bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Multisite generalizability of schizophrenia diagnosis classification based on functional brain connectivity

Pierre Orban[a,b,c,*,1], Christian Dansereau[a,d,1], Laurence Desbois[b], Violaine Mongeau-Pérusse[b], Charles-Édouard Giguère[b], Hien Nguyen[e], Adrianna Mendrek[b,f], Emmanuel Stip[b,c,g], Pierre Bellec[a,d]

[a] Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal, Montréal, Québec, Canada
[b] Centre de Recherche de l'Institut Universitaire en Santé Mentale de Montréal, Montréal, Québec, Canada
[c] Département de Psychiatrie, Université de Montréal, Montréal, Québec, Canada
[d] Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, Montréal, Québec, Canada
[e] School of Mathematics and Physics, University of Queensland, St. Lucia, Australia
[f] Department of Psychology, Bishop's University, Sherbrooke, Québec, Canada
[g] Centre Hospitalier Universitaire de Montréal, Montréal, Québec, Canada

[*] Corresponding author at: Centre de Recherche de l'Institut Universitaire en Santé Mentale de Montréal, 7331 Hochelaga, Montréal H1N 3V2, QC, Canada. Email: pierre.orban@umontreal.ca. Phone: +1-514-251-4000 #3553

[1] Equally contributed to this work.

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Our objective was to assess the generalizability, across sites and cognitive contexts, of schizophrenia classification based on functional brain connectivity. We tested different training-test scenarios combining fMRI data from 191 schizophrenia patients and 191 matched healthy controls obtained at 6 scanning sites and under different task conditions. Diagnosis classification accuracy generalized well to a novel site and cognitive context provided data from multiple sites were used for classifier training. By contrast, lower classification accuracy was achieved when data from a single distinct site was used for training. These findings indicate that it is beneficial to use multisite data to train fMRI-based classifiers intended for large-scale use in the clinical realm.

1. Introduction

Psychiatrists and other mental health professionals could benefit in the not-so-far future from neuroimaging-based classification tools to assist diagnosis and prognosis in mental illness (Huys et al., 2016). Recent developments in the neuroimaging field have led to a shift from group comparisons based on averaging across subjects to machine learning techniques making prediction at the individual level (Dubois and Adolphs, 2016). In this approach, the emphasis is put on the ability of an algorithm to classify individuals into clinical categories with good generalizability to unseen subjects. Over the last decade, hundreds of studies have successfully classified various psychiatric and neurological disorders based on in vivo brain imaging (reviewed in Arbabshirani et al., 2016; Wolfers et al., 2015). For instance, Arbabshirani et al. (2016) identified 30 published studies that distinguished schizophrenia patients from healthy controls with an average accuracy of 83% using functional magnetic resonance imaging (fMRI), either under task or rest states.

To date, however, the vast majority of classification works in mental illness were performed in a research context, using data from single sites of acquisition. Such findings may not generalize to large-scale clinical settings, with patients being scanned at widely-spread sites and possibly under various mental states. In most cases, the performance of classifiers was only assessed for unseen, test subjects with the exact same characteristics as the sample used for training. Yet, using gender as a proof-of-concept target variable, there was initial evidence that classifiers only poorly generalize to data drawn from other site samples (Huf et al., 2014). The inclusion of data from multiple sites during training improved the classifier performance for data of unseen sites.

In schizophrenia, a study pooling fMRI data from two distinct scanning sites reported similar prediction accuracy levels irrespective of whether test data were drawn from the dataset used for training or not, thus suggesting good generalizability (Ska□tun et al., 2017). However, this result appears at odds with a recent fMRI study in autism that showed poorer accuracy for inter-site than intra-site training/test configurations, depending on the ratio of training set used (Abraham et al. 2017). In the case of inter-site testing, data pooled from 4 sites were used for training the classifier, which was tested on data from a fifth site. Yet, none of these two studies specifically evaluated whether using multisite training data could compensate to some extent for the deleterious effect of inter-site testing, by assuming the actual presence of such an effect. In the present work, we sought to address this question based on fMRI brain connectivity in schizophrenia. Since it is impossible to completely control the variations in mental states in realistic clinical situations, we further promoted the complexity of the classification problem by including data obtained in distinct cognitive task conditions across sites. Mass univariate findings have indicated that cognitive state does not further impact on the nature of functional brain connectivity alterations in schizophrenia (Kaufmann et al., 2016; Orban et al., 2016). However, the potential influence of cognitive context on classification performance in a multivariate analysis should not be rejected.

