ARTICLE

Check for updates

Robust and replicable functional brain signatures of 22q11.2 deletion syndrome and associated psychosis: a deep neural network-based multi-cohort study

Kaustubh Supekar[1,2]✉, Carlo de los Angeles[1], Srikanth Ryali[1], Leila Kushan[3,4], Charlie Schleifer[4], Gabriela Repetto[5], Nicolas A. Crossley[6,7], Tony Simon[8,9], Carrie E. Bearden[3,4] and Vinod Menon[1,2]✉

A major genetic risk factor for psychosis is 22q11.2 deletion (22q11.2DS). However, robust and replicable functional brain signatures of 22q11.2DS and 22q11.2DS-associated psychosis remain elusive due to small sample sizes and a focus on small single-site cohorts. Here, we identify functional brain signatures of 22q11.2DS and 22q11.2DS-associated psychosis, and their links with idiopathic early psychosis, using one of the largest multi-cohort data to date. We obtained multi-cohort clinical phenotypic and task-free fMRI data from 856 participants (101 22q11.2DS, 120 idiopathic early psychosis, 101 idiopathic autism, 123 idiopathic ADHD, and 411 healthy controls) in a case-control design. A novel spatiotemporal deep neural network (stDNN)-based analysis was applied to the multi-cohort data to identify functional brain signatures of 22q11.2DS and 22q11.2DS-associated psychosis. Next, stDNN was used to test the hypothesis that the functional brain signatures of 22q11.2DS-associated psychosis overlap with idiopathic early psychosis but not with autism and ADHD. stDNN-derived brain signatures distinguished 22q11.2DS from controls, and 22q11.2DS-associated psychosis with very high accuracies (86–94%) in the primary cohort and two fully independent cohorts without additional training. Robust distinguishing features of 22q11.2DS-associated psychosis emerged in the anterior insula node of the salience network and the striatum node of the dopaminergic reward pathway. These features also distinguished individuals with idiopathic early psychosis from controls, but not idiopathic autism or ADHD. Our results reveal that individuals with 22q11.2DS exhibit a highly distinct functional brain organization compared to controls. Additionally, the brain signatures of 22q11.2DS-associated psychosis overlap with those of idiopathic early psychosis in the salience network and dopaminergic reward pathway, providing substantial empirical support for the theoretical aberrant salience-based model of psychosis. Collectively, our findings, replicated across multiple independent cohorts, advance the understanding of 22q11.2DS and associated psychosis, underscoring the value of 22q11.2DS as a genetic model for probing the neurobiological underpinnings of psychosis and its progression.

Molecular Psychiatry; https://doi.org/10.1038/s41380-024-02495-8

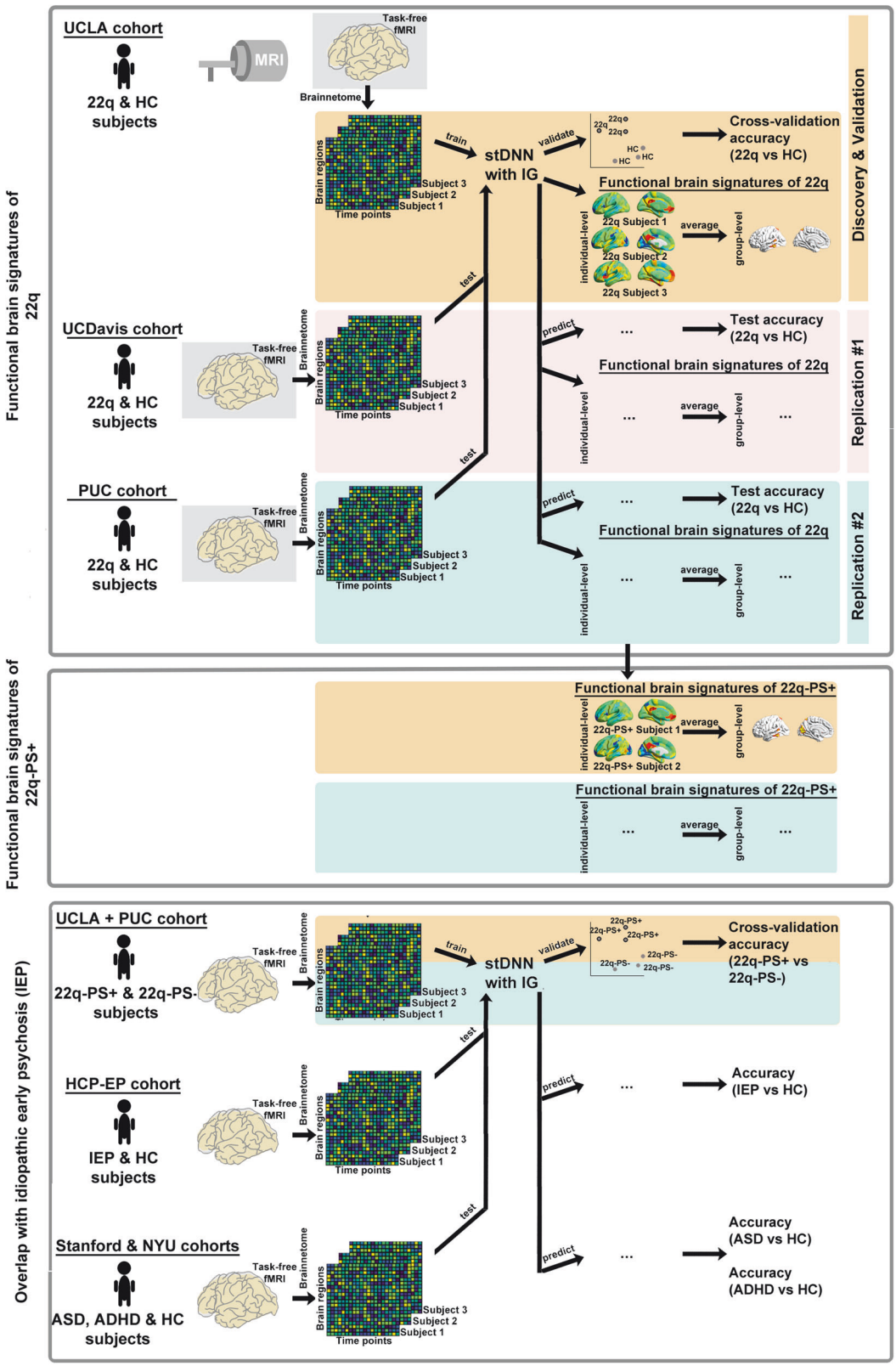
INTRODUCTION

Chromosome 22q11.2 deletion syndrome (22q11.2DS) results from a microdeletion on the long (q) arm of chromosome 22 [1] and occurs in approximately 1 in 2000–4000 live births [2–5]. The microdeletions typically range in size from 1.5 to 3 megabases [6, 7] and include about 46 protein-coding genes, the majority of which are expressed in the brain [8, 9]. Notably, children with 22q11.2DS have a significantly increased risk of developing schizophrenia in adolescence and adulthood [10, 11], making the deletion the highest genetic risk factor for schizophrenia after having an identical twin or two parents with the disorder [12–14]. Thus, 22q11.2DS offers a unique opportunity to investigate early brain signatures in a genetic subtype of schizophrenia that may shed light on the neurodevelopmental mechanisms underpinning the emergence and expression of the disorder [15–17]. Here, we identify robust and replicable functional brain signatures of 22q11.2DS and 22q11.2DS-associated psychosis, as well as their links with idiopathic early psychosis [18], using one of the largest multi-cohort functional brain imaging data to date (see Fig. 1 for study overview).

The first aim of our study was to determine whether individuals with 22q11.2DS differ in their functional brain organization when compared to healthy age-matched controls. Functional brain imaging studies have attempted to identify functional brain signatures of 22q11.2DS. Small studies involving single-site cohorts have found evidence for both increased and decreased functional brain connectivity in 22q11.2DS [19–26]; however, no consistent findings have emerged. These discrepancies may have

[1]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA. [2]Wu Tsai Neurosciences Institute, Stanford University School of Medicine, Stanford, CA, USA. [3]Department of Psychiatry and Behavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA. [4]Department of Psychology, University of California, Los Angeles, Los Angeles, CA, USA. [5]Center for Genetics and Genomics, Facultad de Medicina, Clinica Alemana Universidad del Desarrollo, Santiago, Chile. [6]Department of Psychiatry, Pontificia Universidad Católica de Chile, Santiago, Chile. [7]Department of Psychiatry, University of Oxford, Oxford, UK. [8]Department of Psychiatry and Behavioral Sciences, University of California, Davis, Sacramento, CA, USA. [9]MIND Institute, University of California, Davis, Sacramento, CA, USA. ✉email: ksupekar@stanford.edu; menon@stanford.edu

Received: 19 June 2023 Revised: 22 February 2024 Accepted: 26 February 2024
Published online: 12 April 2024



arisen due to small sample sizes (typically $N = 20–35$ per group) and the focus on a limited set of brain regions [19–21] despite mounting evidence of widespread structural brain abnormalities associated with 22q11.2DS [27–35]. Moreover, most previous 22q11.2DS studies have relied on static functional connectivity measures [19–26]. There is growing evidence that dynamic functional connectivity measures may carry more useful distinguishing features because they are more sensitive to detecting changes in time-varying brain function that often accompany disorders [36–39]. To address these challenges and identify robust

Fig. 1 Study overview. The study identified functional brain signatures of 22q11.2DS and 22q11.2DS-associated psychosis, as well as their links with idiopathic early psychosis, using six independent cohorts and a three-part analysis. In the first part, the functional brain signatures of 22q11.2DS were identified through a discovery/validation analysis in the UCLA cohort and replication analysis in the UCDavis and PUC cohorts. A stDNN model was trained using fMRI timeseries data from brain regions across the entire brain to distinguish between groups, specifically from the UCLA cohort. The performance of the stDNN in distinguishing individuals with 22q11.2DS from healthy controls was evaluated using a cross-validation procedure. The generalization/replicability of the stDNN model trained on the UCLA cohort data to the independent cohort data (UCDavis and PUC) was also examined. The integrated gradients (IG) procedure was then applied to the trained stDNN model to identify functional brain features/signatures that distinguish between 22q11.2DS and healthy controls at individual and group levels in each of the three independent cohorts separately. The second part focused on identifying the functional brain signatures of 22q11.2DS with psychosis spectrum symptoms (22q-PS+). This was accomplished by averaging the IG derived functional brain signatures of 22q-PS+ individuals in each of the three independent cohorts. The third part investigated the overlap between the functional brain signatures of 22q-PS+ and those of idiopathic early psychosis (IEP). A stDNN model was trained de novo, combining data from the UCLA and PUC cohorts, to distinguish between 22q-PS+ and 22q11.2DS without psychosis spectrum symptoms (22q-PS−). Additionally, the stDNN model trained to distinguish between 22q-PS+ and 22q-PS− was tested on the HCP-EP cohort to identify individuals with idiopathic early psychosis. To assess the distinctiveness of functional brain signatures of 22q11.2DS-associated psychosis, Stanford and NYU cohort data were utilized. The stDNN model that differentiated between 22q-PS+ and 22q-PS- was examined to determine its ability to identify individuals with idiopathic autism spectrum disorder and individuals with idiopathic attention-deficit/hyperactivity disorder. 22q Chromosome 22 long arm (q) deletion syndrome, HC healthy controls, 22q-PS+ 22q11.2DS with psychosis spectrum symptoms, 22q-PS- 22q11.2DS without psychosis spectrum symptoms, IEP idiopathic early psychosis, ASD autism spectrum disorder, ADHD attention-deficit hyperactivity disorder, MRI magnetic resonance imaging, fMRI functional MRI, stDNN spatiotemporal deep neural network model, IG integrated gradients, Brainnetome Brainnetome whole brain atlas, UCLA University of California Los Angeles, UCDavis University of California Davis, PUC Pontificia Universidad Católica de Chile, HCP-EP Human Connectome Project-Early Psychosis, NYU New York University.

functional brain signatures of 22q11.2DS, we examined, to the best of our knowledge, one of the largest functional brain imaging data to date of individuals with 22q11.2DS, using a novel end-to-end data-driven spatiotemporal deep neural network (DNN)-based computational framework (Supplementary Fig. 1). Briefly, the stDNN model takes as its input functional magnetic resonance imaging (fMRI) time series data from brain regions of interest spanning the whole brain and models the underlying dynamic spatiotemporal characteristics of brain activity to distinguish between groups with minimal assumptions [40, 41] We trained our stDNN model on large-scale data from the primary cohort and evaluated its performance in distinguishing 22q11.2DS using a cross-validation procedure. Our primary cohort consisted of clinical phenotypic and task-free fMRI data from 22q11.2DS and healthy control participants collected at the University of California, Los Angeles (UCLA). We hypothesized that stDNN would accurately distinguish between individuals with 22q11.2DS and healthy controls in the primary cohort, indicating that the functional brain organization of individuals with 22q11.2DS is distinct from that of healthy controls.

