Identifying Key Targets for Interventions to Improve Psychological Wellbeing: Replicable Results from Four UK Cohorts

Stochl, J.[1,3*+], Soneson, E.[1+], Wagner, A.P.[2,3], Khandaker, G.M.[1], Goodyer, I. [1], Jones, P.B. [1,3]

Financial support:
The NSPN study was supported by the Wellcome Trust Strategic Award (095844/Z/11/Z). JS, APW, and PBJ received support from the NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) East of England (EoE). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. GMK is supported by an Intermediate Clinical Fellowship from the Wellcome Trust (201486/Z/16/Z).

Word count: 3,613

+ joint first authors
[1] Department of Psychiatry, University of Cambridge, Cambridge Biomedical Campus, Herschel Smith Building for Brain & Mind Sciences, Cambridge, CB2 0SZ, Cambridge, UK
[2] Norwich Medical School, University of East Anglia, Norwich, UK
[3] National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East of England (CLAHRC), Cambridge, UK

* *Address for correspondence:*
Department of Psychiatry
University of Cambridge
Cambridge Biomedical Campus
Herchel Smith Building for Brain & Mind Sciences
Cambridge, CB2 0SZ
United Kingdom
Phone: +44 07587146299
Email: js883@cam.ac.uk

Introduction

Mental health and wellbeing are becoming increasingly prominent in national and international health policy (Department of Health and Social Care, 2011, Mehta *et al.*, 2015, World Health Organization, 2002, 2004). At a societal level, they represent important resources closely linked to social, human, and economic capital (Friedli and Parsonage, 2007, Knapp *et al.*, 2011), and are associated with lower levels of inequality, less community violence, and higher life expectancy (Friedli and World Health Organization, 2009). For individuals, mental health and wellbeing are closely connected to normal functioning and quality of life and are instrumental in creating and maintaining good relationships (Jané-Llopis *et al.*, 2005, World Health Organization, 2004). Clinically, the growing evidence for the existence of a 'continuum' of psychopathology (also referred to as 'common mental distress' or the 'general psychopathology factor') (Caspi *et al.*, 2014, Stochl *et al.*, 2015) suggests that improving mental health and wellbeing may also help to prevent the development of mental disorders.

Several approaches have been suggested for improving mental health and wellbeing, including psychological therapies (Fava *et al.*, 1998, Galante *et al.*, 2017, Slade, 2010), school and workplace interventions (Jané-Llopis *et al.*, 2005, Jané-Llopis and Barry, 2005, Knapp *et al.*, 2011, Weare and Nind, 2011), improvement of housing and nutrition, reduction of substance misuse, and prevention of violence (Jané-Llopis *et al.*, 2005, World Health Organization, 2004). Despite their promise, however, many of these approaches have been criticised for their lack of supporting empirical evidence (Mehta *et al.*, 2015). Indeed, current methods used to inform intervention targets are mainly limited to theoretical models (e.g. Ryff's model of wellbeing; general stress theory), literature reviews, and qualitative methods (e.g. interviews with experts and service users), and do not consider any type of quantitative method.

Psychological network analysis is an innovative statistical approach that can complement theoretical knowledge and clinical expertise by providing quantitative evidence for the identification of intervention targets. Essentially, it examines relationships between different items on clinical questionnaires, and determines which items are most 'central' to the condition of interest due to having strong relationships with other items. Central items may then serve as indicators for clinical intervention targets (Fried *et al.*, 2017), as their improvement is most likely to destabilise harmful network structures and prevent exacerbation of other items (Smith *et al.*, 2018). Network analysis has been used to suggest potential intervention targets for depression (van Borkulo *et al.*, 2015), PTSD (Fried *et al.*, 2018), and eating disorders (Smith *et al.*, 2018). Furthermore, it aligns with the clinical characterisation of psychopathology as a system of causal relationships between symptoms, where some symptoms are more influential than others (van Borkulo *et al.*, 2015).

To make valid inferences in network analysis, comprehensive tools to measure mental health and wellbeing, such as the well-established Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS), are crucial. In this study, we have used psychological network analysis to identify items central to the WEMWBS, which we present as potentially optimal targets for interventions aiming to improve mental health and wellbeing.

Methods

Participants

This study sample comprises 47,578 participants from four different UK cohorts.

National Child Development Study (NCDS)

The NCDS (University of London, 2012) is a major longitudinal British cohort study initiated in 1958. As such, this sample is homogeneous for age. At age 53, 8,643 NCDS participants (51.8% women) completed the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) as part of a set of self-report questionnaires. Full details on sampling design and data collection can be found at https://tinyurl.com/y7q2m66z.

Northern Ireland Health Survey (NIHS)

The NIHS (Department of Health Northern Ireland, 2016) covers a range of health topics important to the lives of people in Northern Ireland. The survey has been annually conducted since 2010. Respondents are sampled from those aged 16+ living in private households. The 2010-2011 survey collected wellbeing data from 4,161 individuals of which 3,873 (58.8% women) had complete WEMWBS data. Details about the data collection methodology can be found at https://tinyurl.com/ybfakdsm.

Neuroscience in Psychiatry Network (NSPN)

The NSPN (Kiddle *et al.*, 2018) cohort consists of 2,403 participants, aged 14-25, recruited from Cambridgeshire, London and surrounding areas. The sample analysed here was recruited between November 2012 and July 2017. Study invites were sent through general practice (GP) surgeries and schools with the aim of recruiting 200 women and 200 men for each of five age strata (ages: 14-15; 16-17; 18-19; 20-21; 22-24). Complete WEMWBS data was available from 2,337 individuals (53.8% women).

Scottish Schools Adolescent Lifestyle and Substance Use Survey (SALSUS)

The SALSUS (NHS National Services Scotland, 2013) survey was set up by the Scottish Government to monitor progress on reducing smoking and substance misuse. Information from the survey helps national planning and facilitates the monitoring of policy implementation. The WEMWBS data used in this study were collected in 2010 from 32,725 individuals (49.4%, women) from the second (age 12-14) and fourth (age 14-16) years of secondary school. Full details can be found at https://tinyurl.com/ya66mdq4.

The Warwick–Edinburgh Mental Well-being Scale (WEMWBS)

The WEMWBS (Tennant *et al.*, 2007) is a 14-item, self-report measure designed to assess a range of wellbeing concepts including affective-emotional aspects, cognitive-evaluative dimensions, and psychological functioning in the general population. All items are worded positively and have 5 response categories (1-None of the time; 2-Rarely; 3- Some of the time; 4-Often; 5-All of the time). The wellbeing score is computed as sum of all items (range: 14-70), with higher scores representing better wellbeing. The WEMWBS was found to be a uni-dimensional measure and to have desirable psychometric properties (Tennant *et al.*, 2007). The scale is well-regarded by service users and their carers, who tend to prefer it to other mental health and wellbeing measures (Crawford *et al.*, 2011) for the way that it asks about positive aspects of mental health.

Analysis

Psychological network analysis (Borsboom and Cramer, 2013) conceptualises behaviour as a complex interplay of psychological and other components. Recently, this methodology has become popular in psychometrics partly due to its ability to identify worthwhile items for intervention development in questionnaires and surveys. In typical network analysis applied to questionnaire data (Gaussian graphical models), nodes (representing questionnaire items) are interconnected via edges (representing partial correlations) (Costantini *et al.*, 2015). The use of partial correlations ensures that bivariate relationships between nodes are not confounded by relationships to other variables in the network and provides unbiased computation of centrality indices. Networks in this paper utilise the ‘spring’ layout (Fruchterman and Reingold, 1991b), where nodes are positioned on a plane so that distances between them relate to the size of their partial correlations.

Typically, the network in each cohort is estimated separately and sparsity (and thus improved interpretability) of such networks is achieved by the application of an adaptive graphical LASSO penalty (Friedman *et al.*, 2008). However, recent developments allow for joint estimation of multiple networks using *fused graphical LASSO* (FGL) (Danaher *et al.*, 2014). FGL extends traditional graphical LASSO by extending the penalty function to incorporate differences among corresponding edge-weights estimated across networks. This strategy neither masks nor inflates similarities across networks (Fried *et al.*, 2018). In this study, the optimal value of this penalty was achieved by k-fold cross-validation. A detailed explanation of FGL and its use in psychological networks is given elsewhere (Costantini *et al.*, 2017, Danaher *et al.*, 2014, Fried *et al.*, 2018). The similarity of networks was assessed by calculating the Spearman correlation of edge-weights between each pair of networks (Borsboom, 2017).

The relative importance of questionnaire items is subsequently evaluated using measures from graph theory, using typical centrality indices such as *strength, closeness,* and *betweenness* (Newman, 2010). A *strong* central node (item) (Barrat *et al.*, 2004) is one that can influence many other nodes (or be influenced by them) directly, without considering the mediating role of other nodes (Costantini *et al.*, 2015). As such, *strength* is the crucial index for identification of items for developing the most effective interventions. Nodes with high *closeness* (defined as the inverse of the sum of distances of the focal node to all other nodes in the network) are those whose responses are likely to be quickly affected by changes in other nodes, either directly or indirectly. If nodes with high *betweenness* are removed from a network then the distance among other nodes will generally increase (Costantini *et al.*, 2015). As such, nodes with high *betweenness* speed up the flow of information in networks.

Lack of accuracy and network stability have been recognised as an important issue in psychological networks (Epskamp *et al.*, 2017, Forbes *et al.*, 2017). Thus, bootstrapping procedures have been developed for psychological networks to address this issue and prevent biased inferences about the importance of individual nodes (Epskamp *et al.*, 2017). To evaluate accuracy and stability we have followed recommendations made by Epskamp *et al*. (2017). They proposed the correlation stability (CS) coefficient to investigate the stability of the order of centrality indices after observing only portions of the data. Its computation is based on case dropping bootstrap methods. The CS coefficient can be interpreted as the maximum proportion of cases that can be dropped, such that with 95% probability the correlation between the original centrality indices and the centrality of networks based on subsets is 0.7 or higher (this figure can be changed but is taken as a default based on a simulation study by Epskamp *et al.* (2017)). This coefficient should not drop below 0.25 and should ideally be above 0.5 to justify robust interpretation of centrality indices.

Functions from the R (R Core Team, 2017) packages ‘qgraph’ (Epskamp *et al.*, 2012), ‘EstimateGroupNetwork’ (Costantini and Epskamp, 2017), and ‘mgm’ (Haslbeck and Waldorp, 2016) were used to estimate the network graphs. Given that the WEMWBS items are ordinal, polychoric correlations are used in the input weight matrix. The resulting networks were plotted using the *spring* layout (Fruchterman and Reingold, 1991a) where more related edges are plotted closer together. Bootstrapping of networks was accomplished using the R package ‘bootnet’ (Epskamp *et al.*, 2017). To assess network differences (global network *strength*, edges) with respect to gender, permutation tests implemented in the package ‘NetworkComparisonTest’ (van Borkulo *et al.*, 2016) were used with 5,000 iterations. All p-values were corrected for multiple testing (using Holm-Bonferroni correction), where applicable.