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2. Methods

2.1. Datasets

Brain imaging data from 6 independent studies were obtained through either the SchizConnect and OpenfMRI data sharing platforms (http://schizconnect.org; https://openfmri.org) or local scanning (Çetin et al., 2014; Gollub et al., 2013; Kogan et al., 2016; Orban et al., 2016; Poldrack et al., 2016; Wang et al., 2016). The 6 datasets differed in terms of both scanning site and cognitive context during fMRI data acquisition (resting-state, emotional memory, Sternberg item recognition paradigm, N-back, task-switching and oddball tasks). Classification analyses included fMRI data from 382 subjects, 191 patients diagnosed with schizophrenia and 191 healthy controls. Subjects provided informed consent to participate in their respective studies and ethics approval was obtained at the site of secondary analysis (Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal, Montréal, Canada).

2.2. Subjects matching

Sample size differed between sites (N = 84, 82, 70, 62, 50 and 34). Site samples were obtained after subjects were selected in order to ensure even proportions of schizophrenia patients and controls within each site (N = 42, 41, 35, 31, 25 and 17 subjects per group) and to reduce between-group differences with regards to gender ratio (75% vs 73% males in controls vs. schizophrenia patients), age distribution (32.3 ± 9.8 vs. 33.4 ± 9.5 years old) and motion levels (average frame displacement = 0.15 ± 0.05 vs 0.17 ± 0.06, see Data preprocessing). Matching of schizophrenia and control subjects was achieved based on propensity scores, using the Optmatch R library version 0.9-7 (https://cran.r-project.org/web/packages/optmatch/index.html). The propensity score associated with each participant was defined by the conditional probability of being in the clinical or control group given the confounding covariates (gender, age and motion). Propensity scores were then used to balance those covariates in the two groups. Although we took great care in matching participants with respect to these factors of no interest, it is very likely that other confounds such as medication in schizophrenia patients impacted the reported findings.

2.3. Data preprocessing

Brain imaging data preprocessing and extraction of functional brain connectomes were performed with the NeuroImaging Analysis Kit version 0.12.17 (NIAK, http://niak.simexp-lab.org). Briefly, preprocessing included slice timing correction, estimation of rigid-body motion within the functional runs, nonlinear coregistration of the structural scan in stereotaxic space, individual coregistration between structural and functional scans, resampling of the functional scans at 3mm isotropic resolution in stereotaxic space, scrubbing of volumes with excessive motion (frame displacement greater > 0.5 mm), regression of confounds (slow time drifts, average of conservative white matter and cerebrospinal fluid masks and motion parameters), and smoothing of functional volumes with a 6 mm isotropic Gaussian blurring kernel. A detailed description of the preprocessing pipeline can be found at http://niak.simexp-lab.org/pipe_preprocessing.html.

Individual functional connectomes included 2016 functional connections between 64 brain parcels. The functional brain parcellation was previously obtained by conducting a bootstrap analysis of stable clusters (BASC, Bellec et al. 2010) on an independent fMRI dataset of 200 healthy young subjects (https://doi.org/10.6084/m9.figshare.1285615.v1). In each schizophrenia or control participant, the time series of a brain parcel consisted in the average of the voxel signals in the parcel. Connectivity measures between pairs of parcels were defined by Pearson product-moment correlation coefficients. Individual connectomes were parcel by parcel (64 x 64) symmetrical matrices that summarized connectivity levels in the whole brain. Lower triangular matrices were then vectorized for all subjects in order to form a subject by connections (382 x 2016) matrix.

