

## SYSTEMATIC REVIEW OPEN



# A meta-analysis and systematic review of single vs. multimodal neuroimaging techniques in the classification of psychosis

Alexis Porter<sup>1</sup>✉, Sihan Fei<sup>1</sup>, Katherine S. F. Damme<sup>1,2</sup>, Robin Nusslock<sup>1</sup>, Caterina Gratton<sup>3,7</sup> and Vijay A. Mittal<sup>1,2,4,5,6,7</sup>

© The Author(s) 2023

**BACKGROUND:** Psychotic disorders are characterized by structural and functional abnormalities in brain networks. Neuroimaging techniques map and characterize such abnormalities using unique features (e.g., structural integrity, coactivation). However, it is unclear if a specific method, or a combination of modalities, is particularly effective in identifying differences in brain networks of someone with a psychotic disorder.

**METHODS:** A systematic meta-analysis evaluated machine learning classification of schizophrenia spectrum disorders in comparison to healthy control participants using various neuroimaging modalities (i.e., T1-weighted imaging (T1), diffusion tensor imaging (DTI), resting state functional connectivity (rs-FC), or some combination (multimodal)). Criteria for manuscript inclusion included whole-brain analyses and cross-validation to provide a complete picture regarding the predictive ability of large-scale brain systems in psychosis. For this meta-analysis, we searched Ovid MEDLINE, PubMed, PsychInfo, Google Scholar, and Web of Science published between inception and March 13th 2023. Prediction results were averaged for studies using the same dataset, but parallel analyses were run that included studies with pooled sample across many datasets. We assessed bias through funnel plot asymmetry. A bivariate regression model determined whether differences in imaging modality, demographics, and preprocessing methods moderated classification. Separate models were run for studies with internal prediction (via cross-validation) and external prediction.

**RESULTS:** 93 studies were identified for quantitative review (30 T1, 9 DTI, 40 rs-FC, and 14 multimodal). As a whole, all modalities reliably differentiated those with schizophrenia spectrum disorders from controls (OR = 2.64 (95%CI = 2.33 to 2.95)). However, classification was relatively similar across modalities: no differences were seen across modalities in the classification of independent internal data, and a small advantage was seen for rs-FC studies relative to T1 studies in classification in external datasets. We found large amounts of heterogeneity across results resulting in significant signs of bias in funnel plots and Egger's tests. Results remained similar, however, when studies were restricted to those with less heterogeneity, with continued small advantages for rs-FC relative to structural measures. Notably, in all cases, no significant differences were seen between multimodal and unimodal approaches, with rs-FC and unimodal studies reporting largely overlapping classification performance. Differences in demographics and analysis or denoising were not associated with changes in classification scores.

**CONCLUSIONS:** The results of this study suggest that neuroimaging approaches have promise in the classification of psychosis. Interestingly, at present most modalities perform similarly in the classification of psychosis, with slight advantages for rs-FC relative to structural modalities in some specific cases. Notably, results differed substantially across studies, with suggestions of biased effect sizes, particularly highlighting the need for more studies using external prediction and large sample sizes. Adopting more rigorous and systematized standards will add significant value toward understanding and treating this critical population.

*Molecular Psychiatry* (2023) 28:3278–3292; <https://doi.org/10.1038/s41380-023-02195-9>

## INTRODUCTION

Psychosis is a devastating and heterogeneous disorder with a poorly understood etiology [1–5]. Psychosis symptoms are thought to emerge from network-level abnormalities within the brain as opposed to disruptions in one discrete location [6–12]. Consistent with the neurodevelopmental theories and stress diathesis models of psychosis, whole-brain structural abnormalities such as impaired myelination [13–15], and accelerated

demyelination have been linked with symptom severity and deficits in cognitive function [16, 17]. There are signs of progressive degeneration of other structural measures such as cortical thickness [18] and gray matter volume [19] linked to psychosis. Psychotic disorders have also been characterized as a disruption in the functional communication between brain regions [20, 21] and alterations in the functional strength of connections [22, 23]. These findings suggest that psychosis is

<sup>1</sup>Department of Psychology, Northwestern University, Evanston, IL, USA. <sup>2</sup>Institute for Innovations in Developmental Sciences, Northwestern University, Evanston and Chicago, IL, USA. <sup>3</sup>Department of Psychology, Florida State University, Tallahassee, FL, USA. <sup>4</sup>Department of Psychiatry, Northwestern University, Chicago, IL, USA. <sup>5</sup>Medical Social Sciences, Northwestern University, Chicago, IL, USA. <sup>6</sup>Institute for Policy Research, Northwestern University, Chicago, IL, USA. <sup>7</sup>These authors contributed equally: Caterina Gratton, Vijay A. Mittal. ✉email: alexis.porter1313@gmail.com

Received: 3 October 2022 Revised: 11 July 2023 Accepted: 17 July 2023

Published online: 10 August 2023

characterized by a combination of structural and functional dysfunction across distributed brain systems [8, 23–25].

Non-invasive neuroimaging methods can be used to measure structural (T1-weighted imaging, diffusion imaging) and functional (resting-state functional connectivity) brain networks. Given the link between psychosis and brain system dysfunction, one may ask which specific neuroimaging modalities are best suited for diagnostic purposes, or if a combination of multiple modalities would allow a more holistic and accurate classification of the disorder. While a number of studies have begun to probe this question, this idea has not been tested in a systematic review. This meta-analytic study was designed to directly compare neuroimaging methods (T1-weighted imaging, diffusion imaging, and resting-state functional connectivity) and their combination (multimodal approaches) in their ability to classify psychosis from healthy controls using machine learning data from whole-brain networks. Additionally, this review evaluated whether various statistical, methodological, and demographic information had any moderating effects on classification.

In psychosis, a reduction in gray matter volume and enlargement of the ventricles has been reported through the use of T1-weighted imaging (abbreviated as 'T1' in this manuscript) [26]. Although, this may be due to neurotoxic effects related to medications [27], some evidence suggests volumetric differences are present in never medicated and first-episode patients [28]. This would suggest that gray matter abnormalities may be a risk factor leading up to the onset of psychosis or primary aspects of its etiology. Reductions in gray matter can vary with time and are not always consistent across people [19]. Prior work has demonstrated that gray matter cortical thickness declines with age in participants with psychosis at a higher rate compared to controls, particularly in regions important for cognitive function such as inferior frontal cortex, anterior cingulate cortex, and lateral temporal cortex (for review see [18]).

White matter abnormalities such as decreased expression of oligodendrocytes have been associated with psychosis [29, 30]. Diffusion tensor imaging (DTI) is a non-invasive measure of the myelin integrity of underlying white matter [31–33]. Researchers have found that measures of white matter integrity decrease at higher rates in psychosis compared to controls across the lifespan [16, 17]. The development of myelination also tends to proceed and co-occur with the emergence of symptoms of psychosis during the adolescent time period [34]. This work provides evidence of developmental abnormalities that might be linked with psychosis-specific accelerated aging of white matter pathways. However, the location of disruption in the integrity of white matter pathways has remained inconsistent [35, 36]. A recent meta-analytic study aimed at evaluating white matter integrity in high-risk individuals found significant variation in the integrity of white matter pathways across large tracts such as superior longitudinal fasciculus, inferior longitudinal fasciculus, and inferior fronto-occipital fasciculus [35]. These abnormalities tended to vary across study design and have not been consistently linked to variation in symptom severity.

Resting-state functional connectivity (rs-FC) is a non-invasive way to evaluate large-scale functional networks across brain regions. Changes in these networks may be associated with genetic, cognitive, and developmental factors of psychosis. These functional changes have been found to be more closely linked to the expression of behavioral symptoms used in diagnosis [21, 37]. Rs-FC alterations in networks related to higher-order processes such as attention and executive control [21, 38–43] have particularly been highlighted in individuals with psychosis. However, these results are often inconsistent in the direction and location of dysfunction [23, 44]. The inconsistency in the rs-FC literature could be due to individual variation in clinical characteristics: researchers that have evaluated schizophrenia from an individual-specific approach have identified key characteristics linked to symptoms and behavior

[45–48]. After accounting for these variations, rs-FC may serve as a key factor in distinguishing characteristics that are specific to psychosis in machine learning classification.

The neuroimaging approaches described above have helped to uncover key neurobiological associations of psychosis. However, there are limitations in each non-invasive method for measuring brain systems [49]. One potential solution is to use a multimodal approach in which different imaging modalities are combined. This multimodal approach may bridge the relationship between gray matter, white matter, and functional features of brain networks that would otherwise be lost when evaluating a single modality. Prior work has suggested that using multiple imaging modalities provides a sensitive approach to identify converging areas of dysfunction in schizophrenia [50–52].

To determine the utility of each method in understanding neurobiological features of schizophrenia, we focus here on machine learning approaches. These methods can use multivariate information to identify subtle variations in the brain that may not otherwise be captured using standard univariate methods [53, 54]. Imaging modalities can also be used as features to classify various forms of psychiatric disorders [55]. However, there are several important factors to consider when using machine learning methods with neuroimaging data including improper cross-validation [56], small sample sizes [37] and physiological artifacts [57–60] that are known to produce inflated or misrepresented classification results.

Here, we completed a systematic review and meta-analysis to determine to what extent neuroimaging methods can classify individuals with psychosis. Specifically, we asked whether any method (or their combination) outperforms others in the ability to distinguish participants diagnosed with a schizophrenia spectrum disorder from healthy controls in the context of machine learning classification. We used a bivariate random-effects model assessing the sensitivity and specificity in each study [61]. To reduce the potential for inflated results, we opted for a strict set of criteria for manuscript extraction including cross-validation. Additionally, we evaluated whether other variables moderate the metrics associated with classification such as preprocessing technique, statistical methods, sample size, and participant characteristics.

## METHODS

### Overview

This meta-analysis was conducted following the preferred reporting guidelines for systematic reviews and meta-analysis (PRISMA) [62, 63]. Search criteria limited the analysis to studies that applied classification algorithms to predict clinical status in psychosis participants who met criteria for subtypes within schizophrenia spectrum disorders relative to healthy controls (psychosis v. healthy control). This meta-analysis includes estimates of sensitivity and specificity as calculated based on confusion matrices. A bivariate approach and hierarchical summary receiver operating characteristics (ROC) model were used to estimate sensitivity and specificity across studies. Additionally, we conducted a meta-regression analysis to examine differences between datasets that may contribute to variability found between imaging subgroups (e.g., participant characteristics, statistical methods, and quality of preprocessing methods).

### Search strategy

We searched databases Ovid MEDLINE, PubMed, PsychInfo, Google Scholar, and Web of Science for relevant, peer-reviewed publications. Databases were searched from inception until March 13th 2023. Titles and abstracts were searched using the following keywords: (Schizo\* or psychosis or psychotic) AND/OR (DTI or DSI or white matter or fractional anisotropy or FA) AND/OR (fMRI or functional connectivity or network or resting state or rsfMRI or circuit) AND/OR (structural or T1 or anatomical) AND

(support vector or SVM or classification or categorization or machine learning). We included advanced search terms to only evaluate studies written in English that included human subjects. In the case of insufficient data, authors were contacted via email to provide additional information.

### Study selection

All titles and abstracts of identified publications were screened by authors A.P. and S.F. for eligibility. Articles had to meet the following inclusion criteria: (1) studies had to apply a machine learning classification model to predict clinical status using neuroimaging data as features. (2) Studies were required to have some form of cross-validation (e.g., leave-one-out, k-folds, test-train split) or external dataset validation. Results from internal cross-validation and external dataset validation were separated in analyses. (3) Clinical participants had to meet a diagnosis for a psychotic disorder following the diagnostic statistics manual (DSM) or the international classification of diseases (ICD). This included first-episode psychosis, first-episode schizophrenia, schizophrenia, schizophrenia with comorbidity, and schizophrenia spectrum disorder. (4) Given our focus on large-scale brain systems, we restricted ourselves to studies that included whole-brain analyses, excluding those focused on single regions or networks. All potential studies were carefully screened for review. If a study included both region-specific and whole-brain analysis, then the whole-brain results were kept for reporting. (5) Classification was based on at least one of the following imaging types: DTI, rs-FC, T1, or some combination. As we were interested in focusing on intrinsic brain networks rather than task modulations, task-based FC studies and dynamic FC studies were excluded from this analysis. Use of a simulated or synthetically created dataset were also grounds for exclusion. Each neuroimaging type required at least 5 studies to be used for formal meta-analysis [64]. All neuroimaging types examined reached this criterion.

Publications were excluded based on the following criteria: (1) failure to obtain full text of manuscript online or upon request from authors, (2) insufficient information for quantitative extraction, (3) non- or limited- peer-reviewed publications, including conference proceeding abstracts, and (4) intervention-based study designs. We included multiple studies reporting on the same original dataset. All studies were kept for qualitative analysis. To reduce the likelihood of overfitting due to non-independence across results and dataset decay [65], we calculated the mean classification metrics across studies that used the same dataset and included this combined result in our quantitative analysis. Overfitting from repeated re-use of the same public dataset can lead to minimal increases in prediction of unseen or independent data across different machine learning classifiers within clinical populations [66].

### Data extraction

Our primary estimates extracted from each study included sensitivity, a measurement used to assess the model's ability to accurately predict a psychosis participant correctly, and specificity, which measures the probability of accurately predicting a healthy control. Sensitivity is derived as the number of psychosis participants correctly identified by the classifier divided by the sum of all psychosis participants in the sample. Similarly, specificity is calculated as the number of healthy control participants correctly identified by the classifier divided by the sum of all healthy control participants. From this measurement we can derive the false positive rate ( $FPR = 1 - \text{specificity}$ ); this measurement specifies the probability of incorrectly labeling a healthy control as someone with psychosis.

