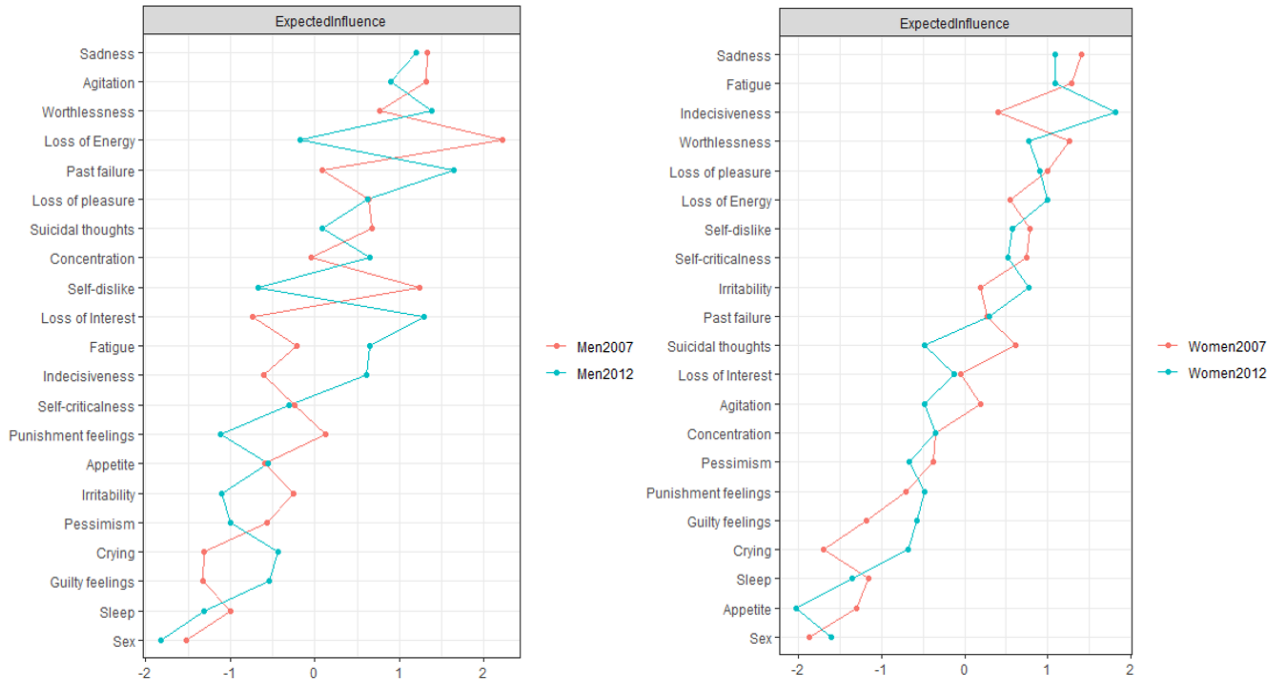

Figure 4: ExpectedInfluence centrality for men and women in two time points. Centrality estimates are shown in z-scores.

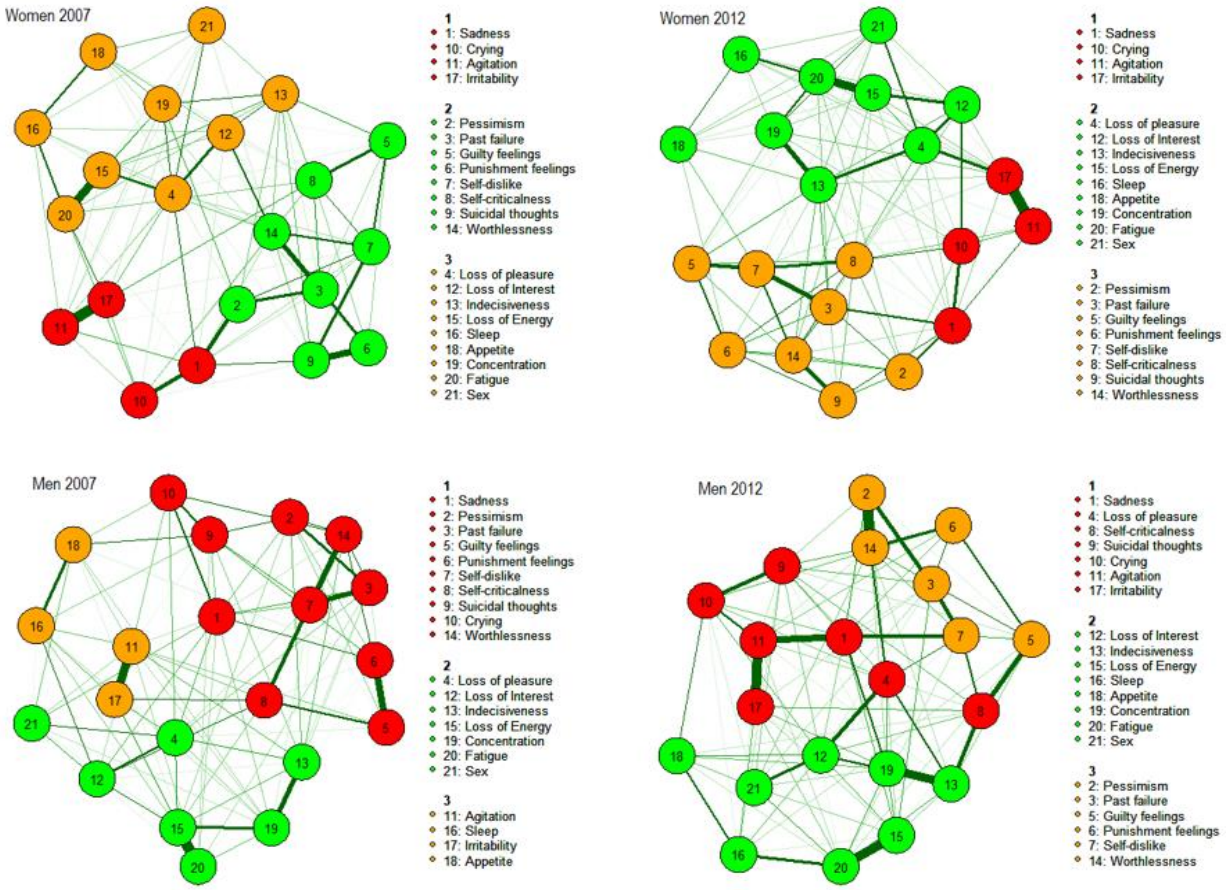

Figure 5: EGA community structure of men and women. Colors reflect EGA clustering results.

Table 1
Descriptive statistics of depressive symptoms of men and women in two time-points.