Ethical approvals

Ethical approval was not required for the present secondary data analysis.

Results

Table 1 shows the basic item descriptive statistics for each cohort.

------------------------------ insert Table 1 about here ------------------------------

Estimated networks are shown in Figure 1. Visual comparison reveals similarities across them: for example, items 8 (*I have been feeling good about myself*) and 14 (*I have been feeling cheerful*) are always central. Item 10 (*I have been feeling confident*) seems to have a more prominent role in both the older (NCDS) and younger adult (NSPN) cohorts. Conversely, items such as 1 (*I have been feeling optimistic about the future*), 2 (*I have been feeling useful*), and 5 (*I have had energy to spare*) are generally on the periphery of the networks and less connected with other items. The formal comparison of networks (using a permutation test) revealed statistically significant differences in global network *strength* between NCDS and SALSUS (network strength NCDS=6.75, network strength SALSUS=6.23, p<0.001) and also between NIHS and SALSUS (network strength NIHS=6.54, network strength SALSUS=6.23, p=0.002). On average, around six edges between each pair of networks are statistically different. Information about significant differences in edge weights is available from the authors upon request. We formally compare centrality indices later in this paper.

------------------------------ insert Figure 1 about here ------------------------------

Comparison of edge-weights and their accuracy

To improve visual comparability of edges, we also estimated the average layout of these four networks and plotted all networks using this layout (see Figure 2). The patterns of relationships among items are similar across samples. Items 8 and 10, which evaluate self-perception, are highly related. The same holds for items 4, 9, and 12, which assess relationships with other people, and items 6, 7, and 11, which deal with processing ideas and problems. It is less clear why items 1 (*I have been feeling optimistic about the future*) and 2 (*I have been feeling useful*) are related.

------------------------------ insert Figure 2 about here ------------------------------

The visual similarity of networks was confirmed by investigating Spearman correlations of edge-weights for all pairs of networks, presented in Supplementary Table 1. They ranged from 0.75 to 0.87, suggesting high similarity across networks.

Supplementary Figure 1 depicts point estimates and bootstrap confidence intervals of the edge values for each network. In general, confidence intervals suggest that the accuracy of edges is satisfactory. As expected, the confidence intervals are smaller in larger samples.

Centrality indices and their stability

Standardised centrality indices for each item, computed for each network, are shown in Figure 3. The indices are remarkably similar across all networks. With respect to *strength* and *closeness*, the three most central items across all networks are items 8, 10, and 14. *Betweenness* of these three items is also highest in NCDS and NSPN. The top three *betweenness* items in NIHS are items 8, 7, and 9. In SALSUS, the top *betweenness* item is item 8 but next highest *betweenness* is indistinguishable for items 4, 9, and 10. These results suggest that the wellbeing intervention targets (as measured by *strength*) replicate well across cohorts. The same holds for *closeness*. Mediating items which speed up influence of changes in the network (*betweenness*) vary only slightly across cohorts.

------------------------------ insert Figure 3 about here ------------------------------

Stability of the centrality indices was assessed using the case dropping bootstrap (Epskamp *et al.*, 2017). The results from are plotted in Supplementary Figure 2, and corresponding CS coefficients are given in Table 2.

These results show that *closeness* and *strength* are very stable (even with only 25% of cases, the order of centrality indices has not considerably changed). *Betweenness* is slightly less stable, but apart from NIHS sample, its confidence intervals are still above the recommended cut-off of 0.5. Therefore, *betweenness* in NIHS should be interpreted with caution.

------------------------------ insert Table 2 about here ------------------------------

Gender differences

Network structures, and thus wellbeing intervention targets, might be different for men and women. We have therefore tested for statistically significant gender differences in global network *strength* and edge-weights. Regardless of cohort, there were no statistically significant differences by gender in global network *strength* (p-values: NCDS=0.163; NIHS=0.422; NSPN=0.696; SALSUS=0.474). No differences in edge-weights were found in the NIHS or NSPN cohorts. In NCDS, a significant difference between men and women was found for the edge between items 8 and 10 (0.33 for men, 0.45 for women, p=0.035). This suggests the link between item 8 (*feeling good about oneself*) and item 10 (*feeling confident*) is stronger for middle-aged women than for middle-aged men. Even in the very large SALSUS cohort (n=32,725), only six edges (out of 91) were significantly different between men and women (p(item 8, item

9)<0.001; p(item 8, item 10)<0.001; p(item 2, item 10)<0.001; p(item 7, item 10)<0.001; p(item 8, item 14)<0.001; p(item 9, item 14)=0.017). In this cohort, links between (1) item 8 (*feeling good about oneself*) and item 10 (*feeling confident*) and (2) item 9 (*feeling close to others*) and item 14 (*feeling cheerful*) were stronger for women than for men. Conversely, links between (1) item 2 (*feeling useful*) and item 10 (*feeling confident*), (2) item 7 (*thinking clearly*) and item 10 (*feeling confident*), (3) item 8 (*feeling good about oneself*) and item 9 (*feeling close to others*), and (4) item 8 (*feeling good about oneself*) and item 14 (*feeling cheerful*) were all stronger for men than for women. On the whole, the relatively small number of significantly different edges suggests that gender differences in these wellbeing networks are minimal.

Discussion

This study aimed to identify the central aspects of psychological wellbeing, which may thus be considered as important intervention targets. Score improvements on these items should have the largest positive impact on other aspects of psychological wellbeing. To find these keystones, we used psychological network analysis to identify the most central items within graph networks created from a well-established psychological wellbeing measure (WEMWBS). The WEMWBS data were obtained from four major UK cohorts varying with respect to age (young people (SALSUS), adolescents and young adults (NSPN), general adult population (NIHS), and middle-aged adults (NCDS)) and location (England, Northern Ireland, and Scotland).

Generally, results were consistent across cohorts. Edge-weights showed very similar patterns across cohorts and were accurate enough to make valid inferences about network architecture. This suggests high replicability of the network structure and high generalisability of findings across ages and geographical locations within the UK.

To highlight optimal targets that maximise intervention effectiveness, the most important items are those central to a network. The top three items, as measured by *strength* are items 8 (*I have been feeling good about myself*), 10 (*I have been feeling confident*), and 14 (*I have been feeling cheerful*). This suggests that positive self-perception and cheerfulness may play the most important role in influencing other aspects of psychological wellbeing. Due to the undirected character of the network it is not surprising that these items demonstrate the highest levels of *closeness*, indicating that they are easily influenced by other network nodes. The least influential items vary slightly across samples, but often include items 1 (*I have been feeling optimistic about the future*), 5 (*I have had energy to spare*), 6 (*I have been dealing with problems well*), and 11 (*I have been able to make up my own mind about things*). This suggests that improving upon processing problems, energy, and future expectations may have the smallest effect on other aspects of wellbeing.

These inferences seem to be robust given the high stability of centrality indices. Apart from *betweenness* in the NIHS cohort (which has questionable interpretability due to poor stability), all other correlation stability coefficients were above the recommended criteria of 0.50 (Epskamp *et al.*, 2017).

Gender differences in network architecture (global *strength*, size of edges) were also assessed to determine if interventions targets might differ for men and women. Omnibus tests of global network *strength* suggested no gender differences in any sample. Given there was only a total of seven edge differences by gender across all four cohorts, our results suggest that interventions targets are unlikely to differ by gender.

Strengths and limitations

A key strength of this study is that it utilises a number of cohorts, addressing the considerable concern about the replicability crisis in network literature (Forbes *et al.*, 2017). In addition, the considered cohorts are large and cover a wide range of age and geographical locations, supporting the generalizability of findings.

A substantial limitation is the use of cross-sectional data, which constrains network analysis to undirected networks. Using undirected networks in turn limits inferences about the direction of influence. Although presented network edges can be interpreted as putative causal paths, it is equally likely that influence flows from A to B as from B to A (other scenarios are also possible including mediation by another node C). Indeed, it seems plausible that feeling good about yourself (item 8), being confident (item 10), and feeling cheerful (item 14) might be the consequence rather than cause of other aspects of wellbeing considered here (e.g. feeling relaxed, loved by others, or thinking positively about the future). An intervention affecting only the end-points of a causality chain, as in this scenario, is likely to have only limited, if any, impact on mental wellbeing (Fried *et al.*, 2018). Experimental studies that intervene directly on the central symptoms are therefore needed to test whether this would indeed affect other symptoms in an expected way (Fried and Cramer, 2017).

Additionally, as clearly described in Fried, Eidhof *et al.* (2018), there are at least two other reasons why using central items as intervention targets should be considered with caution. First, feedback loops, which are difficult to detect in undirected networks, can make central items the most resilient to change. Second, peripheral items should not automatically be regarded as clinically unimportant; their importance should be also considered based on substantive clinical arguments. However, despite all these limitations, Fried, Eidhof *et al.* (2018, page 11) conclude, 'If we had to put our money on selecting a clinical feature as an intervention target in the absence of all other clinical information, […] choosing the most central node might be a viable heuristic.'

Implications for practice

Our findings have implications for the design of national mental health and wellbeing strategies for all ages. Positive self-perception and confidence in children and young people could be improved effectively at schools (e.g. bullying prevention programmes) or at home (e.g. positive parenting programmes), and in adults at the workplace (e.g. through regular training and supervision; fostering positive and supporting working environments). Indeed, the UK government expects schools and employers to play active roles in promoting population mental health and wellbeing (Department of Health and Social Care, 2011). Furthermore, although our findings are based on general population samples, they may be useful for providing care for people seeking treatment for mental disorders. Since evidence suggests that psychological wellbeing and mental ill health exist on a continuum (Böhnke and Croudace, 2016, Caspi *et al.*, 2014, St Clair *et al.*, 2017, Stochl *et al.*, 2015), it is likely that improving wellbeing in mentally unwell individuals may also lead to improvements in their clinical symptoms. Finally, our analysis may also have implications for the development and trialling of psychological therapies as they indicate that interventions that focus on improving self-esteem and confidence may be more effective in increasing overall wellbeing than those that do not focus on these qualities.

Conclusions

In conclusion, our study shows that the most worthwhile intervention targets for improvement of psychological wellbeing are aspects related to positive self-perception and positive mood. Regardless of gender, their improvement is likely to have a positive impact on the remaining aspects of psychological wellbeing, either directly or indirectly.

Acknowledgements

The authors would like to thank Michele Atiemo for her comments on early versions of this paper.

Conflict of interest

None.

References

Barrat, A., Barthelemy, M., Pastor-Satorras, R. & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 3747-52.

