BRIEF REPORT

REVISED Brain structural abnormalities in six major psychiatric disorders: shared variation and network perspectives [version 2; peer review: 1 approved, 2 approved with reservations]

Euclides José de Mendonça Filho[1,2], Márcio Bonesso Alves[1,2], Patricia Pelufo Silveira[1,2]

[1]Department of Psychiatry, McGill University, Montreal, Quebec, H3A 1A1, Canada
[2]Ludmer Centre for Neuroinformatics & Mental Health, Douglas Mental Health University Institute, Montreal, Quebec, H4H 1R3, Canada

v2 **First published:** 07 May 2021, **10**:356
https://doi.org/10.12688/f1000research.51475.1
Latest published: 17 Jun 2022, **10**:356
https://doi.org/10.12688/f1000research.51475.2

Abstract

Common brain abnormalities are a possible explanation for comorbidities in psychiatric disorders. Challenges in understanding these conditions are likely due to the paucity of studies able to analyze the extent and regional distribution of shared morphometric abnormalities between disorders. Recently, Opeal *et al.* presented an elegant rationale to investigate shared and specific morphometric measures of cortical thickness and subcortical gray matter volume between healthy individuals and subjects across six major psychiatric disorders. Although their approach has the potential to systematically portray shared brain alterations, the chosen principal component analysis solution may not address the central question of the observed shared versus specific brain alterations due to misspecification of the number of components. Given how this misspecification can lead to different conclusions, we reanalyzed Opel *et al.* data to thoroughly determine the number of factors to be considered, explore the alternative solution, and visualize the patterns of shared brain matter correlations using network analysis. Our approach suggests that a unidimensional solution was appropriate in this situation. The unidimensional solution indicated that brain alterations in autism spectrum disorder (ASD) had a significant negative component loading, suggesting that brain abnormalities found in ASD covaried with major depressive disorder (MDD), bipolar disorder (BD), schizophrenia (SCZ), and obsessive-compulsive disorder (OCD), a finding not demonstrated by the original work. Network analysis indicated that SCZ had the highest strength, BD the highest closeness, and BD and MDD had the highest betweenness in the network. This work highlights how different component solutions can lead to different conclusions, with important implications for the understanding of overlapped patterns of symptoms among six major

Open Peer Review

Approval Status ✓ ? ?

	1	2	3
version 2 (revision) 17 Jun 2022	✓ view	? view	? view
version 1 07 May 2021	? view		

1. **Takahiro Osada**, Juntendo University School of Medicine, Tokyo, Japan
2. **Esther Walton**, University of Bath, Bath, UK
3. **Y. C. Janardhan Reddy**, National Institute of Mental Health and Neuro Sciences (NIMHANS), Bengaluru, India
Srinivas Balachander, National Institute of Mental Health and Neuro Sciences (NIMHANS), Bangalore, India

Any reports and responses or comments on the article can be found at the end of the article.

psychiatric diseases. The network approach is complementary in indicating central markers of specific psychopathology domains. Investigations using shared-variation and network perspectives are promising for the study of pathophysiological patterns of common brain alterations.

Keywords
Cross-disorder, ENIGMA, Psychiatric disorders, Structural MRI, Principal Component Analysis, Network Analysis

This article is included in the INCF gateway.

Corresponding author: Patricia Pelufo Silveira (patricia.silveira@mcgill.ca)

Author roles: de Mendonça Filho EJ: Conceptualization, Data Curation, Formal Analysis, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Alves MB**: Conceptualization, Data Curation, Methodology, Writing – Original Draft Preparation; **Silveira PP**: Conceptualization, Funding Acquisition, Methodology, Resources, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

How to cite this article: de Mendonça Filho EJ, Alves MB and Silveira PP. **Brain structural abnormalities in six major psychiatric disorders: shared variation and network perspectives [version 2; peer review: 1 approved, 2 approved with reservations]** F1000Research 2022, **10**:356 https://doi.org/10.12688/f1000research.51475.2

First published: 07 May 2021, **10**:356 https://doi.org/10.12688/f1000research.51475.1

REVISED **Amendments from Version 2**

Major alterations in the manuscript are related to additional information on the rationale for retaining a unidimensional solution, and alterations in Figure 2.

Any further responses from the reviewers can be found at the end of the article

Introduction

Challenges to the understanding of heterogeneity and comorbidity of psychiatric disorders have long been acknowledged in medicine. The National Institute of Mental Health's Research Domain Criteria (RDoC) initiative acknowledges that common brain abnormalities can potentially explain psychiatry comorbidities,[1] but few studies were able to systematically investigate the extent and regional distribution of shared morphometric abnormalities between disorders.