	Mean sum score	SD	Network density	Node stability (CS-coefficient)	Negative edges
Women 2007	6.113	7.141	12.49	0.517	19,5 %
Women 2012	5.327	6.683	12.11	0.283	13,1 %
Men 2007	4.316	5.794	17.08	0.439	29,6 %
Men 2012	4.573	6.592	22.78	0.283	33,7 %

5 DISCUSSION

5.1 SYMPTOM ARCHITECTURE OF MEN AND WOMEN

5.1.1 Symptom severity

In line with other studies on depression (Salk, Hyde, & Abramson, 2017), women reported more depressive symptoms and had higher mean sum scores compared to men (Figure 1). Women also had significantly higher means on several items compared to men (Figure 2). Most symptoms, which were reported higher among women were somatic and affective symptoms, including “fatigue,” “changes in appetite,” “irritability,” “loss of energy,” “crying,” “guilty feelings,” and “sadness.” Cognitive symptoms were reported in more similar levels between genders. The results support previous findings that women are more prone to report somatic symptoms and symptoms of atypical depression (Schuch et al., 2014; Marcus et al., 2005; Kroenke & Spitzer, 1998). Although many symptoms showed differences in severity between genders, when looking at the largest possible differences, “fatigue” and “loss of interest in sex” were the only symptoms that were reported consistently higher among women in all comparisons. As opposed to an earlier study with clinically depressed individuals (Schuch et al., 2014) men did not experience more sadness in this general population sample compared to women.

5.1.2 Symptom centrality

The mean value of a symptom was not predictive of symptom centrality. For example, most somatic symptoms in the women’s samples, such as “loss of interest in sex” were among the least central symptoms in the women’s networks. This implies that, for women, somatic symptoms did not share

much connectivity with the rest of the depressive symptoms in this study. “Fatigue” was the only central somatic symptom among women. It could be that, overall, somatic symptoms share more associations with other types of symptoms or factors. For example, it has been suggested that somatic symptoms could be explained by anxiety-related symptoms (Kapfhammer, 2006). Different underlying biological factors could also increase the likelihood of somatic symptoms in women. It has been suggested that the whole gender difference in depression is due to the higher rates of reported somatic symptoms in women (e.g. Silverstein, 2002), but these claims have also been criticized as unlikely to fully explain the different ratios of diagnoses and symptom severity (Delisle et al., 2012).

For women, the centrality estimates were similar across time. Thus, associations between symptoms remained relatively stable as time had passed, even at low levels of symptom severity. For men, centrality estimates varied much more across time (*Figure* 4). For example, symptoms such as “lo of energy,” “past failure,” “self-dislike,” and “loss of interest” ranged from high centrality to low centrality in the men's two networks. Thus for men, at low levels of symptom severity, the relationships between symptoms varied much more compared to women. However, “sadness” and “worthlessness” were consistently high in centrality for both genders in both networks (Figure 4). Furthermore, “fatigue” and “agitation” differed in centrality between genders (Figure 4).

Firstly, “fatigue” was highly central in both women’s networks but had low centrality in the men's networks. The finding that “fatigue” was highly central among women is in line with previous literature on fatigue. In general practice, women are observed to report more fatigue than men (Bensing, Hulsman, & Schreurs, 1999; Cullen, Kearney, & Bury, 2002). Reflecting on the high centrality of fatigue in the symptom network, women might be more prone to fatigue when other symptoms arise, and by experiencing more fatigue, women are more susceptible to develop depressive episodes. There is also evidence from previous network analysis studies that “fatigue” is associated with prospective depressive episodes (Boschloo, van Borkulo, Borsboom, & Schoevers, 2016). A study which investigated the gender difference in “fatigue” based on the Dutch National Survey of Morbidity and Interventions in General Practice (men: n = 4,681; women: n = 4,698) analyzed different biological, psychological, and social factors relating to fatigue (Bensing, Hulsman, & Schreurs, 1999). Several factors were identified, which may underly the higher rates of fatigue among women, including child-care and employment. For men, fatigue was associated with more specific factors that limit functionality, such as having handicaps and severe chronic complaints. The authors emphasized that fatigue should be understood in a multicausal biopsychosocial framework.

Based on these results, enhancing our understanding of the etiology of fatigue (especially in women) could be a valuable direction in supporting the well-being of women.

"Agitation" was the other symptom which showed a clear gender difference in centrality, being highly central for men in both time points, but having low centrality in the women's networks. Interestingly, "agitation" is widely discussed in literature concerning suicide, as individuals who have died by suicide have often been characterized as physically or psychologically agitated (Busch, Fawcett, & Jacobs, 2003). It has long been acknowledged that men have higher suicide rates in most parts of the world; in Europe, the difference between genders is around 5:1 (WHO, 2017). "Agitation" has also been associated with a significantly higher risk of suicide attempt in men but not in women (Bryan et al., 2014). "Agitation" has also been shown to be especially central in a network analysis study on rapid-onset, acute suicidality (Rogers, Jom, & Hoiner, 2019). Thus, the high centrality of agitation among men could relate to processes which associate with suicidality of men. It should be noted that agitation may not in itself increase the risk of suicide. The effect of agitation may become relevant in the context where a person has developed the capacity to commit suicide (Ribeiro, Silva, & Joiner, 2014).

5.1.3 Network density

Network density was higher in both of the men's networks, compared to women's networks (Figure 3). The possibility to inspect networks in two time points revealed that men and women also had distinct developmental features in their symptom network structures. The most visible difference between genders was that, while women had a very similar network structure across time, density in the men's network increased by 20% in the second measuring point compared to the first (Table 1). In contrast to network theory (Borsboom, 2017), symptom severity (sum scores) and network density showed no clear relationship. Although women showed higher mean values compared to men, the symptom networks of women were less dense and had weaker connections as the men's networks (Table 1). Many of the symptoms such as "changes in appetite," "crying," and "sadness" were also among the symptoms that were visibly more connected in the men's networks, although they were reported more by women. Furthermore, although women showed significant changes in the mean scores, there was no change in network density (Table 1). Thus, in this general population sample, network density did not reflect a more pathological network.

Previous results regarding network density have been mixed. There is some evidence that increased network density is associated with recovery. On the other hand, denser baseline networks are associated with later depressive episodes and worse prognosis. Yet, in other cases, density has been higher in the general population compared to clinical samples when other factors outside the network could explain symptom activation more. Sadly, it was not possible to estimate predictability in this study to confirm how much of the variation of a single symptom was explained by the other symptoms in the network. One explanation, which could account for the higher network density in men's network in this study, could relate to the high number of negative connections in the men's networks (Table 1). Negative connections imply that these symptoms are not reported together, as the increase in one symptom would mean a decrease in the other. It has been pointed out previously, that items with many or strong negative connections might not be highly problematic, but instead, diminish the effects of other items in the network it interacts with (Robinaugh, Millner & McNally, 2016). The negative edges in the men's networks might perhaps partly explain why a denser network in men does not result in a more pathological network. It could mean that the negative edges actually reflect a protective factor in men against developing depressive episodes.