Our second aim was to address the replicability crisis in 22q11.2DS research, and in clinical neuroscience more broadly, arising from small single-site cohort studies that fail to capture the high individual variability observed in the disorder [16]. Crucially, because of the focus on individual cohorts, the robustness and replicability of findings across different cohorts are not known. To address this, we leveraged large-scale data from multiple independent cohorts. We investigated whether the stDNN trained using primary cohort data could generalize to untrained independent cohort data. Typically, most approaches fail at this important step [42]. Our secondary and tertiary cohorts consisted of clinical phenotypic and task-free fMRI data from 22q11.2DS and healthy control participants collected at the University of California, Davis (UCDavis) and Pontificia Universidad Católica de Chile (PUC), respectively. We used the stDNN trained on primary UCLA cohort data and evaluated its performance in distinguishing 22q11.2DS from the other two independent cohorts (UCDavis and PUC). This approach enabled us to address the replicability crisis in 22q11.2DS research. We hypothesized that stDNN would generalize well and accurately distinguish 22q11.2DS in previously unseen data from entirely different independent cohort data.

Our third aim was to identify functional brain features/signatures that distinguish between 22q11.2DS and healthy controls. Prior research on DNNs for brain imaging has primarily focused on classification accuracy, disregarding the brain features underlying the classification. To address this issue, we employed the integrated gradients [43] algorithm, which provides a score indicating the contribution of each feature to the final prediction. Consequently, this algorithm offers a ranking of brain features that distinguish between 22q11.2DS and healthy controls. Specifically, we identified individual-level functional brain signatures of 22q11.2DS by estimating the integrated gradients of the stDNN model trained to distinguish 22q11.2DS from healthy controls. We hypothesized that stDNN, along with integrated gradients, would enable us to identify robust and replicable functional brain organization patterns/signatures that differ between 22q11.2DS and healthy controls across the three independent cohorts.

The fourth aim of our study was to identify robust functional brain signatures of 22q11.2DS with psychosis spectrum symptoms (22q-PS+). Despite the clinical significance of investigating functional brain signatures of 22q-PS+ [44], only a limited number of studies have been conducted thus far. One study reported a relationship between the functional organization of the default mode network and psychosis symptoms in 22q11.2DS [22, 25]. However, Padula et al. did not find evidence supporting such a relationship [19]. Other studies that directly compared 22q-PS+ and 22q11.2DS without psychosis spectrum symptoms (22q-PS−) have also yielded inconclusive results. Some studies have reported that default mode network and fronto-temporal functional organization patterns can be distinguished [23, 24], while another study found that 22q-PS+ and 22q-PS− could not be distinguished with sufficient accuracy using functional brain organization features [26]. To identify robust functional brain signatures of 22q-PS+, it is critical to examine large samples and evaluate whether the signatures are consistently observed across cohorts. Accordingly, we leveraged our large multi-cohort dataset and identified functional brain signatures of 22q-PS+ by averaging the integrated gradients derived functional brain signatures of 22q-PS+ individuals. We determined the distinctiveness of functional brain signatures of 22q-PS+ by contrasting them with functional brain signatures of 22q-PS−. We recently proposed a theoretical salience-based model of psychosis that takes into account striatal dysfunction and sensitivity to perceptual and cognitive prediction errors in the insula node of the salience network. This model posits that dysregulated dopamine modulation of salience network–centered processes contributes to the positive symptoms of schizophrenia [45]. Based on this model we hypothesized that salience network nodes and striatal regions would prominently feature in the functional brain signatures of 22q-PS+.

The final aim of our study was to investigate whether the functional brain signatures of 22q-PS+ overlap with those of idiopathic early psychosis [18]. While there is clinical evidence that the presentation of psychosis symptoms in 22q11.2DS is similar to that in idiopathic psychosis [46, 47], little is known about the common functional brain signatures between 22q-PS+ and idiopathic psychosis. This knowledge is critical not only for elucidating the convergent and divergent elements of brain aberrancies underlying 22q11.2DS-associated psychosis and idiopathic psychosis but also for advancing our understanding of the brain mechanisms involved in the development of psychotic disorders in the broader population [16]. We leveraged our DNN-based framework in conjunction with clinical phenotypic and task-free fMRI data from individuals with idiopathic early psychosis obtained from the Human Connectome Project-Early Psychosis (HCP-EP). We investigated whether our stDNN model trained de novo to distinguish between 22q-PS+ and 22q-PS− could also identify individuals with idiopathic early psychosis from the HCP-EP cohort. Lastly, to determine the distinctiveness of functional brain signatures of 22q11.2DS-associated psychosis, we used clinical phenotypic and task-free fMRI data collected at Stanford and New York University (NYU), and investigated whether the stDNN model that distinguished between 22q-PS+ and 22q-PS− could identify individuals with idiopathic autism and individuals with idiopathic attention-deficit/hyperactivity disorder (ADHD)—two sets of neuropsychiatric symptoms that are highly prevalent in the 22q11.2DS population [14, 48, 49].

RESULTS

Classification of 22q11.2DS versus control subjects in the UCLA cohort

We first sought to determine whether there are differences in functional brain organization between individuals with 22q11.2DS and healthy controls. We found that our stDNN model, which models the underlying dynamic spatiotemporal characteristics of brain activity to distinguish between groups using fMRI timeseries, accurately (>94%) distinguishes individuals with 22q11.2DS from healthy controls in the UCLA cohort, outperforming conventional approaches (see Supplementary Materials for details).

Generalization of 22q11.2DS classification in the UCLA cohort to the independent UCDavis and PUC cohorts

Next, we aimed to determine whether differences in the functional brain organization between individuals with 22q11.2DS and healthy controls observed in the UCLA cohort generalize to independent cohorts. We found that stDNN accurately (=84–90%) distinguishes individuals with 22q11.2DS from healthy controls in a robust and consistent manner across two independent cohorts (UCDavis and PUC cohorts) without additional training, outperforming conventional approaches (see Supplementary Materials for details).

Identification of discriminating brain features underlying 22q11.2DS classification in the UCLA cohort

We then sought to identify brain features that distinguish individuals with 22q11.2DS from healthy controls in the UCLA cohort. To accomplish this, we used the stDNN 22q11.2DS classification model and the integrated gradients procedure, which yields a measure of feature strength associated with 22q11.2DS vs. healthy controls classification in each brain region and at each time point for each individual. To identify brain areas that contributed the most to classification, we computed the median of feature scores across the five folds and thresholded them - top 5% of features - based on the distribution of feature scores across all time points and regions. This resulted in the identification of a distributed set of brain areas, including the anterior insula, striatum (nucleus accumbens, ventromedial putamen and caudate), and inferior temporal gyrus as brain areas that contributed most significantly to predicting the 22q11.2DS class label (Fig. 2, Supplementary Table S19).

Generalization of discriminating brain features underlying 22q11.2DS classification in the UCLA cohort to the independent UCDavis and PUC cohorts

To determine the generalizability of the discriminating features underlying 22q11.2DS classification identified in the UCLA cohort, we applied the same procedures described in the previous section to the independent UCDavis and PUC cohort data. These cohort-wise analyses identified the anterior insula, striatum, and inferior temporal gyrus as the brain areas that contributed most significantly to predicting the 22q11.2DS class label (Fig. 2, Supplementary Tables S20–S21) in both cohorts. Cross-cohort comparison analyses confirmed the consistency of 22q11.2DS discriminating features across the three cohorts ($p < 0.001$, Fig. 2). These results demonstrate that stDNN, together with integrated gradients procedures, automatically identifies similar discriminating features as in the UCLA cohort, again without the need for ad hoc feature engineering.

Distinctiveness of brain features underlying 22q11.2DS classification in the UCLA cohort

We then determined the distinctiveness of brain features underlying 22q11.2DS classification in the UCLA cohort by computing the distance between brain signatures in individuals with 22q11.2DS in contrast to controls. The integrated gradients procedure identifies an individual "signature" of predictive features in each participant. Briefly, a "signature" of an individual refers to the unique whole brain pattern of integrated gradients-derived stDNN model feature importance that classifies that individual as having 22q11.2DS or as a healthy control. We examined whether these "signatures" cluster differently in individuals with 22q11.2DS from healthy controls. Specifically, we computed a distance metric across brain features between individuals and compared the distances between individuals in the 22q11.2DS and healthy control groups. We found that the intra-22q11.2DS group distance metrics were significantly shorter compared to distances with the healthy control group ($p < 0.0001$, Fig. 3). These results demonstrate that individualized brain "signatures" mirrored the broader diagnostic discrimination of 22q11.2DS and that they are distinct.

Distinctiveness of brain features underlying 22q11.2DS classification in the UCDavis and PUC cohorts

We then used the same procedures described in the previous section to determine the distinctiveness of brain features underlying 22q11.2DS classification in the UCDavis and PUC cohorts. These cohort-wise analyses revealed individualized brain "signatures" which mirrored the broader diagnostic discrimination of 22q11.2DS ($p < 0.0001$, Fig. 3), in both cohorts. These results demonstrate the distinctiveness of brain features underlying 22q11.2DS classification in the independent UCDavis and PUC cohorts, as in the UCLA cohort.

Identification of brain features predictive of 22q11.2DS-associated psychosis in the UCLA cohort

We next determined brain features predictive of 22q-PS+ in the UCLA cohort using the output of the integrated gradients procedure – individualized brain signature – described in the previous section. To identify brain areas that were most predictive of 22q-PS+, we computed the mean of brain signatures of all 22q-PS+ and contrasted them against the mean of brain signatures of 22q-PS−. This resulted in the identification of a distributed set of brain areas including the anterior insula and the ventral striatum as brain areas that contributed most significantly to predicting 22q-PS+ (Fig. 4, Supplementary Table S22).

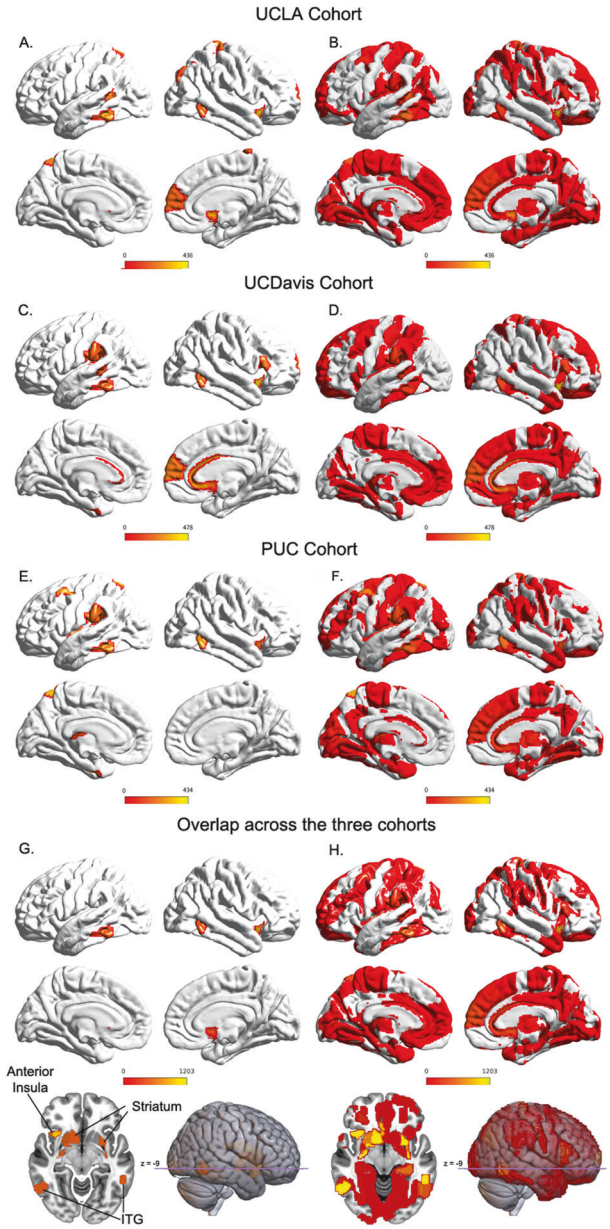


Fig. 2 Group-level functional brain signatures (feature attribution maps) of 22q11.2DS. A Feature attribution map showing the top 5% features that underlie 22q11.2DS vs healthy controls classification in the UCLA cohort. stDNN with integrated gradients identified brain features that distinguish individuals with 22q11.2DS from healthy controls. The algorithm automatically identified distinguishing features in the anterior insula, striatum, and inferior temporal gyrus, which anchor the salience network, dopaminergic reward pathway and ventral visual stream, respectively (see Supplementary Table S19 for a detailed listing of brain areas). **B** Visualization of (unthresholded) feature weights across the whole brain in the UCLA cohort. **C** Feature attribution maps showing the top 5% features showing replication of the predictive the insula, striatum, and inferior temporal gyrus features in the UCDavis cohort (see Supplementary Table S20 for a detailed listing of brain areas). **D** Visualization of (unthresholded) feature weights across the whole brain in the UCDavis cohort. **E** Feature attribution maps showing the top 5% features showing replication of the predictive insula, striatum, and inferior temporal gyrus features in the PUC cohort (see Supplementary Table S21 for a detailed listing of brain areas). **F** Visualization of (unthresholded) feature weights across the whole brain in the PUC cohort. **G** Feature attribution maps showing the top 5% features showing consistency of the predictive insula, striatum, and inferior temporal gyrus features across the three cohorts (UCLA, UCDavis, and PUC). **H** Visualization of (unthresholded) feature weights across the whole brain overlap across the three cohorts. Feature weights scaled for visualization.