Using published meta- and mega-analyses of the Enhancing Neuro Imaging Genetics Through Meta-Analysis (ENIGMA) consortium, Opel *et al.*[2] present an elegant rationale to investigate shared and specific morphometric measures of cortical thickness and subcortical gray matter volume between typical control individuals and subjects with six of the major psychiatric disorders (major depressive disorder [MDD], bipolar disorder [BD], schizophrenia [SCZ], obsessive-compulsive disorder [OCD], attention-deficit/hyperactivity disorder [ADHD], and autism spectrum disorder [ASD]). To address whether brain-structural alterations related to these disorders loaded onto latent variables, shared brain abnormalities among them were examined using principal component analysis (PCA) across all cortical and subcortical regions. Then component scores were compared with the empirical regional effect sizes, allowing the definition of regions in each disorder that were better predicted by a shared variance component. The authors retained three principal components across disorders with a solution defined by MDD, BD, SCZ, and OCD loading on the first component, ADHD on the second, and ASD on the third component, leading to the conclusion that MDD, BD, SCZ, and OCD shared neuro-abnormality patterns, whereas ASD and ADHD exhibited disease-specific alterations.

Although the rationale used to investigate shared and specific morphometric measures of brain matter consists of an opportune strategy for improving the understanding of the pathophysiologic mechanisms of psychiatric disorders,[3] the number of components retained in the analysis, in this case, may not address the central question of the observed shared versus specific brain alterations. The three-factor solution parcels out the weaker residual correlation into minor components that might be of theoretical importance. We recognize that it is often a challenge to define the appropriate number of factors for data reduction, but common recommendations assume that retaining components with eigenvalues >1.0 usually indicates an excessive number of components. Moreover, the consideration of weak or poorly identified factors (i.e., components defined by only one or two variables) is an indication that the number of factors extracted should be reconsidered.[4]

The over-extraction and under-extraction of factors retained in data reduction can have deleterious effects on the results.[5] Given the importance of the number of factors in data reduction and how different component solutions can lead to different conclusions, we leveraged the cross-disease effect sizes reported in Opel *et al.*[2] to thoroughly determine the number of factors to be considered, explore the appropriate solution, and visualize the patterns of shared brain matter correlations using network analysis.

Methods

We reanalyzed the underlying structure of 41 regional measures of cortical thickness and subcortical volumes across the six psychiatric disorders compiled by Opel *et al.*[2] The data consisted of effect sizes obtained by contrasting healthy controls (N = 33,146) and patients (N = 19,578) of published structural neuroimaging mega- and meta-analyses of the ENIGMA consortium in the years 2016 to 2019.[2] The data selection criteria included the availability of effect sizes of psychiatric disorders for all 34 cortical brain regions based on the Desikan-Killiany atlas automated labeling system,[6] and 7 subcortical regions included in the standardized probabilistic information, modeled using the Markov random fields imaging pipeline[7] applied by the ENIGMA consortium. This criteria identified 11 studies of six psychiatric disorders: MDD (N = 2), BD (N = 2), SCZ (N = 2), OCD (N = 2), ADHD (N = 2), and ASD (N = 1). The effect size measures consisted of Cohen's *d* mean differences in each cortical or subcortical region after age, sex, scanner adjustment, and in case of subcortical volume, total intracranial volume.[2]

Data source and preparation

The object of analysis is a dataset comprised of the effect size estimates for six psychiatric disorders and 41 regions of interest – hippocampus, amygdala, thalamus, accumbens, caudate, putamen, pallidum, isthmus cingulate cortex,

posterior cingulate cortex, rostral anterior cingulate cortex, caudal anterior cingulate cortex, lateral orbit frontal cortex, pars opercularis of inferior frontal gyrus, rostral middle frontal gyrus, superior frontal gyrus, medial orbital frontal cortex, pars orbitalis of inferior frontal gyrus, pars triangularis of inferior frontal gyrus, caudal middle frontal gyrus, precentral gyrus, frontal pole, paracentral lobule, insula, lateral occipital cortex, lingual gyrus, cuneus, pericalcarine cortex, inferior parietal cortex, supramarginal gyrus, precuneus, superior parietal, postcentral gyrus fusiform gyrus, middle temporal gyrus, inferior temporal gyrus, banks superior temporal sulcus, superior temporal gyrus, parahippocampal gyrus, transverse temporal gyrus, entorhinal cortex, and temporal pole. This dataset was obtained by copying the effect sizes estimates reported by Opel *et al.*[2] on Supplementary Table S2 of their manuscript.[2] The full manuscript was accessed in https://doi.org/10.1016/j.biopsych.2020.04.027 via the McGill Library Portal (https://www.mcgill.ca/library/) on September 19, 2020. The data extracted from Opel's manuscript[2] Supplementary Table S2 was prepared for statistical manipulation in the SPSS (V.21, IBM Corp., Armonk, NY) statistical environment and is available upon request.