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2.4. Data analysis

Classification analyses were performed with a linear support vector machine (SVM) algorithm, as implemented in the SciKit-Learn python library version 0.18.1 (Abraham et al., 2014). The SVM classifier, a supervised classification algorithm, represented subjects as points in space, mapped so that the subjects of the separate clinical labels were divided by a clear gap (called a margin) that was as wide as possible. The hyperparameter C of the SVM was optimized using nested cross-validation. Each model used the residuals from a regression of confounding variables (gender, age and motion parameters) across connections estimated from the subjects selected for training the model. The evaluation metrics were computed using four main values, namely the number of true and false positive (TP, FP) as well as true and false negatives (TN, FN). Sensitivity was defined as TP/(TP+FN), specificity as TN/(TN+FP) and accuracy as (TP+TN)/(TP+FP+TP+FN). The main analyses evaluated the impact on classification accuracy of the number of site(s) (1, 2, 3, 4 or 5) included in the training set. We evaluated this impact in situations where the test set included only subjects from the same site(s) used during training (intra-site test with 10-fold cross validation) or, alternatively, situations where the test set included only subjects from sites not used during training (inter-site test with “leave-site-out” cross validation). Cross validation ensured that the subjects used for training were never used in the test phase.

The statistical significance of changes in accuracy levels as a function of the number of sites used for training and whether data used for testing were drawn from the same dataset(s) used for training (intra-site vs inter-site) was assessed with binary logistic regressions using the GLM function in R version 3.2.5. These analyses relied on the prediction of categorical outcomes (hit/miss data) based on predictor variables (number of sites used for training, intra-site vs inter-site). Significance threshold in the different contrasts was set at $p < 0.05$.

Complementary analyses were conducted. First, we explored differences in whole brain connectivity between schizophrenia patients and controls using mass univariate statistics for the various training site combinations. Similarly for multivariate classification analyses, we extracted feature weights separately for all site combinations. We then examined the level of correspondence across site combinations for both univariate and multivariate analyses. Second, we aimed at demonstrating the presence of multivariate site effects on functional brain connectivity. To this end, we determined accuracy levels for the classification of scanning sites by performing separate SVM analyses for all pairs of sites, using 10-fold cross validation as in the main analyses.

3. Results

3.1. Correspondence across site combinations

We first report patterns of functional brain dysconnectivity in schizophrenia patients based on mass univariate statistics. For the sake of interpretability, the 64 brain parcels were sorted in relation to 7 large-scale brain networks from the same multiscale functional brain atlas (Figure 1a,b). When pooling data from all subjects and sites, a connectome-wide association analysis revealed widespread decreased connectivity in schizophrenia patients (Figure 1c), with 769 out of 2016 connections exhibiting a significant effect after false discovery rate correction ($q^{FDR} < 0.05$). Differences between schizophrenia patients and controls were further examined separately for each unique combination of 1 to 5 training sites (61 possibilities: 1, 2, 3, 4, 5, 6, 1-2, 1-3, 1-4, …, 1-2-4-5-6, 1-3-4-5-6, 2-3-4-5-6). Results revealed high variability in the nature of mass univariate effects across training site combinations, with small correlation between them when there was no overlap between site combinations (Figure 1c). By contrast, large correlations were observed when site combinations overlapped. Weight matrices, which indicate for each connection the importance of that connection in the decision process, were also extracted for the whole sample as well as each site combination in multivariate classification analyses (Figure 1d). The correspondence between site combinations mimicked the patterns of correlations seen

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

for univariate analyses, with a large correlation between weights for site combinations that overlapped but a small correlation otherwise.

3.2. Classification findings

Classification of sites was performed with high accuracy (84%), indicating a significant multivariate impact of scanning site on functional brain connectivity. Training on data from a single site led to a poor generalization of diagnosis classification to subjects drawn from another site, i.e. classification accuracy was much lower in the inter-site than intra-site configuration when only one site was used for training ($p<0.005$). However, increasing the heterogeneity of the training set by including data from different sites improved accuracy of the classifier applied to another unknown scanning site and cognitive context ($p< 5x10^{-8}$) (Figure 2a). This compensatory effect was such that inter-site classification reached similar accuracy performance to intra-site classification when 5 different sites were used for classifier training ($p = 0.56$), thus suggesting excellent generalization in this context. The benefit of using heterogeneous training data when classifying subjects drawn from the same sites as the training set was much more moderate than for the inter-site training-test configuration, yet was significant ($p <0.05$). Formal testing of an interaction effect revealed a significant effect ($p < 0.05$), thus demonstrating that improved generalization on novel sites following multisite training was not merely a consequence of increasing sample size.

4. Discussion

The present findings highlight a prerequisite for an optimal translation of classification tools from the research to clinical realm. Namely, classifier training should be performed on data that are sufficiently representative of sites and/or mental state variations in order to generalize well for large-scale clinical use. In particular, the accuracy scores reported in most of the existing literature should be interpreted with caution, as they only reflect within-site generalizability and may therefore overestimate the accuracy.