We also collected the following information: year of publication, participants characteristics such as group size, age, gender, antipsychotic medications (as converted to chlorpromazine (CPZ) equivalents), illness duration in months, handedness, nicotine use, symptom severity as measured by the positive and negative

syndrome scale (PANSS [67]), and analysis characteristics such as dataset origin if using publicly available data, neuroimaging modality type (T1, DTI, rs-FC, or multimodal), classification method (e.g., support vector, ridge regression, decision tree), cross-validation procedure (e.g., leave one out, kfold, train test split), and number of features. If studies reported performance from multiple predictive models all measures were initially extracted. In the case of more than one statistical model, the sensitivity and specificity scores averaged across all models were used for the quantitative analysis.

Prediction results were classified as based on internal prediction (within dataset cross-validation) or external prediction (validation in a new dataset); these were used for separate quantitative analyses. When manuscripts incorporated both an internal and an external dataset for, both sets of performance measures were kept.

If multiple different studies used the same dataset for analysis, we recorded the sensitivity and specificity values for each study separately for reporting purposes and qualitative review, and included the mean across studies with that dataset for quantitative analysis to reduce the risk of overfitting (as discussed above). In the case of studies including several different datasets for model training, manuscripts were excluded if datasets were already in use among other studies, otherwise, the average pooled result was included for the quantitative analysis. When examining studies involving more than one clinical subgroup or first-degree relative we extracted classification measures specifically for participants diagnosed with psychosis and healthy control groups for quantitative analysis.

### Statistical analysis and assessment of bias

Cochrane's Q and  $I^2$  tests of heterogeneity were used to determine significant differences between studies and modalities (T1, DTI, rs-FC, multimodal) using a random effects model [68]. Q is used to assess that the proportion of successful classification is equal for all groups (healthy control and psychosis). Q is defined as the weighted sum of squared deviations from individual study effects (log odds ratio) against the pooled effect across studies. To determine if there is heterogeneity within and across imaging groups we formally test whether Q follows a chi-squared distribution with  $k-1$  degrees of freedom. If the null hypothesis is rejected ( $p < 0.05$ ) heterogeneity is likely present. Heterogeneity can also be measured using  $I^2$ . This measure describes the amount of variation present across studies [69, 70]. This procedure is calculated as a percentage of Q minus the degrees of freedom divided by Q. Heterogeneity was operationalized as small ( $I^2 = 25\%$ ), moderate ( $I^2 = 50\%$ ), or large ( $I^2 = 75\%$ ) [71].

To evaluate the potential for systematic bias in published results, several analyses were conducted. First, we created funnel plots for visual inspection of effect sizes for each imaging modality. This figure plots the effect estimates from each study against the standard error of effect estimates. This plot is used to evaluate the variation in classifying psychosis while accounting for sample size. If published effects are unbiased, then one should assume that no correlation exists between standard error and effect estimates after accounting for sample size heterogeneity across studies [72]. However, a correlation between standard error and effect estimates (seen as an asymmetry in the funnel plots), would suggest that there is some form of bias across studies that is not due to random sampling variation. Bias can be due to a number of factors such as publication bias, selective reporting, poor methodological design, and high heterogeneity [72]. Publication bias is just one of many potential reasons for asymmetry and it is impossible to know the precise mechanisms of asymmetry.

To formally test for funnel plot asymmetry, we conducted an Egger's regression test [73] and an alternative test called Peter's test [74]. Egger's test is a linear regression model of the estimates (log diagnostic odds ratio) on their standard errors weighted by

their inverse variance. While commonly used, this method can be problematic for assessing log odds ratio based estimates as the standard error is dependent on the size of the odds ratio even in the absence of small study effects [74]. The Egger's test can produce false positive results when sample sizes across groups (healthy control and psychosis) are not evenly balanced [74]. Peter's test is an alternative test that instead uses the inverse of the total sample size as the independent variable, thereby accounting for heterogeneity across groups (healthy control and psychosis) without increasing the likelihood for Type I errors.

Two sets of meta-analyses were conducted, one in which all studies were used and one in which a subset of outlier studies was excluded. For these meta-analyses, we implemented a bivariate approach in which sensitivity and specificity scores were log-transformed and combined into a bivariate regression model [61]. This approach is useful to assess diagnostic accuracy by accounting for biases in sensitivity and specificity [75]. Due to variation in modeling methods and specific cutoff thresholds for sensitivity and specificity, a random-effects model was applied. Each study was weighted based on sample size to account for variation in effect size. Statistical analyses were conducted using R [76]. To evaluate sensitivity and specificity values and conduct a bivariate meta-regression model to examine moderating effects of statistical methods, participant characteristics, and preprocessing method (when applicable) on the pooled estimates the packages *mada* [77] and *metafor* [78] were implemented. To reduce heterogeneity across studies, analyses were separated into internal prediction (via cross-validation in the same dataset), and external prediction (in a new dataset). A significant main effect was determined based on the use of a likelihood ratio test comparing the derived model to a null model. Based on this result, follow-up pairwise comparisons are conducted to determining the level of significant across factors (e.g., comparing each imaging type).

### Quality assessment of rs-FC preprocessing

To evaluate how denoising influences classification, we derived a quality measure of the denoising procedure at removing motion artifacts (Table 1) and related this rating to classification performance. This analysis was limited to rs-FC datasets, as other modalities did not include as many denoising procedures (Supplemental Table 1–4). The rating was based on results reported by Ciric and colleagues [57], which systematically compared the ability of different processing pipelines to remove motion biases in rs-FC analyses. The score was based on two criteria that measured the two major influences of motion on functional connectivity [57, 59]: (1) the total percent of edges related to head motion in each strategy (Fig. 2 in [57]) and (2) the distance-dependent influences of head motion on functional connectivity (Fig. 4 in [57]). The final score was weighted such that up to 75% of the final score was based on the first criteria and up to 25% of the final score was based on the second criteria (to reflect the relative difference in their impact on functional connectivity values [57]). Note that additional tests were also conducted on each criteria separately.

For the first criteria, each processing strategy was given a score scaled 1–5 with 1=good performance at removing motion artifacts in rs-FC (i.e., 0–10% edges contaminated by motion), 2=moderate performance (10–20%), 3=moderately poor performance (20–30% edges contaminated by motion), 4=poor performance (30–40% contaminated by motion), and 5=extreme contamination (>40% edges contaminated by motion). Manuscripts that included an additional step evaluating or excluding subjects based on framewise displacement (FD), resulted in one point subtracted off the initial edge score (i.e., subject removal, reporting of mean FD, group-related differences in FD, or any other mitigation strategy [60]).

For the second criteria the score was based on the magnitude of distance-dependent motion artifacts with 1 = good performance of minimizing distance dependence ( $r > -0.15$ ), 2=moderate

**Table 1.** Overview of scaling method used for assessing quality of preprocessing methods.

Good performance 1	Moderate performance 2	Moderately poor performance 3	Poor performance 4	Extreme Concern 5
Percent of Edges related to Motion				
(0–10% edges correlated with motion)	(10–20%)	(20–30%)	(30–40%)	(> 40%)
36P+spkreg 36P+despike 36P 36P+scrub	ICA+GSR 9P aCompCor	ICA	wmMean	2P 6P 24P wmLocal tCompCor
Distance dependence of motion effects				
QC-FC corr $r > -0.15$	$r = -0.15$ to $-0.2$	$r = -0.2$ to $-0.25$	$r = -0.25$ to $-0.3$	$r < -0.3$
ICA 36P+scrub	ICA+GSR wmLocal	24P 6P 36P+spkreg wmMean 36P+despike	acompcor 2P	9P tcompcor 36P

Percent of edges related to motion and distance dependence of motion effects are described in detail elsewhere [57]. 36 P = nuisance regressors included 6 motion estimates, mean white matter (WM), mean cerebral spinal fluid (CSF), and mean global signal (GS), along with the derivatives, quadratic terms and squares of these signals [171]. 36 P+despike = includes 36 regressors as described above, with despiking removal of high motion frames [194]. 36 P+spkreg = includes 36 regressors with spike regression of high motion frames [171]. 36 P+scrub = 36 parameters and motion scrubbing of high motion frames [59]. Scrubbing high motion frames were defined using framewise displacement (FD), computed as the sum of the absolute values of the derivatives of translational and rotational motion estimates.  $FD > .2$  mm was flagged as high motion. 2 P = nuisance regression includes mean WM and mean CSF.

6 P = nuisance regression only includes 6 motion estimates from realignment. 9 P + GSR = nuisance regression includes 6 motion estimates, mean WM, mean CSF, and mean GSR [195, 196]. 24 P = nuisance regression includes 6 motion estimates, their temporal derivatives and quadratic expansion terms [197]. aCompCor = nuisance regression includes 5 principal components each from the WM and CSF, in addition to 6 motion parameters and their temporal derivatives [198]. tCompCor = nuisance regression includes 6 principal components from voxels with high variance over time [199]. wmLocal = nuisance regression includes a voxelwise localized WM regressor in addition to 6 motion parameters, and their temporal derivatives and despiking [200]. wmMean = nuisance regression includes mean WM in addition to 6 motion parameters and their temporal derivatives and despiking [200]. ICA = independent component analysis, removal of motion-related variance components from the BOLD data including mean WM and CSF regressors [201].

performance ( $r = -0.15$  to  $-0.2$ ), 3 = moderately poor performance ( $r = -0.2$  to  $-0.25$ ), 4 = poor performance ( $r = -0.25$  to  $-0.3$ ), and 5 = extreme contamination ( $r < -0.3$ ). Each processing strategy was scored based on these criteria as shown in Table 1. Note that for the purposes of scoring, all ICA methods were grouped with ICA-AROMA as the closest comparator; other methods were also grouped with their closest fitting denoising approach. A composite score representing the quality of the denoising pipeline was generated with a 75% weighting from the edges contamination measure and a 25% weighting from the distance-dependent influence of motion.

Any manuscript using rs-FC features for classification was used in this analysis. Each manuscript's denoising methods were scored by 3 independent reviewers (authors A.P., S.F., and C.G.), and that assigned value was used for the quantitative analysis. Inter-rater reliability was high across reviewers (100%).

## RESULTS

The initial search yielded a total of 1003 manuscripts; 684 remained after removing manuscripts that did not involve psychosis-based disorders. Articles were then restricted based on those that included machine learning classification based on the selected MRI imaging modalities (T1, DTI, rs-FC). This led to 224 manuscripts that were analyzed for further review. Full-text publications were assessed for eligibility and after full text review 95 articles were retained for qualitative review and 93 for quantitative review (for a detailed breakdown of inclusion see Fig. 1 and Supplemental Table 1–4). Articles were removed from analysis for the following reasons: FC derived from tasks rather than rest, region/network specific analysis (not whole brain), intervention or longitudinal design, lack of cross-validation, lack of healthy controls, and review or meta-analysis manuscripts.

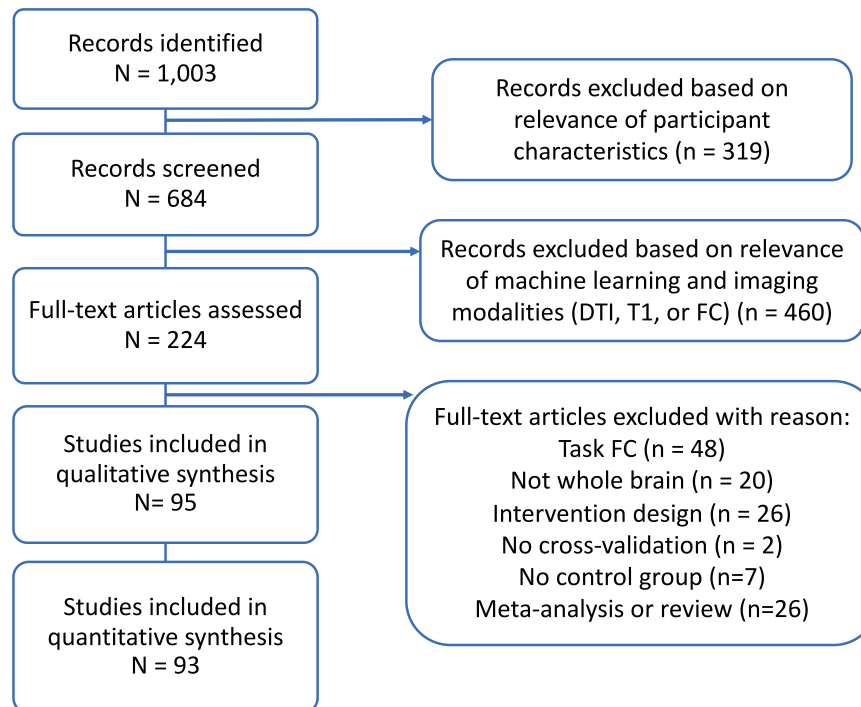
Results were separated for studies using internal validation (cross-validation within the same dataset) vs. external validation (validation within a new independent dataset). We focus first on reporting analyses from results of the larger internal validation

group. This initial analysis consisted of 28 T1 [79–109], 9 DTI [48, 110–117], 38 rs-FC [48, 118–156], and 14 multimodal [157–170] subgroups. From this sample, we identified 21 manuscripts that used the overlapping datasets for classification [86, 97, 102, 105–109, 116, 138–140, 142–154, 156, 166–169, 171]. To decrease the risk of inflation from non-independence and overfitting from dataset decay [70], we calculated the mean sensitivity and specificity scores across all studies that used the same dataset with the same imaging modality for the primary analyses reported in this manuscript. In the case that the same dataset was used with a different imaging modality for classification, we calculated the average measures per modality. This resulted in the inclusion of 3 rs-FC reports (using the W. China [172], Cobre [173], and NAMIC [174] datasets), 2 T1 (BRNO [106], Cobre [173]), and 1 multimodal (Cobre [173]). The reports from each analysis and the represented average are shown in detail in Supplemental Fig. 1.