Data analysis

The number of principal components retained was determined using the scree plot criteria and Horn's parallel analysis.[8] The scree plot shows how much variation each component captures from the data and allows to determine the inflection point in the data where additional components are unnecessary. The number of data points above the inflection is usually the number of components to retain. Horn's parallel analysis compares the eigenvalues randomly generated from the data using Monte-Carlo simulation with the original data. The number of components retained consists of the original eigenvalues that are higher than the simulated eigenvalues.[8] We also verified the number of components in the cross-disease correlation matrix with eigenvalues greater than 1.0, although this procedure is considered one of the least accurate methods for selecting the number of components.[5] At this stage, we used the functions implemented by the *psych* package[9] from the R statistical language (V. 4.0.2). At a second stage, we used the SPSS software to conduct exploratory factor analysis. The principal components algorithm was used for dimensional extraction, and component scores were obtained using the regression weights (Thurstone method), which are obtained by multiplying the inverse of the observed variable correlation matrix by the matrix of factor loadings.[9]

To investigate specific patterns of residual correlations between the psychiatric disorders, we fitted a Gaussian graphical model with a Least Absolute Shrinkage and Selection Operator (*gLASSO*)[10] to the data using the *qgraph*[11] R package. This procedure yields parsimonious partial-correlations of the brain alterations for each pair of psychiatric diseases here represented as nodes. Edges between diseases indicate a regularized partial correlation, after conditioning on all other diseases in the dataset. To assess the importance of nodes in the network, we computed the following centrality measures: strength, a measure of how well a node is directly connected to other diseases, closeness, how well a node is indirectly connected to other diseases, and betweenness, quantification of how important a node is in the average path between two other diseases.

Results

Inspection of the scree plot and Horn's parallel analysis indicated the consideration of one component instead of the original three-component solution. Figure 1 shows a steeper decrease from the first to the second eigenvalues, followed by a flatter pattern for the remaining components. In addition, only the first actual eigenvalue was higher than the resampled eigenvalues (depicted in red) suggesting a unidimensional solution. It can be noticed that the third component obtained an eigenvalue of .98, violating the liberal eigenvalues greater than the 1.0 cut-off. Therefore, we opted to retain one component for subsequent analysis.

The unidimensional solution explained 48.3% percent of the six-brain structural alterations and indicated that the effect size of the differences in brain structure between ASD patients and controls shared a significant negative component loading ($\lambda = -0.30$, $p = 0.04$) with brain abnormalities of the remaining five diseases (Figure 2). Although the unimodal solution explains less of the variance than the original work (the three-factor solution explained 85.14% of variance), the consideration of a unidimensional solution was based on several criteria. Namely, the scree plot of eigenvalues, the exhaustive parallel analysis procedure that identified one component as a recurrent robust solution, and the inspection of the Opel *et al.*[2] solution that indicated additional second and third components defined by only one variable each. Since principal component analysis is a data reduction technique, considering components with only one variable seemed redundant, and an indication that the number of factors extracted should be reconsidered.

While in the original analysis SCZ had the highest shared correlation (indicated by the highest component loading) of the four conditions (MDD, BD, SCZ, and OCD), the unidimensional solution showed that the BD ($\lambda = 0.89$, $p < 0.001$) followed by SCZ ($\lambda = 0.88$, $p < 0.001$), and OCD ($\lambda = 0.87$, $p < 0.001$) had the highest shared correlation with all six disorders (Figure 2B), illustrating how different dimensional reduction solutions are implied in different patterns of

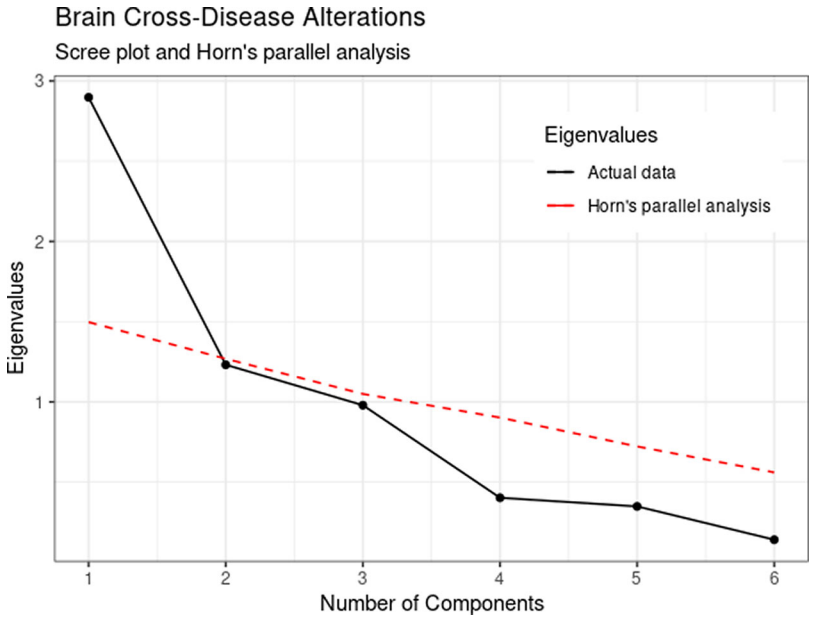

Figure 1. Eigenvalues scree plot and Horn's parallel analysis of brain structural effect size alteration in six psychiatric disorders.

