Original Article

Toward incorporating genetic risk scores into symptom networks of psychosis

Adela-Maria Isvoranu[1], Sinan Guloksuz[2,3], Sacha Epskamp[1], Jim van Os[4], Denny Borsboom[1] and GROUP Investigators[†*]

[1]Department of Psychology, Psychological Methods, University of Amsterdam, Amsterdam, The Netherlands; [2]Department of Psychiatry and Neuropsychology, School of Mental Health and Neuroscience, Maastricht University Medical Center, Maastricht, The Netherlands; [3]Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA and [4]Utrecht University Medical Centre, Utrecht, The Netherlands

Abstract

Background. Psychosis spectrum disorder is a heterogeneous, multifactorial clinical phenotype, known to have a high heritability, only a minor portion of which can be explained by molecular measures of genetic variation. This study proposes that the identification of genetic variation underlying psychotic disorder may have suffered due to issues in the psychometric conceptualization of the phenotype. Here we aim to open a new line of research into the genetics of mental disorders by explicitly incorporating genes into symptom networks. Specifically, we investigate whether links between a polygenic risk score (PRS) for schizophrenia and measures of psychosis proneness can be identified in a network model.

Methods. We analyzed data from $n = 2180$ subjects (controls, patients diagnosed with a non-affective psychotic disorder, and the first-degree relatives of the patients). A network structure was computed to examine associations between the 42 symptoms of the Community Assessment of Psychic Experiences (CAPE) and the PRS for schizophrenia.

Results. The resulting network shows that the PRS is directly connected to the spectrum of positive and depressive symptoms, with the items *conspiracy* and *no future* being more often located on predictive pathways from PRS to other symptoms.

Conclusions. To our knowledge, the current exploratory study provides a first application of the network framework to the field of behavior genetics research. This allows for a novel outlook on the investigation of the relations between genome-wide association study-based PRSs and symptoms of mental disorders, by focusing on the dependencies among variables.

Introduction

Psychosis spectrum disorder is a potentially severe, heterogeneous, and multifactorial disorder (Keshavan *et al.*, 2011; Guloksuz and van Os, 2017). Although twin and family studies indicate substantial heritability, few genetic variants are consistently linked to psychotic disorder (Pardiñas *et al.*, 2018). The difficulty in identifying genes specific to psychosis has been explained in several ways (Sullivan, 2008). One possibility that has received scant attention is that the conceptualization of the phenotype may be suboptimal: typically, genetic studies use *symptom counts* (e.g. total scores defined on questionnaire data) or case–control designs that define cases and controls as polythetic functions of symptom data (i.e. the definition of cases corresponds to many distinct symptom profiles), thus defining the phenotype in a highly simplified fashion. Essentially, these approaches assume that such compound scores are estimates of a single underlying dimension (e.g. a *liability spectrum*) that partly stands under genetic control (Franić *et al.*, 2013). Additionally, genetic research uses clear-cut diagnostic categories, even though dominant liability spectrum theories are more consistent with a

CAMBRIDGE
UNIVERSITY PRESS

spectrum of related phenotypes (van Os and Linscott, 2012), which feature no single set of necessary and sufficient properties (e.g. a *disease entity* or *common pathogenic pathway*).

Several recent papers have argued that this conceptualization of mental disorders may be too simplistic. Instead of reflecting a single disease entity, pathogenic pathway, or unidimensional liability spectrum, disorders may result from causal interactions between symptoms (e.g. delusions→paranoia→social isolation), which constitute a network of symptoms and other components (Cramer *et al.*, 2010; Borsboom and Cramer, 2013; Borsboom, 2017). Within this *network conceptualization*, the focus shifts from investigating mental disorders – such as schizophrenia – as *disease entities* to investigating interactions between symptoms (Borsboom and Cramer, 2013; Borsboom, 2017). Thus, mental disorders do not result from one central dysfunction (e.g. a brain dysfunction) that causally produces symptoms, but from a complex interplay between symptoms, psychological, biological, and sociological components (Epskamp, 2017; Isvoranu *et al.*, 2019).

For psychosis spectrum disorders, which feature a clear genetic signal, this raises the question of how genetic factors could relate to symptom–symptom interactions in a network structure. In our view, genetic risk could plausibly influence symptom networks in two ways. First, genetic risk could act as a *direct main effect* on symptoms. In this case, genetic risk would be conductive to the *liability* to develop *certain symptoms* of schizophrenia; this liability could be thought of in terms of genetic make-up laying the foundations for symptoms to develop. For instance, one's genetic make-up could have an effect on the likelihood that the symptom 'hallucinations' occurs, so that higher genetic risk influences the liability of developing the symptom, which in turn would activate other neighboring symptoms in a network. Second, genetic risk could function as a *moderator* of the network structure: genetic risk may involve genetic factors that increase the *strength* of a causal connection between symptoms. For instance, genetics may predispose a person to experience (more) anxiety in response to hallucinations. In this scenario, genetic factors control part of the *structure* of a network, such that a genetic risk would be expressed as the likelihood that one symptom activates another. Analyses that would optimally represent the moderation hypothesis are not yet fully developed, and therefore the current manuscript is focused on the first approach.

To the extent that the above scenarios are plausible, current genetic studies using total scores defined by these symptoms may rest on an inadequate definition of the phenotype. In particular, genetic markers or polygenic risk scores (PRSs) may be more profitably analyzed in relation to a network phenotype, instead of to a total score or case–control status. The current paper aims to introduce methodology suited to incorporate genetic risk scores – in particular, genome-wide association study (GWAS)-based PRSs for schizophrenia – into symptom networks, thus opening new lines of research into the genetics of mental disorders. PRS is commonly used to summarize genetic effects and represents the sum of trait-associated alleles across many genetic loci, weighted by effect sizes estimated from GWAS (Euesden *et al.*, 2015). PRS is generally used to identify genetic basis of phenotypes and to construct risk prediction models (Dudbridge, 2013). Specifically, here we investigate whether links between PRS for schizophrenia and measures of psychosis proneness can be identified in a network model, by including the PRS as a variable in the symptom network for psychosis spectrum disorder.