Mass univariate analyses evidenced brain dysconnectivity across the entire brain, with significant effects in over a third of brain connections distributed in various large-scale brain networks, from cognitive to primary sensory networks. Abnormally decreased rather than increased functional connectivity in schizophrenia is largely consistent with previous reports in the literature (Pettersson-Yeo et al., 2011). With close to 200 schizophrenia patients and 200 controls, our fMRI connectome-wide association analysis is one of the largest to be reported to date. The present work and previous similar studies conducted in schizophrenia (Cheng et al., 2015; Schilbach et al., 2016; Skatun et al., 2017) underscore the utility of pooling data across multiple sites of acquisition in order to achieve higher sample size and more reliable findings. The marked variability in dysconnectivity patterns detected in each site separately could induce a deleterious effect of multisite data pooling on statistical power as compared to data obtained at a single site. However, there is support to claim that the initial deleterious effect of multisite data pooling can be mitigated by the increase in sample size and number of sites in the context of intra-site mass univariate as well as multivariate analysis (Dansereau et al., 2017). This is in line with our findings suggesting that a similar compensating effect can be obtained using multisite data aggregation in inter-site multivariate prediction, a configuration that is most likely to be found in clinical settings.

Multivariate classification analyses indicate that increasing sample size through multisite data pooling increased diagnosis prediction in schizophrenia, although this effect was of small amplitude. More critically, an additional benefit of including heterogeneous data was that the classifier generalized better to data that were not represented during training, neither in terms of scanning site nor mental content. This demonstration is concordant with a previous report that classified gender as a proof-of-concept application (Huf et al., 2014), and underscores the benefit of pooling multisite data for the purpose of generalizability and clinical use. The observed gain of almost 10% in classification accuracy is appreciable. It is

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

nonetheless noteworthy that the highest accuracy of schizophrenia diagnosis classification was below 70%, which precludes the immediate translation of such machine learning tools in the clinical realm. Beyond the fact that most classification work has to date investigated within-site generalizability, it is notable that most studies relied on small samples. This is likely to be accompanied by a publication bias by which only the most significant findings were published. While the average classification of schizophrenia diagnosis over 30 published studies is above 80%, it was accordingly shown that studies with a large sample size in fact reported lower classification accuracy (Arbabshirani et al., 2016). Besides, low classification accuracy is very likely dependent on the ill-definition of clinical labels, as schizophrenia is a highly heterogeneous psychiatric disorder (Kapur et al., 2012). The stratification of patients into more homogeneous neurobiological subtypes, beyond clinical symptoms, will likely define more precise labels that will lead to improved classification of their diagnosis. The characterization of mental illness heterogeneity through the identification of such different biotypes, in particular based on fMRI brain connectivity, is a topic of burgeoning research in various psychiatric disorders (Costa Dias et al., 2015; Drysdale et al., 2016; Gates et al., 2015).

It is anticipated that neuroimaging-based classification will ultimately assist psychiatrists in not only diagnosis but also prognosis and theragnosis in mental illness, including schizophrenia. The future integration of classifiers into mental health care will require studies with dramatically increased sample size (Dubois and Adolphs, 2016). Most studies indeed suffer from insufficient data, possibly resulting in biased accuracy estimation, under-representation of mental illness heterogeneity and unstable findings (Arbabshirani et al., 2016). Future work will also need to develop novel algorithms with improved capabilities and to better define clinical labels. The present work identifies one specific parameter that will facilitate an optimal translation of supervised machine learning into clinical practice, namely the need to train classifiers on data that are sufficiently representative of heterogeneity with regards to scanning sites and mental contents.