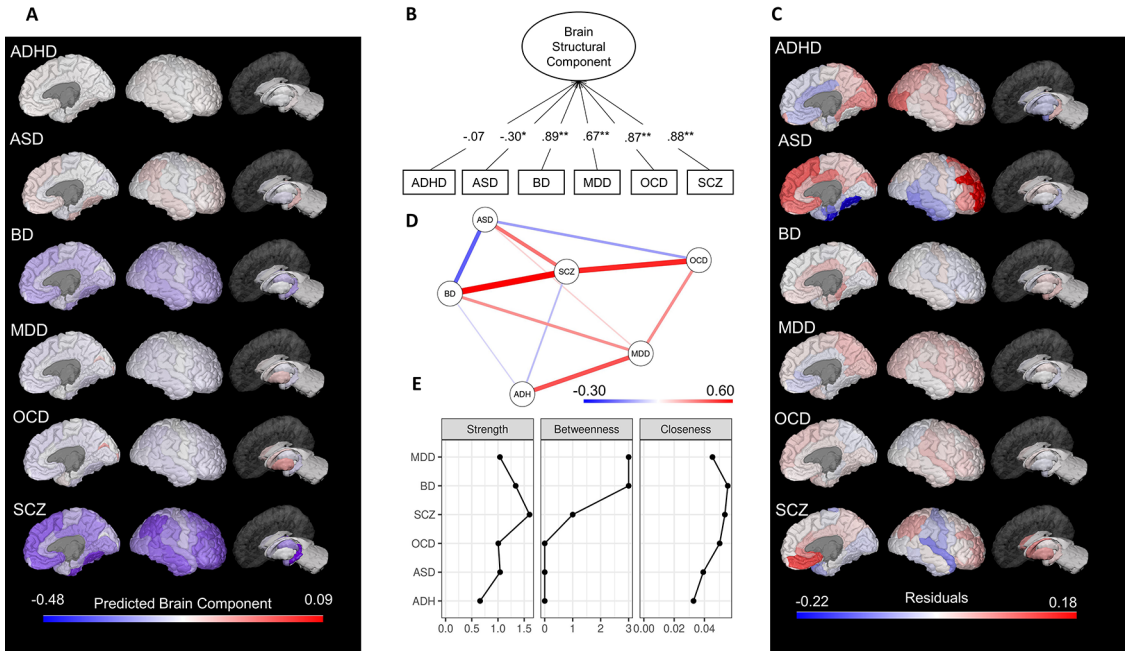

Figure 2. Principal component and network analysis of the effect sizes of brain structural alterations in six psychiatric disorders. (A) Predicted component scores mapped onto brain regions; (B) Path diagram of the principal component unidimensional solution, with number depicting component loadings; (C) Residuals from the regional effect sizes accounting for the predicted shared principal component. As in Opel *et al.*,[2] the absolute size of residuals encompasses the degree of representation through the shared unidimensional component. Lower (negative) residuals represent underestimation and higher (positive) residuals represent overestimation based on the brain-regional shared variance. (D) *gLASSO* correlations network of brain structural alterations in the six psychiatric disorders. Edges represent parsimonious partial-correlations between psychiatric disorders. A stronger correlation (positive = blue; negative = red) results in a thicker and darker edge. (E) Network centrality measures. ADHD, attention-deficit/hyperactivity disorder; ASD, autism spectrum disorder; BD, bipolar disorder; MDD, major depressive disorder; OCD, obsessive-compulsive disorder; SCZ, schizophrenia. All effect sizes used in these results are taken from Opel *et al.*[2] *$p < 0.05$; **$p < 0.001$.

covariance. Similar to the original report, ADHD did not load into the PCA, indicating low shared brain abnormalities with the other disorders.

We also calculated a regional component score (M = 0.0, SD = 1.0) to identify which brain areas were more affected in a cross-disease manner (see Figure 2A for the regional component score map). The hippocampus (-1.64) and the fusiform gyrus (-1.40) exhibited a more prominent shared reduction, whereas the pallidum (2.94) and putamen (2.09) showed a stronger shared increase. In order to explore shared- and disorder-specific morphometric abnormalities, we computed regional effect-sizes residuals from the component score. In contrast to the original work, this allowed the inclusion of ASD and ADHD in the analysis, although ADHD was not further explored due to its non-significant component loading. Regional specificities for MDD, BD, SCZ, and OCD were similar to Opel *et al.*[2] results, in which large absolute residuals were found in the rostral ACC and the medial orbito frontal cortex (OFC) for MDD, parahippocampal gyrus and pallidum for BD and superior temporal gyrus and medial OFC for SCZ (Figure 2C).

ASD showed large residuals especially for the rostral middle (residual [res] = .18) and superior frontal gyrus (res = .15), as well as fusiform (res = -.22) and entorhinal gyrus (res = -.17).