Method

Participants

The sample analyzed was part of the longitudinal observational cohort study 'Genetic Risk and Outcome of Psychosis Project' (GROUP), release database 7.0 (Korver *et al.*, 2012). At baseline, the GROUP sample consisted of 1119 patients diagnosed with a non-affective psychotic disorder, 1059 siblings of these patients, 920 parents, and 586 unrelated healthy controls. The patients were recruited from mental health care institutions across The Netherlands and Belgium and the control subjects were recruited through random mailing. Inclusion criteria for patients were age 16–50 years, Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2000) criteria for a non-affective psychotic disorder, maximum duration of illness of 10 years, and estimated level of intelligence quotient (IQ) above 70. For full details on the GROUP sample please refer to the cited paper (Korver *et al.*, 2012).

The current study used baseline data, restricted to the European white ethnic group, as there is evidence for a differential impact of PRS in different ethnic groups (Marden *et al.*, 2014; van Os *et al.*, 2017). Only subjects for which a PRS was available were included in the analysis. Overall, a baseline sample of $n = 2180$ individuals was analyzed here (patients, siblings, parents, and controls).

Symptomatology

We used baseline data from the Community Assessment of Psychic Experiences (CAPE; Konings *et al.*, 2006), a self-report measure of lifetime psychotic experiences. All items, measuring frequency of positive (20 items), negative (14 items), and depressive symptoms (8 items) were included in the analyses and were scored on a 4-point Likert scale.

Genotyping, imputation, and PRS

A standard procedure for genotyping, imputation, and PRS was used. For a detailed procedure refer to online Supplementary Appendix S1.

Network construction

We fitted a Gaussian graphical model (GGM; Lauritzen and Wermuth, 1989) to the data (i.e. an undirected network). All items of the CAPE and the PRS were represented as *nodes*. An *edge* between any two nodes indicates a partial correlation between the two variables, after conditioning on all other variables in the dataset. To account for the ordinal nature of the CAPE data, we used Spearman correlations when estimating the network structure, as recommended by Epskamp and Fried (2018). Due to the high sample size, we used unregularized model selection rather than regularization techniques commonly used in estimating GGMs[1] (Williams and Rast, 2018). All analyses were performed using R-statistical software, version 3.4.3 (R Development Core Team, 2014). The networks were constructed and visualized using the R-package *qgraph* version 1.4.5 (Epskamp *et al.*, 2012).

Blue (red) edges in the network indicate positive (negative) partial correlations, and the wider and more saturated the edge, the stronger the partial correlation (Epskamp *et al.*, 2012). Because the current paper specifically focuses on edges between

the PRS and symptoms, edges representing associations between PRS and symptoms have been manually unfaded.[2] For constructing the layout of the network, we used a manually specified layout where the PRS is positioned at the center of the network, while allowing the rest of the nodes in the network to cluster based on their associations.

To display how the connectivity of edges is related across genetic and symptomatic levels of analysis, we computed a *predictive path diagram* (Fig. 2). A *predictive path diagram* takes one node (in our case PRS) as a source node and then lists (i) the nodes that have an edge with PRS; (ii) the nodes to which the nodes that have an edge with PRS are connected; (iii) the nodes to which the latter are connected – and so forth. Within the so constructed diagram, the *shortest predictive pathway* can be defined as the pathway that results from traveling the edges that, in each step, maximize the quality of the prediction of the target node from the source node, while controlling for all other variables in the network. To investigate which nodes more often lie on the shortest predictive pathways from PRS to other nodes we computed node-specific predictive betweenness as a centrality measure (Fig. 3). Because betweenness is generally not a very stable centrality measure (Epskamp *et al.*, 2018), we used both nonparametric and case-drop bootstraps to investigate the extent of its variability (Epskamp *et al.*, 2017).

Network stability

As recommended in the literature (Borsboom *et al.*, 2018), to investigate robustness and replicability of results we performed accuracy and stability checks using the R package *bootnet* (Epskamp *et al.*, 2017). In addition, we included a stability analysis to investigate the average node-specific predictive betweenness under case-dropping. For these results we refer the reader to online Supplementary Appendix S2.

Results

The study sample at baseline consisted of 2180 subjects (335 control subjects, 640 siblings, 630 parents, and 575 patients), of which 47% female and 53% male. Within the patient sample, the main diagnosis was *schizophrenia, paranoid type* (53%), followed by *psychotic disorder NOS* (11%), *schizoaffective disorder* (10%) and *schizophreniform disorder* (6%). The mean age of the subjects was 35.97 (14.61) years and the mean IQ was 103.48 (16.17). In the current analyses a GWAS *p* value threshold of 0.05 was used. In total, 194 665 single-nucleotide polymorphisms (SNPs) were included in the PRS.

Network analysis

Figure 1 presents the network depicting positive, negative, and depressive symptoms of the CAPE, as well as the PRS. Results show that the PRS is directly connected to several individual symptoms of the CAPE, but especially related to the spectrum of positive psychotic symptoms. Network analysis identified relations between the PRS and the positive psychotic symptoms C10 (conspiracy), C30 (thought echo), and C41 (Capgras). In addition, we identified positive relations between the PRS and the depressive symptom C12 (no future).[3] Stability analyses show that the network and identified edges are generally stable (see online Supplementary Appendix S2).