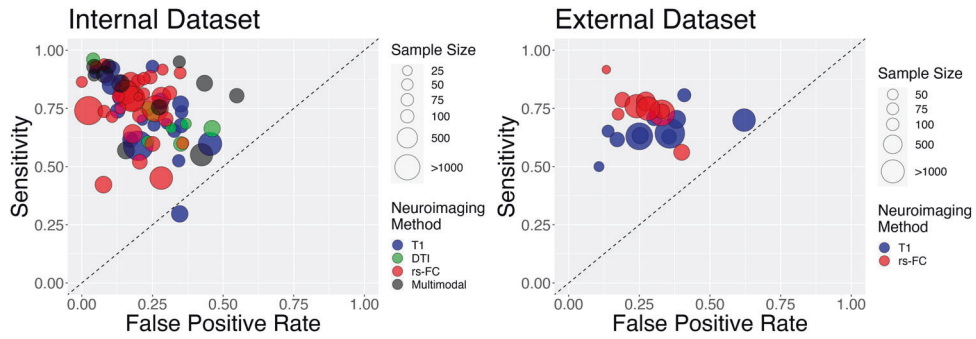
However, we also conducted a parallel set of analyses in which manuscripts that pooled information across overlapping datasets were included separately, in order to provide information based on larger, better powered studies. These manuscripts are also included Supplemental Table 1–4 and results from this parallel analysis are shown in Supplemental Fig. 2 and reported in the results sections below.

### All tested imaging modalities have a moderate ability to predict psychosis

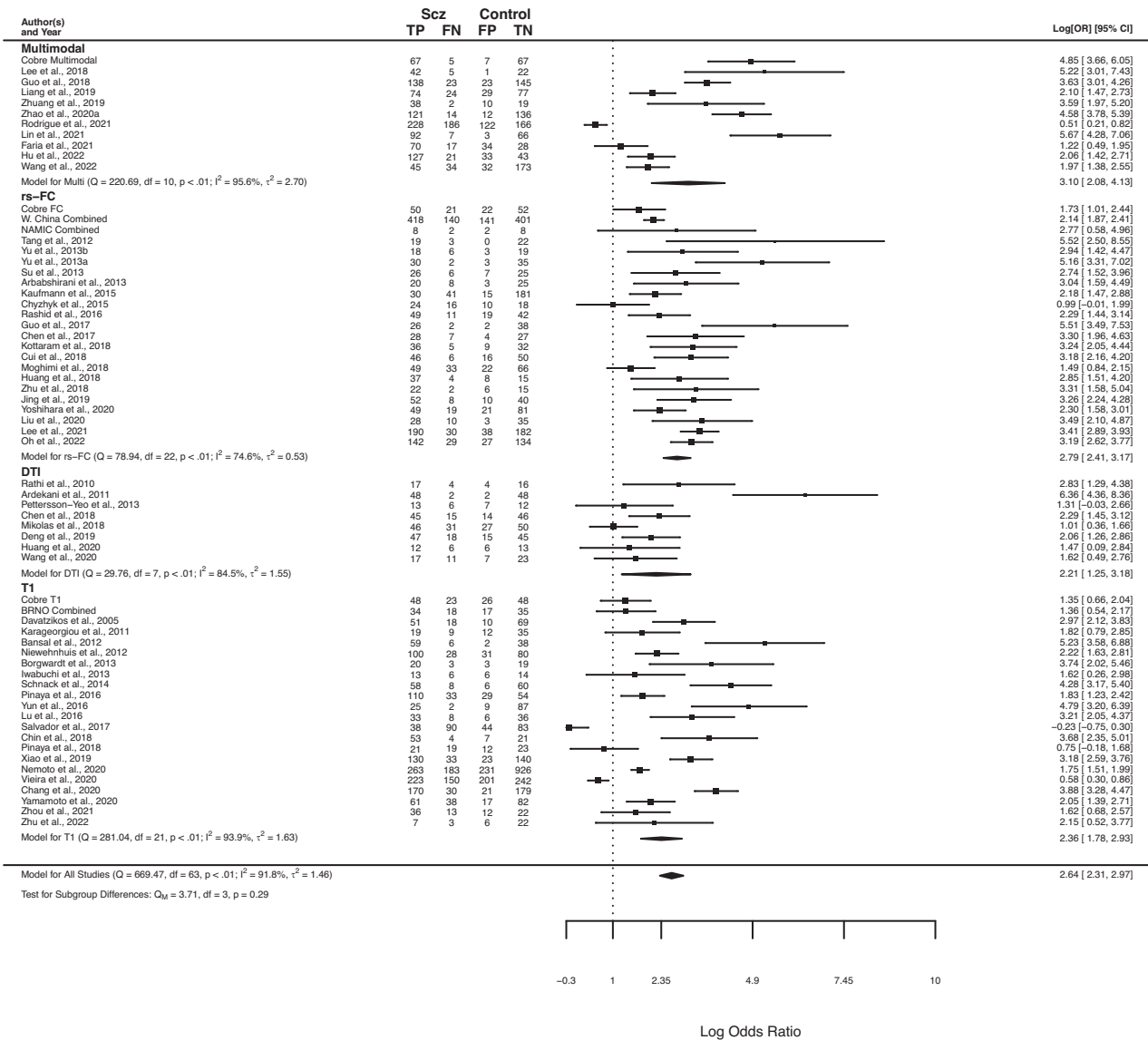
Aside from one T1 study [104] we found that all studies, independent of sample size, were able to reliably differentiate psychosis from healthy controls for all imaging modalities (Fig. 2, Fig. 3). Average sensitivity and specificity measures were modest across all imaging groups (T1: sensitivity =  $0.73 \pm 0.15$ , specificity =  $0.77 \pm 0.11$ ; DTI: sensitivity =  $0.71 \pm 0.11$ , specificity =  $0.73 \pm 0.12$ ; rs-FC: sensitivity =  $0.76 \pm 0.13$ ; specificity =  $0.81 \pm 0.09$ ; Multimodal: sensitivity =  $0.81 \pm 0.14$ , specificity =  $0.79 \pm 0.17$ ). Similar results were seen when larger-sample studies with pooled datasets were included in analysis (Supplemental Fig. 2).



**Fig. 1** Flow diagram mapping the selection of studies for the classification of psychosis-based disorders from neuroimaging modalities, following PRISMA guidelines [64, 65]. Full inclusion and exclusion criteria are listed in *Methods*.



**Fig. 2 Sensitivity plots of psychosis classification performance when tested on independent internal (left; i.e., cross-validation) and external (right) datasets.** Sensitivity and specificity scores were derived using data from the classifier in each manuscript. All manuscripts were able to reliably differentiate participants with psychosis from healthy controls independent of neuroimaging type aside from [169]. The size of points is scaled according to sample size and modality of analysis is shown in various colors.



**Fig. 3 Forest plots for log diagnostic odds ratio, separated by imaging modality.** Summary forest plot for log diagnostic odds ratio for all imaging modalities presented at the bottom of the plot. Multimodal: classification that used at least two of the following: rs-FC, T1, and/or DTI as features. RS-FC: resting state functional connectivity. DTI: diffusion tensor imaging. T1: T1 weighted imaging. The point size of squares and polygons are a function of the precision of the estimates.

We next examined whether specific modalities were better able to classify psychosis. Using a bivariate analysis, we did not find a significant difference in internal classification performance based on imaging modality ( $p > 0.05$ ; Fig. 3). Including pooled datasets resulted in a similar finding ( $p > 0.05$ ; Supplemental Fig. 2). These results suggest that, based on methods in the current literature, combining multiple neuroimaging methods to track psychosis does not provide any major advantage relative to single imaging modalities on average.

We also conducted a separate bivariate analysis using studies that provided classification performance in an external dataset ( $N = 21$ ). Due to the small number of studies, we were limited to quantitatively contrasting the use of T1 (14) or FC (12) imaging modalities. As might be expected, average sensitivity and specificity values were slightly lower compared to the internal results (Fig. 2; T1: sensitivity =  $0.66 \pm 0.07$ , specificity =  $0.69 \pm 0.14$ ; rs-FC: sensitivity =  $0.75 \pm 0.09$ , specificity =  $0.74 \pm 0.09$ ). In this analysis, we did not find a difference in performance by imaging modality. However, when pooled datasets were included we found a statistically significant association between imaging modality and classification performance in external samples (Sensitivity  $z = 0.34$ ;  $p = 0.003$ , Supplemental Fig. 3). This result indicates that rs-FC outperforms T1-based classification of psychosis in external datasets when large pooled datasets are used.

Notably, a close examination of these results indicates that there is substantial heterogeneity in classification performance across studies. We next analyze this heterogeneity in more detail, and ask whether there is bias in reported effects and how this bias affects classification performance.

#### Addressing heterogeneity reveals differences in psychosis prediction across imaging modalities

Neuroimaging subgroups appeared to have high amounts of variability as measured by Cochrane's  $Q$  ( $X^2 = 1,004.06$ ,  $p < 0.0001$ ). We asked whether this variability in reported effects

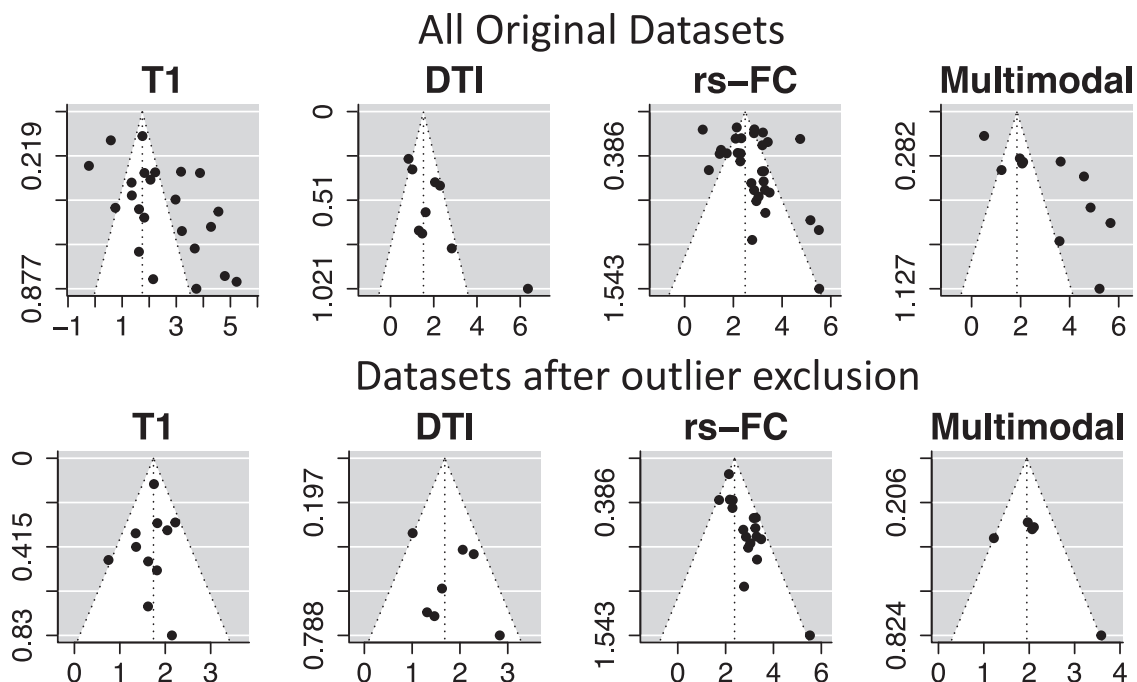
exhibited any evidence of bias. This was evaluated by examining signs of funnel plot asymmetry, Egger's test [73] and an alternative test [74]. Funnel plot asymmetry can indicate a bias of reported effect estimates and standard error after accounting for sample size heterogeneity.

Funnel plots and Egger's test produced evidence of asymmetry in reported effects for all modalities ( $p < 0.05$ ; Fig. 4), suggesting that reported effects may be biased. Bias can occur due to publication bias, selective reporting, poor methodological design, and high heterogeneity of findings [72]. However, use of the Peter's test resulted in no significant evidence of funnel plot asymmetry ( $p > 0.1$ ). These conflicting results are likely due to the large heterogeneity across studies and use of log odds ratio estimates (as described in more detail in ref. [74]).