Network analysis indicated that ASD showed a stronger negative association with BD, suggesting that BD is a bridge node between ASD and the other diseases. Thus, the brain alterations in BD may connect ASD alterations to the other nodes, and if the BD node is removed, the observed shared variance with SCZ, OCD, and MDD is likely to decrease. Interestingly, after controlling for the other diseases, ASD exhibited a positive correlation with SCZ in contrast with the PCA and Opel *et al.*[2] results. ADHD had a weaker partial correlation with the other nodes that were linked by MDD. SCZ, BD, and OCD maintained the pattern of strong positive associations (Figure 2D). Centrality measures of the network indicated that SCZ had the highest strength (direct connections to other diseases), BD had the highest closeness (indirect connection to other diseases). BD and MDD had the highest betweenness (the average path between two other diseases),[5] see Figure 2E.

Discussion and conclusion

Our approach suggests that the brain abnormalities found in ASD co-varied with MDD, BD, SCZ, and OCD, thus indicating that ASD is not independent as portraited by the original three-component solution in Opel *et al.*[2] Indeed, the well-observed pattern of co-occurrence and clinical overlap among ASD and other psychiatric disorders indicates that they share important pathogenic brain mechanisms and risk factors.[12,13] The small association of ASD and the specificity of ADHD abnormalities might be explained by the fact that these are neurodevelopmental disorders, although recent work has supported the idea that the etiology of BD and SCZ also involves a substantial neurodevelopmental basis.[14,15] An alternative and interesting view argue that these disorders could be conceptualized as a neurodevelopmental continuum, in which the symptoms would reflect the severity, timing, and pattern of brain abnormalities, as well as the modulatory effects of genetic and environmental factors.[14] The results we obtained here seem to partially reflect Owen and O'Donavan's[14] theoretical understanding of this phenomenon.

To summarize, Opel *et al.*[2] advance our understanding of brain morphometric features in highly debilitating psychiatric conditions. Notably, different component solutions can lead to different conclusions, but both approaches were categorical in demonstrating the strong alterations of the hippocampus, fusiform gyrus, pallidum, and putamen. These findings may be of special interest given that the overlapped pattern of symptoms among the major psychiatric diseases usually makes it difficult to accurately diagnose and to prescribe tailored treatment. Moreover, the network approach might help to understand specific disease domains of psychopathology. With more neuroimaging studies of psychiatric disorders becoming available, investigations via shared variation and network perspectives are promising venues for understanding the subtypes of shared pathophysiological patterns.

Data availability

Underlying data

The underlying data of the present study is based on Opel *et al.*[2] work and are available on https://doi.org/10.1016/j.biopsych.2020.04.027. The dataset was obtained by copying the effect sizes estimates reported by Opel *et al*[2] on Supplementary Table S2 of their manuscript.[2] The full manuscript was accessed via the McGill Library Portal (https://www.mcgill.ca/library/) on September 19, 2020.

References

1. Xia CH, Ma Z, Ciric R, *et al.*: **Linked dimensions of psychopathology and connectivity in functional brain networks.** *Nat Commun.* 2018; **9**(1): 1–14.
Publisher Full Text | PubMed Abstract | Publisher Full Text | Free Full Text

2. Opel N, Goltermann J, Hermesdorf M, *et al.*: **Cross-Disorder Analysis of Brain Structural Abnormalities in Six Major Psychiatric Disorders: A Secondary Analysis of Mega- and Meta-analytical Findings From the ENIGMA Consortium.** *Biol Psychiatry.* 2020; **88**(9): 678–686.
PubMed Abstract | Publisher Full Text

3. Eckstrand KL: **Shared Versus Disorder-Specific Brain Morphometric Features of Major Psychiatric Disorders in Adulthood.** *Biol Psychiatry.* 2020; **88**(9): e41–e43.
PubMed Abstract | Publisher Full Text

4. Izquierdo I, Olea J, Abad FJ: **Exploratory factor analysis in validation studies: uses and recommendations.** *Psicothema.* 2014; **26**(3): 395–400.
PubMed Abstract | Publisher Full Text

5. Costello AB, Osborne JW: **Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis.** *Pract Assessment, Res Eval.* 2005; **10**(7).
Publisher Full Text

6. Desikan RS, Ségonne F, Fischl B, *et al.*: **An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest.** *Neuroimage.* 2006; **31**(3): 968–980.
PubMed Abstract | Publisher Full Text

7. Fischl B, Salat DH, Busa E, *et al.*: **Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain.** *Neuron.* 2002; **33**(3): 341–355.
PubMed Abstract | Publisher Full Text

8. Horn JL: **A rationale and test for the number of factors in factor analysis.** *Psychometrika.* 1965; **30**(2): 179–185.
PubMed Abstract | Publisher Full Text

9. Revelle W: **psych: Procedures for Personality and Psychological Research.** 2020.
Reference Source

10. Epskamp S, Borsboom D, Fried EI: **Estimating psychological networks and their accuracy: A tutorial paper.** *Behav Res Methods.* 2018; **50**(1): 195–212.
PubMed Abstract | Publisher Full Text | Free Full Text

11. Epskamp S, Cramer AOJ, Waldorp LJ, *et al.*: **Borsboom D. qgraph: Network Visualizations of Relationships in Psychometric Data.** *J Stat Softw.* 2015; **48**(4).
Publisher Full Text