5.1.4 Community structure

The main observation in the community analysis was that women had a more stable community structure across time compared to men. For women, depressive symptoms formed three distinct communities, which remained the same in both measuring points (Figure 5). Women reported these symptoms together more with each other than with other symptoms, although the formation of the community network evolved (greed and orange communities swapped over), and changes occurred in the connections both within and between clusters. Firstly, women reported co-occurring somatic-affective symptoms characterized by low mood, irritation, and agitation (red community, Figure 5). Secondly, cognitive symptoms, such as worthlessness and self-dislike (first green then orange) formed a distinct community. Lastly, women reported somatic-affective symptoms together, which shared characteristics of anhedonia and problems with fatigue, sleep, indecisiveness, and appetite. Thus, the last community was associated most with atypical depression (first orange then green). For men, the community structures were more unstable. Both the number and type of symptoms changed within communities. As with women, cognitive symptoms associated consistently more with each other compared to other symptoms. Somatic and affective symptoms did not group consistently. For example, although sleep and appetite clustered together, atypical symptoms overall did not form a distinct community in both of the men's networks.

5.2 IMPLICATIONS AND FUTURE DIRECTIONS

The network analysis revealed distinct gender differences in the way depressive symptoms associated with each other in the general population. Women had a relatively stable network structure across time, unlike men. Thus, especially the network structure properties of women could form a basis for preventive action and further research in women's depression. In the framework of network theory, the structure of the symptom network reflects how quickly an episode of psychopathology might emerge, and both local and global properties of the network can be inspected. For example, central symptoms of the network could be targeted in preventive actions. Targeting central symptoms is based on the idea that central symptoms can activate other symptoms more easily. By eliminating or hindering the activation of a central symptom from the symptom network, other connections in the network will not be activated or diminish as well.

As "fatigue" was highly central for women, but not in men, it could prove an especially valuable target for preventive action in women's depression. Thus, more research could focus specifically on factors relating to fatigue in women. It may be especially suitable to examine fatigue in a multidimensional model, where many types of factors could be examined together. More holistic networks have also been called for by the founders of the network approach (Fried & Cramer, 2017; Borsboom, Cramer, and Kalis, 2019). In this context, this could mean looking at fatigue along with social, physical or cognitive vulnerabilities and symptoms associated with fatigue across many diagnostic categories. The effects of central symptoms in depression could also be studied more in simulation studies. Simulations could give new insight into how a change in a central symptom would affect the rest of the network. The global structure, which included three distinct symptom communities in the women's networks could provide an interesting topic for future research. For example, the community structure could be used in research relating to biological determinants of depression in projects like the RDoC (Woody & Gibb, 2015; Ahmed et al, 2018).

More broadly, conceptualizing mental disorders as systems of symptoms has implications for clinical work as well. Patients could be given psychoeducation how different symptoms relate to one another, or central symptoms could be targeted in interventions for already depressed individuals. Temporal networks can already be created for individuals by collecting data with digital devices several times a day for a period of a week or two. These longitudinal networks provide specific information on the causal relationships of symptoms and the symptom network structures of individuals. Collecting more longitudinal group data is also important for developing a more accurate picture of how symptom

networks and specific symptoms trigger, maintain and deactivate episodes of psychopathology. Attaining such group data could consist of developing freely available online/mobile apps, where users could attain personal symptom networks and data could be gathered for research purposes.

5.3 LIMITATIONS

There are several limitations to this study. Firstly, there is the question of the appropriateness of the statistical method. The partial correlation network assumes a multivariate normal data. However, the current data is non-normal and strongly skewed to the right, as most responses cumulated on the lower end of the scale. The normality assumptions were relaxed in this study by assuming that the observed data was a transformation of a latent multivariate normally distributed system (Liu et al., 2009). However, the model still assumes the latent variable to be normally distributed. It is quite plausible that at least some latent variables scores underlying observed variable scores, such as suicidal ideation are not normally distributed. Although partial correlation networks are widely used in the network literature, it is still an open question on how to best deal with non-normal data (Epskamp, 2017). One way to solve the issue is to dichotomize data and use network estimation, which is based on Ising models (vanBorkulo et al., 2014). Instead of correlations or partial correlations, network estimation for binary data is based on logistic regression and goodness of fit measures and is free from the normality assumptions linked to Gausal models. The two previous network studies of gender differences on depressive symptoms (one clinical and one non-clinical sample) are both based on Ising models. The two studies have found limited, if any, evidence of gender differences in the depressive symptom networks (vanBorkulo, 2017; Mullarkey, Marchetti & Beevers, 2018).

For comparison, and to check whether the results would remain the same in this study, all networks were also estimated based on an Ising model and compared using the NCT for binary data (vanBorkulo, 2015). Based on NCT with binary data, there were no significant differences in the networks between men and women (results included in the Annex). There were also no negative edges in any network. Naturally, sensitivity is lost with binary data so the question is which method provides a more accurate picture of the true network. It could be that binary data is not sensitive enough to detect the gender differences found in the other analyses. Or, it could be that partial correlations are not suitable for such skewed data and/or when most of the data consists of zeros. In any case, there is a significant difference in the results based on the network estimation method.

Assuming that the partial correlation network gives us a more sensitive and a true reflection of symptom interaction, we still have the uncertainty of what centrality and density estimates actually tell us. According to network theory, associations between symptoms reflect causal relationships, instead of shared latent variables, and higher network density is associated with more pathological networks (Cramer, 2017). Results of this study suggest the opposite. However, direct causality between symptom interactions may not even be possible to assume. There is a multitude of options between the two extremes of systems of direct causality and common cause models (see Bringmann & Eronen, 2018). It has been shown that symptoms could have covariation due to a common factor, and a common cause does not prevent causal interaction between symptoms (Humphry & McGrain, 2010). Also, symptoms do not have clear boundaries, but most likely overlap. For example, symptoms such as "fatigue" and "sleep problems," or "irritation" and "agitation," which had consistent strong connections in this study, might merely measure a common shared factor and inferring causal relationships between them is ambiguous (Bringmann and Eronen, 2018). To overcome this problem, Bringmann and Eronen (2018) have argued that there is no need to contrast latent variable models and network approach models. Instead, the focus should be on the symptom-orientated approach and in studying psychopathology as a dynamic system, which evolves in time. There are already network studies, which focus on the temporal dynamics of symptoms, where causality can be investigated more accurately (e.g. Bringmann et al., 2015).

More fundamental problems of the network analysis have been raised regarding cross-sectional and group level networks. For example, group-level networks might not reflect symptom associations at the level of the individuals (Bos & Wanders, 2016). Furthermore, it has been observed that cross-sectional data does not reflect how symptoms trigger each other over time, as centrality measures are not concurrent when cross-sectional and temporal data are compared in the same study (Bos et al., 2017). Another limitation is how to estimate the meaningfulness of the network. Whether connections to other items explain 1% or 50% of an item's variance is highly relevant for making estimations on the meaningfulness of the network in practice (Haslbeck & Waldorp, 2018). As predictability estimates could not be applied to this data it is unclear how meaningful the estimated networks were.

