Susceptibility or Resilience to Maltreatment Can Be Explained by Specific Differences in Brain Network Architecture

Supplemental Information

Recruitment

Briefly, subjects were recruited by advertisements. Potentially interested participants were phone screened to be medically healthy, right handed, unmedicated (except for contraceptives, hormone replacement and occasional use of albuterol inhalers or non-sedating antihistamines) and between 18-25 years of age. Those who appeared eligible were invited to log onto a HIPAA-compliant online enrollment system to provide detailed information on demographics, medical and psychiatric history, developmental history, life experiences, psychiatric symptomatology and history of childhood maltreatment. Those that meet criteria were invited to the laboratory for further evaluation.

Exclusion criteria included: history of neurologic disease, concussion or head trauma resulting in loss of consciousness for more than 5 minutes, multiple unrelated forms of adversity including natural disaster, motor vehicle accidents, near drowning, house fire, mugging, witnessing or experiencing war, gang violence or murder, riot, or assault with a weapon or animal attack. Additionally, high levels of drug or alcohol use were grounds for exclusion.

Overall, 2188 subjects provided on-line information. From this group we interviewed 670 subjects, and from this interviewed pool we selected the neuroimaging sample. The large number of subjects screened was not specifically required to provide the required number of participants with moderate-to-high levels of maltreatment. Most online screened subjects were eliminated due to histories of head injury / possible concussion, exposure to multiple types of

trauma (e.g., natural disasters, motor vehicle accidents), prematurity or birth complications and binge drinking.

Subjects received $25 for completing the online assessment, $100 per interview and assessment session and $100 for a one-hour MRI protocol.

Additional Information – Maltreatment and Abuse Chronology of Exposure (MACE) Scale

MACE provides ratings on 10 types of abuse and neglect: 1) emotional neglect; 2) physical neglect; 3) witnessing interparental violence; 4) witnessing violence towards siblings; 5) sexual abuse; 6) parental verbal abuse; 7) parental non-verbal emotional abuse; 8) parental physical abuse; 9) peer emotional abuse and 10) peer physical bullying. The MACE provides two measures of overall exposure – MACE Severity (range 0-100), which indicates composite degree of exposure, and MACE Multiplicity (range 0 – 10), which indicates the number of different types of maltreatment experienced. MACE provides excellent overall reliability and good to excellent reliability at each age and to each type of maltreatment. MACE Severity correlated 0.738 with Childhood Trauma Questionnaire (CTQ) score and MACE Multiplicity correlated 0.698 with the ACE scale scores. MACE accounted for 2.00- and 2.07-fold more of the variance in psychiatric symptom ratings than CTQ or ACE, respectively (n=1051) based on variance decomposition. MACE was our primary measure of exposure as it provided data on timing of exposure not available in other instruments. Further, each MACE category fits a Rasch Model which means that each category provides a ‘fundamental measurement’ of exposure in which items are measured on an interval scale with a common unit (1, 2).

MRI Acquisition

Multiple diffusion-weighted images and high-resolution T1-weighted images were acquired using 3T Siemens Trio with 32-channel coil (Siemens Medical Solutions, Erlangen, Germany).

An image analyst observed all scans and sequences with discernible motion artifacts were recollected. Multiple diffusion-weighted images were acquired in 72 directions. Scan parameters were: b=1000 sec/mm2; echo time (TE)/repetition time (TR)=81 msec/6sec; matrix=128x128 on 240mmx240mm field of view (FOV); slices 3.5mm without gap, to yield voxel size of 1.8mmx1.8mmx3.5mm. Anatomical images were acquired using magnetization prepared rapid gradient echo (MPRAGE) sequence. Scan parameters were: sagittal plane, echo time/repetition time/inversion time/flip 1⁄4 2.74 msec/2.1 sec/ 1.1 sec/121; three-dimensional matrix 256x256x128 on 256x256x170 mm field of view; bandwidth of 48.6 kHz; scan time 4:56.

MRI Analysis and Network Construction

First, all scans were inspected for motion artifacts or incomplete coverage. Second, eddy current corrected DTI data were fit to a diffusion tensor model in order to generate FA images using the FMRIB's Diffusion Toolbox (FDT: FSL FMRIB Software Library). Parcellation was conducted in DTI native space (3-5). Each individual's co-registered structural image was normalized to the Montreal Neurological Institute (MNI) template and transformed back to DTI native space along with the AAL template (5) which was used to parcellate each brain into 90 regional nodes (45 for each hemisphere with the cerebellum excluded). Subsequently, diffusion tensor tractography (DTT) was performed and the number of fiber streams interconnecting each node were estimated with the deterministic tractography method implemented in the Diffusion Toolkit and TrackVis (6) using previously reported criteria (7). Two AAL nodes (ROI) were considered to be connected if the reconstructed fiber streams touch these two regions. Fiber streams connecting nodes were defined as edges. Unweighted networks were created by assigning a weight of 1 for edges between nodes with one or more interconnecting fiber streams and 0 for edges between nodes with no interconnecting fiber streams. All graph theory measures presented were for unweighted networks.