12. Guan J, Cai JJ, Ji G, *et al.*: **Commonality in dysregulated expression of gene sets in cortical brains of individuals with autism, schizophrenia, and bipolar disorder.** *Transl Psychiatry.* 2019; **9**(1).
PubMed Abstract | Publisher Full Text | Free Full Text

13. Lai MC, Kassee C, Besney R, *et al.*: **Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis.** *Lancet Psychiatry.* 2019; **6**(10): 819–829.
PubMed Abstract | Publisher Full Text

14. Owen MJ, O'Donovan MC: **Schizophrenia and the neurodevelopmental continuum: evidence from genomics.** *World Psychiatry.* 2017; **16**(3): 227–235.
PubMed Abstract | Publisher Full Text | Free Full Text

15. Kloiber S, Rosenblat JD, Husain MI, *et al.*: **Neurodevelopmental pathways in bipolar disorder.** *Neurosci Biobehav Rev.* 2020; **112**(February): 213–226.
PubMed Abstract | Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ? ?

Version 2

Reviewer Report 13 April 2023

https://doi.org/10.5256/f1000research.135143.r160849

?

Y. C. Janardhan Reddy
National Institute of Mental Health and Neuro Sciences (NIMHANS), Bengaluru, India
Srinivas Balachander
National Institute of Mental Health and Neuro Sciences (NIMHANS), Bangalore, Karnataka, India

This paper presents a re-analysis of a previous cross-disorder ENIGMA paper, to investigate shared or distinct structural brain measures in SCZ, BD, MDD, OCD, ADHD, and ASD.

The main message of the paper is that a slightly different statistical approach, which might have certain merits/demerits over the original paper's (Opel *et al.* (2020)) approach, can lead to significantly different conclusions.

The paper raises an important issue, and demonstrates how analyses can be performed differently, but we couldn't really "see" how different their findings are. The main findings of the regional abnormalities that are shared between disorders (hippocampus, fusiform gyrus and certain sub-cortical regions) are the same as in the original paper (Opel *et al.* (2020)). Also, if the ASD and ADHD groups loaded negatively on Factor 1, can the authors state, with confidence, that the same regional abnormalities are shared within ASD/ADHD as well? We assume that the answer is no.

While the authors have extensively discussed the "statistical" appropriateness, and advantages of the unidimensional solution, we feel that one should also take clinical relevance and meaningfulness into consideration while choosing the number of components.

Some of the possible confounds, as to why the 4 psychiatric disorders (SCZ, BD, MDD, OCD) may have loaded together - these were only discussed in passing by the original authors (Opel *et al.* (2020)):

- Shared controls between the samples at various sites (especially between SCZ, BD and OCD and between the ASD and ADHD groups)

- Medication use in the adult psychiatric disorder cohorts - the original ENIGMA papers had

also reported results in medication-naive sub-samples (all disorders except BD). These indicated that medication status had a very significant impact on the brain morphometry

- Comorbidity within original samples, that was not accounted for in the main analysis (OCD with comorbid MDD, schizo-affective disorder, etc)

While the authors' objective was purely to look at statistical considerations, the authors could also suggest how some of the above clinical factors could be taken into account, in future directions.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am a psychiatrist working on obsessive-compulsive disorder (also my co-reviewer). We are not experts in statistical analysis, therefore a statistical review may be required.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Reviewer Report 21 February 2023

https://doi.org/10.5256/f1000research.135143.r160851

?

Esther Walton
University of Bath, Bath, UK

Filho and colleagues reanalysed data published by Opel *et al.* to characterize the shared (vs unique) variation in brain structural alterations across six psychiatric disorders. In contrast to Opel who reported a 3-factor solution, the authors of this manuscript identify a unidimensional solution. And unlike Opel who reported variation shared across SCZ, BD, OCD and MDD but a unique loading for ASD and ADHD, the current manuscript suggests that there might be a somewhat larger degree of shared variation among all six disorders.

The manuscript presents an interesting exploration of how analysis decisions can impact results / interpretation of findings, and through that is also a valuable contribution to discussions around 'researcher degrees of freedom'.

I have the following comments:

- I suggest citing or reporting in a table the original studies and not just Opel *et al.*, as it is difficult to understand what data the current manuscript / results are based on.

- I assume there is no sample overlap between the original studies (e.g., shared set of controls), which could explain some of the clustering. Might be good to just briefly mention.

- Why is the data only available upon request?

- Why was the data analysed first in R, then SPSS, and then back in R? Wasn't it possible to analyse everything in one workflow / software? Also, I suggest sharing the R code and / or SPSS syntax.

- It was not fully clear to me whether the only difference between Opel and the current work was simply the 1- vs 3-factor solutions. Were all other methods / parameters the same?

- I'm not fully convinced that the results between the two studies are indeed so different. Figure 2B does show a one-factor solution, but it's still quite clear that ASD and ADHD load differently onto this factor compared to the other disorders (i.e. negative or not at all versus positive). This seems to align quite closely with the conclusions made by Opel and colleagues. Can the authors comment on that?