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

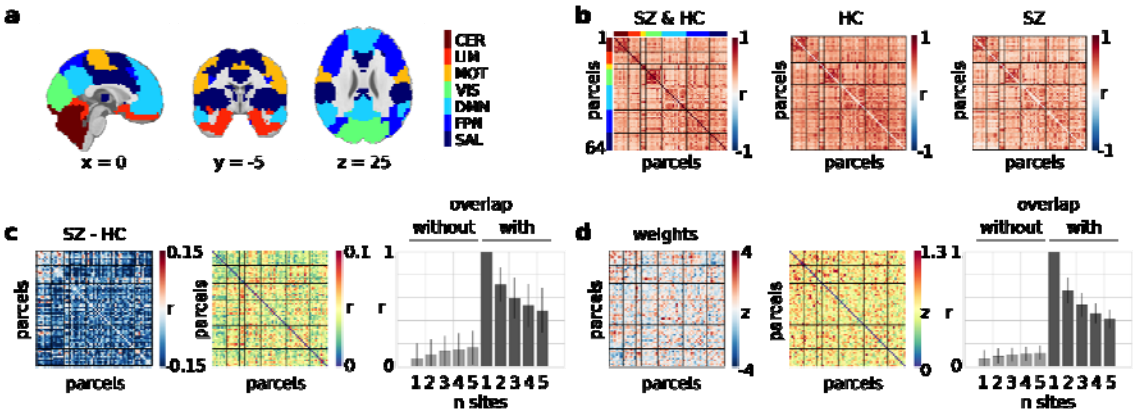

Figure 1. Correspondence between connectomes

Sorting the 64 brain parcels in relation to 7 large-scale brain networks (a) allowed to best reveal the structure of whole-brain connectivity in both schizophrenia patients and controls, here shown when combining data from all sites (b). Results for univariate analyses (c). The left panel reports mean differences in connectivity between schizophrenia patients and controls pooled across all sites. The middle panel shows variations (standard deviation) of between-groups differences across the various site combinations. The right panel reports correlations of univariates effects of schizophrenia between single sites of reference (sites 1 to 6) and various site combinations as a function of whether site combinations included the sites of reference and the number of sites included in the combination (mean and standard deviation across the 6 reference sites). Results for multivariate classification analyses are similarly organized (d). The left panel provides the normalized weights obtained when pooling all subjects from all sites. The middle panel indicates how these weights vary across site combinations (standard deviation). The right panel provides correlations of weight matrices between single sites of reference and various site combinations. Abbreviations for networks are as follows: CER, cerebellum; VIS, visual; LIM, limbic, MOT, motor; SAL, salience; FPN, fronto-parietal; DMN, default-mode.

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

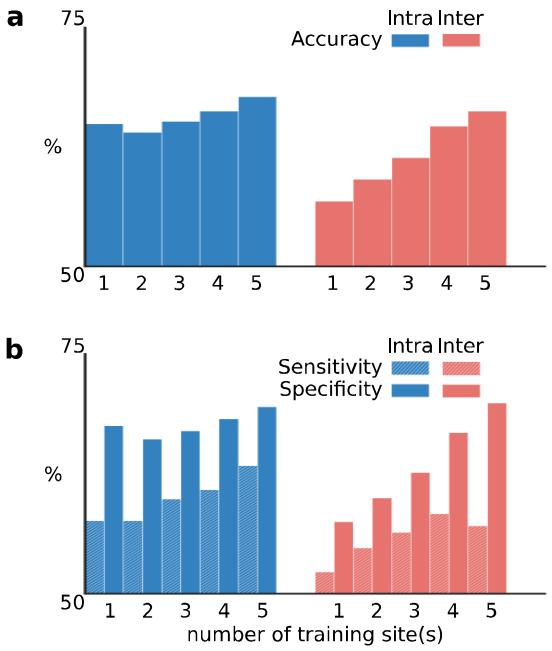

Figure 2. Classification findings
Histograms show the percentages of classification accuracy (a) as well as sensitivity and specificity (b) as a function of the number of sites (1, 2, 3, 4 or 5) from which individual connectomes were drawn for training, and whether testing was performed on subjects drawn from the same site(s) as during training (intra-site test) or not (inter-site test).