Identifying Nodal Differences between Groups: Elastic Net and Random Forest Regression

To identify nodal differences between symptomatic and asymptomatic participants we used elastic net, a form of penalized regression that includes both lasso (least absolute shrinkage and selection operator) and ridge regularization procedures to define an optimal model with a minimal number of predictor variables. Elastic net classification minimizes overfitting by penalizing estimate coefficients reducing the variance of several of the potential predictors to zero and provides more stable estimates for the other predictors. Second, we used random forest regression with conditional inference trees (RFR-CIT) as a powerful machine learning model (8), which does not assume a linear relationship between predictor variables and outcome.

Random forest regression (RFR) predicts outcome by creating a forest of decision trees with each tree generated from a different subset of the data and constrained in the number of predictor it can consider at each decision point (9). This “wisdom of the crowd” strategy provides superior predictions versus conventional regression techniques (10). The tree structure can also model interactions, does not assume a linear relationship between predictor and response and is highly resistant to collinearity. Variable importance is assessed by permuting each variable, and determining how much this degrades model fit (increase in mean square error) (9). Permuting important predictors decreases fit to a large degree whereas permuting unimportant predictors has little impact. We use a variant of Breiman’s approach with conditional inference trees (8) that rectifies a problem in the estimation of importance of predictors with many versus few levels or categories (8). We also found in Monte-Carlo simulations that it is better able to detect the most important predictors in collinear data sets. Statistical significance of the importance measures was assessed through permutation testing by repeatedly fitting the predictors to randomized outcome data.

Multiclass Prediction

We used random forest classifier to create a multiclass prediction model. As with the random forest regression, the model was built by a set of decision trees from randomly selected subsets of the training data. The votes from all of the decision trees are aggregated to decide the final class of the test object. This set of decision trees was used to predict the class. Predictive accuracy was determined using 10-fold cross-validation.

Between Group Differences in Fractional Anisotropy (FA)

Our hypothesis of maltreatment-related alterations in global network architecture was presaged by findings of alterations in FA in corpus callosum, arcuate fasciculus, superior longitudinal fasciculus, inferior longitudinal fasciculus, uncinate, cingulum bundle and fornix (11, 12). A key concern with tractography is that the performance of the algorithm is affected by divergence of fiber streams from major pathways and by potential crossing of fiber tracts in a given voxel. Further, performance is directly proportional to FA values which are thresholded for ending points. As FA values are sensitive to both genetic and psychiatric disorders network findings may stem from overall FA differences between the groups.

Overall, there were no differences between groups in mean FA for the entire brain ($F_{2, 301}$ = 0.223, p > 0.8) nor for the mean of the skeletonized pathways ($F_{2,301}$ = 0.320, p > 0.7). Including FA as a covariate had no effect on the overall findings. For example, MANOVA comparison for network architecture in resilient-susceptible-control comparison was essentially unchanged by including skeletonized FA as a covariate (PBT = 0.088, $F_{10,552}$ = 2.542, p = .005). Likewise, symptomatic, asymptomatic and control group differences were unaffected by including FA as a covariate (without FA $F_{2,301}$ = 8.836, p = 0.00019; with FA as covariate $F_{2,300}$ = 8.798, p = 0.00019) as were group differences in nodal efficiency (e.g., nodal efficiency of right amygdala without FA $F_{2,305}$ = 4.272, p = 0.0148; with mean skeletonized FA $F_{2,304}$ = 4.222, p = 0.0155).

Lack of influence of FA values is likely the result of analyzing unweighted networks in which nodes were either considered connected (1 or more interconnecting fiber streams) or unconnected (no interconnecting fiber streams). FA values may well exert significant influence on weighted networks.

Correlation Between Measures of Nodal Efficiency and Current Symptom Scores

We also assessed in an exploratory analysis whether there were significant associations between measures of nodal efficiency in the 9 identified nodes and current symptoms of anxiety, depression, anger-hostility and somatization in maltreated individuals after controlling for differences in relevant covariates (including degree of exposure to maltreatment, gender and sociodemographic factors). Increased nodal efficiency (NE) in inferior triangularis, amygdala, supplemental motor area and middle cingulate was associated with higher composite symptoms scores (sum anxiety, depression, somatization and anger-hostility). Nodal efficiency in the inferior triangularis was specifically associated with symptoms of depression ($F_{1,173} = 4.25$, $p = .04$) and somatization ($F_{1,173} = 4.53$, $p < .04$). Amygdala NE was associated with somatization ($F_{1,173} = 3.98$, $p < .05$) and anger-hostility ($F_{1,173} = 8.25$, $p < .005$). Supplementary motor area NE was associated with anxiety ($F_{1,172} = 4.58$, $p < .04$) and somatization ($F_{1,173} = 7.89$, $p < .006$) while NE in the mid portion of the cingulate was associated with somatization ($F_{1,174} = 4.19$, $p < 0.05$).