There are also limitations regarding depression research more generally, which relates to the problem of heterogeneity when measuring depression. A meta-analysis of data-driven studies on the factorial structure of common depression check-lists, shows large variation in the number and content of different factorial solutions the different studies yield (vanLoo et al., 2012). Similarly,

common rating scales seem to be lacking unidimensionality and longitudinal measurement invariance, meaning that the factorial structure of depressive symptoms is not the same even among the same people at different points in time (Fried et al, 2016b). This means that it is not clear what is being measured by common depression checklists. Current depression inventories include a mixture of mood, behavioral, somatic and cognitive symptoms in varying numbers and the selection of symptoms in a checklist might be based more on history and tradition rather than science (Santor, Gregus & Welch, 2006). Network analyses on depression check-list structures show that the centrality of items varies: where some items are highly central, some are much less so (Fried et al., 2016; Kendler et al., 2018).

The heterogeneity of measuring depression is especially tricky with network analysis. What items should be included in the analysis? Using the 21 preselected BDI-II symptoms inevitably result in a certain kind of network, which includes symptoms which may, or may not be important in the networks of depressive symptoms for men and/or women. A different selection of symptoms would have resulted in a different network. For example, ‘loneliness' is not included in the most used depression inventories, such as Hamilton Depression Rating Scale (HAM-D), Beck Depression Inventory (BDI), PHQ-9 or the Major Depression Inventory (MDI). However, loneliness is consistently associated with worse mental and physical health (Santini et al., 2016; Cacioppo et a., 2006; Holt-Lunstad & Smit, 2010), and has been reported as a highly central symptom in the networks of adolescents and adults regarding depression (Mullarkey, Marchetti & Beevers, 2018; Santos et al, 2017). Similarly, anxiety-related symptoms are very common among depressed individuals, are more pronounced among women, and associated with worse outcomes (Howland et al., 2009). It has also been suggested that symptoms, such as “crying” and “increased appetite”, which are often reported more by women, might bias assessment on gender differences, as these symptoms could relate more to women in the first place (Romans, 2007). Thus, the results are only valid in the specific network context. There are also limitations associated with all self-reported measures. For example, it could be that differences between genders are due to different response styles between genders. However, research on the gender difference in depression is vast, and the results on the higher depression rates among women are consistent across nations.

Future network analysis studies should focus on enhancing methodological and theoretical grounds, centrality estimates, and network analysis techniques, which could better account for psychological and mental health data. Developing network analysis on more solid theoretical grounds could

include building research on longitudinal temporal group data. Clearly, there are limitations to current inventories and methods, which should be given great attention.

5.4 CONCLUSION

The partial correlation depressive symptom networks of men and women showed distinct differences in connectivity and development across time. Although women showed alterations in the severity of symptoms, unlike men, the network structure remained relatively stable. For men, lower mean scores of depressive symptoms were linked to higher density, more negative connections, more variation in symptom centrality across time, and to a more evolving community structure. In the network theory framework, it could be that the pathways from symptoms to a depressive episode are more direct with women. Women's networks also consisted of three distinct symptom clusters, which may reflect gender-specific emotional, cognitive, social or biological determinants, which make women more susceptible to depression. In contrast, negative connections in the men's networks may hinder the co-occurrence of different symptoms and reflect a protective factor men have in developing depression. "Fatigue" could play an especially important role in the development of depression in women, as it was highly central in the networks of women, but not in the men's networks. Although there are still open questions relating to the network theory, the novel statistical methods have enabled fresh hypotheses and frameworks for understanding psychopathology. The network approach has created new avenues for research and enhanced knowledge about individual symptoms and their differences.

REFERENCES

Ahmed, A. T., Frye, M. A., Rush, A. J., Biernacka, J. M., Craighead, W. E., McDonald, W. M., ... & Hall-Flavin, D. K. (2018). Mapping depression rating scale phenotypes onto research domain criteria (RDoC) to inform biological research in mood disorders. *Journal of affective disorders*, *238*, 1-7.

American Psychiatric Association. (2015). Depressive Disorders: DSM-5® Selections. American Psychiatric Pub.

Angst, J., Gamma, A., Sellaro, R., Zhang, H., & Merikangas, K. (2002). Toward validation of atypical depression in the community: results of the Zurich cohort study. *Journal of affective disorders*, *72*(2), 125-138.

Antypa, N., Giegling, I., Calati, R., Schneider, B., Hartmann, A. M., Friedl, M., ... & Rujescu, D. (2013). MAOA and MAOB polymorphisms and anger-related traits in suicidal participants and controls. *European archives of psychiatry and clinical neuroscience*, *263*(5), 393-403.

Beck, A. T. (2002). Cognitive models of depression. *Clinical advances in cognitive psychotherapy: Theory and application*, *14*(1), 29-61.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. *San Antonio*, *78*(2), 490-498.

Bensing, J. M., Hulsman, R. L., & Schreurs, K. M. (1999). Gender differences in fatigue: biopsychosocial factors relating to fatigue in men and women. *Medical care*, *37*(10), 1078-1083.

van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific reports*, *4*, 5918.

van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B. W., Waldorp, L. J., & Schoevers, R. A. (2015). Association of symptom network structure with the course of depression. *JAMA psychiatry*, *72*(12), 1219-1226.

van Borkulo, C. D., Boschloo, L., Kossakowski, J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2017). Comparing network structures on three aspects: A permutation test. *Manuscript submitted for publication*.

Borsboom, D. (2017). A network theory of mental disorders. World Psychiatry, *16*(1), 5-13.

Borsboom, D., Cramer, A. O., & Kalis, A. (2019). Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences*, *42*(e2), 1–63.

Borsboom, D., & Cramer, A. O. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual review of clinical psychology*, 9, 91-121.

Bos, F. M., Fried, E. I., Hollon, S. D., Bringmann, L. F., Dimidjian, S., DeRubeis, R. J., & Bockting, C. L. (2018). Cross-sectional networks of depressive symptoms before and after antidepressant medication treatment. *Social psychiatry and psychiatric epidemiology*, *53*(6), 617-627.

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

Bos, E. H., & Wanders, R. B. (2016). Group-level symptom networks in depression. *JAMA psychiatry*, *73*(4), 411-411.

Boschloo, L., van Borkulo, C. D., Borsboom, D., & Schoevers, R. A. (2016). A prospective study on how symptoms in a network predict the onset of depression. *Psychotherapy and psychosomatics*, *85*(3), 183-184.

Bringmann, L. F., & Eronen, M. I. (2018). Don't blame the model: Reconsidering the network appr oach to psychopathology. *Psychological Review*, *125*(4), 606.

Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological medicine*, *45*(4), 747-757.

Bryan, C. J., Hitschfeld, M. J., Palmer, B. A., Schak, K. M., Roberge, E. M., & Lineberry, T. W. (2014). Gender differences in the association of agitation and suicide attempts among psychiatric inpatients. *General hospital psychiatry*, *36*(6), 726-731.

Busch, K. A., Fawcett, J., & Jacobs, D. G. (2003). Clinical correlates of inpatient suicide. *The Journal of clinical psychiatry*, *64*(1), 14-9.