Generalization of brain features predictive of 22q11.2DS-associated psychosis in the UCLA cohort to the independent PUC cohort

To determine the generalizability of discriminating features predictive of 22q-PS+ identified in UCLA data, we applied the same procedures as described in the previous section to the independent PUC cohort data. Herein we focused our analysis on the PUC cohort since none of the 22q11.2DS participants in the UCDavis cohort endorsed psychosis spectrum symptoms. This analysis identified the anterior insula and the ventral striatum as the brain areas that contributed most significantly to predicting 22q-PS+ (Fig. 4, Supplementary Table S23). Cross-cohort comparison analyses confirmed the consistency of features predictive of 22q-PS+ across the two cohorts ($p < 0.001$, Fig. 4). These results demonstrate that stDNN, together with integrated gradients procedures, automatically identifies similar features predictive of 22q-PS+ as in the UCLA cohort, again without the need for ad hoc feature engineering.

Distinctiveness of brain features predictive of 22q11.2DS-associated psychosis in the UCLA cohort

We then determined the distinctiveness of brain features predictive of 22q11.2DS-associated psychosis in the UCLA cohort by computing a distance metric across brain features between individuals and comparing the distances between individuals in the 22q-PS+ and 22q-PS− groups. We found that intra-22q-PS+ group distance metrics were significantly shorter compared to distances with the 22q-PS−group ($p < 0.01$, Fig. 5). These results demonstrate the distinctiveness of brain features predictive of 22q-PS + .

Distinctiveness of brain features predictive of 22q11.2DS-associated psychosis in the PUC cohort

We then used the same procedures described in the previous section to determine the distinctiveness of brain features predictive of 22q-PS+ in the PUC cohort. This analysis revealed that the intra-22q-PS+ group distance metrics were significantly shorter compared to distances with the 22q-PS− group ($p < 0.01$, Fig. 5). These results demonstrate the distinctiveness of brain features predictive of 22q-PS+ in the independent PUC cohort, as in the UCLA cohort.

Examination of the relationship between the brain features predictive of 22q11.2DS-associated psychosis and brain features predictive of idiopathic early psychosis

To determine the overlap between 22q-PS+ and idiopathic early psychosis, we computed a distance metric across brain features between individuals and compared the distances between individuals in the 22q-PS+, 22q-PS−, and idiopathic early psychosis groups. We found that distances between the 22q-PS + group and the idiopathic early psychosis group were significantly shorter than distances between the 22q-PS− and the idiopathic early psychosis group ($p < 0.01$) in the combined UCLA and PUC cohort, indicating overlap between brain features of 22q-PS+ and idiopathic early psychosis.

To further examine the overlap between 22q-PS+ and idiopathic early psychosis, we trained a de novo model to distinguish between 22q-PS+ and 22q-PS− by combining 22q11.2DS data from the UCLA and PUC cohorts and then investigated whether the stDNN

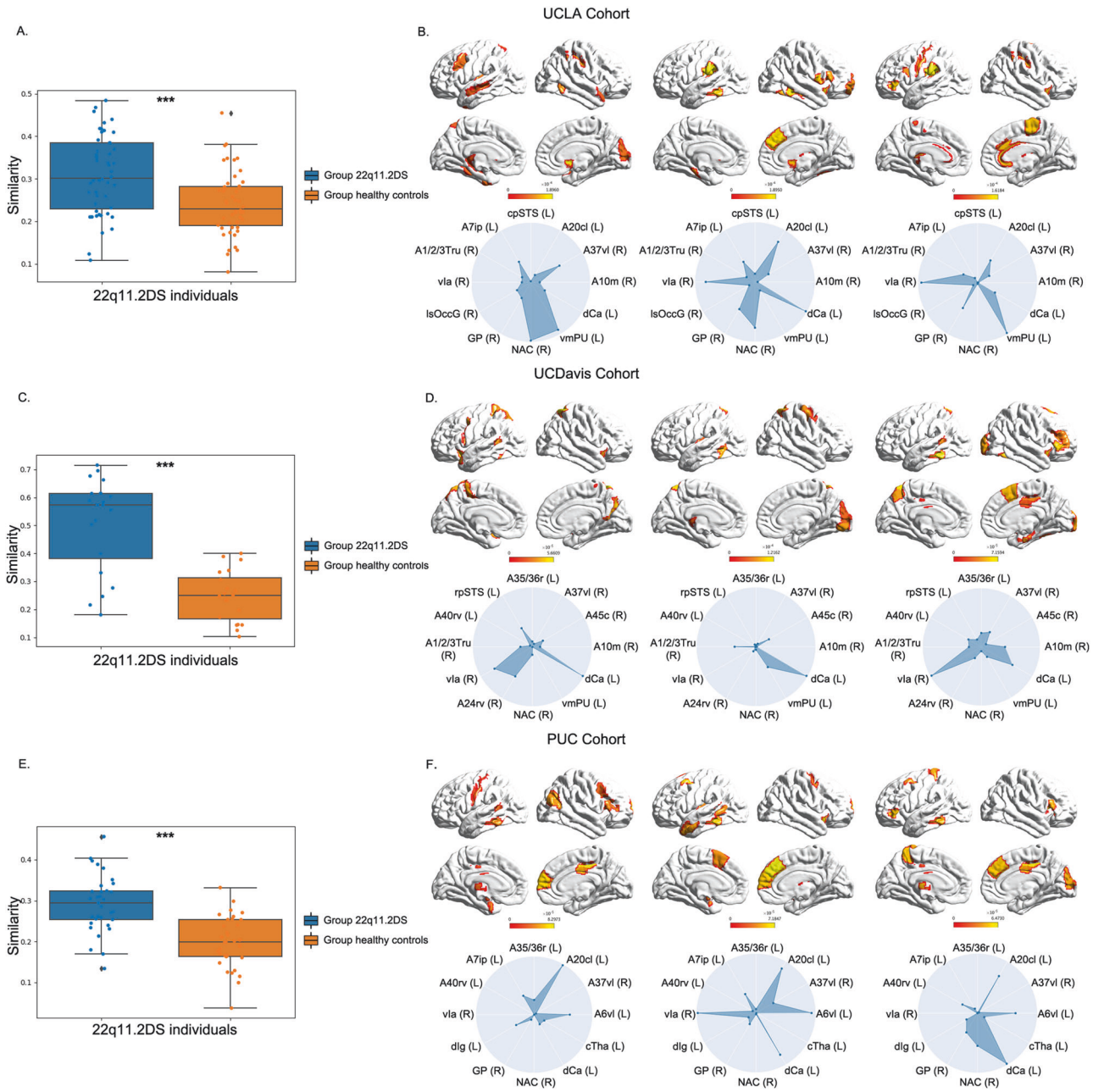


Fig. 3 Individual functional brain signatures (feature attribution maps) of 22q11.2DS and their distinctiveness. A Box plot of similarity between individual 22q11.2DS signatures and group-level 22q11.2DS signatures, and between individual 22q11.2DS signatures and group-level healthy controls signatures in the UCLA cohort, demonstrating that 22q11.2DS signatures are distinct from healthy control signatures. **B** stDNN-derived individual feature attribution maps/signatures in three 22q11.2DS individuals randomly selected from the UCLA cohort. **C** Box plot of similarity between individual 22q11.2DS signatures and group-level 22q11.2DS signatures, and between individual 22q11.2DS signatures and group-level healthy controls signatures in the UCDavis cohort, demonstrating that 22q11.2DS signatures are distinct from healthy control signatures. **D** stDNN-derived individual feature attribution maps/signatures in three 22q11.2DS individuals randomly selected from the UCDavis cohort. **E** Box plot of similarity between individual 22q11.2DS signatures and group-level 22q11.2DS signatures, and between individual 22q11.2DS signatures and group-level healthy controls signatures in the PUC cohort, demonstrating that 22q11.2DS signatures are distinct from healthy control signatures. **F** stDNN-derived individual feature attribution maps/signatures in three 22q11.2DS individuals randomly selected from the PUC cohort. ***: $p < 0.001$. A10m medial area 10, A6vl ventrolateral area 6, A45c caudal area 45, A37vl ventrolateral area 37, A20cl caudolateral of area 20, A35/36r rostral area 35/36, rpSTS rostroposterior superior temporal sulcus, cpSTS caudoposterior superior temporal sulcus, A7ip intraparietal area 7(hIP3), A40rv rostroventral area 40(PFop), A1/2/3tru area1/2/3(trunk region), vIa ventral anterior insular, dIg dorsal granular insular, A24rv rostroventral area 24, lsOccG lateral superior occipital gyrus, GP globus pallidus, NAC nucleus accumbens, vmPu ventromedial putamen, dCa dorsal caudate, cTtha caudal Temporal thalamus, L left, R right.

model trained to distinguish between 22q-PS+ and 22q-PS− can identify individuals with idiopathic early psychosis from the HCP-EP cohort (Supplementary Table S24). In this analysis, we combined 22q11.2DS data from the UCLA and PUC cohorts to increase the sample size. Importantly, we de novo trained our stDNN model here using a novel label-distribution-aware margin (LDAM) loss as our combined dataset has a small number of 22q-PS+ participants compared to 22q-PS− participants. We have previously shown

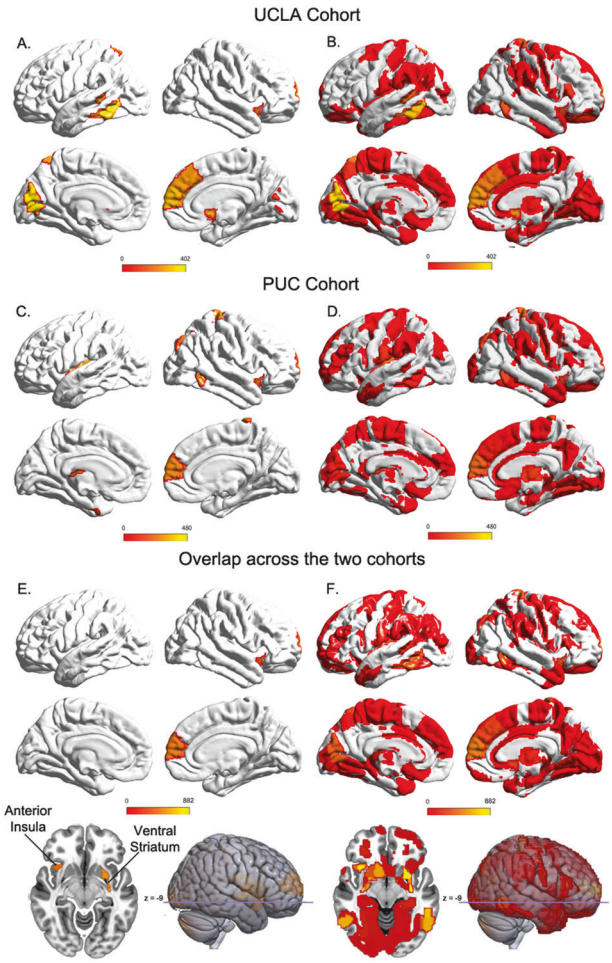


Fig. 4 Group-level functional brain signatures (feature attribution maps) of 22q11.2DS with psychosis. A Feature attribution map showing the top 5% features that underlie prediction of 22q11.2DS with psychosis in the UCLA cohort. Our approach automatically identified features in the anterior insula and the ventral striatum, as those associated with 22q11.2DS with psychosis (see Supplementary Table S22 for a detailed listing of brain areas). **B** Visualization of (unthresholded) feature weights across the whole brain in the UCLA cohort. **C** Feature attribution maps showing the top 5% features showing replication of the anterior insula and ventral striatum features in the PUC cohort (see Supplementary Table S23 for a detailed listing of brain areas). **D** Visualization of (unthresholded) feature weights across the whole brain in the PUC cohort. **E** Feature attribution maps showing the top 5% features showing consistency of the predictive anterior insula and striatum features across the two cohorts (UCLA and PUC). **F** Visualization of (unthresholded) feature weights across the whole brain overlap across the two cohorts. Feature weights scaled for visualization.