Second, the predictive path diagram (Fig. 2) highlights secondary levels of connectivity within the network – level two connections show strong relations between symptoms C10 (conspiracy), C7 (persecution) and C41 (Capgras), as well as between symptoms C12 (no future) and C14 (suicidal). Level three connections display the strongest relations between symptoms C33 (hallucinations) and C34 (voices conversing), C32 (blunted emotions) and C27 (blunted affect), C20 (voodoo) and C15 (telepathy), and C11 (important person) and C13 (special person). Within this level, most symptoms were connected to most of the other symptoms. Notably, both in level two and three, most items that had strong links with other items were in fact linked to an item from the same dimension. Level three of connectivity was formed mostly of the positive symptoms, while level four of connectivity was formed mostly of the negative symptoms and the relations between these. Overall, the predictive path diagram shows that it becomes possible to reach any other symptom in the network from the PRS in no more than four steps. Nonetheless, these results should be interpreted with caution considering the large variability in edge weights and the presence of negative edges.
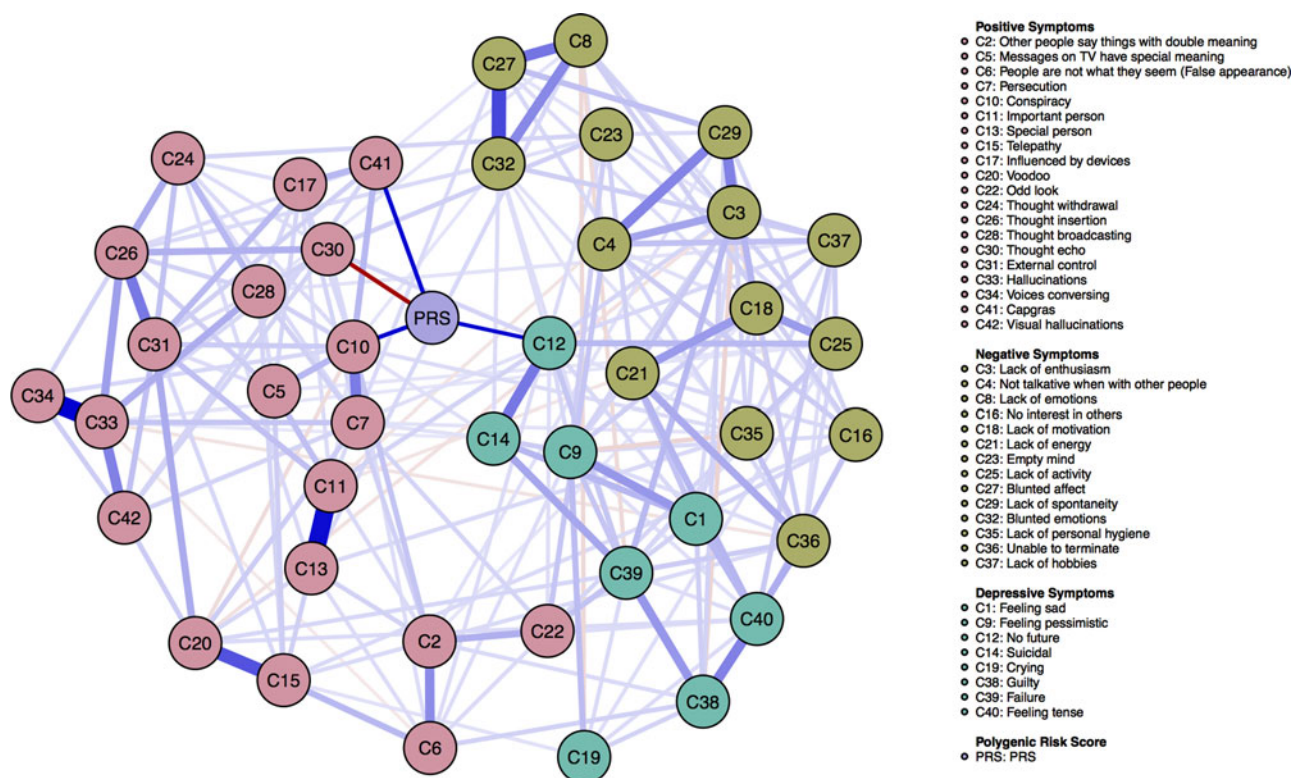
Bootstrapping showed that the estimation of node-specific predictive betweenness (i.e. items that more often lie on the shortest pathways from PRS to other nodes; Fig. 3) was considerably less precise than that of other features of the network. In particular, Fig. 3 shows that several nodes featured high node-specific predictive betweenness in bootstrap samples but not in the centrality of the estimated network structure based on the sample, indicating that the shortest paths from PRS to other nodes can take various forms. A potentially important result that was robustly present across case-drop bootstraps (online Supplementary Fig. S4) is that the most central symptoms in the current sample in terms of node-specific predictive betweenness are symptoms C12 (no future) and C10 (conspiracy). Other central symptoms – even though to a much lesser degree and possibly as secondary items on the pathway between PRS and other nodes – were items C25 (lack of activity), C30 (thought echo), and C39 (failure). Additionally, the nonparametric bootstrapped networks showed items C9 (feeling pessimistic), C21 (lack of energy), and C36 (unable to terminate) as further potentially important nodes, even though these were not specifically captured as high node-specific predictive betweenness items in the current network structure (Fig. 3, black lines).

Discussion

The current study provides, to the best of our knowledge, a first application of the network framework (Borsboom and Cramer, 2013) to the field of behavior genetics research. This allows for a novel outlook on the investigation of the relations between GWAS-based PRSs and symptoms of mental disorders, by focusing on the dependencies among variables.

Most links we identified in our data are links between the PRS and positive psychotic symptoms, especially symptoms related to notions of *conspiracy* and *paranoia*. This can be especially of interest as studies including risk factors in network models of psychosis (e.g. cannabis use, childhood trauma, urbanization) did not identify direct relations between any of these risk factors and positive or negative psychotic symptoms (Isvoranu *et al.*, 2016a, 2016b). In addition, positive psychotic symptoms are generally found to have few connections to other nodes, especially nodes related to real life functioning (Galderisi *et al.*, 2018). In light of this, a plausible hypothesis to put forward would be that different types of risk factors may lead to different pathways to the onset (and development) of psychotic disorders: genetic

Positive Symptoms
- C2: Other people say things with double meaning
- C5: Messages on TV have special meaning
- C6: People are not what they seem (False appearance)
- C7: Persecution
- C10: Conspiracy
- C11: Important person
- C13: Special person
- C15: Telepathy
- C17: Influenced by devices
- C20: Voodoo
- C22: Odd look
- C24: Thought withdrawal
- C26: Thought insertion
- C28: Thought broadcasting
- C30: Thought echo
- C31: External control
- C33: Hallucinations
- C34: Voices conversing
- C41: Capgras
- C42: Visual hallucinations