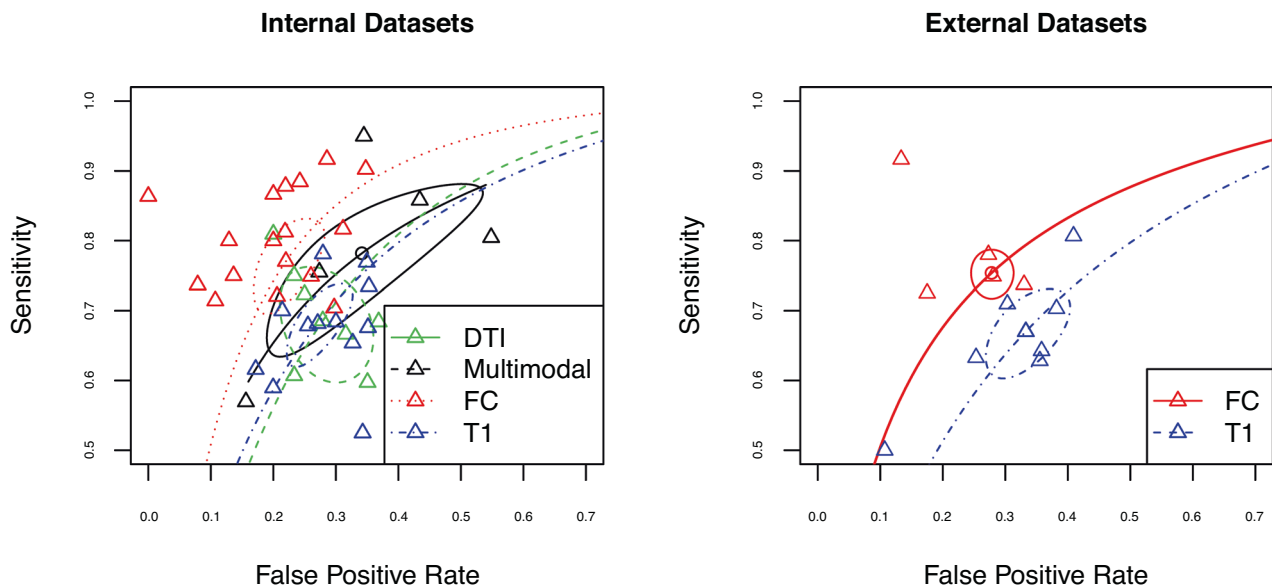
To address potential confounding results due to funnel plot asymmetry, we conducted an additional set of analyses similar to our primary ones, but after removing outlier studies. Outlier studies were identified as lying outside the 95% confidence interval of the pooled effect among each imaging group. This resulted in a final sample of 11 T1, 8 DTI, 21 rs-FC, and 5 multimodal internally-validated studies (Fig. 4, highlighted in Supplemental Table 1–4). After outlier exclusion, a dip test was conducted and did not reveal significant signs of non-unimodal distribution within each imaging group ( $p > 0.1$ ).

In this reduced set of manuscripts, we continued to observe similar classification performance across modalities, with no significant difference in psychosis classification across imaging modalities when classification was tested on independent internal datasets (i.e., via cross-validation;  $p > 0.05$ ; Fig. 5 left). If larger pooled datasets were included in the meta-analysis, studies based on rs-FC were associated with a significant decrease in false positive classification rates compared to studies based on DTI (Supp. Fig. 3;  $z = -0.58$ ,  $p(\text{FDR}) = 0.01$ ).

When classification was tested on external datasets, we also found signs of funnel plot asymmetry ( $p < 0.05$ ). After outlier exclusion,



**Fig. 4** Funnel plots with psychosis classification performance, broken down into subplots based on neuroimaging modality. Note: these results are based on internal validation only, given the higher number of studies in this domain. The top row shows funnel plots for all original internal datasets, while the bottom row shows funnel plots after outlier exclusion. Each plot is centered on a fixed effect summary estimate, the outer dashed lines indicate the 95% confidence interval of the fixed effect estimates. Symmetry is apparent when all studies are randomly dispersed around the dashed vertical line. In contrast, each imaging group showed signs of funnel plot asymmetry. These observations were confirmed by formally testing the correlation between study size and effect estimates using the Egger's test ( $p < .05$ ).



**Fig. 5** Sensitivity plots of psychosis classification performance after the exclusion of outlier studies, when tested on independent internal (left; i.e., cross-validation) and external (right) datasets. Colors represent different imaging modalities; estimated summary receiver operating curves are shown overlaid for each of the imaging modalities with respective confidence interval regions surrounding mean sensitivity and specificity values. No significant differences were found for classification of internal datasets. When external datasets were used for classification, rs-FC studies (red) have a significant advantage in classification relative to T1 studies (blue).

there was still a significant difference between imaging modality and classification, with studies using rs-FC resulting in higher sensitivity rates ( $z = 0.42$ ,  $p(\text{FDR}) < 0.001$ ) compared to T1 imaging (Fig. 5, right). This effect remained consistent when including pooled datasets ( $z = 0.39$ ;  $p(\text{FDR}) < 0.001$ ; Supplemental Fig. 3).

No other differences were present across modalities. Notably, in all versions of analyses, rs-FC studies performed similarly to multimodal studies, and multimodal studies did not outperform other imaging modalities in their prediction of psychosis. These findings suggest that, with the current literature, most imaging modalities perform similarly in psychosis classification, without major advantages for multimodal methods relative to unimodal methods (although rs-FC shows some slight advantages over structural-based approaches).

#### Classification performance is not associated with demographic or analysis covariates

In our final set of analyses, we investigated different potential sources of variation in prediction results (for a detailed overview see Supplemental Table 5). First, we examined participant characteristics: age, gender, CPZ equivalents, illness duration, and PANSS (we did not obtain enough studies that reported handedness or nicotine use to conduct a bivariate model assessing these participant characteristics). When conducting a bivariate analysis using all studies independent of imaging modality, we did not find a moderating effect on sensitivity or specificity measures for any of these factors ( $p > 0.1$ ).

Next, we examined analysis characteristics. We did not find moderating effects based on classification method (e.g., support vector, ridge, decision tree), deep vs. non-deep methods, cross-validation scheme (e.g., leave one out, test train split, kfold), feature size, or publication year ( $p > 0.1$ ) (see Supplemental Fig. 4). We also did not find a significant main effect based on sample size ( $p > 0.1$ ).

Head motion is a major confound in neuroimaging analyses [57–59, 175]. Therefore, we conducted an additional analysis to evaluate whether variance in effects sizes could be related to denoising methods designed to reduce the influence of head motion. This analysis was performed with a subset of studies that were either primarily rs-FC based or multimodal (rs-FC + other

methods) ( $N = 50$ ). We examined the influence of motion related artifacts on rs-FC effect sizes using a manually derived quality assessment score (as described in *Quality assessment of rs-FC preprocessing*). We found that the total quality assessment measure did not have an effect on classification ( $X^2(2,50) = 0.34$ ,  $p = 0.8$ ). This result suggests that motion artifacts were not a major driver of classification performance. However, it is important to note that the majority of rs-FC studies used similar motion denoising techniques.

#### DISCUSSION

We conducted a meta-analysis to test whether there are advantages to combining neuroimaging modalities for classification of individuals with psychosis. This analysis yielded several surprising and important findings. First, we found that all neuroimaging modalities examined (T1, DTI, rs-FC, multimodal) were able to classify individuals with psychosis from healthy controls. Second, we only found limited differences across modalities, primarily in advantages of rs-FC relative to T1 in classification performance in external datasets. Third, there was significant evidence for heterogeneity across studies. The reported effect sizes within each imaging group appear asymmetric, suggesting that systematic bias may be present in past reports of the classification of psychosis. When studies outside of the expected confidence bounds were removed, we continued to detect only limited differences across modalities, primarily associated with rs-FC approaches relative to structural imaging methods (DTI when classification was tested with internal datasets; T1 when classification was tested with external datasets). Notably, across all analyses, no difference was seen between multimodal and rs-FC approaches, which had largely overlapping classification distributions. We further discuss the implications of these findings and provide suggestions for improvements in future studies below.

Given that the extant of psychosis literature has identified changes in both the function and structure of whole brain networks, it is, perhaps, surprising that we did not see large improvements in classification when using multimodal methods



compared to all unimodal methods. Interestingly, in literature predicting behavioral variables from neuroimaging, recent papers have also found that multimodal methods do not provide an increase in performance relative to single modalities [176, 177]. This finding is in direct contrast to prior work arguing for the advantages of multimodal approaches (for review see [178]). This may be specific to these particular classification cases, to limitations in current methods in merging multimodal results, or due to individual differences in the function and structure of networks which introduces noise during classification and does not aid in prediction [179, 180].

When our analyses were restricted to remove outlier studies, we found evidence that rs-FC approaches statistically outperform structural (DTI for internal and T1 for external) studies in classifying psychosis. This suggests that heterogeneity in study results (discussed further below) may also limit our ability to detect modality-based effects. It is interesting that even in this circumstance, multimodal approaches did not differ from functional (rs-FC) approaches and had largely overlapping distributions. It is possible that different advantages between the modalities will be identified that pertain to specific questions and sub-populations, and that differences across modalities will be enhanced with additional methodological and analytical development. While the use of multimodal methods has gained popularity and can, at times, produce advantages to predicting psychopathology and cognition [178, 181], these differences in accuracy may not fully capture the inherent variation in sensitivity and specificity that was found across studies. We look forward to seeing additional methodological development and larger studies in this area that will help expand knowledge in this domain.

Results from Egger's test and the funnel plots demonstrated asymmetry in effect sizes among reported studies. This finding suggests that the ability to predict psychosis is negatively correlated with study precision (as measured by variance per participant group). Asymmetry in reported effects is usually evidence for systematic bias, independent of random sampling variance. This effect may be due to selective biases in reporting, poor methodological design, or inflated effects from small sample sizes. Publication bias, or the selective reporting of results that produce a significant effect, is one potential interpretation of why asymmetry was present in this analysis. Notably, however, removing studies outside of the expected confidence interval bound from our funnel plot analysis did not substantially change results regarding classification performance across modalities, aside from revealing a difference between rs-FC and DTI studies in prediction of internal datasets. We are hopeful that the increase in predictive modeling and popularity in preregistration of projects (e.g. Open Science Framework) will help reduce the effect of systematic bias in publication over time.

It is important to highlight that the Egger's Test can produce false positive results when sample sizes across participant groups are not evenly balanced [74]. Our meta-analysis included manuscripts with a wide range of group-level sample sizes that were not always balanced across participant groups ( $n = 10\text{--}600$  per group). When we conducted an alternative test for asymmetry using Peter's Test we did not find evidence of funnel plot asymmetry. From this analysis, we conclude that the asymmetry of effect sizes is likely at least in part associated with unbalanced participant groups. Future work should seek to align the sample size of participant groups prior to classifying psychosis.

Recent work has shown that separation of machine learning models based on race, gender, and age results in significant differences in classification performance [182, 183]. These findings indicate that there are biases in classification for certain groups, and that the lack of diversity in samples may lead to poor performance in broader and more diverse samples. We did not see a difference in classification based on age or gender within this meta-analysis (Supplemental Fig. 4), but were likely underpowered

for conducting a more rigorous analysis to determine if diversity characteristics had an effect on performance. Future work should evaluate how classification in psychosis varies when a diverse sample of individuals is used.

When evaluating studies that used external datasets for prediction, we found that functional (rs-FC) methods were significantly better at classifying psychosis compared to structural (T1) methods. It is important to note that we were limited to quantitatively contrast the use of T1 or rs-FC imaging modalities due to the small number of studies that use external prediction across other imaging modalities. Future work should place more emphasis on using external datasets to determine the extent to which a model generalizes and to provide an unbiased view of predictive performance across different imaging types. Notably, psychosis classification in external datasets was slightly lower than in internal datasets, suggesting that internal classification is inflating classification prediction abilities.

In addition to sample size, studies over the past decade have reported that head motion can systematically alter rs-FC estimates, and reduction of these biases requires appropriate preprocessing strategies [57–60]. When examining the quality of preprocessing methods in rs-FC we did not find that motion preprocessing impacted classification performance. This result could be due to the limited range of motion denoising methods across the majority of rs-FC based studies. There is substantial evidence that motion, respiration, and other physiological artifacts can significantly bias estimates of rs-FC [57–60]. These biases in rs-FC are of particular concern within psychosis samples [184–186]. Caution should be employed when evaluating classification metrics using rs-FC if preprocessing methods do not properly account for motion. As the field advances in motion filtering techniques, future work will need to reevaluate the effects of motion and classification in the context of psychopathology.

Notably, motion has also been demonstrated to impact DTI and T1 measures and lead to misleading results such as reduced volume and gray matter thickness [187, 188] and distorted measures of FA [189, 190]. Unfortunately, very few DTI and T1 studies addressed motion to a great extent, limiting our ability to analyze whether preprocessing strategies on effect sizes. Future work should evaluate how motion artifacts in these imaging modalities can also influence classification [191].

### Limitations and future directions

The use of neuroimaging-based classification holds considerable promise towards supporting clinicians in the diagnosis of psychopathology. Here, we found that many different neuroimaging methods were able to classify psychosis, but that these methods performed largely similarly, with slight differences observed between functional and structural imaging measures. There are important factors to consider that could influence the outcome of these findings.

Recent work has demonstrated that identification of behavioral phenotypes linked to psychopathology requires very large sample sizes ( $N > 2000$ ) in order to produce replicable results when using rs-FC and structural MRI measures [37]. Studies that met criteria for this meta-analysis varied considerably in sample size (20–1100), but were generally substantially smaller than this recommended size. Due to the limited number of studies available, we were unable to obtain manuscripts that utilized large enough samples and could not formally evaluate how samples larger than 2000 participants performed across imaging modality. As the trend to increase sample size continues, future work should reevaluate whether the combination of neuroimaging modalities provide substantial advantages when sample sizes are sufficiently large.

Notably, several of the largest sample sizes present in classification studies were associated with pooling of (the same) large public datasets. Given their non-independence, and in order to reduce the risk of overfitting and dataset decay [65] for these

large datasets, we included all of these results as a single average classification statistic in our primary quantitative analyses. However, this limited our ability to include datasets with increased sample sizes, which is an important limitation in machine learning studies. Therefore, we included a parallel set of analyses (reported in Supp. Fig. 1–3) that conducted meta-analyses with these large pooled datasets included separately. These results paralleled the findings from the primary analyses, indicating that, at present, the use of dataset pooling is not substantially altering findings in psychosis classification.