LDAM loss to outperform conventional cross-entropy loss under such class-imbalance conditions. For 22q-PS+ vs. 22q-PS− classification, the stDNN model de novo trained with LDAM loss using the combined UCLA PUC cohort data achieved an accuracy of 84.0 ± 4.2% across the 5 folds, and an average precision of 0.87 ± 0.02, recall of 0.84 ± 0.03 and F1 score of 0.81 ± 0.04 (Supplementary Table S25).

Furthermore, the stDNN model that most accurately distinguished between 22q-PS+ and 22q-PS− could identify individuals with idiopathic early psychosis with an accuracy of 77.5% in the HCP-EP cohort. To determine the model's specificity, we investigated whether the stDNN model trained to distinguish between 22q-PS+ and 22q-PS− could identify individuals with idiopathic autism and individuals with idiopathic ADHD – two sets of neuropsychiatric symptoms that are highly prevalent in the 22q11.2DS population – using autism and ADHD data from Stanford (Supplementary Table S26) and NYU (Supplementary Table S27) cohorts respectively. We found that the stDNN model trained to most accurately distinguish between 22q-PS+ and 22q-PS− could identify individuals with idiopathic autism with chance-level accuracy (=48.5%) and identify individuals with idiopathic ADHD with chance-level accuracy (=43.9%), highlighting the uniqueness of the brain features associated with 22q11.2DS-associated psychosis vis-à-vis those associated with idiopathic autism and idiopathic ADHD.

Taken together, these results establish the distinctness of brain features associated with psychosis spectrum symptoms in 22q11.2DS and provide evidence that they specifically overlap with idiopathic early psychosis and not with idiopathic autism and idiopathic ADHD.

DISCUSSION

In this study, one of the most extensive multi-cohort functional brain imaging investigations of 22q11.2DS to date, we identified robust and replicable functional brain signatures of 22q11.2DS and associated psychosis using a novel stDNN-based approach. Specifically, stDNN-derived functional brain features associated with the salience network, striatal dopaminergic pathway, and ventral visual stream accurately differentiated individuals with 22q11.2DS from healthy controls. Furthermore, stDNN-derived brain features associated with the anterior insula and the ventral striatum accurately distinguished 22q11.2DS-associated psychosis, achieving high classification rates between the 22q-PS+ and 22q-PS− subgroups. Notably, the brain features that distinguished 22q-PS+ significantly overlapped with those of idiopathic early psychosis, but not with those of idiopathic autism or ADHD. Crucially, these results were observed in the primary UCLA cohort and replicated in two fully independent cohorts (UCDavis and PUC). Taken together, our robust and reproducible findings provide novel insights into the neurobiology of 22q11.2DS and demonstrate the similarity of functional brain signatures between 22q-PS+ and idiopathic early psychosis. These findings are critical for improving the accuracy of psychosis risk prediction in individuals with 22q11.2DS and identifying brain-circuit targets for clinical interventions to mitigate the risk of developing psychosis.

The first aim of our study was to determine whether individuals with 22q11.2DS differ in their functional brain organization compared to healthy age-matched controls. We developed an stDNN model that takes as input fMRI timeseries from regions of interest spanning the entire brain and determined distinguishing functional brain signatures by modeling latent dynamic time-varying functional interactions among the brain regions. Our model uncovered robust differences between 22q11.2DS and healthy controls, achieving a cross-validation classification accuracy of 93.5 ± 4.04%, and high precision and recall in the primary cohort, outperforming previous studies. Importantly, we observed tight standard deviation bounds on cross-validation classification accuracy across folds, indicating, robust classification. Together, these findings reveal robust differences in functional brain organization between 22q11.2DS and healthy controls, and demonstrate the strengths of our DNN-based approach.

The second aim of our study was to address the replicability crisis in clinical neuroscience [50], and determine if the identified differences between 22q11.2DS and healthy controls could be generalized to independent participant cohorts. To date, reported functional brain organization abnormalities in 22q11.2DS have been inconsistent, and to the best of our knowledge, no studies have attempted model-based validation and replication across multiple independent cohorts. stDNN achieved high classification accuracy in the independent UCDavis and PUC cohorts. Notably,

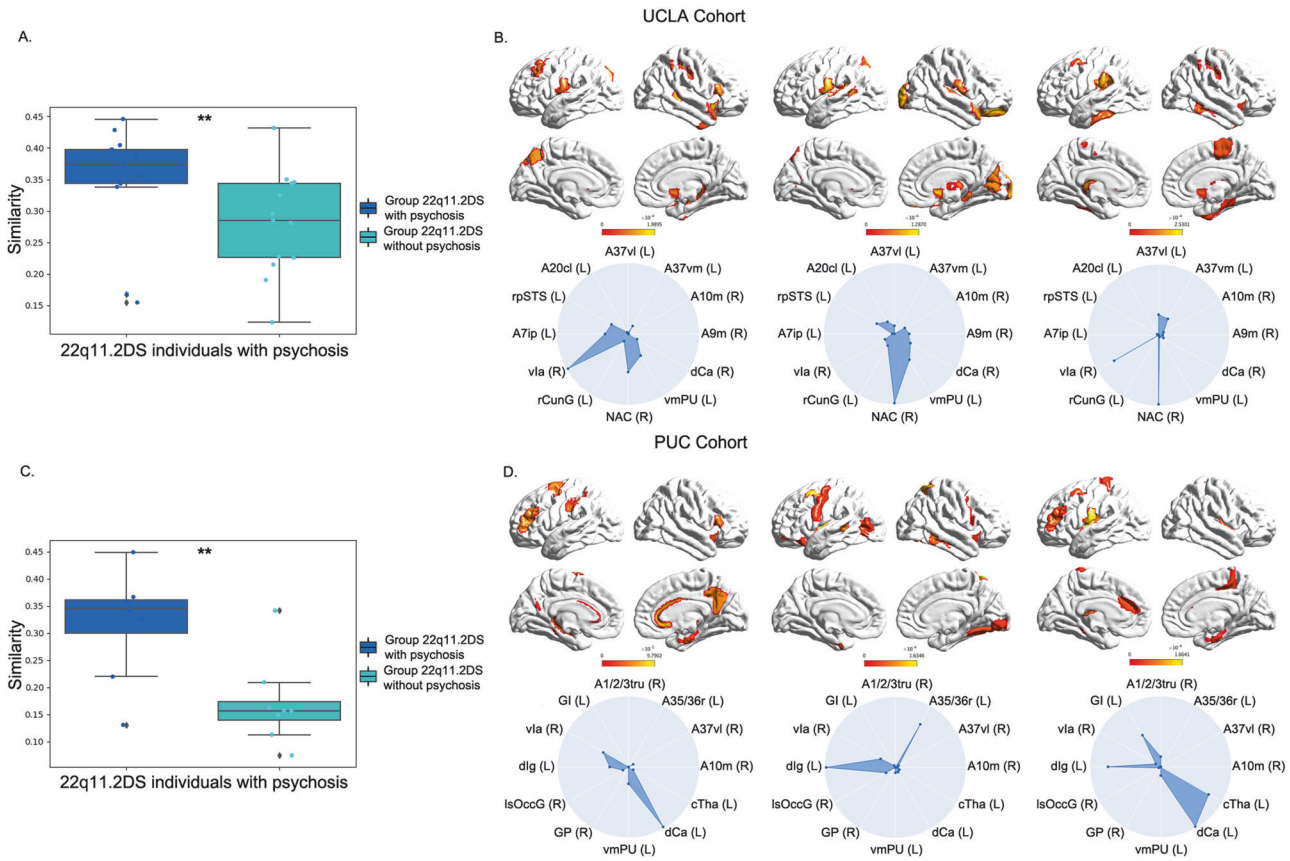


Fig. 5 Individual functional brain signatures (feature attribution maps) of 22q11.2DS with psychosis spectrum symptoms (22q-PS+) and their distinctiveness. A Box plot of similarity between individual 22q-PS+ signatures and group-level 22q-PS+ signatures, and between individual 22q-PS+ signatures and group-level 22q-PS− signatures in the UCLA cohort, demonstrating that 22q-PS+ are distinct from 22q-PS− signatures. **B** stDNN-derived individual feature attribution maps/signatures in three 22q-PS+ individuals randomly selected from the UCLA cohort. **C** Box plot of similarity between individual 22q-PS+ signatures and group-level 22q-PS+ signatures, and between individual 22q-PS+ signatures and group-level 22q-PS− signatures in the PUC cohort, demonstrating that 22q-PS+ signatures are distinct from 22q-PS− signatures. **D** stDNN-derived individual feature attribution maps/signatures in three 22q-PS+ individuals randomly selected from the PUC cohort. **: $p < 0.01$. A9m medial area 9, A10m medial area 10, A37vm ventromedial area37, A37vl ventrolateral area 37, A37vl ventrolateral area 37, A20cl caudolateral of area 20, A35/36r rostral area 35/36, rpSTS rostroposterior superior temporal sulcus, A7ip intraparietal area 7(hIP3), A1/2/3tru area1/2/3(trunk region), G hypergranular insular, vla ventral anterior insular, dlg dorsal granular insular, rCunG rostral cuneus gyrus, lsOccG lateral superior occipital gyrus, GP globus pallidus, NAC nucleus accumbens, vmPu ventromedial putamen, dCa dorsal caudate, cTtha caudal Temporal thalamus, L left, R right.

the UCDavis and PUC cohort data were not used for training the stDNN model, and therefore served as fully independent datasets for demonstrating the generalizability of our stDNN 22q11.2DS model. Furthermore, stDNN significantly outperformed conventional approaches that employed static functional connectivity features and regional time series features for 22q11.2DS classification. These results suggest that the abnormal intrinsic spatiotemporal brain dynamics, captured by our stDNN model, represent a robust and replicable functional brain signature of 22q11.2DS [51].

The third aim of our study was to identify functional brain features that distinguish between individuals with 22q11.2DS and healthy controls. Our stDNN-based integrated gradients analysis identified the anterior insula, which anchors the salience network [52], as a brain area whose intrinsic spatiotemporal dynamics most clearly distinguished between 22q11.2DS and healthy controls. Additionally, the striatum, a crucial node of the reward pathway [53–60], and the inferior temporal gyrus, encompassing the ventral visual stream pathway [61], also emerged as brain areas with distinctive intrinsic spatiotemporal dynamics. Crucially, these features were observed in the UCLA cohort and independently replicated in the independent UCDavis and PUC cohorts, attesting to the robustness and generalizability of our findings. These findings provide evidence of functional circuit alterations associated with previously reported gray matter abnormalities in the insula [33] and striatum [34] as well as structural connectivity between brain regions involved in visuo-spatial processing [62–65]. The evidence from functional brain imaging studies has been mixed: one study found that abnormal time-averaged intrinsic connectivity of the salience network distinguished 22q11.2DS, albeit with poor accuracy in an independent cohort [24], while another study found no abnormalities in intrinsic functional connectivity of the salience network [66]. Our findings offer replicable evidence for abnormal intrinsic functional dynamics of the salience network and provide novel evidence for abnormalities in intrinsic functional dynamics of the mesolimbic reward and ventral visual stream pathways in 22q11.2DS across three independent cohorts.

The fourth major goal of our study was to identify functional brain signatures of 22q11.2DS-associated psychosis. Despite 22q11.2DS being one of the most prominent genetic risk factors for psychosis, the functional brain mechanisms underlying these symptoms remain unclear. To address this knowledge gap, we analyzed the functional brain fingerprints of 22q11.2DS individuals with and without psychosis spectrum symptoms (22q-PS+ and 22q-PS-). Notably, the anterior insula node of the salience network

and the ventral striatum node of the mesolimbic reward pathway distinguished individuals with 22q-PS+.

The anterior insula is involved in attributing salience to both external stimuli and internal events, modulating context-dependent interactions between the frontoparietal network and default mode networks [67]. A proposed key mechanism underlying psychosis is the misattribution of salience, where irrelevant events are assigned undue significance, leading to a cascade of aberrant switching processes and salience network-centered brain dynamics [68, 69]. Recent findings have strongly supported the contribution of these aberrant salience network dynamics to positive psychosis symptoms in individuals with idiopathic schizophrenia [70]. Our current results align with these findings, and highlight the pivotal role of aberrant salience network dynamics in psychosis in 22q11.2DS.