Negative Symptoms
- C3: Lack of enthusiasm
- C4: Not talkative when with other people
- C8: Lack of emotions
- C16: No interest in others
- C18: Lack of motivation
- C21: Lack of energy
- C23: Empty mind
- C25: Lack of activity
- C27: Blunted affect
- C29: Lack of spontaneity
- C32: Blunted emotions
- C35: Lack of personal hygiene
- C36: Unable to terminate
- C37: Lack of hobbies

Depressive Symptoms
- C1: Feeling sad
- C9: Feeling pessimistic
- C12: No future
- C14: Suicidal
- C19: Crying
- C38: Guilty
- C39: Failure
- C40: Feeling tense

Polygenic Risk Score
- PRS: PRS

Fig. 1. Network of the 42 CAPE (Konings *et al.*, 2006) symptoms and the PRS for psychosis (*n* = 2180). Blue (red) lines represent positive (negative) associations between variables and the wider and more saturated the edge, the stronger the association (Epskamp *et al.*, 2012). Please note that since the focus of the paper was to investigate the relations between the PRS and symptoms, the edges between the PRS and symptoms have been manually un-faded, while the edges between the other nodes in the network retain transparency. Symptom groups are differentiated by color.

risk factors may have more of a main effect on the liability to develop positive psychotic symptomatology, while environmental risk factors may have more of a secondary effect on general psychopathology symptoms. This may partly explain why environmental factors are often shared between distinct mental disorder categories and is in line with previous findings which argued that symptoms can have different risk factors (Fried *et al.*, 2014, 2015a, 2015b; Isvoranu *et al.*, 2016a, 2016b).

To further investigate symptom associations at different levels of connectivity, the current study constructed for the first time a *predictive path diagram*. To a certain degree, the predictive path diagram aims to point out toward potential mediating items. Here, the second level of symptom–symptom interactions consisted of mostly positive psychotic symptoms, while the third level of symptom–symptom interactions had a predominance of negative psychotic symptoms. Most of the symptoms connected to other symptoms from the same domain, suggesting that node activation may first spread within the same domain. In addition, since most links were identified between the PRS and positive psychotic symptoms, the second level included most of the positive psychotic symptoms. The few negative symptoms identified at the second level of symptom–symptom interactions are items related to lack of enthusiasm, energy, and (social) activities. The only two positive psychotic symptoms identified at the third level of interaction are *telepathy* and *special person*. Depressive symptoms are shared between the two levels and generally have less strong relations within and between domains. It should be noted that the symptoms with the strongest relations between different levels of connectivity are highly similar in content to each other (e.g. strong relations between items C10: *Do you ever feel as*

if there is a conspiracy against you? and C7: *Do you ever feel as if you are being persecuted in some way?* or between items C32: *Do you ever feel that your emotions are blunted?* and C27: *Do you ever feel that your feelings are lacking in intensity?*). This result may therefore be an artifact of the scale used to measure symptomatology, in which many of the items are contextually similar and therefore are likely highly correlated with each other. Further research using different measurement scales is warranted before drawing strong conclusions based on this methodology.

Finally, for the current study we developed a measure of *node-specific predictive betweenness*, which allows for the investigation of items that more often lie on the shortest pathway from PRS to other nodes in the network. We identified the items *conspiracy* and *no future* as central items and hypothesize that these may be important hubs in the network.

Methodological notes

Several qualifications of the network methodology are in order here. First, the statistical approach of incorporating PRS in a network is suited to pick up direct links between the PRS and psychotic symptoms. As such, this methodology is best interpreted as incorporating the hypothesis that PRS contains genetic factors that increase the liability to develop certain symptoms (i.e. acts through main effects on the symptom thresholds in a network). This is because PRS is modeled as a predictive node from the causal background.