- I'm also not fully convinced that either the one or the other method is superior. The third component (Figure 1) was just minimally below the suggested cutoff of 1 (0.98) and issues around using binary cutoffs are discussed at length in the literature (often in the context of p-values or diagnostic criteria). So, a 3-factor solution is not too far-fetched, in the same way that a one-factor solution can also work with this data. In fact, the first factor in Opel seems to explain the same amount of variance as the one factor presented here (~48%), suggesting convergent findings. Can the authors make it clearer whether they believe that a 3-factor solution is in fact misleading? Also, is it possible to present model fit estimates (like RMSEA, CFI, TLI) for both the one- and three-factor solutions?

- What is a "*stronger shared increase*" / "*shared reduction*"? Stronger compared to what? An explanation on how to interpret these findings would be very helpful.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Structural brain imaging in psychiatry, multi-morbidity / shared signatures across disorders

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 21 June 2022

https://doi.org/10.5256/f1000research.135143.r141125

Takahiro Osada
Department of Neurophysiology, Juntendo University School of Medicine, Tokyo, Japan

The authors have satisfactorily revised the manuscript by addressing the comments raised in the earlier review.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: neuroimaging

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 06 June 2022

https://doi.org/10.5256/f1000research.54649.r139204

?

Takahiro Osada
Department of Neurophysiology, Juntendo University School of Medicine, Tokyo, Japan

The manuscript by de Mendonça Filho *et al*. reanalyzed the data presented in Opel *et al.* to investigate the common brain abnormalities for various psychiatric disorders. Specifically, the authors reconsidered the number of factors in the principal component analysis and demonstrated that a solution of one factor was appropriate while the original study adopted three factors. Based on the unidimensional solution, brain abnormalities in ASD showed a negative component loading, whereas those in MDD, BD, SCZ, and OCD showed positive component loadings. The authors concluded that the same component could explain ASD as well as MDD, BD, SCZ, and MDD. Understanding brain abnormalities related to psychiatric disorders is clinically important, and using a large database of the ENIGA consortium is a plausible way for investigation. However, I have some concerns with the analyses and interpretations of the results in this manuscript.

Major comments:

1. The authors compared the number of components in the principal component analysis

solely based on eigenvalues (Figure 1). I wonder whether the unidimensional analysis explained sufficiently the variance of brain regional effect sizes. According to the original study by Opel *et al.*, the three factors explained 85.14 % of the variance in total. To what extent did this unimodal solution could explain the variance?

2. In the unidimensional analysis, ASD showed a negative component loading while MDD, BD, SCN, and OCD showed positive component loadings. Based on this finding, the authors stated in the Abstract and Discussion that brain abnormalities in ASD had more similarities with the four disorders than the original study. The authors should describe this conclusion with caution. The pattern of brain abnormalities in ASD was not "similar", but rather opposite to that in the four disorders because of the negative and positive values. The same component could explain ASD and the four disorders rather than the two distinct factors suggested in the original study.

3. The authors calculated the partial correlations among the residuals and conducted network analyses (Figures 2D and 2E). I wonder whether these analyses were valid. It would be reasonable to examine the relationships between the residuals after the component with the same tendency was extracted (i.e., the same sign and nearly the same amount of loading). However, the authors compared the residuals between the disorders for negative loadings (ASD and ADHD) and positive loadings (MDD, BD, SCN, and OCD). The network analysis seems interesting, but what is the rationale for examining the residuals from data with different tendencies?

Minor comments:

1. Figure 2A was not cited in the text, and how was the "predicted component score" calculated?

2. For evaluating component loadings, why was "λ" used rather than "r"? Is this different from a simple correlation coefficient?

3. It would be helpful to provide a figure for the regional component score map.

4. In the third paragraph in the Results section, the authors stated, "Regional specificities for MDD, BD, SCZ, and OCD were similar to Opel *et al.* results". Could you elaborate on this?

5. In the fourth paragraph in the Results section, the authors stated, "BD is a bridge node between ASD and the other disease". Please elaborate on this.

6. Figure 2B: it would be helpful to explain the numbers in the figure legend. I assume these are component loadings.

7. Figure 2C and 2D: the color coding is somewhat counterintuitive. In Figure 2A, red means represents while blue represents negative. In Figure 2D, it would be helpful to add a color scale bar.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: neuroimaging

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 14 Jun 2022

Euclides Jose de Mendonca Filho

We thank the reviewer for the thorough revision of our manuscript. Bellow, we address the comments to improve readers' understanding of our findings.

Major comments:
1. The authors compared the number of components in the principal component analysis solely based on eigenvalues (Figure 1). I wonder whether the unidimensional analysis explained sufficiently the variance of brain regional effect sizes. According to the original study by Opel et al., the three factors explained 85.14% of the variance in total. To what extent did this unimodal solution could explain the variance?

Although the unimodal solution explains less of the variance, the consideration of a unidimensional solution was based on three criteria:1) the scree plot of eigenvalues, 2) the exhaustive parallel analysis procedure that identified one component as a recurrent robust solution, and 3) the inspection of the Opel *et al.* solution that indicated that the second and third components were defined by only one variable each. The inspection of the original 3-factors solution was the motivator for the reanalysis presented in our manuscript. Since principal component analysis is a data reduction technique, considering components with only one variable seemed redundant, and an indication that the number of factors extracted should be reconsidered. We added the explained variance in the text.