Böhnke, J. R. & Croudace, T. J. (2016). Calibrating well-being, quality of life and common mental disorder items: psychometric epidemiology in public mental health research. *British Journal of Psychiatry* **209**, 162-168.

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

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

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R. & Moffitt, T. E. (2014). The p Factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science* **2**, 119-137.

Costantini, G. & Epskamp, S. (2017). EstimateGroupNetwork: Perform the Joint Graphical Lasso and Selects Tuning Parameters. R package version 0.1.2.

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.

Costantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S. & Perugini, M. (2017). Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*.

Crawford, M. J., Robotham, D., Thana, L., Patterson, S., Weaver, T., Barber, R., Wykes, T. & Rose, D. (2011). Selecting outcome measures in mental health: the views of service users. *Journal of Mental Health* **20**, 336-46.

Danaher, P., Wang, P. & Witten, D. M. (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **76**, 373-397.

Department of Health and Social Care (2011). *No health without mental health: a cross-government mental health outcomes strategy for people of all ages*. Stationery Office: London, UK.

Department of Health Northern Ireland (2016). Northern Ireland Health Survey, 2010-2011. UK Data Service.

Epskamp, S., Borsboom, D. & Fried, E. I. (2017). Estimating Psychological Networks and their Accuracy: A Tutorial Paper. *Behavior Research Methods*.

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D. & Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software* **48**, 1-18.

Fava, G. A., Rafanelli, C., Cazzaro, M., Conti, S. & Grandi, S. (1998). Well-being therapy. A novel psychotherapeutic approach for residual symptoms of affective disorders. *Psychological Medicine* **28**, 475-80.

Forbes, M. K., Wright, A. G., Markon, K. E. & Krueger, R. F. (2017). Evidence that psychopathology symptom networks have limited replicability. *Journal of abnormal psychology* **126**, 969.

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

Fried, E. I., Eidhof, M. B., Palic, S., Costantini, G., Huisman-van Dijk, H. M., Bockting, C. L., Engelhard, I., Armour, C., Nielsen, A. B. & Karstoft, K.-I. (2018). Replicability and Generalizability of Posttraumatic Stress Disorder (PTSD) Networks: A Cross-Cultural Multisite Study of PTSD Symptoms in Four Trauma Patient Samples. *Clinical Psychological Science*, 2167702617745092.

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

Friedli, L. & Parsonage, M. (2007). Building an economic case for mental health promotion: part I. *Journal of Public Mental Health* **6**, 14-23.

Friedli, L. & World Health Organization (2009). Mental health, resilience and inequalities.

Friedman, J., Hastie, T. & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9**, 432-41.

Fruchterman, T. M. J. & Reingold, E. M. (1991a). Graph drawing by force-directed placement. *Software: Practice and Experience* **21**, 1129-1164.
Fruchterman, T. M. J. & Reingold, E. M. (1991b). Graph drawing by force-directed placement. *Software: Practice and Experience* **21**, 1129-1164.
Galante, J., Dufour, G., Vainre, M., Wagner, A. P., Stochl, J., Benton, A., Lathia, N., Howarth, E. & Jones, P. B. (2017). A mindfulness-based intervention to increase resilience to stress in university students (the Mindful Student Study): a pragmatic randomised controlled trial. *Lancet Public Health* **3**, e72-e81.
Haslbeck, J. M. & Waldorp, L. J. (2016). mgm: Structure Estimation for time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software*.
Jané-Llopis, E., Barry, M., Hosman, C. & Patel, V. (2005). Mental health promotion works: a review. *Promotion & Education* **12**, 9-25.
Jané-Llopis, E. & Barry, M. M. (2005). What makes mental health promotion effective? *Promotion & education* **12**, 47-54.
Kiddle, B., Inkster, B., Prabhu, G., Moutoussis, M., Whitaker, K. J., Bullmore, E. T., Dolan, R. J., Fonagy, P., Goodyer, I. M. & Jones, P. B. (2018). Cohort Profile: The NSPN 2400 Cohort: a developmental sample supporting the Wellcome Trust NeuroScience in Psychiatry Network. *International Journal of Epidemiology* **47**, 18-19g.
Knapp, M., McDaid, D. & Parsonage, M. (2011). Mental health promotion and mental illness prevention: The economic case.
Mehta, N., Croudace, T. & Davies, S. C. (2015). Public mental health: evidenced-based priorities. *Lancet* **385**, 1472-5.
Newman, M. (2010). *Networks: an introduction*. Oxford university press.
NHS National Services Scotland (2013). Scottish Schools Adolescent Lifestyle and Substance Use Survey, 2010. UK Data Service.
R Core Team (2017). R: A Language and Environment for Statistical Computing.
Slade, M. (2010). Mental illness and well-being: the central importance of positive psychology and recovery approaches. *BMC health services research* **10**, 26.
Smith, K. E., Crosby, R. D., Wonderlich, S. A., Forbush, K. T., Mason, T. B. & Moessner, M. (2018). Network analysis: An innovative framework for understanding eating disorder psychopathology. *International Journal of Eating Disorders* **51**, 214-222.
St Clair, M. C., Neufeld, S., Jones, P. B., Fonagy, P., Bullmore, E. T., Dolan, R. J., Moutoussis, M., Toseeb, U. & Goodyer, I. M. (2017). Characterising the latent structure and organisation of self-reported thoughts, feelings and behaviours in adolescents and young adults. *PloS one* **12**, e0175381.
Stochl, J., Khandaker, G. M., Lewis, G., Perez, J., Goodyer, I. M., Zammit, S., Sullivan, S., Croudace, T. J. & Jones, P. B. (2015). Mood, anxiety and psychotic phenomena measure a common psychopathological factor. *Psychological Medicine* **45**, 1483-93.
Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J. & Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health and Quality of Life Outcomes* **5**, 63-63.
University of London (2012). National Child Development Study: Sweep 8, 2008-2009. UK Data Service.
van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B. W. J. H., Waldorp, L. J. & Schoevers, R. A. (2015). Association of symptom network structure with the course of depression. *JAMA psychiatry* **72**, 1219-1226.
van Borkulo, C. D., with contributions from Sacha, E. & Millner, A. (2016). NetworkComparisonTest: Statistical Comparison of Two Networks Based on Three Invariance Measures.
Weare, K. & Nind, M. (2011). Mental health promotion and problem prevention in schools: what does the evidence say? *Health promotion international* **26**, i29-i69.
World Health Organization (2002). Prevention and promotion in mental health.
World Health Organization (2004). Promoting mental health: Concepts, emerging evidence, practice: Summary report.

Figure 1

Click here to access/download;Figure(s);Figure 1.docx

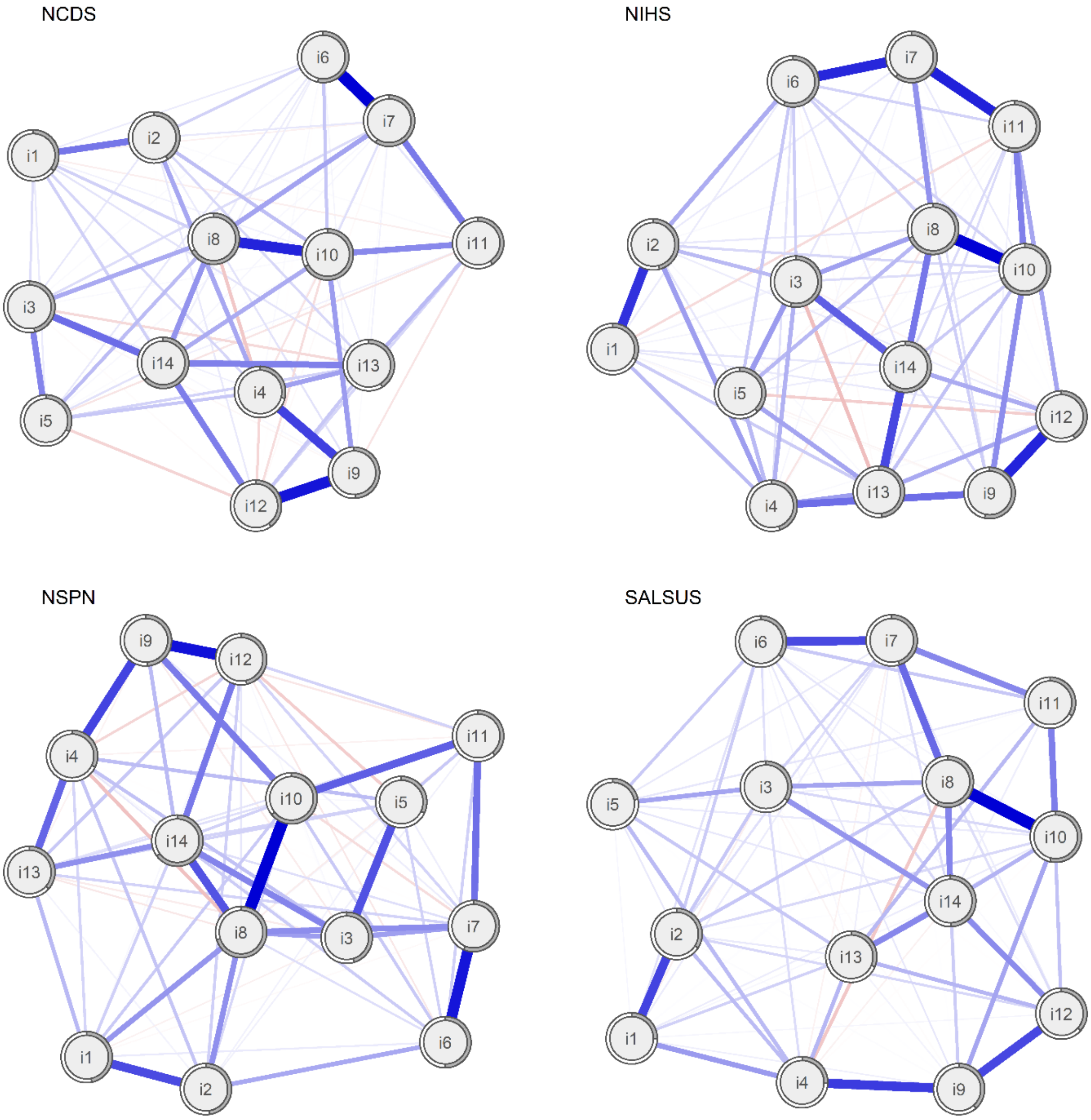


Figure 1: Networks of WEMWBS items in four general population samples. Nodes represent WEMWBS items and edges partial correlations with LASSO penalty. Distances between nodes and the thickness of edges relate to the size of their partial correlations. Grey doughnut charts surrounding each node show its explained variance.