Second, predictive nodes from the causal background are of course unlikely to operate on the same timescale as symptoms in a symptom network. However, the nature of statistical control
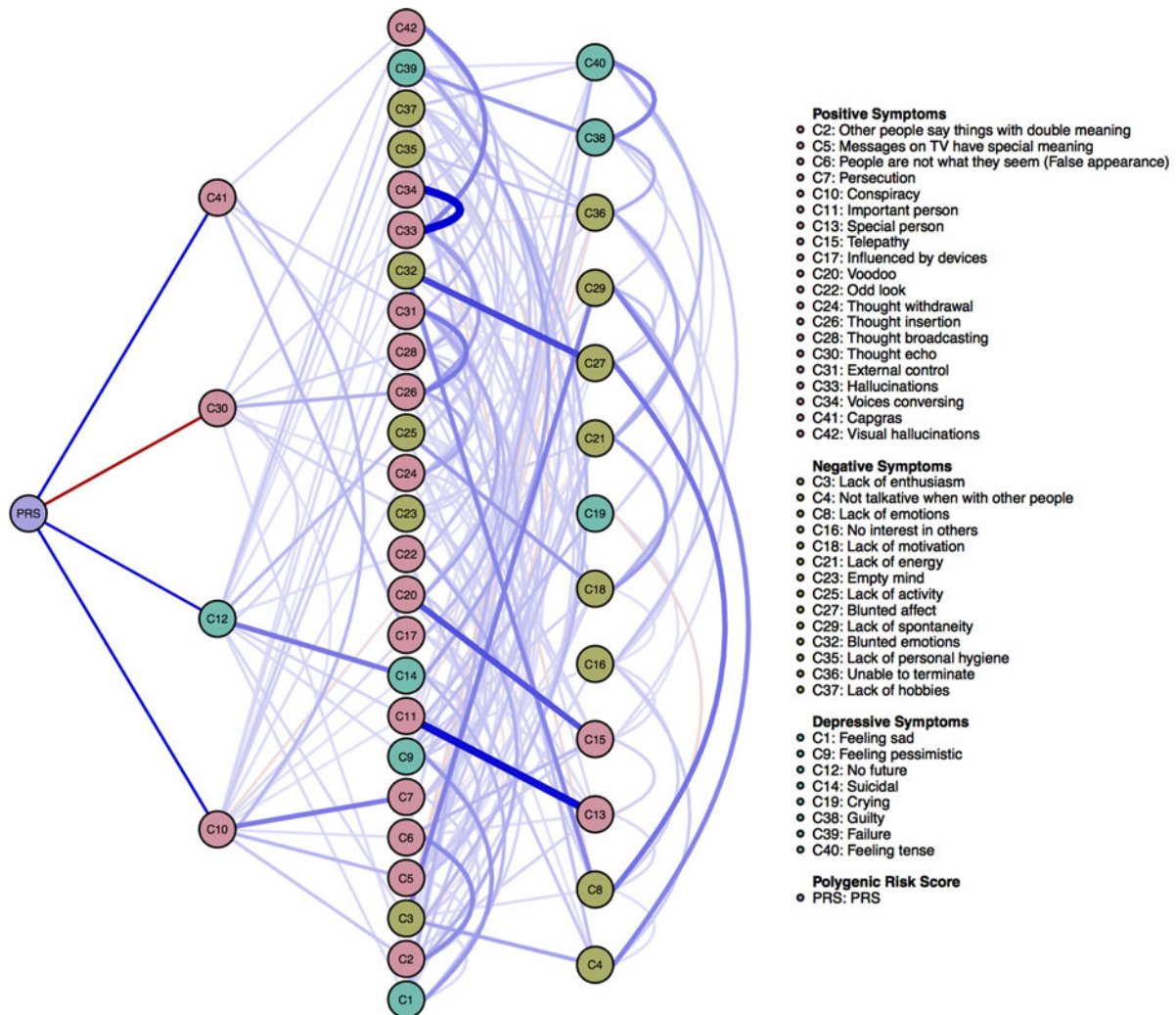
Fig. 2. Predictive path diagram of the 42 CAPE (Konings *et al.*, 2006) symptoms and the PRS for psychosis (*n* = 2180). A predictive path diagram is based on *shortest pathway analysis* (Brandes, 2008; Opsahl *et al.*, 2010) and represents the first three levels of connectivity between variables. Specifically, here we visualize (1) the immediate nodes to which the PRS is connected; (2) the immediate nodes to which the nodes that are connected to the PRS are connected; and (3) the nodes to which the latter are connected. Blue (red) lines represent positive (negative) associations between variables and the wider and more saturated the edge, the stronger the association (Epskamp *et al.*, 2012). Please note that since the focus of the paper was to investigate the relations between the PRS and symptoms, the edges between the PRS and symptoms have been manually un-faded (i.e. within the first level of connectivity), while the edges between the other nodes retained default plotting functions. Symptom groups are differentiated by color.

in a partial correlation network which includes PRS as a node is likely not that different from that of a network that exclusively contains symptoms. If, for instance, PRS was a common cause for two symptoms within a network structure (e.g. increased PRS leads to increased liability to develop symptoms *sadness* and *conspiracy*) but not included in the network (and thus not controlled for), a spurious edge between the two symptoms would be induced. If PRS was indeed linked to symptoms, such spurious associations would then be eliminated due to the inclusion of PRS as a node in the network. Of note, a node only induces spurious connections if it is a common effect in a network (Epskamp *et al.*, 2018), but PRS cannot be a common effect, because it is not caused by symptoms. Therefore, to the extent that stable edges arise between PRS and symptomatology, such edges are consistent with the hypothesis that PRS impinges at the symptom level.

Third, it is widely known that genetic influences account for a small proportion of the explained variance in a mental disorder and even though this proportion is larger for psychosis, the effect

is still modest (Greenwood *et al.*, 2007). Naturally, symptoms tend to cluster together and have a stronger influence on each other, while variables from the external field of a disorder, such as environmental and genetic factors also better cluster together within their domain. Networks therefore often display weaker between-domain links (Isvoranu *et al.*, 2016b; Santos *et al.*, 2017), and may have missed between-domain links due to a lack of statistical power. In addition, recent studies showed that when using estimation techniques relying on the lasso regularization (Tibshirani, 1994) – which many prior network studies have been based on – specificity can be lower than expected in dense networks with many small edges, leading to an increase in false positives (Williams and Rast, 2018). To this end, we chose a different and more conservative estimation technique here based on unregularized model search, as recently recommended by Epskamp (2018). For low-dimensional settings with high sample size, simulation studies using this technique have shown good sensitivity as well as high specificity. We compared this method with the classic lasso regularization method, as well as with a method based on
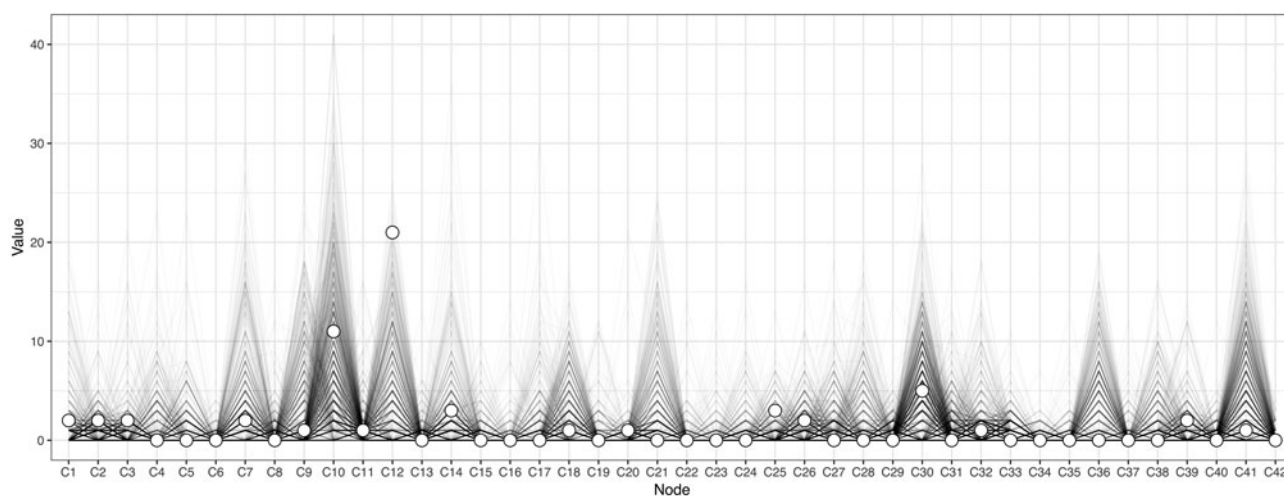
Fig. 3. Node-specific predictive betweenness (i.e. how often a node lies on the pathways between two other nodes, of which one is always the PRS). The white dots represent the node-specific predictive betweenness in the current sample, while the black lines represent the variability of node-specific betweenness when using nonparametric bootstrapping over 1000 iterations.