In our search for large-scale network markers that predict psychosis, we were selective in the studies that met criteria for our analysis, requiring that, for example, they included whole-brain analyses, used cross-validation, and reported classification measures (and for functional connectivity analyses, were based on static rs-FC; for full list of inclusion and exclusion criteria, see *Methods*). We restricted our meta-analysis to static resting-state FC, excluding task-based FC and dynamic approaches (given the wide variation in these, and the desire to measure task/variation effects), and we only included whole brain analyses as opposed to region-specific analyses, as our interest was to evaluate widespread changes in brain networks. These restrictions resulted in a surprisingly large number of rejected studies ( $N = 68$ ). Future work will be needed to systematically contrast different networks/regions and different functional connectivity methods on their ability to classify psychosis; past work has suggested commonalities across these functional network methods [192], but that task effects can also significantly influence prediction [193].

In addition to overlapping datasets, we only included manuscripts that had been peer-reviewed and indexed in PubMed Central or Web of Science. This approach limited our inclusion of conference abstracts, focused on recent advances. A number of abstracts from IEEE conferences [194–199] provide a glimpse into the outlook and growth of advanced modeling techniques for predicting psychosis. Generally, these studies demonstrate similar trends in performance as the studies included in the meta-analysis and highlight the field's continued increase in evaluating the clinical potential associated with predictive models [200, 201].

Our investigation included results from a range of different machine learning classification methods in the prediction of psychosis, including deep learning ( $N = 23$ ). There has been an exciting increase in analyses that have used the analytical potential of deep learning methods for prediction in psychopathology, often adopting collaborative (team-based) approaches [194, 202, 203]. However, recent research has also highlighted the limitations in deep learning approaches relative to more conventional machine learning models [204–208], including potential for overfitting datasets [66] and the lack of consistent quantitative benefits relative to prior work [209]. Consistent with Eital [209], in our meta-analysis, we did not find any significant differences in classification performance with deep learning methods relative to other machine learning methods (see Supplemental Fig. 4). However, this area of investigation is relatively nascent, and we believe that it will be valuable to continue additional investigations regarding the precision of deep learning methods relative to other algorithms.

Furthermore, it is worth noting that T1 and DTI data can be analyzed through a variety of methods (e.g., T1 can be used to analyze surface area, thickness and volume) that do not measure the same neurobiological underpinnings. Due to the variation in T1 measures used across each study within this meta-analysis, and the relatively small number of studies passing our criteria, we were unable to perform a more sophisticated model comparing each measurement type (surface area, cortical thickness, etc.) as it relates to classification. We look forward to seeing the addition of progressively more studies of each of these types taking a whole-brain approach to allow for the evaluation of many different networks in evaluating psychosis.

Finally, this analysis was restricted to only include classification between psychosis and healthy controls. It is possible that differences in network organization between healthy controls and psychosis are quite large resulting in higher classification than would be expected when evaluating more nuanced comparisons (e.g., schizophrenia vs. bipolar vs depression). The scope of this analysis was focused on the interplay of function and structure in brain networks related to psychosis and did not extend to other disorders. Future work should evaluate how various disorders and comorbidities relate to classification and neuroimaging modalities.

## CONCLUSIONS

All imaging techniques were able to classify psychosis from healthy controls. When accounting for variation in funnel plot asymmetry, we found significant evidence that rs-FC methods outperform structural imaging modalities. However, the results did not find significant differences between multimodal and rs-FC, suggesting that rs-FC may provide thorough information for classification. Future work should apply stringent guidelines when evaluating the predictive nature of neuroimaging modalities among psychosis.

## DATA AVAILABILITY

This review was not registered. The data supporting these findings of this study are available within this article and its supplementary materials.

## CODE AVAILABILITY

Availability of code to derive these findings can be found online ([https://github.com/GrattonLab/Porter\\_metaAnalysis2022](https://github.com/GrattonLab/Porter_metaAnalysis2022)).

## REFERENCES

- Andreasen NC, Flaum M. Schizophrenia: the characteristic symptoms. *Schizophr Bull.* 1991;17:27–49.
- Blanchard MM, Jacobson S, Clarke MC, Connor D, Kelleher I, Garavan H, et al. Language, motor and speed of processing deficits in adolescents with sub-clinical psychotic symptoms. *Schizophr Res.* 2010;123:71–6.
- Heaton RK, Gladsjo JA, Palmer BW, Kuck J, Marcotte TD, Jeste DV. Stability and course of neuropsychological deficits in schizophrenia. *Arch Gen Psych.* 2001;58:24–32.
- Lyne J, O'Donoghue B, Roche E, Renwick L, Cannon M, Clarke M. Negative symptoms of psychosis: a life course approach and implications for prevention and treatment. *Early Intervention Psych.* 2018;12:561–71.
- Walther S, Mittal VA. Motor System Pathology in Psychosis. *Curr Psychiatry Rep.* 2017;19:97.
- Collin G, Turk E, Van Den Heuvel MP. Connectomics in schizophrenia: from early pioneers to recent brain network findings. *Biol Psychiatry: Cogn Neurosci Neuroimaging.* 2016;1:199–208.
- Friston K, Brown HR, Siemerkus J, Stephan KE. The dysconnection hypothesis (2016). *Schizophr Res.* 2016;176:83–94.
- Goldman-Rakic PS, Selemon LD. Functional and anatomical aspects of prefrontal pathology in schizophrenia. *Schizophr Bull.* 1997;23:437–58.
- Menon V. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn Sci.* 2011;15:483–506.
- Supekar K, Cai W, Krishnadas R, Palaniyappan L, Menon V. Dysregulated Brain Dynamics in a Triple-Network Saliency Model of Schizophrenia and Its Relation to Psychosis. *Biol Psych.* 2019;85:60–9.
- van den Heuvel MP, Fornito A. Brain networks in schizophrenia. *Neuropsychol Rev.* 2014;24:32–48.
- Yu Q, Allen EA, Sui J, Arbabshirani MR, Pearlson G, Calhoun VD. Brain connectivity networks in schizophrenia underlying resting state functional magnetic resonance imaging. *Curr Top Med Chem.* 2012;12:2415–25.
- Flynn SW, Lang DJ, Mackay AL, Goghari V, Vavasour IM, Whittall KP, et al. Abnormalities of myelination in schizophrenia detected in vivo with MRI, and post-mortem with analysis of oligodendrocyte proteins. *Mol Psych.* 2003;8:811–20.
- Mauney SA, Pietersen CY, Sonntag KC, Woo TUV. Differentiation of oligodendrocyte precursors is impaired in the prefrontal cortex in schizophrenia. *Schizophr Res.* 2015;169:374–80.

15. Takahashi N, Sakurai T, Davis KL, Buxbaum JD. Linking oligodendrocyte and myelin dysfunction to neurocircuitry abnormalities in schizophrenia. *Prog Neurobiol.* 2011;93:13–24.
16. Cetin-Karayumak S, Di Biase MA, Chunga N, Reid B, Somes N, Lyall AE, et al. White matter abnormalities across the lifespan of schizophrenia: a harmonized multi-site diffusion MRI study. *Mol Psych.* 2020;25:3208–19.
17. Seitz-Holland J, Cetin-Karayumak S, Wojcik JD, Lyall A, Levitt J, Shenton ME, et al. Elucidating the relationship between white matter structure, demographic, and clinical variables in schizophrenia—a multicenter harmonized diffusion tensor imaging study. *Mol Psych.* 2021;26:5357–70.
18. Zhao Y, Zhang Q, Shah C, Li Q, Sweeney JA, Li F, et al. Cortical Thickness Abnormalities at Different Stages of the Illness Course in Schizophrenia: A Systematic Review and Meta-analysis. *JAMA Psych.* 2022;79:560–70.
19. Dietsche B, Kircher T, Falkenberg I. Structural brain changes in schizophrenia at different stages of the illness: A selective review of longitudinal magnetic resonance imaging studies. *Aust N. Z J Psych.* 2017;51:500–8.
20. van Dellen E, Börner C, Schutte M, van Montfort S, Abramovic L, Boks MP, et al. Functional brain networks in the schizophrenia spectrum and bipolar disorder with psychosis. *NPJ Schizophr.* 2020;6:22.
21. Xia CH, Ma Z, Ciric R, Gu S, Betzel RF, Kaczkurkin AN, et al. Linked dimensions of psychopathology and connectivity in functional brain networks. *Nat Commun.* 2018;9:3003.
22. Satterthwaite TD, Baker JT. How Can Studies of Resting-state Functional Connectivity Help Us Understand Psychosis as a Disorder of Brain Development? *Curr Opin Neurobiol.* 2015;0:85–91.
23. Sheffield JM, Barch DM. Cognition and resting-state functional connectivity in schizophrenia. *Neurosci Biobehav Rev.* 2016;61:108–20.
24. Robbins TW. The Case for Frontostriatal Dysfunction in Schizophrenia. *Schizophr Bull.* 1990;16:391–402.
25. Schmidt A, Borgwardt S. Disturbed Brain Networks in the Psychosis High-Risk State? In: Diwadkar VA, Eickhoff SB, editors. *Brain Network Dysfunction in Neuropsychiatric Illness: Methods, Applications, and Implications* [Internet]. Cham: Springer International Publishing; 2021 [cited 2022 Jul 28]. p. 217–38. Available from: [https://doi.org/10.1007/978-3-030-59797-9\\_11](https://doi.org/10.1007/978-3-030-59797-9_11).
26. Borgwardt S, McGuire P, Fusar-Poli P. Gray matters!—mapping the transition to psychosis. *Schizophr Res.* 2011;133:63–7.
27. McGlashan T. Schizophrenia in Translation: Is Active Psychosis Neurotoxic? *Schizophr Bull.* 2005;32:609–13.
28. Goff DC, Falkai P, Fleischhacker WW, Girgis RR, Kahn RM, Uchida H, et al. The Long-Term Effects of Antipsychotic Medication on Clinical Course in Schizophrenia. *AJP.* 2017;174:840–9.
29. Konrad A, Winterer G. Disturbed structural connectivity in schizophrenia—primary factor in pathology or epiphenomenon? *Schizophr Bull.* 2008;34:72–92.
30. Segal D, Koschnick JR, Slegers LHA, Hof PR. Oligodendrocyte pathophysiology: a new view of schizophrenia. *Int J Neuropsychopharmacol.* 2007;10:503–11.
31. Klawiter EC, Schmidt RE, Trinkaus K, Liang HF, Budde MD, Naismith RT, et al. Radial diffusivity predicts demyelination in ex vivo multiple sclerosis spinal cords. *Neuroimage.* 2011;55:1454–60.
32. Song SK, Sun SW, Ramsbottom MJ, Chang C, Russell J, Cross AH. Demyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage.* 2002;17:1429–36.
33. Song SK, Yoshino J, Le TQ, Lin SJ, Sun SW, Cross AH, et al. Demyelination increases radial diffusivity in corpus callosum of mouse brain. *Neuroimage.* 2005;26:132–40.
34. Samartzis L, Dima D, Fusar-Poli P, Kyriakopoulos M. White Matter Alterations in Early Stages of Schizophrenia: A Systematic Review of Diffusion Tensor Imaging Studies: White Matter Alterations in Early Schizophrenia. *J Neuroimaging.* 2014;24:101–10.
35. Waszczuk K, Rek-Owodziński K, Tyburski E, Mak M, Misiak B, Samochowiec J. Disturbances in White Matter Integrity in the Ultra-High-Risk Psychosis State—A Systematic Review. *JCM.* 2021;10:2515.
36. Vijayakumar N, Bartholomew C, Whitford T, Hermens DF, Nelson B, Rice S, et al. White matter integrity in individuals at ultra-high risk for psychosis: a systematic review and discussion of the role of polyunsaturated fatty acids. *BMC Psych.* 2016;16:287.
37. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, et al. Reproducible brain-wide association studies require thousands of individuals. *Nature.* 2022;603:654–60.
38. Baker JT, Holmes AJ, Masters GA, Yeo BTT, Krienen F, Buckner RL, et al. Disruption of Cortical Association Networks in Schizophrenia and Psychotic Bipolar Disorder. *JAMA Psych.* 2014;71:109.
39. Lefort-Besnard J, Bassett DS, Smallwood J, Margulies DS, Derntl B, Gruber O, et al. Different shades of default mode disturbance in schizophrenia: Subnodal covariance estimation in structure and function. *Hum Brain Mapp.* 2018;39:644–61.
40. Repovs G, Csernansky JG, Barch DM. Brain network connectivity in individuals with schizophrenia and their siblings. *Biol psychiatry.* 2011;69:967–73.
41. Tu PC, Lee YC, Chen YS, Li CT, Su TP. Schizophrenia and the brain's control network: aberrant within-and between-network connectivity of the frontoparietal network in schizophrenia. *Schizophrenia Res.* 2013;147:339–47.
42. Unschuld PG, Buchholz AS, Varvaris M, Van Zijl PC, Ross CA, Pekar JJ, et al. Prefrontal brain network connectivity indicates degree of both schizophrenia risk and cognitive dysfunction. *Schizophrenia Bull.* 2014;40:653–64.
43. Zhou Y, Liang M, Jiang T, Tian L, Liu Y, Liu Z, et al. Functional dysconnectivity of the dorsolateral prefrontal cortex in first-episode schizophrenia using resting-state fMRI. *Neurosci Lett.* 2007;417:297–302.
44. Cao H, Dixon L, Meyer-Lindenberg A, Tost H. Functional connectivity measures as schizophrenia intermediate phenotypes: advances, limitations, and future directions. *Curr Opin Neurobiol.* 2016;36:7–14.
45. Cole MW, Anticevic A, Repovs G, Barch D. Variable Global Dysconnectivity and Individual Differences in Schizophrenia. *Biol Psychiatry.* 2011;70:43–50.
46. Fan Y, Li L, Peng Y, Li H, Guo J, Li M, et al. Individual-specific functional connectome biomarkers predict schizophrenia positive symptoms during adolescent brain maturation. *Hum Brain Mapp.* 2021;42:1475–84.
47. Nawaz U, Lee I, Beermann A, Eack S, Keshavan M, Brady R. Individual Variation in Functional Brain Network Topography is Linked to Schizophrenia Symptomatology. *Schizophrenia Bull.* 2021;47:180–8.
48. Wang D, Li M, Wang M, Schoepp F, Ren J, Chen H, et al. Individual-specific functional connectivity markers track dimensional and categorical features of psychotic illness. *Mol Psych.* 2020;25:2119–29.
49. Schultz CC, Fusar-Poli P, Wagner G, Koch K, Schachtzabel C, Gruber O, et al. Multimodal functional and structural imaging investigations in psychosis research. *Eur Arch Psychiatry Clin Neurosci.* 2012;262:97–106.
50. Camchong J, MacDonald AW III, Bell C, Mueller BA, Lim KO. Altered functional and anatomical connectivity in schizophrenia. *Schizophrenia Bull.* 2011;37:640–50.
51. Pomarol-Clotet E, Canales-Rodríguez EJ, Salvador R, Sarró S, Gomar JJ, Vila F, et al. Medial prefrontal cortex pathology in schizophrenia as revealed by convergent findings from multimodal imaging. *Mol Psychiatry.* 2010;15:823–30.
52. Tian L, Meng C, Yan H, Zhao Q, Liu Q, Yan J, et al. Convergent Evidence from Multimodal Imaging Reveals Amygdala Abnormalities in Schizophrenic Patients and Their First-Degree Relatives. *PLoS ONE.* 2011;6:e28794.
53. McIntosh AR, Mišić B. Multivariate Statistical Analyses for Neuroimaging Data. *Annu Rev Psychol.* 2013;64:499–525.
54. Woo CW, Chang LJ, Lindquist MA, Wager TD. Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci.* 2017;20:365–77.
55. Nielsen AN, Barch DM, Petersen SE, Schlaggar BL, Greene DJ. Machine learning with neuroimaging: Evaluating its applications in psychiatry. *Biol Psychiatry: Cogn Neurosci Neuroimaging.* 2020;5:791–8.
56. Poldrack RA, Huckins G, Varoquaux G. Establishment of Best Practices for Evidence for Prediction: A Review. *JAMA Psych.* 2020;77:534–40.
57. Ciric R, Wolf DH, Power JD, Roalf DR, Baum GL, Ruparel K, et al. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *Neuroimage.* 2017;154:174–87.
58. Gratton C, Dworetzky A, Coalson RS, Adeyemo B, Laumann TO, Wig GS, et al. Removal of high frequency contamination from motion estimates in single-band fMRI saves data without biasing functional connectivity. *Neuroimage.* 2020;217:116866.
59. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* 2014;84:320–41.
60. Satterthwaite TD, Ciric R, Roalf DR, Davatzikos C, Bassett DS, Wolf DH. Motion artifact in studies of functional connectivity: Characteristics and mitigation strategies. *Hum Brain Mapp.* 2019;40:2033–51.
61. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJP, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol.* 2005;58:982–90.
62. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2007;62:e1–34.
63. Moher D, Liberati A, Tetzlaff J, Altman DG. PRISMA Group\*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med.* 2009;151:264–9.
64. Ioannidis JP, Lau J. Evolution of treatment effects over time: empirical insight from recursive cumulative metaanalyses. *Proc Natl Acad Sci.* 2001;98:831–6.
65. Thompson WH, Wright J, Bissett PG, Poldrack RA. Dataset decay and the problem of sequential analyses on open datasets. *Rodgers P, Baker CI, Holmes N, Baker CI, Rousselet GA, editors. eLife.* 2020;9:e53498.