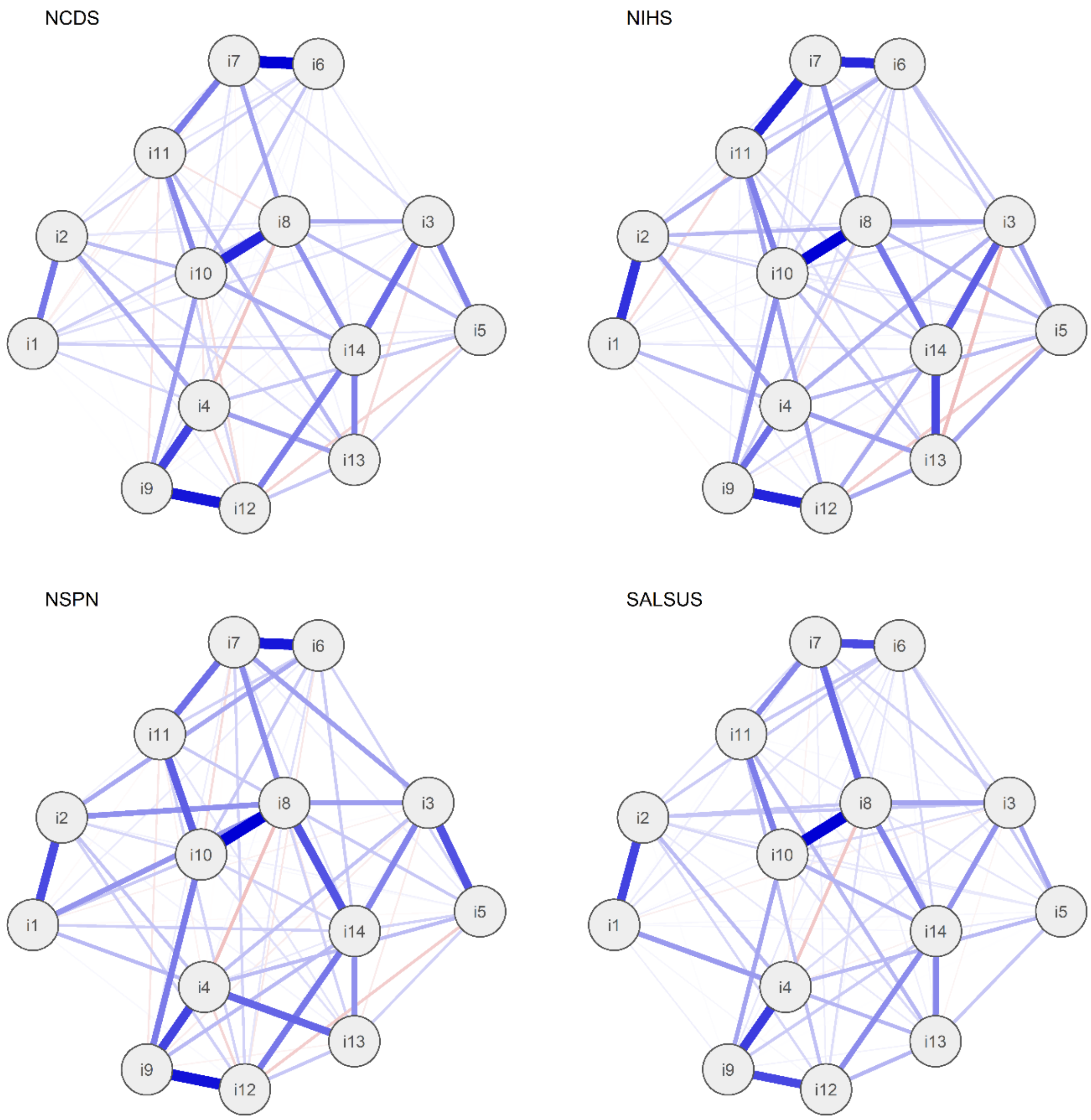


Figure 2: Networks of WEMWBS items in four general population samples using average spring layout. Nodes represent WEMWBS items and edges partial correlations with LASSO penalty. Distances among nodes and thickness of edges relate to size of their partial correlations. Grey doughnut charts surrounding each node show its explained variance.

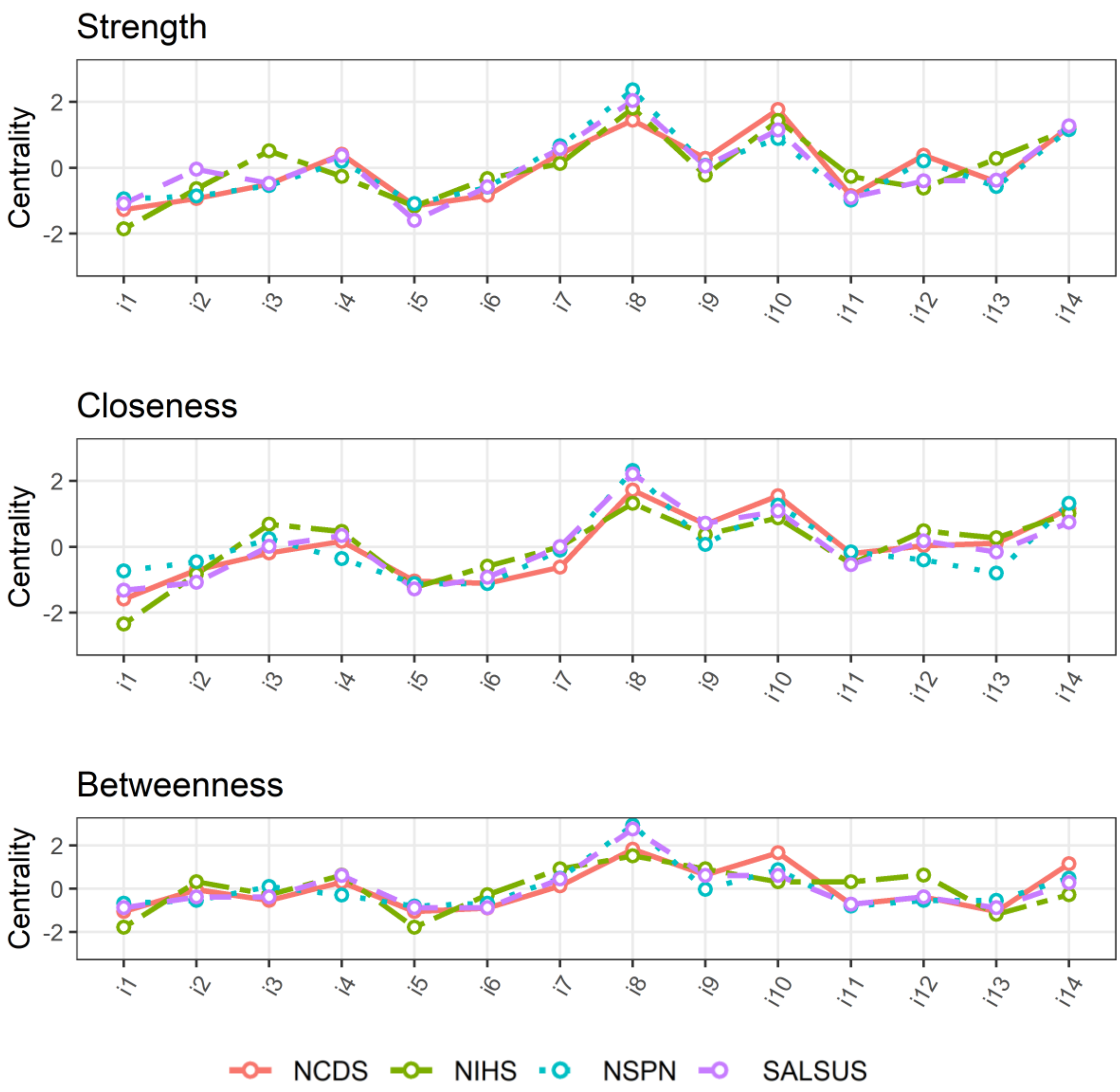


Figure 3: Centrality indices across cohorts

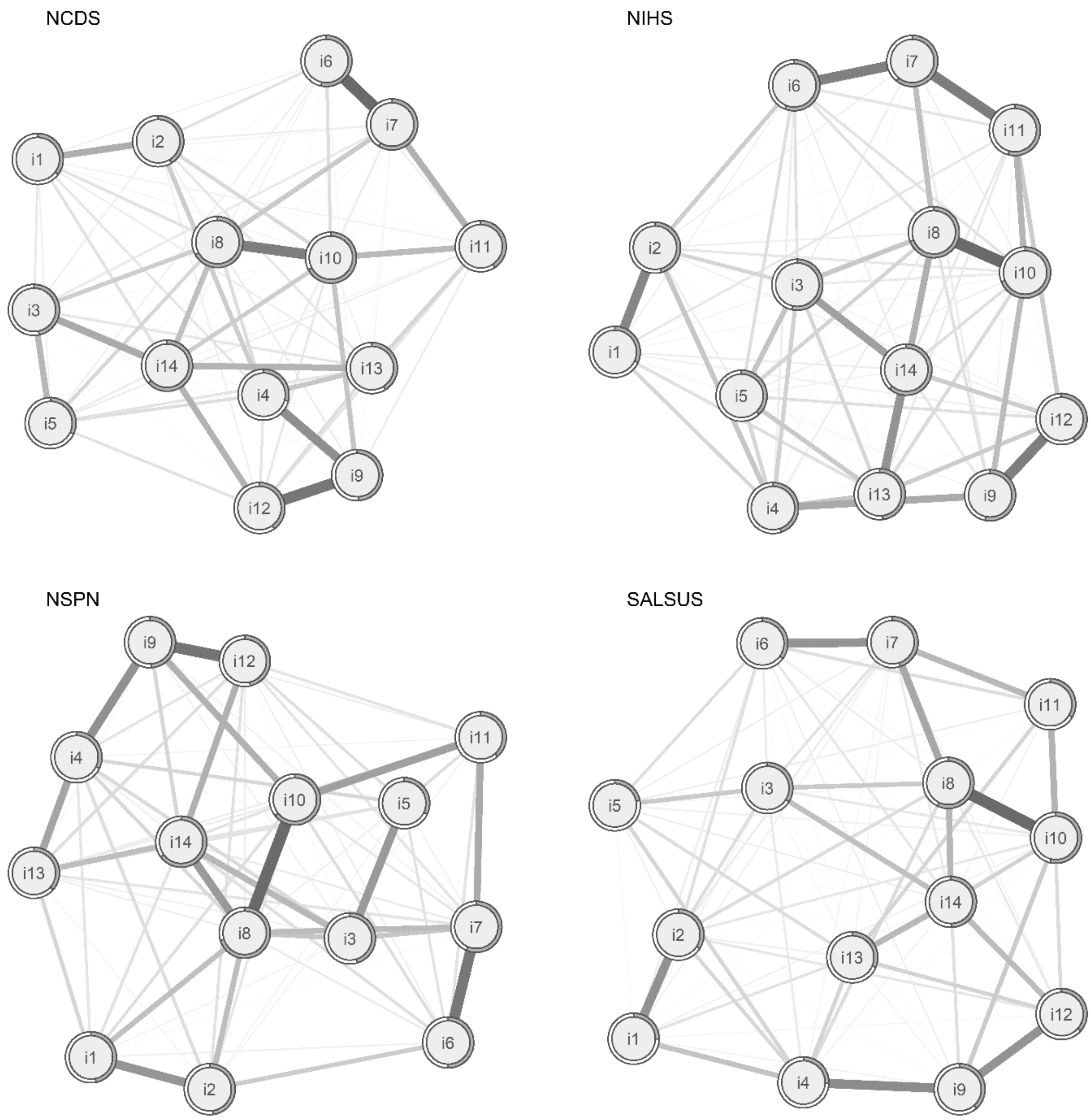


Figure 1: Networks of WEMWBS items in four general population samples. Nodes represent WEMWBS items and edges partial correlations with LASSO penalty. Distances between nodes and the thickness of edges relate to the size of their partial correlations. Grey doughnut charts surrounding each node show its explained variance.