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgments

Data from one study (emotional memory task) were collected thanks to grants from the Canadian Institutes of Health Research, Gender and Health Institute to Dr Mendrek (CIHR grant number 200603MOP-158161-GSH-CFCA-130656). Data from 4 other studies were accessed through the SchizConnect platform (http://schizconnect.org). As such, the investigators within SchizConnect contributed to the design and implementation of SchizConnect and/or provided data but did not participate in analysis or writing of this report. Funding of the SchizConnect project was provided by NIMH cooperative agreement 1U01 MH097435. SchizConnect enabled access to the following data repository: the COllaborative Informatics and Neuroimaging Suite Data Exchange tool (COINS; http://coins.mrn.org/dx). Data from one study (resting-state) was collected at the Mind Research Network and funded by a Center of Biomedical Research Excellence (COBRE) grant 5P20RR021938/P20GM103472 from the NIH to Dr Vince Calhoun. Data from two studies (oddball and Sternberg item recognition paradigm tasks) were obtained from the Mind Clinical Imaging Consortium database. The MCIC project was supported by the Department of Energy under award number DE-FG02-08ER6458. MCIC is the result of efforts of co-investigators from University of Iowa, University of Minnesota, University of New Mexico and Massachusetts General Hospital. Data from a fourth study (N-back task) were obtained from the Neuromorphometry by Computer Algorithm Chicago (NMorphCH) dataset (http://nunda.northwestern.edu/nunda/data/projects/NMorphCH) As such, the investigators within NMorphCH contributed to the design and implementation of NMorphCH and/or provided data but did not participate in analysis or writing of this report. The NMorphCH project was funded by NIMH grant RO1 MH056584. The last study (task-switching) was obtained through the OpenFMRI project (http://openfmri.org)from the Consortium for Neuropsychiatric Phenomics (CNP), which was supported by NIH Roadmap for Medical Research grants UL1-DE019580, RL1MH083268, RL1MH083269, RL1DA024853, RL1MH083270, RL1LM009833, PL1MH083271, and PL1NS062410. Data analysis was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC grant number #436141) to PB. CD is supported by a bursary from the Lemaire foundation.

Conflicts of interest

EM is on the scientific advisory boards for Janssen, Otsuka, Lundbeck, Sunovion and BMS. He has also received grants or research support from Roche. PB is a member of an advisory board for Roche and is currently a part-time consultant for two contract research organizations, Biospective Inc. and NeuroRX research. No other competing interests declared.

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Abraham, A., Milham, M.P., Di Martino, A., Craddock, R.C., Samaras, D., Thirion, B., Varoquaux, G., 2017. Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. Neuroimage 147, 736–745.

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., Varoquaux, G., 2014. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 8, 14.

Arbabshirani, M.R., Plis, S., Sui, J., Calhoun, V.D., 2017. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. Neuroimage 145, 137-165.

Bellec, P., Rosa-Neto, P., Lyttelton, O.C., Benali, H., Evans, A.C., 2010. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. Neuroimage 51, 1126–1139.

Cheng, W., Palaniyappan, L., Li, M., Kendrick, K.M., Zhang, J., Luo, Q., Liu, Z., Yu, R., Deng, W., Wang, Q., Ma, X., Guo, W., Francis, S., Liddle, P., Mayer, A.R., Schumann, G., Li, T., Feng, J., 2015. Voxel-based, brain-wide association study of aberrant functional connectivity in schizophrenia implicates thalamocortical circuitry. Npj Schizophrenia 1:15016.

Costa Dias, T.G., Iyer, S.P., Carpenter, S.D., Cary, R.P., Wilson, V.B., Mitchell, S.H., Nigg, J.T., Fair, D.A., 2015. Characterizing heterogeneity in children with and without ADHD based on reward system connectivity. Dev Cogn Neurosci 11, 155–174.

Çetin, M.S., Christensen, F., Abbott, C.C., Stephen, J.M., Mayer, A.R., Cañive, J.M., Bustillo, J.R., Pearlson, G.D., Calhoun, V.D., 2014. Thalamus and posterior temporal lobe show greater inter-network connectivity at rest and across sensory paradigms in schizophrenia. Neuroimage 97, 117–126.

Dansereau, C., Benhajali, Y., Risterucci, C., Pich, E.M., Orban, P., Arnold, D., Bellec, P., 2017. Statistical power and prediction accuracy in multisite resting-state fMRI connectivity. Neuroimage 149, 220–232.

Drysdale, A.T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., Fetcho, R.N., Zebley, B., Oathes, D.J., Etkin, A., Schatzberg, A.F., Sudheimer, K., Keller, J., Mayberg, H.S., Gunning, F.M., Alexopoulos, G.S., Fox, M.D., Pascual-Leone, A., Voss, H.U., Casey, B.J., Dubin, M.J., Liston, C., 2017. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med 23, 28-38.

Dubois, J., Adolphs, R., 2016. Building a Science of Individual Differences from fMRI. Trends Cogn Sci 20, 425-443.