Now it reads:
"The unidimensional solution explained 48.3% percent of the six-brain structural alterations and indicated that the effect size of the differences in brain structure between ASD patients and controls shared a significant negative component loading (λ = -0.30, p = 0.04) with brain abnormalities of the remaining five diseases (Figure 2). Although the unimodal solution explains less of the variance than the original work (the three-factor solution explained 85.14% of variance), the consideration of a unidimensional solution was based on several criteria. Namely, the scree plot of eigenvalues, the exhaustive parallel analysis procedure that identified one component as a recurrent robust solution, and the inspection of the Opel *et al*.[2] solution that indicated additional second and third components defined by only one variable each. Since principal component analysis is a data reduction technique, considering components with only one variable seemed redundant, and an indication that the number of factors extracted should be reconsidered."

2. In the unidimensional analysis, ASD showed a negative component loading while MDD, BD, SCN, and OCD showed positive component loadings. Based on this finding, the authors stated in the Abstract and Discussion that brain abnormalities in ASD had more similarities with the four disorders than the original study. The authors should describe this conclusion with caution. The pattern of brain abnormalities in ASD was not "similar", but rather opposite to that in the four disorders because of the negative and positive values. The same component could explain ASD and the four disorders rather than the two distinct factors suggested in the original study.

We agree with the reviewer that the observed pattern is not similar but is negatively associated with alterations in SCZ, BD, MDD, and OCD. We clarified this information throughout the manuscript.

3. The authors calculated the partial correlations among the residuals and conducted network analyses (Figures 2D and 2E). I wonder whether these analyses were valid. It would be reasonable to examine the relationships between the residuals after the component with the same tendency was extracted (i.e., the same sign and nearly the same amount of loading). However, the authors compared the residuals between the disorders for negative loadings (ASD and ADHD) and positive loadings (MDD, BD, SCN, and OCD). The network analysis seems interesting, but what is the rationale for examining the residuals from data with different tendencies?

Our framework is based on the RDoC initiative that acknowledges that common brain abnormalities can potentially explain psychiatry comorbidities. We aimed at investigating patterns of shared variance independently of the direction of observed brain alterations. PCA and Network analysis are both appropriate for modelling opposite tendencies, and the use of network analysis is aimed at further interrogating the relationship between brain alterations by investigating the function of the observables in the network. This allowed us to obtain information about which observables are dominant in the network in terms of the strength of relations with other observables, for instance.

Minor comments:
Figure 2A was not cited in the text, and how was the "predicted component score"

calculated?

Thank you for bringing this to our attention, we added Figure2A to the 11th paragraph. We also added information about the estimation of the component score.

Now it reads:
"The principal components algorithm was used for dimensional extraction, and component scores were obtained using the regression weights (Thurstone method), which are obtained by multiplying the inverse of the observed variable correlation matrix by the matrix of factor loadings."

1. For evaluating component loadings, why was "λ" used rather than "r"? Is this different from a simple correlation coefficient?

The λ symbol represents a factor or component loading, thus it is not a Pearson correlation coefficient, although λ reported are standardized yielding a similar interpretation to a simple correlation coefficient.

2. It would be helpful to provide a figure for the regional component score map.

This information is depicted in Figure 2A, we added a sentence to clarify this.

3. In the third paragraph in the Results section, the authors stated, "Regional specificities for MDD, BD, SCZ, and OCD were similar to Opel et al. results". Could you elaborate on this?

Thank you for bringing this to our attention. The sentence now reads:
"Regional specificities for MDD, BD, SCZ, and OCD were similar to Opel *et al*. results, in which large absolute residuals were found in the rostral ACC and the medial orbito frontal cortex (OFC) for MDD, parahippocampal gyrus and pallidum for BD and superior temporal gyrus and medial OFC for SCZ."

4. In the fourth paragraph in the Results section, the authors stated, "BD is a bridge node between ASD and the other disease". Please elaborate on this.

Thank you for bringing this to our attention. The sentence now reads:
"Network analysis indicated that ASD showed a stronger negative association with BD, suggesting that BD is a bridge node between ASD and the other diseases. Thus, the brain alterations in BD may connect ASD alterations to the other nodes, and if the BD node is removed, the observed shared variance with SCZ, OCD, and MDD is likely to decrease"

5. Figure 2B: it would be helpful to explain the numbers in the figure legend. I assume these are component loadings.

The reviewer is correct, we added a sentence to grant a better understanding of Figure 2A

6. Figure 2C and 2D: the color coding is somewhat counterintuitive. In Figure 2A, red means represents while blue represents negative. In Figure 2D, it would be helpful to add a color

scale bar.

Thank you for your suggestion, we altered Figures 2C and 2D for clarity.

Competing Interests: The authors have no competing interests to disclose.