Cacioppo, J. T., Hughes, M. E., Waite, L. J., Hawkley, L. C., & Thisted, R. A. (2006). Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. *Psychology and aging*, *21*(1), 140.

de Azevedo Cardoso, T., Mondin, T. C., Wiener, C. D., Marques, M. B., de Avila Fucolo, B., Pinheiro, R. T., ... & Oses, J. P. (2014). Neurotrophic factors, clinical features and gender differences in depression. *Neurochemical research*, *39*(8), 1571-1578.

Centre for Mental Health (2010) The Economic and Social Costs of Mental Health Problems in 2009/10.http://www.centreformentalhealth.org.uk/pdfs/Economic_and_social_costs_2010.pdf Intro

Contreras, A., Nieto, I., Valiente, C., Espinosa, R., & Vazquez, C. (2019). The study of psychopathology from the network analysis perspective: a systematic review. *Psychotherapy and psychosomatics*, *88*(2), 71-83.

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13-29.

Cramer, A. O., van Borkulo, C. D., Giltay, E. J., van der Maas, H. L., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PloS one*, *11*(12), e0167490.

Cullen, W., Kearney, Y., & Bury, G. (2002). Prevalence of fatigue in general practice. *Irish journal of medical science*, *171*(1), 10.

Delisle, V. C., Beck, A. T., Dobson, K. S., Dozois, D. J., & Thombs, B. D. (2012). Revisiting gender differences in somatic symptoms of depression: much ado about nothing?. *PLoS one*, *7*(2), e32490.

DHSC publication, (2011). No Health Without Mental Health: a cross-Government mental health outcomes strategy for people of all ages - a call to action. Supporting document – The economic case for improving efficiency and quality in mental health. Department of Health and Social Care, United Kindom.https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/215808/dh_123993.pdf

Efron, B. (1982). *The jackknife, the bootstrap, and other resampling plans* (Vol. 38). Siam.

Elovainio, M., Aalto, A. M., Kivimäki, M., Pirkola, S., Sundvall, J., Lönnqvist, J., & Reunanen, A. (2009). Depression and C-reactive protein: population-based Health 2000 Study. *Psychosomatic Medicine*, *71*(4), 423-430.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195-212.

Fava, M., & Rush, A. J. (2006). Current status of augmentation and combination treatments for major depressive disorder: a literature review and a proposal for a novel approach to improve practice. *Psychotherapy and psychosomatics*, *75*(3), 139-153.

Foygel, R., & Drton, M. (2010). Extended Bayesian information criteria for Gaussian graphical models. In *Advances in neural information processing systems*, *1*, 604-612.

Frank, R. I., & Davidson, J. (2014). *The transdiagnostic road map to case formulation and treatment planning: Practical guidance for clinical decision making*. New Harbinger Publications.

Fried, E. I. (2015). Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. *Frontiers in psychology*, *6*, 309.

Fried, E. I., & Cramer, A. (2017). Moving forward: challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science*, *12*(6), 999–1020.

Fried, E. I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016a). 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.

Fried, E. I., & Nesse, R. M. (2015a). Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR* D study. *Journal of affective disorders*, *172*, 96-102.

Fried, E. I., & Nesse, R. M. (2015b). Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC medicine*, *13*(1), 72.

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), 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*(11), 1354.

Fried, E. I., van Borkulo, C. D., Epskamp, S., Schoevers, R. A., Tuerlinckx, F., & Borsboom, D. (2016b). Measuring depression over time... Or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychological Assessment*, *28*(11), 1354.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432-441.

Goldberg, D. (2011). The heterogeneity of "major depression". *World Psychiatry*, *10*(3), 226-228.

Golino, H. F., & Christensen, A. P. (2019). EGAnet: Exploratory Graph Analysis - A framework for estimating the number of dimensions in multivariate data using network psychometrics. R package version 0.6.0.

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PloS one*, *12*(6), e0174035.

Greden, J. F. (2001). The burden of recurrent depression: causes, consequences, and prospects. *Journal of Clinical Psychiatry*, *62*, 5-9.

Halbreich, U., & Kahn, L. S. (2007). Atypical depression, somatic depression and anxious depression in women: are they gender-preferred phenotypes? *Journal of affective disorders*, *102*(1-3), 245-258.

Hartung, T. J., Fried, E. I., Mehnert, A., Hinz, A., & Vehling, S. (2019). Frequency and network analysis of depressive symptoms in patients with cancer compared to the general population. *Journal of affective disorders*, *1*(256), 295-301.

Hasin, D. S., Sarvet, A. L., Meyers, J. L., Saha, T. D., Ruan, W. J., Stohl, M., & Grant, B. F. (2018). Epidemiology of adult DSM-5 major depressive disorder and its specifiers in the United States. *JAMA Psychiatry*, *75*(4), 336-346.

Haslbeck, J. M. B., & Fried, E. I. (2017). How predictable are symptoms in psychopathological networks? A reanalysis of 18 published datasets. *Psychological medicine*, *47*(16), 2767-2776.

Haslbeck, J. M., & Waldorp, L. J. (2018). How well do network models predict observations? On the importance of predictability in network models. *Behavior research methods*, *50*(2), 853-861.

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS medicine*, *7*(7), e1000316.

Holtzheimer III, P. E., & Nemeroff, C. B. (2006). Future prospects in depression research. *Dialogues in clinical neuroscience, 8*(2), 175.

Howell, D. C. (2009). *Statistical methods for psychology*. Cengage Learning.

Howland, R. H. (2014). Multidisciplinary Treatments and Medications for Depressive Disorders and Comorbidity. In The Oxford Handbook of Depression and Comorbidity, edited by C. S. Richards and M. W. O'Hara. Print Publication Date: Jun 2014 Subject: Psychology, Health Psychology, Clinical Psychology Online Publication Date: Aug 2014 DOI: 10.1093/oxfordhb/9780199797004.013.008

Howland, R. H., Rush, A. J., Wisniewski, S. R., Trivedi, M. H., Warden, D., Fava, M., ... & Berman, S. R. (2009). Concurrent anxiety and substance use disorders among outpatients with major depression: clinical features and effect on treatment outcome. *Drug and alcohol dependence, 99*(1-3), 248-260.

Hryhorczuk, C., Sharma, S., & Fulton, S. E. (2013). Metabolic disturbances connecting obesity and depression. *Frontiers in neuroscience*, *7*, 177.

Humphry, S. M., & McGrane, J. A. (2010). Is there a contradiction between the network and latent variable perspectives? *Behavioral and Brain Sciences*, *33*(2-3), 160–161.

Kendler, K. S. (2008). Explanatory models for psychiatric illness. *American Journal of Psychiatry, 165*(6), 695-702.

Kapfhammer, H. P. (2006). Somatic symptoms in depression. *Dialogues in clinical neuroscience, 8* (2), 227.

Kendler, K. S., Aggen, S. H., Flint, J., Borsboom, D., & Fried, E. I. (2018). The centrality of DSM and non-DSM depressive symptoms in Han Chinese women with major depression. *Journal of affective disorders*, *227*, 739-744.

Kessler, R. C. (2012). The costs of depression. *Psychiatric Clinics, 35*(1), 1-14.