thresholded partial correlations. The unregularized model search performed best in terms of goodness-of-fit measures. In addition, the edges from PRS to *conspiracy* and *no future* were identified by two of the three methods, while the edges from PRS to *thought echo* and *capgras* were identified by all three methods (see online Supplementary Fig. S5 and Table S1). Overall, based on these results, the effects identified in this study are unlikely to be simply a result of noise and may hold etiological significance. In addition, bootstrap analyses (see online Supplementary Fig. S2) show the edges from PRS to *C41: capgras* and *C30: thought echo* are the most robust (i.e. showing up in 67% of the bootstraps and 75% of the bootstraps respectively). Notably, all edges from the PRS are identified in at least 50% of the bootstraps. Since the network was not constructed using *p* values but was based on a model fitting approach, which has been shown to be highly effective in type-I error control, multiple testing should not pose as an issue here.

Limitations

The findings of this study need to be considered in light of several limitations. First, a PRS itself can be considered an important limitation. The signal between a PRS and symptoms is generally weak (Greenwood *et al.*, 2007), also resulting in weak network connections, which are not overly consistent across different estimation techniques. In addition, by adding up SNPs to construct the PRS we may obscure details in the same way in which adding up items into sum-scores may obscure the different symptoms – investigating individual SNPs in relation to symptoms may therefore be a promising new avenue. It is further possible that the predictive power of the PRS also becomes weaker in broader mixed samples across the spectrum. The sample analyzed in this paper is a heterogeneous sample, which includes controls, siblings, parents and patients; in addition, as the CAPE is not designed for measuring symptomatology in patients, ceiling effects may occur when applying the CAPE to the patient population. However, given that the strong phenotypic manifestation of the PRS can often be identified in healthy relatives of patients (van Os *et al.*, 2017), we chose to include the whole sample rather than focus on one group only. Ideally it would be possible to carry out analyses and direct network comparisons between different

population groups; this was not possible in the current study due to limited data. Second, it is difficult to investigate to what degree any PRS effect would be reducible to group membership (i.e. patients have much more psychopathology and higher PRS scores). Including a *group membership* variable in a network would be problematic (Fried and Cramer, 2017) due to the categorical nature of the variable and therefore investigating to what extent there would be overlap between the two variables was not possible. A future extension for network modeling of genetic data may be modeling the multilevel structure of the data. Third, the negative association between PRS and thought echo is a puzzling finding, which could be related to some form of measurement bias or rarity, but could also be a true effect given a fixed level of other symptoms. Further investigation showed that, even though the expectation would be that PRS would positively correlate with all other variables, PRS and though echo were actually also weakly marginally negatively correlated ($r = -0.004$), thus ruling out that the result is paradoxical (i.e. a marginal positive correlation was not made negative upon conditioning). One possible alternative explanation to the stronger negative association in the network is that it may be due to the presence of a collider in the data (i.e. both PRS and thought echo cause a third variable in the dataset; Pearl, 2000). Another explanation may be that this is a spurious effect, resulting from the violation of the normality assumption (i.e. the variable 'thought echo' is relatively skewed to the left). Future research should focus on if the link between PRS and thought echo is a genuine negative effect that can be replicated. Finally, the current analyses were based on cross-sectional data and conclusions regarding direction or causality should be drawn with caution. In addition, it may be argued that results from cross-sectional networks may not be generalizable to within-person dynamics (Bos *et al.*, 2017). It should however be noted that longitudinally investigating variables from the external field of a network, such as environmental or genetic risk factors, is methodologically infeasible as these background variables may not change themselves; in addition, as we argued in this manuscript, if such variables act to facilitate the ease with which nodes in the network are activated, the current analysis may in fact be more informative than a within-subject analysis of dynamically changing factors would by itself be. Future

research could expand on our analyses by investigating whether PRS might further have an effect on the interactions between symptoms, using *network moderation* techniques (Haslbeck *et al.*, 2018) or by relating PRS to parameters of networks that govern dynamic changes in symptomatology.

Conclusion

Overall, the results of our study indicate that investigating pathways within a network structure through which genetic components may affect the liability to develop a mental disorder may be promising and may open new research avenues. Here we have taken a first step toward introducing this methodology. Further investigating how (and whether) other PRSs are related to (other) symptoms, as well as whether we can identify certain genetic segments for specific symptoms could lead to biological informative structures. The current study is exploratory in nature; in light of this, we discuss methodological considerations and limitations of the approach and argue further research replicating these results is essential. Ultimately, we aim to provide a first means toward bringing together the field of behavior genetics and the network framework, and provide preliminary results for symptom-specific biological pathways to psychosis.

Note

[1] The estimation procedure first runs the glasso algorithm (Friedman *et al.*, 2008) for 100 different tuning parameters to obtain 100 different network structures (ranging from sparse networks with few edges to dense networks with many edges). Next, the algorithm refits all those networks without regularization and selects the model that minimizes the Bayesian information criterion (BIC). Subsequently, the algorithm adds and removes edges until BIC can no longer be improved. It has been shown that BIC selection of GGMs selects the true model as *N* grows to infinity (Foygel & Drton, 2011). In addition, the stepwise procedure ensures that no single edge can be removed or added in the final model to improve fit.