Dysregulation of the dopaminergic system is another leading theoretical model for understanding schizophrenia [71]. Central to this theory is the idea that impaired prediction errors, arising from aberrancies in dopamine tone, are fundamental mechanisms underlying the misattribution of salience to irrelevant stimuli [72–76]. This can result in the formation of erroneous associations about irrelevant or neutral information, potentially leading to psychosis. Research has shown that both individuals with schizophrenia and those at high risk for developing the disorder exhibit elevated dopamine synthesis capacities and availability in the striatum [77, 78]. Furthermore, recent PET studies have indicated dopamine dysregulation within the striatum in 22q11.2DS and observed that the degree of dysregulation correlates with the severity of psychosis-risk symptoms and predicts the onset of psychosis in individuals with 22q11.2DS [79]. There is growing evidence for the modulatory influence of striatal dopamine on the functional brain network organization of the salience network [80]. Our findings contribute to this body of work by providing novel evidence that aberrant intrinsic dynamics of the striatum and salience network, stemming from dysregulated striatal dopamine modulation, is a robust brain phenotype of psychosis in 22q11.2DS.

The final goal of our study was to determine the extent of overlap between the functional brain signatures of 22q11.2DS-associated psychosis and idiopathic early psychosis. Despite the similar clinical manifestations of psychosis symptoms in 22q11.2DS and idiopathic psychosis [46, 47], the commonalities in their functional brain signatures remain unknown. stDNN enabled us to seamlessly determine whether the functional brain patterns that distinguished the 22q-PS+ and 22q-PS- groups also distinguished individuals with idiopathic early psychosis. Analyzing the similarities between 22q11.2DS and idiopathic early psychosis offers several advantages over earlier investigations that focused on comparing 22q11.2DS with established schizophrenia in older adults [33, 81], as those studies were likely influenced by confounding factors such as age, disease stage, and medication use. Critically, our stDNN modeling provided a unified framework for elucidating whether a subtype of 22q11.2DS could serve as a suitable model for studying the functional neurobiology of psychosis.

We found that the stDNN model initially trained to differentiate individuals in the 22q-PS+ group, could also, without any additional training of the model, distinguish individuals with idiopathic early psychosis but not idiopathic autism and ADHD. Additionally, distances between the functional brain fingerprints of the 22q-PS+ and idiopathic early psychosis groups were significantly shorter compared to those between the 22q-PS− and idiopathic early psychosis groups. These results provide novel evidence for an overlap in functional signatures associated with psychosis in 22q11.2DS and idiopathic early psychosis, and are aligned with findings based on brain anatomy [33, 81–84] and clinical phenotypic features [46, 47].

Our robust methodology allowed us to discover and characterize a pattern in 22q11.2DS-associated psychosis that overlaps with idiopathic early psychosis, reflecting common patterns of abnormal intrinsic salience network and striatum dynamics underlying psychosis. Collectively, these findings lend substantial empirical support to theoretical models of psychosis [45], which suggest that the clinical manifestation of positive symptoms arises from an exaggerated attribution of significance to both external and internal stimuli. This aberrant attribution is thought to be a result of dysregulated striatal dopamine modulation, which consequently impacts the dynamics of the salience network [45].

Limitations

The study has a few limitations that merit discussion. First, despite consolidating data from three independent cohorts to create one of the largest fMRI datasets for 22q11.2DS, the overall sample size remains modest, in part due to the low incidence rate of the disorder. Second, our model did not take account for genetic factors such as deletion size that could potentially influence brain features associated with 22q11.2DS and its related psychosis. Another limitation is the wide age range of participants, which introduces heterogeneity in developmental stages. Further work is needed to explore the impact of genetic and cognitive factors using larger sample sizes that capture heterogeneities within the 22q11.2DS population, including more focused age-specific analyses. Finally, longitudinal data are needed to understand how psychosis in 22q11.2DS progresses over time. Future studies should aim to include longitudinal data to better understand the prognostic value of the brain signatures identified in our study.

CONCLUSIONS

We developed novel DNNs that revealed robust and replicable functional brain signatures of 22q11.2DS and associated psychosis, enriching our understanding of the disorder's neurobiological underpinnings. Aberrancies in the salience network, striatal dopaminergic pathway, and ventral visual stream were prominent features of 22q11.2DS. Moreover, aberrancies in the intrinsic spatiotemporal dynamics of the anterior insula and ventral striatum emerged as a robust functional brain signature of 22q11.2DS-associated psychosis. Importantly, we discovered significant overlap between the functional brain signatures of 22q11.2DS-associated psychosis and idiopathic early psychosis. This commonality points to shared underlying brain mechanisms, which could serve as early biomarkers for psychosis and potential brain targets for effective interventions. Despite the wide range of ages, symptom profiles, disease stages, and data acquisition protocols across cohorts/sites, we successfully replicated our main classification findings and functional brain signatures of 22q11.2DS and 22q11.2DS associated psychosis in multiple independent cohorts, demonstrating their robustness and developmental stability. The robustness of our findings also further underscores the potential of our novel DNN-based approach to uncover commonalities between neurogenetic and psychiatric disorders. Overall, our findings significantly advance our understanding of 22q11.2DS and 22q11.2DS-associated psychosis, and highlight the value of 22q11.2DS as a genetic model for investigating the neurobiological underpinnings of psychosis and its progression.

MATERIALS/SUBJECTS AND METHODS

Study cohorts and participants

UCLA. We leveraged neuroimaging and phenotypic data from 22q11.2DS and healthy control participants collected at the UCLA. The diagnosis of 22q11.2 microdeletion was confirmed molecularly. Exclusion criteria for all study participants were additional neurological or medical condition that might affect neuroimaging measures, insufficient fluency in English, substance or alcohol abuse and/ or dependence within the past

Table 1. Demographic information for the 22q11.2DS and healthy control groups in the UCLA, UCDavis and PUC cohorts.

	UCLA		UCDavis		PUC	
	22q11.2DS (*N* = 48)	**Controls (*N* = 43)**	**22q11.2DS (*N* = 20)**	**Controls (*N* = 17)**	**22q11.2DS (*N* = 33)**	**Controls (*N* = 32)**
Age (in years)	16.5 ± 6.9 (range: 6–39)	13.0 ± 5.1 (range: 6–27)	15.1 ± 2.2 (range: 12–18)	14.7 ± 1.8 (range: 12–18)	18.2 ± 2.2 (range: 15–23)	21.2 ± 2.2 (range: 15–24)
Gender (female/male)	26/22	20/23	8/12	10/7	16/17	22/10
Psychosis Status (no/yes)	35/13	43/0	20/0	17/0	25/8	32/0
Medication Status	9 antidepressant, 5 antipsychotic	0 antidepressant, 0 antipsychotic	3 antidepressant, 0 antipsychotic	0 antidepressant, 0 antipsychotic	0 antidepressant, 4 antipsychotic	0 antidepressant, 0 antipsychotic
IQ	78 ± 2.0 (range: 55–109)	112 ± 3.5 (range: 73–148)	82 ± 2.3 (range: 64–107)	115 ± 2.8 (range: 97–134)	67 ± 2.2 (range: 41–93)	[a]

[a]Data not available.

6 months, and/or any condition that is a contraindication for MRI (pregnancy, claustrophobia, etc.). Healthy controls additionally did not meet criteria for any major mental disorder, based on information gathered during administration of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID), with an additional developmental disorders module, for participants over the age of 16 years, and/or the Computerized Diagnostic Interview Schedule for Children for participants aged ≤16 years. Healthy controls additionally could not meet criteria for a prodromal state, as assessed by the Structured Interview for Prodromal Syndromes (SIPS) [85]. 22q-PS+ were categorically defined based on the presence of any positive symptom rated in the prodromal or psychotic range, that is, a rating of 3 or higher on any item in the positive symptom subscale of the SIPS. 22q-PS− individuals was defined as having no positive symptoms in the prodromal/psychotic range. Table 1 shows demographic information.

UCDavis. An independent cohort of participants recruited and scanned at the University of California Davis was used to determine the robustness of brain signatures of 22q11.2DS and associated psychosis, identified using the UCLA (primary) cohort data. The diagnosis of 22q11.2 microdeletion was confirmed molecularly. Exclusion criteria for all study participants include contraindication for an MRI. Healthy controls additionally did not meet criteria for psychiatric disorders. 22q-PS+ individuals were categorically defined based on the presence of any positive symptom rated in the prodromal or psychotic range, that is, a rating of 3 or higher on any item in the positive symptom subscale of the SIPS [85]. 22q-PS− individuals were defined as having no positive symptoms in the prodromal /psychotic range. Table 1 shows demographic information.

PUC. An independent cohort of participants recruited and scanned at the Pontificia Universidad Católica de Chile was used to determine the robustness of brain signatures of 22q11.2DS and associated psychosis, identified using the UCLA (primary) cohort data. The diagnosis of 22q11.2 microdeletion was confirmed molecularly. Exclusion criteria for all study participants include contraindication for an MRI. Healthy controls additionally did not meet criteria for psychiatric disorders. All healthy controls also underwent molecular analysis to confirm that they were not carriers of the deletion. 22q-PS+ individuals were categorically defined as those 22q11.2DS individuals who fulfilled criteria for a present or lifetime psychotic disorder according to the MINI. Presence of psychotic symptoms when scanned where determined by the Positive and Negative Syndrome Scale (PANSS) [86]. Table 1 shows demographic information.

HCP-EP. An independent cohort of participants acquired by the Human Connectome Project for Early Psychosis (HCP-EP) and shared through the National Institutes of Mental Health Data Archive was used to investigate the overlap between brain signatures of 22q-PS+ is and those associated with idiopathic early psychosis. The cohort information has been described in detail elsewhere (https://www.humanconnectome.org/study/human-connectome-project-for-early-psychosis). Briefly, early psychosis participants were identified as those within the first 3 years of onset of psychotic symptoms. The diagnosis of affective/nonaffective psychosis was based on DSM-V. Supplementary Table S24 shows demographic information.

Stanford. An independent cohort of participants recruited and scanned at Stanford University was used to investigate was used to determine the distinctiveness of the brain signatures of 22q-PS+. The cohort information has been described in detail elsewhere [40]. Briefly, ASD diagnosis was assessed using the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-

Revised (ADI-R) as described in our published studies [87–89]. Supplementary Table S26 shows demographic information.

NYU. An independent cohort of participants recruited and scanned at NYU and made available through the ADHD200 consortium was used to determine the distinctiveness of the brain signatures of 22q-PS+. The cohort information has been described in detail elsewhere [40]. Briefly, the Schedule of Affective Disorders and Schizophrenia for Children–Present and Lifetime Version (KSADS-PL) was used for ADHD diagnosis. Supplementary Table S27 shows demographic information.

fMRI acquisition

UCLA. Structural MRI and resting state functional scans for the UCLA cohort were acquired at either the Ahmanson–Lovelace Brain Mapping Center (BMC) or the Staglin Center for Cognitive Neuroscience. Both sites had an identical 3 T Siemens Tim Trio system, using a 12-channel head coil. 22q11DS subjects and control subjects were split equally between the 2 scanner sites, the scanning protocols implemented at each site were identical and analysis of potential between scanner differences revealed no regions of differential cortical activation between scanner locations. The primary structural scan used for registration purposes consisted of a matched-bandwidth high resolution T1 image (voxel size $1.5 \times 1.5 \times 4.0$ mm^3,echotime[TE] = 34 ms, repetition time [TR] = 5000 ms, echo spacing = 0.89 ms, 34 axial slices, slice thickness 4.0 mm, slice spacing 0 mm, flip angle 90°, field of view [FOV] = 210, matrix size = 128×128). Subsequently, a 5-min resting state functional scan was acquired, during which a black screen was presented and participants were instructed to keep their eyes open, remain relaxed, and attempt to avoid falling asleep. The resting state scan consisted of 152 BOLD 3D images (voxel size $3.0 \times 3.0 \times 4.0$ mm^3,TE = 30 ms, TR = 2000 ms, echo spacing = 0.79 ms, 34 axial slices, slice thickness 4.0 mm, slice spacing 0 mm,flip angle 90°,FOV = 192, matrix size = 64×64).

UCDavis. Structural MRI and resting state functional scans for the UCDavis cohort were acquired at the UCDavis Imaging Research Center on a 3-Tesla Siemens Trio MRI scanner using a 32-channel whole-head coil. The primary structural scan used for registration purposes consisted of magnetization prepared rapid gradient echo (MPRAGE) pulse sequence T1 image (voxel resolution = 1 mm^3, matrix size = 256×256, slice direction = sagittal, number of image slices = 192, TR = 2170 ms, TE = 4.82 ms, and flip angle = 7). Subsequently, a 4-min resting state functional scan was acquired, where the participants were instructed to keep their eyes closed and not move during the scan. The resting state scan consisted of 120 BOLD 3D images of 31 axial slices that were acquired parallel to the plane connecting the anterior and posterior commissures.