66. Traut N, Heuer K, Lemaître G, Beggiano A, Germanaud D, Elmaleh M, et al. Insights from an autism imaging biomarker challenge: Promises and threats to biomarker discovery. *Neuroimage*. 2022;255:119171.
67. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bull*. 1987;13:261–76.
68. Cochran WG. The combination of estimates from different experiments. *Biometrics*. 1954;10:101–29.
69. Higgins CA, Judge TA, Ferris GR. Influence tactics and work outcomes: a meta-analysis. *J Organiz Behav*. 2003;24:89–106.
70. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21:1539–58.
71. Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ*. 2011;342:d549.
72. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002–d4002.
73. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Bmj*. 1997;315:629–34.
74. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of Two Methods to Detect Publication Bias in Meta-analysis. *JAMA* 2006;295:676–80.
75. Gatsonis C, Paliwal P. Meta-analysis of diagnostic and screening test accuracy evaluations: methodologic primer. *Am J Roentgenol*. 2006;187:271–81.
76. R. R. A language and environment for statistical computing. 2013;
77. Doebler P, Bürkner PC, Rucker G. Statistical Packages for Diagnostic Meta-Analysis and Their Application. *Diagnostic Meta-Analysis*. 2018;161–81.
78. Viechtbauer W, Viechtbauer MW Package 'metafor.' The Comprehensive R Archive Network Package 'metafor' <https://cran.r-project.org/web/packages/metafor/metafor.pdf>. 2015.
79. Nieuwenhuis M, van Haren NEM, Hulshoff Pol HE, Cahn W, Kahn RS, Schnack HG. Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. *NeuroImage*. 2012;61:606–12.
80. Schnack HG, Nieuwenhuis M, van Haren NEM, Abramovic L, Scheewe TW, Brouwer RM, et al. Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *NeuroImage*. 2014;84:299–306.
81. Xiao Y, Yan Z, Zhao Y, Tao B, Sun H, Li F, et al. Support vector machine-based classification of first episode drug-naïve schizophrenia patients and healthy controls using structural MRI. *Schizophrenia Res*. 2019;214:11–7.
82. Pinaya WHL, Mechelli A, Sato JR. Using deep autoencoders to identify abnormal brain structural patterns in neuropsychiatric disorders: A large-scale multi-sample study. *Hum Brain Mapp*. 2019;40:944–54.
83. Lu X, Yang Y, Wu F, Gao M, Xu Y, Zhang Y, et al. Discriminative analysis of schizophrenia using support vector machine and recursive feature elimination on structural MRI images. *Medicine*. 2016;95:e3973.
84. Iwabuchi SJ, Liddle PF, Palaniyappan L Clinical Utility of Machine-Learning Approaches in Schizophrenia: Improving Diagnostic Confidence for Translational Neuroimaging. *Front Psychiatry* [Internet]. 2013 [cited 2022 Sep 1];4. Available from: <http://journal.frontiersin.org/article/10.3389/fpsy.2013.00095/abstract>.
85. Yamamoto M, Bagarinao E, Kushima I, Takahashi T, Sasabayashi D, Inada T, et al. Support vector machine-based classification of schizophrenia patients and healthy controls using structural magnetic resonance imaging from two independent sites. *Yamasue H, editor. PLoS ONE*. 2020;15:e0239615.
86. Janousova E, Montana G, Kasparek T, Schwarz D Supervised, Multivariate, Whole-Brain Reduction Did Not Help to Achieve High Classification Performance in Schizophrenia Research. *Front Neurosci* [Internet]. 2016 Aug 25 [cited 2022 Sep 1];10. Available from: <http://journal.frontiersin.org/article/10.3389/fnins.2016.00392>.
87. Chin R, You AX, Meng F, Zhou J, Sim K. Recognition of Schizophrenia with Regularized Support Vector Machine and Sequential Region of Interest Selection using Structural Magnetic Resonance Imaging. *Sci Rep*. 2018;8:13858.
88. Yun J, Nyun Kim S, Young Lee T, Chon M, Soo, Kwon J. Individualized covariance profile of cortical morphology for auditory hallucinations in first-episode psychosis. *Hum Brain Mapp*. 2016;37:1051–65.
89. Borgwardt S, Koutsouleris N, Aston J, Studerus E, Smieskova R, Riecher-Rossler A, et al. Distinguishing Prodromal From First-Episode Psychosis Using Neuroanatomical Single-Subject Pattern Recognition. *Schizophrenia Bull*. 2013;39:1105–14.
90. Zhou Z, Wang K, Tang J, Wei D, Song L, Peng Y, et al. Cortical thickness distinguishes between major depression and schizophrenia in adolescents. *BMC Psychiatry*. 2021;21:361.
91. Chang YW, Tsai SJ, Wu YF, Yang AC. Development of an AI-Based Web Diagnostic System for Phenotyping Psychiatric Disorders. *Front Psych*. 2020;11:542394.
92. Davatzikos C, Ruparel K, Fan Y, Shen D, Acharyya M, Loughhead JW, et al. Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage*. 2005;28:663–8.
93. Zhu Y, Nakatani H, Yassin W, Maikusa N, Okada N, Kunimatsu A, et al. Application of a Machine Learning Algorithm for Structural Brain Images in Chronic Schizophrenia to Earlier Clinical Stages of Psychosis and Autism Spectrum Disorder: A Multiprotocol Imaging Dataset Study. *Schizophrenia Bull*. 2022;48:563–74.
94. Vieira S, Gong QY, Pinaya WHL, Scarpazza C, Tognin S, Crespo-Facorro B, et al. Using Machine Learning and Structural Neuroimaging to Detect First Episode Psychosis: Reconsidering the Evidence. *Schizophrenia Bull*. 2020;46:17–26.
95. Winterburn JL, Voineskos AN, Devenyi GA, Plitman E, de la Fuente-Sandoval C, Bhagwat N, et al. Can we accurately classify schizophrenia patients from healthy controls using magnetic resonance imaging and machine learning? A multi-method and multi-dataset study. *Schizophrenia Res*. 2019;214:3–10.
96. Karageorgiou E, Schulz SC, Gollub RL, Andreasen NC, Ho BC, Lauriello J, et al. Neuropsychological Testing and Structural Magnetic Resonance Imaging as Diagnostic Biomarkers Early in the Course of Schizophrenia and Related Psychoses. *Neuroinform*. 2011;9:321–33.
97. Dwyer DB, Cabral C, Kambeitz-Ilanovic L, Sanfelici R, Kambeitz J, Calhoun V, et al. Brain Typing Enhances The Neuroanatomical Discrimination of Schizophrenia. *Schizophrenia Bull*. 2018;44:1060–9.
98. Schwarz E, Doan NT, Pergola G, Westlye LT, Kaufmann T, Wolfers T, et al. Reproducible grey matter patterns index a multivariate, global alteration of brain structure in schizophrenia and bipolar disorder. *Transl Psych*. 2019;9:12.
99. Nemoto K, Shimokawa T, Fukunaga M, Yamashita F, Tamura M, Yamamori H, et al. Differentiation of schizophrenia using structural MRI with consideration of scanner differences: A real-world multisite study. *Psych Clin Neurosci*. 2020;74:56–63.
100. Bansal R, Staib LH, Laine AF, Hao X, Xu D, Liu J, et al. Anatomical Brain Images Alone Can Accurately Diagnose Chronic Neuropsychiatric Illnesses. *Zhan W, editor. PLoS ONE*. 2012;7:e50698.
101. Davatzikos C, Shen D, Gur RC, Wu X, Liu D, Fan Y, et al. Whole-Brain Morphometric Study of Schizophrenia Revealing a Spatially Complex Set of Focal Abnormalities. *Arch Gen Psych*. 2005;62:1218.
102. Sabuncu MR, Konukoglu E. Alzheimer's Disease Neuroimaging Initiative. Clinical prediction from structural brain MRI scans: a large-scale empirical study. *Neuroinformatics* 2015;13:31–46.
103. Pinaya WHL, Gadelha A, Doyle OM, Noto C, Zugman A, Cordeiro Q, et al. Using deep belief network modelling to characterize differences in brain morphology in schizophrenia. *Sci Rep*. 2016;6:38897.
104. Salvador R, Radau J, Canales-Rodriguez EJ, Solanes A, Sarró S, Goikolea JM, et al. Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS One*. 2017;12:e0175683.
105. Monté-Rubio GC, Falcón C, Pomarol-Clotet E, Ashburner J. A comparison of various MRI feature types for characterizing whole brain anatomical differences using linear pattern recognition methods. *NeuroImage*. 2018;178:753–68.
106. Vyškovský R, Schwarz D, Kašpárek T. Brain Morphometry Methods for Feature Extraction in Random Subspace Ensemble Neural Network Classification of First-Episode Schizophrenia. *Neural Comput*. 2019;31:897–918.
107. Chen Z, Yan T, Wang E, Jiang H, Tang Y, Yu X, et al. Detecting Abnormal Brain Regions in Schizophrenia Using Structural MRI via Machine Learning. *Computational Intell Neurosci*. 2020;2020:1–13.
108. Latha M, Kavitha G. Combined Metaheuristic Algorithm and Radiomics Strategy for the Analysis of Neuroanatomical Structures in Schizophrenia and Schizoaffective Disorders. *IRBM*. 2021;42:353–68.
109. Vyškovský R, Schwarz D, Churová V, Kašpárek T. Structural MRI-Based Schizophrenia Classification Using Autoencoders and 3D Convolutional Neural Networks in Combination with Various Pre-Processing Techniques. *Brain Sci*. 2022;12:615.
110. Mikolas P, Hlinka J, Koch A, Pitra Z, Frodl T, Spaniel F, et al. Machine learning classification of first-episode schizophrenia spectrum disorders and controls using whole brain white matter fractional anisotropy. *BMC Psych*. 2018;18:97.
111. Ardekani BA, Tabesh A, Sevy S, Robinson DG, Bilder RM, Szeszko PR. Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers. *Hum Brain Mapp*. 2011;32:1–9.
112. Deng Y, Hung KSY, Lui SSS, Chui WWH, Lee JCW, Wang Y, et al. Tractography-based classification in distinguishing patients with first-episode schizophrenia from healthy individuals. *Prog Neuro-Psychopharmacol Biol Psych*. 2019;88:66–73.
113. Huang J, Wang M, Xu X, Jie B, Zhang D. A novel node-level structure embedding and alignment representation of structural networks for brain disease analysis. *Med Image Anal*. 2020;65:101755.
114. Hutchison D, Kanade T, Kittler J, Kleinberg JM, Mattern F, Mitchell JC, et al. Biomarkers for Identifying First-Episode Schizophrenia Patients Using Diffusion Weighted Imaging. In: Jiang T, Navab N, Pluim JPW, Viergever MA, editors. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010 [cited 2022 Sep 1]. p. 657–65. (Lecture Notes in Computer Science; vol. 6361). Available from: [http://link.springer.com/10.1007/978-3-642-15705-9\\_80](http://link.springer.com/10.1007/978-3-642-15705-9_80).
115. Pettersson-Yeo W, Benetti S, Marquand AF, Dell'Acqua F, Williams SCR, Allen P, et al. Using genetic, cognitive and multi-modal neuroimaging data to identify ultra-high-risk and first-episode psychosis at the individual level. *Psychol Med*. 2013;43:2547–62.