[2] This is because the software used for constructing the networks automatically fades edges with a smaller effect – an effect between a variable from the external field of a symptom (such as an environmental variable or a genetic risk score; Borsboom, 2017) will likely almost always be smaller than the effect of the variables within the network (e.g. the relations between the symptoms themselves).

[3] Of note, when running the same analysis within the subsample of patients and relatives only, the relations between the PRS and C10 (Conspiracy), and C41 (Capgras) are retained and no other edges emerge. It is likely that edges between PRS and C12 (No future) and C30 (Thought echo) do not withhold due to a loss in power.

Author ORCIDs. (iD) Adela-Maria Isvoranu, 0000-0001-7981-9198

References

American Psychiatric Association (2000) *DSM-IV. Diagnostic and Statistical Manual of Mental Disorders, 4th edition TR*. Washington, DC: American Psychiatric Association.

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

Borsboom D and Cramer AOJ (2013) Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology* **9**, 91–121.

Borsboom D, Robinaugh DJ, Rhemtulla M and Cramer AOJ (2018) Robustness and replicability of psychopathology networks. *World Psychiatry* **17**, 143–144.

Bos FM, Snippe E, De Vos S, Hartmann JA, Simons CJP, Van Der Krieke L, De Jonge P and Wichers M (2017) *Can We Jump from Cross-Sectional to Dynamic Interpretations of Networks? Implications for the Network Perspective in Psychiatry. Psychotherapy and Psychosomatics* **6**, 175–177.

Brandes U (2008) On variants of shortest-path betweenness centrality and their generic computation. *Social Networks* **30**, 136–145.

Cramer AOJ, Waldorp LJ, van der Maas HLJ and Borsboom D (2010) Comorbidity: a network perspective. *The Behavioral and Brain Sciences* **33**, 137–150, discussion 150-93.

Dudbridge F (2013) Power and predictive accuracy of polygenic risk scores. *PLoS Genetics* **9**, e1003348.

Epskamp S (2017) *Network Psychometrics*, PhD dissertation. Available at http://dare.uva.nl/search?identifier=a76273c6-6abc-4cc7-a2e9-3b5f1ae3c29e.

Epskamp S (2018) *New features in qgraph* 1.5. Available at http://psychosystems.org/qgraph_1.5.

Epskamp S and Fried EI (2018) A tutorial on regularized partial correlation networks. *Psychological Methods* **23**, 617–634.

Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD and Borsboom D (2012) Qgraph: network visualizations of relationships in psychometric data. *Journal of Statistical Software* **48**, 1–18.

Epskamp S, Borsboom D and Fried EI (2017) Estimating psychological networks and their accuracy: a tutorial paper. *Behavior Research Methods* **50**, 195–212.

Epskamp S, Waldorp LJ, Mõttus R and Borsboom D (2018) The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research* **53**, 453–480.

Euesden J, Lewis CM and O'Reilly PF (2015) PRSice: Polygenic Risk Score software. *Bioinformatics (Oxford, England)* **31**, 1466–1468.

Foygel R and Drton M (2011) Bayesian model choice and information criteria in sparse generalized linear models. *arXiv preprint arXiv*:1112.5635, 1–37.

Franić S, Dolan CV, Borsboom D, Hudziak JJ, van Beijsterveldt CEM and Boomsma DI (2013) Can genetics help psychometrics? Improving dimensionality assessment through genetic factor modeling. *Psychological Methods* **18**, 406–433.

Fried EI and Cramer AOJ (2017) Moving forward: Challenges and direction for psychopathological network theory and methodology. *Perspectives on Psychological Science* **12**, 999–1020.

Fried EI, Nesse RM, Zivin K, Guille C and Sen S (2014) Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. *Psychological Medicine* **44**, 2067–2076.

Fried EI, Bockting C, Arjadi R, Borsboom D, Amshoff M, Cramer AOJ, Epskamp S, Tuerlinckx F, Carr D and Stroebe M (2015a) From loss to loneliness: the relationship between bereavement and depressive symptoms. *Journal of Abnormal Psychology* 124, 256–265.

Fried EI, Epskamp S, Nesse RM, Tuerlinckx F and Borsboom D (2015b) 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.

Friedman J, Hastie T and Tibshirani R (2008) Sparse inverse covariance estimation with the graphical lasso. *Biostatistics (Oxford, England)* 9, 432–441.

Galderisi S, Rucci P, Kirkpatrick B, Mucci A, Gibertoni D, Rocca P, Rossi A, Bertolino A, Strauss GP, Aguglia E, Bellomo A, Murri MB, Bucci P, Carpiniello B, Comparelli A, Cuomo A, De Berardis D, Dell'Osso L, Di Fabio F, Gelao B, Marchesi C, Monteleone P, Montemagni C, Orsenigo G, Pacitti F, Roncone R, Santonastaso P, Siracusano A, Vignapiano A, Vita A, Zeppegno P and Maj M (2018) Interplay among psychopathologic variables, personal resources, context-related factors, and real-life functioning in individuals with schizophrenia a network analysis. *JAMA Psychiatry* 5, 396–404.

Greenwood TA, Braff DL, Light GA, Cadenhead IS, Calkins ME, Dobie DJ, Freedman R, Green MF, Gur RE, Gur RC, Mintz J, Nuechterlein KH, Olincy A, Radant AD, Seidman LJ, Siever LJ, Silverman JM, Stone WS, Swerdlow NR, Tsuang DW, Tsuang MT, Turetsky BI and Schork NJ (2007) Initial heritability analyses of endophenotypic measures for schizophrenia: The Consortium on the Genetics of Schizophrenia. *Archives of General Psychiatry* 64, 1242–1250.

Guloksuz S and van Os J (2017) The slow death of the concept of schizophrenia and the painful birth of the psychosis spectrum. *Psychological Medicine* 48, 229–244.

Haslbeck J, Borsboom D and Waldorp L (2018) Moderated Network Models. *arXiv preprint* arXiv:1807.02877.