Kessler, R. C., Petukhova, M., & Zaslavsky, A. M. (2011). The role of latent internalizing and externalizing predispositions in accounting for the development of comorbidity among common mental disorders. *Current Opinion in Psychiatry*, *24*(4), 307.

Kornstein, S. G., Schatzberg, A. F., Yonkers, K. A., Thase, M. E., Keitner, G. I., Ryan, C. E., & Schlager, D. (1995). Gender differences in presentation of chronic major depression. *Psychopharmacology Bulletin, 31*(4), 711-718.

Kovacs, M. (1985). The children's depression inventory (CDI). *Psychopharmacol bull*, *21*(4), 995-998.

Kroenke, K., & Spitzer, R. L. (1998). Gender differences in the reporting of physical and somatofor m symptoms. *Psychosomatic Medicine, 60*(2), 150-155.

Lauritzen, S. L. (1996). *Graphical models* (Vol. 17). Clarendon Press.

van Loo, H. M., De Jonge, P., Romeijn, J. W., Kessler, R. C., & Schoevers, R. A. (2012). Data-driven subtypes of major depressive disorder: a systematic review. *BMC medicine, 10*(1), 156.

van Loo, H. M., Van Borkulo, C. D., Peterson, R. E., Fried, E. I., Aggen, S. H., Borsboom, D., & Kendler, K. S. (2018). Robust symptom networks in recurrent major depression across different levels of genetic and environmental risk. *Journal of affective disorders*, *227*, 313-322.

Löwe, B., Spitzer, R. L., Williams, J. B., Mussell, M., Schellberg, D., & Kroenke, K. (2008). Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. *General hospital psychiatry*, *30*(3), 191-199.

Madhoo, M., & Levine, S. Z. (2016). Network analysis of the Quick Inventory of Depressive Symptomatology: Reanalysis of the STAR* D clinical trial. *European Neuropsychopharmacology*, *26*(11), 1768-1774.

Marcus, S. M., Young, E. A., Kerber, K. B., Kornstein, S., Farabaugh, A. H., Mitchell, J., ... & Rush, A. J. (2005). Gender differences in depression: findings from the STAR* D study. *Journal of affective disorders*, *87*(2-3), 141-150.

McElroy, E., Napoleone, E., Wolpert, M., & Patalay, P. (2019). Structure and connectivity of depressive symptom networks corresponding to early treatment response. *EClinicalMedicine*, *8*, 29-36.

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy*, *86*, 95-104.

Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V., & Ustun, B. (2007). Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *The Lancet*, *370*(9590), 851-858.

Mullarkey, M. C., Marchetti, I., & Beevers, C. G. (2019). Using network analysis to identify central symptoms of adolescent depression. *Journal of Clinical Child & Adolescent Psychology*, *48*(4), 656-668.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, *49*(1), 115-132.

NIMH publication (2008). The national institute of mental health strategic plan. NIH Publication No. 08–6368). Retrieved from: https://www.nimh.nih.gov/about/strategic-planning-reports/index.shtml

OECD/EU (2018), *Health at a Glance: Europe 2018: State of Health in the EU Cycle*, OECD Publishing, Paris/EU, Brussels, https://doi.org/10.1787/health_glance_eur-2018-en.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social networks*, *32*(3), 245-251.

Pariante, C. M. (2009). Risk factors for development of depression and psychosis: glucocorticoid receptors and pituitary implications for treatment with antidepressant and glucocorticoids. *Annals of the New York Academy of Sciences*, *1179*, 144-52.

Pariante, C. M., & Miller, A. H. (2001). Glucocorticoid receptors in major depression: relevance to pathophysiology and treatment. *Biological psychiatry*, *49*(5), 391-404.

Pons, P., & Latapy, M. (2005, October). Computing communities in large networks using random walks. In *International symposium on computer and information sciences* (pp. 284-293). Springer, Berlin, Heidelberg.

Ramsey, J. M., Cooper, J. D., Bot, M., Guest, P. C., Lamers, F., Weickert, C. S., ... & Bahn, S. (2016). Sex differences in serum markers of major depressive disorder in the Netherlands study of depression and anxiety (NESDA). *PLoS One*, *11*(5), e0156624.

Regier, D. A., Narrow, W. E., Kuhl, E. A., & Kupfer, D. J. (2009). The conceptual development of DSM-V. *American Journal of Psychiatry*, *166*(6), 645-650.

Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnoses. *American journal of psychiatry, 170*(1), 59-70.

Ribeiro, J. D., Silva, C., & Joiner, T. E. (2014). Overarousal interacts with a sense of fearlessness about death to predict suicide risk in a sample of clinical outpatients. *Psychiatry research*, *218*(1-2), 106-112.

Robinaugh, D. J., Millner, A. J., & McNally, R. J. (2016). Identifying highly influential nodes in the complicated grief network. *Journal of Abnormal Psychology*, *125*(6), 747.

Rogers, M. L., Hom, M. A., & Joiner, T. E. (2019). Differentiating acute suicidal affective disturbance (ASAD) from anxiety and depression Symptoms: A network analysis. *Journal of affective disorders*, *250*, 333-340.

Romans, S. E., Tyas, J., Cohen, M. M., & Silverstone, T. (2007). Gender differences in the symptoms of major depressive disorder. *The Journal of nervous and mental disease*, *195*(11), 905-911.

Rush, A. J. (2007). The varied clinical presentations of major depressive disorder. *The Journal of clinical psychiatry*, 68(8), 4-10.

Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin*, *143*(8), 783.

Santini, Z. I., Fiori, K. L., Feeney, J., Tyrovolas, S., Haro, J. M., & Koyanagi, A. (2016). Social rela tionships, loneliness, and mental health among older men and women in Ireland: A prospective com munity-based study. *Journal of affective disorders*, *204*, 59-69.

Santor, D. A., Gregus, M., & Welch, A. (2006). FOCUS ARTICLE: Eight decades of measurement in depression. *Measurement: Interdisciplinary Research and Perspectives*, *4*(3), 135-155.

Santos Jr, H., Fried, E. I., Asafu-Adjei, J., & Ruiz, R. J. (2017). Network structure of perinatal depressive symptoms in Latinas: relationship to stress and reproductive biomarkers. *Research in Nursing & Health*, *40*(3), 218-228.

Schmidt, H. D., Shelton, R. C., & Duman, R. S. (2011). Functional biomarkers of depression: diagnosis, treatment, and pathophysiology. *Neuropsychopharmacology*, *36*(12), 2375.

Schuch, J. J., Roest, A. M., Nolen, W. A., Penninx, B. W., & De Jonge, P. (2014). Gender differences in major depressive disorder: results from the Netherlands study of depression and anxiety. *Journal of affective disorders*, *156*, 156-163.

Schweren, L., van Borkulo, C. D., Fried, E., & Goodyer, I. M. (2018). Assessment of symptom network density as a prognostic marker of treatment response in adolescent depression. *JAMA psychiatry*, *75*(1), 98-100.