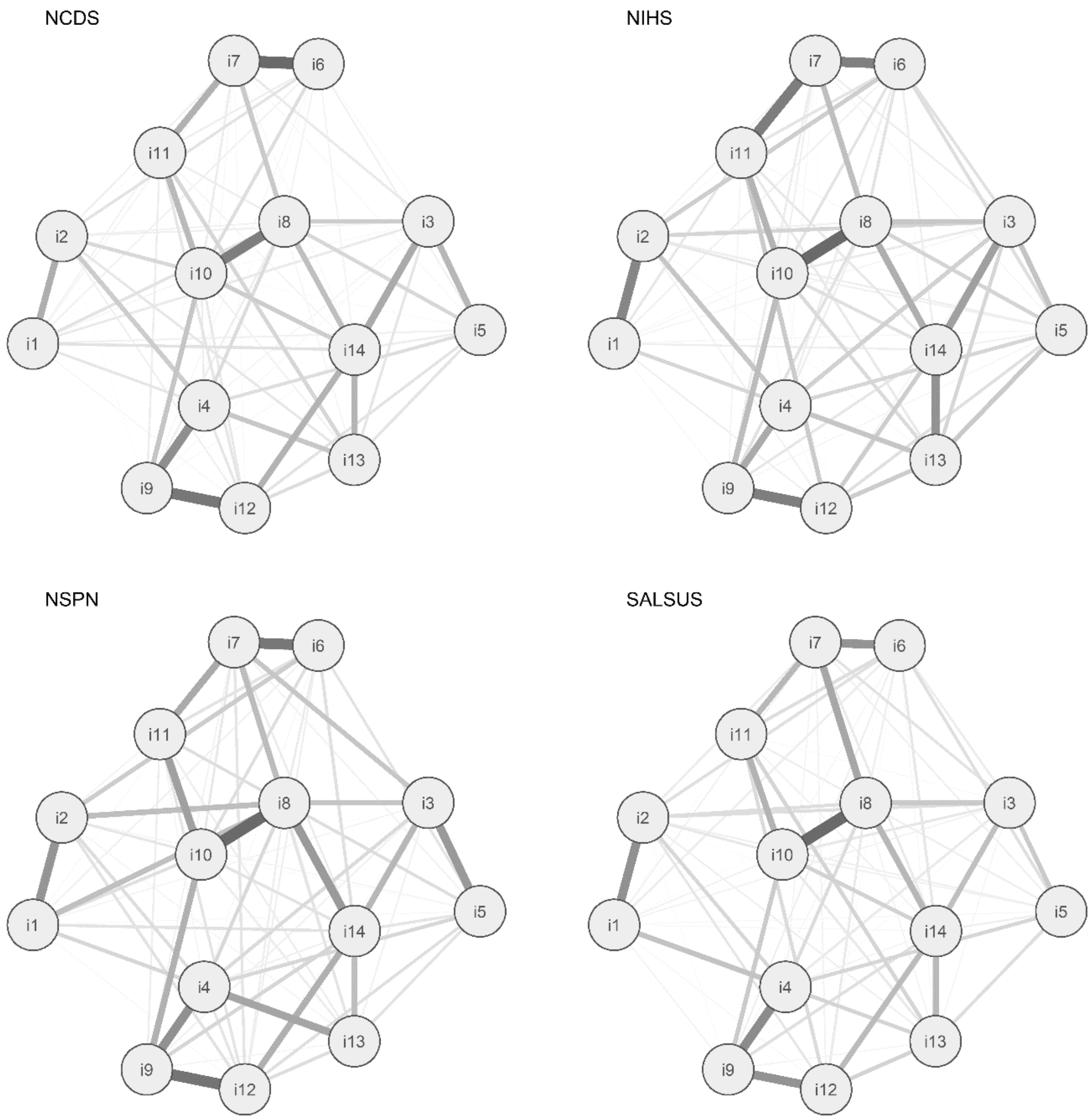


Figure 2: Networks of WEMWBS items in four general population samples using average spring layout. Nodes represent WEMWBS items and edges partial correlations with LASSO penalty. Distances among nodes and thickness of edges relate to size of their partial correlations. Grey doughnut charts surrounding each node show its explained variance.

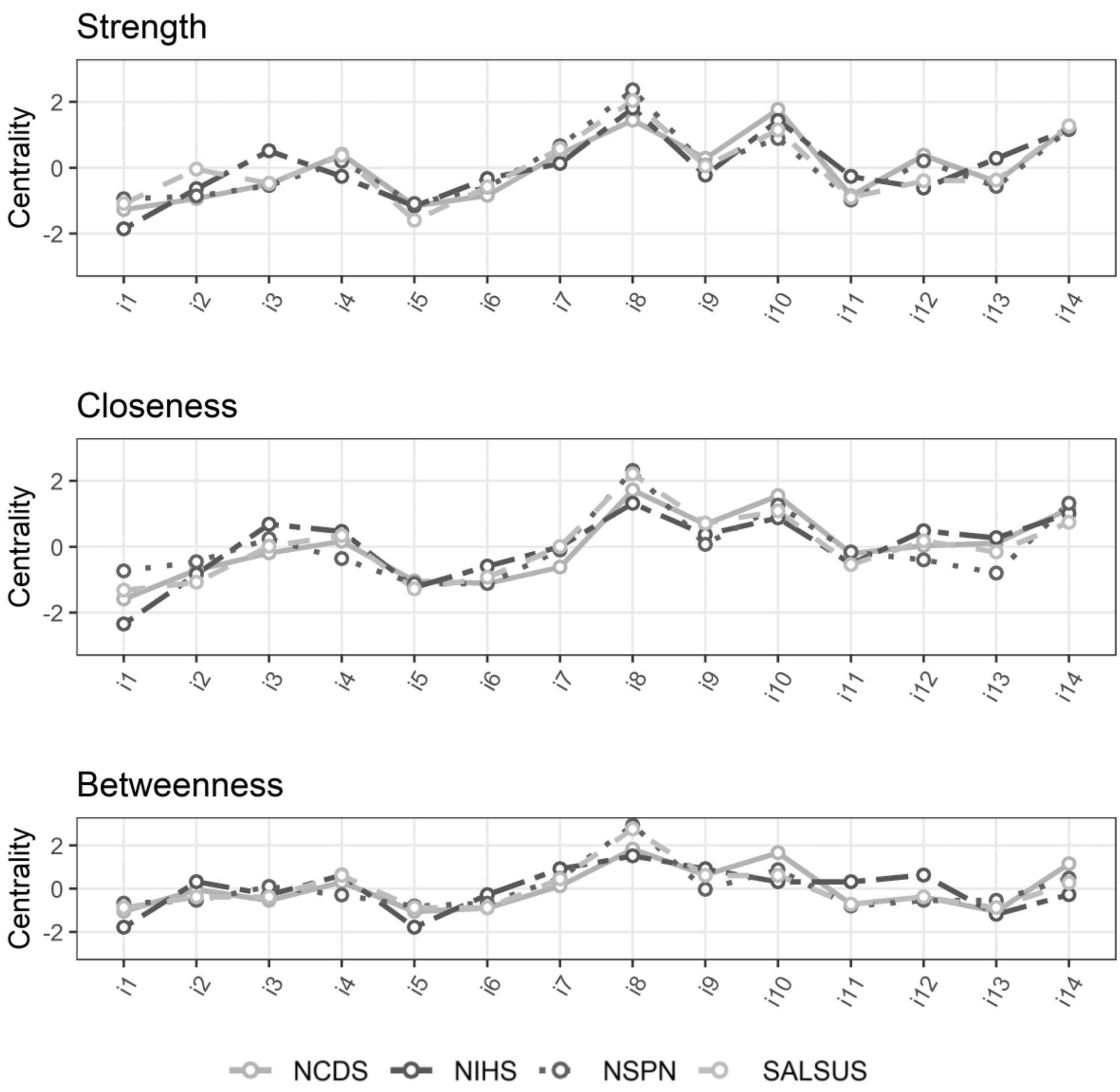


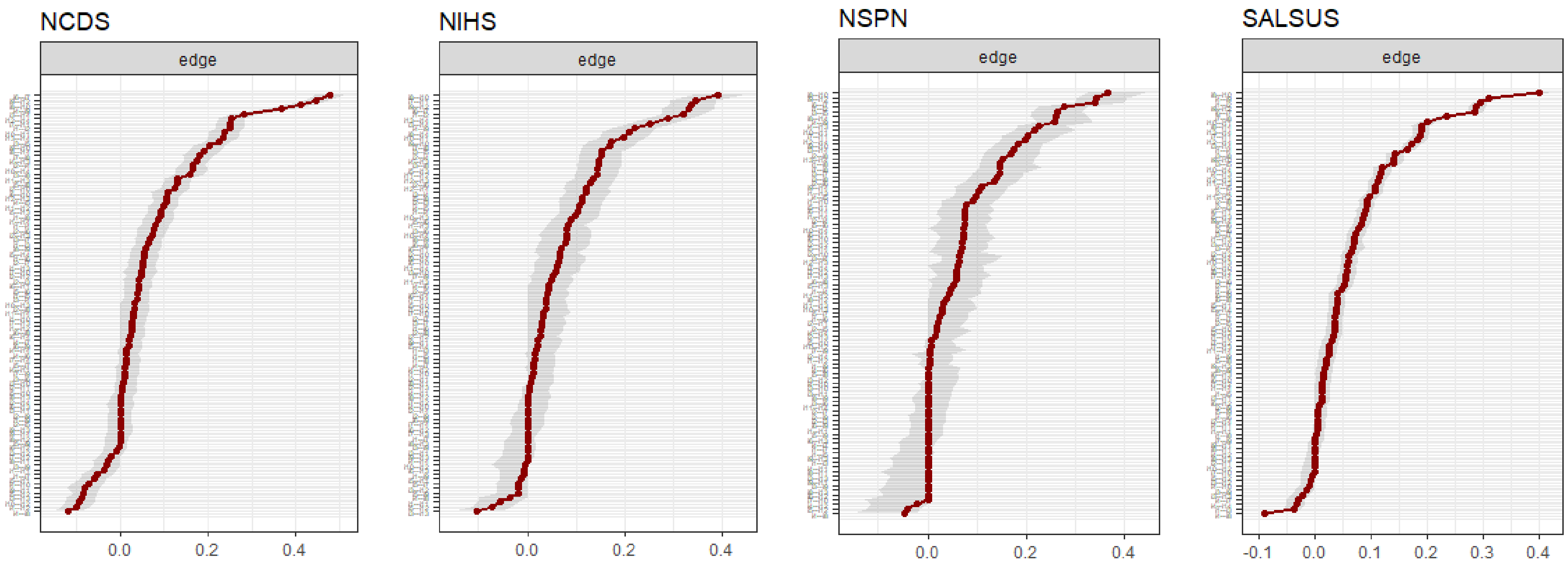
Figure 3: Centrality indices across cohorts

Table 1: WEMWBS item labels, wording, and item means (standard deviations) across samples.