PUC. Structural MRI and resting state functional scans for the PUC cohort were acquired at the Pontificia Universidad Católica de Chile on a 3-Tesla Philips Ingenia scanner. The primary structural scan used for registration purposes consisted of TFE pulse sequence T1 image (voxel resolution = 1 mm^3, number of slices = 341, direction of acquisition = sagittal, TR = 7.7 ms, TE = 3.5 ms, FA = 8° and TI = 965.3 ms). Subsequently, an 8-min resting state functional scan was acquired, where the participants were instructed to remain calm with their eyes opened. The resting state scan consisted of 200 BOLD 3D images (TR = 2.5 s, TE = 32 ms, FA = 82°, FOV = 220×220 mm, acquired voxel size = $2.75 \times 2.75 \times 3.00$ mm^3 and number of slices = 40).

HCP-EP. Structural MRI and resting state functional scans for the HCP-EP cohort were acquired on a 3 T Siemens Magnetom Prisma scanner. The primary structural scan used for registration purposes consisted of T1 image (voxel resolution = 0.8 mm3, TR = 2.4 ms; TE = 2.2 ms; flip angle = 8°; field of view = 256 mm; 208 axial slices (0.8 mm thickness)). Subsequently, a 5.5-min resting state functional scan was acquired, where the participants were instructed to stay awake, keep their eyes closed and try not to move for the duration of the scan. The resting state scan consisted of 410 BOLD 3D images (TR = 800 ms, TE = 37 ms, flip angle = 52°, FOV = 208 mm, multiband, acquired voxel size = $2 \times 2 \times 2$ mm^3 and number of slices = 72).

Stanford. Structural MRI and resting state functional scans for the Stanford cohort were acquired at the Lucas Imaging Center on a 3 T General Electric (GE) Signa scanner using a custom-built head coil. The primary structural scan used for registration purposes consisted of MPRAGE pulse sequence T1 image (TR = 5.9 ms; TE = minimum; flip angle = 11°; field of view = 240 mm; matrix size = 256×192; 170 axial slices (1.0 mm thickness)). Subsequently, a 6-min resting state functional scan was acquired, where the participants were instructed to stay awake, keep their eyes closed and try not to move for the duration of the scan. The resting state scan consisted of 240 BOLD 3D images (TR = 2000 ms, TE = 30 ms, flip angle = 80°, 1 interleave, FOV = 20 cm, and the matrix size was 64×64, providing an in-plane spatial resolution of 3.125 mm).

NYU. Structural MRI and resting state functional scans for the HCP-EP cohort were acquired on a 3 T Siemens Magnetom Allegro syngo MR scanner. The primary structural scan used for registration purposes consisted of T1 image (voxel resolution = $1.3 \times 1.0 \times 1.3$ mm^3, TR = 2530 ms; TE = 3.25 ms; flip angle = 7°; field of view = 256 mm; 208 axial slices (1.33 mm thickness)). Subsequently, a 6-min resting state functional scan was acquired, where the participants were instructed to remain still, close their eyes, think of nothing systematically and not fall asleep. The resting state scan consisted of 240 BOLD 3D images (TR = 2000 ms, TE = 15 ms, flip angle = 90°, FOV = 240 mm, acquired voxel size = $3 \times 3 \times 4$ mm^3 and number of slices = 33).

fMRI preprocessing

All functional MRI data were preprocessed using the SPM12 software package, along with in-house MATLAB scripts. Structural MRI images were segmented into gray matter, white matter (WM), and cerebrospinal fluid (CSF). Prior to preprocessing, quality assurance (QA) of functional and structural MRI was performed, and subjects with poor quality imaging data were excluded from analysis. Resting-state functional MRI (fMRI) data were realigned to the averaged time frame to correct for head motion, slice-time corrected to the first slice, and co-registered to each participant's T1-weighted images. The functional images were then normalized to the standard Montreal Neurological Institute (MNI152) template at 2mm^3. A 6-mm Gaussian kernel was used to spatially smooth the functional images and a band-pass filter ranging from 0.01 to 0.1 Hz was applied. Band-pass filtering of fMRI timeseries was used to remove low frequency artifacts such as scanner drifts and high-frequency components, which do not contain useful information. Critically, band-pass filtering does not remove non-stationarities in the data, and non-stationarities such as time-varying means and covariances can still exist in a band-pass filtered signal. To account for artifacts from motion and nonneural sources, the mean time series from each of the CSF and WM masks as well as six motion parameters, obtained by rigid body registration, were regressed out from the fMRI data. We used the binarized WM and CSF tissue probability maps provided by FSL (https://fsl.fmrib.ox.ac.uk/fsl).

Data input into the stDNN

We used the Brainnetome Atlas (246 regions) and computed the average resting-state fMRI timeseries across the voxels in a given region of interest (ROI). We used Brainnetome as it provides fine-

grained brain-wide parcellations of both cortical and subcortical areas with better anatomical and functional interpretability than most other atlases. Critically, the Brainnetome Atlas is one of the most extensively used atlases, with over 1000 studies using it, enabling the comparison of our method/findings with those from extant related research work as well as those under development elsewhere. Each participant's time series data was represented by a matrix of size $N_C \times N_T$, where N_C is the number of channels or ROIs, and N_T is the number of time points. Critically, because our stDNN model accommodates varying time lengths, subjects in the training and testing cohorts were not required to have the same number of time points, as is common with existing approaches.

stDNN model

We developed an innovative stDNN model to extract informative brain dynamics features that accurately distinguish between 22q11.2DS and neurotypical controls. A key advantage of our approach is that it provides a novel technique to capture latent dynamics without the need for explicit feature engineering [90]. Our stDNN model consists of two 1D convolutional block layers, a "temporal averaging" operation, and then a linear output layer (Supplementary Fig. 2, Supplementary Table S28). Each convolutional block layer consists of a convolutional operation and Parametric ReLU activation. We introduce "batch normalization" and "maxpool" layers after each of the two convolutional block layers. The batch normalization layers help in training the stDNN faster and more stable by normalizing each layers' inputs. The "maxpool" layers help in (a) reducing the temporal dimension of the data, (b) hierarchical representation of the features, and (c) increasing the receptive field of the filter to capture the long-term correlations in the timeseries. Conventionally, after the last convolutional block, the data is flattened and a fully connected layer is connected to an output layer. The fully connected layers typically have the maximum number of parameters to be trained compared to the convolutional layers. In our model, instead of the normal flattening operation, we use a "temporal averaging layer" where we average the temporal features for each filter and therefore the number of inputs to the fully connected layer is just the number of output channels of the second convolution block layer. The temporal averaging layer is a dimensionality reduction step in the latent space and not in the original timeseries space, so is unlikely to cause loss of significant temporal information. The advantages of the averaging layer over the flattening layer are (a) the number of parameters reduced from $N_{C2} \times N_{T2}$ to N_{C2}, where N_{C2} is the number of output channels of the second convolutional block layer and N_{T2} is the temporal dimension of the output of the second "maxpool" layer, (b) with the averaging layer, we can train and test fMRI timeseries with varying time lengths. Varying time length is common with open-source data where the data is acquired with different data acquisition protocols. We introduce a dropout layer (=0.55) after each of the two convolutional block layers to avoid overfitting during the model training process. The stDNN classified participants in the two groups by minimizing the binary cross-entropy cost function. We train the model for up to 250 epochs with a stopping criterion and a learning rate of 0.0003 with a batch size of 16. An Adam optimizer with a weight decay of 6e-7 was used to estimate the stDNN model parameters [91]. Model hyperparameters were determined using Ray Tune.

Classification of 22q11.2DS versus control subjects in the UCLA cohort

To prevent bias and account for low variance, we conducted a five-fold cross-validation to evaluate the performance of our stDNN model. In the five-fold cross-validation approach, we divided the whole dataset into five different parts. We used four parts for training and validation and the fifth part as the test set. We then rotated through the whole dataset five times to select a different section as the test set during each iteration (Supplementary Fig. 3). For each of the five subsets, we evaluated the performance of our stDNN model individually and then averaged over the five subsets to report the mean and standard deviation values of the key performance metrics (accuracy, precision, recall, F1). Using the five-fold cross-validation approach, the performance for every sample from the UCLA data gets accounted, which helps in assessing the effectiveness of the model more robustly instead of just reporting the performance on one-time random split of the data.

Generalization of 22q11.2DS classification in the UCLA cohort to the independent UCDavis and PUC cohorts

Similar to the five-fold cross-validation process used for UCLA, for reporting the performance of our stDNN for UCDavis cohort, we used each of the five stDNN models trained on different subsets of UCLA. Using the five different models, we evaluated each model's performance on the UCDavis cohort data independently (Supplementary Fig. 3) and reported the mean and standard deviation values of the key performance metrics (accuracy, precision, recall, F1).

Similar to the five-fold cross-validation process used for UCLA, for reporting the performance of our stDNN for the PUC cohort, we used each of the five stDNN models trained on different subsets of UCLA. Using the five different models, we evaluate each model's performance on the PUC cohort data independently (Supplementary Fig. 3) and report the mean and standard deviation values of the key performance metrics (accuracy, precision, recall, F1).

Control analysis examining alternative models for classification of 22q11.2DS versus Control subjects

stDNN models the dynamic spatiotemporal characteristics of brain activity to classify 22q11.2DS vs. neurotypical controls using fMRI timeseries without any explicit feature engineering. To demonstrate the advantages of our stDNN model over extant classification approaches, we performed extensive control analyses. We evaluated several commonly used linear and nonlinear classification algorithms, including K-Nearest Neighbor, Decision Tree, Linear SVM, Logistic Regression, Ridge Classifier, LASSO, and Random Forest. fMRI timeseries features were provided as input to these classification algorithms. We explored several commonly used fMRI timeseries feature spaces, including static functional connectivity, amplitude of low-frequency fluctuation (ALFF), BOLD signal variability, and sliding-window functional connectivity variability. Static functional connectivity [92], which represents time-invariant functional interactions between brain regions, was computed by Pearson correlations between regional fMRI timeseries. ALFF [93], which reflects the intensity of spontaneous fluctuations in a brain region, was computed by obtaining the square root of the fMRI timeseries signal across the low-frequency range of 0.01–0.08 Hz for each of the 246 Brainnetome brain regions. BOLD signal variability, which represents the moment-to-moment brain variability, was calculated as the standard deviation of regional fMRI signal across the duration of the fMRI scan. Sliding-window functional connectivity variability, which reflects variability in dynamic functional interactions between brain regions, was computed by first calculating time-varying functional connectivity using a sliding-window approach [94] and then calculating the standard deviation of the time-varying functional connectivity. To evaluate the performance of the aforementioned classification models in distinguishing between 22q11.2DS and healthy controls, we used the same five-fold cross-validation approach we used to evaluate our stDNN model. Briefly, we divided the whole dataset into five different parts, where we used four parts for training and validation and the fifth part as the test set. We then rotated through the whole dataset five times to select a different section as the test set during each iteration. For each of the five subsets, we evaluated the performance of the alternate classification model individually and then averaged over the five subsets to report the mean and standard deviation values of the key performance metrics (accuracy, precision, recall, F1).

Identification of discriminating brain features underlying 22q11.2DS classification

We used an integrated gradients-based feature attribution approach to identify brain features that discriminated between the 22q11.2DS and healthy control groups. A major problem in developing and evaluating feature attribution methods is that it is difficult to distinguish errors from the DNN model and those from feature attribution procedures. Integrated gradients solves this problem by taking an approach that satisfies two fundamental axioms—sensitivity and implementation invariance [43]. Another advantage of integrated gradients is that the gradients can be computed easily for any given network architecture. Integrated gradients estimates the integral of gradients with respect to the *i-th* dimension of the input x along the straight-line path from a given (or random) baseline to the input as follows:

$$IG_i = (x - x') \int_0^1 (x_i - x'_i) \frac{\partial F(x' + \alpha(x - x'))}{\partial x_i} \partial \alpha$$

where, IG_i is the integrated gradient for the *i-th* component of the input x and x' is the baseline input for which the neural network F results in a neutral output. *IG* provides a score of how important each feature contributes to the final prediction. This approach provides insights about important features that predict 22q11.2DS class label. Conventional gradient-based approaches wrongly assign zero attributions for inputs where the function is flat, even when the output of F for such an input is different from the baseline. Integrated gradients avoids this problem by computing an average gradient along a linear path. Our integrated gradients implementation is based on the *"Captum"* (https://captum.ai/docs/introduction.html) module of *Pytorch*. The IG-derived feature importance/weights are computed at an individual level and relative to a baseline that is common across individuals, and therefore were not normalized.