116. Morgan SE, Young J, Patel AX, Whitaker KJ, Scarpazza C, van Amelsvoort T, et al. Functional Magnetic Resonance Imaging Connectivity Accurately Distinguishes Cases With Psychotic Disorders From Healthy Controls, Based on Cortical Features Associated With Brain Network Development. *Biol Psych: Cogn Neurosci Neuroimaging*. 2021;6:1125–34.
117. Chen YJ, Liu CM, Hsu YC, Lo YC, Hwang TJ, Hwu HG, et al. Individualized prediction of schizophrenia based on the whole-brain pattern of altered white matter tract integrity: Individualized Tract-Based Schizophrenia Prediction. *Hum Brain Mapp*. 2018;39:575–87.
118. Arbabshirani MR, Kiehl KA, Pearson GD, Calhoun VD Classification of schizophrenia patients based on resting-state functional network connectivity. *Front Neurosci* [Internet]. 2013 [cited 2022 Aug 26];7. Available from: <http://journal.frontiersin.org/article/10.3389/fnins.2013.00133/abstract>.
119. Zhu Q, Huang J, Xu X. Non-negative discriminative brain functional connectivity for identifying schizophrenia on resting-state fMRI. *BioMed Eng OnLine*. 2018;17:32.
120. Chen H, Uddin LQ, Duan X, Zheng J, Long Z, Zhang Y, et al. Shared atypical default mode and salience network functional connectivity between autism and schizophrenia: Shared atypical FC in ASD and schizophrenia. *Autism: Research*. 2017;10:1776–86.
121. Huang P, Cui LB, Li X, Lu ZL, Zhu X, Xi Y, et al. Identifying first-episode drug naïve patients with schizophrenia with or without auditory verbal hallucinations using whole-brain functional connectivity: A pattern analysis study. *NeuroImage: Clin*. 2018;19:351–9.
122. Su L, Wang L, Shen H, Feng G, Hu D Discriminative analysis of non-linear brain connectivity in schizophrenia: an fMRI Study. *Front Hum Neurosci* [Internet]. 2013 [cited 2022 Aug 26];7. Available from: <http://journal.frontiersin.org/article/10.3389/fnhum.2013.00702/abstract>.
123. Liu W, Zhang X, Qiao Y, Cai Y, Yin H, Zheng M, et al. Functional Connectivity Combined With a Machine Learning Algorithm Can Classify High-Risk First-Degree Relatives of Patients With Schizophrenia and Identify Correlates of Cognitive Impairments. *Front Neurosci*. 2020;14:577568.
124. Jing R, Li P, Ding Z, Lin X, Zhao R, Shi L, et al. Machine learning identifies unaffected first-degree relatives with functional network patterns and cognitive impairment similar to those of schizophrenia patients. *Hum Brain Mapp*. 2019;40:3930–39.
125. Lee LH, Chen CH, Chang WC, Lee PL, Shyu KK, Chen MH, et al. Evaluating the performance of machine learning models for automatic diagnosis of patients with schizophrenia based on a single site dataset of 440 participants. *Eur Psych*. 2022;65:e1.
126. Tang Y, Wang L, Cao F, Tan L. Identify schizophrenia using resting-state functional connectivity: an exploratory research and analysis. *BioMed Eng OnLine*. 2012;11:50.
127. Yu Y, Shen H, Zeng LL, Ma Q, Hu D. Convergent and Divergent Functional Connectivity Patterns in Schizophrenia and Depression. Zang YF, editor. *PLoS ONE*. 2013;8:e68250.
128. Yu Y, Shen H, Zhang H, Zeng LL, Xue Z, Hu D. Functional connectivity-based signatures of schizophrenia revealed by multiclass pattern analysis of resting-state fMRI from schizophrenic patients and their healthy siblings. *BioMed Eng OnLine*. 2013;12:10.
129. Chyzyhyk D, Graña M, Öngür D, Shinn AK. Discrimination of Schizophrenia Auditory Hallucinators by Machine Learning of Resting-State Functional MRI. *Int J Neur Syst*. 2015;25:1550007.
130. Rashid B, Arbabshirani MR, Damaraju E, Cetin MS, Miller R, Pearson GD, et al. Classification of schizophrenia and bipolar patients using static and dynamic resting-state fMRI brain connectivity. *NeuroImage*. 2016;134:645–57.
131. Skåtun KC, Kaufmann T, Doan NT, Alnæs D, Córdova-Palomera A, Jönsson EG, et al. Consistent Functional Connectivity Alterations in Schizophrenia Spectrum Disorder: A Multisite Study. *Schizophr Bull*. 2017;43:914–24.
132. Moghimi P, Lim KO, Netoff TI. Data Driven Classification Using fMRI Network Measures: Application to Schizophrenia. *Front Neuroinform*. 2018;12:71.
133. Kaufmann T, Skåtun KC, Alnæs D, Doan NT, Duff EP, Tønnesen S, et al. Disintegration of Sensorimotor Brain Networks in Schizophrenia. *SCHBUL*. 2015;41:1326–35.
134. Cui LB, Liu L, Wang HN, Wang LX, Guo F, Xi YB, et al. Disease Definition for Schizophrenia by Functional Connectivity Using Radiomics Strategy. *Schizophr Bull*. 2018;44:1053–9.
135. Yoshihara Y, Lisi G, Yahata N, Fujino J, Matsumoto Y, Miyata J, et al. Overlapping but Asymmetrical Relationships Between Schizophrenia and Autism Revealed by Brain Connectivity. *Schizophr Bull*. 2020;46:1210–8.
136. Kottaram A, Johnston L, Ganella E, Pantelis C, Kotagiri R, Zalesky A. Spatiotemporal dynamics of resting-state brain networks improve single-subject prediction of schizophrenia diagnosis. *Hum Brain Mapp*. 2018;39:3663–81.
137. Serin E, Zalesky A, Matory A, Walter H, Kruschwitz JD. NBS-Predict: A prediction-based extension of the network-based statistic. *NeuroImage*. 2021;244:118625.
138. Fekete T, Wilf M, Rubin D, Edelman S, Malach R, Mujica-Parodi LR Combining Classification with fMRI-Derived Complex Network Measures for Potential Neurodiagnostics. Hayasaka S, editor. *PLoS ONE*. 2013;8:e62867.
139. Hu X, Zhu D, Lv P, Li K, Han J, Wang L, et al. Fine-Granularity Functional Interaction Signatures for Characterization of Brain Conditions. *Neuroinform*. 2013;11:301–17.
140. Kim J, Calhoun VD, Shim E, Lee JH. Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *NeuroImage*. 2016;124:127–46.
141. Guo W, Liu F, Chen J, Wu R, Li L, Zhang Z, et al. Using short-range and long-range functional connectivity to identify schizophrenia with a family-based case-control design. *Psychiatry Res: Neuroimaging*. 2017;264:60–7.
142. Zeng LL, Wang H, Hu P, Yang B, Pu W, Shen H, et al. Multi-Site Diagnostic Classification of Schizophrenia Using Discriminant Deep Learning with Functional Connectivity MRI. *EBioMedicine* 2018;30:74–85.
143. Ji G, Chen X, Bai T, Wang L, Wei Q, Gao Y, et al. Classification of schizophrenia by intersubject correlation in functional connectome. *Hum Brain Mapp*. 2019;40:2347–57.
144. Matsubara T, Tashiro T, Uehara K. Deep Neural Generative Model of Functional MRI Images for Psychiatric Disorder Diagnosis. *IEEE Trans Biomed Eng*. 2019;66:2768–79.
145. Yan W, Calhoun V, Song M, Cui Y, Yan H, Liu S, et al. Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site fMRI data. *EBioMedicine* 2019;47:543–52.
146. Yang B, Chen Y, Shao QM, Yu R, Li WB, Guo GQ, et al. Schizophrenia Classification Using fMRI Data Based on a Multiple Feature Image Capsule Network Ensemble. *IEEE Access*. 2019;7:109956–68.
147. Zhu Q, Li H, Huang J, Xu X, Guan D, Zhang D. Hybrid Functional Brain Network With First-Order and Second-Order Information for Computer-Aided Diagnosis of Schizophrenia. *Front Neurosci*. 2019;13:603.
148. Lei D, Pinaya WHL, van Amelsvoort T, Marcelis M, Donohoe G, Mothersill DO, et al. Detecting schizophrenia at the level of the individual: relative diagnostic value of whole-brain images, connectome-wide functional connectivity and graph-based metrics. *Psychol Med*. 2020;50:1852–61.
149. Lei D, Pinaya WHL, Young J, Amelsvoort T, Marcelis M, Donohoe G, et al. Integrating machine learning and multimodal neuroimaging to detect schizophrenia at the level of the individual. *Hum Brain Mapp*. 2020;41:1119–35.
150. Xiang Y, Wang J, Tan G, Wu FX, Liu J. Schizophrenia Identification Using Multi-View Graph Measures of Functional Brain Networks. *Front Bioeng Biotechnol*. 2020;7:479.
151. Zhao J, Huang J, Zhi D, Yan W, Ma X, Yang X, et al. Functional network connectivity (FNC)-based generative adversarial network (GAN) and its applications in classification of mental disorders. *J Neurosci Methods*. 2020;341:108756.
152. Zhu Y, Fu S, Yang S, Liang P, Tan Y. Weighted Deep Forest for Schizophrenia Data Classification. *IEEE Access*. 2020;8:62698–705.
153. Gallos IK, Galaris E, Siettos CI. Construction of embedded fMRI resting-state functional connectivity networks using manifold learning. *Cogn Neurodyn*. 2021;15:585–608.
154. Lei D, Qin K, Pinaya WHL, Young J, Van Amelsvoort T, Marcelis M, et al. Graph Convolutional Networks Reveal Network-Level Functional Dysconnectivity in Schizophrenia. *Schizophr Bull*. 2022;48:881–92.
155. Oh KH, Oh IS, Tsogt U, Shen J, Kim WS, Liu C, et al. Diagnosis of schizophrenia with functional connectome data: a graph-based convolutional neural network approach. *BMC Neurosci*. 2022;23:5.
156. Yuan X, Gu L, Huang J. GK-BSC: Graph Kernel-Based Brain States Construction With Dynamic Brain Networks and Application to Schizophrenia Identification. *IEEE Access*. 2022;10:58558–65.
157. Rodrigue AL, Mastrovito D, Esteban O, Durnez J, Koenis MMG, Janssen R, et al. Searching for Imaging Biomarkers of Psychotic Dysconnectivity. *Biol Psychiatry: Cogn Neurosci Neuroimaging*. 2021;6:1135–44.
158. Guo S, Huang CC, Zhao W, Yang AC, Lin CP, Nichols T, et al. Combining multi-modality data for searching biomarkers in schizophrenia. Hu D, editor. *PLoS ONE*. 2018;13:e0191202.
159. Faria AV, Zhao Y, Ye C, Hsu J, Yang K, Cifuentes E, et al. Multimodal MRI assessment for first episode psychosis: A major change in the thalamus and an efficient stratification of a subgroup. *Hum Brain Mapp*. 2021;42:1034–53.
160. Zhuang H, Liu R, Wu C, Meng Z, Wang D, Liu D, et al. Multimodal classification of drug-naïve first-episode schizophrenia combining anatomical, diffusion and resting state functional resonance imaging. *Neurosci Lett*. 2019;705:87–93.
161. Wang J, Ke P, Zang J, Wu F, Wu K. Discriminative Analysis of Schizophrenia Patients Using Topological Properties of Structural and Functional Brain Networks: A Multimodal Magnetic Resonance Imaging Study. *Front Neurosci*. 2022;15:785595.
162. Zhao W, Guo S, Linli Z, Yang AC, Lin CP, Tsai SJ. Functional, Anatomical, and Morphological Networks Highlight the Role of Basal Ganglia–Thalamus–Cortex Circuits in Schizophrenia. *Schizophr Bull*. 2019;46:422–31.