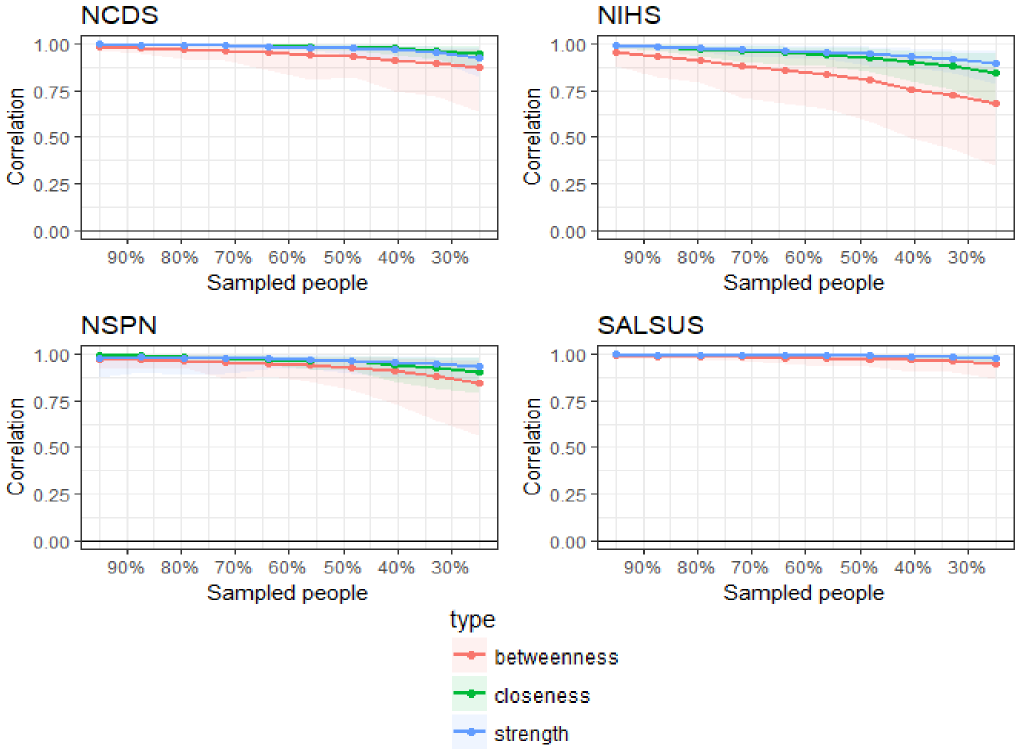
		Mean (standard deviation)			
Item	**Statement**	**NCDS**	**NIHS**	**NSPN**	**SALSUS**
i1	I have been feeling optimistic about the future	3.28 (0.87)	3.23 (1.06)	3.44 (0.98)	3.25 (1.08)
i2	I have been feeling useful	3.56 (0.80)	3.50 (0.99)	3.25 (0.92)	3.21 (0.97)
i3	I have been feeling relaxed	3.30 (0.81)	3.32 (0.96)	3.22 (0.94)	3.41 (0.98)
i4	I have been feeling interested in other people	3.54 (0.82)	3.56 (0.97)	3.56 (0.91)	3.42 (1.04)
i5	I have had energy to spare	2.81 (0.91)	2.85 (1.06)	2.94 (1.02)	3.48 (1.06)
i6	I have been dealing with problems well	3.59 (0.78)	3.59 (0.90)	3.35 (0.95)	3.46 (1.06)
i7	I have been thinking clearly	3.71 (0.75)	3.82 (0.90)	3.54 (0.94)	3.65 (1.00)
i8	I have been feeling good about myself	3.39 (0.88)	3.57 (0.96)	3.40 (1.00)	3.49 (1.08)
i9	I have been feeling close to other people	3.58 (0.84)	3.73 (0.93)	3.53 (0.99)	3.72 (1.02)
i10	I have been feeling confident	3.46 (0.88)	3.52 (0.97)	3.37 (1.02)	3.54 (1.06)
i11	I have been able to make up my own mind about things	3.96 (0.79)	4.01 (0.87)	3.63 (0.98)	4.04 (0.93)
i12	I have been feeling loved	3.91 (0.99)	4.04 (0.98)	3.77 (1.08)	3.93 (1.07)
i13	I have been interested in new things	3.60 (0.90)	3.51 (1.02)	3.68 (1.00)	3.73 (1.02)
i14	I have been feeling cheerful	3.58 (0.81)	3.63 (0.86)	3.57 (0.95)	3.79 (1.00)

Table 2: Correlation stability coefficients

	betweenness	closeness	strength
NCDS	0.67	>0.75	>0.75
NIHS	0.36	>0.75	>0.75
NSPN	0.67	>0.75	>0.75
SALSUS	>0.75	>0.75	>0.75



Supplemental Figure 1: Point estimates (red) and 95% bootstrap confidence intervals (grey) of network edges (representing partial correlations between items).



Supplemental Figure 2: Stability of centrality indices: point estimates and corresponding 95% CIs.

Supplemental Table 1: Spearman correlations between edges.