Cross-cohort comparative analyses of brain features underlying 22q11.2DS classification

To examine the consistency of brain features underlying 22q11.2DS classification across cohorts, we conducted cross-cohort comparative analyses. Specifically, we computed Pearson correlation analysis between the mean brain signatures of all 22q-PS+ individuals in the UCLA cohort, the mean brain signatures of all 22q-PS+ individuals in the UCDavis cohort, and the mean brain signatures of all 22q-PS+ individuals in the PUC cohort. The mean brain signatures are comprised of ROI-level feature strengths; as a result, we did not utilize the spin test, which is typically employed to control for voxel-level spatial similarity, in this analysis. This approach was chosen as the focus was on ROI-level features rather than voxel-level spatial characteristics.

Distinctiveness of brain features underlying 22q11.2DS classification

The aforementioned integrated gradients-based feature attribution procedure also identifies an individual signature of predictive brain features in each participant. We examined whether these "signatures" cluster differently in 22q11.2DS individuals from healthy controls. Specifically, we computed a distance metric across brain features between individuals and compared the distances between individuals in the 22q11.2DS and healthy control groups. The Pearson correlation between individuals' integrated gradients-derived brain feature maps was used to calculate the distance between them.

Identification of brain features predictive of 22q11.2DS-associated psychosis

We next determined the brain features predictive of 22q-PS+ classification using the output of the integrated gradients procedure—individualized brain signature – described in the previous section. To identify brain areas that were most predictive of 22q-PS+ individuals, we computed the mean of brain signatures of all 22q-PS+ individuals and contrasted them against the mean of brain signatures of 22q-PS− individuals.

Cross-cohort comparative analyses of brain features predictive of 22q11.2DS-associated psychosis

To examine the consistency of brain features predictive of 22q11.2DS-associated psychosis across cohorts, we conducted cross-cohort comparative analyses. Specifically, we computed Pearson correlation analysis between the mean brain signatures of all 22q-PS+ individuals in the UCLA cohort and the mean brain signatures of all 22q-PS+ individuals in the PUC cohort.

The mean brain signatures are comprised of ROI-level feature strengths; as a result, we did not utilize the spin test, which is typically employed to control for voxel-level spatial similarity, in this analysis. This approach was chosen as the focus was on ROI-level features rather than voxel-level spatial characteristics.

Distinctiveness of brain features predictive of 22q11.2DS-associated psychosis

We next examined whether these "signatures" cluster differently in 22q-PS+ individuals from 22q-PS− individuals. Specifically, we computed a distance metric across brain features between individuals and compared the distances between individuals in the 22q-PS+ and 22q-PS− groups. The Pearson correlation between individuals' integrated gradients-derived brain feature maps was used to calculate the distance between them.

Examination of the relationship between the brain features predictive of 22q11.2DS-associated psychosis and brain features predictive of idiopathic early psychosis

To determine the overlap between 22q-PS+ and idiopathic early psychosis, we computed a distance metric across brain features between individuals and compared the distances between individuals in the 22q-PS+ and 22q-PS− groups from the combined UCLA and PUC cohort, and individuals in the idiopathic early psychosis group from the HCP-EP cohort. The Pearson correlation between individuals' integrated gradients-derived brain feature maps was used to calculate the distance between them.

To further examine the overlap between 22q-PS+ and idiopathic early psychosis, we trained a de novo stDNN model to distinguish between 22q-PS+ and 22q-PS− by combining 22q11.2DS data from the UCLA and PUC cohorts and then investigated whether the stDNN model trained to distinguish between 22q-PS+ and 22q-PS− can identify individuals with idiopathic early psychosis from the HCP-EP cohort. In this analysis, we combined 22q11.2DS data from the UCLA and PUC cohorts to increase the sample size. Importantly, we de novo trained our stDNN model here using a label-distribution-aware margin (LDAM) loss as our combined dataset has a small number of 22q-PS+ participants compared to 22q-PS−. We have previously shown LDAM loss to outperform the conventional cross-entropy loss under such class-imbalance conditions. Specifically, we first conducted a five-fold cross-validation to evaluate the performance of our stDNN model with LDAM loss in distinguishing between 22q-PS+ and 22q-PS−.

In the five-fold cross-validation approach, we divided the combined UCLA-PUC dataset into five different parts, where we used four parts for training and validation and the fifth part as the test set. We then rotated through the whole dataset five times to select a different section as the test set during each iteration. For each of the five subsets, we evaluated the performance of our stDNN model individually and then averaged over the five subsets to report the mean and standard deviation values of the key performance metrics (accuracy, precision, recall, F1). We next investigated whether the stDNN model trained to distinguish between 22q-PS+ and 22q-PS− can distinguish individuals with idiopathic early psychosis in the HCP-EP cohort. Specifically, we used the stDNN model that most

accurately distinguishes between 22q-PS+ and 22q-PS− in the combined UCLA-PUC dataset. Using the best model, we evaluated its performance on the HCP-EP idiopathic early psychosis data and reported classification accuracy. Lastly, to determine the model's specificity, we investigated whether the stDNN model trained to distinguish between 22q-PS+ and 22q-PS− can distinguish individuals with idiopathic autism in the Stanford cohort and individuals with idiopathic ADHD in the NYU cohort. Specifically, we used the stDNN model that most accurately distinguishes between 22q-PS+ and 22q-PS− in the combined UCLA-PUC dataset. Using the best model, we evaluated its performance on the Stanford autism cohort data as well as NYU ADHD cohort data, and reported classification accuracy.

DATA AVAILABILITY

The processed fMRI regional time-series data along with demographic information will be made available upon reasonable request.

REFERENCES

1. McDonald-McGinn DM, Sullivan KE, Marino B, Philip N, Swillen A, Vorstman JAS, et al. 22q11.2 deletion syndrome. Nat Rev Dis Prim. 2015;1:15071.
2. Oskarsdottir S, Vujic M, Fasth A. Incidence and prevalence of the 22q11 deletion syndrome: a population-based study in Western Sweden. Arch Dis Child. 2004;89:148–51.
3. Shprintzen RJ. Velo-cardio-facial syndrome: 30 Years of study. Dev Disabil Res Rev. 2008;14:3–10.
4. Goodship J, Cross I, LiLing J, Wren C. A population study of chromosome 22q11 deletions in infancy. Arch Dis Child. 1998;79:348–51.
5. Blagojevic C, Heung T, Theriault M, Tomita-Mitchell A, Chakraborty P, Kernohan K, et al. Estimate of the contemporary live-birth prevalence of recurrent 22q11.2 deletions: a cross-sectional analysis from population-based newborn screening. CMAJ Open. 2021;9:E802–9.
6. Shaikh TH, O'Connor RJ, Pierpont ME, McGrath J, Hacker AM, Nimmakayalu M, et al. Low copy repeats mediate distal chromosome 22q11.2 deletions: sequence analysis predicts breakpoint mechanisms. Genome Res. 2007;17:482–91.
7. Edelmann L, Pandita RK, Morrow BE. Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. Am J Hum Genet. 1999;64:1076–86.
8. Maynard TM, Haskell GT, Peters AZ, Sikich L, Lieberman JA, LaMantia AS. A comprehensive analysis of 22q11 gene expression in the developing and adult brain. Proc Natl Acad Sci USA. 2003;100:14433–8.
9. Guna A, Butcher NJ, Bassett AS. Comparative mapping of the 22q11.2 deletion region and the potential of simple model organisms. J Neurodevelopmental Disord. 2015;7:18.
10. Murphy KC, Jones LA, Owen MJ. High rates of schizophrenia in adults with velo-cardio-facial syndrome. Arch Gen Psychiatry. 1999;56:940–5.
11. Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. Nat Genet. 2017;49:27–35.
12. Bassett AS, Chow EW. Schizophrenia and 22q11.2 deletion syndrome. Curr Psychiatry Rep. 2008;10:148–57.
13. Green T, Gothelf D, Glaser B, Debbane M, Frisch A, Kotler M, et al. Psychiatric disorders and intellectual functioning throughout development in velocardiofacial (22q11.2 deletion) syndrome. J Am Acad Child Adolesc Psychiatry. 2009;48:1060–8.
14. Schneider M, Debbane M, Bassett AS, Chow EW, Fung WL, van den Bree M, et al. Psychiatric disorders from childhood to adulthood in 22q11.2 deletion syndrome: results from the International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome. Am J Psychiatry. 2014;171:627–39.
15. Siebner HR, Callicott JH, Sommer T, Mattay VS. From the genome to the phenome and back: linking genes with human brain function and structure using genetically informed neuroimaging. Neuroscience. 2009;164:1–6.
16. Fiksinski AM, Hoftman GD, Vorstman JAS, Bearden CE. A genetics-first approach to understanding autism and schizophrenia spectrum disorders: the 22q11.2 deletion syndrome. Mol Psychiatry. 2023;28:341–53.
17. Jonas RK, Montojo CA, Bearden CE. The 22q11.2 deletion syndrome as a window into complex neuropsychiatric disorders over the lifespan. Biol Psychiatry. 2014;75:351–60.
18. Lewandowski KE, Bouix S, Ongur D, Shenton ME. Neuroprogression across the early course of psychosis. J Psychiatry Brain Sci. 2020;5:e200002.
19. Padula MC, Schaer M, Scariati E, Schneider M, Van De Ville D, Debbane M, et al. Structural and functional connectivity in the default mode network in 22q11.2 deletion syndrome. J Neurodev Disord. 2015;7:23.
20. Schreiner MJ, Karlsgodt KH, Uddin LQ, Chow C, Congdon E, Jalbrzikowski M, et al. Default mode network connectivity and reciprocal social behavior in 22q11.2 deletion syndrome. Soc Cogn Affect Neurosci. 2014;9:1261–7.
21. Zoller D, Schaer M, Scariati E, Padula MC, Eliez S, Van De Ville D. Disentangling resting-state BOLD variability and PCC functional connectivity in 22q11.2 deletion syndrome. Neuroimage. 2017;149:85–97.
22. Debbane M, Lazouret M, Lagioia A, Schneider M, Van De Ville D, Eliez S. Resting-state networks in adolescents with 22q11.2 deletion syndrome: associations with prodromal symptoms and executive functions. Schizophr Res. 2012;139:33–9.
23. Scariati E, Schaer M, Richiardi J, Schneider M, Debbane M, Van De Ville D, et al. Identifying 22q11.2 deletion syndrome and psychosis using resting-state connectivity patterns. Brain Topogr. 2014;27:808–21.
24. Schreiner M, Forsyth JK, Karlsgodt KH, Anderson AE, Hirsh N, Kushan L, et al. Intrinsic connectivity network-based classification and detection of psychotic symptoms in youth with 22q11.2 deletions. Cereb Cortex. 2017;27:3294–306.
25. Mattiaccio LM, Coman IL, Thompson CA, Fremont WP, Antshel KM, Kates WR. Frontal dysconnectivity in 22q11.2 deletion syndrome: an atlas-based functional connectivity analysis. Behav Brain Funct. 2018;14:2.
26. Schleifer C, Lin A, Kushan L, Ji JL, Yang G, Bearden CE, et al. Dissociable disruptions in thalamic and hippocampal resting-state functional connectivity in youth with 22q11.2 deletions. J Neurosci. 2019;39:1301–19.
27. Chow EWC, Zipursky RB, Mikulis DJ, Bassett AS. Structural brain abnormalities in patients with schizophrenia and 22q11 deletion Syndrome. Biol Psychiatry. 2002;51:208–15.
28. Eliez S, Schmitt JE, White CD, Reiss AL. Children and adolescents with velocardiofacial syndrome: a volumetric MRI study. Am J Psychiatry. 2000;157:409–15.
29. Campbell LE, Daly E, Toal F, Stevens A, Azuma R, Catani M, et al. Brain and behaviour in children with 22q11.2 deletion syndrome: a volumetric and voxel-based morphometry MRI study. Brain. 2006;129:1218–28.
30. Schmitt JE, Vandekar S, Yi J, Calkins ME, Ruparel K, Roalf DR, et al. Aberrant cortical morphometry in the 22q11.2 deletion syndrome. Biol Psychiatry. 2015;78:135–43.
31. Jalbrzikowski M, Jonas R, Senturk D, Patel A, Chow C, Green MF, et al. Structural abnormalities in cortical volume, thickness, and surface area in 22q11.2 microdeletion syndrome: relationship with psychotic symptoms. Neuroimage-Clin. 2013;3:405–15.
32. Bearden CE, van Erp TGM, Dutton RA, Tran H, Zimmermann L, Sun DQ, et al. Mapping cortical thickness in children with 22q11.2 deletions. Cereb Cortex. 2007;17:1889–98.
33. Sun D, Ching CRK, Lin A, Forsyth JK, Kushan L, Vajdi A, et al. Large-scale mapping of cortical alterations in 22q11.2 deletion syndrome: Convergence with idiopathic psychosis and effects of deletion size. Mol Psychiatry. 2020;25:1822–34.
34. Ching CRK, Gutman BA, Sun D, Villalon Reina J, Ragothaman A, Isaev D, et al. Mapping subcortical brain alterations in 22q11.2 deletion syndrome: effects of deletion size and convergence with idiopathic neuropsychiatric illness. Am J Psychiatry. 2020;177:589–600.
35. Villalon-Reina JE, Martinez K, Qu X, Ching CRK, Nir TM, Kothapalli D, et al. Altered white matter microstructure in 22q11.2 deletion syndrome: a multisite diffusion tensor imaging study. Mol Psychiatry. 2020;25:2818–31.
36. Ryali S, Supekar K, Chen T, Kochalka J, Cai W, Nicholas J, et al. Temporal dynamics and developmental maturation of salience, default and central-executive network interactions revealed by variational bayes hidden markov modeling. PLoS Comput Biol. 2016;12:e1005138.
37. Taghia J, Cai W, Ryali S, Kochalka J, Nicholas J, Chen T, et al. Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition. Nat Commun. 2018;9:2505.
38. Ryali S, Supekar K, Chen T, Menon V. Multivariate dynamical systems models for estimating causal interactions in fMRI. Neuroimage. 2011;54:807–23.
39. Calhoun VD, Miller R, Pearlson G, Adali T. The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. Neuron. 2014;84:262–74.
40. Supekar K, Ryali S, Yuan R, Kumar D, de Los Angeles C, Menon V. Robust, generalizable, and interpretable artificial intelligence-derived brain fingerprints of autism and social communication symptom severity. Biol Psychiatry. 2022;92:643–53.
41. Supekar K, de Los Angeles C, Ryali S, Cao K, Ma T, Menon V. Deep learning identifies robust gender differences in functional brain organization and their dissociable links to clinical symptoms in autism. Br J Psychiatry. 2022;220:202–9.
42. Durstewitz D, Koppe G, Meyer-Lindenberg A. Deep neural networks in psychiatry. Mol Psychiatry. 2019;24:1583–98.
43. Axiomatic attribution for deep networks. Proceedings of the proceedings of the 34th international conference on machine learning-Volume 702017. JMLR. org.
44. Jalbrzikowski M. Neuroimaging phenotypes associated with risk and resilience for psychosis and autism spectrum disorders in 22q11.2 microdeletion syndrome. Biol Psychiatry Cogn Neurosci Neuroimaging. 2021;6:211–24.