163. Lin X, Li W, Dong G, Wang Q, Sun H, Shi J, et al. Characteristics of Multimodal Brain Connectomics in Patients With Schizophrenia and the Unaffected First-Degree Relatives. *Front Cell Dev Biol.* 2021;9:631864.
164. Lee J, Chon MW, Kim H, Rathi Y, Bouix S, Shenton ME, et al. Diagnostic value of structural and diffusion imaging measures in schizophrenia. *NeuroImage: Clin.* 2018;18:467–74.
165. Liang S, Li Y, Zhang Z, Kong X, Wang Q, Deng W, et al. Classification of First-Episode Schizophrenia Using Multimodal Brain Features: A Combined Structural and Diffusion Imaging Study. *Schizophr Bull.* 2019;45:591–9.
166. Qureshi MNI, Oh J, Cho D, Jo HJ, Lee B. Multimodal Discrimination of Schizophrenia Using Hybrid Weighted Feature Concatenation of Brain Functional Connectivity and Anatomical Features with an Extreme Learning Machine. *Front Neuroinform.* 2017;11:59.
167. Cabral C, Kambeitz-Ilanovic L, Kambeitz J, Calhoun VD, Dwyer DB, von Salder S, et al. Classifying Schizophrenia Using Multimodal Multivariate Pattern Recognition Analysis: Evaluating the Impact of Individual Clinical Profiles on the Neurodiagnostic Performance. *SCHBUL.* 2016;42:S110–7.
168. Qureshi MNI, Oh J, Lee B. 3D-CNN based discrimination of schizophrenia using resting-state fMRI. *Artif Intell Med.* 2019;98:10–7.
169. Masoudi B, Daneshvar S, Razavi SN. A multi-modal fusion of features method based on deep belief networks to diagnosis schizophrenia disease. *Int J Wavel, Multiresolution Inf Process.* 2021;19:2050088.
170. Hu M, Qian X, Liu S, Koh AJ, Sim K, Jiang X, et al. Structural and diffusion MRI based schizophrenia classification using 2D pretrained and 3D naive Convolutional Neural Networks. *Schizophr Res.* 2022;243:330–41.
171. Wang T, Bezerianos A, Cichocki A, Li J. Multikernel Capsule Network for Schizophrenia Identification. *IEEE Trans Cyber.* 2022;52:4741–50.
172. Liu S, Wang H, Song M, Lv L, Cui Y, Liu Y, et al. Linked 4-Way Multimodal Brain Differences in Schizophrenia in a Large Chinese Han Population. *Schizophr Bull.* 2019;45:436–49.
173. Wood D, King M, Landis D, Courtney W, Wang R, Kelly R, et al. Harnessing modern web application technology to create intuitive and efficient data visualization and sharing tools. *Front Neuroinform [Internet].* 2014 Aug 26 [cited 2022 Dec 9];8. Available from: <http://journal.frontiersin.org/article/10.3389/fninf.2014.00071/abstract>.
174. Kapur T, Pieper S, Whitaker R, Aylward S, Jakab M, Schroeder W, et al. The National Alliance for Medical Image Computing, a roadmap initiative to build a free and open source software infrastructure for translational research in medical image analysis. *J Am Med Inf Assoc.* 2012;19:176–80.
175. Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughhead J, Calkins ME, et al. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage.* 2013;64:240–56.
176. Dhamala E, Jamison KW, Jaywant A, Dennis S, Kuceyeski A. Distinct functional and structural connections predict crystallised and fluid cognition in healthy adults. *Hum Brain Mapp.* 2021;42:3102–18.
177. Ooi LQR, Chen J, Zhang S, Kong R, Tam A, Li J, et al. Comparison of individualized behavioral predictions across anatomical, diffusion and functional connectivity MRI. *NeuroImage.* 2022;263:119636.
178. Sui J, Jiang R, Bustillo J, Calhoun V. Neuroimaging-based Individualized Prediction of Cognition and Behavior for Mental Disorders and Health: Methods and Promises. *Biol Psych.* 2020;88:818–28.
179. Llera A, Wolfers T, Mulders P, Beckmann CF. Inter-individual differences in human brain structure and morphology link to variation in demographics and behavior. *Elife.* 2019;8:e44443.
180. Mansour LS, Tian Y, Yeo BTT, Cropley V, Zalesky A. High-resolution connectomic fingerprints: Mapping neural identity and behavior. *NeuroImage* 2021;229:117695.
181. Sui J, Adali T, Yu Q, Chen J, Calhoun VD. A review of multivariate methods for multimodal fusion of brain imaging data. *J Neurosci Methods.* 2012;204:68–81.
182. Li J, Bzdok D, Chen J, Tam A, Ooi LQR, Holmes AJ, et al. Cross-ethnicity/race generalization failure of behavioral prediction from resting-state functional connectivity. *Sci Adv.* 2022;8:eabj1812.
183. Wang R, Chaudhari P, Davatzikos C. Embracing the disharmony in medical imaging: A Simple and effective framework for domain adaptation. *Med Image Anal.* 2022;76:102309.
184. Bassett DS, Xia CH, Satterthwaite TD. Understanding the Emergence of Neuropsychiatric Disorders With Network Neuroscience. *Biol Psych: Cogn Neurosci Neuroimaging.* 2018;3:742–53.
185. Parkes L, Fulcher B, Yücel M, Fornito A. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *NeuroImage.* 2018;171:415–36.
186. Siegel JS, Mitra A, Laumann TO, Seitzman BA, Raichle M, Corbetta M, et al. Data Quality Influences Observed Links Between Functional Connectivity and Behavior. *Cereb Cortex.* 2017;27:4492–502.
187. Backhausen LL, Herting MM, Buse J, Roessner V, Smolka MN, Vetter NC Quality Control of Structural MRI Images Applied Using FreeSurfer—A Hands-On Workflow to Rate Motion Artifacts. *Front Neurosci [Internet].* 2016 Dec 6 [cited 2022 Aug 12];10. Available from: <http://journal.frontiersin.org/article/10.3389/fnins.2016.00558/full>.
188. Reuter M, Tisdall MD, Qureshi A, Buckner RL, van der Kouwe AJW, Fischl B. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *NeuroImage.* 2015;107:107–15.
189. Heim S, Hahn K, Sämman PG, Fahrmeir L, Auer DP. Assessing DTI data quality using bootstrap analysis: Assessing DTI Data Quality. *Magn Reson Med.* 2004;52:582–9.
190. Ling J, Merideth F, Caprihan A, Pena A, Teshiba T, Mayer AR. Head injury or head motion? Assessment and quantification of motion artifacts in diffusion tensor imaging studies. *Hum Brain Mapp.* 2012;33:50–62.
191. Rokham H, Pearlson G, Abrol A, Falakshahi H, Plis S, Calhoun VD. Addressing Inaccurate Nosology in Mental Health: A Multilabel Data Cleansing Approach for Detecting Label Noise From Structural Magnetic Resonance Imaging Data in Mood and Psychosis Disorders. *Biol Psychiatry: Cogn Neurosci Neuroimaging.* 2020;5:819–32.
192. Gratton C, Laumann TO, Nielsen AN, Greene DJ, Gordon EM, Gilmore AW, et al. Functional Brain Networks Are Dominated by Stable Group and Individual Factors, Not Cognitive or Daily Variation. *Neuron.* 2018;98:439–52.e5.
193. Porter A, Nielsen A, Dorn M, Dworetzky A, Edmonds D, Gratton C. Masked features of task states found in individual brain networks. *Cereb Cortex.* 2022;33:2879–900.
194. Silva RF, Castro E, Gupta CN, Cetin M, Arbabshirani M, Potluru VK, et al. The tenth annual MLSP competition: Schizophrenia classification challenge. In: 2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP) [Internet]. Reims, France: IEEE; 2014 [cited 2023 Apr 14]. p. 1–6. Available from: <http://ieeexplore.ieee.org/document/6958889/>.
195. Hu M, Sim K, Zhou JH, Jiang X, Guan C Brain MRI-based 3D Convolutional Neural Networks for Classification of Schizophrenia and Controls. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) [Internet]. Montreal, QC, Canada: IEEE; 2020 [cited 2023 Apr 14]. p. 1742–5. Available from: <https://ieeexplore.ieee.org/document/9176610/>.
196. Rodrigues AF, Barros M, Furtado P Squizofrenia: Classification and correlation from MRI. In: 2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI) [Internet]. Orlando, FL, USA: IEEE; 2017 [cited 2023 Apr 14]. p. 381–4. Available from: <http://ieeexplore.ieee.org/document/7897285/>.
197. Arbabshirani MR, Castro E, Calhoun VD Accurate classification of schizophrenia patients based on novel resting-state fMRI features. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society [Internet]. Chicago, IL: IEEE; 2014 [cited 2023 Apr 14]. p. 6691–4. Available from: <http://ieeexplore.ieee.org/document/6945163/>.
198. GeethaRamani R, Sivaselvi K Data mining technique for identification of diagnostic biomarker to predict Schizophrenia disorder. In: 2014 IEEE International Conference on Computational Intelligence and Computing Research [Internet]. Coimbatore, India: IEEE; 2014 [cited 2023 Apr 14]. p. 1–8. Available from: <http://ieeexplore.ieee.org/document/7238525/>.
199. Castro E, Gupta CN, Martinez-Ramon M, Calhoun VD, Arbabshirani MR, Turner J Identification of patterns of gray matter abnormalities in schizophrenia using source-based morphometry and bagging. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society [Internet]. Chicago, IL: IEEE; 2014 [cited 2023 Apr 14]. p. 1513–6. Available from: <http://ieeexplore.ieee.org/document/6943889/>.
200. Cortes-Briones JA, Tapia-Rivas NI, D'Souza DC, Estevez PA. Going deep into schizophrenia with artificial intelligence. *Schizophrenia Res.* 2022;245:122–40.
201. de Filippis R, Carbone EA, Gaetano R, Bruni A, Pugliese V, Segura-Garcia C, et al. Machine learning techniques in a structural and functional MRI diagnostic approach in schizophrenia: a systematic review. *Neuropsychiatr Dis Treat.* 2019;15:1605.
202. Vieira S, Pinaya WHL, Mechelli A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neurosci Biobehav Rev.* 2017;74:58–75.
203. Plis SM, Hjelm DR, Salakhutdinov R, Allen EA, Bockholt HJ, Long JD, et al. Deep learning for neuroimaging: a validation study. *Front Neurosci [Internet].* 2014 Aug 20 [cited 2023 Apr 14];8. Available from: <http://journal.frontiersin.org/article/10.3389/fnins.2014.00229/abstract>.
204. Flint C, Cearns M, Opel N, Redlich R, Mehler DMA, Emden D, et al. Systematic misestimation of machine learning performance in neuroimaging studies of depression. *Neuropsychopharmacol.* 2021;46:1510–7.
205. Marinescu RV, Oxtoby NP, Young AL, Bron EE, Toga AW, Weiner MW, et al. TAD-POLE Challenge: Accurate Alzheimer's Disease Prediction Through Crowdsourced Forecasting of Future Data. In: Reikik I, Adeli E, Park SH, editors. Predictive Intelligence in Medicine [Internet]. Cham: Springer International Publishing; 2019 [cited

- 2023 Apr 14]. p. 1–10. (Lecture Notes in Computer Science; vol. 11843). Available from: [http://link.springer.com/10.1007/978-3-030-32281-6\\_1](http://link.springer.com/10.1007/978-3-030-32281-6_1).
206. Mihalik A, Brudfors M, Robu M, Ferreira FS, Lin H, Rau A, et al. ABCD Neurocognitive Prediction Challenge 2019: Predicting Individual Fluid Intelligence Scores from Structural MRI Using Probabilistic Segmentation and Kernel Ridge Regression. In: Pohl KM, Thompson WK, Adeli E, Linguraru MG, editors. Adolescent Brain Cognitive Development Neurocognitive Prediction [Internet]. Cham: Springer International Publishing; 2019 [cited 2023 Apr 14]. p. 133–42. (Lecture Notes in Computer Science; vol. 11791). Available from: [http://link.springer.com/10.1007/978-3-030-31901-4\\_16](http://link.springer.com/10.1007/978-3-030-31901-4_16).
207. Schulz MA, Yeo BTT, Vogelstein JT, Mourao-Miranada J, Kather JN, Kording K, et al. Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets. *Nat Commun.* 2020;11:4238.
208. He T, Kong R, Holmes AJ, Nguyen M, Sabuncu MR, Eickhoff SB, et al. Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *NeuroImage* 2020;206:116276.
209. Eitel F, Schulz MA, Seiler M, Walter H, Ritter K. Promises and pitfalls of deep neural networks in neuroimaging-based psychiatric research. *Exp Neurol.* 2021;339:113608.

## ACKNOWLEDGEMENTS

This work was supported by NIH grants R01MH118370 (CG), and T32MH126368 (KSFD). The content of this work is of the sole responsibility of the authors. The sources of support played no role in the commentary of this work. Special thanks to Larry Hedges for his feedback on the initial stages of project development and the ADAPT and Gratton Lab for their feedback on various states of this project.

## AUTHOR CONTRIBUTIONS

AP: Conceptualization, investigation, data collection, software, formal analysis, visualization, writing (original draft, review, and editing). SF: Investigation, data collection, formal analysis. KSFD: Conceptualization, methodology, writing (review and editing). RN: Methodology and writing (original draft, review and editing). CG: Conceptualization, methodology, investigation, writing (original draft, review, and editing), resources, supervision. VAM: Conceptualization, methodology, investigation, writing (original draft, review, and editing), resources, supervision.

## FUNDING

NIH grant R01MH118370 (CG) and T32MH126368 (KSFD).

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41380-023-02195-9>.

**Correspondence** and requests for materials should be addressed to Alexis Porter.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023