Silverstein, B. (2002). Gender differences in the prevalence of somatic versus pure depression: a replication. *American Journal of Psychiatry*, *159*(6), 1051-1052.

Thase, M. E., Entsuah, A. R., & Rudolph, R. L. (2001). Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *The British journal of psychiatry, 178*(3), 234-241.

Uher, R., Payne, J. L., Pavlova, B., & Perlis, R. H. (2014). Major depressive disorder in DSM-5: Implications for clinical practice and research of changes from DSM-IV. *Depression and anxiety, 31*(6), 459-471.

Van de Velde, S., Bracke, P., & Levecque, K. (2010). Gender differences in depression in 23 European countries. Cross-national variation in the gender gap in depression. *Social science & medicine*, *71*(2), 305-313.

Walther, A., Rice, T., Kufert, Y., & Ehlert, U. (2017). Neuroendocrinology of a male-specific pattern for depression linked to alcohol use disorder and suicidal behavior. *Frontiers in psychiatry*, *7*, 206.

Weiss, E. L., Longhurst, J. G., & Mazure, C. M. (1999). Childhood sexual abuse as a risk factor for depression in women: psychosocial and neurobiological correlates. *American journal of psychiatry*, *156*(6), 816-828.

Weissman, M. M., Bland, R., Joyce, P. R., Newman, S., Wells, J. E., & Wittchen, H. U. (1993). Sex differences in rates of depression: cross-national perspectives. *Journal of affective disorders*, *29*(2-3), 77-84.

Woody, M. L., & Gibb, B. E. (2015). Integrating NIMH research domain criteria (RDoC) into depression research. *Current opinion in psychology*, *4*, 6-12.

World Health Organization. (2015). *Health in 2015: from MDGs, millennium development goals to SDGs, sustainable development goals*. World Health Organization.

World Health Organization. (2017). *Depression and other common mental disorders: global health estimates* (No. WHO/MSD/MER/2017.2). World Health Organization.