	NCDS	NIHS	NSPN	SALSUS
NCDS	1	-	-	-
NIHS	0.87	1	-	-
NSPN	0.79	0.75	1	-
SALSUS	0.80	0.82	0.83	1

```
########################################    R code used for analyses    ########################################
# The following code shows the script for our analysis. We had data from four cohorts stored in R as four separate data.frames
objects called wemwbs_ncds, wemwbs_nihs, wemwbs_nspn and wemwbs_sals. The structure of each data.frame is outlined below
and was the same across all cohorts.
# head(wemwbs_ncds)
#   ID        sex     i1  i2  i3  i4  i5  i6  i7  i8  i9  i10  i11  i12  i13  i14
#   N10001N   Female  4   3   5   5   4   5   5   5   4   5    5    2    5    4
#   N10002P   Male    4   4   4   2   4   4   4   4   3   4    5    5    4    4
#   N10007U   Female  4   4   4   4   4   5   5   5   5   5    5    5    5    5
#   N10008V   Male    4   4   4   5   5   5   5   5   4   3    5    5    5    4
#   N10009W   Male    4   4   4   4   3   4   4   4   3   4    5    4    4    4
#   N10011Q   Male    4   4   3   4   3   4   4   4   4   4    4    4    4    4

# sets working directory
setwd("D:/..../") # please insert the path to working directory where you want to save figures

# installs required packages
install.packages(c("qgpraph", "NetworkComparisonTest", "bootnet","networktools", "ggplot2", "gridExtra","EstimateGroupNetwork",
"mgm", "reshape", "lemon", "dplyr"), dependencies=T)

# loads required packages
require(qgraph); require(NetworkComparisonTest); require(bootnet); require(networktools); require(ggplot2); require(gridExtra);
require(EstimateGroupNetwork); require(mgm); require(reshape); require(lemon); require(dplyr)

# excludes individuals with missing data
wemwbs_ncds <- na.omit(wemwbs_ncds)
wemwbs_nihs  <- na.omit(wemwbs_nihs)
wemwbs_nspn <- na.omit(wemwbs_nspn)
wemwbs_sals  <- na.omit(wemwbs_sals)

# makes Table 1: WEMWBS item labels, wording, and item means (standard deviations) across samples
temp1 <- round(colMeans(wemwbs_ncds[,3:16], na.rm = TRUE),2)
temp2 <- round(colMeans(wemwbs_nihs[,3:16], na.rm = TRUE),2)
temp3 <- round(colMeans(wemwbs_nspn[,3:16], na.rm = TRUE),2)
temp4 <- round(colMeans(wemwbs_sals[,3:16], na.rm = TRUE),2)

temp1 <- paste(temp1, " (", round(apply(wemwbs_ncds[,3:16], 2, sd),2), ")", sep="")
temp2 <- paste(temp2, " (", round(apply(wemwbs_nihs[,3:16], 2, sd),2), ")", sep="")
temp3 <- paste(temp3, " (", round(apply(wemwbs_nspn[,3:16], 2, sd),2), ")", sep="")
temp4 <- paste(temp4, " (", round(apply(wemwbs_sals[,3:16], 2, sd),2), ")", sep="")

itemstats <- data.frame('Item label'= paste("i",1:14, sep=""), Statement=c(
  "I have been feeling optimistic about the future",
  "I have been feeling useful",
  "I have been feeling relaxed",
  "I have been feeling interested in other people",
  "I have had energy to spare",
  "I have been dealing with problems well",
  "I have been thinking clearly",
  "I have been feeling good about myself",
  "I have been feeling close to other people",
  "I have been feeling confident",
  "I have been able to make up my own mind about things",
  "I have been feeling loved",
  "I have been interested in new things",
  "I have been feeling cheerful"), 'NCDS'=temp1, 'NIHS'=temp2, 'NSPN'=temp3, 'SALSUS'=temp4)

itemstats

# estimates networks using mgm package and computes node predictability
temp1 <- mgm(wemwbs_ncds[,3:16], type=rep('g', 14), lev=rep(1,14), k=2)
pred_ncds <- predict(temp1, wemwbs_ncds[,3:16], error.continuous='VarExpl')
temp2 <- mgm(wemwbs_nihs[,3:16], type=rep('g', 14), lev=rep(1,14), k=2)
pred_nihs <- predict(temp2, wemwbs_nihs[,3:16], error.continuous='VarExpl')
temp3 <- mgm(wemwbs_nspn[,3:16], type=rep('g', 14), lev=rep(1,14), k=2)
pred_nspn <- predict(temp3, wemwbs_nspn[,3:16], error.continuous='VarExpl')
temp4 <- mgm(wemwbs_sals[,3:16], type=rep('g', 14), lev=rep(1,14), k=2)
pred_sals <- predict(temp4, wemwbs_sals[,3:16], error.continuous='VarExpl')

# computes fused graphical LASSO networks
```

```
groupnetwork_kfold <-
EstimateGroupNetwork(list(wemwbs_ncds[,3:16],wemwbs_nihs[,3:16],wemwbs_nspn[,3:16],wemwbs_sals[,3:16]),inputType =
"list.of.dataframes", covfun = cor_auto, method = "crossvalidation", strategy = "sequential", k = 10, seed=1234, criterion = c("ebic",
"bic", "aic"), count.unique = FALSE, optimize = TRUE, optmethod = "CG", penalty = "fused", weights = "equal", penalize.diagonal =
FALSE, ncores = 4, simplifyOutput = FALSE)

# makes Figure 1: Networks of WEMWBS items in four general population samples
png("Figure 1.png", width=10, height=10, units = "in", res = 600)
par(mfrow=c(2,2))
g1 <- qgraph(groupnetwork_kfold$network[[1]], layout = "spring",theme="colorblind", pie=pred_ncds$errors$Error.R2,
border.width=2, vsize=10, border.color='#555555', label.color="#555555", color="#EEEEEE",DoNotPlot=TRUE); plot(g1);
title("NCDS",adj=0, font.main=1, line=2.5)
g2 <- qgraph(groupnetwork_kfold$network[[2]], layout = "spring", theme="colorblind", pie=pred_nihs$errors$Error.R2,
border.width=2, vsize=10, border.color='#555555', label.color="#555555", color="#EEEEEE",DoNotPlot=TRUE); plot(g2);
title("NIHS",adj=0, font.main=1, line=2.5)
g3 <- qgraph(groupnetwork_kfold$network[[3]], layout = "spring", theme="colorblind", pie=pred_nspn$errors$Error.R2,
border.width=2, vsize=10, border.color='#555555', label.color="#555555", color="#EEEEEE",DoNotPlot=TRUE); plot(g3);
title("NSPN",adj=0, font.main=1, line=2.5)
g4 <- qgraph(groupnetwork_kfold$network[[4]], layout = "spring", theme="colorblind", pie=pred_sals$errors$Error.R2,
border.width=2, vsize=10, border.color='#555555', label.color="#555555", color="#EEEEEE",DoNotPlot=TRUE); plot(g4);
title("SALSUS",adj=0, font.main=1, line=2.5)
dev.off()


# assesses network differences across cohorts
comp_ncds_nihs <- NCT(wemwbs_ncds[,3:16], wemwbs_nihs[,3:16], it=5000, binary.data=FALSE, paired=FALSE,
weighted=TRUE, test.edges=TRUE, progressbar=TRUE, edges='all')
comp_ncds_nspn <- NCT(wemwbs_ncds[,3:16], wemwbs_nspn[,3:16], it=5000, binary.data=FALSE, paired=FALSE,
weighted=TRUE, test.edges=TRUE, progressbar=TRUE, edges='all')
comp_ncds_sals <- NCT(wemwbs_ncds[,3:16], wemwbs_sals[,3:16], it=5000, binary.data=FALSE, paired=FALSE,
weighted=TRUE, test.edges=TRUE, progressbar=TRUE, edges='all')
comp_nihs_nspn <- NCT(wemwbs_nihs[,3:16], wemwbs_nspn[,3:16], it=5000, binary.data=FALSE, paired=FALSE,
weighted=TRUE, test.edges=TRUE, progressbar=TRUE, edges='all')
comp_nihs_sals <- NCT(wemwbs_nihs[,3:16], wemwbs_sals[,3:16], it=5000, binary.data=FALSE, paired=FALSE, weighted=TRUE,
test.edges=TRUE, progressbar=TRUE, edges='all')
comp_nspn_sals <- NCT(wemwbs_nspn[,3:16], wemwbs_sals[,3:16], it=5000, binary.data=FALSE, paired=FALSE,
weighted=TRUE, test.edges=TRUE, progressbar=TRUE, edges='all')

comp_ncds_nihs$glstrinv.sep # shows network strength
comp_ncds_nihs$glstrinv.pval # p-value for global strength difference
comp_ncds_nihs$einv.pvals[comp_ncds_nihs$einv.pvals$`p-value`<0.05,] #shows which edges are statistically significant

comp_ncds_nspn$glstrinv.sep
comp_ncds_nspn$glstrinv.pval
comp_ncds_nspn$einv.pvals[comp_ncds_nspn$einv.pvals$`p-value`<0.05,]

comp_ncds_sals$glstrinv.sep
comp_ncds_sals$glstrinv.pval
comp_ncds_sals$einv.pvals[comp_ncds_sals$einv.pvals$`p-value`<0.05,]

comp_nihs_nspn$glstrinv.sep
comp_nihs_nspn$glstrinv.pval
comp_nihs_nspn$einv.pvals[comp_nihs_nspn$einv.pvals$`p-value`<0.05,]

comp_nihs_sals$glstrinv.sep
comp_nihs_sals$glstrinv.pval
comp_nihs_sals$einv.pvals[comp_nihs_sals$einv.pvals$`p-value`<0.05,]

comp_nspn_sals$glstrinv.sep
comp_nspn_sals$glstrinv.pval
comp_nspn_sals$einv.pvals[comp_nspn_sals$einv.pvals$`p-value`<0.05,]


# computes average layout
Layout <-
averageLayout(groupnetwork_kfold$network[[1]],groupnetwork_kfold$network[[2]],groupnetwork_kfold$network[[3]],groupnetwork_k
fold$network[[4]])

# makes Figure 2: Networks of WEMWBS items in four general population samples using average spring layout
png("Figure 2.png", width=10, height=10, units = "in", res = 600)
```

```
par(mfrow=c(2,2))
g1 <- qgraph(groupnetwork_kfold$network[[1]], layout = Layout, theme="colorblind", pie=pred_ncds$errors$Error.R2,
border.width=2, vsize=10, border.color='#555555', label.color="#555555", color="#EEEEEE",DoNotPlot=TRUE); plot(g1);
title("NCDS",adj=0, font.main=1, line=2.5)
g2 <- qgraph(groupnetwork_kfold$network[[2]], layout = Layout, theme="colorblind", pie=pred_nihs$errors$Error.R2,
border.width=2, vsize=10, border.color='#555555', label.color="#555555", color="#EEEEEE",DoNotPlot=TRUE); plot(g2);
title("NIHS",adj=0, font.main=1, line=2.5)
g3 <- qgraph(groupnetwork_kfold$network[[3]], layout = Layout, theme="colorblind", pie=pred_nspn$errors$Error.R2,
border.width=2, vsize=10, border.color='#555555', label.color="#555555", color="#EEEEEE",DoNotPlot=TRUE); plot(g3);
title("NSPN",adj=0, font.main=1, line=2.5)
g4 <- qgraph(groupnetwork_kfold$network[[4]], layout = Layout, theme="colorblind", pie=pred_sals$errors$Error.R2,
border.width=2, vsize=10, border.color='#555555', label.color="#555555", color="#EEEEEE",DoNotPlot=TRUE); plot(g4);
title("SALSUS",adj=0, font.main=1, line=2.5)
dev.off()

# bootstraps networks
set.seed("12345")

boot_networklasso_ncds <-   bootnet(wemwbs_ncds[,3:16], nBoots = 2500, default = "EBICglasso", type = "nonparametric", nCores
= 4, verbose = TRUE, computeCentrality =TRUE,lambda.min.ratio=0.001)

boot_networklasso_nihs <-   bootnet(wemwbs_nihs[,3:16], nBoots = 2500, default = "EBICglasso", type = "nonparametric", nCores =
4, verbose = TRUE, computeCentrality =TRUE,lambda.min.ratio=0.001)

boot_networklasso_nspn <-   bootnet(wemwbs_nspn[,3:16], nBoots = 2500, default = "EBICglasso", type = "nonparametric", nCores
= 4, verbose = TRUE, computeCentrality =TRUE,lambda.min.ratio=0.001)

boot_networklasso_sals <-   bootnet(wemwbs_sals[,3:16], nBoots = 2500, default = "EBICglasso", type = "nonparametric", nCores =
4, verbose = TRUE, computeCentrality =TRUE,lambda.min.ratio=0.001)