Gates, K.M., Molenaar, P.C.M., Iyer, S.P., Nigg, J.T., Fair, D.A., 2014. Organizing heterogeneous samples using community detection of GIMME-derived resting state functional networks. PLoS ONE 9, e91322.

Gollub, R.L., Shoemaker, J.M., King, M.D., White, T., Ehrlich, S., Sponheim, S.R., Clark, V.P., Turner, J.A., Mueller, B.A., Magnotta, V., O'Leary, D., Ho, B.C., Brauns, S., Manoach, D.S., Seidman, L., Bustillo, J.R., Lauriello, J., Bockholt, J., Lim, K.O., Rosen, B.R., Schulz, S.C., Calhoun, V.D., Andreasen, N.C., 2013. The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. Neuroinformatics 11, 367–388.

Huf, W., Kalcher, K., Boubela, R.N., Rath, G., Vecsei, A., Filzmoser, P., Moser, E., 2014. On the generalizability of resting-state fMRI machine learning classifiers. Front Hum Neurosci 8, 502.

Huys, Q.J.M., Maia, T.V., Frank, M.J., 2016. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci 19, 404–413.

Kapur, S., Phillips, A.G., Insel, T.R., 2012. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? Mol Psychiatry 17, 1174–1179.

Kaufmann, T., Alnæs, D., Brandt, C.L., Doan, N.T., Kauppi, K., Bettella, F., Lagerberg, T.V., Berg, A.O., Djurovic, S., Agartz, I., Melle, I.S., Ueland, T., Andreassen, O.A., Westlye, L.T., 2017. Task modulations and clinical manifestations in the brain functional connectome in 1615 fMRI datasets. Neuroimage 147, 243–252.

bioRxiv preprint doi: https://doi.org/10.1101/141192; this version posted May 23, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Kogan, A., Alpert, K., Ambite, J.L., Marcus, D.S., Wang, L., 2016. Northwestern University schizophrenia data sharing for SchizConnect: A longitudinal dataset for large-scale integration. Neuroimage 124, 1196–1201.

Orban, P., Desseilles, M., Mendrek, A., Bourque, J., Bellec, P., Stip, E., 2016. Altered brain connectivity in patients with schizophrenia is consistent across cognitive contexts. J Psychiatry Neurosci 41, 150247.

Pettersson-Yeo, W., Allen, P., Benetti, S., McGuire, P., Mechelli, A., 2011. Dysconnectivity in schizophrenia: where are we now? Neurosci Biobehav Rev 35, 1110–1124.

Poldrack, R.A., Congdon, E., Triplett, W., Gorgolewski, K.J., Karlsgodt, K.H., Mumford, J.A., Sabb, F.W., Freimer, N.B., London, E.D., Cannon, T.D., Bilder, R.M., 2016. A phenome-wide examination of neural and cognitive function. Scientific Data, 49 3,

Schilbach, L., Hoffstaedter, F., Müller, V., Cieslik, E.C., Goya-Maldonado, R., Trost, S., Sorg, C., Riedl, V., Jardri, R., Sommer, I., Kogler, L., Derntl, B., Gruber, O., Eickhoff, S.B., 2016. Transdiagnostic commonalities and differences in resting state functional connectivity of the default mode network in schizophrenia and major depression. Neuroimage Clin 10, 326–335.

Skatun, K.C., Kaufmann, T., Doan, N.T., Alnæs, D., Córdova-Palomera, A., Jönsson, E.G., Fatouros-Bergman, H., Flyckt, L., KaSP, Melle, I., Andreassen, O.A., Agartz, I., Westlye, L.T., 2016. Consistent Functional Connectivity Alterations in Schizophrenia Spectrum Disorder: A Multisite Study. Schizophr Bull. doi:10.1093/schbul/sbw145

Wang, L., Alpert, K.I., Calhoun, V.D., Cobia, D.J., Keator, D.B., King, M.D., Kogan, A., Landis, D., Tallis, M., Turner, M.D., Potkin, S.G., Turner, J.A., Ambite, J.L., 2016. SchizConnect: Mediating neuroimaging databases on schizophrenia and related disorders for large-scale integration. Neuroimage 124, 1155–1167.

Wolfers, T., Buitelaar, J.K., Beckmann, C.F., Franke, B., Marquand, A.F., 2015. From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. Neurosci Biobehav Rev 57, 328–349.