Predictive Multiclass Model and Potential for Overfitting

As a reviewer pointed out, performing classification based on features which have been selected on the basis that they are already known to show significant differences in the same dataset tends to lead to overfitting (13). This however, depends to a considerable degree on sample size. With small N's the extent of overfitting can be alarming when using a 2-way feature set in a new 2-way classification analysis that contains some of the same participants. With

N=20, and assuming an average AUROC (area under the receiver operating characteristic curve) of 0.7, this approach was estimated (in the 8 models using real data) to overfit by an average of 27%, and in the worst case by 43% (13). However, with N=310 fit using this approach overestimated fit by less than 2% on average, with a worst case of ~ 8% overfit (e.g., AUROC of 0.757 vs AUROC of 0.700).

To further evaluate this concern we assessed the fit of the 5 global architecture and 9 nodal feature model by creating an independent training set (67% of subjects) and test set. The two groups were created by balancing percent of subjects in each class as well as age and gender. The top nine nodes distinguishing symptomatic and asymptomatic participants were selected using RFR-CIT from just the training set and combined with five measures of global network architecture to build a predictive multiclass model that best fit the training set. The test set was then run through this predictive model. The results were almost identical to the cross validated fit showing balanced predicted accuracy of 73.1%, 78.8% and 82.2% for controls, resilient and susceptible groups respectively. Hence, it appears that the three-way cross validated model provided reasonable estimates of fit.

Table S1. Participant's general demographic information

Age (years)	21.7 ± 2.5
Subjects Education (years)	14.7 ± 2.0
Parental Education (years)	15.6 ± 3.2
Financial sufficiency during childhood	
Much less than enough money	5.0%
Less than enough money	20.3%
Enough money	42.5%
More than enough money	28.5%
Much more than enough money	3.2%
Race	
White	68.1%
Asian	14.7%
Black	9.5%
American Indian/Alaska Native/Hawaiian	2.4%
Other	5.2%
Hispanic Ethnicity	14.1%

Table S2. Exploratory analysis of differences in nodal centrality between symptomatic and asymptomatic participants.

Predictor	**Elastic Net**	**Random Forest**		
		Importance	**t-levels**	**p-levels**
Amygdala (R)	2.74	0.38	8.48	$p < 10^{-12}$
Frontal Inferior Triangularis (L)	2.50	0.52	10.4	$p < 10^{-17}$
Mid Cingulum (R)	1.38	0.28	5.53	$p < 10^{-5}$
Paracentral Lobule (R)	1.33	0.48	10.8	$p < 10^{-19}$
Supplemental Motor Area (L)	0.50	0.20	3.76	$p < .05$
Olfactory Cortex (L)	0.0	0.27	5.96	$p < 10^{-6}$
(R)	0.0	0.19	3.85	$p < .01$
Postcentral Gyrus (L)	0.0	0.18	3.76	$p < .05$
(R)	0.0	0.26	5.22	$p < 10^{-4}$

Supplemental References

1. Bond T, bond, Fox CM (2015): Applying the Rasch Model. *Fundamental Measurement in the Human Sciences*, Third ed: Routledge, 2015.
2. Brogden HE (1977): The rasch model, the law of comparative judgment and additive conjoint measurement. *Psychometrika*: Springer-Verlag, pp 631-634.
3. Gong G, He Y, Concha L, Lebel C, Gross DW, Evans AC, et al. (2009): Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cereb Cortex*: Oxford University Press, pp 524-536.
4. Li Y, Liu Y, Li J, Qin W, Li K, Yu C, et al. (2009): Brain anatomical network and intelligence. *PLoS Comput Biol*, pp e1000395.
5. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, pp 273-289.
6. Wang L, Dai Z, Peng H, Tan L, Ding Y, He Z, et al. (2014): Overlapping and segregated resting-state functional connectivity in patients with major depressive disorder with and without childhood neglect. *Human brain mapping*, pp 1154-1166.
7. Ohashi K, Anderson CM, Bolger EA, Khan A, McGreenery CE, Teicher MH (2017): Childhood maltreatment is associated with alteration in global network fiber-tract architecture independent of history of depression and anxiety. *NeuroImage*.
8. Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007): Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC bioinformatics*. 8:25.
9. Breiman L (2001): Random Forests. *Machine Learning*. 45:5-32.
10. Liaw A, and Wiener, M. (2002): Classification and Regression by randomForest. *R News*. 2/3:18-22.
11. Teicher MH, Samson JA (2016): Annual Research Review: Enduring neurobiological effects of childhood abuse and neglect. *J Child Psychol Psychiatry*. 57:241-266.
12. Teicher MH, Samson JA, Anderson CM, Ohashi K (2016): The effects of childhood maltreatment on brain structure, function and connectivity. *Nature reviews Neuroscience*. 17:652-666.
13. Smialowski P, Frishman D, Kramer S (2010): Pitfalls of supervised feature selection. *Bioinformatics*. 26:440-443.