Appendix

R Script

```
# Creating data for women 2007 (W07)
setwd("C:\\Users\\Tarja\\Documents\\Tarja\\Gradu\\R")
library(foreign)
all_data <- read.dta("C:\\Users\\Tarja\\Documents\\Tarja\\Gradu\\R\\saarela.dta")
names <- c("Sadness", "Pessimism", "Past failure", "Loss of pleasure", "Guilty feelings",
"Punishment feelings", "Self-dislike", "Self-criticalness", "Suicidal thoughts", "Crying",
"Agitation", "Loss of Interest", "Indecisiveness", "Worthlessness", "Loss of Energy", "Sleep",
"Irritability", "Appetite", "Concentration", "Fatigue", "Sex")

data.W07 <- na.omit(subset(data, gender == 1,
                           select = colnames(data)[grep("_07", colnames(data))]))
data.W12 <- na.omit(subset(data, gender == 1,
                           select = colnames(data)[grep("_12", colnames(data))]))
data.M07 <- na.omit(subset(data, gender == 2,
                           select = colnames(data)[grep("_07", colnames(data))]))
data.M12 <- na.omit(subset(data, gender == 2,
                           select = colnames(data)[grep("_12", colnames(data))]))

# network estimation
library(bootnet)
networkW07 <- estimateNetwork(data.W07, default="EBICglasso", threshold = FALSE,
corMethod="cor_auto", tuning=0.5)
networkW12 <- estimateNetwork(data.W12, default="EBICglasso", threshold = FALSE,
corMethod="cor_auto", tuning=0.5)
networkM07 <- estimateNetwork(data.M07, default="EBICglasso", threshold = FALSE,
corMethod="cor_auto", tuning=0.5)
networkM12 <- estimateNetwork(data.M12, default="EBICglasso", threshold = FALSE,
corMethod="cor_auto", tuning=0.5)
```

```
# attaining density and the number of negative and positive edges
sum(abs(networkW07$graph))/2 #12.48532
sum(abs(networkW12$graph))/2 #12.10573
sum(abs(networkM07$graph))/2 #17.08223
sum(abs(networkM12$graph))/2 #22.78138

edgevalue_M07 <- networkM07$graph[upper.tri(networkM07$graph)]
edgevalue_M12 <- networkM12$graph[upper.tri(networkM12$graph)]
edgevalue_W07 <- networkW07$graph[upper.tri(networkW07$graph)]
edgevalue_W12 <- networkW12$graph[upper.tri(networkW12$graph)]

sum(edgevalue_M07!= 0)#tot 152
sum(edgevalue_M07> 0) #pos 107
sum(edgevalue_M07<0) #neg 45

sum(edgevalue_M12!= 0)# tot 184
sum(edgevalue_M12> 0) #pos 122
sum(edgevalue_M12<0) #neg 62

sum(edgevalue_W07!= 0) #tot 128
sum(edgevalue_W07> 0) #pos 103
sum(edgevalue_W07<0) #neg 25

sum(edgevalue_W12!= 0) #tot 130
sum(edgevalue_W12> 0) #pos 113
sum(edgevalue_W12<0) #neg 17

# plotting networks with average layout
library(qgraph)
Layout <- averageLayout(networkW07, networkM07, networkW12, networkM12, layout =
"spring")
layout(matrix(1:2, nrow =1))
```

```
plot(networkW07, maximum= 0.62, cut= 0.15, layout = Layout, details= TRUE, labels = TRUE,
vsize=6, title = "Women 2007", nodeNames = names, legend.cex = 0.4)
plot(networkW12, maximum= 0.62, cut= 0.15, layout = Layout, details= TRUE, labels = TRUE,
vsize=6, title = "Women 2012", nodeNames = names, legend.cex = 0.4)
plot(networkM07, maximum= 0.62, cut= 0.15, layout = Layout, details= TRUE, labels = TRUE,
vsize=6, title = "Men 2007", nodeNames = names, legend.cex = 0.4)
plot(networkM12, maximum= 0.62, cut= 0.15, layout = Layout, details= TRUE, labels = TRUE,
vsize=6, title = "Men 2012", nodeNames = names, legend.cex = 0.4)

# centrality analysis
centralityPlot(Year2007 = list(Men2007 = networkM07, Men2012 = networkM12), labels = names,
include =c("ExpectedInfluence"), orderBy = "ExpectedInfluence")
centralityPlot(Year2007 = list(Women2007 = networkW07, Women2012 = networkW12), labels =
names, include =c("ExpectedInfluence"), orderBy = "ExpectedInfluence")

centrality_auto(networkM07)
centrality_auto(networkM12)
centrality_auto(networkW07)
centrality_auto(networkW12)

nodebootM07 <- bootnet(networkM07, statistics =c("ExpectedInfluence", "Strength",
"Betweenness", "Closeness"), type = "case")
nodebootM12 <- bootnet(networkM12, statistics =c("ExpectedInfluence", "Strength",
"Betweenness", "Closeness"), type = "case")
nodebootW07 <- bootnet(networkW07, statistics =c("ExpectedInfluence", "Strength",
"Betweenness", "Closeness"), type = "case")
nodebootW12 <- bootnet(networkW12, statistics =c("ExpectedInfluence", "Strength",
"Betweenness", "Closeness"), type = "case")

# CS-coefficients
corStability(nodebootM07) #ExpectedInfluence: 0.439, strength: 0.128
corStability(nodebootM12) # ExpectedInfluence: 0.283, strenght: 0.128
corStability(nodebootW07) # ExpectedInfluence: 0.517, strength: 0.128
corStability(nodebootW12) # ExpectedInfluence: 0.283, strength: 0.205
```

```
# Community analysis
install.packages("devtools")
library("devtools")
devtools::install_github('hfgolino/EGA')
library("EGAnet")
ega_M07<-EGA(data.M07, model = c("glasso"), plot.EGA = TRUE)
ega_M12<-EGA(data.M12, model = c("glasso"), plot.EGA = TRUE)
ega_W07<-EGA(data.W07, model = c("glasso"), plot.EGA = TRUE)
ega_W12<-EGA(data.W12, model = c("glasso"), plot.EGA = TRUE)

plot(ega_M07, layout = "Spring", labels = TRUE, vsize=6, border.width=1,5, nodeNames = names,
legend.cex = 0.4, border.width=1.5, color=c("red", "green", "orange"), title = "Men 2007")
plot(ega_M12, layout = "Spring", labels = TRUE, vsize=6, border.width=1.5, nodeNames = names,
legend.cex = 0.4, border.width=1.5, color=c("red", "green", "orange"), title = "Men 2012")
plot(ega_W07, layout = "Spring", labels = TRUE, vsize=6, border.width=1.5, nodeNames = names,
legend.cex = 0.4, border.width=1.5, color=c("red", "green", "orange"), title = "Women 2007")
plot(ega_W12, layout = "Spring", labels = TRUE, vsize=6, border.width=1.5, nodeNames = names,
legend.cex = 0.4, border.width=1.5, color=c("red", "green", "orange", "white"), title = "Women
2012")

# Bootstrapped mean differences between W07 and M12
x <- rowSums(data.W07)
y <- rowSums(data.M12)
pooled <- c(x, y)
xt <- x - mean(x) + mean(pooled)
yt <- y - mean(y) + mean(pooled)
boot.t <- c(1:10000)
for(i in 1:10000){
  sample.x <- sample(xt,replace=TRUE)
  sample.y <- sample(yt,replace=TRUE)
  boot.t[i] <- t.test(sample.x,sample.y)$statistic}
p.h0 <-  (1 + sum(abs(boot.t) > abs(t.test(x,y)$statistic))) / (10000+1)
p.h0
```

```
# Bootstrapped CIs
data.list <- list(data.W07, data.W12, data.M07, data.M12)
CI.data <- matrix(rep(0, times = 2*4*21), 8, 21)
boot.means <- matrix(rep(0, times = 21*10000), 10000, 21)
for(i in 1:4){ x.data <- data.list[[i]] for(q in 1:21){x <- x.data[,q]
for(p in 1:10000){boot.means[p,q] <- mean(sample(x, replace = T))}
boot.means[,q] <- sort(boot.means[,q])
CI.data[(i*2)-1, q] <- boot.means[250, q]
CI.data[(i*2), q] <- boot.means[9750, q]}
}
rownames(CI.data) <- c("W07 Lower", "W07 Upper", "W12 Lower", "W12 Upper", "M07 Lower",
"M07 Upper", "M12 Lower", "M12 Upper")

# The Ising model
# creating binary data
data.W07.di <- apply(data.W07, c(1,2), function(x){ifelse(x<1, 0, 1)})
data.W12.di <- apply(data.W12, c(1,2), function(x){ifelse(x<1, 0, 1)})
data.M07.di <- apply(data.M07, c(1,2), function(x){ifelse(x<1, 0, 1)})
data.M12.di <- apply(data.M12, c(1,2), function(x){ifelse(x<1, 0, 1)})

# network estimation and plotting
library(IsingFit)
library(qgraph)
resM07.di <- IsingFit(as.matrix(data.M07.di),gamma = 0.25, plot=FALSE)
resM12.di <- IsingFit(as.matrix(data.M12.di),gamma = 0.25, plot=FALSE)
resW07.di <- IsingFit(as.matrix(data.W07.di),gamma = 0.25, plot=FALSE)
resW12.di <- IsingFit(as.matrix(data.W12.di),gamma = 0.25, plot=FALSE)

GraphW07.di <- qgraph(resW07.di$weiadj, layout="spring", maximum= 2.41, cut=.15, details=
TRUE, labels = TRUE, vsize=6, title = "Women 2007", nodeNames = names, legend.cex = 0.4)
GraphW12.di <- qgraph(resW12.di$weiadj, layout="spring", maximum= 2.41, cut=.15, details=
TRUE, labels = TRUE, vsize=6, title = "Women 2012", nodeNames = names, legend.cex = 0.4)
GraphM07.di <- qgraph(resM07.di$weiadj, layout="spring", maximum= 2.41, cut=.15, details=
TRUE, labels = TRUE, vsize=6, title = "Men 2007", nodeNames = names, legend.cex = 0.4)
```

```
GraphM12.di <- qgraph(resM12.di$weiadj, layout="spring", maximum= 2.41, cut=.15, details= TRUE, labels = TRUE, vsize=6, title = "Men 2012", nodeNames = names, legend.cex = 0.4)

# NCT (Ising model)
library(NetworkComparisonTest)

# Women 2007 - Men 2007:
NCTW07M07 <- NCT(data.M07.di, data.W07.di, it= 1000, binary.data= TRUE, paired= TRUE, test.edges= TRUE, edges= "all")

NCTW07M07$glstrinv.sep # densities 46.24660 49.91756
NCTW07M07$glstrinv.real # difference 3.670964
NCTW07M07$glstrinv.pva # p-value 0.312

# Women 2007 – Men 2012
NCTW07M12 <- NCT(data.M12.di, data.W07.di, it= 1000, binary.data= TRUE, paired= TRUE, test.edges= TRUE, edges= "all")

NCTW07M12$glstrinv.sep # 50.21134 49.91756
NCTW07M12$glstrinv.real # 0.2937802
NCTW07M12$glstrinv.pva # 0.928

# Women 2012 - Men 2007:
NCT_W12_M07 <- NCT(data.M07.di, data.W12.di, it= 1000, binary.data= TRUE, paired= TRUE, test.edges= TRUE, edges= "all")

NCTW12M07$glstrinv.sep # densities 46.24660 49.64927
NCTW12M07$glstrinv.real #difference 3.402677
NCTW12M07$glstrinv.pva #p-value of difference 0.384

# Women 2012 – Men 2012
NCTW12M12 <- NCT(data.M12.di, data.W12.di, it= 1000, binary.data= TRUE, paired= TRUE, test.edges= TRUE, edges= "all")
```

NCTW12M12$glstrinv.sep # densities 50.21134 49.64927

NCTW12M12$glstrinv.real #difference 0.5620673

NCTW12M12$glstrinv.pva #p-value 0.884