45. Menon V, Palaniyappan L, Supekar K. Integrative brain network and salience models of psychopathology and cognitive dysfunction in schizophrenia. Biol Psychiatry. 2023;94:108–20.
46. Bassett AS, Chow EW, AbdelMalik P, Gheorghiu M, Husted J, Weksberg R. The schizophrenia phenotype in 22q11 deletion syndrome. Am J Psychiatry. 2003;160:1580–6.
47. Armando M, Girardi P, Vicari S, Menghini D, Digilio MC, Pontillo M, et al. Adolescents at ultra-high risk for psychosis with and without 22q11 deletion syndrome: a comparison of prodromal psychotic symptoms and general functioning. Schizophr Res. 2012;139:151–6.
48. Richards C, Jones C, Groves L, Moss J, Oliver C. Prevalence of autism spectrum disorder phenomenology in genetic disorders: a systematic review and meta-analysis. Lancet Psychiatry. 2015;2:909–16.
49. Tang KL, Antshel KM, Fremont WP, Kates WR. Behavioral and psychiatric phenotypes in 22q11.2 deletion syndrome. J Dev Behav Pediatr. 2015;36:639–50.
50. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013;14:365–76.
51. Schmitt JE. Dynamic functional connectivity: a new perspective on 22q11.2 deletion syndrome and psychosis. Biol Psychiatry Cogn Neurosci Neuroimaging. 2019;4:852–3.
52. Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, et al. Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci. 2007;27:2349–56.
53. Haber SN, Knutson B. The reward circuit: linking primate anatomy and human imaging. Neuropsychopharmacology. 2010;35:4–26.
54. Hikida T, Morita M, Macpherson T. Neural mechanisms of the nucleus accumbens circuit in reward and aversive learning. Neurosci Res. 2016;108:1–5.
55. Abraham AD, Neve KA, Lattal KM. Dopamine and extinction: a convergence of theory with fear and reward circuitry. Neurobiol Learn Mem. 2014;108:65–77.
56. Lenz JD, Lobo MK. Optogenetic insights into striatal function and behavior. Behav Brain Res. 2013;255:44–54.
57. Lammel S, Lim BK, Malenka RC. Reward and aversion in a heterogeneous midbrain dopamine system. Neuropharmacology. 2014;76:351–9.
58. Berridge KC, Kringelbach ML. Pleasure systems in the brain. Neuron. 2015;86:646–64.
59. Krotewicz M, Romaniuk A. Social interactions, brain monoamines, and GABA alterations in MFB-lesioned cats. Pharm Biochem Behav. 1998;60:533–8.
60. Boylan CB, Blue ME, Hohmann CF. Modeling early cortical serotonergic deficits in autism. Behav Brain Res. 2007;176:94–108.
61. Ungerleider LG, Haxby JV. 'What' and 'where' in the human brain. Curr Opin Neurobiol. 1994;4:157–65.
62. Bearden CE, Woodin MF, Wang PP, Moss E, McDonald-McGinn D, Zackai E, et al. The neurocognitive phenotype of the 22q11.2 deletion syndrome: selective deficit in visual-spatial memory. J Clin Exp Neuropsychol. 2001;23:447–64.
63. Simon TJ, Bearden CE, Mc-Ginn DM, Zackai E. Visuospatial and numerical cognitive deficits in children with chromosome 22q11.2 deletion syndrome. Cortex. 2005;41:145–55.
64. Simon TJ, Wu Z, Avants B, Zhang H, Gee JC, Stebbins GT. Atypical cortical connectivity and visuospatial cognitive impairments are related in children with chromosome 22q11.2 deletion syndrome. Behav Brain Funct. 2008;4:25.
65. Kikinis Z, Makris N, Finn CT, Bouix S, Lucia D, Coleman MJ, et al. Genetic contributions to changes of fiber tracts of ventral visual stream in 22q11.2 deletion syndrome. Brain Imaging Behav. 2013;7:316–25.
66. Padula MC, Schaer M, Scariati E, Maeder J, Schneider M, Eliez S. Multimodal investigation of triple network connectivity in patients with 22q11DS and association with executive functions. Hum Brain Mapp. 2017;38:2177–89.
67. Sridharan D, Levitin DJ, Menon V. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proc Natl Acad Sci USA. 2008;105:12569–74.
68. Palaniyappan L, Liddle PF. Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. J Psychiatry Neurosci. 2012;37:17–27.
69. Menon V. Large-scale brain networks and psychopathology: a unifying triple network model. Trends Cogn Sci. 2011;15:483–506.
70. 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 Psychiatry. 2019;85:60–9.
71. Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. Am J Psychiatry. 2003;160:13–23.
72. Winton-Brown TT, Fusar-Poli P, Ungless MA, Howes OD. Dopaminergic basis of salience dysregulation in psychosis. Trends Neurosci. 2014;37:85–94.
73. Roiser JP, Howes OD, Chaddock CA, Joyce EM, McGuire P. Neural and behavioral correlates of aberrant salience in individuals at risk for psychosis. Schizophr Bull. 2013;39:1328–36.
74. Howes OD, Hird EJ, Adams RA, Corlett PR, McGuire P. Aberrant salience, information processing, and dopaminergic signaling in people at clinical high risk for psychosis. Biol psychiatry. 2020;88:304–14.
75. McCutcheon RA, Krystal JH, Howes OD. Dopamine and glutamate in schizophrenia: biology, symptoms and treatment. World Psychiatry. 2020;19:15–33.
76. McCutcheon RA, Abi-Dargham A, Howes OD. Schizophrenia, dopamine and the striatum: from biology to symptoms. Trends Neurosci. 2019;42:205–20.
77. Howes OD, Montgomery AJ, Asselin M-C, Murray RM, Valli I, Tabraham P, et al. Elevated striatal dopamine function linked to prodromal signs of schizophrenia. Arch Gen Psychiatry. 2009;66:13–20.
78. Kegeles LS, Abi-Dargham A, Frankle WG, Gil R, Cooper TB, Slifstein M, et al. Increased synaptic dopamine function in associative regions of the striatum in schizophrenia. Arch Gen psychiatry. 2010;67:231–9.
79. Rogdaki M, Devroye C, Ciampoli M, Veronese M, Ashok AH, McCutcheon RA, et al. Striatal dopaminergic alterations in individuals with copy number variants at the 22q11. 2 genetic locus and their implications for psychosis risk: a [18F]-DOPA PET study. Mol Psychiatry. 2023;28:1995–2006.
80. McCutcheon RA, Nour MM, Dahoun T, Jauhar S, Pepper F, Expert P, et al. Mesolimbic dopamine function is related to salience network connectivity: an integrative positron emission tomography and magnetic resonance study. Biol Psychiatry. 2019;85:368–78.
81. Cheon EJ, Bearden CE, Sun D, Ching CRK, Andreassen OA, Schmaal L, et al. Cross disorder comparisons of brain structure in schizophrenia, bipolar disorder, major depressive disorder, and 22q11.2 deletion syndrome: a review of ENIGMA findings. Psychiatry Clin Neurosci. 2022;76:140–61.
82. Chow EW, Ho A, Wei C, Voormolen EH, Crawley AP, Bassett AS. Association of schizophrenia in 22q11.2 deletion syndrome and gray matter volumetric deficits in the superior temporal gyrus. Am J Psychiatry. 2011;168:522–9.
83. van Amelsvoort T, Daly E, Henry J, Robertson D, Ng V, Owen M, et al. Brain anatomy in adults with velocardiofacial syndrome with and without schizophrenia: preliminary results of a structural magnetic resonance imaging study. Arch Gen Psychiatry. 2004;61:1085–96.
84. Bakker G, Caan MWA, Vingerhoets WAM, da Silva-Alves F, de Koning M, Boot E, et al. Cortical morphology differences in subjects at increased vulnerability for developing a psychotic disorder: a comparison between subjects with ultra-high risk and 22q11.2 deletion syndrome. Plos One. 2016;11:e0159928.
85. Miller TJ, McGlashan TH, Rosen JL, Cadenhead K, Ventura J, McFarlane W, et al. Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. Schizophr Bull. 2003;29:703–15.
86. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull. 1987;13:261–76.
87. Abrams DA, Padmanabhan A, Chen T, Odriozola P, Baker AE, Kochalka J, et al. Impaired voice processing in reward and salience circuits predicts social communication in children with autism. Elife. 2019;8:e39906.
88. Abrams DA, Lynch CJ, Cheng KM, Phillips J, Supekar K, Ryali S, et al. Underconnectivity between voice-selective cortex and reward circuitry in children with autism. Proc Natl Acad Sci USA. 2013;110:12060–5.
89. Lynch CJ, Uddin LQ, Supekar K, Khouzam A, Phillips J, Menon V. Default mode network in childhood autism: posteromedial cortex heterogeneity and relationship with social deficits. Biol psychiatry. 2013;74:212–9.
90. Davatzikos C. Machine learning in neuroimaging: progress and challenges. Neuroimage. 2019;197:652–6.
91. Kingma DP, Ba J Adam: A method for stochastic optimization. Proceedings of the 3rd international conference on learning representations (ICLR) 1412. (2015).
92. van den Heuvel MP, Hulshoff Pol HE. Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur Neuropsychopharmacol. 2010;20:519–34.
93. Zang YF, He Y, Zhu CZ, Cao QJ, Sui MQ, Liang M, et al. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. Brain Dev. 2007;29:83–91.
94. Allen EA, Damaraju E, Plis SM, Erhardt EB, Eichele T, Calhoun VD. Tracking whole-brain connectivity dynamics in the resting state. Cereb Cortex. 2014;24:663–76.

AUTHOR CONTRIBUTIONS

KS designed and conceptualized the study, contributed methods, performed investigation, wrote the original draft of the manuscript and reviewed and edited it. CDLA helped in data analysis. SR contributed to the development of the deep neural network model. LK, CS, GR, NC, and TS supervised and contributed to data collection. CB supervised data collection and reviewed and edited the manuscript. VM designed and conceptualized the study, and reviewed and edited the manuscript.

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/s41380-024-02495-8.

Correspondence and requests for materials should be addressed to Kaustubh Supekar or Vinod Menon.