Isvoranu A-M, Borsboom D, van Os J and Guloksuz S (2016a) A network approach to environmental impact in psychotic disorder: brief theoretical framework. *Schizophrenia Bulletin* 42, 870–873.

Isvoranu A-M, van Borkulo CD, Boyette LL, Wigman JTW, Vinkers CH and Borsboom D (2016b) A network approach to psychosis: pathways between childhood trauma and psychotic symptoms. *Schizophrenia Bulletin* 43, 187–196.

Isvoranu A-M, Boyette L-L, Guloksuz S and Borsboom D (2019) Chapter G1. Symptom network models of psychosis. In van Os J, Reininghaus U, Ivleva EI and Tamminga CA (eds), *Psychotic Disorders: Comprehensive Conceptualization and Treatments*. Oxford, UK: Oxford University Press.

Keshavan MS, Nasrallah HA and Tandon R (2011) Schizophrenia, 'just the facts' 6. *Moving ahead with the schizophrenia concept: From the elephant to the mouse. Schizophrenia Research* 127, 3–13.

Konings M, Bak M, Hanssen M, Van Os J and Krabbendam L (2006) Validity and reliability of the CAPE: a self-report instrument for the measurement of psychotic experiences in the general population. *Acta Psychiatrica Scandinavica* 114, 55–61.

Korver N, Quee PJ, Boos HB, Simons CJ and de Haan L (2012) Genetic Risk and Outcome of Psychosis (GROUP), a multi-site longitudinal cohort study focused on gene-environment interaction: objectives, sample characteristics, recruitment and assessment methods. *International Journal of Methods in Psychiatric Research* 21, 205–221.

Lauritzen SL and Wermuth N (1989) *Graphical Models for Associations Between Variables, Some of which are Qualitative and Some Quantitative. The Annals of Statistics* 17, 1916–1916.

Marden JR, Walter S, Tchetgen Tchetgen EJ, Kawachi I and Glymour MM (2014) Validation of a polygenic risk score for dementia in black and white individuals. *Brain and Behavior* 4, 687–697.

Opsahl T, Agneessens F and Skvoretz J (2010) Node centrality in weighted networks: generalizing degree and shortest paths. *Social Networks* 32, 245–251.

Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, Legge SE, Bishop S, Cameron D, Hamshere ML, Han J, Hubbard L, Lynham A, Mantripragada K, Rees E, MacCabe JH, McCarroll SA, Baune BT, Breen G, Byrne EM, Dannlowski U, Eley TC, Hayward C, Martin Nicholas G, McIntosh AM, Plomin R, Porteous DJ, Wray Naomi R, Caballero A, Geschwind DH, Huckins LM, Ruderfer DM, Santiago E, Sklar P, Stahl EA, Won H, Agerbo E, Als Thomas D, Andreassen OA, Bækvad-Hansen M, Mortensen PB, Pedersen CB, Børglum AD, Bybjerg GJ, Djurovic S, Durmishi N, Pedersen MG, Golimbet V, Grove J, Hougaard DM, Mattheisen M, Molden E, Mors O, Nordentoft M, Pejovic-Milovancevic M, Sigurdsson E, Silagadze T, Hansen CS, Stefansson K, Stefansson H, Steinberg S, Tosato S, Werge T, Harold D, Sims R, Gerrish A, Chapman J, Abraham R, Hollingworth P, Pahwa J, Denning N, Thomas C, Taylor S, Powell J, Proitsi P, Lupton M, Lovestone S, Passmore P, Craig D, McGuinness B, Johnston J, Todd S, Maier W, Jessen F, Heun R, Schurmann B, Ramirez A, Becker T, Herold C, Lacour A, Drichel D, Nothen M, Goate A, Cruchaga C, Nowotny P, Morris JC, Mayo K, O'Donovan M, Owen M, Williams J, Achilla E, Barr CL, Böttger TW, Cohen D, Curran S, Dempster E, Dima D, Sabes-Figuera R, Flanagan RJ, Frangou S, Frank J, Gasse C, Gaughran F, Giegling I, Hannon E, Hartmann AM, Heißerer B, Helthuis M, Horsdal HT, Ingimarsson O, Jollie K, Kennedy JL, Köhler O, Konte B, Lang M, Lewis C, MacCaba J, Malhotra AK, McCrone P, Meier SM, Mill J, Nöthen MM, Pedersen CB, Rietschel M, Rujescu D, Schwalber A, Sørensen HJ, Spencer B, Støvring H, Strohmaier J, Sullivan P, Vassos E, Verbelen M, Collier DA, Kirov G, Owen MJ, O'Donovan MC and Walters JTR (2018) Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics* 50, 381.

Pearl J (2000) *Causality: Models, Reasoning and Inference*. Cambridge, MA, USA: Cambridge University Press.

R Development Core Team R (2014) *R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing* 1, 409.

Santos H, Fried EI, Asafu-Adjei J and Jeanne Ruiz R (2017) Network structure of perinatal depressive symptoms in Latinas: Relationship to stress and reproductive biomarkers. *Research in Nursing and Health* 40, 218–228.

Sullivan PF (2008) Schizophrenia genetics: the search for a hard lead. *Current Opinion in Psychiatry* 21, 157–160.

Tibshirani R (1994) *Regression Selection and Shrinkage via the Lasso. Journal of the Royal Statistical Society B* 58, 267–288.

Van Os J and Linscott RJ (2012) Introduction: the extended psychosis phenotype – relationship with schizophrenia and with ultrahigh risk status for psychosis. *Schizophrenia Bulletin* 38, 227–230.

van Os J, van der Steen Y, Islam MA, Gülöksüz S, Rutten BP and Simons CJ (2017) Evidence that polygenic risk for psychotic disorder is expressed in the domain of neurodevelopment, emotion regulation and attribution of salience. *Psychological Medicine* 47, 2421–2437.

Williams DR and Rast P (2018) Back to the Basics: Rethinking Partial Correlation Network Methodology. Available at osf.io/c62pb.