# makes Supplementary Figure 1: Point estimates (red) and 95% bootstrap confidence intervals (grey) of network edges
(representing partial correlations between items)
png("Supplementary Figure 1.png", width=10, height=10, units = "in", res = 600)
p1 <- plot(boot_networklasso_ncds, statistics=c("edge"), plot="area", CIstyle="quantiles", order="sample", legend=FALSE) +
ggtitle("NCDS") + theme(axis.text.y = element_text(size=4))
p2 <- plot(boot_networklasso_nihs, statistics=c("edge"), plot="area", CIstyle="quantiles", order="sample", legend=FALSE) +
ggtitle("NIHS") + theme(axis.text.y = element_text(size=4))
p3 <- plot(boot_networklasso_nspn, statistics=c("edge"), plot="area", CIstyle="quantiles", order="sample", legend=FALSE) +
ggtitle("NSPN") + theme(axis.text.y = element_text(size=4))
p4 <- plot(boot_networklasso_sals, statistics=c("edge"), plot="area", CIstyle="quantiles", order="sample", legend=FALSE) +
ggtitle("SALSUS") + theme(axis.text.y = element_text(size=4))

grid.arrange(p1, p2, p3, p4, nrow =1)
dev.off()

# correlations presented in Supplementary Table 1: Spearman correlations between edges
cor(getWmat(g1)[lower.tri(getWmat(g1))], getWmat(g2)[lower.tri(getWmat(g2))],   method="spearman") #0.87
cor(getWmat(g1)[lower.tri(getWmat(g1))], getWmat(g3)[lower.tri(getWmat(g3))], method="spearman") #0.79
cor(getWmat(g1)[lower.tri(getWmat(g1))], getWmat(g4)[lower.tri(getWmat(g4))], method="spearman") #0.80
cor(getWmat(g2)[lower.tri(getWmat(g2))], getWmat(g3)[lower.tri(getWmat(g3))], method="spearman") #0.75
cor(getWmat(g2)[lower.tri(getWmat(g2))], getWmat(g4)[lower.tri(getWmat(g4))], method="spearman") #0.82
cor(getWmat(g3)[lower.tri(getWmat(g3))], getWmat(g4)[lower.tri(getWmat(g4))], method="spearman") #0.83
mean(c(0.87,0.79,0.80,0.75,0.82,0.83)) #computes mean correlation (=0.81)

# makes Figure 3: Centrality indices across cohorts
strength <- as.data.frame(cbind(scale(centrality(g1)$InDegree), scale(centrality(g2)$InDegree), scale(centrality(g3)$InDegree),
scale(centrality(g4)$InDegree)))

closeness <- as.data.frame(cbind(scale(centrality(g1)$Closeness), scale(centrality(g2)$Closeness), scale(centrality(g3)$Closeness),
scale(centrality(g4)$Closeness)))

betweenness <- as.data.frame(cbind(scale(centrality(g1)$Betweenness), scale(centrality(g2)$Betweenness),
scale(centrality(g3)$Betweenness), scale(centrality(g4)$Betweenness)))

strength <- mutate(strength, id = rownames(strength))
closeness <- mutate(closeness, id = rownames(closeness))
betweenness <- mutate(betweenness, id = rownames(betweenness))

colnames(strength)<-c("NCDS", "NIHS", "NSPN", "SALSUS", "Symptoms")
```

```
colnames(closeness)<-c("NCDS", "NIHS", "NSPN", "SALSUS", "Symptoms")
colnames(betweenness)<-c("NCDS", "NIHS", "NSPN", "SALSUS", "Symptoms")

strength_long <- melt(strength, id="Symptoms")
strength_long$Symptoms <- rep(1:14,4)
names(strength_long)[2] <- "Cohorts"

closeness_long <- melt(closeness, id="Symptoms")
closeness_long$Symptoms <- rep(1:14,4)
names(closeness_long)[2] <- "Cohorts"

betweenness_long <- melt(betweenness, id="Symptoms")
betweenness_long$Symptoms <- rep(1:14,4)
names(betweenness_long)[2] <- "Cohorts"

png("Figure 3.png", width=6, height=6, units = "in", res = 600)
p5 <- ggplot(data=strength_long, aes(x=Symptoms, y=value, colour=Cohorts)) + geom_line(size=1, aes(linetype=Cohorts)) +
geom_point(shape = 21, fill = "white", size = 1.5, stroke = 1) + xlab(" ") + ylab("Centrality") + scale_y_continuous(limits = c(-3, 3)) +
scale_x_continuous(breaks=c(1:14),labels=strength$Symptoms) + theme_bw() + theme(panel.grid.minor=element_blank(),
axis.text.x = element_text(angle = 60, hjust = 1),legend.position="none") + ggtitle("Strength") +
scale_linetype_manual(values=c("solid", "twodash", "dotted", "dashed"))

p6 <- ggplot(data=closeness_long, aes(x=Symptoms, y=value, colour=Cohorts)) + geom_line(size=1, aes(linetype=Cohorts)) +
geom_point(shape = 21, fill = "white", size = 1.5, stroke = 1) + xlab(" ") + ylab("Centrality") + scale_y_continuous(limits = c(-3, 3)) +
scale_x_continuous(breaks=c(1:14),labels=closeness$Symptoms) + theme_bw() + theme(panel.grid.minor=element_blank(),
axis.text.x = element_text(angle = 60, hjust = 1),legend.position="none") + ggtitle("Closeness") +
scale_linetype_manual(values=c("solid", "twodash", "dotted", "dashed"))

p7 <- ggplot(data=betweenness_long, aes(x=Symptoms, y=value, colour=Cohorts)) + geom_line(size=1, aes(linetype=Cohorts)) +
geom_point(shape = 21, fill = "white", size = 1.5, stroke = 1) + xlab(" ") + ylab("Centrality") + scale_y_continuous(limits = c(-3, 3)) +
scale_x_continuous(breaks=c(1:14),labels=betweenness$Symptoms) + theme_bw() + theme(panel.grid.minor=element_blank(),
axis.text.x = element_text(angle = 60, hjust = 1),legend.position="none") + ggtitle("Betweenness") +
scale_linetype_manual(values=c("solid", "twodash", "dotted", "dashed"))

p7 <- grid_arrange_shared_legend(p7, position='bottom', plot=FALSE)

grid.arrange(p5, p6, p7, nrow =3)
dev.off()

# case dropping bootstrap
set.seed("12345")

boot_networklasso_centrality_ncds <-   bootnet(wemwbs_ncds[,3:16], nBoots = 2500, default = "EBICglasso", type = "case", nCores
= 4, statistics = c("strength","closeness","betweenness"), model = "GGM", verbose = TRUE, computeCentrality =
TRUE,lambda.min.ratio=0.001)
boot_networklasso_centrality_nihs <-   bootnet(wemwbs_nihs[,3:16], nBoots = 2500, default = "EBICglasso", type = "case", nCores
= 4, statistics = c("strength","closeness","betweenness"), model = "GGM", verbose = TRUE, computeCentrality =
TRUE,lambda.min.ratio=0.001)
boot_networklasso_centrality_nspn <-   bootnet(wemwbs_nspn[,3:16], nBoots = 2500, default = "EBICglasso", type = "case",
nCores = 4, statistics = c("strength","closeness","betweenness"), model = "GGM", verbose = TRUE, computeCentrality =
TRUE,lambda.min.ratio=0.001)
boot_networklasso_centrality_sals <-   bootnet(wemwbs_sals[,3:16], nBoots = 2500, default = "EBICglasso", type = "case", nCores
= 4, statistics = c("strength","closeness","betweenness"), model = "GGM", verbose = TRUE, computeCentrality =
TRUE,lambda.min.ratio=0.001)

# makes Supplementary Figure 2: Stability of centrality indices: point estimates and corresponding 95% CIs
g_legend<-function(a.gplot){
  tmp <- ggplot_gtable(ggplot_build(a.gplot))
  leg <- which(sapply(tmp$grobs, function(x) x$name) == "guide-box")
  legend <- tmp$grobs[[leg]]
  return(legend)}

p1forlegend <- plot(boot_networklasso_centrality_ncds) + ylab("Correlation")
mylegend<-g_legend(p1forlegend)

png("Supplementary Figure 2.png", width=7, height=7, units = "in", res = 600)
p1 <- plot(boot_networklasso_centrality_ncds) + ylab("Correlation") + theme(legend.position="none") + ggtitle("NCDS") +
scale_y_continuous(limits = c(0, 1))
p2 <- plot(boot_networklasso_centrality_nihs) + ylab("Correlation") + theme(legend.position="none") + ggtitle("NIHS") +
scale_y_continuous(limits = c(0, 1))
```

```
p3 <- plot(boot_networklasso_centrality_nspn) + ylab("Correlation") + theme(legend.position="none") + ggtitle("NSPN") +
scale_y_continuous(limits = c(0, 1))
p4 <- plot(boot_networklasso_centrality_sals) + ylab("Correlation") + theme(legend.position="none") + ggtitle("SALSUS") +
scale_y_continuous(limits = c(0, 1))

grid.arrange(p1, p2, p3, p4, mylegend, nrow =3)
dev.off()

# computes cs coefficients
CS_ncds <- corStability(boot_networklasso_centrality_ncds)
CS_nihs  <- corStability(boot_networklasso_centrality_nihs)
CS_nspn <- corStability(boot_networklasso_centrality_nspn)
CS_sals  <- corStability(boot_networklasso_centrality_sals)

# computes values in Table 2: Correlation stability coefficients
CSfinal <- rbind(CS_ncds,CS_nihs,CS_nspn,CS_sals)
rownames(CSfinal) <- c("NCDS","NIHS", "NSPN", "SALSUS")
CSfinal


# assesses gender differences
gendercomparison_ncds <- NCT(wemwbs_ncds[wemwbs_ncds$sex=="Male",3:16],
wemwbs_ncds[wemwbs_ncds$sex=="Female",3:16], it=5000, binary.data=FALSE, paired=FALSE, weighted=TRUE,
test.edges=TRUE, edges='all', progressbar=TRUE)
gendercomparison_ncds$glstrinv.pval # p-value for global strength difference
gendercomparison_ncds$einv.pvals$Var1[gendercomparison_ncds$einv.pvals$`p-value`<0.05] #which items are involved in
significant differences
gendercomparison_ncds$einv.pvals$Var2[gendercomparison_ncds$einv.pvals$`p-value`<0.05] #which items are involved in
significant differences
gendercomparison_ncds$einv.pvals$`p-value`[gendercomparison_ncds$einv.pvals$`p-value`<0.05] #p-value

gendercomparison_nihs <- NCT(wemwbs_nihs[wemwbs_nihs$sex=="Male",3:16],
wemwbs_nihs[wemwbs_nihs$sex=="Female",3:16], it=5000, binary.data=FALSE, paired=FALSE, weighted=TRUE,
test.edges=TRUE, edges='all', progressbar=TRUE)
gendercomparison_nihs$glstrinv.pval
gendercomparison_nihs$einv.pvals$Var1[gendercomparison_nihs$einv.pvals$`p-value`<0.05]
gendercomparison_nihs$einv.pvals$Var2[gendercomparison_nihs$einv.pvals$`p-value`<0.05]
gendercomparison_nihs$einv.pvals$`p-value`[gendercomparison_nihs$einv.pvals$`p-value`<0.05]

gendercomparison_nspn <- NCT(wemwbs_nspn[wemwbs_nspn$sex=="Male",3:16],
wemwbs_nspn[wemwbs_nspn$sex=="Female",3:16], it=5000, binary.data=FALSE, paired=FALSE, weighted=TRUE,
test.edges=TRUE, edges='all', progressbar=TRUE)
gendercomparison_nspn$glstrinv.pval
gendercomparison_nspn$einv.pvals$Var1[gendercomparison_nspn$einv.pvals$`p-value`<0.05]
gendercomparison_nspn$einv.pvals$Var2[gendercomparison_nspn$einv.pvals$`p-value`<0.05]
gendercomparison_nspn$einv.pvals$`p-value`[gendercomparison_nspn$einv.pvals$`p-value`<0.05]

gendercomparison_sals <- NCT(wemwbs_sals[wemwbs_sals$sex=="Male",3:16],
wemwbs_sals[wemwbs_sals$sex=="Female",3:16], it=5000, binary.data=FALSE, paired=FALSE, weighted=TRUE,
test.edges=TRUE, edges='all', progressbar=TRUE)
gendercomparison_sals$glstrinv.pval
gendercomparison_sals$einv.pvals$Var1[gendercomparison_sals$einv.pvals$`p-value`<0.05]
gendercomparison_sals$einv.pvals$Var2[gendercomparison_sals$einv.pvals$`p-value`<0.05]
gendercomparison_sals$einv.pvals$`p-value`[gendercomparison_sals$einv.pvals$`p-value`<0.05]
```